



Rodrigo Barreto Alves

Teorema Central do Limite Para Martingais

Dissertação de Mestrado

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

Orientador : Prof. Simon Griffiths
Coorientadora: Profa. Ana Patrícia Carvalho Gonçalves

Rio de Janeiro
abril de 2017



Rodrigo Barreto Alves

Teorema Central do Limite Para Martingais

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

Prof. Simon Griffiths

Orientador

Departamento de Matemática – PUC-Rio

Profa. Ana Patrícia Carvalho Gonçalves

Coorientadora

Departamento de Matemática – IST ULisboa

Prof. Glauco Valle

Departamento de Métodos Estatísticos – UFRJ

Prof. Robert Morris

Instituto de Matemática Pura e Aplicada – IMPA

Prof. Márcio da Silveira Carvalho

Coordenador Setorial do Centro Técnico Científico – PUC-Rio

Rio de Janeiro, 28 de abril de 2017

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Rodrigo Barreto Alves

Graduou-se em Bacharelado de Ciências Atuárias na Universidade Federal do Rio de Janeiro.

Ficha Catalográfica

Alves, Rodrigo Barreto

Teorema Central do Limite Para Martingais / Rodrigo Barreto Alves; orientador: Simon Griffiths; co-orientadora: Ana Patrícia Carvalho Gonçalves. – 2017.

v., 60 f: il. ; 30 cm

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

Inclui bibliografia

1. Martingais;. 2. Teorema Central do Limite;. 3. Teorema Central do Limite para Martingais;. 4. Taxa de Convergência. I. Griffiths, Simon. II. Gonçalves Carvalho, Ana Patrícia. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Matemática. IV. Título.

CDD: 510

Agradecimentos

Agradeço à minha família por todo apoio nesta minha jornada, por todo seu carinho e amor. Agradeço aos professores, funcionários e companheiros de Departamento. Agradeço especialmente aos meus orientadores Simon Griffiths e Patrícia Gonçalves pela imensurável ajuda nesta tese. A minha família tiradentes agradeço por sempre estar presente nos momentos que guardamos ou que alguns guardam, ano após ano, no seu marco inicial, celebramos juntos. Agradeço a CAPES pelo financiamento.

Resumo

Alves, Rodrigo Barreto; Griffiths, Simon; Gonçalves Carvalho, Ana Patrícia; . **Teorema Central do Limite Para Martingais**. Rio de Janeiro, 2017. 60p. Dissertação de Mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

Esta dissertação é dedicada ao estudo das taxas de convergência no Teorema Central do Limite para Martingais. Começamos a primeira parte da tese apresentando a Teoria de Martingais, introduzindo o conceito de esperança condicional e suas propriedades. Desta forma poderemos descrever o que é um Martingal, mostraremos alguns exemplos, e exporemos alguns dos seus principais teoremas. Na segunda parte da tese vamos analisar o Teorema Central do Limite para variáveis aleatórias, apresentando os conceitos de função característica e convergência em distribuição, que serão utilizados nas provas de diferentes versões do Teorema Central do Limite. Demonstraremos três formas do Teorema Central do Limite, para variáveis aleatórias independentes e identicamente distribuídas, a de Lindeberg-Feller e para uma Poisson. Após, apresentaremos o Teorema Central do Limite para Martingais, demonstrando uma forma mais geral e depois enunciaremos uma forma mais específica a qual focaremos o resto da tese. Por fim iremos discutir as taxas de convergência no Teorema Central do Limite, com foco nas taxas de convergência no Teorema Central do Limite para Martingais. Em particular, exporemos o resultado de [4], o qual determina, até uma constante multiplicativa, a dependência ótima da taxa de um certo parâmetro do martingal.

Palavras-chave

Martingais; Teorema Central do Limite; Teorema Central do Limite para Martingais; Taxa de Convergência

Abstract

Alves, Rodrigo Barreto; Griffiths, Simon (Advisor); Gonçalves Carvalho, Ana Patrícia (Co-Advisor). **Martingale Central Limit Theorem**. Rio de Janeiro, 2017. 60p. Dissertação de Mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

This dissertation is devoted to the study of the rates of convergence in the Martingale Central Limit Theorem. We begin the first part presenting the Martingale Theory, introducing the concept of conditional expectation and its properties. In this way we can describe what a martingale is, present examples of martingales, and state some of the principal theorems and results about them. In the second part we will analyze the Central Limit Theorem for random variables, presenting the concepts of characteristic function and the convergence in distribution, which will be used in the proof of various versions of the Central Limit Theorem. We will demonstrate three different forms of the Central Limit Theorem, for independent and identically distributed random variables, Lindeberg-Feller and for a Poisson distribution. After that we can introduce the Martingale Central Limit Theorem, demonstrating a more general form and then stating a more specific form on which we shall focus. Lastly, we will discuss rates of convergence in the Central Limit Theorems, with a focus on the rates of convergence in the Martingale Central Limit Theorem. In particular, we state results of [4], which determine, up to a multiplicative constant, the optimal dependence of the rate on a certain parameter of the martingale.

Keywords

Martingale; Central Limit Theorem; Martingale Central Limit Theorem; Rates of Convergence

Sumário

Sumário	7
1 Introdução	9
2 Teoria de Martingais	11
2.1 Esperança Condicional	11
2.1.1 Propriedades da Esperança Condicional	12
2.2 Martingais	13
2.2.1 Descrição de Martingais como série de incrementos	19
2.2.2 Alguns Resultados sobre Martingais	19
3 Teorema Central do Limite	21
3.1 Função Característica	22
3.2 Convergência em Distribuição	24
3.3 Teorema Central do Limite para Variáveis Aleatórias	26
3.3.1 Necessidade da condição de Lindeberg	34
4 Teorema Central do Limite para Martingais	37
4.1 Formas do Teorema Central do Limite para Martingais	37
5 Taxas de Convergência no Teorema Central do Limite	45
5.1 Taxa de Convergência do Teorema Central do Limite para variáveis aleatórias	45
5.2 Taxas de Convergência no Teorema Central do Limite para Martingais	46
Referências bibliográficas	60

*All art is at once surface and symbol. Those
who go beneath the surface do so at their peril.*

Oscar Wilde.

1

Introdução

O Teorema Central do limite é um resultado de muito estudo quer na matemática quer na estatística. Este resultado é uma ferramenta extremamente importante com aplicações em diversas áreas que perpassa pela biologia, finanças, atuária dentre outras. Com o Teorema Central do Limite é possível que a soma de variáveis aleatórias, dadas determinadas condições, variando para a versão do Teorema Central do Limite que será usado, convirja em distribuição para uma distribuição conhecida, como por exemplo a Gaussiana.

A Teoria de Martingais é um assunto muito importante dentro do estudo de processos estocásticos. Essa teoria pode ser pensada como um modelo de jogo justo em que o conhecimento do passado não interfere na média de futuros ganhos, apenas o que interessa é o presente. A Teoria de Martingais possui diversas aplicações dentro das finanças, atuária e outros.

O objetivo deste trabalho é analisar o Teorema Central do Limite para Martingais, adentrando um caso especial, as suas taxas de convergência, e por fim observando os seus limites. Para isso passaremos pela Teoria de Martingais, suas propriedades e seus principais resultados e discutiremos o Teorema Central do Limite para variáveis aleatórias.

No segundo capítulo abordaremos a Teoria de Martingais, começando pelo conceito de esperança condicional e suas propriedades. Estas são de extrema importância para a Teoria de Martingais, já que a esperança condicional está na definição de Martingais. Após, descreveremos o que é de fato um martingal e mostraremos alguns exemplos. Também serão explicitados as suas propriedades e seus principais resultados.

No terceiro capítulo vamos apresentar duas versões do Teorema Central do Limite para variáveis aleatórias. Começaremos este capítulo definindo o que é uma função característica e mostrando suas propriedades. Passaremos a abordar a convergência em distribuição e juntaremos os dois assuntos demonstrando as suas relações, que serão de extrema importância para as demonstrações do Teorema Central do Limite para variáveis aleatórias e para Martingais, tópico do próximo capítulo. Por fim, iremos demonstrar quatro versões do Teorema Central do Limite para variáveis aleatórias, apontar as suas diferenças e comentar a condição de Lindeberg, uma das versões será o

Teorema Central do Limite de Lindeberg.

No capítulo quatro discutiremos o Teorema Central do Limite para a Teoria de Martingais. Como no Teorema Central do Limite para variáveis aleatórias, o Teorema Central do Limite para Martingais possui diferentes formas. Neste presente trabalho demonstraremos uma forma mais geral do teorema, mas com a intenção de abordarmos um determinado caso específico.

No último capítulo apresentaremos a taxa de convergência para o Teorema Central do Limite com intuito de analisarmos o artigo [4]. Neste é tratada a taxa de convergência para o Teorema Central do Limite para Martingais. Uma das razões de buscarmos, como mencionado anteriormente, uma versão específica do Teorema Central do Limite para Martingais é justamente para abordarmos o artigo que tratará sobre o Teorema Central do Limite para Martingais restrito à classe de Martingais com incrementos limitados.

Veremos que neste artigo o autor consegue otimizar um dos termos do limite da taxa de convergência para o Teorema Central do Limite para Martingais com incrementos limitados e por fim demonstra que este termo não é possível aprimorá-lo.

2

Teoria de Martingais

Um Martingal é uma sequência de variáveis aleatórias, isto é um processo estocástico, no qual, em um determinado tempo em que a sequência foi realizada até este tempo, a esperança do próximo valor da sequência será igual ao valor presente observado, mesmo com todo o conhecimento de todos os valores observados anteriormente.

Podemos pensar em um Martingal $\{S_n, \mathcal{F}_n, n \geq 1\}$ sendo a quantia que temos no tempo n em um jogo justo, ou seja, o valor esperado de ganho ou perda será igual a zero. Esta variável aleatória S_n é na verdade a soma de outras variáveis aleatórias, X_1, X_2 , até X_n , nos respectivos tempos, de ganho ou perda, o martingal será o valor depois de n jogos justos. Isso explica o uso de Martingais, por exemplo, em finanças

A importância de Martingais está também conectada com a decomposição de Doob, que diz que qualquer submartingal, ainda definiremos, pode ser decomposto como uma única soma de um martingal e um processo previsível que começa em zero, o qual é um processo estocástico H_n em que o valor da variável H_n é conhecido, com as informações no tempo $n - 1$, para todo n .

O que nos interessa é o valor esperado da quantia que teremos num tempo n dado o conjunto de informações que possuímos anteriormente, isto significa as informações passadas.

Desta forma, como veremos adiante, o valor esperado futuro de S_n dada as informações que possuímos até o presente momento, como estamos em um jogo justo, será igual à quantia que possuíamos neste mesmo momento.

Existem ainda o supermartingal, que podemos interpretar 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

Suponha um espaço de probabilidade dado por $(\Omega, \mathcal{F}_0, P)$, uma σ -álgebra $\mathcal{F} \subset \mathcal{F}_0$ e uma variável aleatória X \mathcal{F}_0 -mensurável com $E(|X|) < \infty$, sendo

esta esperança com respeito a medida P .

Definimos a esperança condicional, $E(X|\mathcal{F})$, como a única variável aleatória Y qualquer que atenda aos requisitos abaixo:

- (i) $Y \in \mathcal{F}$, ou seja, Y é \mathcal{F} -mensurável;
- (ii) $E(|Y|) < \infty$;
- (iii) $\forall A \in \mathcal{F}$, teremos $\int_A X dP = \int_A Y dP$.

Podemos dizer que esta variável aleatória Y , que preenche as qualidades supracitadas é uma versão da esperança condicional de X dado \mathcal{F} , $E(X|\mathcal{F})$. A prova de existência e unicidade pode ser vista em [6] página 222.

Podemos pensar em um experimento já realizado e as únicas informações referentes ao ponto $\omega \in \Omega$ estão disponíveis através de uma variável aleatória $Z(\omega)$ que é \mathcal{F} -mensurável.

Desta forma a variável aleatória $Y(\omega) = E(X|\mathcal{F})(\omega)$ é o valor esperado de $X(\omega)$ dado este conjunto de informações.

Uma variável aleatória X que é \mathcal{F} -mensurável, terá como valor da sua $E(X|\mathcal{F})$ ela mesma, pois esta é a sua melhor aproximação

Quando \mathcal{F} é igual à σ -álgebra trivial, $\mathcal{F} = \{\Omega, \emptyset\}$, $E(X|\mathcal{F})$ será igual a $E(X)$. Isso ocorre pois, sabendo que estamos no espaço de probabilidade (Ω, \mathcal{F}, P) ,

$$E[X|\mathcal{F}] = \int_{\Omega} X dP = E[X].$$

2.1.1

Propriedades da Esperança Condicional

Veremos agora algumas das propriedades da esperança condicional, estas propriedades praticamente se mantêm iguais às da esperança ordinária.

Quando escrevemos que uma variável aleatória $X \in \mathcal{L}^p$, para $p \geq 1$, significa que X é p -integrável, isto é, $E(|X|^p) < \infty$.

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

(a) *A esperança condicional é linear, $E(aX + Y|\mathcal{F}) = aE(X|\mathcal{F}) + E(Y|\mathcal{F})$, $\forall a \in \mathbb{R}$.*

(b) *Se $X < Y$ então $E(X|\mathcal{F}) < E(Y|\mathcal{F})$.*

(c) Se $X_n \geq 0$ e $X_n \uparrow X$ (sendo que esta notação significa uma convergência monótona pontual, isto é para cada $\omega \in \Omega$) com $E(X) < \infty$ então $E(X_n|\mathcal{F}) \uparrow E(X|\mathcal{F})$.

(d) Temos uma versão do Lemma de Fatou, se $X_n \geq 0 \forall n$ então:

$$E(\liminf X_n|\mathcal{F}) \leq \liminf E(X_n|\mathcal{F})$$

(e) Uma versão do teorema da convergência dominada, se $X_n \rightarrow X$, quase certamente, e existe $Y \in \mathcal{L}^1$, tal que $|X| < Y \forall n$ então, $E(X_n|\mathcal{F}) \rightarrow E(X|\mathcal{F})$, quase certamente.

(f) Se $X \in \mathcal{F}$ e $E[|Y|], E(|XY|) < \infty$ então, $E(XY|\mathcal{F}) = XE(Y|\mathcal{F})$.

(g) Se $\mathcal{F}_1 \subset \mathcal{F}_2$ então $E(E(X|\mathcal{F}_1)|\mathcal{F}_2) = E(E(X|\mathcal{F}_2)|\mathcal{F}_1) = E(X|\mathcal{F}_1)$, chamada de lei da torre.

(h) (Regra da independência) Se \mathcal{G} é independente de $\sigma(\sigma(X), \mathcal{F})$, então $E[X|\sigma(\mathcal{F}, \mathcal{G})] = E[X|\mathcal{G}]$, quase certamente. Em particular, se X é independente de \mathcal{F} então $E[X|\mathcal{F}] = E[X]$, quase certamente.

Demonstração. As demonstrações dos itens (a), (b), (c) e (f) podem ser encontradas em [6] páginas 193-195.

