Pontifícia Universidade Católica
do Rio de Janeiro

Victor D'Angelo Colacino

Taxa de convergência do Teorema Central do Limite para a expressão martingal de desvio da contagem de subgrafos livres de triângulos em grafos aleatórios $G(n, m)$

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre pelo Programa de Pós-graduação em Matemática da PUC-Rio .

Orientador: Prof. Simon Griffiths

Rio de Janeiro
Janeiro de 2021

Pontifícia Universidade Católica
do Rio de Janeiro

Victor D'Angelo Colacino

Taxa de convergência do Teorema Central do Limite para a expressão martingal de desvio da contagem de subgrafos livres de triângulos em grafos aleatórios $G(n, m)$

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre pelo Programa de Pós-graduação em Matemática da PUC-Rio . Aprovada pela Comissão Examinadora abaixo.

Prof. Simon Griffiths
Orientador
Departamento de Matemática – PUC-Rio

Prof. Silvius Klein
Departamento de Matemática – PUC-Rio

Prof. Matheus Secco Torres da Silva
Instituto de Ciência da Computação – Academia De Ciências Da República Tcheca

Prof. Glauco Valle da Silva Coelho
Departamento de Métodos Estatísticos – UFRJ

Prof. Roberto Imbuzeiro Moraes Felinto de Oliveira
Departamento de Matemática – IMPA

Rio de Janeiro, 28 de Janeiro de 2021

Victor D'Angelo Colacino

Graduado em Ciências Atuariais pela Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ, Brasil.

Ficha Catalográfica

Colacino, Victor

Taxa de convergência do Teorema Central do Limite para a expressão martingal de desvio da contagem de subgrafos livres de triângulos em grafos aleatórios $G(n, m)$ / Victor D'Angelo Colacino; orientador: Simon Griffiths. – Rio de janeiro: PUC-Rio , Departamento de Matemática, 2021.

v., 81 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Matemática.

Inclui bibliografia

1. Martingais;. 2. Teorema Central do Limite;. 3. Taxa de Convergência;. 4. Teoria dos Grafos;. 5. Contagem de Subgrafos.. I. Griffiths, Simon. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Matemática. III. Título.

CDD: 004

A minha família, amigos e Julia,
por todo suporte ao longo desta jornada.

Agradecimentos

Primeiramente, agradeço aos meus pais Fabíola e Vicente e ao meu irmão Lucas, por sempre instigarem minha curiosidade e meu lado estudioso.

Agradeço aos amigos que fiz na PUC-Rio, por toda a ajuda e momentos de descontração nessa etapa tão desafiadora.

Agradeço também às pessoas queridas que me apoiaram e incentivaram ao longo desse processo, em especial Fernando, Juliana, Yago, Udo, Luciano e Julia.

Faço um agradecimento especial ao meu Orientador Simon, que esteve presente em todas as etapas do meu Mestrado e foi essencial para a realização desse trabalho, através de muito suporte, paciência e conhecimento compartilhado. Admiro muito sua inteligência e bom humor.

Agradeço ainda aos membros da banca examinadora: Silvius Klein, Matheus Secco, Glauco Valle e Roberto Imbuzeiro, pelo aceite ao convite e por toda contribuição para este projeto.

Meu sincero agradecimento também dedico a Creuza, assim como ao corpo administrativo do Departamento de Matemática, pelo auxílio e boa vontade em lidar com toda a burocracia necessária.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Colacino, Victor; Griffiths, Simon. **Taxa de convergência do Teorema Central do Limite para a expressão martingal de desvio da contagem de subgrafos livres de triângulos em grafos aleatórios** $G(n,m)$. Rio de Janeiro, 2021. 81p. Dissertação de Mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

Nessa dissertação vamos introduzir, elaborar e combinar ideias da Teoria de martingais, a Teoria de grafos aleatórios e o Teorema Central do Limite. Em particular, veremos como martingais podem ser usados para representar desvios de contagem de subgrafos. Usando esta representação e o Teorema Central do Limite para martingais, conseguiremos demonstrar um Teorema Central do Limite para a contagem de subgrafos livres de triângulos no grafo aleatório Erdős-Rényi $G(n,m)$. Além disso, nossa demonstração também nos trará informação sobre a taxa de convergência, mostrando que a distribuição dos desvios converge rapidamente para a distribuição normal.

Palavras-chave

Martingais; Teorema Central do Limite; Taxa de Convergência; Teoria dos Grafos; Contagem de Subgrafos.

Abstract

Colacino, Victor; Griffiths, Simon (Advisor). **Rate of convergence of the Central Limit Theorem for the martingale expression of deviations of triangle-free subgraph counts in $G(n, m)$ random graphs**. Rio de Janeiro, 2021. 81p. Dissertação de Mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

In this dissertation we shall introduce, elaborate and combine ideas from martingale Theory, random graph Theory and the Central Limit Theorem. In particular, we will see how martingales can be used to represent deviations of subgraph counts. Using this representation and the Central Limit Theorem for martingales, we will be able to demonstrate a Central Limit Theorem for the triangle-free subgraph count in the Erdős-Rényi $G(n, m)$ random graph. Furthermore, our proof also gives us information about the rate of convergence, showing that the distribution of deviations converges rapidly to the normal distribution.

Keywords

Martingales; Central Limit Theorem; Convergence Rate; Graph Theory; Subgraph Count.

Sumário

Lista de figuras

A felicidade está no gosto e não nas coisas; é por ter o que amamos que somos felizes, e não por ter o que os outros acham amável.

La Rochefoucauld, *Reflexões ou sentenças e máximas morais.*

1
Introdução

O objetivo principal dessa dissertação é mostrar um Teorema Central do Limite para contagem de subgrafos, com uma taxa de convergência. Em particular, para qualquer grafo H livre de triângulos, mostraremos um Teorema Central do Limite com uma taxa para a contagem de cópias de H no grafo aleatório Erdős-Rényi $G(n, m)$.

Iniciaremos com uma discussão geral sobre contagem de subgrafos.

O modelo Erdős-Rényi [1] [2] é muito utilizado para a geração de grafos aleatórios por meio de duas variantes: o modelo $G(n, m)$, onde o grafo é escolhido de forma uniforme aleatória da coleção de todos os grafos que possuem n vértices e m arestas, e o modelo $G(n, p)$, onde cada aresta é incluída no grafo de n vértices independentemente com probabilidade p.

Dados grafos H e G, podemos definir $N_H(G)$ como o número de cópias isomórficas de H em G, isto é, o número de funções injetivas $\phi : V(H) \to V(G)$ tal que

$$\phi(u)\phi(v) \in E(G) \qquad \text{para todo} \qquad uv \in E(H)\,.$$

Também, escrevemos $L_H(G) = \mathbb{E}(N_H(G))$. Desse modo, definimos $D_H(G) = N_H(G) - L_H(G)$ como o desvio na contagem de subgrafos H em G.

Em anos recentes, o comportamento desses desvios foi foco de intenso estudo. Foram estabelecidos resultados relativos a grandes desvios (da ordem da média), pequenos desvios (da ordem do desvio padrão) e desvios moderados, que são intermediários.

Correspondendo à primeira categoria, desvios da ordem da média, recomendamos a pesquisa de Chatterjee [3] e suas referências para uma visão mais detalhada do assunto. Mais recentemente, muitos problemas relativos a essa categoria de desvios foram resolvidos através do método de estabilidade entrópica, que pode ser visto em mais detalhes no artigo de Harel, Mousset and Samotij [4].

Para desvios moderados, ou seja, de ordem maior do que o desvio padrão, porém menor do que a média, podemos citar o artigo de Goldschmidt, Griffiths e Scott [5], que será estudado mais a frente nessa dissertação e traz importantes avanços no assunto. Para mais referências sobre o assunto, ver artigos de Féray, Méliot e Nikeghbali [6] e Döring, Eichelsbacher [7].

No presente trabalho, focaremos no problema dos desvios na ordem do desvio padrão, que têm sido estudados com mais afinco desde os anos 80. Agora, vamos tangenciar alguns dos principais resultados relativos a essa classe de desvios.

Sejam $e(H)$ e $v(H)$ o número de arestas e vértices de H, respectivamente. Se G_p tem distribuição $G(n,p)$, então $\mathbb{E}(N_H(G_p)) = p^{e(H)}(n)_{v(H)}$, onde $(n)_k := n(n-1)\dots(n-k+1)$ denota o fatorial decrescente, e $D_H(G_p) = N_H(G_p) - p^{e(H)}(n)_{v(H)}$. Em [8], Ruciński mostra que, para toda a coleção de densidades p tal que $np^{e(H)}, (1-p)n^2 \to \infty$, o número de cópias de um grafo fixo H em $G(n,p)$ é assintoticamente normalmente distribuído e estabelece um Teorema Central do Limite para $D_H(G_p)$.

Para uma variável aleatória Y, escrevemos

$$d(Y) := \sup_{a\in\mathbb{R}} |\mathbb{P}\,[\,Y/\sigma_Y \leq a\,] - \Phi(a)|, \tag{1-1}$$

onde σ_Y é o desvio padrão de Y e $\Phi(a) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a}\exp(\frac{-t^2}{2})dt$ é a função de distribuição acumulada da distribuição normal.

Também é interessante mostrarmos versões quantitativas do Teorema Central do Limite, ou seja, mostrar o quão rápido $d(D_H(G_p)) \to 0$. Em [9], Krokowski, Reichenbachs e Thäle definem $F := (T - \mathbb{E}(T))/\sqrt{Var[T]})$ para $T = T(n,p)$ o número de triângulos em $G(n,p)$ e enunciam o seguinte resultado.

Teorema 1.1 *Seja $p = \theta n^{-\alpha}$ com $\alpha \in [0,1)$ e $\theta \in \mathbb{R}$. Então*

$$d(F) = \begin{cases} \mathcal{O}(n^{-1+\alpha}) & , \textit{se } 0 \leq \alpha \leq 1/2, \\ \mathcal{O}(n^{-3/4+\alpha/2}) & , \textit{se } 1/2 < \alpha \leq 2/3, \\ \mathcal{O}(n^{-5(1-\alpha)/4}) & , \textit{se } 2/3 < \alpha \leq 1, \end{cases}$$

onde $f(x) = \mathcal{O}(g(x))$ se e somente se existe um número real positivo M e um número real x_0 tal que $|f(x)| \leq M|g(x)|$ para todo $x \geq x_0$.

Em particular, se p é constante, isto é, se $\alpha = 0$,

$$d(F) = \mathcal{O}(n^{-1}).$$

No caso $G(n,m)$, Janson [10] prova um Teorema Central do Limite para a evolução do desvio na contagem de subgrafos e mostra que a contagem de subgrafos em $G(n,m)$ também são assintoticamente normalmente distribuídos.

Nessa dissertação, iremos mostrar o resultado quantitativo desse Teorema Central do Limite para $G(n,m)$. Conforme veremos no Capítulo 7, usamos a variável $D_H(G_m)$ para o desvio na contagem de subgrafos H em G_m, onde G_m tem distribuição $G(n,m)$. Essa variável pode ser expressa como um martingal, portanto poderemos estabelecer sua velocidade de convergência no Teorema Central do Limite, obtendo o seguinte resultado.

Utilizaremos $t = m/N$ para denotarmos a densidade do grafo $G(n,m)$, onde $N = \binom{n}{2}$ é o máximo número de arestas possíveis de um grafo com n vértices. Essas notações serão utilizadas a seguir nos capítulos 7 e 8.

Teorema 1.2 *Seja $k \in [1, +\infty)$, $\eta > 0$ e H um grafo livre de triângulos, então existe uma constante $C_{k,H,\eta}$ tal que o seguinte vale. Para todo $n \geq 2$ e todo $\eta N \leq m \leq N/2$ temos*

$$d(D_H(G_m)) \leq C_{k,H,\eta}\left[\left(\frac{\sqrt{\log(n)}}{tn^{1/2}}\right)^{\frac{k}{2k+1}}\right].$$

Daremos agora uma visão geral da dissertação. No segundo capítulo, estudaremos a Teoria de martingais. Inicialmente, apresentaremos o conceito e propriedades da esperança condicional, que é intrínseca à definição de martingal. Em seguida, definiremos martingais, submartingais e supermartingais, e apresentaremos propriedades, principais teoremas e exemplos relacionados.

No terceiro capítulo, apresentaremos resultados auxiliares básicos que nos servirão de apoio posteriormente.

No quarto capítulo, introduziremos o Teorema Central do Limite e provaremos duas de suas versões, a para variáveis aleatórias independentes e identicamente distribuídas e a versão de Lindeberg. Também, abordaremos o conceito e propriedades de funções características e convergência em distribuição, que são de extrema importância para as demonstrações do Teorema Central do Limite.

No quinto capítulo, veremos o Teorema Central do Limite para martingais, que permite dependência entre as variáveis aleatórias. Primeiro, provaremos uma versão geral do Teorema Central do Limite para martingais, e depois veremos uma versão mais específica, que será utilizada no próximo capítulo.

No sexto capítulo, exploraremos o artigo [11], que aborda as taxas de convergência no Teorema Central do Limite para martingais, restrito à classe de martingais com incrementos limitados.

No sétimo capítulo, abordaremos inicialmente noções e definições básicas da Teoria de grafos. Em seguida, iremos discutir o artigo [5], apresentaremos uma expressão martingal para o desvio na contagem de subgrafos $D_H(G_m)$, e uma expressão mais simples que aproxima bem $D_H(G_m)$. Também, abordaremos os conceitos de grau e cograu e, por fim, entenderemos o comportamento da variação quadrática discreta dos martingais utilizados.

No capítulo final, unificaremos todos os temas abordados até então para estabelecer uma taxa de convergência do Teorema Central do Limite para a expressão martingal do desvio na contagem de subgrafos $D_H(G_m)$. Focaremos nosso estudo em subgrafos que não contém triângulos, por motivos que ficarão mais claros a partir do sétimo capítulo.

2
Teoria de Martingais

Um martingal é uma sequência de variáveis aleatórias, isto é, um processo estocástico, no qual, em qualquer tempo, o valor esperado da próxima variável da sequência será igual ao valor observado no momento presente, mesmo dado o conhecimento de todas as variáveis anteriormente observadas.

Podemos pensar em um martingal $\{S_n, n \geq 1\}$ como a quantia que um jogador possui em tempo n num jogo justo, ou seja, o valor esperado de ganho ou perda será zero. A variável aleatória S_n é a soma em tempo n de variáveis $X_1, X_2, \ldots, X_n$, que representam o ganho ou perda em cada momento. Martingais excluem a possibilidade de estratégias de ganho baseadas no histórico do jogo e, portanto, são um modelo de jogos honestos.

Desta forma, como veremos adiante, o valor esperado futuro de S_n dadas as informações que possuímos até o presente momento será igual à quantia que possuímos no presente momento, visto que estamos em um jogo justo.

Ademais, existem ainda o supermartingal, que pode ser interpretado como um jogo desfavorável ao apostador e o submartingal, que seria um jogo favorável ao apostador.

Para darmos a definição formal de martingal será necessário apresentarmos o conceito de esperança condicional.

2.1
Esperança Condicional

Primeiramente, iremos definir o conceito de esperança condicional e, posteriormente, ilustraremos alguns exemplos.

Seja um espaço de probabilidade $(\Omega, \mathcal{F}_0, P)$, uma σ-álgebra $\mathcal{F} \subset \mathcal{F}_0$, e uma variável aleatória X $\mathcal{F}_0$-mensurável com $\mathbb{E}|X| < \infty$. Definimos a esperança condicional de X dado $\mathcal{F}$, $\mathbb{E}(X|\mathcal{F})$ como uma variável aleatória Y tal que:

(i) $Y \in \mathcal{F}$, isto é, Y é $\mathcal{F}$-mensurável.

(ii) $\forall\, \mathcal{A} \in \mathcal{F}$, temos que $\int_{\mathcal{A}} X dP = \int_{\mathcal{A}} Y dP$.

Qualquer variável aleatória Y que satisfaça as condições (i) e (ii) é uma versão de $\mathbb{E}(X|\mathcal{F})$. A esperança condicional existe e é única, a menos de mudanças em conjuntos de medida nula, e a prova de sua existência e unicidade pode ser vista em [13], página 190.

Intuitivamente, podemos pensar em $\mathcal{F}$ como uma descrição das informações que temos à nossa disposição, ou seja, para cada $\mathcal{A} \in \mathcal{F}$, sabemos se $\mathcal{A}$ ocorreu ou não. Deste modo, a variável aleatória $Y = \mathbb{E}(X|\mathcal{F})$ seria nossa "melhor estimativa" para o valor de X dado as informações que temos.

Por exemplo, se $X \in \mathcal{F}$, temos que $\mathbb{E}(X|\mathcal{F}) = X$, ou seja, a melhor estimativa da variável é a própria.

Também, quando $\mathcal{F}$ é a σ-álgebra trivial, $\mathcal{F} = \{\Omega, \emptyset\}$, $\mathbb{E}(X|\mathcal{F})$ será igual a $\mathbb{E}(X)$. Isto ocorre, pois como estamos no espaço de probabilidade $(\Omega, \mathcal{F}_0, P)$,

$$\mathbb{E}(X|\mathcal{F}) = \int_\Omega X dP = \mathbb{E}(X).$$

2.1.1 Propriedades da Esperança Condicional

A esperança condicional mantém muitas das propriedades da esperança convencional. Quando escrevemos que uma variável aleatória $X \in \mathcal{L}^k$, para $k \geq 1$, significa que X é k-integrável, isto é, $\mathbb{E}(|X|^k) < \infty$.

Proposição 2.1 *Sejam X e Y variáveis aleatórias tais que $X, Y \in \mathcal{L}^1$. Então,*

(a) A esperança condicional é linear:

$$\mathbb{E}(aX + Y|\mathcal{F}) = a\mathbb{E}(X|\mathcal{F}) + \mathbb{E}(Y|\mathcal{F}), \quad \forall\, a \in \mathbb{R}.$$

(b) Se $X \leq Y$, então

$$\mathbb{E}(X|\mathcal{F}) \leq \mathbb{E}(Y|\mathcal{F}).$$

(c) Se $X_n \geq 0$ e $X_n \uparrow X$ (esta notação indica convergencia monótona pontual, isto é, para cada $\omega \in \Omega$), então

$$\mathbb{E}(X_n|\mathcal{F}) \uparrow \mathbb{E}(X|\mathcal{F}).$$

(d) *(Propriedade da Torre) Se $\mathcal{G} \subset \mathcal{F}$, ou seja, $\mathcal{G}$ é uma sub σ-álgebra de $\mathcal{F}$, então*

$$\mathbb{E}[\mathbb{E}(X|\mathcal{F})|\mathcal{G}] = \mathbb{E}[\mathbb{E}(X|\mathcal{G})|\mathcal{F}] = \mathbb{E}(X|\mathcal{G}).$$

(e) *Se $Z \in \mathcal{F}$ e $\mathbb{E}|X|, \mathbb{E}|XZ| < \infty$, então*

$$\mathbb{E}(ZX|\mathcal{F}) = Z\mathbb{E}(X|\mathcal{F}).$$

(f) *(Regra da independência) Se $\mathcal{G}$ é independente de $\sigma(\sigma(X), \mathcal{F})$, então*

$$\mathbb{E}[X|\sigma(\mathcal{F}, \mathcal{G})] = \mathbb{E}(X|\mathcal{F}) \quad q.c.$$

Em particular, se X é independente de $\mathcal{G}$, então $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$.

(g) *(Lei da esperança total)*

$$\mathbb{E}[\mathbb{E}[X|\mathcal{F}]] = \mathbb{E}[X].$$

Prova. As demonstrações dos itens (a), (b), (c) e (e) podem ser encontradas em [13] páginas 193-195.

As demonstrações dos itens (d), (f) e (g) podem ser encontradas em [14], páginas 89-90.

■

2.2 Martingais

Nessa seção, definiremos o que são martingais, submartingais e supermartingais. Além disso, apresentaremos propriedades, principais teoremas e exemplos relacionados.

Seja $\mathcal{F}_n$ uma filtração, isto é, uma sequência crescente de σ-álgebras, $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}$. Se $\mathcal{F}_n$ for gerada por uma sequência de variáveis aleatórias $X_1, X_2, \ldots$, então chamamos $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$ para $n \in \mathbb{N}$ de filtração natural.

Uma sequência X_n é dita adaptada à $\mathcal{F}_n$ se X_n é $\mathcal{F}_n$-mensurável para todo n. Intuitivamente, sabemos o valor de $X_n(\omega)$ em cada tempo n.

Se X_n é uma sequência que satisfaz

(i) $\mathbb{E}|X_n| < \infty$,

(ii) X_n é adaptada à $\mathcal{F}_n$,

(iii) $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$ para todo $n \geq 1$,

então, o processo X_n é um martingal (com respeito à $\mathcal{F}_n$).

Um supermartingal é definido de maneira similar, porém, para todo $n \geq 1$, (iii) é substituído por

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq X_n$$

e, num submartingal, para todo $n \geq 1$, (iii) é substituído por

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \geq X_n.$$

Agora, daremos alguns exemplos de martingais e suas aplicações. Em cada caso definiremos o processo e verificaremos que esse é um martingal, pois atende às três condições supracitadas.

Exemplo 2.2 *(Passeio Aleatório Simples) Considere uma sequência de lançamentos de uma moeda justa e seja $X_n = 1$ se o n-ésimo lançamento for cara e $X_n = -1$ se o n-ésimo lançamento for coroa. Definimos $S_n = X_1 + X_2 + \cdots + X_n$, $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$ para todo $n \geq 1$, $S_0 = 0$ e $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Então, $S_n, n \geq 0$ é um martingal com respeito à $\mathcal{F}_n$. Para demonstrar isso, vamos ver que S_n respeita as três condições postas acima.*

(i) $\mathbb{E}|S_n| < \infty$.

Temos que $\mathbb{E}|S_n| = \mathbb{E}|X_1 + X_2 + \cdots + X_n| \leq \mathbb{E}|X_1| + \mathbb{E}|X_2| + \cdots + \mathbb{E}|X_n|$. *Como* $\mathbb{E}|X_i| < \infty$ *para todo* i, *obtemos que* $\mathbb{E}|S_n| < \infty$.

