Euclésio Rangel Waiandt

Alguns Teoremas Limites para Sequências de Variáveis Aleatórias

Vitória-ES, Brasil

Outubro de 2014

Euclésio Rangel Waiandt

Alguns Teoremas Limites para Sequências de Variáveis Aleatórias

Dissertação submetida como parte dos requisitos para a obtenção do título de Mestre em Matemática pelo Programa de Pós-Graduação em Matemática da Universidade Federal do Espírito Santo.

Universidade Federal do Espírito Santo – UFES
Departamento de Matemática
Programa de Pós-Graduação em Matemática

Orientador: Fábio Júlio da Silva Valentim

Vitória-ES, Brasil
Outubro de 2014

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Central da Universidade Federal do Espírito Santo, ES, Brasil)

W138a

Waiandt, Euclésio Rangel, 1982-
Alguns teoremas limites para sequências de variáveis aleatórias / Euclésio Rangel Waiandt. – 2014.
104 f.

Orientador: Fábio Júlio da Silva Valentim.
Dissertação (Mestrado em Matemática) – Universidade Federal do Espírito Santo, Centro de Ciências Exatas.

1. Variáveis aleatórias. 2. Convoluções (Matemática). 3. Distribuição (Teoria da probabilidade). I. Valentim, Fábio Júlio da Silva. II. Universidade Federal do Espírito Santo. Centro de Ciências Exatas. III. Título.

CDU: 51

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
Centro de Ciências Exatas
Programa de Pós-Graduação em Matemática

"Alguns Teoremas Limites Para Sequência de Variáveis Aleatórias"

Euclésio Rangel Waiandt

Dissertação submetida ao Programa de Pós-Graduação em Matemática da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do título de Mestre em Matemática.

Aprovada em 16/10/2014 por:

Fábio Júlio da Silva Valentim – UFES

Ana Patrícia Carvalho Gonçalves – PUC/RJ

Daniela Paula Demuner – UFES

Aos meus filhos, pois a existência deles
é o que me faz querer evoluir como ser humano.

Agradecimentos

À minha esposa, Ervânia, por ter confiado na minha escolha em fazer o Mestrado em Matemática e por ter me apoiado e motivado durante o curso.

Aos meus filhos, Êndrio e Eduarda, por existirem e serem a motivação para seguir sempre em frente.

Ao meu orientador, Prof. Dr. Fábio Júlio da Silva Valentim, por ter confiado na minha capacidade em fazer este trabalho.

Aos meus pais, Helmut e Marineuza, pelos valores que me ensinaram, os quais contribuem nas escolhas que faço.

Resumo

O Teorema Central do Limite e a Lei dos Grandes Números estão entre os mais importantes resultados da teoria da probabilidade. O primeiro deles busca condições sob as quais $\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}}$ converge em distribuição para a distribuição normal com parâmetros 0 e 1, quando n tende ao infinito, onde S_n é a soma de n variáveis aleatórias independentes. Ao mesmo tempo, o segundo estabelece condições para que $\frac{S_n - \mathbb{E}S_n}{n}$ convirja a zero, ou equivalentemente, para que $\frac{S_n}{n}$ convirja para a esperança das variáveis aleatórias, caso elas sejam identicamente distribuídas. Em ambos os casos as sequências abordadas são do tipo $\frac{S_n + b_n}{a_n}$, onde $a_n > 0$ e b_n são constantes reais. Caracterizar os possíveis limites de tais sequências é um dos objetivos dessa dissertação, já que elas não convergem exclusivamente para uma variável aleatória degenerada ou com distribuição normal como na Lei dos Grandes Números e no Teorema Central do Limite, respectivamente. Assim, somos levados naturalmente ao estudo das distribuições infinitamente divisíveis e estáveis, e os respectivos teoremas limites, e este vem a ser o objetivo principal desta dissertação. Para as demonstrações dos teoremas utiliza-se como estratégia principal a aplicação do método de Lyapunov, o qual consiste na análise da convergência da sequência de funções características correspondentes às variáveis aleatórias. Nesse sentido, faremos também uma abordagem detalhada de tais funções neste trabalho.

Palavras-chave: Sequências de Variáveis Aleatórias. Funções Características. Distribuições Infinitamente Divisíveis e Estáveis. Teoremas Limites.

Abstract

The Central Limit Theorem and the Law of Large Numbers are among the most important results of probability theory. The first one seeks conditions under which $\frac{s_n-\mathbb{E}S_n}{\sqrt{VarS_n}}$ converges in distribution to the normal distribution with parameters 0 and 1, when n tends to infinity, where S_n is the sum of n independent random variables. At the same time, the second gives conditions such that $\frac{S_n-\mathbb{E}S_n}{n}$ converges to zero, or equivalently, that $\frac{S_n}{n}$ converges to the expectation of the random variables, if they are identically distributed. In both cases, the sequences discussed are of the type $\frac{S_n+b_n}{a_n}$, where $a_n > 0$ and b_n are real constants. Characterizing the possible limits of such sequences is one of the goals of this dissertation, as they not only converge to a degenerated random variable or a random variable with normal distribution, as the Law of Large Numbers and the Central Limit Theorem, respectively. Thus, we are naturally led to the study of infinitely divisible and stable distributions and their limits theorems. This becomes the main objective of this dissertation. In order to prove the theorems, the method of Lyapunov is applied as the main strategy, which analyzes the convergence of the sequence of characteristic functions related to the random variables. So we carry out a detailed approach of such functions in this research.

Key-words: Sequences of Random Variables. Characteristic Functions. Infinitely Divisible and Stable Distributions. Limits theorems.

Lista de símbolos

$\mathbb{E}X$	Esperança da variável aleatória X
$VarX$	Variância da variável aleatória X
$X_1 \stackrel{D}{=} X_2$	X_1 é igual a X_2 em distribuição
$X_n \xrightarrow{D} X$	X_n converge a X em distribuição
$X_n \xrightarrow{P} X$	X_n converge a X em probabilidade
$X_n \searrow X$	$X_n \geq X$ e $X_n \stackrel{n\to\infty}{\longrightarrow} X$
$X_n \nearrow X$	$X_n \leq X$ e $X_n \stackrel{n\to\infty}{\longrightarrow} X$
$\dot{\bigcup}$	União disjunta
$*$	Convolução
$f(x+)$	$\lim\limits_{y\to x+} f(y)$
$f(x-)$	$\lim\limits_{y\to x-} f(y)$
$f(\infty)$	$\lim\limits_{y\to\infty} f(y)$
$f(-\infty)$	$\lim\limits_{y\to-\infty} f(y)$
$\mathcal{L}$	Família de índices
$\mathcal{B}(\mathbb{R})$	$\sigma-$álgebra de Borel
$\lfloor x \rfloor$	Parte inteira de x
$\liminf\limits_n a_n$	$\lim\limits_{N\to\infty} \inf\limits_{n\geq N} a_n$
$\limsup\limits_n a_n$	$\lim\limits_{N\to\infty} \sup\limits_{n\geq N} a_n$

Sumário

1 Introdução

A origem da teoria da probabilidade se deu em meados do século XVII e está atrelada à forma como Pascal e Fermat lidavam com jogos de azar, a qual não se encaixava nos conceitos matemáticos daquela época. Como as ciências naturais até então eram pouco desenvolvidas, os jogos de azar e os problemas de demografia foram durante muito tempo o seu único objeto de estudo, e os problemas eram resolvidos com o uso de aritmética elementar e métodos combinatórios.

Com a evolução das ciências naturais foi necessário o surgimento de métodos mais sofisticados e eficazes para o tratamento das informações. Na física molecular, por exemplo, descobriu-se que cada substância é composta por pequenas partículas que estão em constante movimento, o que ficou conhecido como movimento browniano. Esta nomenclatura é uma homenagem ao britânico Robert Brown, que foi o primeiro a observar tal fato, em 1827. Em 1905 Albert Einstein propôs uma descrição matemática do movimento a partir das leis da física, mas somente na década de 20 é que tal teoria foi formalizada por Norbert Wiener.

Para termos uma ideia intuitiva do que é o movimento browniano, consideremos uma partícula que a cada unidade de tempo Δt se move para a direita ou para a esquerda, uma distância Δx com probabilidade $\frac{1}{2}$ para ambos os lados. Considere também as variáveis aleatórias independentes $Y_1, Y_2, \cdots$, tais que $P(Y_k = \Delta x) = P(Y_k = -\Delta x) = \frac{1}{2}$, $k = 1, 2, \cdots$. Observe que num tempo t, a partícula salta $\left\lfloor \frac{t}{\Delta t} \right\rfloor$ vezes, onde $\lfloor z \rfloor$ indica a parte inteira de z, e assim, a posição da partícula no instante t é dada por

$$D(t) = Y_1 + \cdots + Y_{\lfloor \frac{t}{\Delta t} \rfloor}.$$

Como $\mathbb{E}D(t) = 0$ e $\mathbb{E}D^2(t) = \left\lfloor \frac{t}{\Delta t} \right\rfloor (\Delta x)^2$, tomamos $\Delta x = \sqrt{\Delta t}$ para que a variância de $D(t)$ seja sempre finita e diferente de zero. Daí, tomando $\Delta t = \frac{1}{n}$, teremos $\Delta x = \frac{1}{\sqrt{n}}$, e logo, $P(Y_k = \frac{1}{\sqrt{n}}) = P(Y_k = -\frac{1}{\sqrt{n}}) = \frac{1}{2}$. Assim, $D(t)$ terá a mesma distribuição que

$$X^n(t) = \frac{Z_1 + \cdots + Z_{\lfloor nt \rfloor}}{\sqrt{n}},$$

onde $P(Z_k = 1) = P(Z_k = -1) = \frac{1}{2}$. Agora, observando que

$$X^n(t) = \sqrt{t} \left(\frac{Z_1 + \cdots + Z_{\lfloor nt \rfloor}}{\sqrt{nt}} \right),$$

basta aplicarmos o Teorema Central do Limite e teremos $X^n(t) \xrightarrow{D} N(0, t)$ quando n vai para o infinito.

A partir daí define-se que um movimento browniano $(B_t)_{t\geq 0}$ é um processo estocástico $B : \Omega \times [0, \infty) \to \mathbb{R}$, tal que valem as três condições abaixo: (BREIMAN, 1992, Capítulo 12)

1. Para quase todo $\omega \in \Omega$ fixado, $B(\omega, \cdot) : [0, \infty) \to \mathbb{R}$ é uma função contínua com $B(\omega, 0) = 0$;

2. Os incrementos $B_t - B_s$, $0 \leq s < t$, são tais que $B_t - B_s \sim N(0, t - s)$;

3. Incrementos em intervalos de tempo disjuntos são independentes, isto é, $B_{t_4} - B_{t_3}$ e $B_{t_2} - B_{t_1}$ são independentes sempre que $0 \leq t_1 < t_2 \leq t_3 < t_4$.