As demonstrações dos itens (d), (e), (g) e (h) podem ser encontradas em [10] página 89-90.

□

2.2 Martingais

Nesta seção o objetivo será definir o que é um martingal, submartingal e supermartingal. Mostrar algumas de suas propriedades e seus principais teoremas, os quais, mais a frente, usaremos para demonstrar o Teorema Central do Limite para martingais.

Antes de abordarmos a teoria de martingais precisamos de duas definições muito importantes, as filtrações e uma sequência adaptada.

Dado um espaço de probabilidade (Ω, \mathcal{F}, P) , definimos uma σ - álgebra gerada por uma coleção \mathcal{C} , $\sigma(\mathcal{C})$, de subconjuntos de Ω , como sendo a menor σ - álgebra de subconjuntos de Ω que contém a coleção \mathcal{C} . Esta menor σ - álgebra que contém a coleção de subconjuntos \mathcal{C} é construída com a interseção

de todas as σ - álgebras que a incluam, ou seja, $\sigma(\mathcal{C}) = \cap X_{\mathcal{C}}$, onde $X_{\mathcal{C}}$ denota a coleção de todas as σ - álgebras de subconjuntos de Ω que incluam \mathcal{C} .

Uma filtração de uma σ - álgebra \mathcal{F} é uma sequência crescente de σ - álgebras, isto é, $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots \subseteq \mathcal{F}$.

Definimos $\mathcal{F}_{\infty} = \sigma(\cup_{n \geq 1} \mathcal{F}_n)$ e $\mathcal{F}_{\infty} \subseteq \mathcal{F}$.

Dizemos que a sequência de variáveis aleatórias X_n é adaptada à filtração \mathcal{F}_n , se X_n é \mathcal{F}_n -mensurável para todo n .

Chamamos de filtração natural a $\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$ para $n \in \mathbb{N}$.

Após definirmos o que é uma filtração e uma sequência adaptada podemos agora tratar da teoria de martingais. Seja \mathcal{F}_n uma filtração e X_n uma sequência de variáveis aleatórias com as características abaixo,

- (i) $X_n \in \mathcal{L}^1$, para todo n ;
- (ii) X_n é adaptado a \mathcal{F}_n ;
- (iii) $E(X_{n+1} | \mathcal{F}_n) = X_n$, para todo n .

Dizemos que o processo X_n é um martingal com respeito a \mathcal{F}_n . Caso substituamos, na última condição, o sinal de igual pelo de menor igual ou pelo de maior igual, teremos um supermartingal ou um submartingal, respectivamente.

Agora daremos alguns exemplos de martingais e suas aplicações. Em cada caso definiremos o processo e verificaremos que este é um martingal, pois atenderá as três condições citadas acima.

Exemplo 2.2. *Pensemos em uma sequência de variáveis aleatórias independentes X_1, X_2, \dots , tal que $X_i = 1$ com probabilidade $1/2$ ou $X_i = -1$ com igual probabilidade, isto para todo $i \in \mathbb{N}$.*

Definimos $S_n = X_1 + X_2 + \dots + X_n$, como a soma destas variáveis aleatórias no tempo n e $\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$ para todo $n \geq 1$, ou seja, \mathcal{F}_n é uma filtração natural. Para $n = 0$ temos, $S_0 = 0$ e $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

Verificaremos agora que S_n é um martingal com respeito a \mathcal{F}_n , para isso S_n deve respeitar as três condições postas acima.

- (i) $S_n \in \mathcal{L}^1$, para todo n ;

Começamos com $E(|S_n|) = E(|X_1 + X_2 + \dots + X_n|) \leq E(|X_1|) + E(|X_2|) + \dots + E(|X_n|)$, como $X_i \in \mathcal{L}^1$ para todo i , ou seja, $E(|X_i|) < \infty$ para todo i .

Então teremos $E(|S_n|) < \infty$, logo $S_n \in \mathcal{L}^1$.

(ii) S_n é adaptado a \mathcal{F}_n ;

X_n é adaptado a \mathcal{F}_n para todo n , isto é, X_n é \mathcal{F}_n -mensurável para todo n .

Façamos $f : \mathbb{R}^n \rightarrow \mathbb{R}$, em que $f(x_1, x_2, \dots, x_n) = x_1 + x_2 + \dots + x_n$, esta função é contínua então é Borel mensurável.

Visto isso temos $S_n = f(X_1, X_2, \dots, X_n)$ é \mathcal{F}_n -mensurável para todo n , ou seja, S_n é adaptado a \mathcal{F}_n .

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

$E(S_{n+1}|\mathcal{F}_n) = E(S_n|\mathcal{F}_n) + E(X_{n+1}|\mathcal{F}_n)$, teremos isto pela propriedade de linearidade da esperança condicional, Proposição 2.1 (a).

Como S_n é \mathcal{F}_n -mensurável, sabemos que $E(S_n|\mathcal{F}_n) = S_n$ e como X_{n+1} é independente de \mathcal{F}_n teremos $E(X_{n+1}|\mathcal{F}_n) = E(X_{n+1})$.

Então $E(S_{n+1}|\mathcal{F}_n) = S_n + E(X_{n+1}) = S_n$, pois $E(X_{n+1}) = 0$.

Verificamos que a sequência de variáveis aleatórias S_n atende as três condições, logo S_n é um martingal.

Exemplo 2.3. Considere agora uma sequência de variáveis aleatórias, independentes e não negativas X_n , $E(X_n) = 1$ para todo n . Defina $M_n = X_1 X_2 \dots X_n$ e \mathcal{F}_n a filtração natural, para todo $n \geq 1$. Para $n = 0$, tem-se $M_0 = 1$ e $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Para podermos chamar M_n de martingal com respeito a \mathcal{F}_n , faremos como no exemplo 1, verificando se o processo M_n se enquadra nas três condições.

(i) $M_n \in \mathcal{L}^1$, para todo n ;

Sabemos que M_n é uma multiplicação de variáveis aleatórias não negativas, desta forma M_n é uma variável não negativa, então $E[|M_n|] = E[X_1 X_2 \dots X_n] = E[X_1 X_2 \dots X_n]$. Como as variáveis aleatórias X_1, X_2, \dots, X_n são independentes temos $E[X_1 X_2 \dots X_n] = E[X_1]E[X_2] \dots E[X_n] = 1$.

Logo $E[|M_n|] < \infty$, isto é, $M_n \in \mathcal{L}^1$, para todo n .

(ii) M_n é adaptado a \mathcal{F}_n ;

X_n é adaptado a \mathcal{F}_n para todo n , isto é, X_n é \mathcal{F}_n -mensurável para todo n . Façamos $f : \mathbb{R}^n \rightarrow \mathbb{R}$, em que $f(x_1, x_2, \dots, x_n) = x_1 x_2 \dots x_n$, esta função é

contínua então é Borel mensurável. Visto isso temos $M_n = f(X_1, X_2, \dots, X_n)$ é \mathcal{F}_n -mensurável para todo n , ou seja, M_n é adaptado a \mathcal{F}_n .

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

$E(M_{n+1}|\mathcal{F}_n) = E(X_{n+1}M_n|\mathcal{F}_n) = M_nE(X_{n+1}|\mathcal{F}_n)$, teremos isso pela Propriedade 2.1 (f). $E(X_{n+1}|\mathcal{F}_n) = E(X_{n+1}) = 1$, temos isto porque X_{n+1} é independente de \mathcal{F}_n . Logo teremos $E(M_{n+1}|\mathcal{F}_n) = M_n$.

Exemplo 2.4 (Urna de Polya). Considere uma urna que possui uma bola branca e uma preta, a cada instante escolhemos uma delas uniformemente, observamos sua cor, a devolvemos para a urna e então adicionamos uma bola da mesma cor daquela observada. Repetimos o processo. Depois de n instantes teremos $n + 2$ bolas. Vamos, neste exemplo, analisar a proporção de bolas pretas em cada instante e esta proporção chamaremos de X_n . O nosso espaço de probabilidade aqui é a urna e a filtração é a natural, ou seja, $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$.

(i) $X_n \in \mathcal{L}^1$, para todo n ;

Trivial, pois X_n é uma proporção.

(ii) X_n é adaptado a \mathcal{F}_n ;

Trivial também, pois a filtração é natural.

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

Vamos denominar P_n o número de bolas pretas na urna no instante n e B_n o número de bolas brancas no instante n , naturalmente $B_n = (n + 2) - P_n$.

O que nos interessa aqui é a proporção $X_n = \frac{P_n}{n+2}$.

Calcularemos o valor esperado do número de bolas pretas dentro da urna no instante $n + 1$, dado as proporções de bolas pretas na urna nos tempos anteriores. Temos que o valor de P_{n+1} poderá ser apenas dois,

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

Visto isso teremos,

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

Desta forma teremos,

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

Exemplo 2.5 (Processo de Galton - Watson). *Um processo de Galton-Watson é um processo estocástico de ramificação em que Z_n é o tamanho da geração no tempo n e $X_{n,i}$ o número de crianças do indivíduo i na geração no tempo n . De uma maneira mais formal 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:*

$$Z_{n+1} = \begin{cases} X_{1,n+1} + \dots + X_{Z_n,n+1} & \text{se } Z_n > 0 \\ 0 & \text{se } Z_n = 0 \end{cases}$$

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

(i) $Y_n \in \mathcal{L}^1$, para todo n ;

Temos que a variável aleatória Y_n é não negativa, pois $\mu \geq 0$, então temos que $E(|Y_n|) = E(Y_n)$.

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

Sabemos que $Z_n = 0$ ou $Z_n = X_{1,n} + \dots + X_{Z_{n-1},n}$, para o primeiro caso teremos $E(Y_n) = 0$.

Já para o segundo caso teremos,

$$E(Y_n) = \frac{1}{\mu^n} E(Z_n) = \frac{1}{\mu^n} E(X_{1,n} + \dots + X_{Z_{n-1},n}) = \frac{1}{\mu^n} [E(X_{1,n}) + \dots + E(X_{Z_{n-1},n})].$$

Como $X_{n,i} \in \mathcal{L}^1$, para todo $i, n \geq 1$, então $E(Y_n) < \infty$. Logo $Y_n \in \mathcal{L}^1$.

(ii) Y_n é adaptado a \mathcal{F}_n ;

Sabemos que $Z_n \in \mathcal{F}_n$, para todo n , pelos mesmos motivos do Exemplo 1.

Como $Y_n = Z_n/\mu^n$, 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) $E(Y_{n+1}|\mathcal{F}_n) = Y_n$, para todo n .

Temos que,

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

pelas Propriedades 2.1 (a) e 2.1 (c). Em $Z_n = k$, fazemos $Z_{n+1} = X_{1,n+1} + \dots + X_{k,n+1}$, visto isso teremos,

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

Pela Propriedade 2.1 (f) obteremos,

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

Como $X_{i,k+1}$ é independente de \mathcal{F}_n , para todo i teremos por fim,

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

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

$$E\left[\frac{Z_{n+1}}{\mu^{n+1}}|\mathcal{F}_n\right] = E[Y_{n+1}|\mathcal{F}_n] = \frac{Z_n}{\mu^n} = Y_n.$$

Logo $E[Y_{n+1}|\mathcal{F}_n] = Y_n$. Concluimos que Y_n é um martingal.

2.2.1

Descrição de Martingais como série de incrementos

No Exemplo 1, vimos que S_n é um martingal. Temos também $X_n = S_n - S_{n-1}$, para todo $n \geq 1$. Desta maneira podemos dar uma outra forma a este martingal, escrevendo-o a partir dos seus incrementos, pois estes possuem média zero, dada uma filtração \mathcal{F}_n :

$$E(X_{n+1}|\mathcal{F}_n) = E(S_{n+1} - S_n|\mathcal{F}_n) = E(S_{n+1}|\mathcal{F}_n) - E(S_n|\mathcal{F}_n) = S_n - S_n = 0$$

Podemos generalizar esta ideia acima para qualquer martingal Y_n , representando-o através de uma sequência de variáveis aleatórias X_1, X_2, \dots , tal que $E(X_{n+1}|\mathcal{F}_n) = 0$, 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 isto definiremos a variância condicional para Martingais e a variância quadrática normalizada, que serão de extrema importância mais a frente.

Podemos definir para qualquer Martingal Y_n e para as variáveis aleatórias X_1, X_2, \dots , tal que $E(X_{n+1}|\mathcal{F}_n) = 0$, para todo n , onde \mathcal{F}_n é uma filtração e $X_n = Y_n - Y_{n-1}$, denotando a diferença de Martingais.

2.2.2

Alguns Resultados sobre Martingais

Daremos agora a definição do tempo de parada, algumas propriedades e o teorema de parada opcional.

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

O tempo de parada pode ser considerado um mecanismo no processo estocástico que estamos analisando para decidirmos se este continua ou pára dado o presente e os eventos passados. Na maioria das vezes o tempo de parada nos levará a decisão de pararmos em um tempo finito.

Teorema 2.7 (Teorema da Convergência para Martingais). *Se X_n é um submartingal com $\sup E(X_n^+) < \infty$, temos que $X_n \rightarrow X$, quase certamente, quando $n \rightarrow \infty$, sendo X uma variável aleatória com $E(|X|) < \infty$.*

Demonstração. A prova deste teorema pode ser encontrada em [6] página 202. □

Definição 2.8. *Uma sequência de variáveis aleatórias X_1, X_2, \dots e uma filtração \mathcal{F}_n , a sequência é dita previsível, se para todo n , X_n é \mathcal{F}_{n-1} - mensurável.*

Esta definição acima é importante para a decomposição de Doob.

Teorema 2.9. *(Decomposição de Doob) Qualquer submartingal X_n , $n \geq 0$, pode ser decomposto de uma única forma, $X_n = M_n + A_n$, onde M_n é um martingal e A_n é uma sequência previsível com $A_0 = 0$.*

Demonstração. A prova deste teorema pode ser encontrada em [6] página 203. \square

3

Teorema Central do Limite

Neste capítulo vamos apresentar o Teorema Central do Limite para variáveis aleatórias. Apresentaremos e provaremos quatro de suas versões, a para variáveis aleatórias independentes e identicamente distribuídas, a versão de Lindeberg, a de Lindeberg-Feller e o Teorema do Limite Central para a distribuição de Poisson.

Para que possamos discutir mais profundamente o Teorema Central do Limite para variáveis aleatórias e para Martingais, ponto do próximo capítulo, e suas demonstrações, abordaremos antes sobre os seguintes assuntos: funções características, suas propriedades e a convergência em distribuição.

Um importante resultado que obteremos com as funções características e a convergência em distribuição é que uma sequência de variáveis aleatórias convergirá em distribuição para uma determinada distribuição, se e somente se, a sequência de funções características convergir pontualmente para a função característica desta mesma determinada distribuição. Este resultado é importante para a prova do Teorema Central do Limite.