(ii) S_n *é adaptada à* $\mathcal{F}_n$.

X_n é adaptada à $\mathcal{F}_n$, ou seja, X_n é $\mathcal{F}_n$-mensurável para todo n. A função $f : \mathbb{R}^n \to \mathbb{R}$, $f(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \cdots + x_n$ é contínua e, portanto, Borel mensurável. Logo, $S_n = f(X_1, X_2, \ldots, X_n)$ é $\mathcal{F}_n$-mensurável para todo n, isto é, S_n é adaptada à $\mathcal{F}_n$.

(iii) $\mathbb{E}(S_{n+1}|\mathcal{F}_n) = S_n$ *para todo* $n \geq 1$.

Como X_{n+1} é independente de $\mathcal{F}_n$, e utilizando a propriedade de linearidade da esperança condicional, Proposição 2.1(a) , temos que

$$\mathbb{E}(S_{n+1}|\mathcal{F}_n) = \mathbb{E}(S_n|\mathcal{F}_n) + \mathbb{E}(X_{n+1}|\mathcal{F}_n) = S_n + \mathbb{E}(X_{n+1}) = S_n,$$

pois $\mathbb{E}(X_{n+1}) = 0$.

Exemplo 2.3 *Seja* $X_1, X_2, \ldots$ *uma sequência de variáveis aleatórias não-negativas independentes com* $\mathbb{E}(X_i) = 1, \forall\, i$. *Definimos* $M_n = X_1X_2\ldots X_n$, $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$ *para todo* $n \geq 1$, $M_0 = 0$ *e* $\mathcal{F}_0 = \{\emptyset, \Omega\}$. *Então,* M_n, $n \geq 0$ *é um martingal com respeito à* $\mathcal{F}_n$. *Para demonstrar isso, vamos novamente ver que* M_n *respeita as três condições postas acima.*

(i) $\mathbb{E}|M_n| < \infty$.

Sabemos que M_n *é uma multiplição de variáveis aleatórias não-negativas, portanto* $\mathbb{E}|M_n| = \mathbb{E}|X_1X_2\ldots X_n| = \mathbb{E}(X_1X_2\ldots X_n) = \mathbb{E}(X_1)\mathbb{E}(X_2)\ldots\mathbb{E}(X_n)$, *pois* $X_1, X_2, \ldots, X_n$ *são independentes. Logo,* $\mathbb{E}|M_n| = 1 < \infty$ *para todo* n.

(ii) M_n *é adaptada à* $\mathcal{F}_n$.

X_n *é adaptada à* $\mathcal{F}_n$, *ou seja,* X_n *é* $\mathcal{F}_n$*-mensurável para todo* n. *A função* $f : \mathbb{R}^n \to \mathbb{R}$, $f(x_1, x_2, \ldots, x_n) = x_1x_2\ldots x_n$ *é contínua e, portanto, Borel mensurável. Logo,* $M_n = f(X_1, X_2, \ldots, X_n)$ *é* $\mathcal{F}_n$*-mensurável para todo* n, *isto é,* M_n *é adaptada à* $\mathcal{F}_n$.

(iii) $\mathbb{E}(M_{n+1}|\mathcal{F}_n) = M_n$ *para todo* $n \geq 1$.

$\mathbb{E}(M_{n+1}|\mathcal{F}_n) = \mathbb{E}(X_{n+1}M_n|\mathcal{F}_n) = M_n\mathbb{E}(X_{n+1}|\mathcal{F}_n)$, *pela Proposição 2.1(e). Como* $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = \mathbb{E}(X_{n+1}) = 1$, *temos que* $\mathbb{E}(M_{n+1}|\mathcal{F}_n) = M_n$.

Exemplo 2.4 *(Urna de Polya) Considere uma urna que contém uma bola branca e uma bola preta. A cada instante, sorteamos uniformemente uma bola da urna, observamos sua cor, a devolvemos à urna, adicionamos uma nova bola da cor observada e repetimos o processo. Após n instantes, teremos n+2 bolas. Nesse exemplo, chamaremos de* X_n *a proporção de bolas pretas na urna, e iremos concluir que se trata de um martingal. A filtração será a natural, ou seja,* $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$.

(i) $\mathbb{E}|X_n| < \infty$.

Trivial, pois X_n *é uma proporção.*

(ii) X_n *é adaptada à* $\mathcal{F}_n$.

Também trivial, pois a filtração é natural.

(iii) $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$ *para todo* $n \geq 1$.

Denominaremos P_n *e* B_n *o número de bolas pretas e brancas na urna no instante* n, *respectivamente. Naturalmente,* $B_n = (n+2) - P_n$. *Estamos* *interessados na proporção* $X_n = \frac{P_n}{n+2}$*. Observe que, dado* P_n*, o número de bolas pretas no instante* $n+1$ *só poderá ser:*

$$P_{n+1} = \begin{cases} P_n \text{ , com probabilidade } \frac{(n+2)-P_n}{n+2} \\ P_n + 1, \text{ com probabilidade } \frac{P_n}{n+2} \end{cases}$$

Portanto, teremos

$$\begin{aligned} \mathbb{E}(P_{n+1}|\mathcal{F}_n) &= (P_n+1)\frac{P_n}{n+2} + P_n\frac{(n+2)-P_n}{n+2} \\ &= \frac{P_n^2 + P_n - P_n^2 + P_n(n+2)}{n+2} \\ &= \frac{P_n(n+3)}{n+2}. \end{aligned}$$

Desta forma,

$$\begin{aligned} \mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \mathbb{E}\left(\frac{P_{n+1}}{n+3}\,\middle|\,\mathcal{F}_n\right) \\ &= \frac{1}{n+3}\mathbb{E}(P_{n+1}|\mathcal{F}_n) \\ &= \frac{1}{n+3}\frac{P_n(n+3)}{n+2} \\ &= \frac{P_n}{n+2} \\ &= X_n. \end{aligned}$$

Exemplo 2.5 *(Processo de Galton-Watson) Um processo de Galton-Watson é um processo estocástico de ramificação, onde* Z_n *é o número de indivíduos da geração em tempo n e* $X_{n,i}$ *o número de descendentes do indivíduo i em tempo n. Formalmente, seja* $X_{n,i}$ *uma sequência não negativa de variáveis aleatórias independentes e identicamente distribuídas e* $X_{n,i} \in \mathcal{L}^1$ *para todo* $i, n \geq 1$*. Temos também a sequência* Z_n*, na qual* $n \geq 0$, $Z_0 = 1$ *e*

$$Z_{n+1} = \begin{cases} X_{n+1,i} + \cdots + X_{n+1,Z_n}, \text{ se } Z_n > 0 \\ 0 \qquad\qquad\qquad\qquad\quad\ \text{, se } Z_n = 0 \end{cases}$$

Neste caso, vamos analisar a sequência $(Y_n)_{n\geq 1}$ *definida por* $Y_n = \frac{Z_n}{\mu^n}$*, onde* $\mu = \mathbb{E}(X_{n,i})$ *e faremos* $\mathcal{F}_n = \sigma(X_{j,i} : i \geq 1, 1 \leq j \leq n)$

(i) $\mathbb{E}|Y_n| < \infty$.

A variável aleatória Y_n *é não negativa, pois* $\mu \geq 0$*, então* $\mathbb{E}|Y_n| = \mathbb{E}(Y_n)$.

$$\mathbb{E}(Y_n) = \mathbb{E}\left(\frac{Z_n}{\mu^n}\right) = \frac{1}{\mu^n}\mathbb{E}(Z_n)$$

Sabemos que $Z_n = 0$ *ou* $Z_n = X_{n,1} + \cdots + X_{n,Z_{n-1}}$ *. Para o primeiro caso, temos* $\mathbb{E}(Y_n) = 0$ *. Para o segundo caso,*

$$\begin{aligned}
\mathbb{E}(Y_n) &= \frac{1}{\mu^n}\mathbb{E}(Z_n) \\
&= \frac{1}{\mu^n}\mathbb{E}(X_{n,1} + \cdots + X_{n,Z_{n-1}}) \\
&= \frac{1}{\mu^n}[\mathbb{E}(X_{n,1}) + \cdots + \mathbb{E}(X_{n,Z_{n-1}})].
\end{aligned}$$

Como $X_{n,i} \in \mathcal{L}^1$ *para todo* $i, n \geq 1$*, então* $\mathbb{E}|Y_n| < \infty$ *.*

(ii) Y_n *é adaptado a* $\mathcal{F}_n$*.*

Sabemos que $Z_n \in \mathcal{F}_n$ *para todo n. Como* $Y_n = \frac{Z_n}{\mu^n}$ *e* μ^n *é uma constante, então* $Y_n \in \mathcal{F}_n$ *para todo n, isto é,* Y_n *é adaptado a* $\mathcal{F}_n$*.*

(iii) $\mathbb{E}(Y_{n+1}|\mathcal{F}_n) = Y_n$ *para todo n.*

Temos que

$$\mathbb{E}(Z_{n+1}|\mathcal{F}_n) = \sum_{k=1}^{\infty} \mathbb{E}(Z_{n+1}\mathbb{1}_{Z_n=k}|\mathcal{F}_n)$$

pelas Propriedades 2.1(a) e 2.1(c). Em $Z_n = k$*, fazemos* $Z_{n+1} = X_{n+1,1} + \cdots + X_{n+1,k}$*, e obtemos*

$$\sum_{k=1}^{\infty} \mathbb{E}(Z_{n+1}\mathbb{1}_{Z_n=k}|\mathcal{F}_n) = \sum_{k=1}^{\infty} \mathbb{E}((X_{n+1,1} + \cdots + X_{n+1,k})\mathbb{1}_{Z_n=k}|\mathcal{F}_n).$$

Pela Propriedade 2.1(e) temos

$$\sum_{k=1}^{\infty} \mathbb{E}((X_{n+1,1}+\cdots+X_{n+1,k})\mathbb{1}_{Z_n=k}|\mathcal{F}_n) = \sum_{k=1}^{\infty} \mathbb{1}_{Z_n=k}\mathbb{E}(X_{n+1,1}+\cdots+X_{n+1,k}|\mathcal{F}_n).$$

Como $X_{k+1,i}$ *é independente de* $\mathcal{F}_n$ *para todo i, temos finalmente*

$$\begin{aligned}
\mathbb{E}(Z_{n+1}|\mathcal{F}_n) &= \sum_{k=1}^{\infty} \mathbb{1}_{Z_n=k}k\mu \\
&= \mu Z_n.
\end{aligned}$$

Dividindo ambos os lados da equação por μ^{n+1}, encontramos

$$\begin{aligned}\mathbb{E}\left(\frac{Z_{n+1}}{\mu^{n+1}}|\mathcal{F}_n\right) &= \mathbb{E}(Y_{n+1}|\mathcal{F}_n)\\ &= \frac{Z_n}{\mu^n}\\ &= Y_n.\end{aligned}$$

Logo, $\mathbb{E}(Y_{n+1}|\mathcal{F}_n) = Y_n$ e concluímos que Y_n é um martingal.

2.2.1 Martingais como série de incrementos

No exemplo 2.2, vimos que S_n é um martingal. Além disso, temos que $X_n = S_n - S_{n-1}$ para todo $n \geq 1$. Podemos, portanto, reformular esse martingal a partir de seus incrementos, pois dada uma filtração $\mathcal{F}_n$, estes possuem média zero:

$$\begin{aligned}\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \mathbb{E}(S_{n+1} - S_n|\mathcal{F}_n)\\ &= \mathbb{E}(S_{n+1}|\mathcal{F}_n) - \mathbb{E}(S_n|\mathcal{F}_n)\\ &= S_n - S_n\\ &= 0.\end{aligned}$$

Podemos generalizar a ideia acima para qualquer martingal Y_n, representando-o através de uma sequência de variáveis aleatórias X_1, X_2, $\dots$, tal que $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = 0$ para todo n, onde $\mathcal{F}_n$ é uma filtração. Estas variáveis aleatórias são os incrementos do martingal, isto é, $X_{n+1} = Y_{n+1} - Y_n$.

Com esse fato, poderemos definir a variância condicional para martingais e a variância quadrática normalizada no Capítulo 6.

2.2.2 Tempo de parada

Definição 2.6 *(Tempo de Parada) Dado um espaço de probabilidade $(\Omega, \mathcal{F}, P)$, uma filtração $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}$ e uma variável aleatória τ, dizemos que τ é um tempo de parada com respeito à filtração $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}$ se $\{\tau \leq n\} = \{\omega : \tau(\omega) \leq n\} \in \mathcal{F}_n$ para todo $n \geq 1$.*

Intuitivamente, o tempo de parada τ é o momento em que decidimos interromper o processo estocástico, ou o "jogo justo" que mencionamos. A decisão de parar ou não após o n-ésimo momento depende apenas do eventos passados até (e incluindo) o tempo n.

3
Resultados Auxiliares

Neste capítulo apresentaremos resultados básicos que serão utilizados ao longo dessa dissertação. Inicialmente, iremos introduzir conceitos que serão úteis a esses resultados e explorados em maior detalhes posteriormente.

Um martingal pode ser definido em termos dos seus incrementos, por isso será útil introduzirmos a seguinte definição.

Seja M_n definido como o conjunto de todas as sequências $\mathbf{X} = (X_1, \dots, X_n)$ de variáveis aleatórias que satisfazem

(i) X_i é quadrado-integrável para todo i;

(ii) $\mathbb{E}[X_i|\mathcal{F}_{i-1}] = 0$ para todo i, onde $\mathcal{F}_i$ é a σ-álgebra gerada por $(X_1, \dots, X_i)$.

Também, introduziremos as seguintes notações:

$$s^2(\mathbf{X}) = \sum_{i=1}^{n} \mathbb{E}[X_i^2],$$

$$V^2(\mathbf{X}) = s^{-2}(\mathbf{X}) \sum_{i=1}^{n} \mathbb{E}[X_i^2|\mathcal{F}_{i-1}],$$

$$S(\mathbf{X}) = \sum_{i=1}^{n} X_i.$$

Chamamos $V^2(\mathbf{X})$ de variação quadrática normalizada de $\mathbf{X}$. Finalmente, definiremos a distância que representa a taxa em que a convergência entre a soma padronizada de variáveis e a distribuição Normal ocorre. Escrevemos

$$D(\mathbf{X}) = \sup_{t \in \mathbb{R}} |\mathbb{P}\left[S(\mathbf{X})/s(\mathbf{X}) \leq t\right] - \Phi(t)|. \tag{3-1}$$

Apresentaremos agora um Lema que estabelece que D e V^2 são invariantes por dilatações.

Lema 3.1 *Seja* $\mathbf{Y} = (Y_1, \dots, Y_n) \in M_n$ *e* $\lambda \in (0, \infty)$. *Então,*

(i) $D(\mathbf{Y}) = D(\lambda \mathbf{Y})$

(ii) $V^2(\mathbf{Y}) = V^2(\lambda\mathbf{Y})$.

Prova.

(i)

$$\begin{aligned} D(\mathbf{Y}) &= \sup_{t\in\mathbb{R}} |[S(\mathbf{Y})/s(\mathbf{Y}) \le t] - \Phi(t)| \\ &= \sup_{t\in\mathbb{R}} |[\lambda S(\mathbf{Y})/\lambda s(\mathbf{Y}) \le t] - \Phi(t)| \\ &= D(\lambda\mathbf{Y}) \quad \forall\, \lambda > 0. \end{aligned}$$

(ii)

$$\begin{aligned} V^2(\lambda\mathbf{Y}) &= s^{-2}(\lambda\mathbf{Y}) \sum_{i=1}^{n} \mathbb{E}[(\lambda Y_i)^2|\mathcal{F}_{i-1}] \\ &= \lambda^{-2} s^{-2}(\mathbf{Y}) \lambda^2 \sum_{i=1}^{n} \mathbb{E}[Y_i^2|\mathcal{F}_{i-1}] \\ &= V^2(\mathbf{Y}) \quad \forall\, \lambda > 0. \end{aligned}$$

■

Agora, iremos apresentar um resultado bem conhecido sobre variância de martingais.

Lema 3.2 *Seja $\mathbf{X} = (X_1, \ldots, X_n) \in M_n$ e $S_n = \sum_{i=1}^n X_i$. Então,*

$$Var(S_n) = \mathbb{E}\Big[\sum_{i=1}^{n} \mathbb{E}\Big[X_i^2|\mathcal{F}_{i-1}\Big]\Big] = s^2(\mathbf{X}).$$

Prova. Primeiramente, por linearidade da esperança, $\mathbb{E}[S_n] = \sum_{i=1}^n \mathbb{E}[X_i] = 0$, uma vez que $\mathbb{E}[X_i] = \mathbb{E}[\mathbb{E}[X_i|\mathcal{F}_{i-1}]] = 0$, então temos que

$$\begin{aligned} Var(S_n) &= \mathbb{E}[S_n^2] \\ &= \sum_{i=1}^{n} \mathbb{E}\Big[S_i^2 - S_{i-1}^2\Big] \\ &= \sum_{i=1}^{n} \mathbb{E}\Big[\mathbb{E}\Big[S_i^2 - S_{i-1}^2|\mathcal{F}_{i-1}\Big]\Big] \\ &= \mathbb{E}\Big[\sum_{i=1}^{n} 2S_{i-1}\mathbb{E}\Big[X_i|\mathcal{F}_{i-1}\Big] + \mathbb{E}\Big[X_i^2|\mathcal{F}_{i-1}\Big]\Big], \end{aligned}$$

pois $S_i^2 - S_{i-1}^2 = (S_{i-1} + X_i)^2 - S_{i-1}^2 = 2S_{i-1}X_i + X_i^2$ e $S_{i-1} \in \mathcal{F}_{i-1}$. Portanto, como $\mathbb{E}[X_i|\mathcal{F}_{i-1}] = 0$, conclui-se que

$$Var(S_n) = \mathbb{E}\Big[\sum_{i=1}^{n}\mathbb{E}\Big[X_i^2|\mathcal{F}_{i-1}\Big]\Big] = \sum_{i=1}^{n}\mathbb{E}\Big[\mathbb{E}\Big[X_i^2|\mathcal{F}_{i-1}\Big]\Big] = s^2(\mathbf{X}).$$

■

Os próximos lemas serão úteis para demonstrações relativas ao Capítulo 4, sobre Teorema Central do Limite

Lema 3.3 *Sejam* $c_1, c_2, \dots,$ *e* c *números complexos tais que* $\lim_{n\to\infty} c_n = c$. *Então*

$$\lim_{n\to\infty}(1 + \frac{c_n}{n})^n = e^c.$$

Prova. O que estamos interessados em provar pode ser escrito da seguinte forma:

$$\lim_{n\to\infty}\frac{(1+\frac{c_n}{n})^n}{e^c} = 1.$$

Temos que, se $a_n \to 0$ quando $n \to \infty$, então $\frac{1+a_n}{e^{a_n}} \to 1$ quando $n \to \infty$. De fato, por quociente de limites,

$$\lim_{x\to 0}\frac{1+x}{e^x} = \frac{\lim_{x\to 0} 1+x}{\lim_{x\to 0} e^x} = 1.$$

Portanto, sendo $\log z$ o logaritmo pricipal de z, definido para $z \neq 0$, escrevemos

$$\begin{aligned}\lim_{n\to\infty}(1+\frac{c_n}{n})^n &= \lim_{n\to\infty}\exp\left(\frac{\log(1+c_n/n)}{c_n/n}c_n\right)\\ &= \lim_{n\to\infty} e^{c_n}\\ &= e^c.\end{aligned}$$

Usamos o fato que $\lim_{z\to 0}\frac{\log(1+z)}{z} = 1$ e que $\lim_{n\to\infty} c_n = c$ para concluir. ■

Lema 3.4 *Sejam* $c_{n,k}$ *números complexos tais que* $\sum_{k=1}^{n} c_{n,k} \to c$ *quando* $n \to \infty$, *onde* c *é um número complexo finito. Se quando* $n \to \infty$, $\max_{1\leq k\leq n}|c_{n,k}| \to 0$ *e* $\sum_{k=1}^{n}|c_{n,k}| \leq M < \infty$, *onde* M *é uma constante que não depende de* n, *então*

$$\prod_{k=1}^{n}(1+c_{n,k}) \to e^c, \quad \textit{quando } n\to\infty.$$

Prova. A demonstração deste lema pode ser vista em [12], página 208. ■

Enunciaremos agora uma desigualdade em forma de proposição, que utilizaremos no capítulo 6.

Proposição 3.5 *Se $\{S_i, \mathcal{F}_i, 1 \leq i \leq n\}$ é um martingal, $X_i = S_i - S_{i-1}$ são os incrementos do martingal e $k > 0$, então existe uma constante C que depende somente de k tal que*

$$\mathbb{E}\Big(\max_{i\leq n}|S_i|^k\Big) \leq C\Bigg\{\mathbb{E}\Bigg[\Bigg(\sum_{i=1}^{n}\mathbb{E}(X_i^2|\mathcal{F}_{i-1})\Bigg)^{k/2}\Bigg] + \mathbb{E}\Big(\max_{i\leq n}|X_i|^k\Big)\Bigg\}.$$

Prova. A demonstração está em [17], página 28. ■

A seguir, apresentaremos resultados que serão úteis para as demonstrações finais do último capítulo e do Teorema 1.2.

