UNIVERSIDADE ESTADUAL DA PARAÍBA
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA

RICARDO BATISTA MEDEIROS

VERSÕES DO TEOREMA CENTRAL DO LIMITE

CAMPINA GRANDE - PB
2019

RICARDO BATISTA MEDEIROS

VERSÕES DO TEOREMA CENTRAL DO LIMITE

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Orientador: Profa. Dra. Divanilda Maia Esteves

CAMPINA GRANDE - PB

2019

M488v Medeiros, Ricardo Batista.
Versões do Teorema Central do Limite [manuscrito] / Ricardo Batista Medeiros. - 2019.
60 p. : il. colorido.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia , 2019.

"Orientação : Profa. Dra. Divanilda Maia Esteves , Coordenação do Curso de Estatística - CCT."

1. Teorema Central do Limite. 2. Teoria assintótica. 3. Teoria de Probabilidade. I. Título

21. ed. CDD 519.5

Elaborada por Giulianne M. Pereira - CRB - 15/714 BC/UEPB

RICARDO BATISTA MEDEIROS

VERSÕES DO TEOREMA CENTRAL DO LIMITE

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Trabalho aprovado em 04 de Dezembro de 2019.

BANCA EXAMINADORA

DMEsteves

Profa. Dra. Divanilda Maia Esteves
Universidade Estadual da Paraíba

Prof. Dr. Sílvio Fernando Alves Xavier Júnior
Universidade Estadual da Paraíba

Prof. Dr. Kleber Napoleão Nunes de Oliveira Barros
Universidade Estadual da Paraíba

Agradecimentos

A Deus por ter me dado saúde e ouvido as minhas preces para que eu continuasse no caminho certo e fazendo as coisas da melhor forma possível.

Aos meus pais, Maria e Carlos, pelo empenho, esforço e dedicação para que eu estudasse em outra cidade.

Ao meu irmão, Rogério, por ter me inscrito nesse curso e me ajudado muito aqui em Campina Grande .

A minha orientadora, Divanilda Maia Esteves, por ter me ensinado da melhor forma possível e pela paciência e gentileza ao tirar as minhas dúvidas.

Aos meus professores que me ajudaram muito na minha caminhada até aqui.

A minha esposa, Luciana Vieira, por estar sempre ao meu lado me ajudando e me apoiando.

Aos meus colegas que me ajudaram e nunca me deixaram só.

Resumo

O presente trabalho tem o objetivo de apresentar e demonstrar algumas das principais versões do Teorema Central do Limite. O Teorema Central do Limite é um dos mais famosos e importantes teoremas na teoria de probabilidade. Segundo ele, sob certas condições, a média amostral tem distribuição assintótica normal. Esse resultado é extremamente importante, porque permite o uso da distribuição normal para aproximar a distribuição da média quando se trata de grandes amostras e isso favorece o uso de diversas técnicas estatísticas paramétricas e não paramétricas mesmo quando a amostra considerada não tem distribuição normal. Foram consideradas aqui a versão clássica desse teorema, atribuída a De Moivre-Laplace, a versão de Liapunov, de Lindeberg, de Hájek-Sidak, além da versão para vetores aleatórios.

Palavras-chaves: Teorema Central do Limite. Teoria assintótica. Teoria de Probabilidade.

Abstract

This work aims to present and demonstrate some of the major versions of the Central Limit Theorem. The Central Limit Theorem is one of the most famous and important theorems in probability theory. According to him, under certain conditions, the sample mean has a normal asymptotic distribution. This result is extremely important because it allows the use of the normal distribution to approximate the mean distribution when dealing with large samples and this favors the use of various parametric and nonparametric statistical techniques even when the sample considered has no normal distribution. We considered here the classical version of this theorem, attributed to De Moivre-Laplace, the Lindeberg version of Liapunov, the Hájek-Sidak, and the random vector version.

Key-words: Central Limit Theorem. Asymptotic Theory. Probability Theory.

Lista de ilustrações

Sumário

1 Introdução

Na Inferência Estatística há interesse em deduzir o comportamento de determinada característica populacional com base em uma amostra. Pode-se dizer que a inferência segue, em geral, dois caminhos: a estimação e os testes de hipótese. Frequentemente, para usar essas técnicas, faz-se necessário fazer suposições sobre a distribuição amostral (métodos paramétricos). Muitos testes e estimadores intervalares são construídos sob a condição de normalidade dos dados, no entanto tal condição nem sempre é observada. Daí provém um dos pontos que torna tão importante o teorema considerado neste trabalho.

O Teorema Central do Limite, também chamado de Teorema do Limite Central, é um dos teoremas mais importantes para a Estatística, pois esse teorema diz que, sob certas condições, a soma de n variáveis aleatórias padronizadas converge em distribuição para uma $N(0,1)$, independentemente da distribuição de onde provém as variáveis. Desta forma, o teorema garante que para amostras grandes pode-se usar os métodos estatísticos paramétricos mesmo quando a amostra observada não parece vir de uma variável normalmente distribuída.

Existem várias versões para o Teorema Central do Limite, mas neste trabalho, serão apresentadas apenas seis das principais versões desse teorema. Entre essas versões deste teorema, a mais geral é a de Lindeberg, que não exige que as variáveis aleatórias sejam identicamente distribuídas, além disso, com a condição do Lindeberg demonstra-se as versões clássica, de De Moivre- Laplace, de Liapunov e a de Hájek-Sidak. A diferença de cada uma dessas versões reside nas suposições feitas sobre a distribuição amostral.

Neste trabalho, inicialmente será feita uma revisão teórica de alguns tópicos de Matemática e Probabilidade que serão utilizados nas demonstrações, em seguida serão apresentadas as versões consideradas do Teorema Central do Limite e suas demonstrações.

2 Fundamentação Teórica

As demonstrações das versões do Teorema Central do Limite requerem conceitos matemáticos e probabilísticos não triviais. Boa parte destes temas são vistos nos cursos de Cálculo Diferencial e Integral e de Probabilidade, mas, em geral, de forma superficial eles serão apresentados neste capítulo para facilitar a leitura do próximo capítulo e pode ser suprimido, se o leitor considerar que tem suficiente familiaridade com os assuntos considerados.

2.1 Séries de Taylor e Maclaurin

Nesta seção será dada apenas uma breve introdução ao polinômio de Taylor. Para mais detalhes consulte, por exemplo, Thomas e Finney (1975) e Lima (2002).

Definição 2.1. *Seja f uma função com derivadas de todas as ordens em algum intervalo contendo x_0 como ponto interior. Então, a série de Taylor gerada por f em $x = x_0$ é*

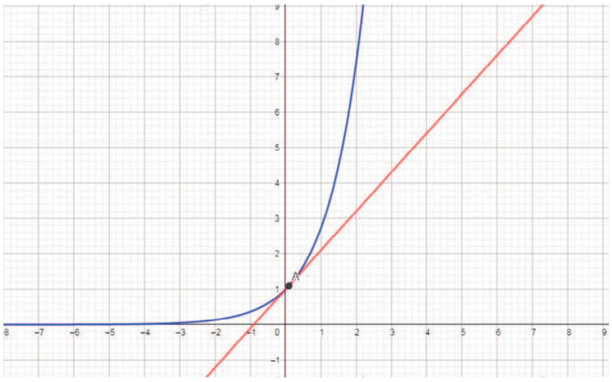
$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)(x-x_0)^k}{k!} = f(x_0) + \frac{f'(x_0)(x-x_0)^1}{1!} + \frac{f''(x_0)(x-x_0)^2}{2!} + \cdots + \frac{f^{(n)}(x_0)(x-x_0)^n}{n!} + \ldots. \tag{2.1}$$

Quando $x_0 = 0$, então, têm-se as séries de Maclaurin, que são um caso particular das séries de Taylor:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)(x)^k}{k!} = f(0) + \frac{f'(0)(x)^1}{1!} + \frac{f''(0)(x)^2}{2!} + \cdots + \frac{f^{(n)}(0)(x)^n}{n!} + ...$$

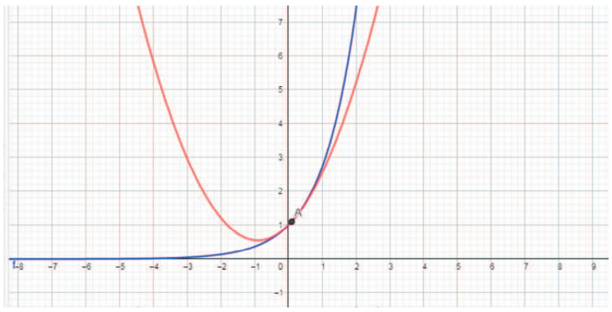
As séries de Taylor, também conhecidas como polinômio de Taylor, surgiram a partir dos trabalhos do matemático Brook Taylor em 1715. A ideia principal é obter estimativas cada vez mais próximas do verdadeiro valor de uma função $f(x)$ em um determinado ponto x_0 aumentando a ordem da derivada, ou seja, aumentando a quantidade de derivadas de uma função $f(x)$ melhora-se cada vez mais a estimativa para o verdadeiro valor de $f(x)$ em um determinado ponto x_0. Sendo assim, a estimativa feita por $f''(x)$ aplicada em um ponto x_0 é melhor do que com $f'(x)$ aplicada em x_0 e assim por diante. As Figuras 1, 2 e 3 ilustram graficamente a ideia do polinômio de Taylor.

Figura 1 – Gráfico de uma função exponencial, em azul, aproximada com o polinômio de Taylor de grau 1, em x=0.



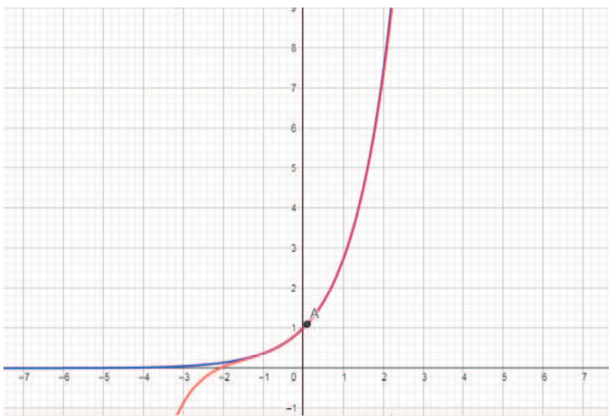
Fonte: Elaborada pelo autor, 2019.

Figura 2 – Gráfico de uma função exponencial, em azul, aproximada com o polinômio de Taylor de grau 2, em x=0.



Fonte: Elaborada pelo autor, 2019.

Figura 3 – Gráfico de uma função exponencial, em azul, aproximada com o polinômio de Taylor de grau 5, em x=0.



Fonte: Elaborada pelo autor, 2019.

A partir da série dada pela Equação (2.1) pode-se obter Séries de Taylor para determinadas funções, desde que as mesmas admitam derivadas de qualquer ordem. A seguir tal processo será exemplificado.

Exemplo 2.1. *Determine as Séries de Taylor para as funções* $f(x) = e^x$, $f(x) = \operatorname{sen} x$ *e* $f(x) = \cos x$ *todas em* $x_0 = 0$.

Demonstração. A função e^x é a única função cuja derivada é sempre ela mesma. Sendo assim,

$$f^{(n)}(x) = e^x \Rightarrow f^{(n)}(0) = 1, \forall n \in \{0, 1, 2, \dots\},$$

logo, aplicando este resultado em (2.1), segue que

$$e^x \ = \ 1 \ + \ \frac{x^1}{1!} \ + \ \frac{x^2}{2!} \ + \ \frac{x^3}{3!} \ + \ \frac{x^4}{4!} \ + \ \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \tag{2.2}$$

Agora, para a $f(x) = \text{sen}(x)$ em $x_0 = 0$, daí tem-se que

$$\begin{aligned}
f(x) = \operatorname{sen} x \quad &\Rightarrow \quad f(0) = 0 \\
f'(x) = \cos x \quad &\Rightarrow \quad f'(0) = 1 \\
f''(x) = -\operatorname{sen} x \quad &\Rightarrow \quad f''(0) = 0 \\
f'''(x) = -\cos x \quad &\Rightarrow \quad f'''(0) = -1 \\
f^{(4)}(x) = \operatorname{sen} x \quad &\Rightarrow \quad f^{(4)}(0) = 0 \\
f^{(5)}(x) = \cos x \quad &\Rightarrow \quad f^{(5)}(0) = 1 \\
f^{(6)}(x) = -\operatorname{sen} x \quad &\Rightarrow \quad f^{(6)}(0) = 0 \\
f^{(7)}(x) = -\cos x \quad &\Rightarrow \quad f^{(7)}(0) = -1
\end{aligned}$$

A partir dessas sete derivadas, percebe-se que as derivadas de ordem par, aplicadas em $x_0 = 0$, são nulas, e as ímpares são não nulas e seguem alternando o sinal em positivo e negativo, nessa ordem, então, substituindo essas informações em (2.1), obtem-se que a Série de Taylor da função $\text{sen}(x)$ é dada por

$$\text{sen}(x) \ = \ \frac{x^1}{1!} - \frac{x^3}{3!} \ + \ \frac{x^5}{5!} \ - \ \frac{x^7}{7!} \ + \ \cdots = \sum_{i=1}^{\infty} (-1)^{i+1} \frac{x^{2n-1}}{(2n-1)!}.$$

Por fim, para a função $f(x) = \cos(x)$ em $x_0 = 0$, a demonstração é feita de forma similar a da $f(x) = \operatorname{sen} x$ em $x_0 = 0$. As sete primeiras derivadas para a $f(x) = \cos x$ aplicadas em $x_0 = 0$ são dadas por

$$\begin{aligned}
f(x) = \cos(x) \quad &\Rightarrow \quad f(0) = 1 \\
f'(x) = -\operatorname{sen}(x) \quad &\Rightarrow \quad f'(0) = 0 \\
f''(x) = -\cos(x) \quad &\Rightarrow \quad f''(0) = -1 \\
f'''(x) = \operatorname{sen}(x) \quad &\Rightarrow \quad f'''(0) = 0 \\
f^{(4)}(x) = \cos(x) \quad &\Rightarrow \quad f^{(4)}(0) = 1 \\
f^{(5)}(x) = -\operatorname{sen}(x) \quad &\Rightarrow \quad f^{(5)}(0) = 0 \\
f^{(6)}(x) = -\cos(x) \quad &\Rightarrow \quad f^{(6)}(0) = -1 \\
f^{(7)}(x) = \operatorname{sen}(x) \quad &\Rightarrow \quad f^{(7)}(0) = 0
\end{aligned}$$

Neste caso, são as derivadas de ordem ímpar que são nulas, em $x_0 = 0$, e as pares são não nulas e os sinais se alternam em positivo e negativo, nessa ordem. Substituindo essas derivadas em (2.1), obtem-se que a série de Taylor do $\cos(x)$ é dada por

$$\cos(x) \;=\; 1 \;-\; \frac{x^2}{2!} \;+\; \frac{x^4}{4!} \;-\; \frac{x^6}{6!} \;+\; \cdots = \sum_{i=0}^{\infty}(-1)^i\frac{x^{2i}}{(2i)!}.$$

■

2.2 Variáveis Aleatórias

O conjunto de todos os resultados possíveis de um experimento aleatório é chamado de **espaço amostral**, que geralmente é denotado por Ω. Por exemplo, ao realizar o experimento que consiste em lançar duas moedas quaisquer e observar as faces que ficaram voltadas para cima, o espaço amostral é $\Omega = \{(c,c),(c,k),(k,c),(k,k)\}$, onde c e k representam, respectivamente, cara e coroa. Por outro lado, se o experimento consiste em lançar um dado com seis faces numeradas de 1 a 6 e observar o número da face voltada para cima, o espaço amostral é $\Omega = \{1,2,3,4,5,6\}$.

No exemplo da moeda, o interesse pode estar na quantidade de coroas que ocorreram ao lançar as moedas. Neste caso, tal quantidade poderia ser facilmente obtida a partir dos resultados do experimento. Entretanto, em experimentos com um espaço amostral maior, isso começa a ficar mais difícil.

Associar números aos resultados possíveis de um determinado experimento aleatório é uma ideia muito interessante e prática, pois, permite tratá-las matematicamente de uma maneira mais simples.

Definição 2.2. *Uma* ***variável aleatória*** *é uma função real definida no espaço amostral Ω de um experimento aleatório. Ou seja, uma variável aleatória é uma função que assume valores reais e que associa a cada resultado do espaço amostral do experimento aleatório um número real.*

Se o conjunto de valores que uma variável aleatória pode assumir for finito ou infinito enumerável ela é chamada de **variável aleatória discreta**. Se o conjunto de valores que uma variável aleatória pode assumir for não enumerável, ou seja, um intervalo de valores reais, então, ela é chamada de variável aleatória contínua. Existem ainda as variáveis mistas (parte discreta, parte contínua) e as singulares (JAMES, 2002; MAGALHÃES, 2006). Daqui em diante serão consideradas apenas as variáveis discretas e as contínuas.