O Teorema Central do Limite basicamente, explicando de uma forma informal, anuncia que a soma de variáveis aleatórias independentes, podendo ou não serem identicamente distribuídas, dependendo do caso, estas com médias e variâncias finitas, após uma padronização, tenderá a uma distribuição normal padrão. Quando dizemos tenderá, significa que haverá uma convergência em distribuição.

Veremos primeiro o Teorema Central do Limite para variáveis aleatórias independentes e identicamente distribuídas, após exploraremos a versão de Lindeberg. Uma diferença entre estas versões é que na primeira, para ocorrer a convergência da soma das variáveis aleatórias, após sofrer uma padronização, para a função de distribuição da normal padrão é necessário, como já está no enunciado do teorema, que as variáveis sejam identicamente distribuídas. Já na versão de Lindeberg isto não é necessário, como veremos mais adiante bastará que a condição de Lindeberg seja satisfeita para que ocorra a convergência em distribuição da soma das variáveis aleatórias, após ser padronizada, para a função de distribuição da normal padrão.

Por fim demonstraremos a versão de Lindeberg-Feller e a convergência à

Poisson. Na primeira versão veremos que se as variáveis aleatórias independentes e com média zero atenderem as duas condições, o seu somatório convergirá em distribuição para uma distribuição normal. Já na convergência à Poisson, veremos que se a sequência de variáveis aleatórias independentes atenderem às hipóteses, o seu somatório convergirá em distribuição para uma distribuição de Poisson.

3.1

Função Característica

Começaremos esta seção definindo as funções características que serão de extrema importância para a convergência em distribuição e através desta ferramenta conseguiremos uma prova para o Teorema do Limite Central.

A função característica, como será visto nesta seção, definirá para qualquer variável aleatória a sua distribuição. Através dela podemos fazer diferentes análises sobre as funções de densidade e as funções de distribuição de variáveis aleatórias. Basicamente para qualquer variável aleatória teremos uma única função característica que a definirá, por isso sua grande importância.

Apresentaremos agora a definição formal das funções características.

Definição 3.1. *Seja X uma variável aleatória, definimos sua função característica, como uma função $\varphi : \mathbb{R} \rightarrow \mathbb{C}$:*

$$\varphi_X(t) = E(\exp(itX)) = E(\cos(tX)) + iE(\sin(tX)), t \in \mathbb{R} \quad (3.1)$$

Como vimos na própria definição de função característica, a variável aleatória complexa $\exp(itX)$ possui esperança finita para qualquer variável aleatória X , pois da fórmula de Euler temos que $\exp(iy) = \cos(y) + i\sin(y)$, $y \in \mathbb{R}$, e as funções $\sin(y)$ e $\cos(y)$ são limitadas. Logo garantimos que a função característica está bem definida.

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

Teorema 3.2. *Para todas as funções características temos, sendo X e Y variáveis aleatórias:*

(a) Assume valor 1 no ponto $t = 0$, isto é $\varphi_X(0) = 1$;

(b) $\overline{\varphi_X(t)} = \varphi_X(-t)$, onde $\overline{\varphi_X(t)}$ é o complexo conjugado de $\varphi_X(t)$;

(c) Se X e Y são independentes então temos que,

$$\varphi_{X+Y}(t) = \varphi_X(t) \cdot \varphi_Y(t), \forall t \in \mathbb{R};$$

(d) Se $Y = aX + b$, então $\varphi_Y(t) = \exp(itb)\varphi_X(at)$;

(e) Temos que φ_X também gera os momentos da seguinte maneira:

$$\left. \frac{\partial^n}{\partial t^n} \varphi_X(t) \right|_{t=0} = i^n E(X^n), n = 1, 2, \dots, \quad \text{se } E(|X|^n) < \infty;$$

(f) $|\varphi_X(t)| \leq 1, \forall t \in \mathbb{R}$.

Demonstração. Para as propriedades (a), (b), (c), e (d) veja [8] páginas 225 e 229.

Para a propriedade (e) [9] veja página 280. \square

A propriedade (c) pode ser generalizada para uma sequência de variáveis aleatórias X_1, X_2, \dots, X_n independentes fazendo indução. Desta forma teremos:

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

Exemplo 3.3. Dada uma variável aleatória X com distribuição normal padrão temos que sua função característica é dada por:

$$\varphi_X(t) = \exp\left(\frac{-t^2}{2}\right), \quad \forall t \in \mathbb{R} \quad (3.2)$$

O cálculo deste exemplo está em [6] página 92.

Como já vimos, por definição, a função de distribuição de uma variável aleatória determina a função característica desta variável aleatória.

Apresentaremos agora a fórmula da inversão, que nos dará a recíproca, isto é, que a função característica de uma variável aleatória determina a função de distribuição desta.

Teorema 3.4. 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:

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

sendo $\tilde{F}(w) = \frac{1}{2}\{F(w) + F(w^-)\}$, para todo $w \in \mathbb{R}$, e os números reais a, b , e c tais que $c > 0$ e $a < b$.

Demonstração. A demonstração deste teorema pode ser vista em [9] página 282. \square

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 3.5. *[Teorema Da Unicidade] Se as variáveis aleatórias X e Y têm a mesma função característica, então elas possuem a mesma função de distribuição.*

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

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

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

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

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

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

A função de distribuição é contínua à direita, tomamos o limite para $b \downarrow c$. Visto isso obteremos,

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

e 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. \square

3.2

Convergência em Distribuição

Existem alguns modos de convergência de sequências de variáveis aleatórias, vamos apresentar aqui a convergência em distribuição. Apesar de ser o modo de convergência mais fraco é um dos mais utilizados e aplicados no Teorema Central do Limite.

Definição 3.6. *Sejam X, X_1, X_2, \dots variáveis aleatórias com funções de distribuição F, F_1, F_2, \dots , respectivamente. Dizemos que X_n converge em distribuição para X , quando $n \rightarrow \infty$, se para todo ponto de continuidade x de F , temos $F_n(x) \rightarrow F(x)$, quando $n \rightarrow \infty$.*

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

Como vimos no Teorema 3.5, há uma relação biunívoca entre a função de distribuição e a função característica. Desta forma apresentaremos dois teoremas que relacionarão o modo de convergência em distribuição com as funções características. De modo que conseguiremos provar que:

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

isto é, F_{X_n} convergirá fracamente para F ($F_{X_n} \xrightarrow{f} F_X$), isto é X_n convergirá em distribuição para X , se e somente se, $\varphi_{X_n}(t) \rightarrow \varphi_X(t), \forall t \in \mathbb{R}$.

A convergência em distribuição de F_1, F_2, \dots para F , isto é $F_{X_n} \xrightarrow{D} F_X$, significa também que estas funções acumuladas convergem fracamente para F .

Teorema 3.7. (Teorema de Helly-Bray) *Sejam F, F_1, F_2, \dots funções de distribuição. Se F_n converge fracamente para F , então*

$$\int f(x) dF_n(x) \longrightarrow \int f(x) dF(x),$$

isto sempre que $n \rightarrow \infty$ e para toda função f contínua e limitada ($f : \mathbb{R} \rightarrow \mathbb{R}$).

Demonstração. A demonstração deste teorema está em [8] página 235. \square

Teorema 3.8. (Teorema da Continuidade de Levy) *Sejam F_1, F_2, \dots funções de distribuição cujas funções características são dadas por $\varphi_1, \varphi_2, \dots$ respectivamente. Se φ_n converge pontualmente para um limite φ e se este é contínuo no ponto zero, então*

(i) *existe F , função de distribuição, tal que $F_n \rightarrow F$ fracamente e,*

(ii) *φ é a função característica de F .*

Demonstração. A demonstração deste teorema está em [8] página 237. \square

Com estes dois teoremas conseguimos estabelecer uma relação de equivalência entre a convergência em distribuição e a convergência das respectivas funções características.

Sejam X_1, X_2, \dots , variáveis aleatórias, se $\varphi_{X_n} \rightarrow \varphi$, para todo $t \in \mathbb{R}$, e a função φ é contínua no ponto zero, logo ela é uma função característica de uma variável aleatória X , faremos então $\varphi = \varphi_X$. Desta forma teremos que $X_n \xrightarrow{D} X$.

3.3

Teorema Central do Limite para Variáveis Aleatórias

O Teorema Central do Limite para variáveis aleatórias é extremamente importante para a probabilidade e a estatística. Através do Teorema Central do Limite podemos provar diversos outros teoremas e geramos outras ferramentas.

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, basicamente:

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

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

Para provarmos o Teorema Central do Limite para variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) precisaremos do seguinte resultado que será anunciado em forma de Lema abaixo.

Lema 3.9. *Sejam c_1, c_2, \dots , e c números complexos tais que $\lim_{n \rightarrow \infty} c_n = c$ então $\lim_{n \rightarrow \infty} (1 + \frac{c_n}{n})^n = e^c$.*

Demonstração. O que queremos provar podemos escrever da seguinte forma

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

Se $a_n \rightarrow 0$, quando $n \rightarrow \infty$, então $\frac{1+a_n}{e^{a_n}} \rightarrow 1$, quando $n \rightarrow \infty$.

De fato, por L' Hôpital,

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

Como $c_n/n \rightarrow 0$, quando $n \rightarrow \infty$ e $a_n \rightarrow 0$ quando $n \rightarrow \infty$, tomaremos $a_n = c_n/n$. Visto isso teremos,

$$\lim_{n \rightarrow \infty} \left(1 + \frac{c_n}{n}\right)^n = \lim_{n \rightarrow \infty} (1 + a_n)^n = \lim_{n \rightarrow \infty} (e^{a_n})^n = \lim_{n \rightarrow \infty} (e^{\frac{c_n}{n}})^n = \lim_{n \rightarrow \infty} e^{c_n} = e^c.$$

□

Teorema 3.10. *(Teorema do Central do Limite para variáveis aleatórias i.i.d.) Sejam X_1, X_2, \dots , variáveis aleatórias independentes e identicamente distribuídas com média μ e variância σ^2 , sendo $\sigma < \infty$. Então para $S_n = X_1 + X_2 + \dots + X_n$, teremos,*

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

onde Z é uma variável aleatória com distribuição normal padrão.

Demonstração. Podemos supor, sem perda de generalidade que $\mu = 0$, ou poderíamos definir a variável aleatória $Y_n = X_n - \mu$ que possui média zero.

Como as variáveis aleatórias X_1, X_2, \dots , são identicamente distribuídas, para facilitarmos a notação, faremos $\varphi_{X_i} = \varphi$ para todo i . Aplicando do Teorema 3.2 as propriedades (d) e (c), respectivamente, teremos:

$$\varphi_{\frac{S_n}{\sigma\sqrt{n}}}(t) = \varphi_{S_n}\left(\frac{t}{\sigma\sqrt{n}}\right) = \prod_{i=1}^n \varphi_i\left(\frac{t}{\sigma\sqrt{n}}\right) = \left(\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n.$$

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 $\lim_{x \rightarrow 0} \frac{o(x)}{x} = 0$.

Utilizando o Teorema 3.2 as propriedades (a) e (f), teremos $\varphi(0) = 1$, $\varphi'(0) = i\mu = 0$ e $\varphi'' = i^2\mu = -\sigma^2$ e, desta maneira,

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

Temos que

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

Utilizando o Lemma 3.9 obteremos o seguinte resultado:

$$\lim_{n \rightarrow \infty} \varphi\left(\frac{t}{\sigma\sqrt{n}}\right) = \lim_{n \rightarrow \infty} \left\{1 - \frac{t^2}{2n} \left[1 - \frac{o\left(\frac{t^2}{n}\right)}{\frac{t^2}{2n}}\right]\right\}^n = e^{\frac{-t^2}{2}}$$

Desta maneira pelo Teorema 3.8, podemos concluir que:

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

□

Concluimos assim a prova do Teorema Central do Limite para variáveis aleatórias i.i.d., abaixo iremos enunciar o Teorema Central do Limite de Lindeberg, o qual dará condições mais gerais para a convergência à normal.

Com o 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_n , respectivamente, tais que $E(X_n) = \mu_n$ e $Var(X_n) = \sigma_n^2$, onde $\sigma_n^2 < \infty$, e pelo menos um $\sigma_n^2 > 0$. A condição de Lindeberg diz que:

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

Esta condição de Lindeberg implica por [8] páginas 267 e 268, que:

$$\max_{1 \leq i \leq n} \frac{\sigma_i^2}{s_n^2} \rightarrow 0 \quad \text{quando} \quad n \rightarrow \infty.$$

Isto quer dizer que a contribuição de uma variável aleatória X_i , em que $1 \leq i \leq n$, para a variância s_n^2 é arbitrariamente pequena para um n suficientemente grande.

Antes de demonstrar o teorema apresentaremos um Lema que será utilizado na sua prova.

Lema 3.11. *Sejam $c_{n,k}$ números complexos tais que $\sum_{k=1}^n c_{n,k} \rightarrow c$ quando $n \rightarrow \infty$. Se $\max_{1 \leq k \leq n} |c_{n,k}| \rightarrow 0$ quando $n \rightarrow \infty$ e $\sum_{k=1}^n |c_{n,k}| \leq M < \infty$, onde M é uma constante que não depende de n , então,*

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

Demonstração. A demonstração deste lema está em [5] página 199. \square

Teorema 3.12. *(Teorema Central do Limite de Lindeberg) Sejam X_1, X_2, \dots , variáveis aleatórias independentes, com funções de distribuição F_n , respectivamente, tais que $E(X_n) = \mu_n$ e $Var(X_n) = \sigma_n^2$, onde $\sigma_n^2 < \infty$, e pelo menos um $\sigma_n^2 > 0$. Sejam $S_n = X_1 + \dots + X_n$ e $s_n = \sqrt{Var(S_n)} = \sqrt{\sigma_1^2 + \dots + \sigma_n^2}$. Então para que tenhamos:*

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

quando $n \rightarrow \infty$, dado que Z é uma variável aleatória com distribuição normal padrão, é suficiente que a condição de Lindeberg,

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

esteja satisfeita.

Demonstração. 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:

$$e^{itx} = 1 + itx + \alpha_1(x) \frac{t^2 x^2}{2}, \quad \text{onde } |\alpha_1(x)| \leq 1, \quad (3.3)$$

$$e^{itx} = 1 + itx - \frac{t^2 x^2}{2} + \alpha_2(x) \frac{t^3 x^3}{6}, \quad \text{onde } |\alpha_2(x)| \leq 1 \quad (3.4)$$

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

$$e^{itx} = 1 + itx - \frac{t^2 x^2}{2} + r_\epsilon(x), \text{ onde:}$$

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

Teremos desta forma:

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

Pela fórmula de Taylor:

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

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

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

onde $a_{n,k}$ satisfaz,

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

Desta forma teremos,

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

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

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

Faremos uma sequência de ϵ 's que convergirão para zero. Escolhemos $\epsilon = \frac{1}{m}$ para $m \in \mathbb{N}$, então existe n_m tal que $n \geq n_m$,

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

Desta forma há uma sequência de inteiros positivos $n_1 < n_2 < n_3 < \dots$ tal que,

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

para $n_m \leq n < n_{m+1}$, onde para estes valores de n , os restos $a_{n,k}$ são baseados em $\epsilon = 1/m$. Por economia de notação, a dependência em m não será mais expressa, mas basta lembrar que os valores dos restos $a_{n,k}$, que são determinados por ϵ , dependem da posição de n em relação a n_m .