Lema 3.6 *Sejam Y variável aleatória, $\hat{F}_Y(x) = F_Y(x\sigma_Y)$, onde $F_Y(y) = \mathbb{P}(Y \leq y)$, $\delta > 0$ e $a, b \in \mathbb{R}$. Então,*

$$\Big|\hat{F}_Y(a) - \hat{F}_Y(b)\Big| \leq \Big|\Phi(a) - \Phi(b)\Big| + 2d(Y) \leq |a-b| + 2d(Y).$$

Além disso, se $(1-\delta)a \leq b \leq (1+\delta)a$, então,

$$\Big|\hat{F}_Y(a) - \hat{F}_Y(b)\Big| \leq \delta + 2d(Y).$$

Prova. Inicialmente,

$$\begin{aligned}\Big|\hat{\mathrm{F}}_Y(a) - \hat{\mathrm{F}}_Y(b)\Big| &= \Big|(\hat{\mathrm{F}}_Y(a) - \Phi(a)) + \Phi(a) - \Phi(b) + (\Phi(b) - \hat{\mathrm{F}}_Y(b))\Big| \\ &\leq \Big|\Phi(a) - \Phi(b)\Big| + 2d(Y) \\ &\leq |a-b| + 2d(Y),\end{aligned}$$

pois $\frac{1}{\sqrt{2\pi}}\int_a^b exp(\frac{-x^2}{2})dx \leq \frac{b-a}{\sqrt{2\pi}}$.

Além disso, caso $a \leq b \leq (1+\delta)a$, então

$$\frac{1}{\sqrt{2\pi}}\int_a^{(1+\delta)a} exp(\frac{-x^2}{2})dx \leq \frac{exp(-a^2/2)a\delta}{\sqrt{2\pi}} \leq \delta,$$

uma vez que $exp(-a^2/2)a \leq \sqrt{2\pi}$ para todo a. O caso onde $(1-\delta)a \leq b \leq a$ é análogo. ■

Lema 3.7 *Sejam Y, Y' variáveis aleatórias e $\epsilon \geq 0$ tais que $\mathbb{P}(|Y - Y'| > \epsilon\sigma_Y) \leq \delta$ e $\Big|\sigma_Y - \sigma_{Y'}\Big| \leq \delta\sigma_{Y'}$. Então,*

$$d(Y) \leq 3d(Y') + 2\delta + \epsilon.$$

Prova. Observamos que

$$\mathbb{P}(|Y - Y'| > \epsilon\sigma_Y) \leq \delta \Rightarrow F_Y(a) \leq F_{Y'}(a + \epsilon\sigma_Y) + \delta.$$

Isso é verdade, pois

$$Y \leq a \subseteq (Y, Y' \leq a + \epsilon\sigma_Y) \cup (Y \leq a, Y' > a + \epsilon\sigma_Y).$$

Portanto,

$$\mathbb{P}(Y \leq a) \leq \mathbb{P}(Y' \leq a + \epsilon\sigma_Y) + \mathbb{P}(|Y - Y'| > \epsilon\sigma_Y).$$

Assim,

$$\begin{aligned}
\hat{\mathrm{F}}_Y(a) &= F_Y(a\sigma_Y) \\
&\leq F_{Y'}(\sigma_Y(a + \epsilon)) + \delta \\
&= F_{Y'}((a + \epsilon)\sigma_{Y'}) + (F_{Y'}((a + \epsilon)\sigma_Y) - F_{Y'}((a + \epsilon)\sigma_{Y'})) + \delta \\
&= \hat{\mathrm{F}}_{Y'}(a + \epsilon) + (\hat{\mathrm{F}}_{Y'}((a + \epsilon)\frac{\sigma_Y}{\sigma_{Y'}}) - \hat{\mathrm{F}}_{Y'}(a + \epsilon)) + \delta.
\end{aligned}$$

Pelo lema 3.6, temos que

$$\left|\hat{\mathrm{F}}_{Y'}((a + \epsilon)\frac{\sigma_Y}{\sigma_{Y'}}) - \hat{\mathrm{F}}_{Y'}(a + \epsilon)\right| \leq \delta + 2d(Y').$$

Além disso, observamos que

$$\begin{aligned}
|\hat{\mathrm{F}}_Y(a) - \Phi(a)| &\leq |\hat{\mathrm{F}}_Y(a) - \hat{\mathrm{F}}_{Y'}(a + \epsilon)| + |\hat{\mathrm{F}}_{Y'}(a + \epsilon) - \Phi(a + \epsilon)| + |\Phi(a + \epsilon) - \Phi(a)| \\
&\leq |2d(Y') + 2\delta| + d(Y') + \epsilon \\
&= 3d(Y') + 2\delta + \epsilon.
\end{aligned}$$

O outro lado da desigualdade é análogo. Logo, concluímos que

$$d(Y) \leq 3d(Y') + 2\delta + \epsilon.$$

■

Corolário 3.8 *Sejam* $\mathbf{Y} = (Y_1, \ldots, Y_m), \mathbf{Y}' = (Y'_1, \ldots, Y'_m) \in M_m$ *tais que* $\mathbb{P}(\exists i \leq m,\, Y_i \neq Y'_i) \leq \delta$ *e* $||\mathbf{Y}||_\infty, ||\mathbf{Y}'||_\infty \leq T$*, então,*

$$D(\mathbf{Y}) \leq 3D(\mathbf{Y}') + 2Tm\delta^{1/2} + m\delta.$$

Prova. Como $||\mathbf{Y}||_\infty, ||\mathbf{Y}'||_\infty \leq T$, então

$$\left|\sum_{i=1}^{m} Y_i - \sum_{i=1}^{m} Y_i'\right| \leq 2Tm.$$

Desse modo,

$$\begin{aligned}\left|\left\|\sum_{i=1}^{m} Y_i\right\|_2 - \left\|\sum_{i=1}^{m} Y_i'\right\|_2\right| &\leq \left\|\sum_{i=1}^{m} Y_i - \sum_{i=1}^{m} Y_i'\right\|_2 \\ &\leq \left\|2Tm\mathbb{1}_{\sum_{i=1}^{m} Y_i \neq \sum_{i=1}^{m} Y_i'}\right\|_2 \\ &= (4T^2m^2\delta)^{1/2} \\ &= 2Tm\delta^{1/2}.\end{aligned}$$

Esta desigualdade é utilizada para controlar as condições da variância do Lema 3.7, para que possamos aplicá-lo. Portanto,

$$\begin{aligned}D(\mathbf{Y}) &\leq 3D(\mathbf{Y}') + 2Tm\delta^{1/2} + \sum_{i=1}^{m} \mathbb{P}(Y_i \neq Y_i') \\ &\leq 3D(\mathbf{Y}') + 2Tm\delta^{1/2} + m\delta.\end{aligned}$$

■

4
Teorema Central do Limite

O Teorema Central do Limite nos remete à convergência de somas de variáveis aleatórias para uma distribuição normal e é considerado, pela sua importância na teoria e em aplicações, como o teorema básico mais central da probabilidade.

Neste capítulo iremos apresentar o Teorema Central do Limite para variáveis aleatórias. Apresentaremos e provaremos duas de suas versões, a para variáveis aleatórias independentes e identicamente distribuídas e a versão de Lindeberg.

Para que possamos discutir mais profundamente o Teorema Central do Limite e suas demonstrações, necessitamos de dois tópicos importantes: as funções características e a convergência em distribuição.

Intuitivamente, o Teorema Central do Limite diz que a soma de variáveis aleatórias independentes, com média e variância finitas, após uma padronização, tenderá a uma distribuição normal. Em outras palavras, qualquer que seja a distribuição da variável de interesse para grandes amostras, a distribuição das médias amostrais será aproximadamente normalmente distribuída, e tenderá a uma distribuição normal à medida que o tamanho da amostra crescer. Quando dizemos "tenderá", isso significa que haverá uma convergência em distribuição.

Inicialmente, veremos o Teorema Central do Limite para variáveis aleatórias independentes e identicamente distribuídas, e em seguida exploraremos a versão de Lindeberg. Uma diferença entre essas versões é que na primeira, para que ocorra a convergência da soma das variáveis aleatórias para a função de distribuição da normal padrão, após a padronização, é necessário, como já está no enunciado do teorema, que as variáveis sejam identicamente distribuídas. Já na versão de Lindeberg, basta que a condição de Lindeberg seja satisfeita para que ocorra a convergência supracitada.

4.1 Função Característica

Começaremos essa seção definindo a função característica, que será de extrema importância para demonstrações a respeito do Teorema Central do Limite.

Em teoria da probabilidade e estatística, a função característica de qualquer variável aleatória é útil para se fazer análises sobre sua função de densidade e de distribuição. Além disso, para cada variável aleatória, existe uma única função característica que a define.

Apresentaremos agora a definição formal de funções características.

Definição 4.1 *Seja X uma variável aleatória. Definimos sua função característica como uma função $\varphi : \mathbb{R} \to \mathbb{C}$ tal que*

$$\varphi_X(t) = \mathbb{E}(exp(itX)) = \mathbb{E}(cos(tX)) + i\mathbb{E}(sen(tX)), \quad t \in \mathbb{R}.$$

Da fórmula de Euler, temos que $exp(iy) = cos(y) + isen(y)$, $y \in \mathbb{R}$, e as funções $sen(y)$ e $cos(y)$ são limitadas, portanto a variável aleatória complexa $exp(itX)$ possui esperança finita para qualquer variável aleatória X. Assim, garantimos que a função característica está bem definida.

Além disso, como a função característica é determinada pela sua função de distribuição, temos que se X e Y são identicamente distribuídos então $\varphi_X = \varphi_Y$.

Mostraremos agora algumas propriedades importantes da função característica. Sejam X e Y variáveis aleatórias.

Teorema 4.2 *Todas as funções características possuem as seguintes propriedades:*

(a) $\varphi(0) = 1$.

(b) $\varphi(-t) = \overline{\varphi(t)}$.

(c) $|\varphi(t)| = |\mathbb{E}[e^{itX}]| \leq \mathbb{E}|e^{itX}| = 1$.

(d) $|\varphi(t+h) - \varphi(t)| \leq \mathbb{E}|e^{-ihX} - 1|$, *logo φ é uniformemente contínua em* $(-\infty, \infty)$.

(e) $\mathbb{E}[e^{it(aX+b)}] = e^{itb}\varphi(at)$.

Prova. As demonstrações dos itens podem ser encontradas em [13], página 126. ∎

Teorema 4.3 *Se X e Y são independentes e possuem funções características φ_X e φ_Y, respectivamente, então $X + Y$ tem função característica $\varphi_X(t)\varphi_Y(t)$.*

Prova.

$$\begin{aligned}\varphi_{X+Y}(t) &= \mathbb{E}[e^{it(X+Y)}]\\ &= \mathbb{E}[e^{itX}e^{itY}]\\ &= \mathbb{E}[e^{itX}]\mathbb{E}[e^{itY}]\\ &= \varphi_X(t)\varphi_Y(t),\end{aligned}$$

pois e^{itX} e e^{itY} são independentes. ∎

O Teorema 4.3 pode ser generalizado para uma sequência de variáveis aleatórias $X_1, X_2, \ldots, X_n$ independentes através de indução. Desta forma, teremos

$$\varphi_{X_1+\cdots+X_n}(t) = \prod_{i=1}^{n} \varphi_{X_i}(t)\,, \quad \forall\, t \in \mathbb{R}.$$

Exemplo 4.4 *Dada uma variável aleatória X com distribuição normal padrão, sua função característica é dada por*

$$\varphi_X(t) = exp\left(\frac{-t^2}{2}\right), \quad \forall\, t \in \mathbb{R}.$$

O cálculo deste exemplo está em [13], página 92.

Como já vimos, por definição, a função de distribuição de uma variável aleatória determina sua função característica. Apresentaremos agora a fórmula da inversão, que nos dará a recíproca, isto é, que a função característica de uma variável aleatória determina sua função de distribuição.

Teorema 4.5 *Seja X uma variável aleatória qualquer. Então, sua função característica $\varphi_X(t)$ determina a função de distribuição de X, através da seguinte fórmula de inversão*

$$\widetilde{F}_X(b) - \widetilde{F}_X(a) = \lim_{c\to\infty} \frac{1}{2\pi}\int_{-c}^{c} \frac{e^{-iat} - e^{-ibt}}{it}\varphi_X(t)dt\,,$$

sendo $\widetilde{F}(w) = \frac{1}{2}[F(w) + F(w^-)]$*, onde* $F(w^-) = \lim_{x\to w^-} F(x)$*, para todo* $w \in \mathbb{R}$*, e os números reais* a*,* b *e* c *tais que* $c > 0$ *e* $a < b$*.*

Prova. A demonstração deste teorema pode ser vista em [15], página 282. ■

Com a fórmula da inversão podemos provar o Teorema da Unicidade que nos garantirá que variáveis aleatórias possuindo a mesma função característica terão a mesma função de distribuição.

Teorema 4.6 *(Teorema da Unicidade) Se as variáveis aleatórias* X *e* Y *têm a mesma função característica, então elas possuem a mesma função de distribuição.*

Prova. Sabemos que X e Y possuem a mesma função característica, e pela fórmula de inversão, para quaisquer a e b reais e $a < b$,

$$\widetilde{F}_X(b) - \widetilde{F}_X(a) = \widetilde{F}_Y(b) - \widetilde{F}_Y(a).$$

Para $a \to -\infty$, temos que $\widetilde{F}_X(a) \to 0$ e $\widetilde{F}_Y(a) \to 0$ e a igualdade acima será

$$\widetilde{F}_X(b) = \widetilde{F}_Y(b)\,, \quad \forall\, b \in \mathbb{R}.$$

Consideremos agora $c \in \mathbb{R}$, tal que $c < b$. Pela definição de $\widetilde{F}(.)$, temos

$$F_X(c) \leq \widetilde{F}_X(b) \leq F_X(b) \;\text{ e }\; F_Y(c) \leq \widetilde{F}_Y(b) \leq F_Y(b).$$

A função de distribuição é contínua à direita, portanto tomamos o limite para $b \downarrow c$ e obtemos

$$\lim_{b\downarrow c} \widetilde{F}_X(b) = F_X(c) \text{ e } \lim_{b\downarrow c} \widetilde{F}_Y(b) = F_Y(c).$$

Isto implicará $F_X(c) = F_Y(c)$. Como o argumento vale para qualquer c real, podemos concluir que as funções de distribuição de X e Y são iguais. ■

4.2 Convergência em Distribuição

Nesta seção, vamos discutir algumas características da convergência em distribuição ou convergência fraca das medidas de probabilidades.

Definição 4.7 *Sejam $X, X_1, X_2, \ldots$ variáveis aleatórias com funções de distribuição $F, F_1, F_2, \ldots$, respectivamente. Dizemos que X_n converge em distribuição para X, quando $n \to \infty$, se $F_n(x) \to F(x)$ para todo ponto de continuidade x de F.*

A notação que utilizaremos para esta convergência será $X_n \xrightarrow{D} X$.

A convergência em distribuição de $F_1, F_2, \ldots$ para F, isto é, $F_n \xrightarrow{D} F$, significa que estas funções acumuladas convergem fracamente para F.

Agora, apresentaremos dois teoremas que relacionarão a convergência em distribuição com as funções características.

Teorema 4.8 *(Teorema de Helly-Bray) Sejam $F, F_1, F_2, \ldots$ funções de distribuição. Se F_n converge fracamente para F, então para toda função $g : \mathbb{R} \to \mathbb{R}$ contínua e limitada, temos que*

$$\int g(x) dF_n(x) \xrightarrow[n\to\infty]{} \int g(x) dF(x).$$

Prova. A demonstração deste teorema pode ser vista em [16], página 232. ∎

Teorema 4.9 *(Teorema da Continuidade de Paul Lévy) Sejam $F_1, F_2, \ldots$ funções de distribuição e $\varphi_1, \varphi_2, \ldots$, respectivamente, suas funções características. Se φ_n converge pontualmente para um limite φ e se φ é contínua no ponto zero, então*

(i) existe uma função de distribuição F tal que $F_n \to F$ fracamente e

(ii) φ é a função característica de F.

Prova. A demonstração deste teorema pode ser vista em [16], página 234. ∎

Como vimos no Teorema 4.6 , há uma relação biunívoca entre a função de distribuição e a função característica. Com os dois últimos teoremas estabelecemos uma relação de equivalência entre a convergência em distribuição e a convergência pontual das respectivas funções características. Deste modo, conseguimos provar que

$$X_n \xrightarrow{D} X \iff \varphi_{X_n}(t) \to \varphi_X(t)\,, \quad \forall\, t \in \mathbb{R}.$$

4.3 Teorema Central do Limite para Variáveis Aleatórias

O Teorema Central do Limite enuncia que a soma de uma sequência de variáveis aleatórias independentes, se estas atenderem a certas condições, após uma conveniente padronização, convergirá em distribuição para uma função de distribuição normal padrão, ou seja

$$\frac{S_n - \mathbb{E}(S_n)}{\sqrt{Var(S_n)}} \xrightarrow{D} Z, \quad \text{quando } n \to \infty$$

onde $S_n = X_1 + X_2 + \cdots + X_n$ e Z é uma variável aleatória com distribuição normal padrão.

Com todas as ferramentas necessárias em mãos, iremos primeiramente demonstrar o Teorema Central do Limite para sequências de variáveis aleatórias independentes e identicamente distribuídas.

Teorema 4.10 *(Teorema Central do Limite para variáveis aleatórias i.i.d.) Seja $X_1, X_2, \ldots$ uma sequência de variáveis aleatórias i.i.d. com $\mathbb{E}(X_i) = \mu$, $Var(X_i) = \sigma^2 \in (0, \infty)$. Se $S_n = X_1 + X_2 + \cdots + X_n$ então*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} Z,$$

onde $Z \sim N(0,1)$ é uma variável aleatória com distribuição normal padrão.

Prova. Podemos supor, sem perda de generalidade que $\mu = 0$, ou poderíamos definir a variável $Y_n = X_n - \mu$ que possui média zero. Para facilitarmos a notação, usaremos $\varphi_{X_i} = \varphi$ para todo i, uma vez que as variáveis aleatórias $X_1, X_2, \ldots$ são identicamente distribuídas.

Aplicando a propriedade (e) do Teorema 4.2 e o Teorema 4.3, respectivamente, teremos

$$\begin{aligned}
\varphi_{\frac{S_n}{\sigma\sqrt{n}}}(t) &= \varphi_{S_n}\left(\frac{t}{\sigma\sqrt{n}}\right) \\
&= \prod_{i=1}^{n} \varphi_{X_i}\left(\frac{t}{\sigma\sqrt{n}}\right) \\
&= \left(\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n.
\end{aligned}$$

Pela fórmula de Taylor em torno de zero, expandindo até o termo de segunda ordem, obteremos

$$\varphi\left(\frac{t}{\sigma\sqrt{n}}\right) = \varphi(0) + \frac{t}{\sigma\sqrt{n}}\varphi'(0) + \frac{1}{2}\left(\frac{t}{\sigma\sqrt{n}}\right)^2 \varphi''(0) + o\left(\frac{t^2}{n}\right).$$

A notação $o(x)$ indica funções tais que $\lim_{x\to 0}\frac{o(x)}{x} = 0$.

Utilizando as propriedades (a) e (c) do Teorema 4.2, teremos $\varphi(0) = 1$, $\varphi'(0) = i\mu = 0$ e $\varphi''(0) = i^2\mathbb{E}(X_1^2) = -\sigma^2$. Deste modo,

$$\begin{aligned}\varphi\left(\frac{t}{\sigma\sqrt{n}}\right) &= 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\\ &= 1 - \frac{t^2}{2n}\left[1 - \frac{o(\frac{t^2}{n})}{\frac{t^2}{2n}}\right]\end{aligned}$$

Também, temos que

$$\lim_{n\to\infty}\left[1 - \frac{o(\frac{t^2}{n})}{\frac{t^2}{2n}}\right] = 1.$$

Utilizando o Lema 3.3 obteremos o seguinte resultado

$$\begin{aligned}\lim_{n\to\infty}\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)^n &= \lim_{n\to\infty}\left\{1 - \frac{t^2}{2n}\left[1 - \frac{o(\frac{t^2}{n})}{\frac{t^2}{2n}}\right]\right\}^n\\ &= e^{\frac{-t^2}{2}}\\ &= \varphi_Z(t).\end{aligned}$$

Desta maneira, pelo Teorema 4.9, podemos concluir que

$$\frac{S_n}{\sigma\sqrt{n}} \xrightarrow{D} Z.$$

■

Concluída a prova do Teorema Central do Limite para variáveis aleatórias i.i.d., iremos agora enunciar o Teorema Central do Limite de Lindeberg, que dará condições mais gerais para a validade da convergência à distribuição normal.

Através do Teorema Central do Limite de Lindeberg poderemos trabalhar com variáveis aleatórias independentes com variância finita, precisando que pelo menos uma delas seja maior que zero. Se estas atenderem à condição chamada de condição de Lindeberg, então a sua soma padronizada convergirá para uma normal padrão.

Sejam $X_1, X_2, \dots$ variáveis aleatórias independentes, com funções de distribuição $F_1, F_2, \dots$, respectivamente, tais que $\mathbb{E}(X_n) = \mu_n$ e $Var(X_n) = \sigma_n^2$, onde $\sigma_n^2 < \infty$ e pelo menos um $\sigma_n^2 > 0$. Sejam $S_n = X_1 + \cdots + X_n$ e $s_n = \sqrt{Var(S_n)} = \sqrt{\sigma_1^2 + \cdots + \sigma_n^2}$. A condição de Lindeberg diz que

$$\forall\, \varepsilon > 0, \quad \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^{n} \int_{|x-\mu_k|>\varepsilon s_n} (x-\mu_k)^2 dF_k(x) = 0. \tag{4-1}$$

A condição de Lindeberg significa, basicamente, que as parcelas $\frac{X_k-\mu_k}{s_n}$ da soma $\frac{S_n-\mathbb{E}(S_n)}{s_n}$ são uniformemente pequenas para n grande. Por exemplo, a condição de Lindeberg implica

$$\max_{1\leq k\leq n} \frac{\sigma_k^2}{s_n^2} \to 0, \qquad \text{quando } n\to\infty,$$

ou seja, para n grande, as variâncias das parcelas são uniformemente pequenas em relação à variância da soma. Esse fato pode ser observado da seguinte forma.