Na construção do movimento browniano que fizemos acima fica claro que as noções elementares da teoria da probabilidade não são suficientes para resolver tal problema. Foi necessário construir uma sequência de variáveis aleatórias e aplicar o Teorema Central do Limite a ela. Pascal e Fermat já tinham a ideia de que propriedades importantes podiam aparecer a partir da análise de um grande número de eventos aleatórios, mas somente na década de 1910, quando E. Borel percebeu que seria interessante conectar a teoria da probabilidade com a teoria métrica das funções de uma variável real, é que tal abordagem se tornou efetiva. Na década seguinte, Khintchin, Kolmogorov, Paul Lévy, entre outros, desenvolveram essas ideias, fazendo com que a teoria da probabilidade passasse a destacar-se pela rapidez do seu desenvolvimento e pela sua grande aplicabilidade.

Além do Teorema Central do Limite, que utilizamos acima, um outro teorema muito importante a respeito do limite de uma sequência de variáveis aleatórias é a Lei dos Grandes Números, que além de ser fundamental na teoria de processos estocásticos, pode ser utilizado também para a obtenção de resultados em outras áreas do conhecimento. Vejamos uma pequena aplicação de tal teorema à análise: sendo f uma função contínua em $[0, 1]$, e para cada $n \in \mathbb{N}$, $p_n(x) = \sum\limits_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k}$ o polinômio de Bernstein de f, então $\{p_n\}_{n\geq 1}$ converge pontualmente a f em $[0, 1]$, quando $n \to \infty$. De fato, fixado $x \in [0, 1]$, considere as variáveis aleatórias independentes $X_1, X_2, \cdots$, todas com distribuição de Bernoulli com probabilidade x de sucesso, isto é, $P(X_n = 1) = x$ e $P(X_n = 0) = 1 - x$. Sendo $S_n = X_1 + \cdots + X_n$, temos da teoria básica de probabilidade que

$$P(S_n = k) = \binom{n}{k}x^k(1-x)^{n-k}, \quad 0 \leq k \leq n.$$

Logo, $P(\frac{S_n}{n} = \frac{k}{n}) = \binom{n}{k}x^k(1-x)^{n-k}$, donde segue que $\mathbb{E}(\frac{S_n}{n}) = \sum\limits_{k=0}^{n} \frac{k}{n}\binom{n}{k}x^k(1-x)^{n-k}$, e consequentemente,

$$\mathbb{E}f\left(\frac{S_n}{n}\right) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k} = p_n(x).$$

Como $X_1, X_2, \cdots$ têm esperança igual a x, temos pela Lei Fraca dos Grandes Números de Khintchin que $\frac{S_n}{n} \xrightarrow{P} x$, quando $n \to \infty$, e como f é contínua em $[0,1]$, segue que $f\left(\frac{S_n}{n}\right) \xrightarrow{P} f(x)$, quando $n \to \infty$ (CHUNG, 2001, Seção 4.1). Além disso, como $[0,1]$ é compacto, f é uniformemente contínua em $[0,1]$, e como $|\frac{S_n}{n}| \leq 1$ segue que $\{f\left(\frac{S_n}{n}\right)\}_{n\geq 1}$ é limitada. Assim, pelo Teorema da Convergência Dominada segue que $\mathbb{E}f\left(\frac{S_n}{n}\right) \overset{n\to\infty}{\longrightarrow} \mathbb{E}f(x) = f(x)$, isto é, $p_n(x) \overset{n\to\infty}{\longrightarrow} f(x)$.

No exemplo acima mostramos que p_n converge pontualmente a f, e usando a desigualdade de Chebyshev, poderíamos mostrar também que a convergência é uniforme (CHUNG, 2001, Teorema 5.5.4). Mas esse não é o nosso objetivo aqui, bem como não estávamos preocupados em definir rigorosamente o movimento browniano acima. Na verdade, a ideia é ressaltar a importância que têm os teoremas limites para sequências de variáveis aleatórias, destacando o Teorema Central do Limite e a Lei dos Grande Números, como fizemos. Estes são dois dos mais célebres resultados da teoria de probabilidade. Informalmente, o primeiro assegura que, sob condições bem gerais, médias amostrais de variáveis aleatórias, devidamente normalizadas em média e variância, convergem em distribuição para a normal com parâmetros 0 e 1 quando o tamanho da amostra tende ao infinito. Da mesma forma, o segundo assegura que a média de um grande número de variáveis aleatórias normalizadas pelas suas respectivas médias, converge a zero em probabilidade. Em ambos os casos, as sequências que aparecem são da forma $\frac{S_n+b_n}{a_n}$, onde $a_n > 0$ e b_n são sequências de números reais, e somos levados naturalmente à seguinte questão: além da distribuição degenerada e da distribuição normal, quais são as outras distribuições que podem aparecer como limite das sequências $\{\frac{S_n+b_n}{a_n}\}_{n\geq 1}$? Na procura da resposta de tal questão somos levados a estudar as classes das distribuições infinitamente divisíveis e estáveis, bem como os correspondentes teoremas limites, o que vem a ser o objetivo principal desta dissertação.

O segundo capítulo da dissertação preocupar-se-á em fazer as definições e enunciar os principais teoremas que serão utilizados durante a dissertação. As referências para tal capítulo são (CHUNG, 2001) ou (JAMES, 2006).

O terceiro capítulo tratará dos diferentes tipos de convergência que temos para uma sequência de variáveis aleatórias e da relação que existe entre elas, enunciando-se os principais teoremas a respeito. O roteiro descrito é basicamente o encontrado em (JAMES, 2006), exceto a seção que trata das famílias de medidas de probabilidade rígidas e relativamente compactas. Tal teoria pode ser encontrada em (SHIRYAEV, 1996). Nesse capítulo são enunciados o Teorema de Poisson e algumas versões do Teorema Central do Limite e da Lei dos Grandes Números. Para a demonstração deles utiliza-se o método de Lyapunov, o qual consiste na análise da convergência da sequência de funções características e se destaca pela sua simplicidade e aplicabilidade.

No quarto capítulo é que estudaremos as classes das distribuições infinitamente divisíveis e estáveis, e os correspondentes teoremas limites. Nesse intuito faremos primeiro um estudo detalhado sobre as funções características, pois isto se faz necessário para identificar as distribuições infinitamente divisíveis. As demonstrações feitas nesse capítulo têm como referência principal (CHUNG, 2001) e (SHIRYAEV, 1996).

Os resultados aqui demonstrados não podem ser todos obtidos a partir de uma única referência citada. A maior parte deles está enunciada e demonstrada em (CHUNG, 2001). No entanto, tal obra não aborda as distribuições estáveis e nem a convergência dos arranjos triangulares para as distribuições infinitamente divisíveis. Tais resultados podem ser encontrados em (SHIRYAEV, 1996). Ressalto também que as demonstrações nesta dissertação foram feitas tentando sempre detalhar e simplificar os cálculos o máximo possível, visando que o texto fique acessível também para alunos de graduação. Em particular, a demonstração do Teorema 4.2.3 aqui feita a partir do Teorema 4.2.1 é muito mais simples do que a feita em (CHUNG, 2001), já que lá o Teorema 4.2.1 não é enunciado. Outros resultados, como por exemplo as Proposições 4.1.1 e 4.2.2, nos livros são apenas enunciados, enquanto que aqui são também demonstrados, e o Exemplo 3.2.1 foi por mim elaborado para enfatizar que as somas normalizadas nem sempre convergem para a distribuição normal.

Aos interessados, resultados similares ou que complementam os que aqui estão, podem ser encontrados em (GNEDENKO; KOLMOGOROV, 1954).

2 Noções Preliminares

Para compreender a teoria aqui descrita é necessário o conhecimento prévio da teoria básica de probabilidade. Nesse sentido, este capítulo preocupar-se-á em fazer as definições e enunciar os principais teoremas que serão utilizados durante a dissertação, dando-se destaque ao teorema que garante a existência de infinitas variáveis aleatórias identicamente distribuídas, já que tal fato é que dá sentido à noção de distribuição infinitamente divisível e estável. Por uma questão de objetividade algumas demonstrações serão ocultadas, mas poderão ser encontradas em (CHUNG, 2001) ou (JAMES, 2006).

2.1 Espaços de Probabilidade e Variáveis Aleatórias

Como já foi dito na introdução, quando E. Borel percebeu que seria interessante conectar a teoria da probabilidade com a teoria métrica das funções de uma variável real é que a teoria da probabilidade passou a destacar-se como ciência. Nascia assim o conceito de variável aleatória, o qual abordaremos nesta seção. Como já foi dito também, por uma questão de objetividade, apenas enunciaremos as proposições e teoremas, sem fazer as respectivas demonstrações, as quais podem ser encontradas em (JAMES, 2006) e (CHUNG, 2001).

Fixado um conjunto Ω não vazio, uma $\sigma-$álgebra $\mathcal{F}$ de Ω é uma coleção de subconjuntos de Ω satisfazendo:

i. $\emptyset \in \mathcal{F}$;

ii. $A^c \in \mathcal{F}$ para todo $A \in \mathcal{F}$;

iii. $A_j \in \mathcal{F}$, $j \in \mathbb{N}$, implica que $\bigcup\limits_{j\in\mathbb{N}} A_j \in \mathcal{F}$.

O par $(\Omega, \mathcal{F})$ é chamado de espaço mensurável.

Dado um espaço mensurável $(\Omega, \mathcal{F})$, uma **medida de probabilidade** é uma função $P : \mathcal{F} \to [0, 1]$ tal que:

i. $P(A) \geq 0$;

ii. $P(\Omega) = 1$;

iii. Se $A_1, A_2, \cdots$ são elementos disjuntos pertencentes a $\mathcal{F}$, então $P(\bigcup\limits_{j\in\mathbb{N}} A_j) = \sum\limits_{j\in\mathbb{N}} P(A_j)$.

Em particular, temos $P(A^c) = 1 - P(A)$ para todo $A \in \mathcal{F}$.

Definição 2.1.1 (Espaço de Probabilidade)**.** *O trio* $(\Omega, \mathcal{F}, P)$ *é chamado de espaço de probabilidade.*

Em Probabilidade comumente refere-se ao conjunto Ω como *espaço amostral* e aos elementos da σ-álgebra como *eventos aleatórios*.