Definição 2.3. *Seja X uma variável aleatória discreta. A* ***função de distribuição de probabilidades*** *de X é a função $P(X = x_i)$ que associa, a cada valor possível x_i, com*

$i = 1, 2, \cdots$ *de* X*, sua respectiva probabilidade*

$$p(x_i) = P(X = x_i) = \sum_{\{\omega \in \Omega : X(\omega) = x_i\}} P(\omega).$$

Por construção, tal função deve ser tal que

$$0 \leq p(x_i) \leq 1 \quad \text{e} \sum_{i=1}^{\infty} P(X = x_i) = 1.$$

Definição 2.4. *Seja* X *uma variável aleatória contínua, assumindo valores* $x \in S_x$*, onde* S_x *denota um intervalo real. A* ***função de densidade de probabilidade*** *é uma função* $f(x)$ *que satisfaz as seguintes propriedades:*

$$f(x) \geq 0 \quad e \quad \int_{x \in S_x} f(x)dx = 1.$$

Além disso, para um intervalo $(a; b] \subset \mathbb{R}$ *tem-se*

$$P(a < X \leq b) = \int_a^b f(x)dx.$$

Definição 2.5. *Dada uma variável aleatória* X *qualquer, a* ***função de distribuição acumulada*** *de* X*, denotada por* $F_X(x)$*, é definida por*

$$F_X(x) \; = \; P(X \leq x), \quad \forall \, x \in \mathbb{R}.$$

Em particular, se X *é uma variável aleatória contínua, então*

$$F_X(x) = \int_{-\infty}^{x} f(t)dt. \tag{2.3}$$

Observe que a função de distribuição acumulada é definida não apenas sobre o conjunto dos valores que a variável assume, mas para todo valor real. Além disso, sendo uma probabilidade, deve ser tal que

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

Além do mais, como consequência da Equação (2.3), segue que, no caso específico das variáveis contínuas,

$$\frac{dF_X(x)}{dx} = f(x).$$

A função de distribuição acumulada de uma variável aleatória X satisfaz as seguintes propriedades:

(i) $\lim_{x \to -\infty} F_X(x) = 0$ e $\lim_{x \to \infty} F_X(x) = 1$;

(ii) Se $a < b$ são dois números reais, então $F(a) \leq F(b)$;

(iii) Se a < b, então $P(a < X \leq b) = F(b) - F(a)$;

(iv) $\lim\limits_{x \to a^+} F_X(x) = F(a)$.

As demonstrações de tais propriedades foram omitidas aqui, mas podem ser encontradas em James (2002).

Na sequência, serão apresentadas duas quantidades importantes associadas às variáveis aleatórias, bem como algumas de suas propriedades. As demonstrações das propriedades podem ser encontradas em diversos livros clássicos de probabilidade, tais como James (2002), Magalhães (2006), Dantas (2013), Ross (2014). A primeira quantidade é a esperança, que também pode ser chamada esperança matemática, média, valor esperado ou expectância, e a segunda é a variância.

A esperança de uma variável aleatória é uma medida de tendência central dessa variável, podendo ser interpretada como sendo o centro de gravidade da distribuição (DANTAS, 2013).

Definição 2.6. *Seja X uma variável aleatória discreta que assume os valores x_i e probabilidades $P(x_i)$, com $i = 1, 2, \cdots$. A **esperança matemática** de X é definida como*

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

desde que a série seja convergente.

Definição 2.7. *Seja X uma variável aleatória contínua, com função de densidade de probabilidade, denotada por $f(x)$, então, a sua **esperança matemática** é definida como*

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx,$$

desde que a integral esteja bem definida.

Tanto para o caso discreto quanto para o contínuo, valem propriedades do valor esperado que serão vistas a seguir. Sejam a e b constantes reais diferentes de zero e

1. $\mathbb{E}(a) = a$;
2. $\mathbb{E}(X + a) = \mathbb{E}(X) + a$;
3. $\mathbb{E}(bX) = b\mathbb{E}(X)$;
4. $\mathbb{E}(a + bX) = a + b\mathbb{E}(X)$;
5. $\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i)$;

6. Se as variáveis $X_1, X_2, \cdots, X_n$ são independentes $\mathbb{E}\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} \mathbb{E}(X_i)$.

O resultado a seguir trata do cálculo de esperança de uma função de uma variável aleatória.

Lema 2.1. *Seja X uma variável aleatória, onde g é uma função real e $Y = g(X)$. Então, se X é discreta,*

$$\mathbb{E}(Y) = \sum_i g(x_i)P(X = x_i),$$

e se X é contínua

$$\mathbb{E}(Y) = \int_{S_x} g(x)f(x)dx.$$

É possível também calcular a esperança usando a função de distribuição acumulada, como pode ser visto no resultado abaixo.

Lema 2.2. *Para uma variável aleatória não negativa Y,*

$$\mathbb{E}(Y) = \int_0^{+\infty} P(Y > y)dy = \int_0^{+\infty} (1 - F(y))dy.$$

A variância é uma medida de dispersão que mostra a variabilidade dos dados em torno da média. A definição de variância e as suas propriedades são dadas a seguir. Assim como no caso da esperança, as demonstrações serão omitidas, mas podem ser encontradas em diversos livros, como por exemplo, James (2002), Magalhães (2006), Dantas (2013), Ross (2010), Ross (2014).

Definição 2.8. *A **variância** de uma variável aleatória X qualquer é definida como*

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}(X))^2]. \tag{2.4}$$

*Define-se a partir daí, o **desvio-padrão** de uma variável aleatória X como*

$$\mathrm{dp}(X) = \sqrt{\mathrm{Var}(X)}.$$

Usando propriedades de esperança vistas anteriormente, mostra-se que a variância pode ser calculada como

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2.$$

Seja X uma variável aleatória e a e b constantes reais diferentes de zero. Neste caso, a variância satisfaz as seguintes propriedades.

1. $\mathrm{Var}(X) \geq 0$;

2. $\mathrm{Var}(a) = 0$;

3. $\mathrm{Var}(X + a) = \mathrm{Var}(X)$;

4. $\mathrm{Var}(bX) = b^2 \cdot \mathrm{Var}(X)$;

5. $\mathrm{Var}(a + bX) = b^2 \cdot \mathrm{Var}(X)$.

Foi visto que é possível calcular esperança de uma função de uma variável aleatória. Um caso particularmente importante é aquele em que se considera uma potência da variável aleatória.

Definição 2.9. *O momento de ordem k de uma variável aleatória X é definido como $E(X^k)$ e pode ser calculado da seguinte maneira, para $k \in \{1, 2, \dots\}$*

$$\mathbb{E}(X^k) = \begin{cases} \displaystyle\sum_{i=1}^{n} x_i^k P(X = x_i), & \textit{para o caso discreto,} \\ \\ \displaystyle\int_{-\infty}^{+\infty} x^k f(x)dx, & \textit{para o caso contínuo.} \end{cases}$$

2.3 Modelos Probabilísticos

Muitos experimentos aleatórios apresentam certas características que se repetem em outros experimentos aleatórios. Estas características definem famílias de distribuições de probabilidade/funções de densidade indexadas por um ou mais parâmetros, as quais são chamadas de modelos probabilísticos. Há diversos modelos discretos e contínuos, no entanto serão apresentados apenas três deles, por serem os modelos que serão citados/usados neste trabalho.

2.3.1 A distribuição de Bernoulli e a distribuição Binomial

A distribuição de Bernoulli é um modelo discreto que é utilizado em experimentos onde só ocorrem dois resultados possíveis, um sucesso ou um fracasso. Esse tipo de experimento é conhecido como ensaio de Bernoulli.

Se X é uma variável aleatória com distribuição Bernoulli, denotado por $X \sim Bernoulli(p)$, sua distribuição de probabilidades é dada por:

$$P(X = x) = p^x(1 - p)^{1-x}, \quad \text{com } x = 0, 1,$$

em que p é a probabilidade de obter sucesso e $(1 - p)$ é a probabilidade de obter fracasso.

A partir da distribuição de probabilidades é possível mostrar que $\mathbb{E}(X) = p$ e $\mathrm{Var}(X) = p \cdot (1 - p)$.

A distribuição Binomial é obtida quando se conta o número de sucessos observados em n ensaios (ou tentativas) independentes de Bernoulli, ou seja, define-se uma variável aleatória X que conta o números de sucessos nos n ensaios independentes de Bernoulli.

Se uma variável aleatória X tem distribuição Binomial com parâmetros n e p, denotado por $X \sim Bin(n, p)$, com $0 < p < 1$ e $n \in \{1, 2, \ldots\}$, sua distribuição de probabilidade é dada por

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

Usando a distribuição de probabilidade e binômios de Newton, mostra-se que, neste caso,

$$\mathbb{E}(X) = np \quad \text{e} \quad \mathrm{Var(X)} = np(1-p).$$

2.3.2 A distribuição Normal

A distribuição Normal é um dos principais modelos probabilísticos, por sua ampla aplicabilidade em estatística. Muitas técnicas estatísticas são desenvolvidas sob a suposição de normalidade.

Uma variável aleatória X tem distribuição normal com parâmetros $\mu \in \mathbb{R}$ e $\sigma^2 > 0$, denotado por $X \sim N(\mu, \sigma^2)$ quando sua função densidade de probabilidade é dada por

$$f(x) = \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}.$$

Quando $\mu = 0$ e $\sigma^2 = 1$, então se diz que a variável tem distribuição **normal padrão** ou normal reduzida. A Transformação de uma normal com média μ e variância σ^2 em uma normal com média 0 e variância 1 é chamada de padronização. Usando técnicas de probabilidade, mostra-se que se $X \sim N(\mu, \sigma^2)$, então

$$\frac{X-\mu}{\sigma} \sim N(0, 1).$$

A função de densidade $f_X(x)$ da distribuição normal não possui uma primitiva, o que implica que valores de probabilidades acumuladas - $F(x)$ - só podem ser calculados computacionalmente. Com a padronização, basta calcular tais valores para a distribuição normal padrão e, a partir daí, é possível calcular probabilidades associadas a variáveis normalmente distribuídas com quaisquer parâmetros.

Se $X \sim N(\mu, \sigma^2)$, então $\mathbb{E}(X) = \mu$ e $\mathrm{Var}(X) = \sigma^2$. As demonstrações de tais resultados são não triviais e requerem várias técnicas de cálculo diferencial e integral.

2.4 Função Geradora de Momentos

A função geradora de momentos, ou função geratriz de momentos, como o próprio nome sugere, possibilita o cálculo dos momentos de uma variável aleatória contínua ou discreta. Essa função também é útil para identificar a distribuição de uma variável aleatória, pois ela identifica de maneira única um modelo probabilístico.

Definição 2.10. *Se $M_X(t)$ é a função geradora de momentos (f.g.m.) de uma variável aleatória X, então*

$$M_X(t) = \mathbb{E}\left[e^{tX}\right], \quad \forall\, t \in \mathbb{R}, \tag{2.5}$$

desde que $\mathbb{E}\left[e^{tX}\right]$ seja finita.

Se X for uma variável discreta,

$$M_X(t) = \sum_i e^{tx_i} P(X = x_i),$$

e se for contínua

$$M_X(t) = \int_{-\infty}^{+\infty} e^{tx} f(x) dx.$$

A seguir veremos alguns resultados importantes sobre a fgm.

Teorema 2.1. *Seja X uma variável aleatória. Como $\mathbb{E}(e^{tX})$ está definido para $t \in (-h, h)$ para algum $h \in \mathbb{R}$, então:*

$$M_X^{(n)}(0) \;=\; \mathbb{E}(X^n).$$

Demonstração. Para esta demonstração utilize a série de Taylor da função exponencial (Equação 2.2). Fazendo $X = tX$, obtem-se

$$e^{tX} \;=\; 1 \;+\; \frac{t^1X^1}{1!} \;+\; \frac{t^2X^2}{2!} \;+\; \frac{t^3X^3}{3!} \;+\; \frac{t^4X^4}{4!} \;+\; \cdots \;+\; \frac{t^nX^n}{n!} + \ldots$$

Como esta série está em função de X, aplicando a esperança em ambos os lados da igualdade e supondo que a soma do lado direito da igualdade seja convergente, segue que:

$$\mathbb{E}(e^{tX}) \;=\; 1 \;+\; t\cdot\mathbb{E}(X) \;+\; \frac{t^2}{2!}\cdot\mathbb{E}(X^2) \;+\; \frac{t^3}{3!}\cdot\mathbb{E}(X^3) \;+\; \frac{t^4}{4!}\cdot\mathbb{E}(X^4) \;+\; \cdots \;+\; \frac{t^n}{n!}\cdot\mathbb{E}(X^n)+\ldots,$$

ou seja

$$M_X(t) \;=\; 1 \;+\; t\cdot\mathbb{E}(X) \;+\; \frac{t^2}{2!}\cdot\mathbb{E}(X^2) \;+\; \frac{t^3}{3!}\cdot\mathbb{E}(X^3) \;+\; \frac{t^4}{4!}\cdot\mathbb{E}(X^4) \;+\; \cdots \;+\; \frac{t^n}{n!}\cdot\mathbb{E}(X^n)+\ldots$$

Perceba que a f.g.m. é função de t, então, derivando em relação a t,

$$M'_X(t) \;=\; \mathbb{E}(X) \;+\; t\cdot\mathbb{E}(X^2) \;+\; \frac{t^2}{2!}\cdot\mathbb{E}(X^3) \;+\; \frac{t^3}{6}\cdot\mathbb{E}(X^4) \;+\; \cdots \;+\; \frac{n\cdot t^{n-1}}{n!}\cdot\mathbb{E}(X^n)+\ldots$$

Consequentemente,

$$M'_X(0) = \mathbb{E}(X).$$

Derivando novamente a f.g.m em relação a t,

$$M''_X(t) = \mathbb{E}(X^2) \;+\; t\cdot\mathbb{E}(X^3) \;+\; \frac{t^2}{2}\cdot\mathbb{E}(X^4) \;+\; \cdots \;+\; \frac{n\cdot(n-1)\cdot t^{n-2}}{n!}\cdot\mathbb{E}(X^n)+\ldots$$

de onde se conclui que

$$M''_X(0) = \mathbb{E}(X^2).$$

Recursivamente, chega-se à conclusão que

$$M_X^{(n)}(0) = \mathbb{E}(X^n).$$

■

Conforme o teorema (2.1), percebe-se que ao derivar a função geradora de momentos e aplicar nessa derivada o ponto $t = 0$ obtem-se os momentos referentes à ordem da derivada. É claro que tal resultado está condicionado ao fato de a derivada da função no ponto $t = 0$ esteja bem definida.

Teorema 2.2. *Sejam $X_1, X_2, \ldots, X_n$ v.a's independentes, com f.g.m. $M_{X_i}(t), i = 1, \ldots, n$. Neste caso, a função geradora de momentos de $Z = X_1 + X_2 + \cdots + X_n$ é*

$$M_Z(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot \cdots \cdot M_{X_n}(t).$$

Demonstração. Definindo $Z = X_1 + X_2 + \cdots + X_n$, segue que

$$\begin{aligned} M_Z(t) &= \mathbb{E}[e^{tZ}] = \mathbb{E}\left[e^{t(X_1+X_2+\cdots+X_n)}\right] \\ &= \mathbb{E}\left[e^{tX_1+tX_2+\cdots+tX_n}\right] \\ &= \mathbb{E}\left[e^{tX_1} \cdot e^{tX_2} \ldots e^{tX_n}\right]. \end{aligned}$$

Por hipótese, as variáveis aleatórias $X_1, X_2, \ldots, X_n$ são independentes, então

$$\begin{aligned} M_Z(t) &= \mathbb{E}(e^{tX_1}) \cdot \mathbb{E}(e^{tX_2}) \cdot \cdots \cdot \mathbb{E}(e^{tX_n}) \\ &= M_{X_1}(t) \cdot M_{X_2}(t) \cdot \cdots \cdot M_{X_n}(t). \end{aligned}$$

■

O Teorema (2.2) é muito útil para obter propriedades aditivas para alguns modelos discretos ou contínuos.

Teorema 2.3. *Seja X uma variável aleatória cuja função geradora de momentos é $M_X(t)$. Seja $Y = aX + b$, onde a e b são constantes reais, e por fim, seja $M_Y(t)$ a função geradora de momentos de Y. Então, para todo valor de t tal que $M_X(at)$ existe, a fgm de y é dada por*

$$M_Y(t) = e^{bt} M_X(at). \tag{2.6}$$

Demonstração. Dado que $Y = aX + b$,

$$\begin{aligned} M_Y(t) &= \mathbb{E}\left[e^{tY}\right] = \mathbb{E}\left[e^{t(aX+b)}\right] \\ &= \mathbb{E}\left[e^{(taX+tb)}\right] = \mathbb{E}\left[e^{taX} \cdot e^{tb}\right]. \end{aligned}$$

Como e^{tb} é uma constante, segue que

$$M_Y(t) = e^{tb} \cdot \mathbb{E}\left(e^{(at)X}\right) = e^{tb} \cdot M_X(at).$$

■

A seguir, serão obtidas as funções geradoras de momentos associadas aos modelos probabilísticos vistos anteriormente - Bernoulli, Binomial e Normal.