Notemos que para todo k

$$\begin{aligned}
\frac{\sigma_k^2}{s_n^2} &= \frac{1}{s_n^2} \int_{|x-\mu_k|\leq\varepsilon s_n} (x-\mu_k)^2 dF_k(x) + \frac{1}{s_n^2} \int_{|x-\mu_k|>\varepsilon s_n} (x-\mu_k)^2 dF_k(x) \\
&\leq \frac{1}{s_n^2} \int_{|x-\mu_k|\leq\varepsilon s_n} \varepsilon^2 s_n^2 dF_k(x) + \frac{1}{s_n^2} \sum_{k=1}^{n} \int_{|x-\mu_k|>\varepsilon s_n} (x-\mu_k)^2 dF_k(x) \\
&\leq \frac{1}{s_n^2} \int_{-\infty}^{\infty} \varepsilon^2 s_n^2 dF_k(x) + \frac{1}{s_n^2} \sum_{k=1}^{n} \int_{|x-\mu_k|>\varepsilon s_n} (x-\mu_k)^2 dF_k(x).
\end{aligned}$$

Como a primeira parcela é igual a ε^2 que não depende de k, então temos que

$$\max_{1\leq k\leq n} \frac{\sigma_k^2}{s_n^2} \leq \varepsilon^2 + \frac{1}{s_n^2} \sum_{k=1}^{n} \int_{|x-\mu_k|>\varepsilon s_n} (x-\mu_k)^2 dF_k(x),$$

o qual converge para ε^2, pela condição de Lindeberg. Logo, como vale para todo $\varepsilon > 0$, temos $\max_{1\leq k\leq n} \frac{\sigma_k^2}{s_n^2} \to 0$.

Teorema 4.11 *(Teorema Central do Limite de Lindeberg) Sejam $X_1, X_2, \dots$ variáveis aleatórias independentes, com funções de distribuição $F_1, F_2, \dots$ tais que $\mathbb{E}(X_n) = \mu_n$ e $Var(X_n) = \sigma_n^2$, onde $\sigma_n^2 < \infty$ e pelo menos um $\sigma_n^2 > 0$. Sejam $S_n = X_1 + \cdots + X_n$ e $s_n = \sqrt{Var(S_n)} = \sqrt{\sigma_1^2 + \cdots + \sigma_n^2}$. Então, para que tenhamos*

$$\frac{S_n - \mathbb{E}(S_n)}{s_n} \xrightarrow{D} Z,$$

quando $n \to \infty$, dado que Z é uma variável aleatória com distribuição normal padrão, é suficiente que a condição de Lindeberg 4-1 esteja satisfeita.

Prova. Mostraremos que as funções características das somas parciais padronizadas convergem pontualmente para a função característica da normal padrão. Para isso fixaremos um $t \in \mathbb{R}$. Usaremos duas versões da fórmula de Taylor:

(i) $e^{itx} = 1 + itx + \alpha_1(x)\frac{t^2x^2}{2}$, onde $|\alpha_1(x)| \leq 1$,

(ii) $e^{itx} = 1 + itx - \frac{t^2x^2}{2} + \alpha_2(x)\frac{t^3x^3}{6}$, onde $|\alpha_2(x)| \leq 1$

A primeira fórmula será usada para $|x| > \varepsilon$ e a segunda para $|x| \leq \varepsilon$, desta forma poderemos escrever e^{itx} da seguinte forma:

$$e^{itx} = 1 + itx + \alpha_1(x)\frac{t^2x^2}{2} + r_\varepsilon(x),$$

onde

$$r_\varepsilon(x) = \begin{cases} \{1 + \alpha_1(x)\}\frac{t^2x^2}{2}, \text{ se } |x| > \varepsilon \\ \alpha_2(x)\frac{t^3x^3}{6} \qquad\qquad , \text{ se } |x| \leq \varepsilon \end{cases}$$

Teremos, deste modo

$$\mathbb{E}\left(exp\left\{it\left(\frac{X_k - \mu_k}{s_n}\right)\right\}\right) = \int_{-\infty}^{\infty} exp\left\{it\left(\frac{x - \mu_k}{s_n}\right)\right\}dF_k(x).$$

Pela fórmula de Taylor, a expressão acima é igual a

$$\begin{aligned}\int_{-\infty}^{\infty}&\left\{1 + it\left(\frac{x - \mu_k}{s_n}\right) - \frac{t^2}{2}\left(\frac{x - \mu_k}{s_n}\right)^2 + r_\varepsilon\left(\frac{x - \mu_k}{s_n}\right)\right\}dF_k(x) \\ &= 1 + it\mathbb{E}\left(\frac{X_k - \mu_k}{s_n}\right) - \frac{t^2}{2}\mathbb{E}\left[\left(\frac{X_k - \mu_k}{s_n}\right)^2\right] \\ &+ \frac{t^2}{2}\int_{|x-\mu_k|>\varepsilon s_n}\left\{1 + \alpha_1\left(\frac{x - \mu_k}{s_n}\right)\right\}\left(\frac{x - \mu_k}{s_n}\right)^2 dF_k(x) \\ &+ \frac{t^3}{6}\int_{|x-\mu_k|\leq\varepsilon s_n}\alpha_2\left(\frac{x - \mu_k}{s_n}\right)\left(\frac{x - \mu_k}{s_n}\right)^3 dF_k(x).\end{aligned}$$

Como $\mathbb{E}(X_k) = \mu_k$ e $Var(X_k) = \sigma_k^2$, temos

$$\mathbb{E}\left[exp\left\{it\left(\frac{X_k - \mu_k}{s_n}\right)\right\}\right] = 1 - \frac{t^2\sigma_k^2}{2s_n^2} + a_{n,k},$$

no qual o resto $a_{n,k}$ satisfaz a seguinte desigualdade:

$$\begin{aligned}|a_{n,k}| &\leq t^2 \int_{|x-\mu_k|>\varepsilon s_n} \left(\frac{x-\mu_k}{s_n}\right)^2 dF_k(x) + \frac{|t|^3}{6} \int_{|x-\mu_k|\leq\varepsilon s_n} \varepsilon \left(\frac{x-\mu_k}{s_n}\right)^2 dF_k(x) \\ &\leq \frac{t^2}{s_n^2} \int_{|x-\mu_k|>\varepsilon s_n} (x-\mu_k)^2 dF_k(x) + \frac{\varepsilon|t|^3}{6s_n^2} \int_{-\infty}^{\infty} (x-\mu_k)^2 dF_k(x).\end{aligned}$$

Assim, teremos

$$\sum_{k=1}^{n} |a_{n,k}| \leq \frac{t^2}{s_n^2} \sum_{k=1}^{n} \int_{|x-\mu_k|>\varepsilon s_n} (x-\mu_k)^2 dF_k(x) + \frac{\varepsilon|t|^3}{6}.$$

Pela condição de Lindeberg, a primeira parcela da soma vai para zero quando $n \to \infty$. Então, para n suficientemente grande, obtemos

$$\sum_{k=1}^{n} |a_{n,k}| \leq \frac{\varepsilon|t|^3}{3}.$$

Agora, vamos tomar uma sequência de ε's que converge para zero. Tome $\varepsilon = \frac{1}{j}$ para $j \in \mathbb{N}$, então existe n_j tal que para todo $n \geq n_j$,

$$\sum_{k=1}^{n} |a_{n,k}| \leq \frac{|t|^3}{3j} \to 0,$$

no qual os restos $a_{n,k}$ são determinados pela fórmula baseada em $\varepsilon = \frac{1}{j}$. Assim,

$$\begin{aligned}\varphi_{\frac{S_n-\mathbb{E}(S_n)}{s_n}}(t) &= \prod_{k=1}^{n} \mathbb{E}\left(exp\left\{it\left(\frac{X_k-\mu_k}{s_n}\right)\right\}\right) \\ &= \prod_{k=1}^{n} \left(1 - \frac{t^2\sigma_k^2}{2s_n^2} + a_{n,k}\right).\end{aligned}$$

Assim, provaremos que o último termo da equação acima converge para a função característica de uma variável aleatória que possui distribuição normal padrão.

Se Z é uma variável aleatória com distribuição normal padrão, então sua função característica será $\varphi_Z(t) = e^{\frac{-t^2}{2}}$, pelo Exemplo 4.4. Aplicando o Lema 3.4, faremos $c_{n,k} = -\left(\frac{t^2\sigma_k^2}{2s_n^2}\right) + a_{n,k}$ e $c = \frac{-t^2}{2}$. Portanto,

$$\sum_{k=1}^{n} |c_{n,k}| \leq \frac{t^2}{2} + \sum_{k=1}^{n} |a_{n,k}| \xrightarrow{n\to\infty} \frac{t^2}{2},$$

logo $\sum_{k=1}^{n} |c_{n,k}|$ é uniformemente limitado.

Para aplicarmos o Lema 3.4, basta verificarmos a condição do máximo:

$$\begin{aligned}\max_{1\leq k\leq n}|c_{n,k}| &\leq \max_{1\leq k\leq n}\frac{t^2\sigma_k^2}{2s_n^2}+\max_{1\leq k\leq n}|a_{n,k}|\\ &\leq \frac{t^2}{2}\max_{1\leq k\leq n}\frac{\sigma_k^2}{s_n^2}+\sum_{k=1}^{n}|a_{n,k}|\end{aligned}$$

Como o segundo termo tende para zero e a condição de Lindeberg implica

$$\max_{1\leq k\leq n}\frac{\sigma_k^2}{s_n^2}\xrightarrow{n\to\infty}0,$$

concluímos que

$$\varphi_{\frac{S_n-\mathbb{E}(S_n)}{s_n}}(t)\xrightarrow{n\to\infty}e^{\frac{-t^2}{2}}.$$

■

A seguir, veremos um exemplo para compreendermos melhor o Teorema Central do Limite na prática.

Exemplo 4.12 *Um jogo de Roleta num cassino tem números de 1-36 (18 pretos e 18 vermelhos), além do 0 e 00, pintados de verde. Os jogadores podem apostar \$1 que o número sorteado será vermelho (ou preto) e ganhar \$1 caso acertem. Se denotarmos X_i o valor do ganho na i-ésima rodada, então $X_1, X_2, \ldots$ são variáveis aleatórias i.i.d. com $\mathbb{P}(X_i = 1) = 18/38$ e $\mathbb{P}(X_i = -1) = 20/38$. Então*

$$\mathbb{E}(X_i) = -1/19$$

e

$$\begin{aligned}Var(X_i) &= \mathbb{E}[X_i^2]-(\mathbb{E}[X_i])^2\\ &= 1-\left(\frac{1}{19}\right)^2\\ &= 0,9972\end{aligned}$$

Estamos interessados em

$$\mathbb{P}(S_n\geq 0)=\mathbb{P}\left(\frac{S_n-n\mu}{\sigma\sqrt{n}}\geq\frac{-n\mu}{\sigma\sqrt{n}}\right).$$

Tomando $n = 361 = 19^2$ e adotando $\sigma = 1$ para facilitar os cálculos, temos

$$\frac{-n\mu}{\sigma\sqrt{n}}=\frac{361\cdot(1/19)}{\sqrt{361}}=1.$$

Portanto, o Teorema Central do Limite e a tabela da distribuição normal nos dizem que

$$\mathbb{P}(S_n \geq 0) \approx \mathbb{P}(Z \geq 1) = 1 - 0,8413 = 0,1587.$$

Ou seja, após 361 jogadas na roleta, o cassino terá ganhado \$19 em média, mas o jogador tem probabilidade de aproximadamente 16% de estar no lucro.

5
Teorema Central do Limite para Martingais

As versões de Teorema Central do Limite que vimos até agora exigiam independência entre as variáveis aleatórias. O Teorema Central do Limite para martingais é uma generalização dos resultados anteriores, levando em consideração a dependência entre as variáveis.

Neste capítulo veremos diferentes versões do Teorema Central do Limite para martingais. Primeiro provaremos um caso mais geral, para depois entrarmos numa versão específica que nos interessará mais, pois é a abordada em um dos artigos que queremos discutir.

De maneira informal, podemos dizer que as versões do Teorema Central do Limite para martingais expostas neste capítulo enunciam que, dada uma série de condições, se estas forem satisfeitas, uma delas a condição de Lindeberg, teremos que o martingal após uma certa padronização irá convergir em distribuição para uma função de distribuição normal.

Veremos a seguir três exemplos para entendermos intuitivamente as principais ideias apresentadas nesse capítulo.

Exemplo 5.1 *Em um jogo de Roleta justa num cassino, um jogador pode apostar* \$1 *que o número sorteado será vermelho (ou preto) e ganhar* \$1 *caso acerte. Seja Y_i o valor do ganho na i-ésima rodada. As variáveis aleatórias $(Y_i)_{i\geq 1}$ são independentes com distribuição $\mathbb{P}(Y_i = 1) = \mathbb{P}(Y_i = -1) = 1/2$. Seja $S_n = \sum_{i=1}^n Y_i$.*

Nesse exemplo, S_n é simplesmente um passeio aleatório simples em $\mathbb{Z}$. Consideramos agora dois processos parecidos em que o apostador muda sua estratégia dependendo dos resultados anteriores.

Exemplo 5.2 *Agora, o jogador aposta* \$2 *se ganhou na última jogada. Podemos definir o ganho em termos das variáveis aleatórias definidas no último exemplo:*

$$X_i = (1 + \mathbb{1}_{Y_{i-1}=1})Y_i \,.$$

Seja $S'_n = \sum_{i=1}^n X_i$.

Exemplo 5.3 *Neste exemplo, o jogador aposta \$2 se seu lucro total até o momento é positivo. Podemos definir o ganho em termos das variáveis aleatórias definidas no exemplo 5.1:*

$$W_i = (1 + \mathbb{1}_{S''_{i-1}>0})Y_i ,$$

onde $S''_n = \sum_{i=1}^n W_i$.

Os três exemplos acima são exemplos de martingais. O primeiro corresponde a um caso que se enquadra nos teoremas do último capitulo, porém os outros dois exemplos não são dessa forma. Assim, é natural nos perguntarmos se o Teorema Central do Limite ainda se aplica nesses casos. Veremos em breve que a resposta depende do comportamento da variação quadrática do processo.

Podemos ver que a variância quadrática do Exemplo 5.2 é concentrada. Além disso, parece intuitivo que exista um Teorema Central do Limite para o martingal desse exemplo. Já no Exemplo 5.3, a variância quadrática do martingal não será concentrada, e o Teorema Central do Limite não se aplica ao martingal em questão.

Portanto, podemos dizer que se a variância quadrática do martingal for concentrada, e se outras condições técnicas forem garantidas, o valor final da soma dos incrementos S_n terá distribuição aproximadamente gaussiana. O objetivo desse capítulo será demonstrarmos esse resultado, que será formalizado pelo Corolário 5.12.

5.1 Versões do Teorema Central do Limite para Martingais

Primeiramente, vamos demonstrar uma versão geral do Teorema Central do Limite para martingais, para depois adentrarmos no caso específico que nos interessa mais para o presente trabalho.

Trabalharemos com vetores de martingais $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq n, n \geq 1\}$ de média zero e quadrado integráveis, que são derivados de martingais ordinários $\{S_n, \mathcal{F}_n, 1 \leq n < \infty\}$, da seguinte maneira: definimos $n = n$,$\mathcal{F}_{n,i} = \mathcal{F}_n$ e $S_{n,i} = s_n^{-1} S_i$, onde s_n^{-1} é o desvio padrão de S_n. Esta notação é utilizada em [17].

Antes de começarmos a discutir o Teorema Central do Limite para martingais, precisamos do resultado enunciado no lema abaixo. Para este lema, usaremos a convergência em probabilidade e em $\mathcal{L}^k$ e integrabilidade uniforme, que serão definidas a seguir.

Definição 5.4 *(Convergência em Probabilidade) Sejam $Y, Y_1, Y_2, \ldots$ variáveis aleatórias definidas no mesmo espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. Então, Y_n converge para Y em probabilidade se para todo $\epsilon > 0$,*

$$\mathbb{P}(|Y_n - Y| \geq \epsilon) \to 0 \quad \textit{quando } n \to \infty.$$

A notação que utilizaremos será $Y_n \xrightarrow{p} Y$.

A ideia da convergência em probabilidade é que, quando n é arbitrariamente grande, a probabilidade da diferença $|Y_n - Y|$ ser maior do que qualquer número positivo ϵ tende a zero.

Definição 5.5 *(Convergência em $\mathcal{L}^k$) Seja $k \in [1, \infty)$. A sequência de variáveis aleatórias Y_n converge em $\mathcal{L}^k$ para Y, se $Y_n \in \mathcal{L}^k, Y \in \mathcal{L}^k$ e*

$$\lim_{n \to \infty} \mathbb{E}[|Y_n - Y|^k] = 0$$

A notação que utilizaremos será $Y_n \xrightarrow{\mathcal{L}^k} Y$.

Para $k = 1$, podemos pensar na convergência em $\mathcal{L}^k$ como sendo uma convergência em média das variáveis. Este será o caso do lema a seguir.

Definição 5.6 *(Integrabilidade Uniforme) Uma família de variáveis aleatórias $(X_i)_{i \in I}$ é uniformemente integrável se $\forall\, \epsilon > 0$, existe $K > 0$ tal que $\mathbb{E}[|X_i| \mathbb{1}_{|X_i| \geq K}] \leq \epsilon$ para todo $i \in I$.*

Lema 5.7 *Seja η^2 uma constante positiva e $X_{n,m}$ um array triangular de variáveis aleatórias quaisquer onde $1 \leq m \leq n$. Suponha que, quando $n \to \infty$,*

$$\max_{1 \leq m \leq n} |X_{n,m}| \xrightarrow{p} 0, \tag{5-1}$$

$$\sum_{m=1}^{n} X_{n,m}^2 \xrightarrow{p} \eta^2, \tag{5-2}$$

e para todo t real, sendo $T_n(t) = \prod_{j=1}^{n} (1 + itX_{n,j})$, temos

$$T_n(t) \xrightarrow{\mathcal{L}^1} 1 \tag{5-3}$$

quando $n \to \infty$.

Então, teremos $S_{n,n} = X_{n,1} + \cdots + X_{n,n} \xrightarrow{d} Z$, onde $Z \sim N(0, \eta)$ é uma variável aleatória com distribuição normal.

Prova. Conforme [18], temos que

$$exp(ix) = (1+ix)exp\bigg(-\frac{1}{2}x^2 + r(x)\bigg), \tag{5-4}$$

onde $|r(x)| \leq |x|^3$ para $|x| \leq 1$.

Seja $I_n = exp(itS_{n,n})$ e

$$W_n = exp\bigg(-\frac{1}{2}t^2\sum_{m=1}^{n} X_{n,m}^2 + \sum_{m=1}^{n} r(tX_{n,m})\bigg).$$

Então

$$\begin{aligned} I_n &= exp(it\sum_{m=1}^{n} X_{n,m}) \\ &= \prod_{m=1}^{n} exp(itX_{n,m}). \end{aligned}$$

Utilizando a fórmula 5-4 para I_n, teremos

$$\begin{aligned} I_n &= \prod_{m=1}^{n}(1+itX_{n,m})exp(-\frac{1}{2}t^2X_{n,m}^2 + r(tX_{n,m})) \\ &= \bigg[\prod_{m=1}^{n}(1+itX_{n,m})\bigg]\bigg[\prod_{m=1}^{n} exp(-\frac{1}{2}t^2X_{n,m}^2 + r(tX_{n,m}))\bigg] \\ &= \bigg[\prod_{m=1}^{n}(1+itX_{n,m})\bigg]exp\bigg(-\frac{1}{2}t^2\sum_{m=1}^{n} X_{n,m}^2 + \sum_{m=1}^{n} r(tX_{n,m})\bigg) \\ &= T_nW_n + T_nexp(-\frac{1}{2}\eta^2t^2) - T_nexp(-\frac{1}{2}\eta^2t^2) \\ &= T_nexp(-\frac{1}{2}\eta^2t^2) + T_n(W_n - exp(-\frac{1}{2}\eta^2t^2)). \end{aligned}$$

Como vimos no capítulo anterior, uma variável aleatória X_n converge em distribuição para X se, e somente se, sua função característica $\varphi_{X_n}(t)$ converge para a função característica de X, $\varphi_X(t)$, para todo $t \in \mathbb{R}$. Portanto, é suficiente provar que

$$\mathbb{E}(I_n) \to exp(-\frac{1}{2}\eta^2t^2). \tag{5-5}$$

Como $exp(-\frac{1}{2}\eta^2t^2)$ é limitada, 5-3 garante que, quando $n \to \infty$,

$$\mathbb{E}(T_nexp(-\frac{1}{2}\eta^2t^2)) \to \mathbb{E}(exp(-\frac{1}{2}\eta^2t^2)) = exp(-\frac{1}{2}\eta^2t^2). \tag{5-6}$$

Ademais, qualquer sequência de variáveis aleatórias que converge em $\mathcal{L}^1$ é uniformemente integrável, portanto a sequência

$$T_n(W_n - exp(-\frac{1}{2}\eta^2t^2)) = I_n - T_nexp(-\frac{1}{2}\eta^2t^2)$$

é uniformemente integrável. (A integrabilidade uniforme de I_n segue do fato que $|I_n| = 1$.) As condições 5-1 e 5-2 implicam que, quando $\max_{1\leq m\leq n}|X_{n,m}| \leq 1$,

$$\begin{aligned}\left|\sum_{m=1}^{n} r(tX_{n,m})\right| &\leq |t|^3 \sum_{m=1}^{n} |X_{n,m}|^3 \\ &\leq |t|^3\Big(\max_{1\leq m\leq n}|X_{n,m}|\Big)\left(\sum_{m=1}^{n} X_{n,m}^2\right) \xrightarrow{p} 0.\end{aligned}$$

Segue então que $W_n - exp(-\frac{1}{2}\eta^2 t^2) \xrightarrow{p} 0$, e pela integrabilidade uniforme teremos

$$\mathbb{E}\Big(T_n(W_n - exp(-\frac{1}{2}\eta^2 t^2))\Big) \to 0. \tag{5-7}$$

As condições 5-6 e 5-7 implicam 5-5. ■

Agora, iremos trabalhar com $\{X_{n,m}\}$ sendo o vetor de diferenças de martingais.