A propriedade acima que diz que $P(\bigcup_{j\in\mathbb{N}} A_j) = \sum_{j\in\mathbb{N}} P(A_j)$ é denominada $\sigma-$aditividade da medida P. Quando acontece de termos apenas $P(\bigcup_{j=1}^{n} A_j) = \sum_{j=1}^{n} P(A_j)$ dizemos que a probabilidade P é finitamente aditiva. Note que verificar se uma medida P é finitamente aditiva é mais simples do que verificar a σ-aditividade. Nesse sentido temos uma valiosa proposição:

Proposição 2.1.1. *Uma probabilidade finitamente aditiva é uma probabilidade (σ-aditiva) se, e somente se, é contínua no vazio, isto é, se $P(A_n) \searrow 0$ quando $A_n \searrow \emptyset$.*

Definição 2.1.2 (Variável Aleatória)**.** *Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade. Uma função $X : \Omega \to \mathbb{R}$ será chamada de variável aleatória se for uma função mensurável, isto é, se para todo $x \in \mathbb{R}$, $[X \leq x] = \{\omega \in \Omega : X(\omega) \leq x\}$ é um evento aleatório.*

De outro modo, denotando por $\mathcal{B}(\mathbb{R})$ a σ-álgebra de Borel para $\mathbb{R}$ (isto é, a σ-álgebra gerada pelos intervalos), podemos dizer também que $X : \Omega \to \mathbb{R}$ é uma variável aleatória se $X^{-1}(A) \in \mathcal{F}$ para todo $A \in B_0$, onde B_0 é qualquer família de subconjuntos de $\mathbb{R}$ que gera $\mathcal{B}(\mathbb{R})$, por exemplo $B_0 = \{(-\infty, r] \subset \mathbb{R} : r \in \mathbb{R}\}$ ou $B_0 = \{(a, b) \subset \mathbb{R} : a, b \in \mathbb{R}\}$.

Definição 2.1.3 (Eventos Aleatórios Independentes)**.** *Os eventos aleatórios $A_1, A_2, \cdots, A_n$ são ditos independentes se*

$$P(\bigcap_{k=1}^{n} A_k) = \prod_{k=1}^{n} P(A_k).$$

Da mesma forma, dizemos que $\{A_\lambda\}_{\lambda\in\mathcal{L}}$ é uma família de eventos aleatórios independentes se para todo $n \in \mathbb{N}$ os eventos aleatórios $A_{\lambda_1}, \cdots, A_{\lambda_n}$ são independentes para quaisquer $\lambda_1, \cdots, \lambda_n \in \mathcal{L}$.

Definição 2.1.4 (Variáveis Aleatórias Independentes)**.** *As variáveis aleatórias $X_1, \cdots, X_n$ são ditas independentes se para todo $B_1, \cdots, B_k \in \mathcal{B}(\mathbb{R})$ tivermos*

$$P(X_1 \in B_1, \cdots, X_n \in B_n) = \prod_{k=1}^{n} P(X_k \in B_k).$$

Informalmente podemos dizer que as variáveis aleatórias são independentes se os eventos aleatórios determinados por variáveis aleatórias distintas, $X_k^{-1}(B)$ com $B \in \mathcal{B}(\mathbb{R})$, são sempre independentes. Daí podemos considerar a definição acima válida também para um número infinito de variáveis aleatórias.

Definição 2.1.5 (Função de Distribuição)**.** *Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade e $X : \Omega \to \mathbb{R}$ uma variável aleatória. Definimos a função de distribuição de X, $F : \mathbb{R} \to \mathbb{R}$, por*

$$F(x) = P[X \leq x].$$

Dada uma função $F : \mathbb{R} \to \mathbb{R}$, temos que ela é uma função de distribuição se, e somente se, F é não-decrescente e contínua à direita com $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0$ e $F(\infty) = \lim\limits_{x \to \infty} F(x) = 1$. Outra propriedade importante é que os pontos de continuidade de uma função de distribuição são densos em $\mathbb{R}$, como mostra a seguinte proposição:

Proposição 2.1.2. *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função monótona. Então, o conjunto dos pontos de descontinuidade de f é enumerável. Em particular, o conjunto dos pontos de descontinuidade de uma função de distribuição é enumerável, enquanto que o conjunto dos pontos de continuidade é denso em $\mathbb{R}$.*

Demonstração. Sendo $x_1 < x_2$ pontos de descontinuidade de f, considere os intervalos $I_{x_1} = \{y \in \mathbb{R};\ f(x_1-) \leq y \leq f(x_1+)\}$ e $I_{x_2} = \{y \in \mathbb{R};\ f(x_2-) \leq y \leq f(x_2+)\}$. Como f é monótona, dado $\tilde{x}$ tal que $x_1 < \tilde{x} < x_2$, temos

$$f(x_1+) \leq f(\tilde{x}) \leq f(x_2-),$$

se f é não decrescente, ou então,

$$f(x_2-) \leq f(\tilde{x}) \leq f(x_1+),$$

se f é não crescente. Segue que I_{x_1} e I_{x_2} são intervalos disjuntos, e como em cada intervalo aberto há um número racional, podemos construir uma função injetiva associando cada x que é ponto de descontinuidade de f a algum número racional pertencente a $I_x = \{y \in \mathbb{R};\ f(x-) \leq y \leq f(x+)\}$. Segue que o conjunto dos pontos de descontinuidade de f é enumerável. $\square$

Definição 2.1.6. *Sejam $X_1, \cdots, X_n$ variáveis aleatórias definidas no espaço de probabilidade $(\Omega, \mathcal{F}, P)$. A **função de distribuição conjunta** de $X_1, \cdots, X_n$ é a função $F : \mathbb{R}^n \to \mathbb{R}$ dada por*

$$F_{X_1,\cdots,X_n}(x_1, \cdots, x_n) = P(X_1 \leq x_1, \cdots, X_n \leq x_n).$$

A partir da definição acima temos o seguinte resultado a respeito da independência de variáveis aleatórias:

Proposição 2.1.3 (Critério Para Independência)**.** *Sejam $X_1, \cdots, X_n$ variáveis aleatórias definidas no espaço de probabilidade $(\Omega, \mathcal{F}, P)$.*

i. Se $X_1, \cdots, X_n$ são independentes, então

$$F_{X_1,\cdots,X_n}(x_1, \cdots, x_n) = \prod_{j=1}^{n} F_{X_j}(x_j), \text{ para todo } (x_1, \cdots, x_n) \in \mathbb{R}^n.$$

ii. Reciprocamente, se existem funções $F_1, \cdots, F_n$ *tais que*

$$\lim_{x\to\infty} F_j(x) = 1 \text{ para todo } j, \text{ e}$$

$$F_{X_1,\cdots,X_n}(x_1,\cdots,x_n) = \prod_{j=1}^{n} F_j(x_j), \text{ para todo } (x_1,\cdots,x_n) \in \mathbb{R}^n,$$

então $X_1, \cdots, X_n$ *são independentes e* $F_j = F_{X_j}$ *para todo* $j = 1, \cdots, n$.

Definição 2.1.7. *Seja* $(\Omega, \mathcal{F}, P)$ *um espaço de probabilidade e* $X : \Omega \to \mathbb{R}$ *uma variável aleatória.*

- *Dizemos que* X *é uma* ***variável aleatória discreta*** *se existe um conjunto enumerável* $\{x_1, x_2, \cdots\} \subset \mathbb{R}$ *tal que* $X(\omega) \in \{x_1, x_2, \cdots\}$ *para todo* $\omega \in \Omega$. *Se o conjunto for finito com* n *elementos, dizemos que a função de distribuição de* X *possui* n *átomos.*

- *Dizemos que* X *é uma* ***variável aleatória absolutamente contínua*** *se existe uma função* $f : \mathbb{R} \to \mathbb{R}^+$ *tal que a função de distribuição de* X *é dada por*

$$F(x) = \int_{-\infty}^{x} f(t)dt, \quad x \in \mathbb{R}$$

  Nesse caso, f *é chamada* de ***densidade de probabilidade*** de X, *ou simplesmente, densidade de* X.

- *Dizemos que* X *é uma* ***variável aleatória singular*** *se a sua função de distribuição* F *é contínua com* $F'(x) = 0$ *em quase toda a parte, isto é, se* $F'(x) \neq 0$ *apenas num conjunto de medida nula.*

Note que uma função $f : \mathbb{R} \to \mathbb{R}^+$ é densidade de alguma variável aleatória se, e somente se,

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

Por definição, dada uma medida de probabilidade, existe uma função de distribuição associada a ela. A recíproca de tal fato também é verdadeira, isto é, dada uma função de distribuição F, obtemos uma medida de probabilidade μ em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. A medida μ nos intervalos é dada por $\mu(a, b] = F(b) - F(a)$ e se estende pelo Teorema de Extensão de Medidas de Caratheodory (SHIRYAEV, 1996, Capítulo 2, Seção 3) a uma medida de probabilidade μ definida em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

2.2 Existência de Infinitas Variáveis Aleatórias Independentes

Um dos principais objetivos desta dissertação é investigar o comportamento de uma sequência de variáveis aleatórias. Nesse contexto, precisamos da existência de n

variáveis aleatórias independentes e identicamente distribuídas definidas num espaço de probabilidade adequado, qualquer que seja o inteiro positivo n. Assim, antes de qualquer coisa, é preciso que tenhamos certeza de que tais variáveis aleatórias existem, e isso fica claro no teorema seguinte (CHUNG, 2001, Capítulo 3, Teorema 3.3.4).

Teorema 2.2.1. *Seja $\{\mu_n\}_{n\geq 1}$ uma sequência finita ou infinita de medidas de probabilidade definidas em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Então existe um espaço de probabilidade $(\Omega, \mathcal{F}, P)$ e uma sequência de variáveis aleatórias independentes $\{X_n\}_{n\geq 1}$ em $(\Omega, \mathcal{F}, P)$ tais que μ_n é a medida de probabilidade de X_n para todo n.*

Demonstração. Sem perda de generalidade podemos supor que a sequência $\{\mu_n\}_{n\geq 1}$ é infinita, uma vez que a prova nesse caso também servirá para o caso finito. Observe que para todo n existe um espaço de probabilidade $(\Omega_n, \mathcal{F}_n, P_n)$ e uma variável aleatória X_n nele definida com μ_n sendo a medida de probabilidade a ela associada. De fato, basta tomar $(\Omega_n, \mathcal{F}_n, P_n) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_n)$ e X_n como a função identidade em $\mathbb{R}$, isto é, $X_n(\omega) = \omega$. Vamos fixar então um espaço $(\Omega_n, \mathcal{F}_n, P_n)$ e uma variável aleatória X_n para cada n. Resta-nos construir o espaço $(\Omega, \mathcal{F}, P)$ no qual as variáveis aleatórias estão definidas e provar que elas são independentes. Para o espaço amostral vamos considerar

$$\Omega = \times_{n=1}^{\infty} \Omega_n,$$

que é o conjunto formado pelos pontos da forma $\omega = (\omega_1, \omega_2, ...)$ com $\omega_n \in \Omega_n$ para todo n. Passaremos agora à determinação de uma σ-álgebra de subconjuntos de Ω.