Teremos então que,

$$\sum_{k=1}^n |a_{n,k}| \xrightarrow{n \rightarrow \infty} 0. \quad (3.5)$$

Sabemos que:

$$\varphi_{\frac{S_n - E(S_n)}{s_n}}(t) = \prod_{k=1}^n E \left(\exp \left\{ it \left(\frac{X_k - \mu_k}{s_n} \right) \right\} \right) = \prod_{k=1}^n \left(1 - \frac{t^2 \sigma_k^2}{2s_n^2} + a_{n,k} \right) \quad (3.6)$$

Agora provaremos que o último termo da equação acima converge para a função característica de uma variável aleatória que possui distribuição normal padrão. Para isso utilizaremos o Lemma 3.11.

Utilizando a mesma notação do Lemma 3.11 faremos,

$$c_{n,k} = - \left\{ \frac{t^2 \sigma_k^2}{2s_n^2} \right\} + a_{n,k} \quad \text{e} \quad c = \frac{t^2}{2}.$$

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 3.3.

Aplicaremos o Lema 3.11, em nosso caso faremos $c_{n,k} = -[t^2 \sigma_k^2 / (2s_n^2)] + a_{n,k}$ e $c = -t^2/2$. Por (3.5), temos:

$$\sum_{k=1}^n |c_{n,k}| \leq \frac{t^2}{2} + \sum_{k=1}^n |a_{n,k}| \xrightarrow{n \rightarrow \infty} \frac{t^2}{2}, \quad (3.7)$$

logo $\sum_{k=1}^n |c_{n,k}|$ é uniformemente limitado, para aplicarmos o Lemma 3.11 resta verificarmos a condição do máximo:

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

como já visto em (3.5) o segundo termo tende para zero. A condição de Lindeberg, como já visto, implica em:

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

Desta forma a prova está terminada. □

Podemos dizer que se a condição de Lindeberg for satisfeita, vale a convergência em distribuição para uma função de distribuição normal padrão.

Demonstraremos agora o Teorema de Lindeberg-Feller, em que usaremos a condição de Lindeberg. Neste teorema teremos uma sequência de variáveis aleatórias organizadas na forma de vetor triangular.

Pensemos para cada $n \in \mathbb{N}$, façamos $X_{n,m}$, para $1 \leq m \leq n$ variáveis aleatórias independentes, desta forma faremos com que na linha de cada sequência as variáveis aleatórias continuem independentes entre si. Podemos visualizar o vetor triangular da seguinte forma,

$$\begin{aligned} &X_{1,1} \\ &X_{2,1} \ X_{2,2} \quad \rightarrow \quad \text{independentes entre si} \\ &X_{3,1} \ X_{3,2} \ X_{3,3} \rightarrow \quad \text{independentes entre si} \end{aligned}$$

e assim continuará as sequências até $m = n$.

Para a demonstração do Teorema de Lindeberg-Feller, precisaremos de dois Lemas que apresentaremos abaixo.

Lema 3.13. *Seja X uma variável aleatória e $n \in \mathbb{N}$ temos que,*

$$\left| E(e^{itX}) - \sum_{m=0}^n \frac{E[(itX)^m]}{m!} \right| \leq E \left[\min(|tX|^{n+1}, 2|tX|^n) \right].$$

Demonstração. A demonstração pode ser vista em [6] página 116. \square

Lema 3.14. *Sejam z_1, \dots, z_n e w_1, \dots, w_n números complexos dentro do círculo unitário e para $n \in \mathbb{N}$, teremos,*

$$\left| \prod_{m=1}^n z_m - \prod_{m=1}^n w_m \right| \leq \sum_{m=1}^n |z_m - w_m|.$$

Demonstração. A demonstração está em [6] página 125. \square

O Teorema de Lindeberg-Feller nos diz, como veremos abaixo, que uma grande quantidade de observações independentes que possuem média zero, quando somadas irão convergir em distribuição para uma distribuição normal.

Antes de demonstrarmos o Teorema de Lindeberg-Feller, iremos introduzir uma notação. Quando temos $E[|X_{n,m}|^2; |X_{n,m}| > \epsilon]$, isto significa que o limite de integração para o cálculo do valor esperado de $|X_{n,m}|^2$, será de $(\epsilon, +\infty)$ e $(-\infty, -\epsilon)$.

Teorema 3.15. *(Teorema de Lindeberg-Feller) Para cada n , seja $X_{n,m}$, para $1 \leq m \leq n$, variáveis aleatórias independentes e com $E(X_{n,m}) = 0$. Suponha,*

$$(i) \sum_{m=1}^n E(X_{n,m}^2) \rightarrow \sigma^2 > 0 \quad n \rightarrow \infty;$$

$$(ii) \forall \epsilon > 0, \lim_{n \rightarrow \infty} \sum_{m=1}^n E[|X_{n,m}|^2; |X_{n,m}| > \epsilon] = 0.$$

Então $S_n = X_{n,1} + \dots + X_{n,n} \xrightarrow{D} \sigma Z$, quando $n \rightarrow \infty$, onde a variável aleatória Z tem distribuição normal padrão.

Demonstração. Seja $\varphi_{n,m}(t) = E(\exp(itX_{n,m}))$ e $\sigma_{n,m}^2 = E(X_{n,m}^2)$. Para provarmos o teorema basta mostrar que,

$$\prod_{m=1}^n \varphi_{n,m}(t) \rightarrow \exp(-t^2\sigma^2/2).$$

Façamos $z_{n,m} = \varphi_{n,m}(t)$ e $w_{n,m} = (1 - t^2\sigma_{n,m}^2/2)$. Pelo Lema 3.13 temos,

$$\left| E(\exp(itX_{n,m})) - \sum_{a=0}^b \frac{E[(itX_{n,m})^a]}{a!} \right| \leq E \left[\min(|tX_{n,m}|^{b+1}, 2|tX_{n,m}|^b) \right].$$

Pelo Lema 3.13, $b \in \mathbb{N}$, então faremos $b = 2$, visto isso teremos,

$$\left| E(\exp(itX_{n,m}) - \sum_{a=0}^2 \frac{E[(itX_{n,m})^a]}{a!}) \right| \leq E \left[\min(|tX_{n,m}|^3, 2|tX_{n,m}|^2) \right].$$

Sabemos que $E(\exp(itX_{n,m})) = \varphi_{n,m}(t)$, vamos agora analisar o segundo termo do lado esquerdo da equação anterior,

$$\begin{aligned} \sum_{a=0}^2 \frac{E[(itX_{n,m})^a]}{a!} &= 1 + itE(X_{n,m}) - \frac{t^2 E[X_{n,m}^2]}{2!} \\ &= 1 - \frac{t^2 \sigma_{n,m}^2}{2} = w_{n,m} \end{aligned}$$

Desta forma teremos,

$$|z_{n,m} - w_{n,m}| \leq E \left[\min(|tX_{n,m}|^3, 2|tX_{n,m}|^2) \right].$$

Observando o lado direito da desigualdade anterior faremos,

$$E \left[\min(|tX_{n,m}|^3, 2|tX_{n,m}|^2) \right] \leq E \left[|tX_{n,m}|^3; |X_{n,m}| \leq \epsilon \right] + E \left[2|tX_{n,m}|^2; |X_{n,m}| > \epsilon \right].$$

Podemos explicar esta desigualdade supondo primeiro que $\min(|tX_{n,m}|^3, 2|tX_{n,m}|^2) = |tX_{n,m}|^3$. Visto isso teremos então,

$$\begin{aligned} |z_{n,m} - w_{n,m}| &\leq E \left[\min(|tX_{n,m}|^3, 2|tX_{n,m}|^2) \right] \\ &\leq E \left[|tX_{n,m}|^3; |X_{n,m}| \leq \epsilon \right] + E \left[2|tX_{n,m}|^2; |X_{n,m}| > \epsilon \right] \quad (3.8) \\ &\leq \epsilon t^3 E \left[|X_{n,m}|^2; |X_{n,m}| \leq \epsilon \right] + 2t^2 E \left[|X_{n,m}|^2; |X_{n,m}| > \epsilon \right]. \end{aligned}$$

Antes de continuar a demonstração vamos analisar as suposições (i) e (ii). Na suposição (i) observemos, sabendo que $|X_{n,m}|^2 = X_{n,m}^2$,

$$\sum_{m=1}^n E(X_{n,m}^2) = \sum_{m=1}^n \left[E \left[|X_{n,m}|^2; |X_{n,m}| > \epsilon \right] \right] + \sum_{m=1}^n \left[E \left[|X_{n,m}|^2; |X_{n,m}| \leq \epsilon \right] \right]. \quad (3.9)$$

Fazendo $n \rightarrow \infty$, pela suposição (ii), o primeiro termo do lado direito da equação (3.9) é igual a 0. Então teremos,

$$\lim_{n \rightarrow \infty} \sum_{m=1}^n E(X_{n,m}^2) = \lim_{n \rightarrow \infty} \sum_{m=1}^n E \left[|X_{n,m}|^2; |X_{n,m}| \leq \epsilon \right].$$

Então pela suposição (i),

$$\lim_{n \rightarrow \infty} \sum_{m=1}^n E \left[|X_{n,m}|^2; |X_{n,m}| \leq \epsilon \right] = \sigma^2 > 0. \quad (3.10)$$

Voltemos à equação (3.8). Somando $m = 1$ até n , deixando $n \rightarrow \infty$, usando (i), (ii) e (3.10) teremos,

$$\limsup_{n \rightarrow \infty} \sum_{m=1}^n |z_{n,m} - w_{n,m}| \leq \epsilon t^3 \sigma^2.$$

Como $\epsilon > 0$ é arbitrário, segue que a sequência acima converge para 0. Nosso próximo passo será, usando o Lema 3.14, mostrar que

$$\left| \prod_{m=1}^n \varphi_{n,m}(t) - \prod_{m=1}^n \left(1 - t^2 \sigma_{n,m}^2 / 2 \right) \right| \rightarrow 0.$$

Para isso vamos conferir a hipótese do Lema 3.14. Pela Propriedade 3.2 (f) sabemos que $|\varphi_{n,m}(t)| \leq 1$ para todo n, m . Para os termos do segundo produto temos que,

$$\sigma_{n,m}^2 \leq \epsilon^2 + E \left(|X_{n,m}|^2; |X_{n,m}| > \epsilon \right)$$

e ϵ é arbitrário então por (ii) temos que $\sup_m \sigma_{n,m}^2 \rightarrow 0$. Portanto se n é grande, temos $1 \geq 1 - t^2 \sigma_{n,m}^2 / 2 > -1$, para todo m .

Para completar a prova usaremos o Lema 3.11. Mostramos agora que $\sup_m \sigma_{n,m}^2 \rightarrow 0$ e a suposição (i) implica,

$$\sum_{m=1}^n -t^2 \sigma_{n,m}^2 / 2 \rightarrow -t^2 \sigma^2 / 2.$$

Desta forma $\prod_{m=1}^n \left(1 - t^2 \sigma_{n,m}^2 / 2 \right) \rightarrow \exp(t^2 \sigma^2 / 2)$. Completando assim a prova □

3.3.1

Necessidade da condição de Lindeberg

A segunda suposição do Teorema 3.15 é a condição de Lindeberg. No caso deste teorema, como as variáveis aleatórias possuem média zero a integral da condição fica igual ao valor esperado da variável aleatória elevado ao quadrado.

Sem a condição de Lindeberg veremos que é possível uma convergência de variáveis aleatórias em arrays independentes identicamente distribuídas cujo limite não é uma distribuição normal.

Iremos agora demonstrar a convergência em distribuição a uma Poisson, esta convergência também é chamada como a "lei dos eventos raros". Este nome vem do fato que o limite da soma dos indicadores do evento possuem

probabilidade pequena, por isso eventos raros, e esta soma convergirá em distribuição para uma distribuição de Poisson.

Para demonstrarmos a convergência em distribuição a uma Poisson precisaremos do seguinte resultado,

Lema 3.16. *Se b é um número complexo com $|b| \leq 1$ então $|e^b - (1 + b)| \leq |b|^2$.*

Demonstração. a demonstração deste Lema está em [6] página 125. \square

Teorema 3.17. *Para cada n , sejam $X_{n,m}$, onde $1 \leq m \leq n$, variáveis aleatórias independentes com $P(X_{n,m} = 1) = p_{n,m}$ e $P(X_{n,m} = 0) = 1 - p_{n,m}$. Suponha,*

(i) $\sum_{m=1}^n p_{n,m} \rightarrow \lambda \in (0, \infty)$; e

(ii) $\max_{1 \leq m \leq n} p_{n,m} \rightarrow 0$;

Então $S_n = X_{n,1} + \dots + X_{n,n} \xrightarrow{D} Z$, onde Z tem distribuição Poisson com média λ .

Demonstração. $\varphi_{n,m}(t) = E(\exp(itX_{n,m})) = (1 - p_{n,m}) + p_{n,m}e^{it} = 1 + p_{n,m}(e^{it} - 1)$.

Calcularemos a função característica. Então pela propriedade 3.2 (c) temos que a função característica de S_n será,

$$E[\exp(itS_n)] = \prod_{m=1}^n (1 + p_{n,m}(e^{it} - 1)).$$

Usaremos agora o Lema 3.14. Para isso temos que primeiro verificar as suas condições. Façamos $z_m = \exp(p_{n,m}(e^{it} - 1))$ e $w_m = 1 + p_{n,m}(e^{it} - 1)$.

Deixe Y ser uma variável aleatória com distribuição Poisson com média $p > 0$. Temos então que $\varphi_Y(t) = \exp(p(e^{it} - 1))$ é a sua função característica.

Para z_m temos que se $p_{n,m} = 0$ então $|z_m| = 1$, logo $|z_m| \leq 1$. Agora caso $p_{n,m} > 0$, z_m será uma função característica de uma distribuição Poisson com média $p_{n,m}$ então pelo Teorema 3.2 (f) $|z_m| \leq 1$.