Teorema 5.8 *Seja $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq n\}$ um vetor de martingais com média zero, quadrado integráveis, isto é, $S_{n,i} \in \mathcal{L}^2$, e com incrementos $X_{n,i}$. Seja η^2 uma constante real positiva e suponha que*

$$\max_{1\leq i\leq n}|X_{n,i}| \xrightarrow{p} 0, \tag{5-8}$$

$$\sum_{i=1}^{n} X_{n,i}^2 \xrightarrow{p} \eta^2, \tag{5-9}$$

$$\mathbb{E}\Big(\max_i X_{n,i}^2\Big) \quad \textit{é limitada uniformemente em } n, \tag{5-10}$$

e as σ-álgebras são da forma

$$\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i} \tag{5-11}$$

para $1 \leq i \leq n$.

Então $S_{n,n} = \sum_{i=1}^{n} X_{n,i} \xrightarrow{D} Z$, onde a variável aleatória Z tem função característica $\varphi_Z(t) = exp(-\frac{1}{2}\eta^2 t^2)$, ou seja, uma distribuição normal.

Prova. Façamos $X'_{n,i} = X_{n,i}\mathbb{1}_{(\sum_{j=1}^{i-1} X_{n,j}^2 \leq 2\eta^2)}$ e $S'_{n,i} = \sum_{j=1}^{i} X'_{n,j}$. Então, temos que $\{S'_{n,i}, \mathcal{F}_{n,i}\}$ é um vetor de martingais. Além disso,

$$\mathbb{P}(\exists i,\ X'_{n,i} \neq X_{n,i}) \leq \mathbb{P}(U_{n,n}^2 > 2\eta^2) \to 0 \tag{5-12}$$

para $i \leq n$, onde $U_{n,n}^2 = \sum_{i=1}^{n} X_{n,i}^2$. Isso implica que $\mathbb{P}(S'_{n,n} \neq S_{n,n}) \to 0$, então

$$\mathbb{E}|exp(itS'_{n,n}) - exp(itS_{n,n})| \to 0.$$

Consequentemente, $S_{n,n} \xrightarrow{D} Z$, se e somente se $S'_{n,n} \xrightarrow{D} Z$. Pela condição 5-12, as diferenças de martingal $X'_{n,i}$ satisfazem as condições 5-1 e 5-2 do Lema 5.7. Devemos, portanto, verificar a condição 5-3.

Seja $T'_n = \prod_{j=1}^{n}(1 + itX'_{n,j})$ e

$$J_n = \begin{cases} \min\{i \leq n : U_{n,i}^2 > 2\eta^2\}, \text{ se } U_{n,n}^2 > 2\eta^2 \\ n \qquad\qquad\qquad\qquad\qquad\quad, \text{ caso contrário.} \end{cases}$$

Sejam $x, y \in \mathbb{R}$, temos que $|x + iy|^2 = x^2 + y^2$. Então

$$\begin{aligned} \mathbb{E}|T'_n|^2 &= \mathbb{E}\Big[\Big(\prod_{j=1}^{n} |1 + itX'_{n,j}|\Big)^2\Big] \\ &= \mathbb{E}\Big[\prod_{j=1}^{n}(1 + t^2X'^2_{n,j})\Big] \\ &\leq \mathbb{E}\Big[\Big(exp\Big(t^2\sum_{j=1}^{J_n-1} X'^2_{n,j}\Big)\Big)(1 + t^2X^2_{n,J_n})\Big] \\ &\leq \Big[exp(2\eta^2t^2)\Big]\Big(1 + t^2\mathbb{E}(X^2_{n,J_n})\Big), \end{aligned}$$

onde a desigualdade vem do fato de $1 + x \leq \exp(x)$, é limitado uniformemente em n pela condição 5-10. Consequentemente, $\{T'_n\}$ é uniformemente integrável.

Seja $m \geq 1$ fixo e $E \in \mathcal{F}_{m,m}$. Então pela condição 5-11, $E \in \mathcal{F}_{n,n}$ para todo $n \geq m$, tal que em n,

$$\begin{aligned} \mathbb{E}(T'_n\mathbb{1}(E)) &= \mathbb{E}\Big(\mathbb{1}(E)\prod_{j=1}^{n}(1 + itX'_{n,j})\Big) \\ &= \mathbb{E}\Big(\mathbb{1}(E)\prod_{j=1}^{m}(1 + itX'_{n,j})\prod_{j=m+1}^{n}\mathbb{E}\Big[(1 + itX'_{n,j})|\mathcal{F}_{n,j-1}\Big]\Big) \\ &= \mathbb{E}\Big(\mathbb{1}(E)\prod_{j=1}^{m}(1 + itX'_{n,j})\Big) \\ &= \mathbb{P}(E) + R_n, \end{aligned}$$

onde o termo R_n consiste em no máximo $2^m - 1$ termos na forma

$$\mathbb{E}[\mathbb{1}(E)(it)^rX'_{n,j_1}X'_{n,j_2}\cdots X'_{n,j_r}],$$

onde $1 \leq r \leq m$ e $1 \leq j_1 \leq j_2 \leq \cdots \leq m$.

Como temos

$$\left|X'_{n,j_1}\cdots X'_{n,j_r}\right|^2 \leq \left(\sum_{j=1}^{J_n-1} X'^2_{n,j}\right)^{r-1}\Big(\max_i X'^2_{n,i}\Big) \leq (2\eta^2)^{r-1}\Big(\max_i X'^2_{n,i}\Big),$$

segue que

$$|R_n| \leq (2^m-1)(2\eta^2)^{m/2}\mathbb{E}\Big(\max_i |X'_{n,i}|\Big).$$

Mas, para todo $\varepsilon > 0$, como $\max_i |X_{n,i}| \leq \varepsilon + \max_i |X_{n,i}|\mathbb{1}(|X_{n,i}| > \varepsilon)$, temos pela linearidade da esperança

$$\mathbb{E}\Big(\max_i |X_{n,i}|\Big) \leq \varepsilon + \mathbb{E}\Big(\max_i |X_{n,i}|\mathbb{1}(|X_{n,i}| > \varepsilon)\Big).$$

Utilizando a desigualdade de Cauchy-Schwarz para limitar $\mathbb{E}(\max_i |X_{n,i}|\mathbb{1}(|X_{n,i}| > \varepsilon))$, teremos

$$\mathbb{E}\Big(\max_i |X_{n,i}|\Big) \leq \varepsilon + \left[\mathbb{E}\Big(\max_i X^2_{n,i}\Big)\mathbb{P}\Big(\max_i |X_{n,i}| > \varepsilon\Big)\right]^{1/2}.$$

Como $\mathbb{E}\Big(\max_i X^2_{n,i}\Big)$ é limitada em n e $\mathbb{P}\Big(\max_i |X_{n,i}| > \varepsilon\Big) \to 0$, temos que

$$\mathbb{E}\Big(\max_i |X_{n,i}|\Big) \leq 2\varepsilon$$

para todo n suficientemente grande.

Como ε é arbitrário, segue que $\mathbb{E}(\max_i |X_{n,i}|) \to 0$. Sendo assim, $R_n \to 0$. Desta forma, $\mathbb{E}[T'_n\mathbb{1}(E)] \to \mathbb{P}(E)$.

Seja $\mathcal{F}_\infty = \vee_{n=1}^{\infty}\mathcal{F}_{n,n}$ a σ-álgebra gerada pela $\cup_{n=1}^{\infty}\mathcal{F}_n$. Para qualquer $E' \in \mathcal{F}_\infty$ e algum $\varepsilon > 0$, existe um m e um $E \in \mathcal{F}_{m,m}$ tal que $\mathbb{P}(E'\triangle E) < \varepsilon$, onde $\triangle$ denota a diferença simétrica. Como $\{T'_n\}$ é uniformemente integrável e

$$|\mathbb{E}[T'_n\mathbb{1}(E')] - \mathbb{E}[T'_n\mathbb{1}(E)]| \leq \mathbb{E}[|T'_n|\mathbb{1}(E\triangle E')],$$

temos que $\sup_n |\mathbb{E}[T'_n\mathbb{1}(E')] - \mathbb{E}[T'_n\mathbb{1}(E)]|$ pode ser arbitrariamente pequeno, escolhendo-se ε suficientemente pequeno.

Temos de $\mathbb{E}[T'_n\mathbb{1}(E)] \to \mathbb{P}(E)$, que para qualquer $E' \in \mathcal{F}_\infty$, $\mathbb{E}[T'_n\mathbb{1}(E')] \to \mathbb{P}(E')$. Isto implica que para qualquer variável aleatória limitada X $\mathcal{F}_\infty$-mensurável, obteremos $\mathbb{E}[T'_nX] \to \mathbb{E}(X)$. Finalmente, se $E \in \mathcal{F}$, então

$$\mathbb{E}[T'_n\mathbb{1}(E)] = \mathbb{E}[T'_n\mathbb{E}(\mathbb{1}(E)|\mathcal{F}_\infty)] \to \mathbb{E}[\mathbb{E}(\mathbb{1}(E)|\mathcal{F}_\infty) = \mathbb{P}(E).$$

Isso implica que $T'_n \xrightarrow{\mathcal{L}^1} 1$. Com isso, provamos que a condição 5-3 do Lema 5.7 se aplica, logo a prova está concluída. ■

Corolário 5.9 *Se as condições 5-8 e 5-10 forem substituídas pela condição de Lindeberg*

$$\sum_i \mathbb{E}[X_{n,i}^2 \mathbb{1}(|X_{n,i}| > \varepsilon)|\mathcal{F}_{n,i-1}] \xrightarrow{p} 0, \quad \forall \varepsilon > 0,$$

e a condição 5-9 for substituída por uma condição análoga para a variância condicional

$$V_{n,i}^2 = \sum_{j=1}^{i} \mathbb{E}[X_{n,j}^2|\mathcal{F}_{n,j-1}] \xrightarrow{p} \eta^2,$$

e se a condição 5-11 se mantiver, então a conclusão do Teorema 5.8 permanece.

Para verificarmos o Corolário 5.9 do Teorema acima, devemos analisar a relação entre a variância condicional $V_{n,i}^2$ e o somatório $\sum_j X_{n,j}^2$. Para isso, precisaremos da seguinte proposição enunciada abaixo.

Proposição 5.10 *Seja* $\{S_{n,i} = \sum_{j=1}^{i} X_{n,j}, \mathcal{F}_{n,i}, 1 \leq i \leq n, n \geq 1\}$ *um vetor de martingal de média zero e suponha que as variâncias condicionais* $V_{n,i}^2$ *são apertadas, isto é,*

$$\sup_n \mathbb{P}\Big(V_{n,n}^2 > \lambda\Big) \to 0 \quad \text{quando } \lambda \to \infty,$$

e a condição de Lindeberg está assegurada. Então

$$\max_{i \leq n} \left| \sum_{j=1}^{i} X_{n,j}^2 - V_{n,i}^2 \right| \xrightarrow{p} 0.$$

Prova. A demonstração está em [17], página 45. ■

Como, no Corolário 5.9, a variância condicional $V_{n,n}^2 = \sum_{j=1}^{n} \mathbb{E}[X_{n,j}^2|\mathcal{F}_{n,j-1}] \xrightarrow{p} \eta^2$, onde η^2 é uma constante positiva, a sequência de variâncias condicionais será apertada. Além disso, a condição de Lindeberg está assegurada no Corolário 5.9, logo podemos aplicar a Proposição 5.10. Deste modo, temos que

$$\left| \sum_{j=1}^{n} X_{n,j}^2 - V_{n,n}^2 \right| \xrightarrow{p} 0.$$

Portanto, a condição 5-9 do Teorema 4.5 será equivalente a $V_{n,n}^2 \xrightarrow{p} \eta^2$ e as condições 5-8 e 5-10 serão equivalentes à condição de Lindeberg. Assim, o Corolário 5.9 se mantém.

Teorema 5.11 *Seja* $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq n, n \geq 1\}$ *um vetor de martingais com média zero, quadrado integráveis, isto é,* $S_{n,i} \in \mathcal{L}^2$*, e com incrementos* $X_{n,i}$*. Seja* η^2 *uma constante real positiva e suponha que*

$$\max_{1\leq i\leq n} |X_{n,i}| \xrightarrow{p} 0,$$

$$\sum_{i=1}^{n} X_{n,i}^2 \xrightarrow{p} \eta^2,$$

$$\mathbb{E}\Big(\max_i X_{n,i}^2\Big) \quad \textit{é limitada em } n,$$

e as σ*-álgebras são da forma*

$$\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$$

para $1 \leq i \leq n, n \geq 1$*. Então,*

$$\frac{S_{n,n}}{\left(\sum_j^n X_{n,j}^2\right)^{1/2}} \xrightarrow{D} N(0,1). \qquad (5\text{-}13)$$

Prova. Suponha que Z é uma variável aleatória com função característica $\varphi_Z(t) = exp(-\frac{1}{2}\eta^2 t^2)$. Vimos no Teorema 5.8 que $S_{n,n} \xrightarrow{D} Z$, então para qualquer t real teremos

$$exp(itS_{n,n}) \to exp(-\frac{1}{2}\eta^2 t^2).$$

Portanto,

$$\eta^{-1} S_{n,n} \xrightarrow{D} N(0,1).$$

Da condição 5-9 do Teorema 5.8, presente no enunciado atual, obtemos:

$$\frac{S_{n,n}}{\left(\sum_j^n X_{n,j}^2\right)^{1/2}} \xrightarrow{D} N(0,1).$$

■

Corolário 5.12 *Seja* $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq n, n \geq 1\}$ *um vetor de martingais com média zero, quadrado integráveis, isto é,* $S_{n,i} \in \mathcal{L}^2$*, e com incrementos* $X_{n,i}$*. Seja* η^2 *uma constante real positiva e suponha que*

$$\max_{1\leq i\leq n} |X_{n,i}| \xrightarrow{p} 0,$$

$$V_{n,n}^2 \xrightarrow{p} \eta^2,$$

$$\mathbb{E}\Big(\max_{i\leq n} X_{n,i}^2\Big) \quad \textit{é limitada em } n,$$

e as σ*-álgebras são da forma*

$$\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$$

para $1 \leq i \leq n, n \geq 1$. *Então,*

$$\frac{S_{n,n}}{V_{n,n}} \xrightarrow{D} N(0,1),$$

onde $V_{n,n}^2 = \sum_{j=1}^{n} \mathbb{E}[X_{n,j}^2 | \mathcal{F}_{n,j-1}]$.

Prova. O Corolário 5.12 é semelhante ao Teorema 5.11, porém substituímos $\sum_{j=1}^{n} X_{n,j}^2$ por $V_{n,n}^2$ na segunda condição e no resultado final. Contudo, essas mudanças são consequência da Proposição 5.10, uma vez que

$$\left| \sum_{j=1}^{n} X_{n,j}^2 - V_{n,n}^2 \right| \xrightarrow{p} 0.$$

Isso é suficiente para demonstrar o corolário. ■

Com os Teoremas e Corolários demonstrados acima podemos tratar do Teorema Central do Limite para martingais de uma forma mais específica, da mesma maneira que o artigo [11], que exploraremos no próximo capítulo, o fez.

6
Taxas de Convergência no Teorema Central do Limite para Martingais

No capítulo anterior, estudamos o Teorema Central do Limite para martingais, que atesta que um martingal, sob certas condições, converge em distribuição para uma distribuição normal.

Nos interessa, portanto, determinar as taxas em que essa convergência ocorre. Informalmente, estamos interessados em identificar a velocidade com que a distribuição em questão converge para uma distribuição normal padrão. Para ilustrar esse ponto, se considerarmos um vetor de martingais $\{S_n, \mathcal{F}_n, n \geq 1\}$, queremos saber o quão grande n deve ser para que a aproximação entre as distribuições ocorra.

Neste capítulo, iremos trabalhar com o artigo [11], que trata justamente das taxas de convergência no Teorema Central do Limite para martingais. Primeiramente, retomaremos a notação apresentada na Introdução e que é abordada no artigo, para posteriormente analisarmos e discutirmos os teoremas apresentados e suas conclusões.

Podemos pensar que o limite da taxa de convergência no Teorema Central do Limite pode ser escrita como soma de dois termos, $A_k + B_k$, para qualquer $k \geq 1$, onde $A_k = ||V^2 - 1||_k^{k/(2k+1)}$ e V^2 é a variação quadrática normalizada do martingal. O foco do artigo é discutir a otimização desse termo A_k para $k \geq 1$, focando na classe restrita de martingais de incrementos limitados.

6.1
Noções Introdutórias

Seja $\mathbf{X} = (X_1, \ldots, X_n)$ uma sequência de variáveis aleatórias quadrado-integráveis, isto é, $X_i \in \mathcal{L}^2$ para todo i. Além disso, X_i satisfaz $\mathbb{E}[X_i|\mathcal{F}_{i-1}] = 0$ para todo i, onde $\mathcal{F}_i$ é a σ-álgebra gerada por $(X_1, \ldots, X_i)$. Em outras palavras, $\mathbf{X}$ é uma sequência de martingais.

Escrevemos $\mathcal{M}_n$ como o conjunto de todas as possíveis sequências supracitadas de tamanho n, e reintroduzimos as seguintes notações:

$$s^2(\mathbf{X}) = \sum_{i=1}^{n} \mathbb{E}[X_i^2], \tag{6-1}$$

$$V^2(\mathbf{X}) = s^{-2}(\mathbf{X}) \sum_{i=1}^{n} \mathbb{E}[X_i^2 | \mathcal{F}_{i-1}], \tag{6-2}$$

$$S(\mathbf{X}) = \sum_{i=1}^{n} X_i. \tag{6-3}$$

Chamamos $V^2(\mathbf{X})$ de variação quadrática normalizada de $\mathbf{X}$. Seja $(\mathbf{X}_n)_{n \in \mathbb{N}}$ tal que, para todo n, $\mathbf{X}_n \in \mathcal{M}_n$.

Fixaremos agora a constante η^2, que utilizamos nos teoremas e corolários do capítulo 5, como sendo igual a 1.

Deste modo, ao analisarmos a variação quadrática normalizada, se

$$V^2(\mathbf{X}_n) \xrightarrow{p} 1, \quad \text{quando } n \to \infty \tag{6-4}$$

e a condição de Lindeberg for satisfeita, então a soma $S(\mathbf{X}_n)/s(\mathbf{X}_n)$ converge em distribuição para uma variável aleatória normal padrão, isto é,

$$\forall t \in \mathbb{R}, \quad \mathbb{P}[S(\mathbf{X}_n)/s(\mathbf{X}_n) \le t] \to \Phi(t), \quad \text{quando } n \to \infty, \tag{6-5}$$

onde $\Phi(t) = (2\pi)^{-1/2} \int_{-\infty}^{t} e^{-x^2/2} dx$.

6.2 Taxas de Convergência

Estamos interessados nos limites da velocidade de convergência nesse Teorema Central do Limite. Uma maneira natural de fortalecer a convergência em probabilidade (1.1) é alterá-la para a convergência em $\mathcal{L}^k$, para $k \in [1, +\infty]$. Isso decorre do fato da convergência em $\mathcal{L}^k$ implicar a convergência em probabilidade.

De fato, estimativas quantitativas do termo $||V^2 - 1||_k$ parecem mais convenientes para aplicações práticas dos resultados. Escrevemos

$$D(\mathbf{X}) = \sup_{t \in \mathbb{R}} |[S(\mathbf{X})/s(\mathbf{X}) \le t] - \Phi(t)|,$$

para denotar a maior distância entre as distribuições, conforme definimos no Capítulo 3. Além disso, utilizaremos a notação

$$||\mathbf{X}||_k = \max_{1 \leq i \leq n} ||X_i||_k \quad (k \in [1, +\infty]).$$

Sobre as taxas de convergência no Teorema Central do Limite para martingais, foram obtidos diversos resultados, sob uma variedade de premissas. Veremos agora um teorema que foi demonstrado por [19].

Teorema 6.1 *Seja $\gamma \in (0, +\infty)$. Existe uma constante $C_\gamma > 0$ tal que para todo $n \geq 2$ e para todo $\mathbf{X} \in \mathcal{M}_n$, satisfazendo $||\mathbf{X}||_\infty \leq \gamma$ e $V^2(\mathbf{X}) = 1$ q.c.,*

$$D(\mathbf{X}) \leq C_\gamma \frac{n \log(n)}{s^3(\mathbf{X})}.$$

Tipicamente, $s(\mathbf{X})$ é da ordem $\sqrt{n}$ quando $\mathbf{X} \in \mathcal{M}_n$. Sob tais circunstâncias, o Teorema 6.1 nos dá uma taxa de ordem $\log(n)/\sqrt{n}$. Relaxando a condição de $V^2(\mathbf{X}) = 1$ q.c., [19] demonstrou o seguinte Corolário.

Corolário 6.2 *Seja $\gamma \in (0, +\infty)$. Existe uma constante $\overline{C}_\gamma > 0$ tal que para todo $n \geq 2$ e para todo $\mathbf{X} \in \mathcal{M}_n$, satisfazendo $||\mathbf{X}||_\infty \leq \gamma$,*

$$D(\mathbf{X}) \leq \overline{C}_\gamma \left[\frac{n \log(n)}{s^3(\mathbf{X})} + min\{||V^2(\mathbf{X}) - 1||_1^{1/3}, ||V^2(\mathbf{X}) - 1||_\infty^{1/2}\} \right]. \tag{6-6}$$

Foi desenvolvida por [19] uma estratégia para provar que o termo $||V^2(\mathbf{X}) - 1||_1^{1/3}$ é de fato ótimo, inclusive na classe restrita dos martingais de incrementos limitados consideradas pelo Corolário 6.2. Isso gerou uma resposta satisfatória à questão de otimização do termo $A_k = ||V^2 - 1||_k^{k/(2k+1)}$ para $k = 1$, ou seja, a estratégia usada mostrou que este termo é o que vai mais rapidamente para zero, para $k = 1$.

O artigo [11] tem por objetivo estender essa estratégia para $k \geq 1$, justificando a otimização do termo A_k. Para isso, será necessário dar um novo limite para a taxa de convergência no Teorema Central do Limite para martingais com incrementos limitados.