Exemplo 2.2. *Seja X uma variável aleatória com distribuição Bernoulli(p). Então, a f.g.m. de X é*

$$M_X(t) = \left[pe^t + 1 - p\right].$$

Demonstração. Inicialmente, lembre que, se $X \sim Bernoulli(p)$, então

$$P(X=0) = 1-p, \quad P(X=1) = p \quad \text{e} \quad P(X=x) = 0, \text{para } x \neq 0, 1.$$

Desta maneira,

$$\begin{aligned} M_X(t) &= e^{t\cdot 0} \cdot (1-p) + e^t \cdot 1 \cdot p \\ &= (1-p) + e^t \cdot p, \end{aligned}$$

em que

$$M_X(t) = 1 - p + pe^t, \quad t \in \mathbb{R}.$$

A partir daí, aplicando o Teorema (2.1), pode-se encontrar os momentos da Bernoulli. A primeira derivada da f.g.m. é

$$M'_X(t) = pe^t$$

e portanto

$$\mathbb{E}(X) = M'_X(0) = pe^0 = p.$$

Para obter o segundo momento, deriva-se mais uma vez, obtendo

$$M''_X(t) = pe^t,$$

consequentemente

$$\mathbb{E}(X^2) = M''_X(0) = pe^0 = p.$$

Assim, calcula-se a variância como

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = p - p^2 = p(1-p).$$

■

Exemplo 2.3. *Seja X uma variável aleatória com distribuição $Bin(n,p)$, então, a fgm de X é $M_X(t) = [pe^t + 1 - p]^n$.*

Demonstração. Sabendo que, se $X \sim Bin(n,p)$, então

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x \in \{0, 1, \ldots, n\},$$

segue que

$$\begin{aligned} M_X(t) &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x}. \end{aligned}$$

Aplicando a Fórmula do Binômio de Newton, chega-se a

$$M_X(t) = (1 - p + pe^t)^n, \quad t \in \mathbb{R}.$$

Analogamente ao que foi feito no caso da distribuição Bernoulli, a seguir serão calculados os dois primeiros momentos da binomial e a variância. A primeira derivada a função geradora de momentos é

$$M'_X(t) = npe^t(1 - p + pe^t)^{n-1}$$

e aplicando em $t = 0$, conclui-se que

$$\mathbb{E}(X) = M'_X(0) = np.$$

A segunda derivada da f.g.m. é

$$M''_X(t) = npe^t(1 - p + pe^t)^{n-1} + np^2e^{2t}(n-1)(1 - p + pe^t)^{n-2}$$

e aplicando no ponto $t = 0$ encontra-se

$$\mathbb{E}(X^2) = M''_X(0) = np + n^2p^2 - np^2.$$

Conhecendo os dois primeiros momentos é possível calcular a variância da variável

$$\begin{aligned} \mathrm{Var}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \\ &= np + n^2p^2 - np^2 - n^2p^2 \\ &= np - np^2, \end{aligned}$$

ou seja,

$$\mathrm{Var}(X) = np(1-p).$$

$\blacksquare$

Exemplo 2.4. *Seja X uma variável aleatória com distribuição $N(0,1)$, então, a função geradora de momentos de X é $M_X(t) = e^{t^2/2}$.*

Demonstração. Se $X \sim N(0,1)$, então

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

Consequentemente,

$$\begin{aligned} M_X(t) = E(e^{tx}) &= \int_{-\infty}^{+\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}+tx} dx. \end{aligned}$$

Observe que

$$\left(-\frac{x^2}{2} + tx\right) = -\frac{1}{2}\left(x^2 - 2tx\right) = -\frac{1}{2}\left((x-t)^2 - (t)^2\right) = -\frac{1}{2}(x-t)^2 + \frac{t^2}{2},$$

sendo assim,

$$M_X(t) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2 + \frac{t^2}{2}} dx = e^{\frac{t^2}{2}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-t)^2}{2}} dx.$$

No entanto, fazendo a mudança de variável $z = x - t$ na integral, obtem-se

$$M_X(t) = e^{\frac{t^2}{2}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.$$

Assim, chega-se a uma integral cujo integrando é a função densidade de probabilidade de uma variável aleatória com distribuição $N(0,1)$ e, portanto, quando integrada em todo seu domínio o resultado da integral é igual a 1. Desta forma,

$$M_X(t) = e^{\frac{t^2}{2}}, \quad t \in \mathbb{R}.$$

Agora que se conhece a f.g.m. da $N(0,1)$, serão determinados seus dois primeiros momentos e a variância

$$\begin{aligned} M_X'(t) = te^{\frac{t^2}{2}} &\quad \Rightarrow \quad \mathbb{E}(X) = M_X'(0) = 0 \\ M_X''(t) = e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}} &\quad \Rightarrow \quad \mathbb{E}(X^2) = M_X''(0) = 1, \end{aligned}$$

e daí

$$\operatorname{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 1.$$

■

Exemplo 2.5. *Seja Y uma variável aleatória com distribuição $N(\mu, \sigma^2)$, então, a f.g.m. de Y é $M_Y(t) = \exp\left\{t\mu + \frac{t^2\sigma^2}{2}\right\}$.*

Demonstração. Se $Y \sim N(\mu, \sigma^2)$, então

$$X = \frac{Y - \mu}{\sigma} \sim N(0,1).$$

Assim, pelo Teorema (2.3) e observando que $Y = \sigma X + \mu$,

$$M_Y(t) = e^{t\mu} \cdot M_X(t\sigma) = e^{t\mu} \cdot e^{\frac{t^2\sigma^2}{2}},$$

de onde segue que

$$M_Y(t) = \exp\left\{t\mu + \frac{t^2\sigma^2}{2}\right\}, \quad t \in \mathbb{R}.$$

Da mesma forma que foi feito nos exemplos anteriores, é possível usar a f.g.m. para calcular os dois primeiros momentos e, a partir daí explicitar a esperança e a variância da distribuição.

■

A função geradora de momentos é bastante útil para encontrar os momentos de uma variável aleatória, especialmente naqueles casos em que encontrá-los pela definição envolve longos cálculos. A desvantagem dessa função é que a esperança que a define pode não ser finita. Algumas vezes, a f.g.m. não existe (a esperança não é finita) e, em outros casos, ela não está bem definida para $t = 0$. Uma alternativa é o uso de funções características, que serão vistas na seção a seguir.

2.5 Função Característica

Vimos que a função geradora de momentos facilita, muitas vezes, a obtenção dos momentos de uma determinada distribuição, mas o inconveniente dela é que sua esperança pode não ser finita dependendo da distribuição. Desta forma, uma alternativa muito interessante para escapar desse problema é a função característica. A função característica é uma função definida para todo número real t e que toma valores no conjunto dos complexos. Assim como a função geradora de momentos, através da função característica também pode-se obter os momentos de uma determinada distribuição, mas a grande vantagem da função característica sobre a f.g.m. é que ela sempre está definida para qualquer distribuição de probabilidade, o que não acontece com a função geradora de momentos.

Definição 2.11. *Seja X uma variável aleatória. Então a função característica de X é uma função $\mathbb{R} \to \mathbb{C}$ definida por:*

$$\varphi(t) = \varphi_X(t) = \mathbb{E}(e^{itX}) = \int_{\mathbb{R}} e^{(itx)} dF_X(x), \tag{2.7}$$

para $t \in \mathbb{R}$ e $i = \sqrt{-1}$.

Usando a Fórmula de Euler (Apêndice B - Equação B.2) é possível perceber o motivo da função característica sempre estar definida para qualquer distribuição. Isso acontece devido ao fato de as funções seno e cosseno serem funções limitadas.

Antes de estudar as propriedades das funções características serão apresentados alguns teoremas e corolários que estão presentes em algumas demonstrações. Esses resultados podem ser encontrados mais detalhadamente em livros de probabilidade como James (2002), Magalhães (2006) e Ash (2008).

Teorema 2.4 (**Teorema da convergência dominada**). *Seja* $\{X_n\}_{n\geq 1}$ *uma sequência de variáveis aleatórias tal que* $X_n \to X$ *quase certamente e existe uma variável aleatória* Y *não negativa tal que* $|X_n| \leq Y$ *com* $\mathbb{E}[Y] < \infty$, *então para* $\mathbb{E}[X] < \infty$,

$$\mathbb{E}[X_n] \to \mathbb{E}[X], \quad \textit{quando} \;\; n \to \infty.$$

O teorema a seguir (**Fórmula da Inversão**) permite encontrar a função de distribuição/densidade de um modelo probabilístico qualquer, ainda que nem sempre seja uma tarefa simples.

Teorema 2.5 (**Fórmula da Inversão**). *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:*

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

sendo a, b *e* c *números reais tais que* $c > 0$ *e* $a < b$.

Demonstração. Para esta demonstração considere

$$H = \frac{1}{2\pi}\int_{-c}^{c} \frac{1}{it}\left[e^{-iat} - e^{-ibt}\right]\varphi_X(t)dt.$$

Aplicando a definição de função característica, obtem-se que

$$H = \frac{1}{2\pi}\int_{-c}^{c} \frac{1}{it}\left[e^{-iat} - e^{-ibt}\right]\int_{-\infty}^{\infty} e^{itx} dF_X(x)\; dt.$$

Mudando a ordem de integração e simplificando as exponenciais, segue que

$$\begin{aligned}
H &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-c}^{c} \frac{1}{it}\left[e^{itx-iat} - e^{itx-ibt}\right] dt\; dF_X(x) \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-c}^{c} \frac{1}{it}\left[e^{it(x-a)} - e^{it(x-b)}\right] dt\; dF_X(x) \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-c}^{c} \frac{1}{it}\left[\cos(tk_1) + i\,\mathrm{sen}(tk_1) - \cos(tk_2) - i\,\mathrm{sen}(tk_2)\right] dt\; dF_X(x),
\end{aligned}$$

onde $k_1 = (x - a)$ e $k_2 = (x - b)$, sendo que a última igualdade foi obtida aplicando a fórmula de Euler.

Mas, note que

$$\begin{aligned}
&\int_{-c}^{c} \frac{1}{it}\left[\cos(tk_1) + i\,\mathrm{sen}(tk_1) - \cos(tk_2) - i\,\mathrm{sen}(tk_2)\right] dt \\
&\quad = \int_{-c}^{c} \frac{\cos(tk_1)}{it}dt + \int_{-c}^{c} \frac{\mathrm{sen}(tk_1)}{t}dt - \int_{-c}^{c} \frac{\cos(tk_2)}{it}dt - \int_{-c}^{c} \frac{\mathrm{sen}(tk_2)}{t}dt,
\end{aligned}$$

sendo que aquelas da forma $\frac{\cos(kt)}{t}$ são ambas iguais à zero, pois são funções ímpares. As outras duas integrais, que são da forma $\frac{\text{sen}(kt)}{t}$ são funções pares. Fazendo as simplificações,

$$\begin{aligned} H &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[2\int_0^c \frac{\text{sen}[t(x-a)]}{t}dt - 2\int_0^c \frac{\text{sen}[t(x-b)]}{t}dt\right]dF_X(x) \\ &= \mathbb{E}\left[\frac{1}{\pi}\int_0^c \frac{\text{sen}[t(x-a)]}{t}dt - \frac{1}{\pi}\int_0^c \frac{\text{sen}[t(x-b)]}{t}dt\right]. \end{aligned}$$

Defina

$$Y = \frac{1}{\pi}\int_0^c \frac{\text{sen}[t(x-a)]}{t}dt - \frac{1}{\pi}\int_0^c \frac{\text{sen}[t(x-b)]}{t}dt.$$

e então

$$\lim_{c\to\infty} Y = \frac{1}{\pi}\lim_{c\to\infty}\int_0^c \frac{\text{sen}[t(x-a)]}{t}dt - \frac{1}{\pi}\lim_{c\to\infty}\int_0^c \frac{\text{sen}[t(x-b)]}{t}dt.$$

O cálculo desse limite será feito considerando cinco situações:

(i) $x < a$ e $x < b$

Neste caso, $x-a<0$ e $x-b<0$. Daí

$$\begin{aligned} \lim_{c\to\infty} Y &= -\frac{1}{\pi}\lim_{c\to\infty}\int_0^c \frac{\text{sen}[-t(x-a)]}{t}dt + \frac{1}{\pi}\lim_{c\to\infty}\int_0^c \frac{\text{sen}[-t(x-b)]}{t}dt \\ &= \frac{(x-a)}{\pi}\lim_{c\to\infty}\int_0^c \frac{\text{sen}[-t(x-a)]}{-t(x-a)}dt - \frac{(x-b)}{\pi}\lim_{c\to\infty}\int_0^c \frac{\text{sen}[-t(x-b)]}{-t(x-b)}dt. \end{aligned}$$

Fazendo a mudança de variável $u=-t(x-a)$ na primeira integral e $u=-t(x-b)$ na segunda, tem-se

$$\begin{aligned} \lim_{c\to\infty} Y &= \frac{(x-a)}{\pi}\lim_{c\to\infty}\int_0^{-c(x-a)} s\,du - \frac{(x-b)}{\pi}\lim_{c\to\infty}\int_0^{-c(x-b)} w\,du \\ &= -\frac{1}{\pi}\lim_{c\to\infty}\int_0^{-c(x-a)} \frac{\text{sen}(u)}{u}du + \frac{1}{\pi}\lim_{c\to\infty}\int_0^{-c(x-b)} \frac{\text{sen}(u)}{u}du \end{aligned}$$

onde

$$s = \frac{\text{sen}(u)}{u}\left(\frac{-1}{(x-a)}\right) \quad \text{e} \quad w = \frac{\text{sen}(u)}{u}\left(\frac{-1}{(x-b)}\right).$$

Usando a integral de Dirichlet (Apêndice B, relação B.5), conclui-se que

$$\begin{aligned} \lim_{c\to\infty} Y &= -\frac{1}{\pi}\lim_{c\to\infty}\int_0^{-c(x-a)} \frac{\text{sen}(u)}{u}du + \frac{1}{\pi}\lim_{c\to\infty}\int_0^{-c(x-b)} \frac{\text{sen}(u)}{u}du \\ &= -\frac{1}{\pi}\cdot\frac{\pi}{2} + \frac{1}{\pi}\cdot\frac{\pi}{2} = 0. \end{aligned}$$

(ii) $x > b$ e $x > a$

Aqui, $x-b>0$ e $x-a>0$ e portanto

$$\begin{aligned} \lim_{c\to\infty} Y &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^c \frac{\text{sen}[t(x-a)]}{t}dt - \frac{1}{\pi}\lim_{c\to\infty}\int_0^c \frac{\text{sen}[t(x-b)]}{t}dt \\ &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^c \frac{\text{sen}[t(x-a)]}{t(x-a)}(x-a)dt - \frac{1}{\pi}\lim_{c\to\infty}\int_0^c \frac{\text{sen}[t(x-b)]}{t(x-b)}(x-b)dt. \end{aligned}$$

Considerando $u = t(x-a)$ e $u = t(x-b)$ na segunda integral, segue que

$$\begin{aligned}\lim_{c\to\infty} Y &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c(x-a)}\frac{\text{sen}(u)}{u}du - \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c(x-b)}\frac{\text{sen}(u)}{u}du\\ &= \frac{1}{\pi}\cdot\frac{\pi}{2} - \frac{1}{\pi}\cdot\frac{\pi}{2} = 0,\end{aligned}$$

sendo que a última linha foi obtida usando a integral de Dirichlet (Apêndice B, relação B.5).

(iii) $x = a$

Tem-se, então, $x - a = 0$ e, sendo $a < b$ por hipótese, $x - b < x - a = 0$. Desta maneira,

$$\begin{aligned}\lim_{c\to\infty} Y &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}(0)}{t}dt + \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}[-t(x-b)]}{t}dt\\ &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}[-t(x-b)]}{t(x-b)}(x-b)dt.\end{aligned}$$

Fazendo a mudança de variáveis $u = -t(x-b)$ e usando a integral de Dirichlet (Apêndice B, relação B.5),

$$\begin{aligned}\lim_{c\to\infty} Y &= -\frac{1}{\pi}\lim_{c\to\infty}\int_0^{-c(x-b)}\frac{\text{sen}(u)}{u}du\\ &= -\frac{1}{\pi}\cdot\frac{\pi}{2} = -\frac{1}{2}.\end{aligned}$$

(iv) $x = b$

Nesta situação, $x - b = 0$ e, dado que $a < b$, então $x - a > 0$. Logo,

$$\begin{aligned}\lim_{c\to\infty} Y &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}[t(x-a)]}{t}dt - \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}[t(x-b)]}{t}dt\\ &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}[t(x-a)]}{t(x-a)}(x-a)dt - \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}(0)}{t}dt\\ &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}[t(x-a)]}{t(x-a)}(x-a)dt.\end{aligned}$$

Considerando $u = t(x-b)$ e aplicando a integral de Dirichlet (Apêndice B, relação B.5),

$$\begin{aligned}\lim_{c\to\infty} Y &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c(x-a)}\frac{\text{sen}(u)}{u}du\\ &= \frac{1}{\pi}\cdot\frac{\pi}{2} = \frac{1}{2}.\end{aligned}$$

(v) $a < x < b$

Neste contexto, $x - a > 0$ e $x - b < 0$, de onde segue que

$$\begin{aligned}\lim_{c\to\infty} Y &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}[t(x-a)]}{t}dt + \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}[-t(x-b)]}{t}dt\\ &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}[t(x-a)]}{t(x-a)}(x-a)dt + \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c}\frac{\text{sen}[-t(x-b)]}{t(x-b)}(x-b)dt.\end{aligned}$$

Agora, tomando $u = (x-a)t$ na primeira integral e $u = -(x-b)t$ na segunda e pela integral de Dirichlet (Apêndice B, relação B.5), obtem-se

$$\begin{aligned}\lim_{c\to\infty} Y &= \frac{1}{\pi}\lim_{c\to\infty}\int_0^{c(x-a)}\frac{\operatorname{sen}(u)}{u}du + \frac{1}{\pi}\lim_{c\to\infty}\int_0^{-c(x-b)}\frac{\operatorname{sen}(u)}{u}du\\ &= \frac{1}{\pi}\cdot\frac{\pi}{2}+\frac{1}{\pi}\cdot\frac{\pi}{2} = 1.\end{aligned}$$

Portanto

$$Y = \begin{cases} -\frac{1}{2}, & \text{se } x = a,\\ 0, & \text{se } x < a \text{ e } x < b \text{ ou } x > b \text{ e } x > a,\\ \frac{1}{2}, & \text{se } x = b,\\ 1, & \text{se } a < x < b.\end{cases}$$