Como podemos ver w_m é a função característica da variável aleatória $X_{n,m}$, então pelo Teorema 3.2 (f) temos que $|w_m| \leq 1$. Desta maneira podemos aplicar o Lema 3.14 e teremos assim,

$$\begin{aligned} \left| \exp\left(\sum_{m=1}^n p_{n,m}(e^{it} - 1)\right) - \prod_{m=1}^n [1 + p_{n,m}(e^{it} - 1)] \right| \\ \leq \sum_{m=1}^n \left| \exp(p_{n,m}(e^{it} - 1)) - [1 + p_{n,m}(e^{it} - 1)] \right| \quad (3.11) \end{aligned}$$

Utilizaremos agora o Lema 3.16, que será válido quando $\max_m p_{n,m} \leq 1/2$ desde que $|e^{it} - 1| \leq 2$. Teremos então,

$$\sum_{m=1}^n \left| \exp(p_{n,m}(e^{it} - 1)) - [1 + p_{n,m}(e^{it} - 1)] \right| \leq \sum_{m=1}^n p_{n,m}^2 |e^{it} - 1|^2.$$

Temos que $|e^{it} - 1| \leq 2$, então,

$$\sum_{m=1}^n \left| \exp(p_{n,m}(e^{it} - 1)) - [1 + p_{n,m}(e^{it} - 1)] \right| \leq 4 \left(\max_{1 \leq m \leq n} p_{n,m} \right) \sum_{m=1}^n p_{n,m}.$$

Pela suposições (i) e (ii) teremos que,

$$4 \left(\max_{1 \leq m \leq n} p_{n,m} \right) \sum_{m=1}^n p_{n,m} \rightarrow 0 \quad n \rightarrow \infty.$$

Por esta conclusão e pela suposição (i) teremos que,

$$E[\exp(itS_n)] \rightarrow \exp(\lambda(e^{it} - 1)),$$

Desta forma concluimos a prova, pois esta é a função característica de Z .

□

Podemos notar que neste teorema, em contraste com o de Lindeberg e Lindeberg-Feller em que nenhum termo contribui muito para o somatório, aqui a contribuição, quando ocorre é positiva e assim contribuirá para o valor do somatório.

O Teorema Central do Limite para Martingais é uma generalização do resultado para variáveis aleatórias só que agora no contexto de Martingais.

Neste capítulo veremos versões do Teorema Central do Limite para Martingais. Existem diferentes versões do Teorema Central do Limite para Martingais, aqui vamos provar um caso mais geral, para entrarmos em uma versão mais específica que nos interessará mais, pois é a abordada no artigo que queremos discutir.

Podemos dizer, de uma maneira informal, que as versões do Teorema Central do Limite para Martingais expostas neste capítulo, anunciam 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 padrão.

4.1

Formas do Teorema Central do Limite para Martingais

Nesta seção vamos demonstrar uma versão geral do Teorema Central do Limite para Martingais, depois entraremos no caso mais específico que nos interessa mais para o presente trabalho.

Trabalharemos, em grande parte, com vetores de Martingais que são geralmente derivados de Martingais ordinários, $\{S_n, \mathcal{F}_n, 1 \leq n \leq +\infty\}$, da seguinte maneira, podemos definir $k_n = n$, $\mathcal{F}_{n,i} = \mathcal{F}_i$ e $S_{n,i} = s_n^{-1} S_i$, onde s_n^{-1} é o desvio padrão de S_n . Neste caso $E(S_{n,k_n}^2) = 1$, e é comum fazer esta suposição para arbitrários vetores de Martingais. Desta forma teremos o vetor de Martingais como $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq n, n \geq 1\}$. Esta notação ela é utilizada em [7].

Antes de começarmos a discutir o Teorema Central do Limite para Martingais, precisamos do resultado anunciado no Lema abaixo, neste Lema, usaremos a convergência em probabilidade e em \mathcal{L}^p , que será definida abaixo.

Definição 4.1. *Sejam Y, Y_1, Y_2, \dots variáveis aleatórias definidas no mesmo espaço de probabilidade (Ω, \mathcal{F}, P) . Y_n converge para Y em probabilidade se para todo $\epsilon > 0$,*

$$P(|Y_n - Y| \geq \epsilon) \rightarrow 0 \quad n \rightarrow \infty.$$

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

A convergência em probabilidade significa que para valores grandes de n , as variáveis Y_n e Y , com probabilidade muito alta, possuem valores aproximadamente iguais.

Definição 4.2. A sequência de variáveis aleatórias Y_n converge em \mathcal{L}^p para Y , sendo $0 < p < \infty$, se $Y_n \in \mathcal{L}^p$, $Y \in \mathcal{L}^p$ e

$$\lim_{n \rightarrow \infty} E[|Y_n - Y|^p] = 0.$$

Podemos pensar na convergência em \mathcal{L}^p , para $p = 1$, como sendo uma convergência em media das variáveis, que será o caso no Lema abaixo.

Uma ultima definição importante para o Lema 4.4 é afirmarmos que uma coleção de variáveis aleatórias é uniformemente integrável.

Definição 4.3. Uma coleção de variáveis aleatórias $(X_i)_{i \in I}$ é uniformemente integrável se dado $\epsilon > 0$, existe $K \in [0, \infty)$ tal que $E[X_i | \mathbf{1}_{|X_i| \geq K}] \leq \epsilon$ para todo X_i .

Lema 4.4. Seja η^2 uma constante positiva e $X_{n,m}$ um array triangular de variáveis aleatórias onde $1 \leq m \leq n$ e $n \geq 1$. Suponha que:

$$\max_{1 \leq m \leq n} |X_{n,m}| \xrightarrow{p} 0. \quad (4.1)$$

$$\sum_{m=1}^n X_{n,m}^2 \xrightarrow{p} \eta^2, \quad n \rightarrow \infty \quad (4.2)$$

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

$$T_n(t) \rightarrow 1 \quad (\text{converge em } \mathcal{L}^1) \quad (4.3)$$

Então teremos $S_n = X_{n,1} + \dots + X_{n,n} \xrightarrow{d} Z$, onde Z é uma variável aleatória com a seguinte função característica $\varphi_Z(t) = \exp(-\frac{1}{2}\eta^2 t^2)$.

Demonstração. Como em [3] temos que:

$$e^{ix} = (1+ix)\exp\left(-\frac{1}{2}x^2 + r(x)\right), \quad \text{onde temos } |r(x)| \leq |x|^3 \quad \text{para } |x| \leq 1. \quad (4.4)$$

Faremos $I_n = \exp(itS_n)$ e

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

Desta forma podemos fazer:

$$I_n = \exp(itS_n) = \exp(it \sum_{m=1}^n X_{n,m}) = \prod_{m=1}^n \exp(itX_{n,m}),$$

Utilizando a fórmula (4.4) para I_n , teremos:

$$\begin{aligned} I_n &= \prod_{m=1}^n (1 + itX_{n,m}) \exp\left(-t^2/2X_{n,m}^2 + r(tX_{n,m})\right) \\ &= \left[\prod_{m=1}^n (1 + itX_{n,m}) \right] \left[\prod_{m=1}^n \exp\left(-t^2/2X_{n,m}^2 + r(tX_{n,m})\right) \right] \\ &= \prod_{m=1}^n (1 + itX_{n,m}) \exp\left(-t^2/2 \sum_{m=1}^n X_{n,m}^2 + \sum_{m=1}^n r(tX_{n,m})\right) \\ &= T_n W_n + T_n \exp(-\eta^2 t^2/2) - T_n \exp(-\eta^2 t^2/2) \\ &= T_n \exp(-\eta^2 t^2/2) + T_n (W_n - \exp(-\eta^2 t^2/2)). \end{aligned} \quad (4.5)$$

Como já vimos, 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 $t \in \mathbb{R}$. Desta forma é suficiente provar que:

$$E(I_n) \rightarrow \exp(-\eta^2 t^2/2). \quad (4.6)$$

Como $\exp(-\eta^2 t^2/2)$ é limitado e com a condição (4.3) teremos que:

$$E(T_n \exp(-\eta^2 t^2/2)) \rightarrow E(\exp(-\eta^2 t^2/2)) \quad (4.7)$$

Qualquer sequência de variáveis aleatórias, que converge em \mathcal{L}^1 , é uniformemente integrável, e a sequência $T_n(W_n - \exp(-\eta^2 t^2/2))$ é uniformemente integrável. As condições (4.1) e (4.2) implicam que quando $\max_{1 \leq m \leq n} |X_{n,m}| \leq 1$,

$$\left| \sum_{m=1}^n r(X_{n,m}) \right| \leq |t|^3 \sum_{m=1}^n |X_{n,m}|^3 \leq |t|^3 \left(\max_{1 \leq m \leq n} |X_{n,m}| \right) \left(\sum_{m=1}^n X_{n,m}^2 \right) \xrightarrow{p} 0,$$

quando $n \rightarrow \infty$.

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

$$E\left(T_n(W_n - \exp(-\eta^2 t^2/2))\right) \rightarrow 0. \quad (4.8)$$

Com as condições (4.7) e (4.8), teremos (4.6).

□

Agora faremos que o vetor $\{X_{n,m}\}$ seja um vetor de diferenças de Martingais.

Teorema 4.5. *Seja $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq k_n, n \geq 1\}$, um vetor de Martingais com média zero, quadrado integrável, isto é $S_{n,i} \in \mathcal{L}^2$, e com incrementos $X_{n,i}$. Seja η^2 uma constante real positiva. Suponha que:*

$$\max_{1 \leq i \leq k_n} |X_{n,i}| \xrightarrow{p} 0, \quad (4.9)$$

$$\sum_{i=1}^{k_n} X_{n,i}^2 \xrightarrow{p} \eta^2, \quad (4.10)$$

$$E \left(\max_i X_{n,i}^2 \right) \text{ é limitada em } n, \quad (4.11)$$

e as σ -álgebras são da forma, para $1 \leq i \leq k_n, n \geq 1$,

$$\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i} \quad (4.12)$$

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

Demonstração. Façamos $X'_{n,i} = X_{n,i} I(\sum_{j=1}^{i-1} X_{n,j}^2 \leq 2\eta^2)$ 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. Uma vez que:

$$P(X'_{n,i} \neq X_{n,i}) \leq P(U_{n,k_n}^2 > 2\eta^2) \rightarrow 0 \quad (4.13)$$

para $i \leq k_n$ e onde $U_{n,n}^2 = \sum_{i=1}^n X_{n,i}^2$. Temos também que $P(S'_{n,k_n} \neq S_{n,k_n}) \rightarrow 0$, e então:

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

Consequentemente $S_{n,k_n} \xrightarrow{D} Z$ se e somente se $S'_{n,k_n} \xrightarrow{D} Z$. Pela condição (4.13) as diferenças de martingal $X'_{n,i}$ satisfazem as condições (4.1) e (4.2) do Lema 4.4, devemos averiguar a condição (4.3).

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

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

Seja $x, y \in \mathbb{R}$, temos que $|x + iy|^2 = x^2 + y^2$. Então $E|T'_n|^2 = E \left[\left(\prod_{j=1}^{k_n} |1 + itX'_{n,j}| \right)^2 \right] = E \left[\prod_{j=1}^{k_n} (1 + t^2 X_{n,j}'^2) \right]$.

Assim sendo:

$$\begin{aligned} E|T'_n|^2 &= E \left(\prod_{j=1}^{k_n} (1 + t^2 X_{n,j}'^2) \right) \\ &\leq E \left[\left(\exp \left(t^2 \sum_{j=1}^{J_n-1} X_{n,j}'^2 \right) \right) (1 + t^2 X_{n,J_n}'^2) \right] \\ &\leq [\exp(2\eta^2 t^2)] (1 + E(X_{n,J_n}^2)) \end{aligned}$$

é limitado uniformemente em n pela condição (4.11). Desta forma T'_n é uniformemente integrável.

Fixamos um $m \geq 1$ e deixe $A \in \mathcal{F}_{m,k_m}$, pela condição (4.12), temos $A \in \mathcal{F}_{n,k_n}$ para todo $n \geq m$, tal que em n ,

$$\begin{aligned} E(T'_n I(A)) &= E \left(I(A) \prod_{j=1}^{k_n} (1 + itX'_{n,j}) \right) \\ &= E \left(I(A) \prod_{j=1}^{k_m} (1 + itX'_{n,j}) \prod_{j=k_m+1}^{k_n} E(1 + itX'_{n,j}) | \mathcal{F}_{n,j-1} \right) \\ &= E \left(I(A) \prod_{j=1}^{k_m} (1 + itX'_{n,j}) \right) \\ &= P(A) + R_n, \end{aligned}$$

onde o termo R_n consiste nos $2^{k_m} - 1$ termos na forma, $E(I(A)(it)^r X'_{n,j_1} X'_{n,j_2} \dots X'_{n,j_r})$, sendo $1 \leq r \leq k_m$ e $1 \leq j_1 \leq j_2 \leq \dots \leq j_r \leq k_m$.

Uma vez que,

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

Daí temos que,

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

Agora, para qualquer $\epsilon > 0$, como $\max_i |X_{n,i}| \leq \epsilon + \max_i |X_{n,i}| I(|X_{n,i}| > \epsilon)$, temos pela propriedade da esperança:

$$E \left(\max_i |X_{n,i}| \right) \leq \epsilon + E \left(\max_i |X_{n,i}| I(|X_{n,i}| > \epsilon) \right).$$

Agora utilizando a desigualdade de Cauchy para limitar $E(\max_i |X_{n,i}| I(|X_{n,i}| > \epsilon))$, em que $p = q = 2$, teremos:

$$E \left(\max_i |X_{n,i}| \right) \leq \epsilon + \left[E \left(\max_i X_{n,i}^2 \right) P \left(\max_i |X_{n,i}| > \epsilon \right) \right]^{1/2},$$

então,

$$E \left(\max_i |X_{n,i}| \right) \leq \epsilon \quad \text{quando } n \rightarrow \infty.$$

Como ϵ é tão pequeno quanto se queira, segue então que $E(\max_i |X_{n,i}|) \rightarrow$

0 e $R_n \rightarrow 0$. Desta forma $E[T'_n I(A)] \rightarrow P(A)$.

Façamos $\mathcal{F}_\infty = \bigvee_1^\infty \mathcal{F}_{n,k_n}$ ser a σ -álgebra gerada pela $\bigcup_{n=1}^\infty \mathcal{F}_n$. Para qualquer $B \in \mathcal{F}_\infty$ e algum $\epsilon > 0$, existe um m e um $A \in \mathcal{F}_{m,k_m}$ tal que $P(A \Delta B) < \epsilon$, onde Δ denota a diferença simétrica. Uma vez que T'_n é uniformemente integrável e $|E[T'_n I(B)] - E[T'_n I(A)]| \leq E[|T'_n| I(A \Delta B)]$, temos que $\sup_n |E[T'_n I(B)] - E[T'_n I(A)]|$, pode ser arbitrariamente muito pequeno, escolhendo um ϵ suficientemente pequeno.

Temos de $E[T'_n I(A)] \rightarrow P(A)$, que para qualquer $B \in \mathcal{F}_\infty$, $E[T'_n I(B)] \rightarrow P(B)$. Isto implica que para qualquer variável aleatória X \mathcal{F}_∞ -mensurável, obteremos $E[T'_n X] \rightarrow E(X)$. Finalmente, se $A \in \mathcal{F}$ então,

$$E[T'_n I(A)] = E[T'_n E(I(A)|\mathcal{F}_\infty)] \rightarrow E[E(I(A)|\mathcal{F}_\infty)] = P(A).$$

Com isso provamos que a condição (4.3) do Lema 4.4 se aplica logo a prova está concluída. \square

Corolário 4.6. *Se as condições (4.9) e (4.11) forem substituídas pela condição de Lindeberg:*

$\sum_{i=1}^{k_n} E(X_{n,i}^2 I(|X_{n,i}| > \epsilon) | \mathcal{F}_{n,i-1}) \xrightarrow{p} 0$, para todo $\epsilon > 0$, também a condição (4.10) for substituída por uma condição análoga com a variância condicional:

$V_{n,k_n}^2 = \sum_{i=1}^{k_n} E(X_{n,i}^2 | \mathcal{F}_{n,i-1}) \xrightarrow{p} \eta^2$, e se a (4.12) se manter, então a conclusão do Teorema 4.5 permanece.