Teorema 6.3 *Seja $k \in [1, +\infty)$ e $\gamma \in (0, +\infty)$. Existe uma constante $C_{k,\gamma} > 0$ tal que para todo $n \geq 2$ e para todo $\mathbf{X} \in \mathcal{M}_n$, satisfazendo $||\mathbf{X}||_\infty \leq \gamma$,*

$$D(\mathbf{X}) \leq C_{k,\gamma} \left[\frac{n \log(n)}{s^3(\mathbf{X})} + (||V^2(\mathbf{X}) - 1||_k^k + s^{-2k}(\mathbf{X}))^{1/(2k+1)} \right]. \tag{6-7}$$

Prova. Seja $\mathbf{X} = (X_1, \ldots, X_n) \in \mathcal{M}_n$ tal que $||\mathbf{X}||_\infty \leq \gamma$. A ideia da prova é expandir a sequência para um $\mathbf{X}' \in M_{2n}$ tal que $V^2(\mathbf{X}') = 1$ q.c., preservando a propriedade de $||\mathbf{X}'||_\infty \leq \gamma$, e aplicar o Teorema 6.1 nessa sequência maior.

Seja

$$\tau = sup\left\{ j \leq n : \sum_{i=1}^{j} \mathbb{E}[X_i^2|\mathcal{F}_{i-1}] \leq s^2(\mathbf{X}) \right\}.$$

Para $i \leq \tau$, definimos $X_i' = X_i$. Seja r o maior inteiro que não excede

$$\frac{s^2(\mathbf{X}) - \sum_{i=1}^{\tau} \mathbb{E}[X_i^2|\mathcal{F}_{i-1}]}{\gamma^2}.$$

Uma vez que $||\mathbf{X}||_\infty \leq \gamma$, claramente $r \leq n$. Condicional em $\mathcal{F}_\tau$ e para $1 \leq i \leq r$, sejam X_i' variáveis aleatórias independentes tais que $\mathbb{P}[X_{\tau+i}' = \pm\gamma] = 1/2$. Se $\tau + r < 2n$, então definimos $X_{\tau+r+1}'$ tal que

$$\mathbb{P}\left[X_{\tau+r+1}' = \pm\left(s^2(\mathbf{X}) - \sum_{i=1}^{\tau} \mathbb{E}[X_i^2|\mathcal{F}_{i-1}] - r\gamma^2 \right)^{1/2} \right] = \frac{1}{2},$$

com o sinal determinado independentemente. Finalmente, se $\tau + r + 1 < 2n$, definimos $X_{\tau+r+i}' = 0$ para todo $i \geq 2$.

Ampliando as σ-álgebras, podemos assumir que X_i' é $\mathcal{F}_i$-mensurável para $i \leq n$, e definimos $\mathcal{F}_i$ como a σ-álgebra gerada por $\mathcal{F}_n$ e $X_{n+1}', \ldots, X_{n+i}'$ se $i > n$. Por construção, teremos

$$\sum_{i=\tau+1}^{2n} \mathbb{E}[X_i'^2|\mathcal{F}_{i-1}] = \mathbb{E}[X_{\tau+1}'^2|\mathcal{F}_\tau] + \cdots + \mathbb{E}[X_{\tau+r}'^2|\mathcal{F}_{\tau+r-1}]$$

$$+\mathbb{E}[X_{\tau+r+1}'^2|\mathcal{F}_{\tau+r}] + \cdots + \mathbb{E}[X_{2n}'^2|\mathcal{F}_{2n-1}].$$

Pela construção das variáveis aleatórias, temos que

$$\mathbb{E}[X_i'^2|\mathcal{F}_{i-1}] = \begin{cases} \gamma^2 & \text{, se } \tau + 1 \leq i \leq \tau + r \\ s^2(\mathbf{X}) - \sum_{i=1}^{\tau} \mathbb{E}[X_i^2|\mathcal{F}_{i-1}] - r\gamma^2 & \text{, se } i = \tau + r + 1 \\ 0 & \text{, se } i > \tau + r + 1. \end{cases}$$

O termo γ^2 se repete r vezes, portanto teremos

$$\sum_{i=\tau+1}^{2n} \mathbb{E}[X_i'^2|\mathcal{F}_{i-1}] = r\gamma^2 + s^2(\mathbf{X}) - \sum_{i=1}^{\tau} \mathbb{E}[X_i^2|\mathcal{F}_{i-1}] - r\gamma^2$$
$$= s^2(\mathbf{X}) - \sum_{i=1}^{\tau} \mathbb{E}[X_i^2|\mathcal{F}_{i-1}],$$

que pode ser reescrito como

$$\sum_{i=1}^{2n} \mathbb{E}[X_i'^2|\mathcal{F}_{i-1}] = s^2(\mathbf{X}).$$

Consequentemente, $s^2(\mathbf{X}') = s^2(\mathbf{X})$ e $V^2(\mathbf{X}') = 1$ q.c. Desse modo, a sequência $\mathbf{X}'$ satisfaz as premissas do Teorema 6.1, portanto

$$D(\mathbf{X}') \leq 4C_\gamma \frac{n\log(n)}{s^3(\mathbf{X})}. \tag{6-8}$$

Para todo $x > 0$, temos

$$\mathbb{P}\left[\frac{S(\mathbf{X})}{s(\mathbf{X})} \leq t\right] \leq \mathbb{P}\left[\frac{S(\mathbf{X})}{s(\mathbf{X})} \leq t, \frac{|S(\mathbf{X}) - S(\mathbf{X}')|}{s(\mathbf{X})} \leq x\right] + \mathbb{P}\left[\frac{|S(\mathbf{X}) - S(\mathbf{X}')|}{s(\mathbf{X})} \geq x\right]$$
$$\leq \mathbb{P}\left[\frac{S(\mathbf{X}')}{s(\mathbf{X})} \leq t + x\right] + \frac{1}{x^{2k}}\mathbb{E}\left[\left|\frac{S(\mathbf{X}) - S(\mathbf{X}')}{s(\mathbf{X})}\right|^{2k}\right], \tag{6-9}$$

pela desigualdade de Markov.

Por causa de 6-8, o primeiro termo da soma do lado direito de 6-9 é menor do que

$$\Phi(t+x) + 4C_\gamma \frac{n\log(n)}{s^3(\mathbf{X})} \leq \Phi(t) + \frac{x}{\sqrt{2\pi}} + 4C_\gamma \frac{n\log(n)}{s^3(\mathbf{X})}. \tag{6-10}$$

Isso decorre do fato que

$$\Phi(t+x) = \Phi(t) + \int_t^{t+x} f(y)dy \leq \Phi(t) + \frac{x}{\sqrt{2\pi}},$$

pois $f(y) \leq \frac{1}{\sqrt{2\pi}}, \forall y$.

Para controlar o segundo termo da soma do lado direito de 6-9, notamos primeiro que

$$S(\mathbf{X}) - S(\mathbf{X}') = \sum_{i=\tau+1}^{2n} (X_i - X_i'), \tag{6-11}$$

onde colocamos $X_i = 0$ para $i > n$. Dado que $\tau + 1$ é um tempo de parada, condicional em τ, $(X_i - X_i')_{i \geq \tau+2}$ ainda forma uma sequência de diferença de martingais.

Assim, aplicando a Proposição 3.5 à sequência de diferença de martingais $(X_i - X_i')_{i \geq \tau+2}$, temos

$$\frac{1}{C}\mathbb{E}\left[\left|\sum_{i=\tau+2}^{2n}(X_i - X_i')\right|^{2k}\right] \leq \mathbb{E}\left[\left(\sum_{i=\tau+2}^{2n}\mathbb{E}[(X_i - X_i')^2|\mathcal{F}_{i-1}]\right)^k\right] + \mathbb{E}\left[\max_{\tau+2\leq i\leq 2n}|X_i - X_i'|^{2k}\right], \tag{6-12}$$

e podemos descartar a parcela da soma indexada por $\tau+1$ que aparece em 6-11, que é uniformemente limitada.

O máximo no lado direito de 6-12 é limitado por $2\gamma^{2k}$. Então

$$\frac{1}{C}\mathbb{E}\left[\left|\sum_{i=\tau+2}^{2n}(X_i - X_i')\right|^{2k}\right] \leq \mathbb{E}\left[\left(\sum_{i=\tau+2}^{2n}\mathbb{E}[(X_i - X_i')^2|\mathcal{F}_{i-1}]\right)^k\right] + 2\gamma^{2k}. \tag{6-13}$$

Quanto ao segundo termo da soma, como X_i e X_i' são variáveis aleatórias ortogonais, temos

$$\begin{aligned}\sum_{i=\tau+1}^{2n}\mathbb{E}[(X_i - X_i')^2|\mathcal{F}_{i-1}] &= \sum_{i=\tau+1}^{2n}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}] + \sum_{i=\tau+1}^{2n}\mathbb{E}[X_i'^2|\mathcal{F}_{i-1}]\\ &= \sum_{i=1}^{2n}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}] - \sum_{i=1}^{\tau}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}] + s^2(\mathbf{X})\\ &- \sum_{i=1}^{\tau}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}]\\ &= \sum_{i=1}^{n}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}] + s^2(\mathbf{X}) - 2\sum_{i=1}^{\tau}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}],\end{aligned}$$

pois $\sum_{i=1}^{2n}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}] = \sum_{i=1}^{n}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}]$ uma vez que $X_i = 0$ para todo $i > n$. Das notações que foram apresentadas na primeira seção do capítulo, temos que

$$\sum_{i=\tau+1}^{2n}\mathbb{E}[(X_i - X_i')^2|\mathcal{F}_{i-1}] = s^2(\mathbf{X})V^2(\mathbf{X}) + s^2(\mathbf{X}) - 2\sum_{i=1}^{\tau}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}]. \tag{6-14}$$

Agora, se $\tau = n$, o termo $\sum_{i=1}^{\tau}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}]$ é equivalente à $s^2(\mathbf{X})V^2(\mathbf{X})$, por definição. Caso contrário, $\sum_{i=1}^{\tau+1}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}] > s^2(\mathbf{X})$. Mas como os incrementos são limitados, então necessariamente

$$\sum_{i=1}^{\tau}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}] \geq s^2(\mathbf{X}) - \gamma^2.$$

Por conseguinte

$$\sum_{i=1}^{\tau}\mathbb{E}[X_i^2|\mathcal{F}_{i-1}] \geq \min\{s^2(\mathbf{X})V^2(\mathbf{X}), s^2(\mathbf{X}) - \gamma^2\}.$$

Assim sendo, de 6-13 e 6-14, obtemos

$$\sum_{i=\tau+1}^{2n}\mathbb{E}[(X_i - X_i')^2|\mathcal{F}_{i-1}] \leq |s^2(\mathbf{X})V^2(\mathbf{X}) - s^2(\mathbf{X})| + 2\gamma^2. \tag{6-15}$$

De 6-15, pela monotonicidade da esperança, e como $k \in [1, +\infty)$,

$$\mathbb{E}\left[\left(\sum_{i=\tau+1}^{2n} \mathbb{E}[(X_i - X_i')^2 | \mathcal{F}_{i-1}]\right)^k\right] \leq \mathbb{E}\left[\left(|s^2(\mathbf{X})V^2(\mathbf{X}) - s^2(\mathbf{X})| + 2\gamma^2\right)^k\right].$$

Combinando o resultado obtido com 6-13, concluímos que

$$\begin{aligned}\frac{1}{C}\mathbb{E}\left[\left|\sum_{i=\tau+2}^{2n} (X_i - X_i')\right|^{2k}\right] &\leq \mathbb{E}\left[\left(|s^2(\mathbf{X})V^2(\mathbf{X}) - s^2(\mathbf{X})| + 2\gamma^2\right)^k\right] + 2\gamma^{2k}\\ &\leq \mathbb{E}\left[\left(s^2(\mathbf{X})|V^2(\mathbf{X}) - 1| + 2\gamma^2\right)^k\right] + 2\gamma^{2k}\\ &\leq \mathbb{E}\left[k\left(s^{2k}(\mathbf{X})|V^2(\mathbf{X}) - 1|^k + 2^k\gamma^{2k}\right)\right] + 2\gamma^{2k}\\ &\leq ks^{2k}(\mathbf{X})||V^2(\mathbf{X}) - 1||_k^k + \gamma^{2k}(2^k k + 2).\end{aligned}$$

Logo, teremos

$$\mathbb{E}\left[\left|\sum_{i=\tau+2}^{2n} (X_i - X_i')\right|^{2k}\right] \leq C\left(s^{2k}(\mathbf{X})||V^2(\mathbf{X}) - 1||_k^k + \gamma^{2k}\right).$$

Combinando este resultado com 6-11, chegamos a

$$\frac{1}{x^{2k}}\mathbb{E}\left[\left|\frac{S(\mathbf{X}) - S(\mathbf{X}')}{s(\mathbf{X})}\right|^{2k}\right] \leq \frac{C}{x^{2k}}\left(||V^2(\mathbf{X}) - 1||_k^k + \frac{\gamma^{2k}}{s^{2k}(\mathbf{X})}\right). \tag{6-16}$$

Finalmente, através das equações 6-16, 6-10 e 6-9, obteremos

$$\mathbb{P}\left[\frac{S(\mathbf{X})}{s(\mathbf{X})} \leq t\right] - \Phi(t) \leq \frac{x}{\sqrt{2\pi}} + 4C_\gamma \frac{n\log(n)}{s^3(\mathbf{X})} + \frac{C}{x^{2k}}\left(||V^2(\mathbf{X}) - 1||_k^k + \frac{\gamma^{2k}}{s^{2k}(\mathbf{X})}\right).$$

Otimizando isso para $x > 0$ teremos a estimativa correta do Teorema. O limite inferior, ou seja, $\Phi(t) - \mathbb{P}\left[\frac{S(\mathbf{X})}{s(\mathbf{X})} \leq t\right]$, é controlado da mesma maneira. ■

Para nossa aplicação, será útil usarmos uma versão alternativa desse resultado, que será apresentada abaixo em forma de corolário.

Corolário 6.4 *Seja $k \in [1, +\infty)$ e $R \in (0, +\infty)$. Seja $\mathbf{Y} = (Y_1, \ldots, Y_n) \in \mathcal{M}_n$ tal que $||\mathbf{Y}||_\infty \leq R$. Então, existe uma constante $C_k > 0$ tal que para todo $n \geq 2$,*

$$D(\mathbf{Y}) \leq C_k\left[\frac{R^3 n\log(n)}{s^3(\mathbf{Y})} + ||V^2(\mathbf{Y}) - 1||_k^{\frac{k}{2k+1}} + R^{\frac{2k}{2k+1}} s^{\frac{-2k}{2k+1}}(\mathbf{Y})\right]. \tag{6-17}$$

Prova. Seja $Y_i' = \frac{Y_i}{R}$. Então, $\mathbf{Y}' = (Y_1', \ldots, Y_n') \in \mathcal{M}_n$ e $||\mathbf{Y}'||_\infty \leq 1$.

Portanto, podemos aplicar o Teorema 6.3 para essa sequência. Assim, existe uma constante $C_k' > 0$ tal que para todo $n \geq 2$,

$$\begin{aligned} D(\mathbf{Y}') &\leq C_k' \left[\frac{n \log(n)}{s^3(\mathbf{Y}')} + (||V^2(\mathbf{Y}') - 1||_k^k + s^{-2k}(\mathbf{Y}'))^{1/(2k+1)} \right] \\ &\leq C_k' \left[\frac{n \log(n)}{s^3(\mathbf{Y}')} + ||V^2(\mathbf{Y}') - 1||_k^{\frac{k}{2k+1}} + s^{\frac{-2k}{2k+1}}(\mathbf{Y}') \right], \end{aligned}$$

pois $(a+b)^x \leq a^x + b^x$ para todo $x < 1$. Por meio da aplicação do lema 3.1, temos que

$$D(\mathbf{Y}) \leq C_k' \left[\frac{n \log(n)}{s^3(\mathbf{Y}')} + ||V^2(\mathbf{Y}) - 1||_k^{\frac{k}{2k+1}} + s^{\frac{-2k}{2k+1}}(\mathbf{Y}') \right]$$

Finalmente, uma vez que $s^2(\mathbf{Y}') = \frac{\sum_{i=1}^n \mathbb{E}[Y_i'^2]}{R^2}$, então $s^2(\mathbf{Y}) = R^2 s^2(\mathbf{Y}')$. Assim,

$$D(\mathbf{Y}) \leq C_k \left[\frac{R^3 n \log(n)}{s^3(\mathbf{Y})} + ||V^2(\mathbf{Y}) - 1||_k^{\frac{k}{2k+1}} + R^{\frac{2k}{2k+1}} s^{\frac{-2k}{2k+1}}(\mathbf{Y}) \right].$$

■

Finalmente, vamos justificar a otimização do termo $||V^2 - 1||_k^{k/(2k+1)}$, que aparece no lado direito da desigualdade em 6-7.

Teorema 6.5 *Seja $k \in [1, +\infty)$ e $\alpha \in (1/2, 1)$. Existe uma sequência de elementos $\mathbf{X}_n \in \mathcal{M}_n$, tal que*

$$||\mathbf{X}_n||_\infty \leq 2,$$

$$s(\mathbf{X}_n) \simeq \sqrt{n},$$

$$||V^2 - 1||_k^{k/(2k+1)} = O(n^{(\alpha-1)/2}),$$

$$\limsup_{n\to\infty} n^{(1-\alpha)/2} D(\mathbf{X}_n) > 0.$$

Prova. A prova desse teorema pode ser encontrado em [11], página 7. ■

Escrevemos $a_n \simeq b_n$ se existe $C > 0$ tal que $a_n/C \leq b_n \leq C a_n$ para todo n suficientemente grande.

A ideia por trás do Teorema 6.5 é que há sequência de elementos que estão dentro das condições do Teorema 6.3 e que com esta não é possível encontrar um termo mais forte que $||V^2 - 1||_k^{k/(2k+1)}$, ou seja, um termo que vá para zero mais rapidamente do que este.

7
Contagem de subgrafos em grafos aleatórios Erdós-Rényi

Neste capítulo, vamos falar do artigo [5], que aborda o assunto de desvios moderados na contagem de subgrafos nos modelos de grafos aleatórios Erdős-Rényi.

Inicialmente, veremos noções e definições básicas da Teoria de grafos para entendermos um pouco sobre a natureza dessas estruturas. Deste modo, poderemos desenvolver as intuições que serão utilizadas ao longo do artigo.

Uma contribuição importante do artigo mencionado é a elaboração de uma expressão martingal para contar desvios de um determinado subgrafo H no modelo $G(n, m)$. Além disso, será importante entendermos o comportamento da variação quadrática discreta desses martingais. Deduziremos que essa variação é previsível e, com alta probabilidade, muito perto de uma particular função determinística.

7.1
Grafos Aleatórios

Primeiramente, iremos definir os conceitos de grafos e subgrafos, além de apresentar os dois modelos de grafos aleatórios que são vistos com maior frequência.

Definição 7.1 *Um **grafo** G é um par (V, E), onde V é um conjunto de objetos denominados **vértices** e E é um conjunto de subconjuntos de V de ordem dois chamados de **arestas**.*

Tendo em vista que frequentemente utilizaremos a letra e para denotar o número de arestas de grafos, não faremos uso do símbolo e para expressar a base do logaritmo neperiano, ao qual usaremos $\exp(1)$.

Definição 7.2 *Um **subgrafo** de G é um grafo G' tal que $V(G') \subset V(G)$ e $E(G') \subset E(G)$.*

Agora, iremos analisar os modelos de grafos aleatórios $G(n, p)$ e $G(n, m)$.

No modelo $G(n,p)$, onde $0 < p < 1$, um grafo aleatório com o conjunto de vértices $V = \{1, 2, \dots, n\}$ é construído incluindo cada aresta independentemente e com probabilidade p. Em outras palavras, se H é um grafo com o conjunto de vértices V e com m arestas, então

$$\mathbb{P}(G = H) = p^m (1-p)^{N-m},$$

onde $N = \binom{n}{2}$. (Utilizaremos essa notação daqui em diante para denotar o máximo número de arestas possíveis de um grafo com n vértices.)

O modelo $G(n,m)$ consiste em todos os grafos com o conjunto de vértices $V = \{1, 2, \dots, n\}$ tendo m arestas, onde cada grafo G é selecionado uniformemente. Assim, $0 \leq m \leq N$, e $G(n,m)$ possui $\binom{N}{m}$ elementos e cada elemento ocorre com probabilidade $\binom{N}{m}^{-1}$.

Nosso foco será no modelo de grafos aleatórios $G(n,m)$, pois esse modelo apresenta um contexto mais natural para se estudar desvios moderados em contagens de subgrafos. O estudo desses desvios é o foco do artigo [5], que iremos trabalhar na seção seguinte.

Para maior aprofundamento referente ao estudo de grafos aleatórios, recomenda-se a leitura de [20].

7.2 Desvios em Contagens de Subgrafos

Escrevemos $N_H(G)$ para o número de cópias isomórficas de um subgrafo H em um grafo G, onde H é um subgrafo fixo e G é um grande grafo aleatório. Deste modo, para um subgrafo H com v vértices e e arestas, o valor esperado do número de cópias de H em $G(n,m)$ é

$$L_H(m) \coloneqq \frac{(n)_v (m)_e}{(N)_e}, \tag{7-1}$$

onde $(n)_k \coloneqq n(n-1)\dots(n-k+1)$ denota o fatorial decrescente.

Uma maneira natural de gerar $G(n,m)$ é adicionando as arestas uma a uma. O processo Érdos-Rényi de grafos aleatórios $(G_i : i = 0, \dots, N)$ é definido da seguinte maneira:

Seja G_0 o grafo vazio, e para cada $i \geq 0$, G_{i+1} é obtido adicionando-se uniformemente uma aresta em G_i. Claramente, G_m tem distribuição $G(n, m)$. O processo termina com G_N sendo o grafo completo K_n.

Portanto, escrevemos $D_H(G_m)$ para o desvio na contagem de H em G_m. Ou seja,

$$D_H(G_m) \coloneqq N_H(G_m) - L_H(m). \tag{7-2}$$

Enunciaremos agora um teorema que fornece um limite geral, porém não tão preciso, sobre os desvios de contagens de subgrafos.