Pelo Teorema da Convergência Dominada, quando $c \to \infty$, então $\mathbb{E}(H) \to \mathbb{E}(Y)$, portanto

$$\begin{aligned}\mathbb{E}(H) &= -\frac{1}{2}P(X=a) + \frac{1}{2}P(X=b) + 1\cdot P(a<X<b)\\ &= \frac{1}{2}P(X=b) - \frac{1}{2}P(X=a) + P(a<X<b)\\ &= \frac{1}{2}\left[F(b)-F(b^-)\right] - \frac{1}{2}\left[F(a)-F(a^-)\right] + F(b) - F(a).\end{aligned}$$

■

Seja X uma variável aleatória qualquer, com função característica dada por $\varphi_X(t)$, então são válidas as seguintes propriedades:

1. $\varphi_X(0) = 1$, ou seja, a função característica assume o valor 1 no ponto $t = 0$;

   Demonstração. Usando a definição,

   $$\varphi_X(0) = \mathbb{E}(e^{iX\cdot 0}) = \mathbb{E}(e^0) = \mathbb{E}(1) = 1.$$

   ■

2. A função característica é limitada por 1, ou seja, $|\varphi_X(t)| \leq 1, \forall\, t \in \mathbb{R}$.

   Demonstração. Usando a fórmula de Euler(Apêndice B, relação B.2), segue que

   $$\begin{aligned}|\varphi_X(t)| &= |\mathbb{E}(\cos(tX) + i\operatorname{sen}(tX))|\\ &= |\mathbb{E}(\cos(tX)) + i\mathbb{E}(\operatorname{sen}(tX))|.\end{aligned}$$

   Usando a Desigualdade de Jensen(Apêndice A, relação (A.1)), obtem-se

   $$\begin{aligned}|\varphi_X(t)| &= \sqrt{[\mathbb{E}(\cos(tX))]^2 + [\mathbb{E}(\operatorname{sen}(tX))]^2}\\ &\leq \sqrt{\mathbb{E}(\cos^2(tX)) + \mathbb{E}(\operatorname{sen}^2(tX))} = \sqrt{\mathbb{E}[\cos^2(tX) + \operatorname{sen}^2(tX)]}\end{aligned}$$

conclui-se que

$$|\varphi_X(t)| \leq \sqrt{\mathbb{E}(1)} = 1.$$

■

3. $\overline{\varphi_X(t)} = \varphi_X(-t)$, ou seja, o conjugado da função característica é igual à $\varphi_X(-t)$;

   Demonstração. Para esta demonstração recorde que o $\cos(x) = \cos(-x)$, pois, o cosseno é uma função par e $-\operatorname{sen}(x) = \operatorname{sen}(-x)$, pois, o seno é uma função ímpar. Logo

$$\varphi_X(-t) = \mathbb{E}(\cos(-tX)) + i\mathbb{E}(\operatorname{sen}(-tX)) = \mathbb{E}(\cos(tX)) - i\mathbb{E}(\operatorname{sen}(tX)) = \overline{\mathbb{E}(e^{itX})}.$$

■

4. Se $Y = aX+b$, com a e b sendo constantes reais quaisquer, então a $\varphi_Y(t) = e^{itb}\varphi_X(at)$;

   Demonstração. Dado que $Y = aX + b$,

$$\begin{aligned}
\varphi_Y(t) &= \mathbb{E}[e^{itY}] = \mathbb{E}[e^{it(aX+b)}] \\
&= \mathbb{E}[e^{itaX+itb}] = \mathbb{E}[e^{itaX}e^{itb}] \\
&= e^{itb}\mathbb{E}[e^{itaX}] \\
&= e^{itb}\varphi_X(at).
\end{aligned}$$

■

5. Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias independentes com funções características dadas respectivamente por $\varphi_{X_1}(t)$, $\varphi_{X_2}(t)$, $\ldots$, $\varphi_{X_n}(t)$ e seja $Z = X_1 + \cdots + X_n$, então $\varphi_Z(t) = \varphi_{X_1}(t) \cdot \varphi_{X_2}(t) \cdot \ldots \cdot \varphi_{X_n}(t)$;

   Demonstração. Por hipótese, as variáveis são independentes, então

$$\begin{aligned}
\varphi_Z(t) &= \mathbb{E}\left[e^{it(X_1+\cdots+X_n)}\right] \\
&= \mathbb{E}\left[e^{itX_1} \cdot \ldots \cdot e^{itX_n}\right] \\
&= \mathbb{E}\left[e^{itX_1}\right] \cdot \ldots \cdot \mathbb{E}\left[e^{itX_n}\right] \\
&= \varphi_{X_1}(t) \cdot \varphi_{X_2}(t) \cdot \ldots \cdot \varphi_{X_n}(t).
\end{aligned}$$

■

6. A variável aleatória X tem distribuição simétrica em torno de zero se, e somente se, $\varphi_X(t)$ é real;

Demonstração. Sabe-se que X é simétrica em torno de zero se, e somente se,

$$P(X \leq x) = P(-X \leq x), \quad \forall x \in \mathbb{R}.$$

Sendo assim, $F_X = F_{-X}$ e X e $-X$ são identicamente distribuídas. Portanto, pela propriedade 3

$$\begin{aligned}
\varphi_X(t) &= \varphi_{-X}(t) \\
&= \overline{\varphi_{-X}(-t)} \\
&= \overline{\mathbb{E}(e^{i(-t)(-X)})} \\
&= \overline{\mathbb{E}(e^{itX})}.
\end{aligned}$$

No entanto, um número complexo só é igual ao seu conjugado se ele for real logo X é simétrica em torno de zero se e somente se $\varphi_X(t)$ é real. ∎

7. $\varphi_X(t)$ é uniformemente contínua na reta;

Demonstração. Observe que

$$\begin{aligned}
|\varphi_X(t+h) - \varphi_X(t)| &= |\mathbb{E}(e^{i(t+h)X} - e^{itX})| \\
&\leq \mathbb{E}(|e^{itX}||e^{ihX} - 1|) = \mathbb{E}(|e^{ihX} - 1|), \quad |e^{itX}| = 1.
\end{aligned}$$

Então,

$$\lim_{h \to 0} |\varphi_X(t+h) - \varphi_X(t)| \leq \lim_{h \to 0} \mathbb{E}(|e^{ihX} - 1|).$$

Pelo Teorema da Convergência Dominada, é possível trocar a ordem entre limite e esperança

$$\lim_{h \to 0} |\varphi_X(t+h) - \varphi_X(t)| \leq \mathbb{E}\left(\lim_{h \to 0}(|e^{ihX} - 1|)\right),$$

ou seja

$$\lim_{h \to 0} |\varphi_X(t+h) - \varphi_X(t)| \leq 0.$$

Como a majoração não depende de t, a continuidade é uniforme. ∎

8. φ_X é positiva definida, isto é, para todo $n = 1, 2, \ldots$ vale $\sum_{j=1}^{n} \sum_{k=1}^{n} \varphi_X(t_j - t_k) z_j \overline{z}_k \geq 0$. Para quaisquer números reais $t_1, \ldots, t_n$ e complexos $z_1, \ldots z_n$.

Demonstração. Para verificar que $\varphi_X(t)$ é positiva definida, considere $n \in \mathbb{N}$. Neste caso,

$$\begin{aligned}
\sum_{j=1}^{n}\sum_{k=1}^{n}\varphi_X(t_j - t_k)z_j\overline{z}_k &= \sum_{j=1}^{n}\sum_{k=1}^{n}\mathbb{E}(e^{it_jX-it_kX})z_j\overline{z}_k \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n}\mathbb{E}(z_j\overline{z}_k e^{it_jX-it_kX}) \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n}\mathbb{E}\left(z_je^{it_jX}\overline{z}_ke^{-it_kX}\right) \\
&= \mathbb{E}\left(\sum_{j=1}^{n}\sum_{k=1}^{n}z_je^{it_jX}\overline{z_ke^{-it_kX}}\right) \\
&= \mathbb{E}\left(\left(\sum_{j=1}^{n}z_je^{it_jX}\right)\overline{\left(\sum_{k=1}^{n}z_je^{it_kX}\right)}\right) \\
&= \mathbb{E}\left[\left|\sum_{j=1}^{n}z_je^{it_jX}\right|^2\right] \qquad \geq 0.
\end{aligned}$$

■

9. A função característica de uma variável aleatória X determina a função de distribuição de X, ou seja, dado a sua função característica podemos determinar qual é a sua função de distribuição acumulada. Esta propriedade decorre da fórmula da inversão.

10. Se $\mathbb{E}|X|^n < \infty$, então φ_X possui n derivadas contínuas e $\varphi_X^k(t) = \int_{\mathbb{R}}(ix)^k e^{itX}dF_X(x)$, $k = 1, 2, \ldots, n$. em particular $\varphi_X^k(0) = i^k\mathbb{E}(X^k)$.

Para encerrar esta seção, será mostrado que a função característica de uma v.a $X \sim N(0,1)$, é dada por

$$\varphi_X(t) = e^{-\frac{t^2}{2}}. \tag{2.8}$$

Este resultado será muito importante para as demonstrações de algumas versões do Teorema Central do Limite.

Demonstração. Por definição,

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2+2itx}{2}}dx.$$

Completando o quadrado de $\frac{-x^2+itx}{2}$, obtem-se que

$$\varphi_X(t) = \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-[(x-it)^2+t^2]}{2}}dx = e^{-\frac{t^2}{2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-(x-it)^2}{2}}dx.$$

Substituindo x por $z = x - it$,

$$\varphi_X(t) = e^{-\frac{t^2}{2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}dz,$$

em que

$$\varphi_X(t) = e^{-\frac{t^2}{2}}.$$

Essa última passagem é feita a partir do Teorema da integral de Cauchy. O teorema diz que se f é uma função analítica numa região R, ao longo de um caminho fechado C, então essa integral é igual a zero. A partir daí pode-se dividir essa integral em quatro integrais, pois são 4 caminhos, e dizer que a soma das quatro integrais é igual à zero, e daí manipulando essas integrais chega-se a conclusão que a integral é igual a 1.

Uma outra maneira de mostrar que essa integral é igual à 1, é simplesmente considerar que z é real e daí tem-se uma fdp de uma $N(0,1)$ (MAGALHÃES, 2006).

■

2.6 Convergência em Distribuição

A seguir será apresentado um conceito de convergência de sequência de variáveis aleatórias. Não é o único, mas é aquele que será usado nos teoremas que são o foco deste trabalho.

Definição 2.12. *Diz-se que $\{X_n, n \geq 0\}$ converge para X em distribuição, denotado por $X_n \xrightarrow{D} X$, se*

$$P(X_n \leq x) \rightarrow P(X \leq x),$$

quando $n \rightarrow \infty$, para todo ponto x em que $F(x) = P(X \leq x)$ é contínua.

O teorema a seguir será bastante utilizado nas demonstrações feitas no próximo capítulo.

Teorema 2.6 (**Teorema da Continuidade de Paul Lévy**)**.** *Seja $\{X_n, n \geq 0\}$ uma sequência de variáveis aleatórias com respectivas funções características $\varphi_1(t), \varphi_2(t), \ldots$. Neste caso, X_n converge em distribuição para uma variável aleatória X se, e somente se, $\varphi_n(t) \rightarrow \varphi_X(t)$.*

3 Teorema Central do Limite

O Teorema Central do Limite é um dos principais teoremas na teoria de probabilidade. Ele surgiu a partir dos trabalhos de De Moivre e de Laplace com a distribuição binomial. Na época, a demonstração feita pelo Laplace utilizando a função característica não era muito rigorosa, por causa disso, o Teorema Central do Limite foi deixado de lado por muitos anos. Entre 1900 e 1910, as demonstrações feitas pelo Laplace foram encontradas e estudadas pelos matemáticos russos Liapunov e Markov, que depois deram uma demonstração mais rigorosa. Em 1920, Sergey Bernstein apresentou uma versão p-dimensional para esse teorema. Em 1922, o Finlandês Lindeberg demonstrou uma versão mais geral do Teorema Central do Limite, pois nesta não era necessário que as variáveis fossem identicamente distribuídas. A versão dada por Lindeberg era válida para certas condições suficientes; as condições necessárias foram demonstradas pelo matemático William Feller em 1935. Existem outros matemáticos que apresentaram versões diferentes para o Teorema Central do Limite, como por exemplo Liapunov e Paul Lévy. Sendo assim, ainda existem outras versões do Teorema Central do Limite, além das seis apresentadas neste trabalho.

A teoria apresentada deste trabalho não é original e pode ser encontrada em diversos livros, como por exemplo Ash e Doléans-Dade (2000), Durret (2005), Lehmann (1999) e Cordeiro (1999), além de trabalhos acadêmicos como as dissertações de Alves (2016) e Waiandt (2014). O mérito maior aqui foi entender as demonstrações e apresentá-las de uma forma mais detalhada.

Teorema 3.1 (**Teorema Central do Limite para variáveis aleatórias i.i.d.**). *Seja* $\{X_n\}_{n\geq 1}$ *uma sequência de variáveis aleatórias i.i.d. com* $\mu = \mathbb{E}(X_1)$ *e* $\sigma^2 = Var(X_1)$, $0 < \sigma^2 < \infty$. *Então:*

$$\frac{T_n - \mathbb{E}(T_n)}{s_n} = \frac{T_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} N(0,1).$$

onde $T_n = \sum_{i=1}^{n} X_i$.

Demonstração. Pela definição de função característica, dada em (2.7), e fazendo

$$Z_n = \frac{T_n - \mathbb{E}(T_n)}{s_n} = \frac{T_n - n\mu}{\sigma\sqrt{n}},$$

segue que

$$\varphi_{Z_n}(t) = \mathbb{E}[e^{itZ_n}] = \mathbb{E}[e^{it(\frac{T_n - n\mu}{\sigma\sqrt{n}})}].$$

Mas,

$$T_n = \sum_{i=1}^{n} X_i \quad \text{e} \quad \sum_{i=1}^{n} X_i - n\mu = \sum_{i=1}^{n}(X_i - \mu),$$

logo

$$\varphi_{Z_n}(t) = \mathbb{E}\left(e^{\frac{it\sum_{i=1}^{n}(X_i-\mu)}{\sigma\sqrt{n}}}\right).$$

Agora, expandindo o somatório do expoente da exponencial e lembrando que v.a's são i.i.d, segue que

$$\begin{aligned}
\varphi_{Z_n}(t) &= \mathbb{E}\left[e^{\frac{it}{\sigma\sqrt{n}}(X_1-\mu)+\ldots+\frac{it}{\sigma\sqrt{n}}(X_n-\mu)}\right]\\
&= \mathbb{E}\left[e^{\frac{it}{\sigma\sqrt{n}}(X-\mu)}\times\ldots\times e^{\frac{it}{\sigma\sqrt{n}}(X-\mu)}\right]\\
&= \mathbb{E}\left(e^{\frac{it}{\sigma\sqrt{n}}(X-\mu)}\right)\cdot\ldots\cdot\mathbb{E}\left(e^{\frac{it}{\sigma\sqrt{n}}(X-\mu)}\right)\\
&= \left(\mathbb{E}\left[e^{\frac{it}{\sigma\sqrt{n}}(X-\mu)}\right]\right)^n,
\end{aligned}$$

ou seja,

$$\varphi_{Z_n}(t) = \left[\varphi_{(X-\mu)}\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n.$$

Usando a série de Taylor da exponencial, com $x = \frac{it}{\sigma\sqrt{n}}(X-\mu)$, até a ordem 2 e denotando as ordens maiores que dois de $\theta\left(\frac{t}{\sigma\sqrt{n}}\right)$, então, a $e^{\frac{it}{\sigma\sqrt{n}}(X-\mu)}$ será dada por

$$e^{\frac{it}{\sigma\sqrt{n}}(X-\mu)} = 1 + \frac{\left(\frac{it(X-\mu)}{\sigma\sqrt{n}}\right)^1}{1!} + \frac{\left(\frac{it(X-\mu)}{\sigma\sqrt{n}}\right)^2}{2!} + \theta\left(\frac{t}{\sigma\sqrt{n}}\right).$$

Aplicando a esperança em ambos os lados da igualdade, obtem-se que

$$\mathbb{E}\left[e^{\frac{it}{\sigma\sqrt{n}}(X-\mu)}\right] = 1 + \frac{it\mathbb{E}(X-\mu)}{\sigma\sqrt{n}} + \frac{i^2t^2\mathbb{E}(X-\mu)^2}{2\sigma^2 n} + \theta\left(\frac{t}{\sigma\sqrt{n}}\right).$$

Mas,

$$E(X-\mu) = E(X) - \mu = \mu - \mu = 0 \quad \text{e} \quad E(X-\mu)^2 = Var(X) = \sigma^2.$$

A partir daí e simplificando as potências de i, segue que

$$\varphi_{(X-\mu)}\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 + 0 - \frac{t^2\sigma^2}{2\sigma^2 n} + \theta\left(\frac{t}{\sigma\sqrt{n}}\right).$$

Porém, $\sigma^2 > 0$, então fazendo as simplificações, chega-se a

$$\varphi_{(X-\mu)}\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 - \frac{t^2}{2n} + \theta\left(\frac{t}{\sigma\sqrt{n}}\right).$$

Sendo assim, $\varphi_{Z_n}(t)$ pode ser escrita como sendo

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

Aplicando o limite com $n\to\infty$ em ambos os lados da igualdade

$$\lim_{n\to\infty}\varphi_{Z_n}(t) = \lim_{n\to\infty}\left[1 - \frac{t^2}{2n} + \theta\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n.$$

Existem muitas maneiras de resolver este limite, a solução que será dada aqui é mais simples e intuitiva. Analisando primeiro o que acontece com a expressão entre colchetes, recorde que os termos de ordem superior a dois são dados por $\theta\left(\frac{t}{\sigma\sqrt{n}}\right)$, e estes convergem para zero mais rapidamente do que os outros termos desse limite, pois, em cada denominador de θ tem-se um n com grau maior do que $\frac{t^2}{2n}$, que tem um n de grau 1. Sendo assim, pode-se analisar esse limite sem o termo $\theta\left(\frac{t}{\sigma\sqrt{n}}\right)$, já que a convergência deste vai para zero muito rapidamente. Portanto

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

Mas esse é o limite fundamental exponencial, dado em (Apêndice B, relação B.1), com $b = -\frac{t^2}{2}$, sendo assim

$$\lim_{n\to\infty} \varphi_{Z_n}(t) = e^{-\frac{t^2}{2}}.$$

Neste contexto, para todo $t \in \mathbb{R}$, $\varphi_{Z_n}(t) = \left[1 - \frac{t^2}{2n} + \theta\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n \to e^{-\frac{t^2}{2}}$, quando $n \to \infty$. Sendo assim, $\varphi_{Z_n}(t)$ tem a mesma função característica de uma normal padrão, o que implica que elas terão também a mesma função de distribuição. Desta forma, o Teorema Central do Limite para variáveis aleatórias independentes e identicamente distribuídas está provado.