Dizemos que $E \subset \Omega$ é um *conjunto produto finito* se

$$E = \times_{n=1}^{\infty} F_n \stackrel{def}{=} \{(\omega_1, \omega_2, \cdots); \omega_n \in F_n\},$$

onde $F_n \in \mathcal{F}_n$ para todo n e $F_n \neq \Omega_n$ para um número finito de $n's$. Vamos representar a classe de tais conjuntos por E_0, isto é,

$$E_0 = \{E \in \Omega : E \text{ é um conjunto produto finito}\}$$

Seja $\mathcal{F}_0$ uma classe de subconjuntos de Ω cujos elementos são a união disjunta de um número finito de conjuntos produto finito, isto é,

$$\mathcal{F}_0 = \{A \subset \Omega : A = \bigcup_{k=1}^{n} E^{(k)} \text{ com } E^{(1)}, \cdots, E^{(n)} \in E_0 \text{ e } E^{(i)} \cap E^{(j)} = \emptyset \text{ se } i \neq j\}.$$

Vamos verificar que $\mathcal{F}_0$ é uma álgebra, definir uma medida de probabilidade P, σ-aditiva em $\mathcal{F}_0$, e usando o Teorema de Caratheodory, estenderemos $\mathcal{F}_0$ de maneira única a uma σ-álgebra $\mathcal{F}$ e P a uma medida de probabilidade em $\mathcal{F}$.

Claramente $\Omega \in \mathcal{F}_0$. Além disso, dados A_1 e $A_2 \in \mathcal{F}_0$, temos

$$A_1 = \bigcup_{k=1}^{n_1} E^{(k)}, \text{ e } A_2 = \bigcup_{j=1}^{n_2} F^{(j)},$$

onde os $E^{(k)}$ e os $F^{(j)}$ pertencem a E_0, $E^{(k)} \cap E^{(k')} = \emptyset$ se $k \neq k'$ e $F^{(j)} \cap F^{(j')} = \emptyset$ se $j \neq j'$, e logo,

$$A_1 \cup A_2 = \bigcup_{k=1}^{n_1} E^{(k)} \cup \bigcup_{j=1}^{n_2} F^{(j)} = \bigcup_{i=1}^{n_1+n_2} G^{(i)},$$

onde $G^{(i)} = E^{(i)}$ para $1 \leq i \leq n_1$ e $G^{(i)} = F^{(i-n_1)}$ para $n_1 + 1 \leq i \leq n_1 + n_2$. Note que os $G^{(i)\prime}s$ não são necessariamente disjuntos, mas se usarmos $H^{(i)} = G^{(i)} - [G^{(1)} \cup \cdots \cup G^{(i-1)}]$, teremos

$$\bigcup_{i=1}^{n_1+n_2} G^{(i)} = \bigcup_{i=1}^{n_1+n_2} H^{(i)},$$

onde os $H^{(i)\prime}s$ são disjuntos e pertencem a E_0. Assim, $A_1 \cup A_2 \in \mathcal{F}_0$, e para mostrarmos que $\mathcal{F}_0$ é uma álgebra de conjuntos de Ω, falta mostrar que $\mathcal{F}_0$ é fechado em relação à complementação. De fato, dado $E^{(k)} \in E_0$, temos

$$E^{(k)} = \times_{n=1}^{\infty} F_n,$$

onde $F_n \in \mathcal{F}_n$ para todo n e $F_n \neq \Omega_n$ para um número finito de $n's$. Logo, $E^{(K)c}$ será dado pela união finita de conjuntos da forma $\times_{n=1}^{\infty} \tilde{F}_n$, onde $\tilde{F}_n = F_n$ ou $\tilde{F}_n = F_n^c$ para $n \leq n_0$ e $F_n = \Omega_n$ para $n > n_0$. Assim, $E^{(k)c}$ se escreve como união de um número finito de elementos de E_0, e logo temos que,

$$E^{(k)c} \in \mathcal{F}_0.$$

Para o caso geral, dado $A \in \mathcal{F}_0$, temos

$$A = \bigcup_{k=1}^{n} E^{(k)}.$$

Logo,

$$A^c = \bigcap_{k=1}^{n} E^{(k)c} \in \mathcal{F}_0,$$

uma vez que a interseção de um número finito de elementos de E_0 resulta em um elemento de E_0 (basta observar que $[A_1 \times B] \cap [A_2 \times B] = [A_1 \cap A_2] \times B$).

Assim, verificamos que $\mathcal{F}_0$ é uma álgebra de subconjuntos de Ω. Agora vamos definir uma medida de probabilidade P em $\mathcal{F}_0$ da seguinte maneira:

Se $E \in E_0$ (temos $E = \times_{n=1}^{\infty} F_n$, onde $F_n \in \mathcal{F}_n$ para todo n e $F_n \neq \Omega_n$ para um número finito de $n's$) vamos definir

$$P(E) = \prod_{n=1}^{\infty} P_n(F_n).$$

Note que nesse caso existe $n_0 \in \mathbb{N}$ tal que $P_n(F_n) = P_n(\Omega_n) = 1$ para todo $n > n_0$, e assim temos

$$P(E) = \prod_{n=1}^{n_0} P_n(F_n).$$

Para o caso geral, se $E \in \mathcal{F}_0$ (temos $E = \bigcup_{k=1}^{n} E^{(k)}$ com $E^{(1)}, \cdots, E^{(n)} \in E_0$ e $E^{(j)} \cap E^{(k)} = \emptyset$ se $j \neq k$) definimos

$$P(E) = \sum_{k=1}^{n} P(E^{(k)}). \tag{2.1}$$

Trivialmente temos $P(E) \geq 0$ para todo $E \in \mathcal{F}_0$ e $P(\Omega) = 1$. É fácil verificar também que se $A, B \in \mathcal{F}_0$, com $A \cap B \neq \emptyset$, temos $A = \bigcup_{k=1}^{n_1} E^{(k)}$ e $B = \bigcup_{j=1}^{n_2} F^{(j)}$, e logo, usando $G^{(i)} = E^{(i)}$ se $1 \leq i \leq n_1$ e $G^{(i)} = F^{(i-n_1)}$ se $n_1 + 1 \leq i \leq n_2$, segue que

$$\begin{aligned} P(A \cup B) &= P\left(\bigcup_{k=1}^{n_1} E^{(k)} \cup \bigcup_{j=1}^{n_2} F^{(j)}\right) = P\left(\bigcup_{i=1}^{n_1+n_2} G^{(i)}\right) \\ &= \sum_{i=1}^{n_1+n_2} P\left(G^{(i)}\right) = \sum_{k=1}^{n_1} P\left(E^{(k)}\right) + \sum_{j=1}^{n_2} P\left(F^{(j)}\right) \\ &= P(A) + P(B). \end{aligned}$$

Assim, P é uma medida de probabilidade finitamente aditiva em $\mathcal{F}_0$. Para verificar que P é σ-aditiva, vamos mostrar que P é contínua no vazio, conforme a Proposição 2.1.1. Para tanto, considere uma sequência $\{C_n\}_{n\geq 1}$ em $\mathcal{F}_0$ de forma que exista $\delta > 0$ tal que para todo n tenhamos $C_{n+1} \subset C_n$ e $P(C_n) \geq \delta$. Precisamos mostrar que $\bigcap_{n=1}^{\infty} C_n \neq \emptyset$. Recorde que para todo $E \in \mathcal{F}_0$ existe $k \in \mathbb{N}$ tal que E é determinado por k coordenadas, isto é, dados $\omega = (\omega_1, \omega_2, \cdots)$ e $\omega' = (\omega'_1, \omega'_2, \cdots)$ com $\omega_j = \omega'_j$ para $1 \leq j \leq k$ temos que ambos pertencem a E, ou nenhum dos dois pertence. Note também que sendo k_1 e k_2 o número mínimo de coordenadas de E_1 e E_2, respectivamente, temos que $E_2 \subset E_1$ implica que $k_1 \leq k_2$. Além disso, se E é determinado por k coordenadas, será determinado também por $k + j$ coordenadas, qualquer que seja o inteiro positivo j. Assim, tomando uma subsequência se necessário, podemos supor que para n grande, C_n é determinado por n coordenadas. Agora, dado $\lambda_1 \in \Omega_1$, defina

$$(E|\lambda_1) = \Omega_1 \times E_1,$$

onde

$$E_1 = \{(\omega_2, \omega_3, \cdots) \in \times_{n=2}^{\infty} \Omega_n : (\lambda_1, \omega_2, \omega_3, \cdots) \in E\}.$$

Note que se λ_1 não for a primeira coordenada de nenhum ponto de E, então $(E|\lambda_1) = \emptyset$. Outro fato é que se $E \in \mathcal{F}_0$, o mesmo ocorre com $(E|\lambda_1)$. Além disso, se $E \in E_0$, com $E = \times_{n=1}^{\infty} F_n$, temos imediatamente que

$$(E|\lambda_1) = \begin{cases} \Omega_1 \times F_2 \times F_3 \times \cdots, & \text{se } \lambda_1 \in F_1 \\ \emptyset, & \text{se } \lambda_1 \notin F_1 \end{cases}.$$

Vamos mostrar que existe $\lambda_1 \in \Omega_1$ tal que $P((C_n|\lambda_1)) \geq \frac{\delta}{2}$ para todo n. Para tanto, observe primeiramente que

$$P(C_n) = \int_{\Omega_1} P((C_n|\omega_1)) dP_1(\omega_1).$$

Isso é imediato se $C_n \in E_0$, pois nesse caso temos $C_n = \times_{k=1}^{\infty} F_k$ com $F_k \in \mathcal{F}_k$ para todo $k, F_k = \Omega_k$ se $k > n$, e

$$(C_n|\omega_1) = \begin{cases} \Omega_1 \times F_2 \times F_3 \times \cdots, & \text{se } \omega_1 \in F_1 \\ \emptyset, & \text{se } \omega_1 \notin F_1 \end{cases}.$$

Daí segue que

$$P((C_n|\omega_1)) = I_{F_1}(\omega_1) \prod_{n=2}^{\infty} P_n(F_n),$$

onde

$$I_{F_1}(\omega_1) = \begin{cases} 1, & \text{se } \omega_1 \in F_1 \\ 0, & \text{se } \omega_1 \notin F_1 \end{cases},$$

e logo,

$$\int_{\Omega_1} P((C_n|\omega_1))dP_1(\omega_1) = \int_{\Omega_1} I_{F_1}(\omega_1) \prod_{n=2}^{\infty} P_n(F_n)dP_1(\omega_1) = \prod_{n=1}^{\infty} P_n(F_n) = P((C_n)).$$

No caso geral, se $C_n \in \mathcal{F}_0$, então $C_n = \bigcup\limits_{k=1}^{n} E^{(k)}$, e logo temos $(C_n|\omega_1) = \bigcup\limits_{k=1}^{n} (E^{(k)}|\omega_1)$, onde $(E^{(k)}|\omega_1) \cap (E^{(j)}|\omega_1) = \emptyset$ se $k \neq j$. Segue que

$$\begin{aligned} P(C_n) &= \sum_{k=1}^{n} P(E^{(k)}) = \sum_{k=1}^{n} \int_{\Omega_1} P(E^{(k)}|\omega_1)dP_1(\omega_1) \\ &= \int_{\Omega_1} \left(\sum_{k=1}^{n} P((E^{(k)}|\omega_1)) \right) dP_1(\omega_1) = \int_{\Omega_1} P\left(\bigcup_{k=1}^{n} (E^{(k)}|\omega_1) \right) dP_1(\omega_1) \\ &= \int_{\Omega_1} P((C_n|\omega_1))dP_1(\omega_1). \end{aligned}$$