Teorema 7.3 *Seja H um grafo com v vértices e e arestas. Então existe uma constante $c = c(H)$ tal que para todo $t = t(n) \in (0, 1)$, onde $t = m/N$, e para todo $\alpha, n \geq c^{-1}$, temos*

$$\mathbb{P}\left(|D_H(G_m)| > \alpha n^{v-3/2}\right) \leq \exp(-c\alpha \min\{\alpha, n^{1/2}\}).$$

Prova. A demonstração pode ser encontrada em [5], página 44. ■

Veremos que caminhos de tamanho dois, que denotaremos como $\wedge$, e triângulos, que denotaremos como $\triangle$, terão um papel particularmente importante. Escrevemos $\binom{H}{\wedge}$ para o número de caminhos de tamanho dois em H, e $\binom{H}{\triangle}$ para o número de triângulos em H. É importante mencionar que estes números são diferentes dos números de cópias isomórficas de H definidas no Capítulo 1.

Nas próximas seções, analisaremos os resultados obtidos por [5] em relação a dois objetivos principais: Estabelecer uma expressão martingal para o desvio na contagem de subgrafos $D_H(G_m)$ e entender o comportamento da variação quadrática discreta dessa expressão martingal.

7.3 Expressão Martingal para $D_H(G_m)$

Seja n fixo e seja $(G_m : m = 0, \dots, N)$ a realização do processo Erdős-Rényi de grafos aleatórios em n vértices. Definimos

$$A_H(G_m) \coloneqq N_H(G_m) - N_H(G_{m-1}) \tag{7-3}$$

como o número de cópias isomórficas de H criadas com a adição da m-ésima aresta. Nossa expressão martingal será baseada na versão centralizada dessas variáveis aleatórias. Seja

$$X_H(G_m) \coloneqq A_H(G_m) - \mathbb{E}[A_H(G_m)|G_{m-1}]. \tag{7-4}$$

Note que $X_H(G_m)$ é obtido de $A_H(G_m)$, alterando-o tal que $\mathbb{E}[X_H(G_m)|G_{m-1}] = 0$. Podemos, então, enunciar a expressão martingal para $D_H(G_m)$.

Teorema 7.4 *Seja H um subgrafo com v vértices e e arestas. Então*

$$D_H(G_m) = \sum_{i=1}^{m} \sum_{F \subseteq E(H)} \frac{(N-m)_{e(F)}(m-i)_{e-e(F)}}{(N-i)_e} X_F(G_i), \tag{7-5}$$

onde a soma interna é tomada sobre todos os 2^e subgrafos F com $V(F) = V(H)$ e $E(F) \subseteq E(H)$.

Repare que $X_H(G_m)$ é um martingal com respeito à filtração de $G_0, \ldots, G_N$. Como

$$\sum_{F \subseteq E(H)} \frac{(N-m)_{e(F)}(m-i)_{e-e(F)}}{(N-i)_e} X_F(G_i)$$

é combinação linear das variáveis aleatórias $X_F(G_i)$, isso também é martingal e, portanto, 7-5 é, de fato, martingal.

Antes de provarmos o Teorema 7.4, enunciaremos um lema a respeito de $\mathbb{E}[A_H(G_m)|G_{m-1}]$, o valor esperado de cópias de H criadas com a adição da m-ésima aresta, dado o grafo G_{m-1}.

Lema 7.5 *No processo Erdős-Rényi de grafos aleatórios $(G_m : m = 0, \ldots, N)$,*

$$\begin{aligned}\mathbb{E}[A_H(G_m)|G_{m-1}] &= \frac{1}{N-m+1} \sum_{f \in E(H)} (N_{H\setminus f}(G_{m-1}) - N_H(G_{m-1})) \\ &= (L_H(m) - L_H(m-1)) \\ &\quad + \frac{1}{N-m+1} \sum_{f \in E(H)} (D_{H\setminus f}(G_{m-1}) - D_H(G_{m-1})),\end{aligned}$$

onde $H\setminus f$ representa o grafo obtido ao se remover a aresta f de H.

Prova.(Lema 7.5) A demonstração pode ser encontrada em [5], página 13. ■

Prova.(Teorema 7.4) A prova é por indução em $e = e(H)$, e em $m \in \{0, \dots, N\}$. Se $e = 1$ ou $m = 0$, o resultado é trivial. Agora, consideramos um grafo H com v vértices e $e \geq 2$ arestas e $m \in \{1, \dots, N\}$. Então, podemos expandir $D_H(G_m)$ da seguinte maneira:

$$\begin{aligned}
D_H(G_m) &= N_H(G_m) - L_H(G_m) \\
&= N_H(G_{m-1}) + A_H(G_m) - L_H(m-1) - (L_H(m) - L_H(m-1)) \\
&= D_H(G_{m-1}) + A_H(G_m) - (L_H(m) - L_H(m-1)) \\
&= D_H(G_{m-1}) + X_H(G_m) + \mathbb{E}[A_H(G_m)|G_{m-1}] - (L_H(m) - L_H(m-1)) \\
&= D_H(G_{m-1}) + X_H(G_m) \\
&+ \frac{1}{N-m+1} \sum_{f \in E(H)} (D_{H \setminus f}(G_{m-1}) - D_H(G_{m-1})),
\end{aligned}$$

onde usamos a definição 7-4 na quarta linha e o Lema 7.5 na última linha.

Assim, temos uma expressão para $D_H(G_m)$ em termos de $X_H(G_m)$ e uma combinação linear dos desvios $D_F(G_{m-1})$ com $F \subseteq H$. Pela hipótese de indução, cada um desses desvios pode ser expresso como uma combinação linear de $X_F(G_i)$ com $F \subseteq H$ e $1 \leq i \leq m$. Pode-se verificar que a expressão resultante de $D_H(G_m)$ é a afirmada pelo Teorema.

Mostrando que os coeficientes que multiplicam $X_F(G_i)$ são iguais entre $D_H(G_m)$ e o outro lado da igualdade, é fácil ver que a igualdade se mantém.

Por exemplo, no caso em que $F = H$, se $i = m$, o coeficiente é 1. Se $i \leq m-1$, o coeficiente em ambos os lados será $\frac{(N-m)_e}{(N-i)_e}$.

Os casos onde $F \neq H$ podem ser verificados da mesma maneira. ■

Apresentaremos agora uma expressão mais simples que aproxima $D_H(G_m)$ muito bem. Para um grafo H com v vértices e e arestas, definimos

$$\begin{aligned}
\mathbb{X}_H(G_i; t) := n^{v-3} \Bigg(& t^{e-2} \binom{H}{\wedge} \frac{(1-t)^2}{(1-s)^2} X_\wedge(G_i) \\
& + t^{e-3} \binom{H}{\triangle} \frac{(1-t)^3}{(1-s)^3} (X_\triangle(G_i) - 3s X_\wedge(G_i)) \Bigg),
\end{aligned} \tag{7-6}$$

onde $s = i/N$ e $t = m/N$, notações que serão usadas daqui em diante, e definimos

$$\Lambda_H(G_m) := \sum_{i=1}^{m} \mathbb{X}_H(G_i; t). \tag{7-7}$$

Teorema 7.6 *Seja H um grafo com v vértices e e arestas. Existe uma constante $C = C(H)$ tal que para todo $t = t(n) \in (0,1)$, temos*

$$\mathbb{P}\Big(|D_H(G_m) - \Lambda_H(G_m)| > Cbt^{1/2}n^{v-2}\Big) \le \exp(-b) \tag{7-8}$$

para todo $3\log n \le b \le t^{1/2}n$. *Além disso,*

$$\mathbb{P}\Big(|D_H(G_m) - \Lambda_H(G_m)| > Cbn^{v-2}\Big) \le \exp(-b) \tag{7-9}$$

para todo $b \ge 3\log n$.

Prova. A demonstração pode ser encontrada em [5], página 35. ■

Intuitivamente, o Teorema 7.6 nos diz que $D_H(G_m)$ pode ser aproximado por $\Lambda_H(G_m)$, pois a probabilidade de seus valores serem significativamente diferentes é muito baixa, para n grande.

Ademais, existe um motivo para expressarmos os termos $\mathbb{X}_H(G_i;t)$ de $\Lambda_H(G_m)$ como combinação linear de $X_\wedge$ e $X_\triangle - 3sX_\wedge$ ao invés de diretamente como combinação linear de $X_\wedge$ e $X_\triangle$. Essa escolha é considerada natural porque $X_\wedge$ e $X_\triangle - 3sX_\wedge$ são assintoticamente ortogonais, ou seja, não-correlacionados, no sentido de que

$$\mathbb{E}\Big[X_\wedge(G_i)(X_\triangle(G_i) - 3sX_\wedge(G_i))|G_{i-1}\Big]$$

é tipicamente $o(n)$, enquanto suas variâncias individuais são $\Theta(n)$.

7.4
Graus e Cograus em $G(n,m)$

Denotaremos $d_u(G)$ para o **grau** de um vértice u em um grafo G. No caso $G_m \sim G(n,m)$, o grau esperado de u é $2m/n$, e portanto

$$D_u(G_m) := d_u(G_m) - \frac{2m}{n} \tag{7-10}$$

é o desvio entre o grau de u e sua média.

Também consideraremos os **cograus**, escrevendo $d_{u,v}(G)$ para o número de vizinhos em comum entre os vértices u e v num grafo G, e

$$D_{u,v}(G_m) := d_{u,v}(G_m) - \frac{(n-2)(m)_2}{(N)_2} \tag{7-11}$$

para o desvio entre $d_{u,v}(G_m)$ e sua média.

Primeiramente, mostraremos um resultado sobre o desvio máximo de graus. Seja

$$D_{\max}(G_m) \coloneqq \max_u D_u(G_m).$$

Lema 7.7 *Para todo $b \geq \log n$ e todo $m \leq N$, temos*

$$\mathbb{P}\Big(D_{\max}(G_m) > 4b^{1/2}t^{1/2}n^{1/2} + 4b\Big) \leq \exp(-b).$$

Prova. Fixe um vértice $u \in V(G_m)$ e seja $a = 4b^{1/2}t^{1/2}n^{1/2} + 4b$. O grau $d_u(G_m)$ tem distribuição hipergeométrica com parâmetros $N, n-1, m$. Portanto, temos

$$\begin{aligned}\mathbb{P}(D_u(G_m) > a) &\leq \exp\left(\frac{-a^2}{2tn + a}\right)\\ &\leq \exp(-2b).\end{aligned}$$

Isso decorre de propriedades da distribuição hipergeométrica. Dados N, K, m, a variável aleatória S_m tem distribuição hipergeométrica com parâmetros N, K, m se $\mathbb{P}(S_m = k) = \binom{K}{k}\binom{N-K}{m-k}/\binom{N}{m}$. Se $\mu = \mathbb{E}[S_m] = Km/N$, temos o seguinte limite superior, derivado da desigualdade de Chernoff:

$$\begin{aligned}\mathbb{P}(S_m \geq \mu + a) &\leq \exp\left(\frac{-a^2}{2\mu + 2a/3}\right)\\ &\leq \exp\left(\frac{-a^2}{2\mu + a}\right).\end{aligned}$$

Deste modo, por sub-aditividade, a união sobre os $n \leq \exp(b)$ vértices completa a prova da afirmação principal. ■

Analogamente, podemos enunciar um resultado sobre o desvio máximo de cograus. Seja

$$D'_{\max}(G_m) \coloneqq \max_{u,v} D_{u,v}(G_m).$$

Lema 7.8 *Para todo $b \geq 2\log n$ e todo $m \leq N$, temos*

$$\mathbb{P}\Big(D'_{\max}(G_m) > 4b^{1/2}t^{1/2}n^{1/2} + 8b\Big) \leq \exp(-b).$$

Prova. A prova desse resultado é essencialmente idêntica à do Lema 7.7. ■

7.5 Variância e Covariância dos incrementos $X_F(G_i)$

O principal objetivo desta subseção é mostrar que a variância condicional $Var(X_F(G_i)|G_{i-1})$ de $X_F(G_i)$ e a covariância condicional

$$\mathbb{E}[X_F(G_i)X_{F'}(G_i)|G_{i-1}] \tag{7-12}$$

de $X_F(G_i)$ e $X_{F'}(G_i)$ são previsíveis, no sentido que são geralmente próximas de certas funções determinísticas.

Dados dois grafos, F com v vértices e e arestas, e F' com v' vértices e e' arestas, definimos

$$V_{F,F'}(i,n) \coloneqq n^{v+v'-5}s^{e+e'-4}(1-s)\Big(s\theta_1(F,F') + (1-s)\theta_2(F,F')\Big), \tag{7-13}$$

onde

$$\theta_1(F,F') \coloneqq 8\binom{F}{\wedge}\binom{F'}{\wedge} \quad \text{e} \quad \theta_2(F,F') \coloneqq 36\binom{F}{\triangle}\binom{F'}{\triangle}.$$

Proposição 7.9 *Sejam F, F' grafos com v, v' vértices (respectivamente) e e, e' arestas (respectivamente) e seja $t \in (0,1)$. Existe uma constante $C = C(F,F',t)$ tal que, para todo $1 \le i \le tN$ e todo $3\log n \le b \le n/2C$, temos*

$$\mathbb{P}\left(\Big|\mathbb{E}[X_F(G_i)X_{F'}(G_i)|G_{i-1}] - V_{F,F'}(i,n)\Big| > Cb^{1/2}n^{v+v'-11/2}\right) \le \exp(-b).$$

Como o caminho de tamanho dois e o triângulo são de particular importância, é interessante notar que nesses casos temos

$$V_{\wedge,\wedge}(i,n) = 8ns(1-s),$$

$$V_{\wedge,\triangle}(i,n) = 24ns^2(1-s) \text{ e}$$

$$V_{\triangle,\triangle}(i,n) = 36ns^2(1-s^2).$$

De fato, [5] demonstra um resultado ainda mais preciso que inclui um termo de segunda ordem relativo ao desvio $D_\wedge(G_{i-1})$. Definimos

$$W_{F,F'}(G_{i-1}) \coloneqq 8n^{v+v'-7}s^{e+e'-4}\binom{F}{\wedge}\binom{F'}{\wedge}D_\wedge(G_{i-1}). \tag{7-14}$$

Proposição 7.10 *Sejam F, F' grafos com v, v' vértices (respectivamente) e e, e' arestas (respectivamente) e seja $t \in (0,1)$. Existe uma constante $C = C(F, F', t)$ tal que, para todo $1 \le i \le tN$ e todo $3\log n \le b \le n/2C$, temos*

$$\mathbb{P}\Big(\Big|\mathbb{E}[X_F(G_i)X_{F'}(G_i)|G_{i-1}] - \Big(V_{F,F'}(i,n) + W_{F,F'}(G_{i-1})\Big)\Big| > Cbn^{v+v'-6}\Big)$$
$$\le \exp(-b).$$

Prova. A demonstração pode ser encontrada em [5], página 48. ■

O termo $W_{F,F'}(G_{i-1})$ é geralmente muito menor do que o termo principal $V_{F,F'}(i,n)$. Isso segue do fato que $D_\wedge(G_{i-1})$ é geralmente muito menor do que n^2, que segue do Teorema 7.3.

Observamos que, enquanto $V_{F,F'}(i,n)$ é um valor determinístico, $W_{F,F'}(G_{i-1})$ é uma variável aleatória que depende do processo de geração de $G(n,m)$. Intuitivamente, a variável $W_{F,F'}(G_{i-1})$ é uma forma de medição da irregularidade dos graus dos vértices, ou seja, mede a variação entre os graus dos vértices do grafo.

Observamos agora que a Proposição 7.9 segue da Proposição 7.10 e do Teorema 7.3.

Prova.(Proposição 7.9) Sejam F, F' e t fixos. Escrevendo C_1 para a constante da Proposição 7.10 e c para a constante associada a $H = \wedge$ do Teorema 7.3, definimos

$$C = 2\Big(C_1 + 8c^{-1}\binom{F}{\wedge}\binom{F'}{\wedge}\Big).$$

Pela desigualdade triangular,

$$\Big|\mathbb{E}[X_F(G_i)X_{F'}(G_i)\,\Big|\,G_{i-1}] \;-\; V_{F,F'}(i,n)\Big|$$
$$\le \Big|\mathbb{E}[X_F(G_i)X_{F'}(G_i)\,\Big|\,G_{i-1}] \;-\; \Big(V_{F,F'}(i,n) + W_{F,F'}(G_{i-1})\Big)\Big|$$
$$+\; |W_{F,F'}(i,n)|\,,$$

e portanto o evento

$$\Big|\mathbb{E}[X_F(G_i)X_{F'}(G_i)|G_{i-1}] - V_{F,F'}(i,n)\Big| > Cb^{1/2}n^{v+v'-11/2} \tag{7-15}$$

só pode ocorrer se

$$\begin{aligned}\left|\mathbb{E}[X_F(G_i)X_{F'}(G_i)|G_{i-1}] - \Big(V_{F,F'}(i,n) + W_{F,F'}(G_{i-1})\Big)\right| &> 2C_1 b^{1/2} n^{v+v'-11/2} \\ &\geq 2C_1 b n^{v+v'-6}\end{aligned}$$

ou

$$|D_\wedge(G_{i-1})| > 2c^{-1} b^{1/2} n^{3/2}.$$

Pela Proposição 7.10 e o Teorema 7.3, respectivamente, cada um desses eventos tem probabilidade no máximo $\exp(-2b)$. Por sub-aditividade, o evento 7-15 tem probabilidade no máximo $2\exp(-2b) \leq \exp(-b)$, como se queria demonstrar. ■

8
Taxa de convergência do Teorema Central do Limite para desvios de subgrafos

Nos capítulos anteriores, estudamos martingais e Teorema Central do Limite, focando na relação entre ambos e na velocidade em que a convergência para uma distribuição normal padrão ocorre, isto é, nas taxas de convergência no Teorema Central do Limite para martingais. Além disso, abordamos os desvios moderados na contagem de subgrafos nos modelos de grafos aleatórios Erdős-Rényi e estudamos a expressão martingal que conta tais desvios.

Portanto, a ideia central deste capítulo será unificar os temas abordados até então, ou seja, estabelecer uma taxa de convergência no Teorema Central do Limite para a expressão martingal específica do desvio na contagem de subgrafos $D_H(G_m)$, ou seja, o resultado do Teorema 1.2.

Isso seguirá facilmente do resultado a seguir, que trata da variável aleatória $\Lambda_H(G_m)$ que aproxima bem $D_H(G_m)$. Para o caso em que H é livre de triângulos, a expressão para $\Lambda_H(G_m)$ é

$$\Lambda_H(G_m) \;=\; n^{v-3}t^{e-2}\binom{H}{\wedge}\sum_{i=1}^{m}\frac{(1-t)^2}{(1-s)^2}X_{\wedge}(G_i)\,.$$

Teorema 8.1 *Seja $k \in [1, +\infty)$, $t = m/N$ e $\eta > 0$. Então, existe uma constante $C_{k,\eta}$ tal que o seguinte vale. Para todo grafo H livre de triângulos, todo $n \geq 2$ e todo $\eta N \leq m \leq N/2$, temos*

$$d(\Lambda_H(G_m)) \leq C_{k,\eta}\left[\left(\frac{\sqrt{\log(n)}}{tn^{1/2}}\right)^{\frac{k}{2k+1}}\right].$$

Munidos desse resultado, estaremos prontos para a demonstração do Teorema 1.2.

Prova.(do Teorema 1.2)

Pelo Teorema 7.6 aplicado com $b = 3\log(n)$, temos que

$$\mathbb{P}\Big(|D_H(G_m) - \Lambda_H(G_m)| > C\log(n)n^{v-2}\Big) \leq n^{-3},$$

onde v é o número de vértices de H.

Observamos que o σ_{Λ_H} do desvio $\Lambda_H(G_m)$ é da ordem de magnitude

$$\begin{aligned}\sigma_{\Lambda_H} &= \Theta(n^{v-3}t^{e-2}\sigma_{X_\wedge})\\ &= \Theta(n^{v-3/2}t^{e-1}),\end{aligned}$$

pois $\sigma_{X_\wedge} = \Theta(tn^{3/2})$, uma vez que

$$\begin{aligned}\sum_{i=1}^{m} \mathbb{E}[X_\wedge^2(G_i)|G_{i-1}] &= \Theta(nstn^2)\\ &= \Theta(t^2n^3).\end{aligned}$$

Deste modo, aplicando o Lema 3.7 com $\delta = n^{-3}$ e $\epsilon = \frac{C\log(n)n^{v-2}}{n^{v-3/2}t^{e-1}} = \frac{C\log(n)}{n^{1/2}t^{e-1}}$, obtemos

$$\begin{aligned}d(D_H(G_m)) &\leq 3d(\Lambda_H(G_m)) + 2n^{-3} + \frac{C\log(n)}{n^{1/2}t^{e-1}}\\ &\leq C_{k,H,\eta}\left[\left(\frac{\sqrt{\log(n)}}{tn^{1/2}}\right)^{\frac{k}{2k+1}}\right],\end{aligned}$$

pelo Teorema 8.1. ■

A princípio, estamos quase prontos para deduzir o Teorema 8.1 do Teorema 6.3, a versão quantificada do Teorema Central do Limite para martingais. Contudo, em alguns casos os incrementos $X_\wedge(G_i)$ podem ser muito grandes, o que poderia piorar a cota obtida. Por isso, será útil considerarmos uma operação parecida com um truncamento, o que irá reduzir o máximo valor absoluto dos incrementos.

Assim, precisamos definir uma variável aleatória alternativa, $X_\wedge^*(G_i)$, em função de $X_\wedge(G_i)$ e de um tempo de parada τ, que veremos a seguir.

Primeiramente, definimos E_i como o evento

$$\{\exists v, \quad |d_v(G_i) - s(n-1)| > \sqrt{tn}\log(n)\}, \tag{8-1}$$

onde novamente $s = i/N$ e $t = m/N$. Informalmente, esse evento representa a existência de um vértice com um comportamento improvável, em que seu grau é muito diferente de sua média.