■

A próxima versão do Teorema Central do Limite é um caso particular da versão clássica do Teorema Central do Limite, para uma sequência de variáveis aleatórias independentes e identicamente distribuídas. Na verdade esta versão de De Moivre e de Laplace foi a primeira do Teorema Central do Limite, mas para um caso particular, onde eles supõem que as variáveis têm distribuição de Bernoulli.

Teorema 3.2 (Teorema Central do Limite de De Moivre e Laplace). *Seja T_n o número de sucessos em n ensaios de Bernoulli independentes com probabilidade de sucesso p, $0 < p < 1$, em cada ensaio. Então*

$$\frac{T_n - np}{\sqrt{np(1-p)}} \xrightarrow{D} N(0,1).$$

Demonstração. Sendo um caso particular do teorema anterior, basta considerar $\mu = p$ e $\sigma = p(1-p)$, dado que as variáveis têm distribuição Bernoulli. ■

A versão mais geral do Teorema Central do Limite é a de Lindeberg e basicamente diz que sob certas condições suficientes $(T_n - \mathbb{E}(T_n))/s_n \xrightarrow{D} N(0,1)$. Essa versão mais geral é chamada assim porque não é necessário que as v.a's sejam identicamente distribuídas, elas só precisam ser independentes. Antes de demonstrar essa versão do teorema será apresentado um lema dos números complexos que é utilizado na demonstração desse teorema (JAMES, 2002).

Lema 3.1. *Seja $c_{n,k}$ uma sequência de números complexos tais que $\sum_{k=1}^{n} c_{n,k} \to c$ quando $n \to \infty$. Se*

$$\max_{1\leq k\leq n} |c_{n,k}| \to 0 \ \ quando \ \ n \to \infty$$

e, sendo M uma constante que não depende de n

$$\sum_{k=1}^{n} |c_{n,k}| \leq M < \infty,$$

então

$$\prod_{k=1}^{n} (1 + c_{n,k}) \to e^{c} \ \ quando \ \ n \to \infty. \tag{3.1}$$

Teorema 3.3 (**Teorema Central do Limite de Lindeberg**)**.** *Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias independentes com $\mu_n = \mathbb{E}(X_n)$ e $\sigma_n^2 = Var(X_n) < \infty$, onde pelo menos um $\sigma_n^2 > 0$. Se, para todo $\epsilon > 0$,*

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|x-\mu_j|>\epsilon s_n} (x - \mu_j)^2 dF_{X_j}(x) = 0, \tag{3.2}$$

então,

$$(T_n - \mathbb{E}(T_n))/s_n \xrightarrow{D} N(0, 1),$$

onde $T_n = X_1 + \cdots + X_n$, $s_n^2 = \sigma_1^2 + \cdots + \sigma_n^2$ e $s_n = \sqrt{\sigma_1^2 + \cdots + \sigma_n^2}$.

Demonstração. Considerando

$$Z_n = (T_n - \mathbb{E}(T_n))/s_n,$$

basta mostrar que Z_n tem a mesma função característica de uma normal padrão.

Por definição,

$$\varphi_{Z_n}(t) = \mathbb{E}[e^{itZ_n}] = \mathbb{E}[e^{it(\frac{T_n-\mu_i}{s_n})}].$$

Como $T_n - \mu_i = \sum_{i=1}^{n}(X_i - \mu_i)$ e as variáveis aleatórias são independentes, então

$$\begin{aligned}
\varphi_{Z_n}(t) &= \mathbb{E}\left[e^{\frac{it}{s_n}\sum_{i=1}^{n}(X_i-\mu_i)}\right] \\
&= \mathbb{E}\left[e^{\frac{it}{s_n}(X_1-\mu_1)} \cdot \ldots \cdot e^{\frac{it}{s_n}(X_n-\mu_n)}\right] \\
&= \mathbb{E}[e^{\frac{it}{s_n}(X_1-\mu_1)}] \cdot \ldots \cdot \mathbb{E}[e^{\frac{it}{s_n}(X_n-\mu_n)}],
\end{aligned}$$

isto é,

$$\varphi_{Z_n}(t) = \prod_{k=1}^{n} \mathbb{E}[e^{\frac{it(X_k-\mu_k)}{s_n}}]. \tag{3.3}$$

Utilizando a série de Taylor da exponencial, dada em (2.2), e expandindo até a ordem dois e os outros termos da série serão dados por $\theta_1(x)\frac{t^2x^2}{2}$, tem-se

$$e^{itx} = 1 + itx + \theta_1(x)\frac{t^2x^2}{2},$$

sendo $|\theta_1(x)| \leq 1$. Expandindo a série de Taylor da exponencial, dada em (2.2), até a ordem três e os outros termos da série serão dados por $\theta_2(x)\frac{t^3x^3}{6}$, conclui-se que

$$e^{itx} = 1 + itx - \frac{t^2x^2}{2} + \theta_2(x)\frac{t^3x^3}{6},$$

em que $|\theta_2(x)| \leq 1$. Com essas duas séries de Taylor é possível escrever uma única série da exponencial em função das outras séries, logo

$$r_\epsilon(x) = \begin{cases} (1+\theta_1(x))\frac{t^2x^2}{2}, & \text{se } |x| > \epsilon, \\ \theta_2(x)\frac{t^3x^3}{6}, & \text{se } |x| \leq \epsilon. \end{cases} \tag{3.4}$$

Deste modo, reescreve-se a esperança dada em (3.3) como sendo

$$\mathbb{E}[e^{\frac{it(X_k-\mu_k)}{s_n}}] = \int_{\mathbb{R}} e^{\frac{it(X_k-\mu_k)}{s_n}} dF_k(x).$$

Por (2.2), com $x = \frac{X_k-\mu_k}{s_n}$, e por (3.4), a série de Taylor da exponencial torna-se

$$\exp\left[\frac{it(X_k-\mu_k)}{s_n}\right] = 1 + it\left(\frac{X_k-\mu_k}{s_n}\right) - \frac{t^2}{2}\left(\frac{X_k-\mu_k}{s_n}\right)^2 + r_\epsilon\left(\frac{X_k-\mu_k}{s_n}\right),$$

o que implica que

$$\mathbb{E}\left[e^{\frac{it(X_k-\mu_k)}{s_n}}\right] = \int_{\mathbb{R}}\left[1 + it\left(\frac{X_k-\mu_k}{s_n}\right) - \frac{t^2}{2}\left(\frac{X_k-\mu_k}{s_n}\right)^2 + r_\epsilon\left(\frac{X_k-\mu_k}{s_n}\right)\right] dF_k(x).$$

Considerando $R = \mathbb{E}[e^{\frac{it(X_k-\mu_k)}{s_n}}]$ e pelas propriedades da integral definida, pode-se escrever que

$$R = \int_{\mathbb{R}} dF_k(x) + \int_{\mathbb{R}} it\ p\ dF_k(x) - \int_{\mathbb{R}} \frac{t^2}{2}\ p^2\ dF_k(x) + \int_{\mathbb{R}} r_\epsilon\ p\ dF_k(x).$$

Onde

$$p = \left(\frac{X_k-\mu_k}{s_n}\right).$$

Observe que a primeira integral é igual à 1 (pela definição de f.d.p.) e que a segunda e a terceira integrais são esperanças, logo

$$R = 1 + it\mathbb{E}\left(\frac{X_k-\mu_k}{s_n}\right) - \frac{t^2}{2}\mathbb{E}\left(\frac{X_k-\mu_k}{s_n}\right)^2 + \int_{\mathbb{R}} r_\epsilon\left(\frac{X_k-\mu_k}{s_n}\right) dF_k(x).$$

Agora, suponha que $\mathbb{E}(X_k) = \mu_k = 0$ e que $\mathbb{E}(X_k^2) = \text{Var}(X_k) = \sigma_k^2$. Neste caso,

- $\mathbb{E}\left(\frac{X_k-\mu_k}{s_n}\right) = \frac{1}{s_n}[\mathbb{E}(X_k) - \mu_k] = 0;$
- $\mathbb{E}\left(\frac{X_k-\mu_k}{s_n}\right)^2 = \frac{1}{s_n^2}\mathbb{E}(X_k-\mu_k)^2 = \frac{\sigma_k^2}{s_n^2}.$

Daí,

$$\begin{aligned} R &= 1 + it\cdot 0 - \frac{t^2}{2}\cdot\frac{\sigma_k^2}{s_n^2} + \int_{\mathbb{R}} r_\epsilon\left(\frac{X_k-\mu_k}{s_n}\right) dF_k(x) \\ &= 1 - \frac{t^2\sigma_k^2}{2s_n^2} + \int_{\mathbb{R}} r_\epsilon\left(\frac{X_k-\mu_k}{s_n}\right) dF_k(x). \end{aligned}$$

Por (3.4), segue que

$$r_\epsilon\left(\frac{X_k-\mu_k}{s_n}\right)=\begin{cases}\left[1+\theta_1\left(\frac{X_k-\mu_k}{s_n}\right)\right]\frac{t^2}{2}\cdot\left(\frac{X_k-\mu_k}{s_n}\right)^2 & \text{, se } |\frac{X_k-\mu_k}{s_n}|>\epsilon,\\ \\ \theta_2\left(\frac{X_k-\mu_k}{s_n}\right)\frac{t^3}{6}\left(\frac{X_k-\mu_k}{s_n}\right)^3 & \text{, se } |\frac{X_k-\mu_k}{s_n}|\leq\epsilon.\end{cases} \tag{3.5}$$

e como

$$\left|\frac{X_k-\mu_k}{s_n}\right|>\epsilon\Rightarrow|X_k-\mu_k|>s_n\epsilon,$$

e

$$\left|\frac{X_k-\mu_k}{s_n}\right|\leq\epsilon\Rightarrow|X_k-\mu_k|\leq s_n\epsilon.$$

e denominando a última integral de $A_{n,k}$, então, obtem-se

$$A_{n,k}=\int_{\mathbb{R}}r_\epsilon\left(\frac{X_k-\mu_k}{s_n}\right)dF_k(x),$$

então, por (3.5), essa integral será dada por:

$$A_{n,k}=L+G.$$

Onde

$$\begin{aligned}L&=\int_{|X_k-\mu_k|>s_n\epsilon}\left[1+\theta_1\left(\frac{X_k-\mu_k}{s_n}\right)\right]\frac{t^2}{2}\cdot\left(\frac{X_k-\mu_k}{s_n}\right)^2 dF_k(x)\\&=\frac{t^2}{2s_n^2}\int_{|X_k-\mu_k|>s_n\epsilon}\left[1+\theta_1\left(\frac{X_k-\mu_k}{s_n}\right)\right]\cdot(X_k-\mu_k)^2\,dF_k(x)\end{aligned}$$

e

$$\begin{aligned}G&=\int_{|X_k-\mu_k|\leq s_n\epsilon}\theta_2\left(\frac{X_k-\mu_k}{s_n}\right)\frac{t^3}{6}\left(\frac{X_k-\mu_k}{s_n}\right)^3 dF_k(x)\\&=\frac{t^3}{6s_n^2}\int_{|X_k-\mu_k|\leq s_n\epsilon}\left(\frac{X_k-\mu_k}{s_n}\right)\theta_2\left(\frac{X_k-\mu_k}{s_n}\right)(X_k-\mu_k)^2\,dF_k(x).\end{aligned}$$

Daí segue que:

$$R=1-\frac{t^2\sigma_k^2}{2s_n^2}+A_{n,k}.$$

Entretanto, $R=\mathbb{E}\left[e^{\frac{it(X_k-\mu_k)}{s_n}}\right]$ então

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

Desta forma, a Equação (3.3) se torna

$$\varphi_{Z_n}(t)=\prod_{k=1}^{n}\left[1-\frac{t^2\sigma_k^2}{2s_n^2}+A_{n,k}\right].$$

Agora veja o que acontece com $A_{n,k}$

- Se $|\theta_1(x)| \leq 1$, ou equivalentemente, se $-1 \leq \theta_1(x) \leq 1$, ou ainda, $0 \leq \theta_1(x) + 1 \leq 2$ e fazendo $x = \frac{X_k - \mu_k}{s_n}$, então

$$0 \leq 1 + \theta_1 \left(\frac{X_k - \mu_k}{s_n} \right) \leq 2.$$

- Se $|\theta_2(x)| \leq 1$, multiplicando ambos os lados dessa desigualdade por $\epsilon > 0$ chega-se a $\epsilon|\theta_2(x)| \leq \epsilon$. Dado que $|\frac{X_k - \mu_k}{s_n}| \leq \epsilon$ e fazendo $x = \frac{X_k - \mu_k}{s_n}$ segue que

$$\left| \frac{X_k - \mu_k}{s_n} \right| \cdot \left| \theta_2 \left(\frac{X_k - \mu_k}{s_n} \right) \right| \leq \epsilon.$$

Tomando o módulo de $A_{n,k}$, obtem-se que:

$$|A_{n,k}| = |L + G|.$$

Mas

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx, \text{ se } a \leq b$$

de onde se conclui que

$$|A_{n,k}| = |L + G| \leq D + F.$$

Onde:

$$D = \frac{t^2}{2s_n^2} \int_{|X_k - \mu_k| > s_n \epsilon} \left| 1 + \theta_1 \left(\frac{X_k - \mu_k}{s_n} \right) \right| \cdot (X_k - \mu_k)^2 \, dF_k(x)$$

e

$$F = \frac{|t^3|}{6s_n^2} \int_{|X_k - \mu_k| \leq s_n \epsilon} \left| \frac{X_k - \mu_k}{s_n} \right| \left| \theta_2 \left(\frac{X_k - \mu_k}{s_n} \right) \right| (X_k - \mu_k)^2 \, dF_k(x).$$

Mas, $0 \leq 1 + \theta_1 \left(\frac{X_k - \mu_k}{s_n} \right) \leq 2$ e $\left| \frac{X_k - \mu_k}{s_n} \right| \left| \theta_2 \left(\frac{X_k - \mu_k}{s_n} \right) \right| \leq \epsilon$, então, usando essas restrições, segue que

$$D = \frac{t^2}{2s_n^2} \int_{|X_k - \mu_k| > s_n \epsilon} 2 \cdot (X_k - \mu_k)^2 \, dF_k(x) = \frac{t^2}{s_n^2} \int_{|X_k - \mu_k| > s_n \epsilon} (X_k - \mu_k)^2 \, dF_k(x)$$

e

$$F = \frac{|t^3|}{6s_n^2} \int_{|X_k - \mu_k| \leq s_n \epsilon} \epsilon \cdot (X_k - \mu_k)^2 \, dF_k(x) = \frac{\epsilon |t^3|}{6s_n^2} \int_{|X_k - \mu_k| \leq s_n \epsilon} (X_k - \mu_k)^2 \, dF_k(x).$$

Portanto

$$|A_{n,k}| = |L + G| \leq \frac{t^2}{s_n^2} \int_{|X_k - \mu_k| > s_n \epsilon} (X_k - \mu_k)^2 \, dF_k(x) + \frac{\epsilon |t^3|}{6s_n^2} \int_{|X_k - \mu_k| \leq s_n \epsilon} (X_k - \mu_k)^2 \, dF_k(x)$$

Sendo assim,

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

Fazendo o k variar de 1 até n, conclui-se que

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

Quando n for suficientemente grande, a primeira integral será igual à condição de Lindeberg, dada em (3.2), e utilizando (3.8) a segunda integral será igual à 1. Portanto

$$\sum_{k=1}^{n} |A_{n,k}| \leq t^2 \cdot 0 + \frac{\epsilon |t^3|}{6} \cdot 1 \Rightarrow \sum_{k=1}^{n} |A_{n,k}| \leq \frac{\epsilon |t^3|}{6}.$$

Agora, deve-se escolher um valor para ϵ de modo que este limite seja fácil de manipular e que convirja para zero. Considere $\epsilon = \frac{1}{m}$. Aplicando o limite quando $m \to \infty$ em ambos os lados da desigualdade e fazendo $\epsilon = \frac{1}{m}$

$$\lim_{m\to\infty} \sum_{k=1}^{n} |A_{n,k}| \leq \lim_{m\to\infty} \frac{|t^3|}{6m}.$$

Como $\epsilon = \frac{1}{m} \to 0$ quando $m \to \infty$ e $\sum_{k=1}^{n} |A_{n,k}|$ não é negativo, então

$$\sum_{k=1}^{n} |A_{n,k}| \to 0 \ \text{ quando } \ n \to \infty. \tag{3.6}$$

Antes de aplicar o Lema (3.1), é preciso mostrar que as condições desse lema estão satisfeitas, fazendo com que a convergência de (3.1) seja válida. Faça $c_{n,k} = -\frac{t^2\sigma_k^2}{2s_n^2} + A_{n,k}$ e $c = -\frac{t^2}{2}$, sendo assim, devemos verificar se:

- $\max\limits_{1\leq k\leq n} |c_{n,k}| \to 0 \ \text{ quando } \ n \to \infty$;

$$\begin{aligned}
\max_{1\leq k\leq n} |c_{n,k}| &\leq \max_{1\leq k\leq n} \left| -\frac{t^2\sigma_k^2}{2s_n^2} + A_{n,k} \right| \\
&\leq \max_{1\leq k\leq n} \left| -\frac{t^2\sigma_k^2}{2s_n^2} \right| + \max_{1\leq k\leq n} |A_{n,k}| \\
&\leq \frac{t^2}{2} \max_{1\leq k\leq n} \left| \frac{\sigma_k^2}{s_n^2} \right| + \max_{1\leq k\leq n} |A_{n,k}|.
\end{aligned}$$

  Desta maneira, o máximo de $|c_{n,k}|$ converge para zero quando $n \to \infty$.