Defina agora o conjunto $B_n \subset \Omega_1$ por

$$B_n = \{\omega_1 \in \Omega_1 : P((C_n|\omega_1)) \geq \frac{\delta}{2}\}.$$

Note que para $\omega_1 \in B_n$ temos $\frac{\delta}{2} \leq P((C_n|\omega_1)) \leq 1$, e para $\omega_1 \notin B_n$ temos $0 \leq P((C_n|\omega_1)) < \frac{\delta}{2}$. Segue daí que

$$\begin{aligned} \delta &\leq P(C_n) = \int_{\Omega_1} P((C_n|\omega_1))dP_1(\omega_1) \\ &= \int_{B_n} P((C_n|\omega_1))dP_1(\omega_1) + \int_{B_n^c} P((C_n|\omega_1))dP_1(\omega_1) \\ &\leq \int_{B_n} dP_1(\omega_1) + \int_{B_n^c} \frac{\delta}{2} dP_1(\omega_1) = \int_{B_n} \left(1 - \frac{\delta}{2}\right) dP_1(\omega_1) + \int_{\Omega} \frac{\delta}{2} dP_1(\omega_1) \\ &= \left(1 - \frac{\delta}{2}\right) P_1(B_n) + \frac{\delta}{2}, \end{aligned}$$

donde temos

$$\left(1 - \frac{\delta}{2}\right) P_1(B_n) \geq \frac{\delta}{2},$$

e logo

$$P_1(B_n) \geq \frac{2\delta}{2(2-\delta)} \geq \frac{\delta}{2},$$

qualquer que seja $n \geq 1$. Observe que B_n decresce junto com C_n, uma vez que $\omega_1 \in B_{n+1}$ implica $P((C_n|\omega_1)) \geq P((C_{n+1}|\omega_1)) \geq \frac{\delta}{2}$, e consequentemente $\omega_1 \in B_n$. Assim, temos $P_1(\bigcap_{n=1}^{\infty} B_n) \geq \frac{\delta}{2}$, e logo, $\bigcap_{n=1}^{\infty} B_n \neq \emptyset$. Tomando $\lambda_1 \in \bigcap_{n=1}^{\infty} B_n$, temos que para todo $n \geq 1$ vale

$$P((C_n|\lambda_1)) \geq \frac{\delta}{2}.$$

Resumindo, partindo de $P(C_n) \geq \delta$ para todo $n \geq 1$, chegamos à existência de $\lambda_1 \in \Omega_1$ tal que $P((C_n|\lambda_1)) \geq \frac{\delta}{2}$ para todo $n \geq 1$. Repetindo o argumento, partindo de $P((C_n|\lambda_1)) \geq \frac{\delta}{2}$ para todo $n \geq 1$, obteremos a existência de $\lambda_2 \in \Omega_2$ tal que $P((C_n|\lambda_1, \lambda_2)) \geq \frac{\delta}{4}$ para todo $n \geq 1$, e por recorrência, obteremos para todo $k \geq 1$ a existência de $\lambda_k \in \Omega_k$ de forma que para todo $n \geq 1$ vale

$$P((C_n|\lambda_1, \cdots, \lambda_k)) \geq \frac{\delta}{2^k}.$$

Segue daí que $(C_k|\lambda_1, \cdots, \lambda_k) \neq \emptyset$ para todo k, isto é, sempre existe um ponto em C_k cujas k primeiras coordenadas são $\lambda_1, \cdots, \lambda_k$, respectivamente. Logo, todo ponto cujas primeiras k coordenadas são $\lambda_1, \cdots, \lambda_k$, respectivamente, está em C_k, qualquer que seja k. Assim, tomando $\lambda = (\lambda_1, \lambda_2, \cdots)$, temos que $\lambda \in C_k$ para todo k, e daí

$$\lambda \in \bigcap_{k=1}^{\infty} C_k,$$

como queríamos demonstrar.

Assim, P é uma medida de probabilidade definida em $\mathcal{F}_0$, e logo, do Teorema de Extensão de Medidas de Caratheodory (SHIRYAEV, 1996, Capítulo 2, Seção 3) temos que existe uma única σ-álgebra $\mathcal{F}$ que estende $\mathcal{F}_0$ e uma medida de probabilidade que estende P definida em $\mathcal{F}$. Dessa forma fica provado que existe um espaço de probabilidade $(\Omega, \mathcal{F}, P)$ no qual está definida a variável aleatória $X = (X_1, X_2, \cdots)$, e além disso, como $P(\times_{n=1}^{\infty} F_n) = \prod_{n=1}^{\infty} P_n(F_n)$ para todo $F_n \in \mathcal{F}_n$, temos que as variáveis aleatórias $X_1, X_2, \cdots$ são independentes. □

Em nenhum momento se fez necessário que as medidas de probabilidade fossem todas distintas, e portanto, bastava que existisse uma quantidade finita delas. Em particular, se tomarmos as medida de probabilidade todas iguais, obteremos a existência de infinitas variáveis aleatórias independentes e identicamente distribuídas.

Outra fato que é importante observar é que na equação (2.1) definimos $P(E)$ para $E \in \mathcal{F}_0$ usando a decomposição $E = \dot{\bigcup}_{k=1}^{n} E^{(k)}$ com $E^{(1)}, \cdots, E^{(n)} \in E_0$. É importante ressaltar que essa decomposição não é única, mas esse fato não influencia a definição de

$P(E)$, pois se tivermos

$$E = \dot{\bigcup_{k=1}^{n_1}} E^{(k)} = \dot{\bigcup_{j=1}^{n_2}} \widehat{E}^{(j)}, \tag{2.2}$$

teremos

$$\sum_{k=1}^{n_1} P(E^{(k)}) = \sum_{j=1}^{n_2} P(\widehat{E}^{(j)}). \tag{2.3}$$

De fato, dado $E \in \mathcal{F}_0$, consideremos as representações de E como na equação (2.2). Note que para todo $k = 1, \cdots, n_1$ teremos $E^{(k)} = \dot{\bigcup}_{j=1}^{n_2} (E^{(k)} \cap \widehat{E}^{(j)})$, e portanto, $P(E^{(k)}) = \sum_{j=1}^{n_2} P(E^{(k)} \cap \widehat{E}^{(j)})$. Segue que

$$\sum_{k=1}^{n_1} P(E^{(k)}) = \sum_{k=1}^{n_1} \sum_{j=1}^{n_2} P(E^{(k)} \cap \widehat{E}^{(j)}).$$

Analogamente teremos também para todo $j = 1, \cdots, n_2$ que $\widehat{E}^{(j)} = \dot{\bigcup}_{k=1}^{n_1} (E^{(k)} \cap \widehat{E}^{(j)})$, e portanto

$$\sum_{j=1}^{n_2} P(\widehat{E}^{(j)}) = \sum_{j=1}^{n_2} \sum_{k=1}^{n_1} P(E^{(k)} \cap \widehat{E}^{(j)}),$$

verificando-se assim a equação (2.3), como queríamos.

2.3 Esperança Matemática

Para falarmos sobre a Lei dos Grandes Números e o Teorema Central do Limite é necessário definirmos primeiro o que são a esperança e a variância de uma variável aleatória. Informalmente, a esperança é o valor que se espera obter ao escolher aleatoriamente um elemento de $X(\Omega)$, e a variância mede quanto que a esperança de X^2 se afasta do quadrado da esperança de X. Com isso em mente, é natural pensarmos que no caso da variável aleatória discreta que assume apenas valores positivos, a esperança coincide com a média aritmética ponderada dos valores de $X(\Omega)$, e em (CHUNG, 2001, Capítulo 3) ela é definida exatamente dessa maneira. No caso geral, ela é obtida por aproximação a partir de uma sequência de variáveis aleatórias discretas, analisando-se separadamente as partes positiva e negativa, chegando-se assim aos resultados que seguem.

Definição 2.3.1 (Esperança). *Seja $X : \Omega \mapsto \mathbb{R}$ uma variável aleatória cuja função de distribuição é F. A esperança de X é definida como*

$$\mathbb{E}X = \int_{-\infty}^{\infty} x dF(x) = \int_{\Omega} X dP = \int_{\Omega} X(\omega) dP(\omega),$$

desde que as integrais estejam bem definidas. Se $\mathbb{E}X = \mu < \infty$, dizemos X é integrável.

Note que

$$\mathbb{E}X = \int_{-\infty}^{\infty} x dF(x) = \int_{-\infty}^{0} x dF(x) + \int_{0}^{\infty} x dF(x) = I + II,$$

de onde segue que

i. X é integrável se I e II são finitas;
ii. $\mathbb{E}X = \infty$ se I é finita e $II = \infty$;
iii. $\mathbb{E}X = -\infty$ se II é finita e $I = -\infty$;
iv. $\mathbb{E}X$ não está definida se $I = -\infty$ e $II = \infty$.

Proposição 2.3.1. *Seja X é uma variável aleatória.*

i. Se X é discreta, tomando os valores $\{x_1, x_2, \cdots\}$, temos

$$\mathbb{E}X = \sum_i x_i P(X = x_i).$$

ii. Se X é contínua com densidade $f(x)$, temos

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

iii. Para toda função mensurável $\varphi : \mathbb{R} \mapsto \mathbb{R}$, a esperança da variável aleatória $\varphi(X)$ é dada por

$$\mathbb{E}\varphi(X) = \int_{-\infty}^{\infty} \varphi(x) dF(x).$$

No caso em que a variável aleatória X for discreta, tomando os valores $\{x_1, x_2, \cdots\}$, teremos

$$\mathbb{E}\varphi(X) = \sum_i \varphi(x_i) P(X = x_i).$$

No caso em que a variável aleatória X for absolutamente contínua, com densidade $f(x)$, teremos

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

Proposição 2.3.2. *A variável aleatória X é integrável se, e somente se, $|X|$ é integrável.*

Proposição 2.3.3 (Propriedades da Esperança)**.** *Sejam $X, Y : \Omega \to \mathbb{R}$ variáveis aleatórias integráveis. Então vale que*

- *Se $X \equiv c$, então $\mathbb{E}X = c$;*
- *E(aX+bY)=aEX+bEY;*
- *Se $X \leq Y$, então $\mathbb{E}X \leq \mathbb{E}Y$;*

- *Se* $X_1, \cdots, X_n$ *são variáveis aleatórias independentes, então*

$$\mathbb{E}(X_1 \cdots X_n) = \mathbb{E}X_1 \cdots \mathbb{E}X_n.$$

Definição 2.3.2 (Momentos)**.** *Seja* X *uma variável aleatória e* $b \in \mathbb{R}$*. Definimos o k-ésimo momento de* X *em torno de* b *como sendo* $\mathbb{E}(X-b)^k$*. Se* $b = 0$*, então* $\mathbb{E}X^k$ *é o k-ésimo momento de* X*, simplesmente. Se* $b = \mathbb{E}X$*, então* $\mathbb{E}(X - \mathbb{E}X)^k$ *é o k-ésimo momento central de* X*.*

Obviamente, o primeiro momento de X é $\mathbb{E}X$, e o primeiro momento central é $\mathbb{E}(X - \mathbb{E}X) = 0$.