Assim, seja $\tau = \min\{i : G_i \in E_i\}$ e definimos

$$X^*_\wedge(G_i) = \begin{cases} X_\wedge(G_i) \text{ , se } i \leq \tau \\ W_i, \text{ se } i > \tau \end{cases} \tag{8-2}$$

onde $W_1, W_2, \dots$ é uma sequência de variáveis aleatórias i.i.d. com distribuição $\mathbb{P}(W_i = \sqrt{8ns(1-s)}) = \mathbb{P}(W_i = -\sqrt{8ns(1-s)}) = 1/2$, ou seja, $W_1, W_2, \dots$ corresponde a um passeio aleatório simples, independente da variável $X_\wedge(G_i)$, com $Var(W_i) = 8ns(1-s)$.

Proposição 8.2 *Seja $X^*_\wedge(G_i)$ definido acima. Então, para todo $n > exp(33)$,*

$$\mathbb{P}(\tau \leq m) \leq n^{-8}.$$

Em particular, $\mathbb{P}[\,\exists i \leq m,\, X^*_\wedge(G_i) \neq X_\wedge(G_i)] \leq n^{-8}$.

Prova. Em cada etapa $i = 1, \dots, m$, temos que $d_v(G_i)$ segue uma distribuição Hipergeométrica com parâmetros $N, n-1, i$ e $\mu = \mathbb{E}[d_v(G_i)] = s(n-1)$. Pela desigualdade de Chernoff, observamos que

$$\mathbb{P}(|d_v(G_i) - \mu| > a) \leq 2exp\left(\frac{-a^2}{3\mu}\right),$$

logo

$$\begin{aligned}\mathbb{P}(|d_v(G_i) - s(n-1)| > \sqrt{tn}\log(n)) &\leq 2exp\left(\frac{-tn(\log^2(n))}{3s(n-1)}\right) \\ &\leq 2exp\left(\frac{-(\log^2(n))}{3}\right) \\ &\leq 2n^{-11} \qquad \forall\, n > exp(33).\end{aligned}$$

Assim, como

$$\begin{aligned}\{\tau = i\} &\subseteq \bigcup_{v \in V(G_i)} \{|d_v(G_i) - s(n-1)| > \sqrt{tn}\log(n)\} \\ \Rightarrow \{\tau \leq m\} &\subseteq \bigcup_{v \in V(G_i)} \bigcup_{i \leq m} \{|d_v(G_i) - s(n-1)| > \sqrt{tn}\log(n)\} \\ \Rightarrow \mathbb{P}(\tau \leq m) &\leq \sum_{v \in V(G_i)} \sum_{i \leq m} \mathbb{P}(|d_v(G_i) - s(n-1)| > \sqrt{tn}\log(n)) \\ &\leq n^{-8}, \qquad \text{pois } m \leq n^2/2.\end{aligned}$$

Portanto,

$$\mathbb{P}[\,\exists i \leq m,\, X^*_\wedge(G_i) \neq X_\wedge(G_i)] \leq n^{-8}.$$

■

Agora, apresentaremos um resultado que será necessário para a Proposição seguinte. Essa Proposição será útil para a demonstração do Teorema 8.1.

Lema 8.3 *Seja $(\kappa_n)_{n\geq 1}$ uma sequência de números reais não-negativos, D constante real positiva, $\overline{\kappa} = \frac{\sum_{i=1}^n \kappa_i}{n}$ e $\kappa = \max_i \kappa_i$. Então, se $\kappa \leq D\overline{\kappa}$*

$$\sum_{i=1}^n i\kappa_i^2 \geq \frac{\kappa^2 n^2}{32D^4}.$$

Prova. Definimos l como o número de elementos da sequência $(\kappa_n)_{n\geq 1}$ tal que $\kappa_i \geq \overline{\kappa}/2$. Assim,

$$\begin{aligned}\sum_{i=1}^n \kappa_i &= n\overline{\kappa}\\ &\leq l\kappa + \frac{n\overline{\kappa}}{2},\end{aligned}$$

logo

$$l \geq \frac{n\overline{\kappa}}{2\kappa} \geq \frac{n}{2D}.$$

Seja $L = \{i_1, \ldots, i_l\}$ o conjunto dos índices tais que $\kappa_i \geq \overline{\kappa}/2$. Então, $i_j \geq j \quad \forall\, j = 1, \ldots, l$, portanto

$$\begin{aligned}\sum_{i=1}^n i\kappa_i^2 &\geq \sum_{j=1}^l i_j \left(\frac{\overline{\kappa}}{2}\right)^2\\ &\geq \frac{\overline{\kappa}^2}{4}\sum_{j=1}^l j\\ &\geq \frac{\overline{\kappa}^2 l^2}{8}\\ &\geq \frac{\overline{\kappa}^2 n^2}{32D^2}\\ &\geq \frac{\kappa^2 n^2}{32D^4}.\end{aligned}$$

■

Proposição 8.4 *Dado $D \in \mathbb{R}$, existe $C = C(D)$ tal que o seguinte vale: Seja $\kappa_1, \ldots, \kappa_m \in \mathbb{R}$ uma sequência de números reais não-negativos, $\kappa = \max_i \kappa_i$ e $\kappa \leq D\overline{\kappa}$ e seja $\mathbf{X} = (\kappa_1 X_\wedge(G_1), \ldots, \kappa_m X_\wedge(G_m)) \in M_m$. Então, para todo $r \in (3, \infty)$ e $t \geq 1/n$,*

$$\mathbb{P}\left(\left|V^2(\mathbf{X}) - 1\right| > \frac{Cr^{1/2}\sqrt{\log(n)}}{tn^{1/2}}\right) \leq n^{-r}.$$

Prova. Seja $C' > 0$ a constante dada pela Proposição 7.9 aplicada com $F = F' = \wedge$ e seja A o evento

$$\left\{\exists i, \quad \left|\mathbb{E}[X_\wedge^2(G_i)|G_{i-1}] - 8ns(1-s)\right| > C'r^{1/2}\sqrt{n\log(n)}\right\}. \tag{8-3}$$

Utilizando a Proposição 7.9, aplicada com $b = (r+2)\log(n)$, temos que

$$\mathbb{P}\left(\left|\mathbb{E}[X_\wedge^2(G_i)|G_{i-1}] - 8ns(1-s)\right| > C'r^{1/2}\sqrt{n\log(n)}\right) \leq n^{-r-2}.$$

Por sub-aditividade,

$$\begin{aligned}\mathbb{P}(A) &\leq \sum_{i\leq m}\mathbb{P}\left(\left|\mathbb{E}[X_\wedge^2(G_i)|G_{i-1}] - 8ns(1-s)\right| > C'r^{1/2}\sqrt{n\log(n)}\right)\\ &\leq n^2 n^{-r-2}\\ &= n^{-r}.\end{aligned}$$

Assim,

$$\begin{aligned}\left|\sum_{i=1}^m \mathbb{E}[X_\wedge^2(G_i)|G_{i-1}] - \sum_{i=1}^m 8ns(1-s)\right| &\leq \sum_{i=1}^m \left|\mathbb{E}[X_\wedge^2(G_i)|G_{i-1}] - 8ns(1-s)\right|\\ &\leq \sum_{i=1}^m \kappa_i^2\left|\mathbb{E}[X_\wedge^2(G_i)|G_{i-1}] - 8ns(1-s)\right|\\ &\leq \sum_{i=1}^m \kappa^2 C'r^{1/2}\sqrt{n\log(n)} + \sum_{i=1}^m \kappa^2(2n)^2\mathbb{1}_A\\ &\leq \kappa^2 C'r^{1/2}tn^{5/2}\sqrt{\log(n)} + 4\kappa^2 n^4\mathbb{1}_A.\end{aligned}$$

Além disso, pelo Lema 3.2, temos que

$$s^2(\mathbf{X}) = \mathbb{E}\left[\sum_{i=1}^m \mathbb{E}\left[\kappa_i^2 X_\wedge^2(G_i)|G_{i-1}\right]\right],$$

portanto

$$\begin{aligned}\left|s^2(\mathbf{X}) - \sum_{i=1}^m \kappa_i^2 8ns(1-s)\right| &\leq \mathbb{E}\left|\sum_{i=1}^m \mathbb{E}\left[\kappa_i^2 X_\wedge^2(G_i)|G_{i-1}\right] - \sum_{i=1}^m \kappa_i^2 8ns(1-s)\right|\\ &\leq \kappa^2 C'r^{1/2}tn^{5/2}\sqrt{\log(n)} + 4\kappa^2 n^4\mathbb{P}(A)\\ &\leq 2\kappa^2 C'r^{1/2}tn^{5/2}\sqrt{\log(n)}.\end{aligned}$$

Então, usando o Lema 8.3 para controlar $\sum_{i=1}^n i\kappa_i^2$, obtemos

$$
\begin{aligned}
V^2(\mathbf{X}) &= \frac{\sum_{i=1}^m \mathbb{E}\Big[\kappa_i^2 X_\wedge^2(G_i)|G_{i-1}\Big]}{s^2(\mathbf{X})} \\
&\leq \frac{\sum_{i=1}^m \kappa_i^2 8ns(1-s) + \kappa^2 C' r^{1/2} t n^{5/2}\sqrt{\log(n)} + 4\kappa^2 n^4 \mathbb{1}_A}{\sum_{i=1}^m \kappa_i^2 8ns(1-s) - 2\kappa^2 C' r^{1/2} t n^{5/2}\sqrt{\log(n)}} \\
&\leq \frac{\sum_{i=1}^m \kappa_i^2 8ns(1-s)\Big(1 + f + 4\kappa^2 n^4 \mathbb{1}_A\Big)}{\sum_{i=1}^m \kappa_i^2 8ns(1-s)\Big(1-2f\Big)} \\
&\leq \frac{\Big(1 + f + 4\kappa^2 n^4 \mathbb{1}_A\Big)}{\Big(1-2f\Big)},
\end{aligned}
$$

onde $f = \frac{32D^4C'r^{1/2}\sqrt{\log(n)}}{tn^{1/2}}$.

Portanto,

$$V^2(\mathbf{X}) \leq 1 + \frac{Cr^{1/2}\sqrt{\log(n)}}{tn^{1/2}} + 4\kappa^2 n^4 \mathbb{1}_A,$$

onde C é a constante em função de D e C'. Pelo mesmo argumento, temos que

$$V^2(\mathbf{X}) \geq 1 - \frac{Cr^{1/2}\sqrt{\log(n)}}{tn^{1/2}} - 4\kappa^2 n^4 \mathbb{1}_A.$$

Finalmente, segue dessas duas desigualdades que

$$\left\{\Big|V^2(\mathbf{X}) - 1\Big| > \frac{Cr^{1/2}\sqrt{\log(n)}}{tn^{1/2}}\right\} \subseteq A,$$

logo

$$\mathbb{P}\left(\Big|V^2(\mathbf{X}) - 1\Big| > \frac{Cr^{1/2}\sqrt{\log(n)}}{tn^{1/2}}\right) \leq n^{-r}.$$

■

Com estes resultados, podemos enfim alcançar a conclusão sobre a taxa de convergência do Teorema Central do Limite para o desvio na contagem de subgrafos em grafos livres de triângulos. Iremos demonstrar o Teorema 8.1 sobre a taxa de convergência para a variável aleatória $\Lambda_H(G_m)$, que aproxima bem $D_H(G_m)$.

Prova.(do Teorema 8.1) Como $d(W) = d(\alpha W)$ para todo $\alpha \in \mathbb{R}$ e para qualquer variável aleatória W, e $\Lambda_H(G_m) = n^{v-3}t^{e-2}\binom{H}{\wedge}(1-t)^2\sum_{i=1}^m(1-s)^{-2}X_\wedge(G_i)$, para demonstrarmos o teorema basta mostrarmos que

$$D(\mathbf{Y}) \leq C_{k,\eta}\left[\left(\frac{\sqrt{\log(n)}}{tn^{1/2}}\right)^{\frac{k}{2k+1}}\right], \tag{8-4}$$

onde $Y_i = \frac{X_\wedge(G_i)}{(1-s)^2}$.

Além disso, pelo Corolário 3.8, o fato que $\mathbb{P}[\exists i \leq m, X^*_\wedge(G_i) \neq X_\wedge(G_i)] \leq n^{-8}$ e $||\sum_{i=1}^m X_\wedge(G_i) - \sum_{i=1}^m X^*_\wedge(G_i)||_\infty \leq n^3$, será suficiente mostrarmos que o resultado vale para $D(\mathbf{Y}^*)$, onde

$$Y_i^* = \frac{X^*_\wedge(G_i)}{(1-s)^2}.$$

Pela construção da variável aleatória $X^*_\wedge(G_i)$, sabemos que $|X^*_\wedge(G_i)| \leq \sqrt{tn}\log(n)$ para todo $i = 1, \ldots, m$. Também, como $t \leq 1/2$, então $s \leq 1/2$, portanto $\frac{1}{(1-s)^2} \leq 4$.

Deste modo, $|\frac{X^*_\wedge(G_i)}{(1-s)^2}| \leq 4\sqrt{tn}\log(n)$ para todo $i = 1, \ldots, m$, então $||\mathbf{Y}^*||_\infty \leq 4\sqrt{tn}\log(n)$, onde $\mathbf{Y}^* = (\frac{X^*_\wedge(G_1)}{(1-1/N)^2}, \ldots, \frac{X^*_\wedge(G_m)}{(1-t)^2})$.

Logo, aplicando o Corolário 6.4,

$$D(\mathbf{Y}^*) \leq C'_{k,\eta}\left[\frac{R^3 m\log(m)}{s^3(\mathbf{Y}^*)} + ||V^2(\mathbf{Y}^*) - 1||_k^{\frac{k}{2k+1}} + R^{\frac{2k}{2k+1}} s^{\frac{-2k}{2k+1}}(\mathbf{Y}^*)\right], \tag{8-5}$$

onde $R = 4\sqrt{tn}\log(n)$.

Controlando o primeiro termo da equação acima, temos que

$$\begin{aligned}\frac{R^3 m\log(m)}{s^3(\mathbf{Y}^*)} &\leq \frac{R^3 m\log(m)}{t^3 n^{9/2}} \\ &\leq \frac{64 t^{3/2} n^{3/2}\log^3(n) m\log(m)}{m^{3/2} t^{3/2} n^{3/2}} \\ &\leq \frac{128\log^4(n)}{m^{1/2}}.\end{aligned} \tag{8-6}$$

A equivalência entre $\log(n)$ e $\log(m)$ deriva da condição $t > 1/n$, que implica $\log(m) \leq 2\log(n)$.

Similarmente, para o terceiro termo obtemos

$$\begin{aligned}\left(\frac{R}{s(\mathbf{Y}^*)}\right)^{\frac{2k}{2k+1}} &\leq \left(\frac{4\sqrt{tn}\log(n)}{tn^{3/2}}\right)^{\frac{2k}{2k+1}} \\ &\leq \left(\frac{4\log(n)}{m^{1/2}}\right)^{\frac{2k}{2k+1}}.\end{aligned} \tag{8-7}$$

Por último, nos resta encontrar uma cota superior para o termo do meio.

Pela Proposição 8.4, temos que

$$\left|V^2(\mathbf{Y}^*) - 1\right| \leq \frac{Cr^{1/2}\sqrt{\log(n)}}{tn^{1/2}} + 4\kappa^2 n^4 \mathbb{1}_A,$$

onde $\mathbb{P}(A) \leq n^{-r}$. Portanto,

$$\begin{aligned}||V^2(\mathbf{Y}^*) - 1||_k^{\frac{k}{2k+1}} &= \mathbb{E}\Big[\left|V^2(\mathbf{Y}^*) - 1\right|^k\Big]^{\frac{1}{2k+1}} \\ &\leq \mathbb{E}\Big[\Big(\frac{2Cr^{1/2}\sqrt{\log(n)}}{tn^{1/2}}\Big)^k + (8\kappa^2 n^4)^k \mathbb{1}_A\Big]^{\frac{1}{2k+1}},\end{aligned}$$

pois $(a+b)^k \leq 2^k a^k + 2^k b^k$ para todo $a, b \geq 0$. Assim, para $r > 10k$, temos que $\mathbb{E}[(8\kappa^2 n^4)^k \mathbb{1}_A]^{\frac{1}{2k+1}} \leq (n^{-r}\kappa^{2k} n^{4k})^{\frac{1}{2k+1}} \leq 1/n$. Logo, por concavidade,

$$||V^2(\mathbf{Y}^*) - 1||_k^{\frac{k}{2k+1}} \leq \Big(\frac{C' r^{1/2}\sqrt{\log(n)}}{tn^{1/2}}\Big)^{\frac{k}{2k+1}} + \frac{1}{n}. \tag{8-8}$$

Finalmente, de 8-5, 8-6, 8-7 e 8-8 obtemos que

$$\begin{aligned}D(\mathbf{Y}^*) &\leq C'_{k,\eta}\left[\frac{\log^4(n)}{m^{1/2}} + \Big(\frac{C' r^{1/2}\sqrt{\log(n)}}{tn^{1/2}}\Big)^{\frac{k}{2k+1}} + \frac{1}{n} + \Big(\frac{\log(n)}{m^{1/2}}\Big)^{\frac{2k}{2k+1}}\right] \\ &\leq C_{k,\eta}\left[\Big(\frac{\sqrt{\log(n)}}{tn^{1/2}}\Big)^{\frac{k}{2k+1}}\right],\end{aligned}$$

pois, para n suficientemente grande,

$$\Big(\frac{\sqrt{\log(n)}}{tn^{1/2}}\Big)^{\frac{k}{2k+1}} = \max\left[\frac{\log^4(n)}{m^{1/2}}, \Big(\frac{\sqrt{\log(n)}}{tn^{1/2}}\Big)^{\frac{k}{2k+1}}, \frac{1}{n}, \Big(\frac{\log(n)}{m^{1/2}}\Big)^{\frac{2k}{2k+1}}\right].$$

■

Conforme elucidado nos capítulos anteriores, o Teorema Central do Limite evidencia que um martingal, após uma padronização, irá convergir em distribuição para uma função de distribuição Normal. Além disso, essa convergência ocorre a uma determinada velocidade, como vimos no Capítulo 6.

Deste modo, o exposto no Teorema 1.2 nos mostra que a variável aleatória $D_H(G_m)$, que representa o desvio na contagem de subgrafos H livres de triângulos em G_m, converge em distribuição para Normal a uma certa taxa. Em particular, para qualquer densidade $t \geq n^{-1/2+\delta}$, para algum $\delta > 0$, a taxa de convergência de $D_H(G_m)$ para uma Normal é mais rápida que $1/(tn^{1/2})^{(1-\varepsilon)/2}$ para qualquer $\varepsilon > 0$.

9
Referências bibliográficas

[1] ERDŐS, P.;RÉNYI, A.. **On random graphs i**. Publ. Math., p. 290–297, 1959. 1

[2] ERDŐS, P.;RÉNYI, A.. **On the evolution of random graphs**. Bull. Inst. Internat. Statist, 38(4):343–347, 1960. 1

[3] CHATERJEE, S.. **An introduction to large deviations for random graphs**. Amer. Math. Soc., 53(4):617–642, 2016. 1

[4] HAREL, M.;MOUSSET, F. W.. **Upper tails via high moments and entropic stability**. arXiv:1904.08212 [math.PR], 2019. 1

[5] GOLDSCHMIDT, C.; GRIFFITHS, S. S. A.. **Moderate Deviations of Subgraph Counts in the Erdos-rényi Random graphs $G(n,m)$ and $G(n,p)$**. Trans. Amer. Math. Soc., (373):5517–5585, 2020. 1, 1, 7, 7.1, 7.2, 7.2, 7.3, 7.3, 7.5, 7.5

[6] FÉRAY, V; MÉLIOT, P.-L. A.. **Mod-ϕ convergence: Normality zones and precise deviations**. Springer Briefs in Probability and Mathematical Statistics, 2016. 1

[7] DŐRING, H; EICHELSBACHER, P.. **Moderate deviations via cumulants**. J. Theoret. Probab, 26(2):360–385, 2013. 1

[8] RUCIŃSKI, A.. **When are small subgraphs of a random graph normally distributed?** Probab.Theory Related Fields, 78(1):1–10, 1988. 1

[9] KROKOWSKI, K.; REINCHENBACHS, A.-C.. **Discrete malliavin–stein method: Berry–esseen bounds for random graphs and percolation**. Ann. Probab., 45(2):1071–1109, 2017. 1

[10] JANSON, S.. **Orthogonal decompositions and functional limit theorems for random graph statistics**. Mem. Amer. Math. Soc., 111(534), 1994. 1

[11] MOURRAT, J.-C.. **On the rate of convergence in the martingale central limit theorem**. Bernoulli, 19(2):633–645, 2013. 1, 5.1, 6, 6.2, 6.2

[12] CHUNG, K.. **A Course in Probability Theory, 2nd ed**. Academic Press, 1974. 3

[13] DURRETT, R.. **Probability: Theory and Examples, 4th ed**. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2013. 2.1, 2.1.1, 4.1, 4.4

[14] WILLIAMS, D.. **Probability with Martingales**. Cambridge University Press, 1991. 2.1.1

[15] MAGALHÃES, M.. **Probabilidade e Variáveis Aleatórias, 3 ed**. EdUSP, 2015. 4.1

[16] JAMES, B.. **Probabilidade: Um Curso em Nível Intermediário, 4 ed**. Coleção projeto Euclides. IMPA, 2015. 4.2, 4.2

[17] HALL, P. ; HEYDE, C.. **Martingale Limit Theory and its Application**. Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic Press, 1980. 3, 5.1, 5.1

[18] MCLEISH, D. L.. **Dependent central limit theorems and invariance principles**. Annals of Probability, 2(4):620–628, 1974. 5.1

[19] BOLTHAUSEN, E.. **Exact convergence rate in some martingale central limit theorem**. The Annals of Probability, 10(3):672–688, 1982. 6.2, 6.2, 6.2

[20] BOLLOBÁS, B.. **Random Graphs, 2nd Edition**. Cambridge University Press, 2001. 7.1