- $\sum\limits_{k=1}^{n} |c_{n,k}| \leq M < \infty$.

$$\begin{aligned}
\sum_{k=1}^{n} |c_{n,k}| &\leq \sum_{k=1}^{n} \left(-\frac{t^2\sigma_k^2}{2s_n^2} + A_{n,k} \right) \\
&\leq \sum_{k=1}^{n} -\frac{t^2\sigma_k^2}{2s_n^2} + \sum_{k=1}^{n} A_{n,k} \\
&\leq \frac{-t^2}{2} \cdot \frac{1}{s_n^2} \cdot \sum_{k=1}^{n} \sigma_k^2 + \sum_{k=1}^{n} A_{n,k}.
\end{aligned}$$

  Por (3.6) e por (3.7), $\sum\limits_{k=1}^{n} |c_{n,k}|$ converge para $-\frac{t^2}{2}$ quando $n \to \infty$, ou seja,

$$\sum_{k=1}^{n} |c_{n,k}| \leq M.$$

Como todas as condições do Lema (3.1) estão satisfeitas, conclui-se que

$$\varphi_{Z_n}(t) = \prod_{k=1}^{n}\left[1 - \frac{t^2\sigma_k^2}{2s_n^2} + A_{n,k}\right] \to e^{-\frac{t^2}{2}} \text{ quando } n \to \infty.$$

Sendo assim, $\varphi_{Z_n}(t)$ tem a mesma função característica de uma normal padrão, logo, pelo Teorema da Unicidade, elas terão também a mesma função de distribuição, portanto, o Teorema Central do Limite de Lindeberg está provado.

■

Observação 3.1. *A condição (3.2) é chamada de condição de Lindeberg e como, para todo $\epsilon > 0$,*

$$s_n^2 = \sum_{j=1}^{n} \sigma_j^2 \tag{3.7}$$

mas, utilizando a definição de variância para o caso contínuo, pode-se escrever

$$s_n^2 = \sum_{j=1}^{n} \int_{\mathbb{R}} (x - \mu_j)^2 dF_{Xj}(x).$$

Daí, segue que, a integral do lado direito da igualdade pode ser dividida em duas, uma com limites de integração para valores menores ou iguais a zero e a outra com limites de integração para valores maiores do que zero, assim

$$s_n^2 = \sum_{j=1}^{n} \int_{|xx-\mu_j| \leq \epsilon s_n} (x - \mu_j)^2 dF_{Xj}(x) + \sum_{j=1}^{n} \int_{|xx-\mu_j| > \epsilon s_n} (x - \mu_j)^2 dF_{Xj}(x)$$

multiplicando ambos os lados da igualdade por $\frac{1}{s_n^2}$, segue que

$$\frac{1}{s_n^2} \cdot s_n^2 = \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|x-\mu_j| \leq \epsilon s_n} (x - \mu_j)^2 dF_{Xj}(x) + \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|x-\mu_j| > \epsilon s_n} (x - \mu_j)^2 dF_{Xj}(x)$$

$$1 = \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|x-\mu_j| \leq \epsilon s_n} (x - \mu_j)^2 dF_{Xj}(x) + \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|x-\mu_j| > \epsilon s_n} (x - \mu_j)^2 dF_{Xj}(x)$$

aplicando o limite quando $n \to \infty$, logo

$$1 = \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|xx-\mu_j| \leq \epsilon s_n} (x - \mu_j)^2 dF_{Xj}(x) + \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|xx-\mu_j| > \epsilon s_n} (x - \mu_j)^2 dF_{Xj}(x)$$

portanto, pela condição de Lindeberg, a última integral é igual à zero, então

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|x-\mu_j| \leq \epsilon s_n} (x - \mu_j)^2 dF_{Xj}(x) = 1. \tag{3.8}$$

Isso significa que a condição de Lindeberg é equivalente a esse último resultado.

Proposição 3.1. *A condição de Lindeberg implica que*

$$\max_{1\leq j\leq n} \frac{\sigma_j^2}{s_n^2} \to 0 \quad \textit{quando} \quad n \to \infty.$$

Demonstração. Para todos $\epsilon > 0$, $n \geq 1$ e j, $1 \leq j \leq n$

$$\begin{aligned}
\frac{\sigma_j^2}{s_n^2} &= \frac{1}{s_n^2} \cdot \sigma_j^2 = \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{\mathbb{R}} (x-\mu_j)^2 dF_{Xj}(x) \\
&= \frac{1}{s_n^2} \int_{|x-\mu_j| \leq \epsilon s_n} (x-\mu_j)^2 dF_{Xj}(x) + \frac{1}{s_n^2} \int_{|x-\mu_j| > \epsilon s_n} (x-\mu_j)^2 dF_{Xj}(x) \\
&\leq \frac{1}{s_n^2} \int_{\mathbb{R}} \epsilon^2 s_n^2 dF_{Xj}(x) + \frac{1}{s_n^2} \int_{|x-\mu_j| > \epsilon s_n} (x-\mu_j)^2 dF_{Xj}(x).
\end{aligned}$$

Simplificando e aplicando o limite quando $n \to \infty$, segue que

$$\lim_{n\to\infty} \frac{\sigma_j^2}{s_n^2} \leq \epsilon^2 + \lim_{n\to\infty} \frac{1}{s_n^2} \int_{|x-\mu_j| > \epsilon s_n} (x-\mu_j)^2 dF_{Xj}(x).$$

Pela condição de Lindeberg, a última integral é igual a zero, logo

$$\lim_{n\to\infty} \frac{\sigma_j^2}{s_n^2} \leq \epsilon^2.$$

Aplicando o máximo, tem-se que

$$\lim_{n\to\infty} \max_{1 \leq j \leq n} \frac{\sigma_j^2}{s_n^2} \leq \epsilon^2.$$

Como esse limite não pode ser menor do que zero e ϵ é arbitrário, então

$$\max_{1 \leq j \leq n} \frac{\sigma_j^2}{s_n^2} \to 0 \text{ quando } n \to \infty.$$

■

Proposição 3.2. *A condição de Lindeberg implica que, para todo $\epsilon > 0$,*

$$\max_{1 \leq j \leq n} P(|X_j - \mu_j| > \epsilon s_n) \to 0 \;\; \textit{quando} \;\; n \to \infty.$$

Demonstração. Considere que para todos $\epsilon > 0$, $n \geq 1$ e $1 \leq j \leq n$

$$\begin{aligned}
\frac{\sigma_j^2}{s_n^2} &= \frac{1}{s_n^2} \int_{\mathbb{R}} (x-\mu_j)^2 dF_{Xj}(x) \\
&\geq \frac{1}{s_n^2} \int_{|x-\mu_j| > \epsilon s_n} (x-\mu_j)^2 dF_{Xj}(x) \\
&\geq \frac{1}{s_n^2} \int_{|x-\mu_j| > \epsilon s_n} \epsilon^2 s_n^2 dF_{Xj}(x) = \epsilon^2 \int_{|x-\mu_j| > \epsilon s_n} dF_{Xj}(x).
\end{aligned}$$

Observando que

$$\epsilon^2 \int_{|x-\mu_j| > \epsilon s_n} dF_{Xj}(x) = \epsilon^2 P(|X_j - \mu_j| > \epsilon s_n),$$

conclui-se que

$$P(|X_j - \mu_j| > \epsilon s_n) \leq \frac{1}{\epsilon} \cdot \frac{\sigma_j^2}{s_n^2},$$

ou, equivalentemente, que

$$\max_{1\leq j\leq n} P(|X_j - \mu_j| > \epsilon s_n) \leq \max_{1\leq j\leq n} \frac{1}{\epsilon}\cdot\frac{\sigma_j^2}{s_n^2}.$$

Desta forma, quando $n \to \infty$, pela Proposição (3.1), segue que

$$\max_{1\leq j\leq n} P(|X_j - \mu_j| > \epsilon s_n) \leq \frac{1}{\epsilon}\max_{1\leq j\leq n}\frac{\sigma_j^2}{s_n^2} \to 0 \text{ quando } n\to\infty.$$

■

Proposição 3.3. *Seja $\{X_n\}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, com $\mu = \mathbb{E}(X_1)$ e $\sigma^2 = \mathrm{Var}(X_1), 0 < \sigma^2 < \infty$, então, para todo $\epsilon > 0$*

$$\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) = 0 \quad \textit{quando } n\to\infty.$$

Demonstração. Por definição,

$$s_n^2 = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2,$$

mas, as variáveis aleatórias são independentes e identicamente distribuídas, logo,

$$s_n^2 = n\sigma^2 \quad \text{e} \quad s_n = \sigma\sqrt{n}.$$

Sendo assim, pode-se escrever que

$$\begin{aligned}\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) &= \frac{1}{n\sigma^2}\sum_{j=1}^{n}\int_{|x-\mu|>\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF_{X_1}(x)\\ &= \frac{n}{n\sigma^2}\int_{|x-\mu|>\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF_{X_1}(x)\\ &= \frac{1}{\sigma^2}\int_{|x-\mu|>\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF_{X_1}(x).\end{aligned}$$

Usando o complementar da última integral, pode-se escrever que

$$\begin{aligned}\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) &= 1 - \frac{1}{\sigma^2}\int_{|x-\mu|\leq\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF_{X_1}(x)\\ &= 1 - \frac{1}{\sigma^2}\left[\int_{\mu-\epsilon\sigma\sqrt{n}}^{\mu}(x-\mu)^2 dF_{X_1}(x) + \int_{\mu}^{\mu+\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF_{X_1}(x)\right].\end{aligned}$$

Daí segue que

$$\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) = 1 - \frac{1}{\sigma^2}\left[\int_{\mu-\epsilon\sigma\sqrt{n}}^{\mu+\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF_{X_1}(x)\right]$$

aplicando o limite quando $n \to \infty$, logo

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) &= 1 - \lim_{n\to\infty}\frac{1}{\sigma^2}\left[\int_{\mu-\epsilon\sigma\sqrt{n}}^{\mu+\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF_{X_1}(x)\right]\\ &= 1 - \frac{1}{\sigma^2}\left[\int_{-\infty}^{+\infty}(x-\mu)^2 dF_{X_1}(x)\right].\end{aligned}$$

Observe que esta última integral é igual a $\mathbb{E}(X-\mu)^2 = \mathrm{Var}(X) = \sigma^2$, logo

$$\lim_{n\to\infty} \frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) = 1 - \frac{1}{\sigma^2}\cdot\sigma^2 = 0.$$

Como a condição de Lindeberg está satisfeita, então conclui-se que o TCL de v.a's i.i.d é um caso particular do TCL de Lindeberg. ■

A próxima versão do Teorema Central do Limite é a de Liapunov. Na verdade esta versão é uma consequência da versão de Lindeberg. O Teorema Central do Limite de Liapunov é muito útil no caso de as variáveis X_n terem momentos finitos de ordem maior do que dois.

Teorema 3.4 (**Teorema Central do Limite de Liapunov**). *Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias independentes com $\mu_n = \mathbb{E}(X_n)$ e $\sigma_n^2 = Var(X_n) < \infty$, onde pelo menos um $\sigma_n^2 > 0$, e suponhamos que, para algum $\delta > 0$, $\mathbb{E}|X_n-\mu_n|^{2+\delta} < \infty$, $n = 1, 2, \ldots$. Se*

$$\lim_{n\to\infty} s_n^{-2-\delta}\sum_{j=1}^{n}\mathbb{E}|X_j-\mu_j|^{2+\delta} = 0 \tag{3.9}$$

então,

$$(T_n - \mathbb{E}(T_n))/s_n \xrightarrow{D} N(0,1).$$

Demonstração. Para esta demonstração, é suficiente mostrar que a condição de Liapunov implica a de Lindeberg. Se $|x-\mu_j| > \epsilon s_n \Rightarrow |x-\mu_j|^\delta > \epsilon^\delta s_n^\delta \Rightarrow \frac{|x-\mu_j|^\delta}{\epsilon^\delta s_n^\delta} > 1$.

Considere que vale a condição (3.9). Neste caso, usando a condição de Lindeberg, que diz que

$$\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) = 1.$$

Pode-se escrever uma desigualdade que envolva a condição de Lindeberg, mas que seja positiva e maior do que 1. Logo para todo $\epsilon > 0$, segue que

$$\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) \leq \frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2\frac{|x-\mu_j|^\delta}{\epsilon^\delta s_n^\delta} dF_{X_j}(x).$$

A integral do lado esquerdo da desigualdade é igual a 1 e a integral do lado direito da desigualdade é maior do que 1, pois $\frac{|x-\mu_j|^\delta}{\epsilon^\delta s_n^\delta} > 1$. Além disso, $|x|^2 = x^2$, então, pode-se reescrever a integral do lado direito da desigualdade como sendo

$$\begin{aligned}
\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) &\leq \frac{1}{\epsilon^2 s_n^{2+\delta}}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}|x-\mu_j|^{2+\delta} dF_{X_j}(x)\\
&\leq \frac{1}{\epsilon^2 s_n^{2+\delta}}\sum_{j=1}^{n}\int_{\mathbb{R}}|x-\mu_j|^{2+\delta} dF_{X_j}(x)\\
&\leq \frac{1}{\epsilon^2 s_n^{2+\delta}}\sum_{j=1}^{n}\mathbb{E}|x-\mu_j|^{2+\delta}.
\end{aligned}$$

Considerando
$$R = \frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x)$$
e tomando o limite quando $n \to \infty$, pode-se escrever que
$$\lim_{n\to\infty} R \leq \lim_{n\to\infty}\frac{1}{\epsilon^2 s_n^{2+\delta}}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}|x-\mu_j|^{2+\delta}dF_{X_j}(x) \leq \lim_{n\to\infty}\frac{1}{\epsilon^2 s_n^{2+\delta}}\sum_{j=1}^{n}\mathbb{E}|x-\mu_j|^{2+\delta}.$$
Aplicando a condição de Liapunov, segue que
$$\lim_{n\to\infty} R \leq \lim_{n\to\infty}\frac{1}{\epsilon^2 s_n^{2+\delta}}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}|x-\mu_j|^{2+\delta}dF_{X_j}(x) \leq 0,$$
ou seja,
$$\lim_{n\to\infty}\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) \leq \lim_{n\to\infty}\frac{1}{\epsilon^2 s_n^{2+\delta}}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}|x-\mu_j|^{2+\delta}dF_{X_j}(x) \leq 0.$$
Portanto,
$$\lim_{n\to\infty}\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) \leq 0.$$
Mas esta última integral não pode ser menor do que zero, já que é positiva, então ela só pode ser igual à zero, sendo assim
$$\lim_{n\to\infty}\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon s_n}(x-\mu_j)^2 dF_{X_j}(x) = 0.$$
Dado que a condição de Lindeberg está satisfeita, o Teorema Central do Limite de Liapunov está provado.

■

A seguir veremos uma versão do Teorema Central do Limite que é muito utilizada em análise de regressão, esta versão é chamada de Teorema Central do Limite de Hájek-Sidak.