Definição 2.3.3 (Variância)**.** *A variância da variável aleatória* X *corresponde ao segundo momento central de* X*, isto é,*

$$VarX = \mathbb{E}(X - \mathbb{E}X)^2.$$

Note que pela Proposição 2.3.3 vale que

$$\begin{aligned} VarX &= \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}(X^2 - 2X\mathbb{E}X + (\mathbb{E}X)^2) \\ &= \mathbb{E}X^2 - 2\mathbb{E}X\mathbb{E}X + (\mathbb{E}X)^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2. \end{aligned}$$

Proposição 2.3.4. *Propriedades da Variância:*

1. *Se* $X \equiv c$*, então* $VarX = 0$*;*
2. $Var(aX + b) = a^2 VarX$*;*
3. *Se* X *e* Y *são independentes, então* $Var(X + Y) = VarX + VarY$*;*
4. $VarX = 0$ *se, e somente se,* $P(X = k) = 1$ *para algum* $k \in \mathbb{R}$*.*

Teorema 2.3.1 (Teorema da Convergência Dominada)**.** *Sejam* $X, X_1, X_2, \cdots$ *variáveis aleatórias definidas num mesmo espaço de probabilidade tais que* $\{X_n\}_{n\geq 1}$ *converge em quase toda parte para* X*. Se existe uma variável aleatória integrável* Y *tal que para todo* $n \geq 1$ *tenhamos* $|X_n| \leq Y$ *em quase toda parte, então* X *e* X_n *são integráveis e* $\mathbb{E}X_n \overset{n\to\infty}{\longrightarrow} \mathbb{E}X$*.*

2.4 Funções Características

Analisar o comportamento de uma sequência de variáveis aleatórias não costuma ser uma tarefa fácil. No entanto, dada uma variável aleatória X, existe uma função complexa $\varphi : \mathbb{R} \to \mathbb{C}$ unicamente determinada por X que chamaremos de **função característica de X**. Desta forma, dada uma sequência de variáveis aleatórias $\{X_n\}_{n\geq 1}$, existe uma sequência de funções características $\{\varphi_n\}_{n\geq 1}$ associada a ela, cujo comportamento costuma ser mais fácil de analisar. Nesse contexto, vamos agora estudar algumas propriedades de tal função.

Definição 2.4.1 (Função Característica). *Seja X uma variável aleatória cuja função de distribuição é F. Define-se como função característica de X, a função $\varphi_X : \mathbb{R} \to \mathbb{C}$ definida por*

$$\varphi_X(t) = \mathbb{E}e^{itX} = \mathbb{E}\cos(tX) + i\mathbb{E}\sin(tX).$$

Da Proposição 2.3.1 segue imediatamente que

$$\varphi_X(t) = \int_{-\infty}^{\infty} \cos(tx)dF(x) + i\int_{-\infty}^{\infty} \sin(tx)dF(x) = \int_{-\infty}^{\infty} e^{itx}dF(x) = \int_{\Omega} e^{itX(\omega)}dP(\omega).$$

Quando não houver risco de confusão, usaremos simplesmente φ para nos referirmos à função característica φ_X.

Proposição 2.4.1 (Propriedades da função característica). *Seja φ uma função característica. Então vale que:*

- $|\varphi(t)| \leq \varphi(0) = 1$ *e* $\varphi(-t) = \overline{\varphi(t)}$*, qualquer que seja* $t \in \mathbb{R}$*;*
- $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$ *e* $\varphi_{-X}(t) = \overline{\varphi_X(t)}$*, quaisquer que sejam* $a, b, t \in \mathbb{R}$*;*
- φ *é uniformemente contínua em* $\mathbb{R}$*;*
- *Se* $\varphi_1, \cdots, \varphi_n$ *são funções características, então* $\varphi = \varphi_1 \cdots \varphi_n$ *também é uma função característica;*
- *A variável aleatória* X *tem distribuição simétrica em torno de zero (isto é,* $P(X \leq x) = P(X \geq -x), x \in \mathbb{R}$*) se, e somente se,* $\varphi_X(t) \in \mathbb{R}$ *para todo* $t \in \mathbb{R}$*.*

Demonstração. A primeira propriedade segue imediatamente das igualdades

$$\varphi(0) = \mathbb{E}e^{i0X} = \mathbb{E}1 = 1,$$

$$|\varphi(t)| = \left|\int_{-\infty}^{\infty} e^{itx}dF(x)\right| \leq \int_{-\infty}^{\infty} |e^{itx}|dF(x) = \int_{-\infty}^{\infty} dF(x) = 1, \text{ e,}$$

$$\varphi(-t) = \int_{-\infty}^{\infty} e^{-itx}dF(x) = \int_{-\infty}^{\infty} \overline{e^{itx}}dF(x) = \overline{\varphi(t)}.$$

Para verificar a segunda propriedade, temos das propriedades da esperança que

$$\varphi_{aX+b}(t) = \mathbb{E}e^{it(aX+b)} = e^{itb}\mathbb{E}e^{itaX} = e^{itb}\varphi_X(at),$$

e consequentemente,

$$\varphi_{-X}(t) = \varphi_X(-t) = \overline{\varphi(t)}.$$

Vamos provar agora que φ é uniformemente contínua em $\mathbb{R}$. Para tanto, dado $\epsilon > 0$, precisamos mostrar que existe $\delta > 0$ tal que $|h| < \delta$ implica $|\varphi(t+h) - \varphi(t)| < \epsilon$, e que o valor de δ não depende do valor de t. Observemos primeiro que

$$\begin{aligned} |\varphi(t+h) - \varphi(t)| &= \left| \int_{-\infty}^{\infty} e^{i(t+h)x} dF(x) - \int_{-\infty}^{\infty} e^{itx} dF(x) \right| = \left| \int_{-\infty}^{\infty} [e^{ihx} e^{itx} - e^{itx}] dF(x) \right| \\ &\leq \int_{-\infty}^{\infty} |e^{itx}| |e^{ihx} - 1| dF(x) = \int_{-\infty}^{\infty} |e^{ihx} - 1| dF(x). \end{aligned}$$

Como $|e^{ihx} - 1|$ converge a zero quando h tende a zero e $|e^{ihx} - 1| \leq 2$, temos pelo Teorema da Convergência Dominada que $\int_{-\infty}^{\infty} |e^{ihx} - 1| dF(x)$ converge a zero quando h tende a zero. Logo existe $\delta > 0$ tal que $|h| < \delta$ implica $\int_{-\infty}^{\infty} |e^{ihx} - 1| dF(x) < \epsilon$, e daí segue que

$$|h| < \delta \Rightarrow |\varphi(t+h) - \varphi(t)| \leq \int_{-\infty}^{\infty} |e^{ihx} - 1| dF(x) < \epsilon,$$

qualquer que seja $t \in \mathbb{R}$, como queríamos demonstrar.

Quanto à quarta propriedade, sejam $X_1, \cdots, X_n$ as variáveis aleatórias tais que $\varphi_j(t) = \mathbb{E}e^{itX_j}$ são as respectivas funções características. Pelo Teorema 2.2.1 podemos supor que elas são independentes, e logo,

$$\varphi(t) = \varphi_1(t) \cdots \varphi_n(t) = \mathbb{E}e^{itX_1} \cdots \mathbb{E}e^{itX_n} = \mathbb{E}e^{itX_1} \cdots e^{itX_n} = \mathbb{E}e^{it(X_1 + \cdots + X_n)}.$$

Assim, φ é a função característica da variável aleatória $X_1 + \cdots + X_n$.

A última propriedade segue das propriedades anteriores, das propriedades das funções de distribuição e do Corolário 2.4.1, segundo as quais

$$\begin{aligned} P(X \leq x) = P(X \geq -x) &\Leftrightarrow F_X(x) = P(X \leq x) = P(-X \leq x) = F_{-X}(x) \\ &\Leftrightarrow \varphi_X(t) = \varphi_{-X}(t) = \overline{\varphi_X(t)} \\ &\Leftrightarrow \varphi_X(t) \in \mathbb{R}. \end{aligned}$$

□

Um importante resultado sobre funções características é que existe uma correspondência biunívoca com as funções de distribuição. Para provarmos este resultado começamos com o lema seguinte:

Lema 2.4.1 (Integral Clássica de Dirchlet)**.**

$$\int_0^{\infty} \frac{\sin(t)}{t} dt = \frac{\pi}{2}.$$

Demonstração. Observe inicialmente que

$$\begin{aligned} \int_0^{\infty} \frac{\sin(t)}{t} dt &= \int_0^{\infty} \sin(t) \left(\frac{e^{-tu}}{-t} \Bigg|_{u=0}^{u=\infty} \right) dt = \int_0^{\infty} \sin(t) \left(\int_0^{\infty} e^{-tu} du \right) dt \\ &= \int_0^{\infty} \int_0^{\infty} \sin(t) e^{-tu} du dt. \end{aligned}$$

Agora, note que

$$\int_0^\infty \frac{\sin(t)}{t}dt = \sum_{k=0}^{\infty}\int_{k\pi}^{(k+1)\pi} \frac{\sin(t)}{t}dt$$

é uma série alternada cujos termos convergem para zero, uma vez que dado $t \in [k, k+1]$, temos $\sin(t) \geq 0$ se k é par, e $\sin(t) \leq 0$ se k é ímpar, ao mesmo tempo que $\frac{\sin(t)}{t} \overset{t\to\infty}{\longrightarrow} 0$. Logo, $\int_0^\infty \frac{\sin(t)}{t}dt = M < \infty$, e daí, o integrando $\sin(t)e^{-tu}$ é contínuo e integrável. Portanto podemos trocar a ordem de integração e temos

$$\begin{aligned}
\int_0^\infty \frac{\sin(t)}{t}dt &= \int_0^\infty\int_0^\infty \sin(t)e^{-tu}dtdu = \int_0^\infty \frac{-e^{-tu}(\cos(t) + u\sin(t))}{1+u^2}\Bigg|_{t=0}^{t=\infty} du \\
&= \int_0^\infty \frac{1}{1+u^2}du = \arctan(u)\Bigg|_{u=0}^{u=\infty} = \frac{\pi}{2}.
\end{aligned}$$

□

Proposição 2.4.2 (Fórmula da Inversão)**.** *Seja X uma variável aleatória, F sua função de distribuição e φ sua função característica. Se $x < y$, então*

$$\frac{1}{2\pi}\lim_{u\to\infty}\int_{-u}^{u} \frac{e^{-itx} - e^{-ity}}{it}\varphi(t)dt = \frac{F(y) + F(y-)}{2} - \frac{F(x) + F(x-)}{2}.$$