Para verificarmos o Corolário 4.6 do teorema acima, devemos analisar a relação entre a esperança condicional, $V_{n,i}^2$ e o somatório $\sum_{i=1}^{k_n} X_{n,i}^2$.

Veremos que a variância condicional com frequência pode ser aproximada pelo somatório $\sum_{i=1}^{k_n} X_{n,i}^2$. Para isso precisamos da seguinte proposição anunciada abaixo:

Proposição 4.7. *Seja $\{S_{n,i} = \sum_{i=1}^{k_n} X_{n,i}, \mathcal{F}_{n,i}, n \geq 1, 1 \leq i \leq k_n\}$ um martingal com média zero. Suponha que as variâncias condicionais, $V_{n,k_n}^2 = \sum_{i=1}^{k_n} E[X_{n,i}^2 | \mathcal{F}_{n,i-1}]$, são justas, ou seja:*

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

e a condição de Lindeberg está assegurada, então:

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

Demonstração. Demonstração está em [7] página 45. \square

Com a condição de Lindeberg assegurada e as variâncias condicionais, V_{n,k_n}^2 , justas, implicam, pela Proposição 4.7, que

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

Para o Corolário 4.6 ser verdadeiro, basta demonstrar que as suas condições são equivalentes as condições do Teorema 4.5.

Como a condição do Corolário 4.6 mesmo determina, $V_{n,k_n}^2 = \sum E(X_n^2 | \mathcal{F}_{n,i-1}) \xrightarrow{p} \eta^2$, onde η^2 é uma constante positiva, a sequência de variâncias condicionais serão justas. No Corolário 4.6 a condição de Lindeberg está assegurada logo poderemos aplicar a Proposição 4.7.

Desta forma a condição (4.10) do Teorema 4.5 é equivalente a $V_{n,k_n}^2 \xrightarrow{p} \eta^2$.

As condições (4.9) e (4.11) do Teorema 4.5 são equivalentes a condição de Lindeberg. Logo o Corolário 4.6 se mantém.

Teorema 4.8. *Suponha que todas as condições do Teorema 4.5 estão asseguradas e que η^2 é uma constante real positiva. Então:*

$$\frac{S_{n,k_n}}{(\sum_j^{k_n} X_{n,j}^2)^{1/2}} \xrightarrow{D} N(0, 1) \quad (4.14)$$

Corolário 4.9. *Suponha que as condições do Corolário 4.8 estejam asseguradas e que η^2 é uma constante real positiva. Então:*

$$\frac{S_{n,k_n}}{V_{n,k_n}} \xrightarrow{D} N(0, 1) \quad (4.15)$$

Demonstração. (Prova do Teorema 4.8) Suponha Z uma variável aleatória com função característica $\varphi(t)_Z = \exp(-\frac{1}{2}\eta^2 t^2)$. Visto no Teorema 4.5 que $S_{n,k_n} \xrightarrow{D} Z$, para qualquer t real teremos:

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

Desta maneira temos que:

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

Da condição (4.10) do Teorema 4.5 obteremos o seguinte resultado:

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

\square

Como já vimos na Proposição 4.7,

$$\left| \sum_j^{k_n} X_{n,j}^2 - V_{n,k_n}^2 \right| \xrightarrow{p} 0,$$

fazemos a prova do Corolário 4.9 da mesma maneira, apenas substituindo $\sum_j^{k_n} X_{n,j}^2$ por V_{n,k_n}^2 .

Com os teoremas e corolários demonstrados acima podemos tratar do Teorema Central do Limite para Martingais de uma forma mais específica, da mesma que o artigo [4], que exploraremos mais à frente, o fez. Exporemos a notação utilizada em [4].

Pensemos em $\mathbf{X} = (X_1, \dots, X_n)$ sendo uma sequência de variáveis aleatórias tais que, para todo i , $X_i \in \mathcal{L}^2$ e X_i satisfaz $E(X_i | \mathcal{F}_{i-1}) = 0$, onde \mathcal{F}_i é uma filtração natural. Desta maneira \mathbf{X} é um martingal com séries de incrementos.

Usaremos a partir de agora as seguintes notações, todas indexadas em n :

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

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

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

sendo esta última a variância condicional normalizada.

Fixaremos agora a constante η^2 , que utilizamos nos teoremas e corolários acima, como sendo igual a 1. Como já vimos na Proposição 4.7,

$$\left| \sum_j^{k_n} X_{n,j}^2 - V_{n,k_n}^2 \right| \xrightarrow{p} 0,$$

desta forma, ao analisarmos a variância condicional normalizada podemos verificar que se:

$$V^2(\mathbf{X}_n) \xrightarrow{p} 1 \quad \text{quando} \quad n \rightarrow \infty$$

e se a condição de Lindeberg for satisfeita, então teremos que

$$S(\mathbf{X}_n)/s(\mathbf{X}_n) \xrightarrow{D} N(0, 1).$$

Para uma prova formal podemos verificá-la em [2] e [3].

Discutiremos neste capítulo as taxas de convergência do Teorema Central do Limite e por fim vamos fazer uma análise do artigo [4], o qual discute justamente a taxa de convergência no Teorema Central do Limite para Martingais.

As taxas de convergência são uma importante ferramenta para verificarmos com que velocidade a distribuição em questão está se aproximando da função de distribuição normal padrão. Podemos pensar, de forma informal, para termos uma intuição, onde \mathcal{F}_n é uma filtração natural, em um martingal $\{S_n, \mathcal{F}_n, n \geq 1\}$, com média zero, que é uma soma de incrementos X_i 's, quão grande terá de ser o n para que esta aproximação ocorra. Estas taxas, como veremos mais a frente, são definidos como o supremo da distância da distribuição em questão à função de distribuição normal padrão, sendo o supremo dado em todos os pontos do domínio destas distribuições.

No artigo [4] é discutida a otimização de um dos termos do limite da taxa de convergência, focando no caso restrito de martingais com incrementos limitados. Para demonstrar esta otimização do termo, no artigo é demonstrado um novo limite para a taxa de convergência no Teorema Central do Limite para martingais com incrementos limitados. Veremos que esta otimização na verdade gerará uma generalização deste termo.

A discussão em [4] começa na seção 5.2. Para dar contexto, apresentaremos na seção 5.1 o resultado da taxa de convergência do Teorema Central do Limite para o caso específico de $S_n = \sum_{i=1}^n X_i$, onde as variáveis aleatórias X_n são independentes.

5.1

Taxa de Convergência do Teorema Central do Limite para variáveis aleatórias

Nesta seção vamos discutir as taxas de convergência para o Teorema Central do Limite para variáveis aleatórias. O que nos interessa agora, de forma informal, é saber o quão rapidamente a distribuição que estamos estudando se aproxima da distribuição normal padrão.

Para fazermos esta análise vamos recorrer a uma versão do Teorema de Berry-Esseen, neste iremos observar a distância entre a distribuição em questão

e distribuição normal.

Teorema 5.1. *Seja X_1, X_2, \dots uma sequência de variáveis aleatórias independentes e identicamente distribuídas com $E[X_i] = 0$, $E[X_i^2] = \sigma^2$ e $E|X_i^3| = \rho < \infty$. Se $F_n(x)$ é a função de distribuição de $(X_1 + \dots + X_n)/\sigma n^{1/2}$ e $\Phi(x)$ é a distribuição normal padrão, então*

$$|F_n(x) - \Phi(x)| \leq 3n^{1/2}\rho/\sigma^3.$$

Demonstração. Demonstração está em [6] página 137. \square

A inequação do Teorema 5.1 é válida para todo x e n , desta forma podemos pensá-la da seguinte maneira,

$$\sup_{-\infty < x < \infty} |F_n(x) - \Phi(x)| \leq 3n^{1/2}\rho/\sigma^3.$$

Visto isso podemos pensar informalmente que procuraremos o ponto nas duas distribuições em que ocorra a maior distância possível e buscaremos os seus melhores limites.

5.2

Taxas de Convergência no Teorema Central do Limite para Martingais

No artigo [4] veremos que o objetivo será dada uma limitada taxa de convergência no Teorema Central do Limite para Martingais poderemos discutir a otimização de um dos termos deste limite e a sua generalização.

Relembrando que agora usaremos as seguintes notações, semelhantes a [4].

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

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

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

sendo esta última a variância condicional normalizada e todas indexadas em n .

Nesta seção vamos discutir os dois teoremas demonstrados no artigo [4], para isso consideremos um martingal em que sua variância condicional normalizada se aproxima de 1 e as condições de Lindeberg são satisfeitas. Desta

forma, como já vimos anteriormente, haverá a convergência em distribuição para uma função de distribuição normal padrão.

Considerando as notações colocadas na seção anterior, formalmente teremos:

$$V^2(\mathbf{X}_n) \xrightarrow{p} 1 \quad (5.1)$$

quando $n \rightarrow \infty$ e com a condição de Lindeberg satisfeita, a soma, $S(\mathbf{X}_n)/s(\mathbf{X}_n)$ convergirá para uma distribuição de normal padrão, isto é,

$$\forall t \in \mathbb{R}, \quad P[S(\mathbf{X}_n)/s(\mathbf{X}_n) \leq t] \rightarrow \Phi(t)$$

, quando $n \rightarrow \infty$.

O nosso interesse está na velocidade da convergência do Teorema Central do Limite e nos seus limites, para isso iremos observar a maior distância das distribuições, isto é,

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

Podemos pensar que o limite da taxa de convergência no Teorema Central do Limite pode ser escrita como soma de dois termos, $A_p + B_p$, para qualquer $p \geq 1$, onde faremos $A_p = \|V^2 - 1\|_p^{p/(2p+1)}$.

O foco do artigo é discutir a otimização do termo A_p para qualquer $p \geq 1$, focando na classe restrita de Martingais com incrementos limitados. Para isso será necessário dar um novo limite para a taxa de convergência no Teorema Central do Limite para Martingais com incrementos limitados.

Podemos substituir a convergência em probabilidade em (5.1) por convergência em L^p para algum $p \in [1, +\infty]$, pois se converge em probabilidade temos convergência em L^p . Além disso fazer estimativas do termo $\|V^2 - 1\|_p$ parece mais conveniente para aplicações em situações práticas.

Antes de apresentarmos o teorema que nos dará melhor limite de aproximação das distribuições devemos introduzir uma notação,

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

e um teorema que é apresentado no artigo [4] e será utilizado,

Teorema 5.2. *[[1]] Seja $\gamma \in (0, +\infty)$. Existe uma constante $C_\gamma > 0$ tal que para qualquer $n > 2$ e algum $\mathbf{X} = (X_1, \dots, X_n)$, que satisfaça $\|\mathbf{X}\|_\infty \leq \gamma$ e $V^2(\mathbf{X}) = 1$ quase certamente, temos,*

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

O objetivo dos próximos teoremas será buscar uma generalização do Corolário abaixo apresentado:

Corolário 5.3. *[[1]] Seja $\gamma \in (0, +\infty)$. Existe uma constante $C'_\gamma > 0$ tal que para qualquer $n \geq 2$ e algum $\mathbf{X} = (X_1, \dots, X_n)$ satisfazendo $\|\mathbf{X}\|_\infty \leq \gamma$, temos,*

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

No artigo [4], em referência a [1], é dito que o termo limitado $\|V^2(\mathbf{X}) - 1\|_1^{1/3}$ será de fato o ótimo, mesmo para todas as classes de Martingais com incrementos limitados consideradas no Corolário 5.3. Desta maneira o Corolário 5.3 responde de forma satisfatória a questão da otimização do termo para $p = 1$.

Buscaremos agora uma generalização deste resultado para qualquer $p \in [1, +\infty)$.

Teorema 5.4. *Seja $p \in [1, +\infty)$ e $\gamma \in (0, +\infty)$. Existe uma constante $C_{p,\gamma} > 0$ tal que para qualquer $n \geq 2$ e qualquer \mathbf{X} que satisfaça $\|X\|_\infty \leq \gamma$,*

$$D(\mathbf{X}) \leq C_{p,\gamma} \left[\frac{n \log(n)}{s^3(\mathbf{X})} + \left(\|V^2(\mathbf{X}) - 1\|_p^p + s^{-2p}(\mathbf{X}) \right)^{\frac{1}{2p+1}} \right] \quad (5.2)$$

Demonstração. Seja $\mathbf{X}' = (X'_1, \dots, X'_{2n})$ Façamos,

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

Para $i \leq \tau$, definimos $X'_i = X_i$. Deixamos r ser a parte inteira de,

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

Como $\|\mathbf{X}\|_\infty \leq \gamma$, então $r \leq n$. Condicionado em \mathcal{F}_τ e para $1 \leq i \leq r$, seja X'_i uma variável aleatória independente tal que $P[X'_{\tau+i} = \pm\gamma] = 1/2$. Se $\tau + r < 2n$, então teremos $X'_{\tau+r+1}$ tal que,

$$P \left[X'_{\tau+t+i} = \pm \left(s^2(\mathbf{X}) - \sum_{i=1}^{\tau} E(X_i^2 | \mathcal{F}_{i-1}) - r\gamma^2 \right)^{1/2} \right] = \frac{1}{2},$$

com o sinal determinando independência de todo resto. Finalmente, se $\tau + r + 1 < 2n$, então teremos $X'_{\tau+r+i} = 0$ para $i \geq 2$.