Teorema 3.5 (Teorema Central do Limite de Hájek-Sidak). *Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias i.i.d. com $\mu = \mathbb{E}(X_1)$ e $\sigma^2 = Var(X_1)$, $0 < \sigma^2 < \infty$. Seja $\{c_n\}_{n\geq 1} = \{(c_{n1}, c_{n2}, \ldots, c_{nn})\}_{n\geq 1}$ uma sequência de vetores reais não nulos. Se $\max\limits_{1\leq j\leq n} c_{nj}^2 / \sum\limits_{k=1}^{n} c_{nk}^2 \to 0$ quando $n \to \infty$ (condição de Noether), então:*
$$Z_n = \frac{\sum_{j=1}^{n} c_{nj}(X_j-\mu)}{\left(\sigma^2\sum_{j=1}^{n} c_{nj}^2\right)^{\frac{1}{2}}} \xrightarrow{D} N(0,1).$$

Demonstração. Sejam $Y_{nj} = c_{nj}X_j$, $n \geq 1$ e $1 \leq j \leq n$. Então $\{Y_{nj}, 1 \leq j \leq n\}_{n\geq 1}$ é uma disposição triangular de v.a's tal que, para cada $n \geq 1$, $Y_{n1}, \ldots, Y_{nn}$ são v.a's independentes. Seja $T_n = \sum_{j=1}^{n} Y_{nj}$, $\ n \geq 1$. Neste caso,
$$\mathbb{E}(T_n) = \mu\sum_{j=1}^{n} c_{nj} \quad \text{e} \quad s_n^2 = \mathrm{Var}(T_n) = \sigma^2\sum_{j=1}^{n} c_{nj}^2$$

e, para mostrar que

$$Z_n = (T_n - \mathbb{E}(T_n))/s_n \xrightarrow{D} N(0,1),$$

basta mostrar que a condição de Lindeberg está satisfeita. Ou seja, basta mostrar que, para todo $\epsilon > 0$,

$$\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|y-c_{nj}\mu|>\epsilon s_n}(y-c_{nj}\mu)^2 dF_{Y_{nj}}(y) \to 0 \text{ quando } n \to \infty.$$

Suponha que

$$\max_{1\leq j\leq n} c_{nj}^2 / \sum_{k=1}^{n} c_{nk}^2 \to 0 \quad \text{quando} \quad n \to \infty.$$

Denotando por I_A a função indicadora de um evento A, segue que, para todo $\epsilon > 0$,

$$\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|y-c_{nj}\mu|>\epsilon s_n}(y-c_{nj}\mu)^2 dF_{Y_{nj}}(y) = \frac{1}{s_n^2}\sum_{j=1}^{n}\mathbb{E}\left[(Y_{nj}-c_{nj}\mu)^2 I_{(|Y_{nj}-c_{nj}\mu|>\epsilon s_n)}\right].$$

Mas $Y_{nj} = c_{nj}X_j$, logo

$$\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|y-c_{nj}\mu|>\epsilon s_n}(y-c_{nj}\mu)^2 dF_{Y_{nj}}(y) = \frac{1}{s_n^2}\sum_{j=1}^{n}\mathbb{E}\left[(c_{nj}X_j-c_{nj}\mu)^2 I_{(|c_{nj}X_j-c_{nj}\mu|>\epsilon s_n)}\right].$$

Daí

$$(c_{nj}X_j - c_{nj}\mu)^2 = (c_{nj}[X_j-\mu])^2 = c_{nj}^2(X_j-\mu)^2$$

e a função indicadora pode ser reescrita como

$$I_{|c_{nj}X_j-c_{nj}\mu|>\epsilon s_n} = I_{|c_{nj}||X_j-\mu|>\epsilon s_n} = I_{c_{nj}^2(X_j-\mu)^2>\epsilon^2 s_n^2}.$$

Portanto

$$\begin{aligned}
\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|y-c_{nj}\mu|>\epsilon s_n}(y-c_{nj}\mu)^2 dF_{Y_{nj}}(y) &= \frac{1}{s_n^2}\sum_{j=1}^{n}\mathbb{E}\left[c_{nj}^2(X_j-\mu)^2 I_{\left(c_{nj}^2(X_j-\mu)^2>\epsilon^2 s_n^2\right)}\right]\\
&= \frac{1}{s_n^2}\sum_{j=1}^{n}c_{nj}^2\mathbb{E}\left[(X_j-\mu)^2 I_{\left(c_{nj}^2(X_j-\mu)^2>\epsilon^2 s_n^2\right)}\right].
\end{aligned}$$

Denominando essa esperança de R e tomando o valor máximo, segue que

$$R \leq \frac{1}{s_n^2}\sum_{j=1}^{n}c_{nj}^2\mathbb{E}(X_j-\mu)^2 I_{\left((X_j-\mu)^2>\epsilon^2 s_n^2/\max\limits_{1\leq j\leq n} c_{nj}^2\right)}.$$

Como as v.a's são i.i.d.,

$$s_n^2 = \sigma^2\sum_{j=1}^{n}c_{nj}^2$$

o que implica que

$$\frac{1}{\sigma^2} = \frac{1}{s_n^2}\sum_{j=1}^{n}c_{nj}^2.$$

Então

$$R \leq \frac{1}{\sigma^2}\mathbb{E}(X-\mu)^2 I_{\left((X-\mu)^2>\epsilon^2\sigma^2\sum_{j=1}^{n} c_{nj}^2/\max_{1\leq j\leq n} c_{nj}^2\right)}$$
$$\leq \frac{1}{\sigma^2}\mathbb{E}(X-\mu)^2 I_{\left(|X-\mu|>\epsilon\sigma\left(\sum_{j=1}^{n} c_{nj}^2/\max_{1\leq j\leq n} c_{nj}^2\right)^{\frac{1}{2}}\right)}.$$

Porém,

$$R = \frac{1}{s_n^2}\sum_{j=1}^{n} c_{nj}^2\mathbb{E}\left[(X_j-\mu)^2 I_{\left(c_{nj}^2(X_j-\mu)^2>\epsilon^2 s_n^2\right)}\right] = \frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|y-c_{nj}\mu|>\epsilon s_n}(y-c_{nj}\mu)^2 dF_{Y_{nj}}(y),$$

e, neste caso,

$$\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|y-c_{nj}\mu|>\epsilon s_n}(y-c_{nj}\mu)^2 dF_{Y_{nj}}(y) \leq \frac{1}{\sigma^2}\mathbb{E}(X-\mu)^2 I_t.$$

Aplicando o limite quando $n \to \infty$ em ambos os lados da desigualdade, obtem-se que

$$\lim_{n\to\infty}\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|y-c_{nj}\mu|>\epsilon s_n}(y-c_{nj}\mu)^2 dF_{Y_{nj}}(y) \leq \lim_{n\to\infty}\frac{1}{\sigma^2}\mathbb{E}(X-\mu)^2 I_t.$$

onde

$$t = \left(|X-\mu| > \epsilon\sigma\left(\sum_{j=1}^{n} c_{nj}^2/\max_{1\leq j\leq n} c_{nj}^2\right)^{\frac{1}{2}}\right).$$

Inicialmente considerou-se que

$$\max_{1\leq j\leq n} c_{nj}^2/\sum_{k=1}^{n} c_{nk}^2 \to 0 \quad \text{quando} \quad n\to\infty.$$

Consequentemente

$$\sum_{j=1}^{n} c_{nj}^2/\max_{1\leq j\leq n} c_{nj}^2 \to \infty \quad \text{quando} \quad n\to\infty.$$

Como a variância tem que ser finita, ou seja, $0 < \sigma^2 < \infty$, então

$$\lim_{n\to\infty}\frac{1}{\sigma^2}\mathbb{E}(X-\mu)^2 I_{\left(|X-\mu|>\epsilon\sigma\left(\sum_{j=1}^{n} c_{nj}^2/\max_{1\leq j\leq n} c_{nj}^2\right)^{\frac{1}{2}}\right)} = 0,$$

pois a variância não será finita. E como a integral de Lindeberg é positiva, então, ela não pode ser menor do que zero, logo, ela só poder ser igual à zero, portanto

$$\lim_{n\to\infty}\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|y-c_{nj}\mu|>\epsilon s_n}(y-c_{nj}\mu)^2 dF_{Y_{nj}}(y) = 0.$$

Assim, pela condição de Lindeberg, o teorema está demonstrado.

$\blacksquare$

O próximo Teorema Central do Limite é referente aos vetores aleatórios i.i.d., mas antes de enunciá-lo, é muito importante conhecer dois teoremas que serão úteis na demonstração dessa versão do Teorema Central do Limite. A seguir serão enunciados esses dois teoremas.

Teorema 3.6 (Teorema de Cramér-Wold)**.** *Sejam* $X_n = (X_{n1}, X_{n2}, \ldots, X_{nk})$ *e* $X = (X_{01}, X_{02}, \ldots, X_{0k})$ *vetores aleatórios k-dimensionais.* $X_n \xrightarrow{D} X$ *se, e somente se,*

$$\sum_{j=1}^{k} t_j X_{nj} \xrightarrow{D} \sum_{j=1}^{k} t_j X_{0j} \quad \textit{quando} \quad n \to \infty,$$

para todo $(t_1, t_2, \ldots, t_k) \in \mathbb{R}^k$*.*

Teorema 3.7. $\boldsymbol{Z} = (Z_1, Z_2, \ldots, Z_p) \sim N(\mu, \sum)$ *se, e somente se,* $\forall\ t = (t_1, t_2, \ldots, t_p) \in \mathbb{R}^p$*,*

$$tZ' = \sum_{j=1}^{p} t_j Z_j \sim N(t\mu', t \sum t').$$

O Teorema Central do Limite para sequências de vetores aleatórios i.i.d. é uma extensão, para o caso p dimensional, do Teorema Central do Limite para v.a's independentes e identicamente distribuídas. A seguir será enunciado esse teorema e a sua demonstração, que será feita com base nos dois últimos teoremas.

Teorema 3.8 (**Teorema Central do Limite para sequências de vetores aleatórios i.i.d.**)**.** *Seja* $\{X_n\}_{n\geq 1} = \{(X_{n1}, X_{n2}, \ldots, X_{np})\}_{n\geq 1}$ *sequência de vetores aleatórios p-dimensionais i.i.d. Sejam* μ *o vetor de médias e* Σ *a matriz de covariância de* X_1*, onde* $\Sigma = (\sigma_{ij})_{1\leq i,\, j\leq p}$ *é finita. Então:*

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) \xrightarrow{D} N(0, \Sigma).$$

Demonstração. Para esta demonstração é necessário utilizar os teoremas de Cramer-Wold e o TCL para o caso i.i.d. e daí basta mostrar que, para todo $t = (t_1, t_2, \ldots, t_p) \in \mathbb{R}^p$

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{p} t_j (X_{ij} - \mu_j) \xrightarrow{D} N(0, t \sum t').$$

Considere $Y_i = \sum_{j=1}^{p} t_j (X_{ij} - \mu_j)$, $i = 1, 2, \ldots, n$. Como $\{Y_n\}_{n\geq 1}$ é uma sequência de v.a's i.i.d. com $\mathbb{E}(Y_i) = 0$ e $\mathrm{Var}(Y_i) = \mathbb{E}(Y_i^2)$, reescrevendo essa variância, segue que

$$\mathrm{Var}(Y_i) = \mathbb{E}\left[\left(\sum_{j=1}^{p} t_j (X_{ij} - \mu_j)\right)^2\right]$$

e

$$\mathbb{E}\left[\left(\sum_{j=1}^{p} t_j(X_{ij}-\mu_j)\right)^2\right] = \mathbb{E}\left[\left(\sum_{j=1}^{p} t_j(X_{ij}-\mu_j)\right)\left(\sum_{k=1}^{p} t_k(X_{ik}-\mu_k)\right)\right]$$
$$= \mathbb{E}\left[\left(\sum_{j=1}^{p}\sum_{k=1}^{p} t_j t_k(X_{ij}-\mu_j)(X_{ik}-\mu_k)\right)\right]$$
$$= \left[\left(\sum_{j=1}^{p}\sum_{k=1}^{p} t_j t_k \mathbb{E}[(X_{ij}-\mu_j)(X_{ik}-\mu_k)]\right)\right].$$

Como $\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))]$, então

$$\mathbb{E}\left[\left(\sum_{j=1}^{p} t_j(X_{ij}-\mu_j)\right)^2\right] = \sum_{j=1}^{p}\sum_{k=1}^{p} t_j t_k \,\mathrm{Cov}(X_{ij},X_{ik}) = \sum_{j=1}^{p}\sum_{k=1}^{p} t_j t_k \sigma_{jk} = t\sum t'.$$

Pelo TLC de v.a's i.i.d, segue que

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i \xrightarrow{D} N(0, t\sum t').$$

Daí,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{p} t_j(X_{ij}-\mu_j) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i \xrightarrow{D} N(0, t\sum t').$$

e, portanto, o teorema está demonstrado.

■

4 Conclusão

O presente trabalho apresentou uma revisão geral sobre as séries de Taylor e alguns tópicos de probabilidade, incluindo funções características. Com a revisão de assuntos tão importantes, construiu-se o conhecimento necessário para entender e demonstrar várias propriedades das funções características e do Teorema Central do Limite. As funções características geram os momentos e sempre estão definidas para qualquer variável aleatória, ao contrário do que acontece com as funções geradoras de momentos, pois esta última, pode não ter esperança finita para um determinado modelo. Além do que, com as funções características pode-se encontrar a função de probabilidade ou de densidade de um determinado modelo a partir da fórmula da inversão, mesmo que nem sempre seja fácil fazer isso.

O Teorema Central do Limite possui várias versões e aqui neste trabalho foram apresentadas seis delas. A versão clássica do Teorema Central do Limite é a mais conhecida a e ela diz que a variável Z_n converge para uma normal padrão, desde que a sequência de variáveis aleatórias seja i.i.d. Quando essa sequência de variáveis aleatórias tem distribuição de Bernoulli, então tem-se a primeira versão do Teorema Central do Limite, dada por De Moivre e Laplace, que nada mais é do que um caso particular da versão clássica.

A condição de Lindeberg diz que a influência da maior variância é muito pequena em relação a soma das variâncias quando o n é suficientemente grande, ou seja, a maior variância não afeta a divisão pelas somas das variâncias quando o n for suficientemente grande. A partir da condição de Lindeberg, demonstra-se a versão clássica do Teorema Central do Limite, a versão de De Moivre e Laplace e a versão de Liapunov. As versões do Teorema Central do Limite de Liapunov e de Hájek-Sidak são utilizadas em situações mais especificas, por exemplo: para momentos finitos de ordem maior do que dois e na análise de regressão. A versão mais geral do Teorema Central do Limite é a de Lindeberg, e nessa versão desse teorema as v.a's não precisam ser identicamente distribuídas, como acontece nas outras versões. A extensão da versão clássica do Teorema Central do Limite para o caso p-dimensional diz que uma sequência de vetores aleatórios i.i.d converge para uma $N(0, \sum)$. Cada uma das versões do Teorema Central do Limite, apresentadas neste trabalho, são enunciadas e demonstradas de maneiras diferentes, mas, todas elas dizem que somas de variáveis aleatórias padronizadas convergem em distribuição para a normal padrão, uma das mais importantes distribuições de probabilidade.

Referências

ALVES, J. E. C. *Teorema Central do Limite: Compreendendo e Aplicando*. Dissertação (Mestrado) — PROFMAT - UFRB, 2016. Citado na página 31.

ASH, R. B. *Basic Probability Theory*. [S.l.]: Dover Publications, 2008. Citado na página 23.

ASH, R. B.; DOLÉANS-DADE, C. A. *Probability & Measure Theory*. 2a. ed. [S.l.]: Harcourt Academic Press, 2000. Citado na página 31.

CORDEIRO, G. M. Introdução à teoria assintótica. In: *XXII Colóquio Brasileiro de Matemática*. [S.l.: s.n.], 1999. Citado na página 31.

DANTAS, C. A. B. *Probabilidade: Um Curso Introdutório*. 3a. ed. [S.l.]: Edusp, 2013. Citado 2 vezes nas páginas 13 e 14.

DURRET, R. *Probability: Theory and Examples*. 3a. ed. [S.l.]: Duxbury, 2005. Citado na página 31.

IEZZI, G. Fundamentos de matemática elementar, vol. 6: Complexos. *Polinomios, Equaçoes (Atual, Sao Paulo, 1977),*, 2004. Citado na página 51.

JAMES, B. R. *Probabilidade: um curso em nível intermediário*. [S.l.]: IMPA, 2002. (Projeto Euclides). Citado 5 vezes nas páginas 11, 13, 14, 23 e 33.

LEHMANN, E. L. *Elements of Large-Sample Theory*. [S.l.]: Springer, 1999. Citado na página 31.

LIMA, E. L. *Curso de Análise*. [S.l.]: IMPA, 2002. v. 1. Citado na página 8.

MAGALHÃES, M. N. *Probabilidade e Variáveis Aleatórias*. 2a. ed. [S.l.]: Edusp, 2006. Citado 5 vezes nas páginas 11, 13, 14, 23 e 30.

ROSS, S. *Probabilidade: Um curso moderno com aplicações*. 8a. ed. [S.l.]: Bookman, 2010. Citado na página 14.

ROSS, S. M. *Introduction to probability models*. [S.l.]: Academic press, 2014. Citado 2 vezes nas páginas 13 e 14.

THOMAS, G. B.; FINNEY, R. *Cálculo*. [S.l.]: Ltc, 1975. Citado na página 8.

WAIANDT, E. R. *Alguns Teoremas Limite para Sequências de Variáveis Aleatórias*. Dissertação (Mestrado) — Programa de Pós-graduação em Matemática - Universidade Federal do Espírito Santo, 2014. Citado na página 31.

Apêndices

APÊNDICE A – Primeiro apêndice

A.1 Números Complexos

O matemático Tartaglia teve uma contribuição muito importante para a história dos números complexos, pois, certa vez ele foi desafiado a resolver a equação $x^3+bx+c=0$, e depois de certo tempo, o Tartaglia conseguiu deduzir uma fórmula para resolver equações desse tipo. Mas surgiu um problema, que na época ficou muito famoso, que era a equação $x^3-15x-4=0$, onde uma das raízes é 4, quando esta equação era submetida ao método do Tartaglia a solução continha raízes quadradas de números negativos, então, como obter a raiz $x=4$, sendo que na fórmula, aparecem raízes quadradas de números negativos ? A partir daí os matemáticos começaram a investigar essas raízes, tentando obter métodos para lidar com raízes desse tipo. Um dos primeiros matemáticos que desenvolveu métodos para lidar com essas raízes foi Bombelli, já Euler, foi o primeiro a utilizar a notação i para a unidade imaginária. Argand e Gauss foram responsáveis pela representação dos números complexos em um plano.

Aqui será dada apenas uma revisão muito resumida sobre os números complexos, então, para mais detalhes sobre os números complexos consulte (IEZZI, 2004).