Demonstração. Seja

$$I(u) = \int_{-u}^{u} \frac{e^{-itx} - e^{-ity}}{it}\varphi(t)dt,$$

onde $x < y$ são pontos de continuidade de F. Logo,

$$\begin{aligned}
I(u) &= \int_{-u}^{u} \frac{e^{-itx} - e^{-ity}}{it}\left[\int_{-\infty}^{\infty} e^{itz}dF(z)\right]dt \\
&= \int_{-u}^{u}\int_{-\infty}^{\infty} \frac{e^{-itx} - e^{-ity}}{it}e^{itz}dF(z)dt.
\end{aligned}$$

Por L'Hôpital temos

$$\lim_{t\to 0}\frac{e^{-itx} - e^{-ity}}{it} = \lim_{t\to 0}\frac{-ixe^{-itx} + iye^{-ity}}{i} = y - x,$$

e logo

$$I(u) = \int_{-u}^{u}\int_{-\infty}^{\infty} f(z,t)dF(z)dt,$$

onde

$$f(z,t) = \begin{cases} y - x, & \text{se } t = 0 \\ \frac{e^{-itx} - e^{-ity}}{it}e^{itz}, & \text{se } t \neq 0 \end{cases}.$$

Como $|f(z,t)| \overset{t\to\infty}{\longrightarrow} 0$, temos f limitada, exceto possivelmente em uma vizinhança de $t = 0$. Mas vimos que f é contínua em $t = 0$, e logo, $|f(z,t)| \leq c < \infty$ para todo $z \in \mathbb{R}$ e $t \in [-u, u]$, de onde segue que

$$\begin{aligned}
\left|\int_{-u}^{u}\int_{-\infty}^{\infty} f(z,t)dF(z)dt\right| &\leq \int_{-u}^{u}\int_{-\infty}^{\infty} |f(z,t)|dF(z)dt \leq \int_{-u}^{u}\int_{-\infty}^{\infty} cdF(z)dt \\
&= c\int_{-u}^{u}\int_{-\infty}^{\infty} dF(z)dt = c\int_{-u}^{u} dt = ct\Bigg|_{t=-u}^{t=u} = 2uc.
\end{aligned}$$

Assim podemos trocar a ordem de integração, e logo,

$$\begin{aligned} I(u) &= \int_{-\infty}^{\infty}\int_{-u}^{u}\frac{e^{it(z-x)}-e^{it(z-y)}}{it}dtdF(z) \\ &= \int_{-\infty}^{\infty}\int_{-u}^{u}\frac{\cos(t(z-x))+i\sin(t(z-x))-\cos(t(z-y))-i\sin(t(z-y))}{it}dtdF(z) \\ &= \int_{-\infty}^{\infty}\int_{-u}^{u}\left[\frac{\cos(t(z-x))-\cos(t(z-y))}{it}+\frac{\sin(t(z-x))-\sin(t(z-y))}{t}\right]dtdF(z). \end{aligned}$$

Como $\frac{\cos(at)}{t}$ é uma função ímpar e $\frac{\sin(at)}{t}$ é uma função par, segue que

$$I(u) = 2\int_{-\infty}^{\infty}\int_{0}^{u}\frac{\sin(t(z-x))-\sin(t(z-y))}{t}dtdF(z) = \mathbb{E}g_u(X),$$

onde

$$g_u(z) = 2\int_0^u \frac{\sin(t(z-x))}{t}dt - 2\int_0^u\frac{\sin(t(z-y))}{t}dt.$$

Vamos agora fazer as verificações necessárias para aplicar o Teorema da Convergência Dominada à sequência $(g_u(x))_{u\geq 1}$. Pela Integral de Dirichlet, dado $a\in\mathbb{R}$, vale que:

$$\begin{aligned} a>0 &\Rightarrow \int_0^u\frac{\sin(at)}{t}dt = \int_0^{au}\frac{\sin(t)}{t}dt \Rightarrow \lim_{u\to\infty}\int_0^u\frac{\sin(at)}{t}dt = \frac{\pi}{2}, \\ a<0 &\Rightarrow \int_0^u\frac{\sin(at)}{t}dt = -\int_0^u\frac{\sin(-at)}{t}dt \\ &\Rightarrow \lim_{u\to\infty}\int_0^u\frac{\sin(at)}{t}dt = -\lim_{u\to\infty}\int_0^u\frac{\sin(-at)}{t}dt = -\frac{\pi}{2}, \\ a=0 &\Rightarrow \lim_{u\to\infty}\int_0^u\frac{\sin(at)}{t}dt = 0. \end{aligned}$$

Assim, lembrando que $x<y$, temos

$$\lim_{u\to\infty}g_u(z) = \begin{cases} 0, & \text{se } z<x \\ \pi, & \text{se } z=x \\ 2\pi, & \text{se } x<z<y \\ \pi, & \text{se } z=y \\ 0, & \text{se } z>y \end{cases}.$$

Em termos de variáveis aleatórias isso significa que

$$g_u(X) \overset{u\to\infty}{\longrightarrow} \pi I_{[X=x]} + 2\pi I_{[x<X<y]} + \pi I_{[X=y]}.$$

Agora, note que a função $\int_0^u \frac{\sin(t)}{t}dt$ é contínua em $[0,\infty]$ e converge para $\frac{\pi}{2}$ quando $u\to\infty$, e portanto é limitada em $[0,\infty]$. Daí,

$$\left|\int_0^u\frac{\sin(at)}{t}dt\right| \leq \sup_{u>0}\left|\int_0^u\frac{\sin(t)}{t}dt\right| = M<\infty.$$

Assim, temos para todo $u>0$ que

$$|g_u(z)| \leq 2\left|\int_0^u\frac{\sin(t(z-x))}{t}dt\right| + 2\left|\int_0^u\frac{\sin(t(z-y))}{t}dt\right| \leq 4M,$$

e logo, do Teorema da Convergência Dominada segue que

$$\begin{aligned}\lim_{u\to\infty}\mathbb{E}g_u(X) &= \mathbb{E}\lim_{u\to\infty}g_u(X) = \pi\mathbb{E}I_{[X=x]} + 2\pi\mathbb{E}I_{[x<X<y]} + \pi\mathbb{E}I_{[X=y]}\\ &= \pi P(X=x) + 2\pi P(x<X<y) + \pi P(X=y),\end{aligned}$$

e daí temos

$$\begin{aligned}\lim_{u\to\infty}\frac{1}{2\pi}I(u) &= \frac{P(X=x)}{2} + P(x<X<y) + \frac{P(X=y)}{2}\\ &= \frac{F(x)-F(x-)}{2} + F(y-) - F(x) + \frac{F(y)-F(y-)}{2}\\ &= \frac{F(y)+F(y-)}{2} - \frac{F(x)+F(x-)}{2}.\end{aligned}$$

□

Corolário 2.4.1. *Sejam X_1 e X_2 variáveis aleatórias cujas funções de distribuição são F_1 e F_2 e as funções características são φ_1 e φ_2, respectivamente. Então temos $F_1 = F_2$ se, e somente se, $\varphi_1 = \varphi_2$.*

Demonstração. Se $F_1 = F_2$ temos por definição que

$$\varphi_1(t) = \int_{-\infty}^{\infty} e^{itx}dF_1(x) = \int_{-\infty}^{\infty} e^{itx}dF_2(x) = \varphi_2(t).$$

Reciprocamente, se $\varphi_1 = \varphi_2$ temos da fórmula da inversão que se y é um ponto de continuidade de F_1 e F_2, então

$$\begin{aligned}F_1(y) &= \lim_{x\to-\infty}\left[\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-itx}-e^{-ity}}{it}\varphi_1(t)dt + F_1(x)\right]\\ &= \lim_{x\to-\infty}\left[\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-itx}-e^{-ity}}{it}\varphi_2(t)dt + F_2(x)\right]\\ &= F_2(y).\end{aligned}$$

Como F_1 e F_2 são contínuas à direita, segue que para todo $z \in \mathbb{R}$,

$$\begin{aligned}F_1(z) &= \lim_{y\searrow z}\lim_{x\to-\infty}\left[\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-itx}-e^{-ity}}{it}\varphi_1(t)dt + F_1(x)\right]\\ &= \lim_{y\searrow z}\lim_{x\to-\infty}\left[\lim_{u\to\infty}\frac{1}{2\pi}\int_{-u}^{u}\frac{e^{-itx}-e^{-ity}}{it}\varphi_2(t)dt + F_2(x)\right]\\ &= F_2(z),\end{aligned}$$

e daí, $F_1 = F_2$. □

Definição 2.4.2 (Convolução)**.** *Dadas as funções de distribuição F_1 e F_2, a função de distribuição F é dita convolução de F_1 e F_2 se, para todo $x \in \mathbb{R}$,*

$$F(x) = \int_{-\infty}^{\infty} F_1(x-y)dF_2(y).$$

Nesse caso escrevemos $F = F_1 * F_2$.

Teorema 2.4.1. *Se X_1 e X_2 são variáveis aleatórias independentes com funções de distribuição F_1 e F_2, e funções características φ_1 e φ_2, respectivamente, então a variável aleatória $X = X_1 + X_2$ tem função de distribuição $F = F_1 * F_2$ e função característica $\varphi = \varphi_1\varphi_2$.*

Demonstração. Pela independência de X_1 e X_2 temos que para todo $t \in \mathbb{R}$,

$$\varphi(t) = \mathbb{E}e^{itX} = \mathbb{E}e^{it(X_1+X_2)} = \mathbb{E}(e^{itX_1}e^{itX_2}) = \mathbb{E}e^{itX_1}\mathbb{E}e^{itX_2} = \varphi_1(t)\varphi_2(t).$$

Para verificar que $F = F_1 * F_2$, vamos definir a função $f : \mathbb{R}^2 \to \mathbb{R}$ tal que

$$f(x_1, x_2) = \begin{cases} 1, & \text{se } x_1 + x_2 \leq x \\ 0, & \text{se } x_1 + x_2 > x \end{cases}.$$

Logo, pela independência de X_1 e X_2, e pela Proposição 2.1.3, vale que

$$\begin{aligned} F(x) &= P(X_1 + X_2 \leq x) = \int_{\mathbb{R}^2} f(x_1, x_2) dF(x_1, x_2) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2) dF_1(x_1) dF_2(x_2) = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{x-x_2} dF_1(x_1)\right] dF_2(x_2) \\ &= \int_{-\infty}^{\infty} F_1(x - x_2) dF_2(x_2). \end{aligned}$$

□

Corolário 2.4.2. *Se $X = \sum_{j=1}^{n} X_j$, onde $X_1, \cdots, X_n$ são variáveis aleatórias independentes, então X tem função de distribuição $F = F_1 * \cdots * F_n$ e função característica $\varphi = \varphi_1 \cdots \varphi_n$, onde F_j é a função de distribuição e φ_j é a função característica de X_j, $j = 1, \cdots, n$.*

Demonstração. Basta aplicar $n - 1$ vezes o Teorema 2.4.1. □

Corolário 2.4.3. *Se φ é uma função característica, então o mesmo ocorre com $|\varphi|^2$.*