Possivelmente alargando as σ -álgebras, podemos assumir que X'_i é \mathcal{F}_i -mensurável para $i \leq n$, e definimos $\mathcal{F}_i = \sigma(\mathcal{F}_n, X'_{n+1}, \dots, X'_{n+i})$ se $i \geq n$. Teremos,

$$\begin{aligned} \sum_{i=\tau+1}^{2n} E(X_i'^2 | \mathcal{F}_{i-1}) &= E(X_{\tau+1}'^2 | \mathcal{F}_\tau) + E(X_{\tau+2}'^2 | \mathcal{F}_{\tau+1}) + \dots + E(X_{\tau+r}'^2 | \mathcal{F}_{\tau+r-1}) \\ &\quad + E(X_{\tau+r+1}'^2 | \mathcal{F}_{\tau+r}) + \dots + E(X_{2n}'^2 | \mathcal{F}_{2n-1}). \end{aligned}$$

Para a construção desta igualdade podemos pensar em partes da parcela da soma,

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

O primeiro termo na expressão acima tem r termos, desta maneira teremos,

$$\begin{aligned} \sum_{i=\tau+1}^{2n} E(X_i'^2 | \mathcal{F}_{i-1}) &= r\gamma^2 + s^2(\mathbf{X}) - \sum_{i=1}^{\tau} E(X_i^2 | \mathcal{F}_{i-1}) - r\gamma^2 + 0 \\ \sum_{i=\tau+1}^{2n} E(X_i'^2 | \mathcal{F}_{i-1}) &= s^2(\mathbf{X}) - \sum_{i=1}^{\tau} E(X_i^2 | \mathcal{F}_{i-1}). \end{aligned} \quad (5.3)$$

Podemos reescrever a expressão como,

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

Consequentemente, $s^2(\mathbf{X}') = s^2(\mathbf{X})$ e $V^2(\mathbf{X}') = 1$ quase certamente. Sendo assim a sequência \mathbf{X}' satisfaz aos requisitos do Teorema 5.2 e portanto,

$$D(\mathbf{X}') \leq 4C_\gamma \frac{n \log(n)}{s^3(\mathbf{X})}. \quad (5.4)$$

Para qualquer $x > 0$, temos,

$$P \left[\frac{S(\mathbf{X})}{s(\mathbf{X})} \leq t \right] \leq P \left[\frac{S(\mathbf{X})}{s(\mathbf{X})} \leq t, \frac{|S(\mathbf{X}) - S(\mathbf{X}')|}{s(\mathbf{X})} \leq x \right] + P \left[\frac{|S(\mathbf{X}) - S(\mathbf{X}')|}{s(\mathbf{X})} \geq x \right],$$

No segundo termo da soma a direita podemos usar a desigualdade de Markov, e

$$P \left[\frac{|S(\mathbf{X}) - S(\mathbf{X}')|}{s(\mathbf{X})} \geq x \right] \leq \frac{1}{x^{2p}} E \left[\left| \frac{S(\mathbf{X}) - S(\mathbf{X}')}{s(\mathbf{X})} \right|^{2p} \right].$$

Dessa forma obteremos,

$$P \left[\frac{S(\mathbf{X})}{s(\mathbf{X})} \leq t \right] \leq P \left[\frac{S(\mathbf{X}')}{s(\mathbf{X})} \leq t + x \right] + \frac{1}{x^{2p}} E \left[\left| \frac{S(\mathbf{X}) - S(\mathbf{X}')}{s(\mathbf{X})} \right|^{2p} \right]. \quad (5.5)$$

Agora seja $f(y)$ a função de densidade da normal padrão, teremos,

$$\Phi(t + x) = \Phi(t) + \int_t^{t+x} f(y) dy \leq \Phi(t) + \frac{x}{\sqrt{2\pi}}. \quad (5.6)$$

Isto ocorre porque $f(y) \leq \frac{1}{\sqrt{2\pi}} \forall y$.

Então por (5.4) e (5.6), o primeiro termo do lado direito de (5.5) é menor do que,

$$\Phi(t + x) + 4C_\gamma \frac{n \log(n)}{s^3(\mathbf{X})} \leq \Phi(t) + \frac{x}{\sqrt{2\pi}} + 4C_\gamma \frac{n \log(n)}{s^3(\mathbf{X})} \quad (5.7)$$

Iremos agora controlar o segundo termo de (5.5). Para isso note que,

$$S(\mathbf{X}) - S(\mathbf{X}') = \sum_{i=\tau+1}^{2n} (X_i - X'_i), \quad (5.8)$$

onde colocamos $X_i = 0$ para $i > n$. Dado que $\tau + 1$ é um tempo de parada, condicionado em τ , a sequência $(X_i - X'_i)_{i \leq \tau+2}$ formará uma sequência de diferença de martingais.

Antes de continuarmos com a demonstração vamos apresentar uma proposição que utilizaremos.

Proposição 5.5. *Se $\{S_i = \sum_{j=1}^i X_j, \mathcal{F}_i, 1 \leq i \leq n\}$ é um martingal e $p > 0$, então existe uma constante C dependendo somente de p tal que,*

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

Demonstração. Demonstração pode ser encontrada em [7] página 28. \square

Desta forma com a sequência $(X_i - X'_i)_{i \leq \tau+2}$ poderemos usar a Proposição 5.5, descartaremos a parcela da soma indexada por $\tau + 1$ que aparece em (5.8), pois esta é uniformemente limitada, isto nos fornecerá,

$$\begin{aligned} & \frac{1}{C} E \left[\left| \sum_{i=\tau+2}^{2n} (X_i - X'_i) \right|^{2p} \right] \\ & \leq E \left[\left(\sum_{i=\tau+2}^{2n} E[(X_i - X'_i)^2 | \mathcal{F}_{i-1}] \right)^p \right] + E \left[\max_{\tau+2 \leq i \leq 2n} |X_i - X'_i|^{2p} \right]. \quad (5.9) \end{aligned}$$

O máximo no lado direito da equação (5.9) é limitado por $2\gamma^{2p}$, pela definição de X_t .

$$\frac{1}{C} E \left[\left| \sum_{i=\tau+2}^{2n} (X_i - X'_i) \right|^{2p} \right] \leq E \left[\left(\sum_{i=\tau+2}^{2n} E[(X_i - X'_i)^2 | \mathcal{F}_{i-1}] \right)^p \right] + 2\gamma^{2p}. \quad (5.10)$$

Quanto ao outro termo temos que X_i e X'_i são variáveis aleatórias ortogonais, então,

$$\sum_{i=\tau+1}^{2n} E[(X_i - X'_i)^2 | \mathcal{F}_{i-1}] = \sum_{i=\tau+1}^{2n} E[X_i^2 | \mathcal{F}_{i-1}] + \sum_{i=\tau+1}^{2n} E[X_i'^2 | \mathcal{F}_{i-1}]. \quad (5.11)$$

Vamos agora analisar separadamente as duas parcelas de (5.11). Como já vimos em (5.3), para o segundo termo temos,

$$\sum_{i=\tau+1}^{2n} E[X_i'^2 | \mathcal{F}_{i-1}] = s^2(\mathbf{X}) - \sum_{i=1}^{\tau} E[X_i^2 | \mathcal{F}_{i-1}]. \quad (5.12)$$

Para a primeira parcela de (5.11), teremos,

$$\sum_{i=\tau+1}^{2n} E[X_i^2 | \mathcal{F}_{i-1}] = \sum_{i=1}^{2n} E[X_i^2 | \mathcal{F}_{i-1}] - \sum_{i=1}^{\tau} E[X_i^2 | \mathcal{F}_{i-1}], \quad (5.13)$$

onde sabemos que,

$$\sum_{i=1}^{2n} E[X_i^2 | \mathcal{F}_{i-1}] = \sum_{i=1}^n E[X_i^2 | \mathcal{F}_{i-1}], \quad \text{pois } X_i = 0 \quad \text{para } i > n.$$

Desta forma podemos reescrever (5.13) da seguinte maneira

$$\sum_{i=\tau+1}^{2n} E[X_i^2 | \mathcal{F}_{i-1}] = \sum_{i=1}^n E[X_i^2 | \mathcal{F}_{i-1}] - \sum_{i=1}^{\tau} E[X_i^2 | \mathcal{F}_{i-1}] \quad (5.14)$$

Das notações que foram rerepresentadas no começo desta seção temos que,

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

$$\sum_{i=1}^n E(X_i^2 | \mathcal{F}_{i-1}) = s^2(\mathbf{X})V^2(\mathbf{X}). \quad (5.15)$$

Substituindo em (5.14), temos

$$\sum_{i=\tau+1}^{2n} E[X_i^2 | \mathcal{F}_{i-1}] = s^2(\mathbf{X})V^2(\mathbf{X}) - \sum_{i=1}^{\tau} E[X_i^2 | \mathcal{F}_{i-1}] \quad (5.16)$$

Vamos agora voltar a (5.11), nesta equação substituiremos suas parcelas por (5.16) e (5.12), respectivamente, nos dando,

$$\sum_{i=\tau+1}^{2n} E[(X_i - X'_i)^2 | \mathcal{F}_{i-1}] = s^2(\mathbf{X})V^2(\mathbf{X}) + s^2(\mathbf{X}) - 2 \sum_{i=1}^{\tau} E[X_i^2 | \mathcal{F}_{i-1}] \quad (5.17)$$

Agora vamos observar o termo $\sum_{i=1}^{\tau} E[X_i^2 | \mathcal{F}_{i-1}]$ da equação (5.17). Se $\tau = n$, então, por (5.15), temos,

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

caso contrário, isto significará que $\sum_{i=1}^{\tau+1} E[X_i^2 | \mathcal{F}_{i-1}]$ excederá $s^2(\mathbf{X})$. Temos que os incrementos são limitados logo,

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

Desta forma podemos fazer,

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

Consequentemente de (5.17), obteremos,

$$\sum_{i=\tau+1}^{2n} E[(X_i - X'_i)^2 | \mathcal{F}_{i-1}] \leq |s^2(\mathbf{X})V^2(\mathbf{X}) - s^2(\mathbf{X})| + 2\gamma^2. \quad (5.18)$$

De (5.18), pela propriedade da esperança e como $p \in [1, +\infty)$,

$$E \left[\left(\sum_{i=\tau+1}^{2n} E[(X_i - X'_i)^2 | \mathcal{F}_{i-1}] \right)^p \right] \leq E \left[\left(|s^2(\mathbf{X})V^2(\mathbf{X}) - s^2(\mathbf{X})| + 2\gamma^2 \right)^p \right] \quad (5.19)$$

Iremos agora reorganizar a equação (5.10), com o resultado obtido em (5.19),

$$\begin{aligned}
\frac{1}{C} E \left[\left| \sum_{i=\tau+2}^{2n} (X_i - X'_i) \right|^{2p} \right] &\leq E \left[\left(|s^2(\mathbf{X})V^2(\mathbf{X}) - s^2(\mathbf{X})| + 2\gamma^2 \right)^p \right] + 2\gamma^{2p} \\
&\leq E \left[\left(|s^2(\mathbf{X})(V^2(\mathbf{X}) - 1)| + 2\gamma^2 \right)^p \right] + 2\gamma^{2p} \\
&\leq E \left[\left(s^2(\mathbf{X})|V^2(\mathbf{X}) - 1| + 2\gamma^2 \right)^p \right] + 2\gamma^{2p} \\
&\leq E \left[p \left(s^{2p}(\mathbf{X})|V^2(\mathbf{X}) - 1|^p + 2^p \gamma^{2p} \right) \right] + 2\gamma^{2p} \\
&\leq p s^{2p}(\mathbf{X}) \|V^2(\mathbf{X}) - 1\|_p^p + 2^p \gamma^{2p} p + 2\gamma^{2p} \\
&\leq p s^{2p}(\mathbf{X}) \|V^2(\mathbf{X}) - 1\|_p^p + \gamma^{2p} (2^p p + 1).
\end{aligned} \tag{5.20}$$

Então teremos,

$$E \left[\left| \sum_{i=\tau+2}^{2n} (X_i - X'_i) \right|^{2p} \right] \leq C \left(s^{2p}(\mathbf{X}) \|V^2(\mathbf{X}) - 1\|_p^p + \gamma^{2p} \right). \tag{5.21}$$

Agora por (5.8) e (5.21), obtemos,

$$\frac{1}{x^{2p}} E \left[\left| \frac{S(\mathbf{X}) - S(\mathbf{X}')}{s(\mathbf{X})} \right|^{2p} \right] \leq \frac{C}{x^{2p}} \left(\|V^2(\mathbf{X}) - 1\|_p^p + \frac{\gamma^{2p}}{s^{2p}(\mathbf{X})} \right). \tag{5.22}$$

Agora vamos combinar as equações (5.22), (5.7) e (5.5), obteremos assim,

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

Otimizando isto sobre $x > 0$ teremos a correta estimativa. O limite inferior é encontrado da mesma maneira. \square

Agora justificaremos a otimização do termo $\|V^2(\mathbf{X}) - 1\|_p^{p/(2p+1)}$ que aparece no lado direito de (5.2).

Teorema 5.6. *Seja $p \in [1, +\infty)$ e $\alpha \in (1/2, 1)$. Existe uma sequência de elementos \mathbf{X}_n tal que,*

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

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

$$\|V^2(\mathbf{X}_n) - 1\|_p^{p/(2p+1)} = O(n^{(\alpha-1)/2}),$$

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

Se escrevemos $a_n \simeq b_n$, significa que existe $C > 0$, tal que $a_n/C \leq b_n \leq Ca_n$, para todo n suficientemente grande.

A ideia no Teorema 5.6 é mostrar que há sequência de elementos que estão dentro das condições do Teorema 5.4 e que com esta não é possível encontrar um termo mais forte que $\|V^2(\mathbf{X}_n) - 1\|_p^{p/(2p+1)}$, ou seja, um termo que vá mais rapidamente para zero que este. Podemos ver isso como consequência do Teorema 5.6, mesmo para um valor fixo de $\alpha \in (1/2, 1)$. Por exemplo, com $\alpha = 3/4$, temos o seguinte resultado,

Teorema 5.7 (Caso especial do Teorema 5.6). *Seja $p \in [1, +\infty)$ e $\alpha = 3/4$. Existe uma sequência de elementos \mathbf{X}_n tal que,*

$$\begin{aligned} \|\mathbf{X}_n\|_\infty &\leq 2, \\ s(\mathbf{X}_n) &\simeq \sqrt{n}, \\ \|V^2(\mathbf{X}_n) - 1\|_p^{p/(2p+1)} &= O(n^{1/8}), \\ \limsup_{n \rightarrow +\infty} n^{1/8} D(\mathbf{X}_n) &> 0. \end{aligned}$$

O Teorema 5.7 implica a existência de uma constante $\epsilon > 0$ e uma sequência X_n tal que,

$$\begin{aligned} \|\mathbf{X}_n\|_\infty &\leq 2, \\ s(\mathbf{X}_n) &\simeq \sqrt{n}, \\ \|V^2(\mathbf{X}_n) - 1\|_p^{p/(2p+1)} &\leq Cn^{1/8}, \\ D(\mathbf{X}_n) &> \epsilon n^{-1/8} \end{aligned}$$

para infinitos n , desta forma temos que,

$$D(X_n) > \frac{\epsilon}{C} \|V^2(\mathbf{X}_n) - 1\|_p^{p/(2p+1)}.$$

Confirmando que o termo $\|V^2(\mathbf{X}_n) - 1\|_p^{p/(2p+1)}$ é o melhor possível até multiplicado por uma constante. Faremos agora a prova do Teorema 5.6, que consequentemente prova o Teorema 5.7.