A.1.1 Formas de Representar um Número Complexo

As principais formas de representar um número complexo são:

1. **Par ordenado**: $z=(a;b)$;

   Considere os pares ordenados $(a;b)$ e $(c;d)$ com a, b, c, $d \in \mathbb{R}$, então, são válidas as seguintes propriedades:

   a) Igualdade: $(a;b)=(c;d) \iff a=c \text{ e } b=d$;

   b) Adição: $(a;b)+(c;d)=((a+c);(b+d))$, com as seguintes propriedades:

      i. $z_1+(0;0)=z_1$ (elemento neutro);

      ii. $z_1+z_2=z_2+z_1$ (comutativa);

      iii. $z_1+(z_2+z_3)=(z_1+z_2)+z_3$(associativa).

   c) Multiplicação: $(a;b)\cdot(c;d)=((ac-bd);(ad+bc))$, com as seguintes propriedades:

      i. $z_1\cdot(1;0)$ (elemento neutro);

      ii. $z_1\cdot z_2=z_2\cdot z_1$ (comutativa);

      iii. $z_1\cdot(z_2\cdot z_3)=(z_1\cdot z_2)\cdot z_3$ (associativa).

2. **Forma algébrica**: $z = a + bi$, onde: a é a parte real e b é a parte imaginária;

   Quando $b = 0$ temos um número real e quando $a = 0$ e $b \neq 0$ tem-se um número imaginário puro.

3. **Forma trigonométrica ou polar**: $Z = \rho(\cos(\theta) + i\sin(\theta))$.

A.1.2 Conjugado de um Número Complexo

Se $z = a + bi$, defini-se $\overline{z}$, em que $\overline{z} = a - bi$, como sendo o conjugado do número complexo z. Com as seguintes propriedades:

1. $\overline{\overline{z}} = z$;

2. $\overline{z_1} \cdot \overline{z_2} = \overline{z_1 \cdot z_2}$;

3. $\overline{z_1} + \overline{z_2} = \overline{z_1 + z_2}$;

4. $z + \overline{z} = 2a$;

5. $z - \overline{z} = 2bi$;

6. $z = \overline{z} \Rightarrow z \in \mathbb{R}$.

A.1.3 Potências de i

$i^2 = -1$, é o artificio utilizado para resolver equações que não possuem soluções reais, a demonstração deste resultado é feita utilizando a fórmula de multiplicação dos números complexos em forma de par ordenado. logo, $i^2 = i \cdot i = (0;1) \cdot (0;1) = ((0 \cdot 0 - 1 \cdot 1);(0 \cdot 1 + 1 \cdot 0)) = ((0 - 1);(0 + 0)) = (-1;0) = -1 + 0 \cdot i = -1$.

As principais potências de i são: $i^0 = 1$, $i^1 = i$, $i^2 = -1$ e $i^3 = -i$. Para obter o valor de uma potência de i com expoente maior do que três basta fazer a divisão inteira do expoente em questão por quatro, pois, a sequência sempre irá se repetir a cada quatro valores, e fazer i elevado ao resto dessa divisão inteira. Exemplos:

1. i^{73};

   $\frac{73}{4} = 18$, com resto inteiro igual à 1, logo, $i^{73} = i^1 = i$.

2. i^{100}.

   $\frac{100}{4} = 25$, com resto inteiro igual à 0, logo, $i^{100} = i^0 = 1$.

A.1.4 Plano de Argand-Gauss

O plano de Argand-Gauss ou plano complexo é uma forma de representar um número complexo em um gráfico, onde o eixo das abscissas será o eixo real e o eixo das ordenadas será o eixo imaginário.

A.1.5 Módulo ou Norma de um Número Complexo

O módulo ou a norma de um número complexo, denotado por $|z|$ ou ρ, representa a distância de um ponto P até a origem $(0,0)$. O módulo ou a norma de um número complexo $z = a + bi$ é dado por

$$|z| = \sqrt{a^2 + b^2} \text{ ou } \rho = \sqrt{a^2 + b^2}.$$

A.2 Desigualdade de Jensen

Teorema A.1 (Desigualdade de Jensen)**.** *Se f é uma função convexa, então*

$$\mathbb{E}(f(x)) \geq f(\mathbb{E}(x)). \tag{A.1}$$

Demonstração. Suponha que f é uma função convexa e que possa ser expandida em séries de Taylor e seja $\mu = \mathbb{E}(x)$, então, a $f(x)$ no ponto $x_0 = \mu$ é dada por

$$\begin{aligned}
f(x) &= f(\mu) + f'(\mu)(x-\mu) \\
f(x) &\geq f(\mu) + f'(\mu)(x-\mu) \\
\mathbb{E}(f(x)) &\geq \mathbb{E}[f(\mu) + f'(\mu)(x-\mu)] \\
\mathbb{E}(f(x)) &\geq \mathbb{E}(f(\mu)) + \mathbb{E}(f'(\mu)(x-\mu)) \\
E(f(x)) &\geq f(\mu) + f'(\mu)\mathbb{E}(x-\mu) \\
E(f(x)) &\geq f(\mu) + f'(\mu)(\mathbb{E}(x)-\mu) \\
E(f(x)) &\geq f(\mu) + f'(\mu)(\mu-\mu) \\
E(f(x)) &\geq f(\mu) \\
E(f(x)) &\geq f(\mathbb{E}(x)).
\end{aligned}$$

■

APÊNDICE B – Segundo apêndice

B.1 Função Par e Função Ímpar

Para identificar a paridade de uma determinada função $f(x)$ pode-se utilizar a análise do gráfico da mesma, pois, uma função par possui o gráfico simétrico em relação ao eixo das ordenadas, já as funções que são ímpares possuem o gráfico simétrico em relação a origem do plano cartesiano. Caso uma função não tenha o gráfico com as características anteriores ela é chamada de função sem paridade ou função nem par nem ímpar. Uma alternativa mais simples para identificar a paridade de uma função é feita a partir das seguintes relações:

1. Se $f(x) = f(-x)$, então a função é par;
2. Se $f(x) = -f(-x) \Rightarrow -f(x) = f(-x)$, então a função é ímpar.

No caso das integrais, onde os limites de integração são simétricos, é útil identificar a paridade da função do integrando, pois, muitas integrais são complicadas e outras não possuem uma primitiva, então, identificar a paridade de funções deste tipo facilita a resolução dessas integrais. Então:

1. Se $f(x)$ é uma função par $\Rightarrow \displaystyle\int_{-a}^{a} f(x)dx = 2 \cdot \int_{0}^{a} f(x)dx$;
2. Se $f(x)$ é uma função ímpar $\Rightarrow \displaystyle\int_{-a}^{a} f(x)dx = 0$.

Demonstração. Considere a integral: $\displaystyle\int_{-a}^{a} f(x)dx = \int_{-a}^{0} f(x)dx + \int_{0}^{a} f(x)dx$.

Por substituição, faça: $x = -t \Rightarrow dx = -dt$ na primeira integral do lado direito da igualdade e quando $x = 0 \Rightarrow t = 0$, quando $x = -a \Rightarrow t = a$ e na segunda integral faça: $x = t \Rightarrow dx = dt$, aqui os limites de integração permanecem os mesmos. Então:

$$\int_{-a}^{a} f(x)dx = -\int_{a}^{0} f(-t)dt + \int_{0}^{a} f(t)dt.$$

invertendo os limites de integração da primeira integral do lado direito da igualdade, o sinal da integral fica invertido, logo

$$\int_{-a}^{a} f(x)dx = \int_{0}^{a} f(-t)dt + \int_{0}^{a} f(t)dt.$$

Se $f(t)$ for uma função par, então $f(t) = f(-t)$, portanto:

$$\int_{-a}^{a} f(x)dx = \int_{0}^{a} f(t)dt + \int_{0}^{a} f(t)dt = 2\cdot\int_{0}^{a} f(t)dt = 2\cdot\int_{0}^{a} f(x)dx.$$

Se $f(t)$ for uma função ímpar, então $f(t) = -f(-t)$, portanto:

$$\int_{-a}^{a} f(x)dx = \int_{0}^{a} f(-t)dt - \int_{0}^{a} f(-t)dt = 0.$$

$\blacksquare$

B.2 Limite Fundamental Exponencial

O limite fundamental exponencial é dado por:

$$\lim_{n\to\infty}\left(1 + \frac{b}{n}\right)^{n} = e^{b}. \tag{B.1}$$

Demonstração. Para esta demonstração recorde que: $e^{\ln x} = x$, logo:

$$\lim_{n\to\infty}\left(1 + \frac{b}{n}\right)^{n} = \lim_{n\to\infty}\left(e^{\ln\left(1 + \frac{b}{n}\right)}\right)^{n}$$

agora, usando a propriedade de potência de potência, segue que:

$$\lim_{n\to\infty}\left(1 + \frac{b}{n}\right)^{n} = \lim_{n\to\infty} e^{n\cdot\ln\left(1 + \frac{b}{n}\right)}.$$

Como a exponencial é uma função contínua, este limite pode ser reescrito da seguinte forma:

$$\lim_{n\to\infty}\left(1 + \frac{b}{n}\right)^{n} = e^{\lim\limits_{n\to\infty}\left(n\cdot\ln\left(1 + \frac{b}{n}\right)\right)},$$

percebe-se que existe uma indeterminação do tipo: $\infty\cdot 0$, portanto, em situações deste tipo, pode-se inverter uma das funções e com isso obter uma indeterminação do tipo: $\frac{0}{0}$, fazendo isso, resulta que:

$$\lim_{n\to\infty}\left(1 + \frac{b}{n}\right)^{n} = e^{\lim\limits_{n\to\infty}\left(\frac{\ln\left(1 + \frac{b}{n}\right)}{\frac{1}{n}}\right)},$$

como existe uma indeterminação do tipo: $\frac{0}{0}$, pode-se utilizar a regra de L'Hospital, sendo assim, derivando o numerador e o denominador em relação a n e fazendo as devidas simplificações, chega-se a:

$$\lim_{n\to\infty}\left(1 + \frac{b}{n}\right)^n = e^{\lim\limits_{n\to\infty}\left(\frac{b\cdot n}{n+b}\right)},$$

novamente existe uma indeterminação do tipo: $\frac{\infty}{\infty}$, sendo assim, utilizando mais uma vez a regra de L'Hospital, obtem-se que:

$$\lim_{n\to\infty}\left(1 + \frac{b}{n}\right)^n = e^{\lim\limits_{n\to\infty}(b)},$$

portanto, como b é uma constante, o limite de uma constante é a própria constante, logo:

$$\lim_{n\to\infty}\left(1 + \frac{b}{n}\right)^n = e^b.$$

■

B.3 Fórmula de Euler

A fórmula de Euler é dada por:

$$e^{i\theta} = \cos(\theta) + i\cdot\text{sen}(\theta). \tag{B.2}$$

Para $\theta = \pi$, tem-se a famosa Identidade de Euler, dada por: $e^{i\pi} = -1 \Rightarrow e^{i\pi} + 1 = 0$.

Demonstração. Utilizando a série de Taylor da exponencial com $x = i\theta$, tem-se que:

$$e^{i\theta} = 1 + \frac{i\theta}{1!} + \frac{i^2\theta^2}{2!} + \frac{i^3\theta^3}{3!} + \frac{i^4\theta^4}{4!} + \cdots$$

calculando as potências de i e expandindo um pouco mais os termos desta série e organizando-os, com os termos que não dependem de i de um lado e do outro os termos que dependem das potências de i colocadas em evidência, obtem-se que:

$$e^{i\theta} = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots + i\cdot\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots\right).$$

Note que do lado direito da igualdade aparecem as séries de Taylor do cosseno e do seno, respectivamente, portanto, pode-se reescrever esta série da seguinte forma:

$$e^{i\theta} = \cos(\theta) + i\cdot\text{sen}(\theta).$$

Para $\theta = \pi$, tem-se que $\cos(\pi) = -1$ e $\text{sen}(\pi) = 0$, portanto, $e^{i\pi} = -1 \Rightarrow e^{i\pi} + 1 = 0$.

■

B.4 Funções Seno e Cosseno Complexas

A função seno e a função cosseno complexas são dadas respectivamente por:

$$\operatorname{sen}(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i} \quad \text{e} \quad \cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2}. \tag{B.3}$$

Demonstração. Para esta demonstração recorde que a função seno é ímpar, ou seja $-f(x) = f(-x)$, e a função cosseno é par logo $f(x) = f(-x)$, sendo assim, pela fórmula de Euler, dada em (Apêndice B, relação B.2), pode-se escrever que

$$e^{-i\theta} = \cos(-\theta) + i\operatorname{sen}(-\theta) \Rightarrow e^{-i\theta} = \cos(\theta) - i\operatorname{sen}(\theta),$$

com este resultado e com a fórmula de Euler, chega-se ao seguinte sistema de equações

$$\begin{cases} e^{i\theta} = & \cos(\theta) + i\operatorname{sen}(\theta) \\ e^{-i\theta} = & \cos(\theta) - i\operatorname{sen}(\theta) \end{cases}$$

Somando estas duas equações, segue que: $e^{i\theta} + e^{-i\theta} = \cos(\theta) + i\operatorname{sen}(\theta) + \cos(\theta) - i\operatorname{sen}(\theta) \Rightarrow e^{i\theta} + e^{-i\theta} = 2 \cdot \cos(\theta) \Rightarrow \cos(\theta) = \frac{e^{i\theta}+e^{-i\theta}}{2}$. Subtraindo estas duas equações, segue que: $e^{i\theta} - e^{-i\theta} = \cos(\theta) + i\operatorname{sen}(\theta) - \cos(\theta) + i\operatorname{sen}(\theta) \Rightarrow e^{i\theta} - e^{-i\theta} = 2i \cdot \operatorname{sen}(\theta) \Rightarrow \operatorname{sen}(\theta) = \frac{e^{i\theta}-e^{-i\theta}}{2i}$.

■

B.5 Integral Gaussiana

A Integral Gaussiana é dada por:

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx = \sqrt{2\pi}. \tag{B.4}$$

Demonstração. Para esta demonstração, faça $\mathrm{I} = \displaystyle\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx.$

Elevando ambos os lados da igualdade ao quadrado, obtem-se que:

$$\mathrm{I}^2 = \left(\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx\right)^2$$

Reescrevendo essa integral como sendo o produto de duas integrais, logo:

$$\mathrm{I}^2 = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx \cdot \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx.$$

Agora, por substituição, faça: $x = y \Rightarrow dx = dy$, na segunda integral do lado direito da igualdade, note que os limites de integração permanecem os mesmos, logo:

$$\mathrm{I}^2 = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx \cdot \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}}\, dy.$$

Reescrevendo essas integrais como sendo uma integral dupla, daí segue que:
$\mathrm{I}^2 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} \cdot e^{-\frac{y^2}{2}}\, dx\, dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{(x^2+y^2)}{2}}\, dx\, dy.$

Por coordenadas polares faça:

$x = r\cos\theta$, $y = r\,\mathrm{sen}\,\theta$, $x^2 + y^2 = r^2$ e $dxdy = rdrd\theta$, com limites de integração dados por $0 < r < \infty$ e $0 < \theta < 2\pi$. Logo, fazendo essas mudanças de variáveis e simplificando, portanto:

$\mathrm{I}^2 = \int_0^{2\pi}\int_0^{+\infty} r\ e^{-\frac{r^2}{2}}\ dr\ d\theta = \int_0^{2\pi} d\theta = 2\pi$, aplicando a raiz quadrada em ambos os lados da igualdade e lembrando que θ é positivo, então $\mathrm{I} = \sqrt{2\pi}$, mas, I $= \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx$, portanto: $\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx = \sqrt{2\pi}.$

■

B.6 Integral de Dirichlet

A Integral de Dirichlet é dada por:

$$\int_0^{+\infty} \frac{\mathrm{sen}\,x}{x}\, dx = \frac{\pi}{2}. \tag{B.5}$$

Demonstração. Para demonstrar este resultado, primeiro deve-se mostrar que

$$\int_0^{+\infty} e^{-xy} dy = \frac{1}{x},$$

sendo assim, multiplicando o numerador e o denominador por x, com $x > 0$, então: $\frac{1}{x}\int_0^{+\infty} xe^{-xy}dy$, como $x > 0$ e é uma constante, percebe-se que esta integral é a fdp de uma exponencial com parâmero x, sendo assim, segue que: $\frac{1}{x}\int_0^{+\infty} xe^{-xy}dy = \frac{1}{x}$.

Com esse resultado, pode-se usá-lo para demonstrar o valor da integral B.5, logo:

$$\int_0^{+\infty} \frac{\mathrm{sen}\,x}{x}\, dx = \int_0^{+\infty} \mathrm{sen}\,x \cdot \frac{1}{x}\, dx,$$

mas, $\frac{1}{x} = \int_0^{+\infty} e^{-xy}dy$, sendo assim, pode-se reescrever o integrando de B.5 como sendo:

$$\int_0^{+\infty} \frac{\mathrm{sen}\,x}{x}\, dx = \int_0^{+\infty} \mathrm{sen}\,x \cdot \int_0^{+\infty} e^{-xy}\, dy\, dx,$$

mudando a ordem de integração, segue que:

$$\int_0^{+\infty} \frac{\mathrm{sen}\,x}{x}\, dx = \int_0^{+\infty}\int_0^{+\infty} \mathrm{sen}\,x \cdot e^{-xy}\, dx\, dy,$$

a integral mais interna pode ser resolvida facilmente utilizando a integração por partes ou a integração tabular e utilizando o teorema do confronto para avaliá-la nos limites de integração, sendo assim:

$$\int_0^{+\infty} \frac{\operatorname{sen} x}{x}\, dx \;=\; \int_0^{+\infty} \frac{1}{y^2 \;+\; 1}\, dy,$$

esta integral é conhecida, trata-se do $\arctan y$, então, avaliando-a com estes limites de integração, segue que:

$$\int_0^{+\infty} \frac{\operatorname{sen} x}{x}\, dx \;=\; \frac{\pi}{2}.$$

■