Demonstração. Sejam $p \geq 1$ e $\alpha \in (1/2, 1)$ fixos. Façamos $(X_{n,i})_{1 \leq i \leq n-n^\alpha}$ serem variáveis aleatórias independentes com $P[X_{n,i} = \pm 1] = 1/2$. A subsequência $(X_{n,i})_{n-n^\alpha < i \leq n}$ é definida recursivamente. Deixe,

$$\lambda_{n,i} = \sqrt{n - i + \kappa_n^2},$$

onde $\kappa_n = n^{1/4}$. Assumindo que $X_{n,1}, \dots, X_{n,i-1}$ foram definidos, escrevemos $\mathcal{F}_{n,i-1}$ para a σ -álgebra que eles geram, e fazamos,

$$S_{n,i-1} = \sum_{j=1}^{i-1} X_{n,j}.$$

Antes de continuar a prova, definiremos δ_x como a massa de Dirac no ponto x , ou seja,

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

Para qualquer i tal que $n - n^\alpha < i \leq n$, construímos $X_{n,i}$ tal que

$$P[X_{n,i} \in \cdot | \mathcal{F}_{n,i-1}] = \begin{cases} \delta_{-\sqrt{3/2}} + \delta_{\sqrt{3/2}} & \text{se } S_{n,i-1} \in [\lambda_{n,i}, 2\lambda_{n,i}], \\ \delta_{-\sqrt{1/2}} + \delta_{\sqrt{1/2}} & \text{se } S_{n,i-1} \in [-2\lambda_{n,i}, \lambda_{n,i}], \\ \delta_{-1} + \delta_1 & \text{caso contrário.} \end{cases}$$

Escreveremos $\mathbf{X}_n = (X_{n,1}, \dots, X_{n,n})$ e $\mathbf{X}_{n,i} = (X_{n,1}, \dots, X_{n,i})$ para qualquer $i \leq n$. Seja

$$\delta(i) = \sup_{n \leq i} D(\mathbf{X}_{n,i}).$$

Proposição 5.8. *Uniformemente sobre n ,*

$$\|V^2(\mathbf{X}_{n,i}) - 1\|_p = O(i^{(\alpha-1)(1+1/2p)}) \quad (i \rightarrow \infty) \quad (5.23)$$

e

$$\delta(i) = O(i^{(\alpha-1)/2}) \quad (i \rightarrow \infty) \quad (5.24)$$

A prova seguirá da seguinte forma, primeiro limitaremos $\|V^2(\mathbf{X}_{n,i}) - 1\|_p$ em termos de $(\delta(j))_{j \leq i}$ no Lema 5.9. Este nos fornece uma desigualdade com a sequência $(\delta(i))_{i \in \mathbb{N}}$ através do Teorema 5.4, do qual deduziremos (5.24), e então (5.23).

Lema 5.9. *Seja $K_i = \max_{j \leq i} \delta(j)j^{(i-\alpha)/2}$. Para qualquer n e i , as seguintes inequações são verdadeiras:*

$$|E[X_{n,i}^2] - 1| \leq \begin{cases} 0 & \text{se } i \leq n - n^\alpha, \\ 2\delta(i-1) & n - n^\alpha < i \leq n, \end{cases} \quad (5.25)$$

$$|s^2(\mathbf{X}_{n,i}) - i| \leq \begin{cases} 0 & \text{se } i \leq n - n^\alpha, \\ Ci^{(3\alpha-1)/2} K_i \leq Ci^\alpha & n - n^\alpha < i \leq n, \end{cases} \quad (5.26)$$

$$\|V^2(\mathbf{X}_{n,i}) - 1\|_p \leq \begin{cases} 0 & \text{se } i \leq n - n^\alpha, \\ Ci^{(\alpha-1)(1+1/2p)}(1 + K_i)^{1/p} + Ci^{(3\alpha-3)/2} K_i & \text{caso contrário.} \end{cases} \quad (5.27)$$

Demonstração. Na inequação (5.25) é óbvia para $i \leq n - n^\alpha$. Escrevamos,

$$I_{n,i}^+ = [\lambda_{n,i}, 2\lambda_{n,i}] \quad \text{e} \quad I_{n,i}^- = [-2\lambda_{n,i}, -\lambda_{n,i}]. \quad (5.28)$$

Para $n - n^\alpha < i \leq n$ da definição (5.2), temos que,

$$\begin{aligned} E[X_{n,i}^2] &= P(S_{n,i-1} \in I_{n,i}^-)E[X_{n,i}^2 | S_{n,i-1} \in I_{n,i}^-] + P(S_{n,i-1} \in I_{n,i}^+)E[X_{n,i}^2 | S_{n,i-1} \in I_{n,i}^+] \\ &\quad + P(S_{n,i-1} \notin I_{n,i}^- \cup I_{n,i}^+)E[X_{n,i}^2 | S_{n,i-1} \notin I_{n,i}^- \cup I_{n,i}^+] \\ &= \frac{1}{2}P(S_{n,i-1} \in I_{n,i}^-) + \frac{3}{2}P(S_{n,i-1} \in I_{n,i}^+) + P(S_{n,i-1} \in I_{n,i}^- \cup I_{n,i}^+) \\ &= \frac{1}{2}P(S_{n,i-1} \in I_{n,i}^-) + \frac{3}{2}P(S_{n,i-1} \in I_{n,i}^+) \\ &\quad + (1 - P(S_{n,i-1} \in I_{n,i}^-) - P(S_{n,i-1} \in I_{n,i}^+)) \\ &= 1 - \frac{1}{2}P(S_{n,i-1} \in I_{n,i}^-) + \frac{1}{2}P(S_{n,i-1} \in I_{n,i}^+) \end{aligned} \quad (5.29)$$

A variável aleatória $S_{n,i-1}/s(\mathbf{X}_{n,i-1})$ é aproximadamente Gaussiana, controlado pelo erro $\delta(i-1)$. Mais precisamente,

$$\left| P(S_{n,i-1} \in I_{n,i}^+) - \int_{I_{n,i}^+/s(\mathbf{X}_{n,i-1})} d\Phi \right| \leq 2\delta(i-1).$$

Temos que, pelo fato da simetria da distribuição Gaussiana,

$$\int_{I_{n,i}^+/s(\mathbf{X}_{n,i-1})} d\Phi = \int_{I_{n,i}^-/s(\mathbf{X}_{n,i-1})} d\Phi.$$

De (5.29), obtemos,

$$|E[X_{n,i}^2] - 1| = 1/2|P(S_{n,i-1} \in I_{n,i}^+) - P(S_{n,i-1} \in I_{n,i}^-)| \quad (5.30)$$

Vamos analisar o lado direito da equação (5.30),

$$\begin{aligned}
|P(S_{n,i-1} \in I_{n,i}^+) - P(S_{n,i-1} \in I_{n,i}^-)| &= |P(S_{n,i-1} \in I_{n,i}^+) - \int_{I_{n,i}^+/s(\mathbf{X}_{n,i-1})} d\Phi \\
&\quad - P(S_{n,i-1} \in I_{n,i}^-) + \int_{I_{n,i}^-/s(\mathbf{X}_{n,i-1})} d\Phi| \\
&\leq \left| P(S_{n,i-1} \in I_{n,i}^+) - \int_{I_{n,i}^+/s(\mathbf{X}_{n,i-1})} d\Phi \right| \\
&\quad + \left| P(S_{n,i-1} \in I_{n,i}^-) - \int_{I_{n,i}^-/s(\mathbf{X}_{n,i-1})} d\Phi \right| \\
&\leq 2\delta(i-1) + 2\delta(i-1) \\
&\leq 4\delta(i-1).
\end{aligned} \tag{5.31}$$

Desta forma com (5.30) e (5.31) obteremos (5.25).

Temos que, sabendo que para $i \leq n - n^\alpha$, obteremos $|E[X_{n,i}^2] - 1| = 0$,

$$\begin{aligned}
|s^2(\mathbf{X}_{n,i}) - i| &= \left| \sum_{j=1}^i (E[X_{n,j}^2] - 1) \right| \\
&\leq \sum_{j=1}^i |E[X_{n,j}^2] - 1| = \sum_{n-n^\alpha < j \leq i} |E[X^2]_{n,j} - 1| \\
&\leq 2 \sum_{n-n^\alpha < j \leq i} \delta(j-1).
\end{aligned} \tag{5.32}$$

Desta maneira teremos que,

$$|s^2(\mathbf{X}_{n,i}) - i| \leq \begin{cases} 0 & \text{se } i \leq n - n^\alpha, \\ 2 \sum_{n-n^\alpha < j \leq i} \delta(j-1) & \text{se } n - n^\alpha < i \leq n, \end{cases}$$

Relembrando que $\alpha < 1$, obtemos (5.26), notando que $n - n^\alpha < i \leq n$,

$$\sum_{n-n^\alpha < j \leq i} \delta(j-1) \leq n^\alpha (n - n^\alpha)^{(\alpha-1)/2} K_i.$$

Em particular, segue que,

$$s^2(\mathbf{X}_{n,i}) = i(1 + o(1)). \tag{5.33}$$

Em (5.27), $\|V^2(\mathbf{X}_{n,i}) - 1\|_p$ é claramente igual a 0 para $i \leq n - n^\alpha$. Vamos agora assumir o contrário. Teremos então

$$\begin{aligned}
\|V^2(\mathbf{X}_{n,i}) - 1\|_p &= s^{-2}(\mathbf{X}_{n,i}) \left\| \sum_{j=1}^i E[X_{n,j}^2 | \mathcal{F}_{n,j-1}] - s^2(\mathbf{X}_{n,i}) \right\|_p \\
&\leq \frac{1}{s^2(\mathbf{X}_{n,i})} \sum_{j=1}^i \|E[X_{n,j}^2 | \mathcal{F}_{n,j-1}] - 1\|_p + \frac{|s^2(\mathbf{X}_{n,i}) - i|}{s^2(\mathbf{X}_{n,i})} \\
&\leq \frac{1}{2s^2(\mathbf{X}_{n,i})} \sum_{n-n^\alpha < j \leq i} \left(P(S_{n,i-1} \in I_{n,i}^- \cup I_{n,i}^+) \right)^{1/p} + \frac{|s^2(\mathbf{X}_{n,i}) - i|}{s^2(\mathbf{X}_{n,i})}.
\end{aligned} \tag{5.34}$$

Consideraremos separadamente os termos de (5.34). Primeiro da definição de δ , teremos,

$$\left| P(S_{n,i-1} \in I_{n,i}^- \cup I_{n,i}^+) - \int_{I_{n,i}^+/s(\mathbf{X}_{n,i-1} \cup I_{n,i}^+)} d\Phi \right| \leq 2\delta(j-1).$$

A equação (5.33) implica que, uniformemente sobre $j > n - n^\alpha$,

$$\int_{I_{n,i}^+/s(\mathbf{X}_{n,i-1} \cup I_{n,i}^+)} d\Phi = (2\pi)^{-1/2} \frac{2\lambda_{n,j}}{s(\mathbf{X}_{n,j-1})} (1 + o(1)) \leq Cn^{(\alpha-1)/2},$$

e então o primeiro termo de (5.34) é limitado por

$$\begin{aligned}
&\frac{C}{i} \sum_{n-n^\alpha < j \leq i} \left(n^{(\alpha-1)/2} + 2\delta(j-1) \right)^{1/p} \\
&\leq \frac{C}{i} \sum_{n-n^\alpha < j \leq i} \left(n^{(\alpha-1)/2} + n^\alpha (n - n^\alpha)^{(\alpha-1)/2} K_i \right)^{1/p} \\
&\leq Ci^{(\alpha-1)(1+1/2p)} (1 + K_i)^{1/p}.
\end{aligned} \tag{5.35}$$

O segundo termo em (5.34) é controlado por (5.26), desta forma obtemos a inequação (5.27). \square

Demonstração da Proposição 5.8. Aplicaremos o Teorema 5.4 e com as informações do Lema 5.9, obteremos que $D(\mathbf{X}_{n,i})$ é limitado da seguinte maneira, lembrado que $s(\mathbf{X}_n) \simeq \sqrt{n}$ e a constante multiplicativa não depende de n e $i \leq n$,

$$\begin{aligned}
D(\mathbf{X}_{n,i}) &\leq C \left[\frac{i \log(i)}{i^{3/2}} + \left(\left(C i^{(\alpha-1)(1+\frac{1}{2p})} (1+K_i)^{\frac{1}{p}} + C i^{(3\alpha-3)/2} K_i \right)^p + i^{-p} \right)^{\frac{1}{2p+1}} \right], \\
&\leq C \left[\frac{\log(i)}{i^{1/2}} + \left(C i^{(\alpha-1)(1+\frac{1}{2p})} (1+K_i)^{\frac{1}{p}} + C i^{(3\alpha-3)/2} K_i \right)^{\frac{p}{2p+1}} + i^{\frac{-p}{2p+1}} \right], \\
&\leq C \left[\frac{\log(i)}{i^{1/2}} + C i^{(\alpha-1)/2} (1+K_i)^{1/(2p+1)} + C i^{\frac{-3(1-\alpha)p}{4p+2}} K_i^{p/(2p+1)} + i^{\frac{-p}{2p+1}} \right].
\end{aligned} \tag{5.36}$$

O primeiro termo da soma dentro dos colchetes pode ser descartado, porque é dominado por $i^{-p/(2p+1)}$. Note que para $p \geq 1$, temos,

$$\frac{3p(1-\alpha)}{4p+2} \geq \frac{1-\alpha}{2},$$

e como $\alpha > 1/2 > 1/(2p+1)$, também teremos,

$$\frac{p}{2p+1} \geq \frac{1-\alpha}{2}.$$

Multiplicando (5.36) por $i^{(1-\alpha)/2}$, obteremos,

$$K_i \leq C(1+K_i)^{1/(2p+1)} + C K_i^{p/(2p+1)}.$$

onde C é uma constante que não depende de i . Observando que o conjunto $\{x \geq 0 : x \leq C(1+x)^{1/(2p+1)} + C x^{p/(2p+1)}\}$ é limitado, temos então que K_i é uma sequência limitada, então (5.24) está provado.

A relação (5.23) então segue de (5.24) e (5.27). \square

Proposição 5.10. *Temos que*

$$\limsup_{i \rightarrow +\infty} i^{(1-\alpha)/2} \delta(i) > 0.$$

Demonstração. A demonstração desta proposição está em [4] página 642. \square

Finalizamos assim a prova. \square

Referências bibliográficas

- [1] BOLTHAUSEN, E. Exact convergence rates in some martingale central limit theorem. *The Annals of Probability* 10, 3 (1982), 672 –688.
- [2] BROWN, B.M. The martingale central limit theorem. *The Annals of Mathematical Statistics* 42, 1 (1971), 59–66.
- [3] MCLEISH, D.L. Dependent central limit theorems and invariance principles. *The Annals of Probability* 2, 4 (1974), 620–628.
- [4] MOURRAT, J-C. On the rate of convergence in the martingale central limit theorem. *Bernoulli* 2013 19, 2 (2013), 633–645.
- [5] CHUNG, K. *A Course in Probability Theory*, 2nd ed. Academic Press, 1974.
- [6] DURRETT, R. *Probability: Theory and Examples*, 4th ed. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2013.
- [7] HALL, P. ; HEYDE, C.C. *Martingale Limit Theory and its Application*. Academic Press, 1980.
- [8] JAMES, B. *Probabilidade: Um Curso em Nível Intermediário*, 3 ed . Coleção projeto Euclides. Instituto de Matemática Pura e Aplicada, 2013.
- [9] MAGALHÃES, M. *Probabilidade e Variáveis Aleatórias*, 2 ed. EdUSP, 2006.
- [10] WILLAMS, D. *Probability with Martingale*. Cambridge University Press, 1990.