Demonstração. Seja X uma variável aleatória cuja função característica é φ. Do Teorema 2.2.1 temos que existe uma variável aleatória Y tal que X e Y são independentes e identicamente distribuídas, e portanto têm também a mesma função característica. Segue daí que a função característica da variável aleatória $X - Y$ será

$$\mathbb{E}e^{it(X-Y)} = \mathbb{E}e^{itX}\mathbb{E}e^{-itY} = \varphi(t)\varphi(-t) = \varphi(t)\overline{\varphi(t)} = |\varphi(t)|^2.$$

□

Teorema 2.4.2. *Seja X uma variável aleatória cuja função de distribuição é F e a função característica é φ. Então $\mathbb{E}|X|^n < \infty$ se, e somente se, φ possui n derivadas contínuas, valendo*

$$\varphi^{(k)}(t) = \int_{-\infty}^{\infty} (ix)^k e^{itx} dF(x).$$

Demonstração. Supondo inicialmente que $\mathbb{E}|X|^n < \infty$, vamos provar a ida por indução. Se $n = 1$, temos $\mathbb{E}|X| < \infty$, e queremos mostrar primeiramente que para todo $t \in \mathbb{R}$ existe

$$\varphi'(t) = \int_{-\infty}^{\infty} ixe^{itx} dF(x),$$

isto é,

$$\frac{\varphi(t+h) - \varphi(t)}{h} \xrightarrow{h\to 0} \int_{-\infty}^{\infty} ixe^{itx} dF(x).$$

Primeiramente observe que para $h \neq 0$ temos

$$\begin{aligned}\frac{\varphi(t+h) - \varphi(t)}{h} &= \int_{-\infty}^{\infty} \frac{e^{i(t+h)x} - e^{itx}}{h} dF(x) = \int_{-\infty}^{\infty} e^{itx}\left(\frac{e^{ihx}-1}{h}\right) dF(x) \\ &= \mathbb{E}\left(e^{itX}\left(\frac{e^{ihX}-1}{h}\right)\right).\end{aligned}$$

Como $\frac{e^{ihx}-1}{h} \xrightarrow{h\to 0} ix$ para todo $x \in \mathbb{R}$, temos que $e^{itX}\left(\frac{e^{ihX}-1}{h}\right)$ converge pontualmente para a variável aleatória iXe^{itX}, quando $h \to 0$. Por outro lado temos

$$\left|\frac{e^{ihx}-1}{h}\right| = \left|\frac{\int_0^h ixe^{isx} ds}{h}\right| \leq |x| \left|\frac{\int_0^h ie^{isx} ds}{h}\right| \leq |x| \frac{\int_0^{|h|} |ie^{isx}| ds}{|h|} \leq |x| \frac{\int_0^{|h|} ds}{|h|} = |x|,$$

e logo,

$$\left| e^{itX}\left(\frac{e^{ihX}-1}{h}\right)\right| \leq |X|.$$

Como |X| é integrável, segue do Teorema da Convergência Dominada que

$$\mathbb{E}\left(e^{itX}\left(\frac{e^{ihX}-1}{h}\right)\right) \xrightarrow{h\to 0} \mathbb{E}(iXe^{itX}),$$

e assim,

$$\varphi'(t) = \lim_{h\to 0} \frac{\varphi(t+h) - \varphi(t)}{h} = \lim_{h\to 0} \mathbb{E}\left(e^{itX}\left(\frac{e^{ihX-1}}{h}\right)\right) = \mathbb{E}(iXe^{itX}) = \int_{-\infty}^{\infty} ixe^{itx} dF(x).$$

A continuidade de φ' também decorre do Teorema da Convergência Dominada, uma vez que temos $ixe^{isx} \xrightarrow{s\to t} ixe^{itx}$ para todo $x \in \mathbb{R}$ e $|ixe^{isx}| = |x|$, donde segue que iXe^{isX} converge pontualmente para iXe^{itX} quando $s \to t$, e $|iXe^{isX}| = |X|$, e logo, como $|X|$ é integrável, temos que

$$\varphi'(s) = \mathbb{E}iXe^{isX} \xrightarrow{s\to t} \mathbb{E}iXe^{itX} = \varphi'(t).$$

Suponha agora que valha $\varphi^{(j)}(t) = \int_{-\infty}^{\infty} (ix)^j e^{itx} dF(x)$ e que $\mathbb{E}|X|^{(j+1)} < \infty$. Logo,

$$\begin{aligned}\frac{\varphi^{(j)}(t+h) - \varphi^{(j)}(t)}{h} &= \int_{-\infty}^{\infty} (ix)^j \left(\frac{e^{i(t+h)x} - e^{itx}}{h} dF(x)\right) \\ &= \int_{-\infty}^{\infty} (ix)^j e^{itx}\left(\frac{e^{ihx}-1}{h}\right) dF(x) \\ &= \mathbb{E}\left((iX)^j e^{itX}\left(\frac{e^{ihX}-1}{h}\right)\right).\end{aligned}$$

De modo análogo ao caso em que $k = 1$, temos que $(iX)^j e^{itX}\left(\frac{e^{ihX}-1}{h}\right)$ converge pontualmente para $(iX)^{j+1}e^{itX}$, e,

$$\left|(iX)^j e^{itX}\left(\frac{e^{ihX}-1}{h}\right)\right| = |(iX)^j||e^{itX}|\left|\frac{e^{ihX}-1}{h}\right| \leq |X|^j|X| = |X|^{j+1}.$$

Daí, pelo Teorema da Convergência Dominada,

$$\mathbb{E}\left((iX)^j e^{itX}\left(\frac{e^{ihX}-1}{h}\right)\right) \xrightarrow{h\to 0} \mathbb{E}((iX)^{j+1}e^{itX}).$$

Assim,

$$\begin{aligned}
\varphi^{(j+1)}(t) &= \lim_{h\to 0}\frac{\varphi^{(j)}(t+h)-\varphi^{(j)}(t)}{h} = \lim_{h\to 0}\mathbb{E}\left((iX)^j e^{itX}\left(\frac{e^{ihX-1}}{h}\right)\right) \\
&= \mathbb{E}((iX)^{j+1}e^{itX}) = \int_{-\infty}^{\infty}(ix)^{j+1}e^{itx}dF(x).
\end{aligned}$$

Novamente de modo análogo ao caso em que $k = 1$, também temos $(iX)^{j+1}e^{isX}$ convergindo pontualmente para $(iX)^{j+1}e^{itX}$ e $|(iX)^{j+1}e^{isX}| = |X|^{j+1}$, donde segue do Teorema da Convergência Dominada que

$$\varphi^{(j+1)}(s) = \mathbb{E}((iX)^{j+1}e^{isX}) \xrightarrow{s\to t} \mathbb{E}((iX)^{j+1}e^{itX}) = \varphi^{(j+1)}(t).$$

Quanto à recíproca, se

$$\varphi^{(k)}(0) = \int_{-\infty}^{\infty}(ix)^k dF(x) = i\mathbb{E}(X^k),$$

e φ possui k derivadas contínuas, segue que X^k é integrável, e portanto, $|X|^k$ também é. □

Corolário 2.4.4. *Seja X uma variável aleatória cuja função característica é φ. Então X é integrável se, e somente se, φ é uma função de classe C^1. Ao mesmo tempo, X tem variância finita se, e somente se, φ é uma função de classe C^2. Em particular, $\mathbb{E}X = -i\varphi'(0)$, e, $VarX = -\varphi''(0) + \varphi'(0)^2$.*

Demonstração. O corolário é apenas uma aplicação do Teorema 2.4.2 para $n = 1$ e $n = 2$. Quanto ao cálculo da esperança e da variância, basta observar que

$$\varphi'(0) = \int_{-\infty}^{\infty} ixdF(x) = i\mathbb{E}X \Rightarrow \mathbb{E}X = -i\varphi'(0),$$

ao mesmo tempo que

$$\varphi''(0) = \int_{-\infty}^{\infty}(ix)^2 dF(x) = -\mathbb{E}(X^2) \Rightarrow VarX = \mathbb{E}(X^2) - (\mathbb{E}X)^2 = -\varphi''(0) + \varphi'(0)^2.$$

□

2.5 Exemplos de Distribuições

Nesta seção vamos apresentar alguns exemplos de distribuições importantes e de grande aplicabilidade. Além da distribuição normal e da distribuição degenerada que aparecem respectivamente no Teorema Central do Limite e na Lei dos Grandes Números, abordaremos também as distribuições de Poisson e de Cauchy, as quais são dois exemplos interessantes de distribuições infinitamente divisíveis, conforme definiremos adiante.

2.5.1 Distribuição Degenerada

Definição 2.5.1. *A variável aleatória X tem distribuição degenerada, se existe $\lambda \in \mathbb{R}$ tal que*

$$P(X = \lambda) = 1.$$

Pela Proposição 2.3.1 é imediato que a função característica correspondente à distribuição degenerada é

$$\varphi(t) = e^{it\lambda},$$

e logo, pelo Corolário 2.4.4, temos $\mathbb{E}X = \lambda$ e $VarX = 0$. Um fato que ocorre em ralação à função característica da variável aleatória com distribuição degenerada, é que, em módulo, ela é constante e igual a 1, e a recíproca também é verdadeira, como segue.

Proposição 2.5.1. *Seja X uma variável aleatória, e $\varphi : \mathbb{R} \to \mathbb{C}$ a sua função característica. Se $|\varphi(t)| \equiv 1$, então a distribuição de X é degenerada.*

Demonstração. A demonstração será feita a partir das seguintes afirmações:

i. Se $|\varphi(t_0)| = 1$ para algum $t_0 \neq 0$, então existe $\lambda \in \mathbb{R}$ tal que X tem massa concentrada nos pontos da forma $\lambda + \frac{2\pi n}{t_0}$, $n \in \mathbb{Z}$, isto é,

$$\sum_{n\in\mathbb{Z}} P\left(X = \lambda + \frac{2\pi n}{t_0}\right) = 1.$$

ii. Se $|\varphi(t_0)| = |\varphi(\alpha t_0)| = 1$, onde $t_0 \neq 0$ e α é um número irracional, então a distribuição de X é degenerada.

Note que, provada a segunda afirmação, a proposição estará demonstrada.

Para provar a primeira afirmação, observe que como $|\varphi(t_0)| = 1$ e $t_0 \neq 0$, existe $\lambda \in \mathbb{R}$ tal que $\varphi(t_0) = e^{it_0\lambda}$. Logo, sendo F a função de distribuição de X, temos

$$e^{it_0\lambda} = \int_{-\infty}^{\infty} e^{it_0x}dF(x) \quad\Rightarrow\quad 1 = \int_{-\infty}^{\infty} e^{it_0(x-\lambda)}dF(x) = \int_{-\infty}^{\infty} \cos(t_0(x-\lambda))dF(x)$$
$$\Rightarrow\quad \int_{-\infty}^{\infty} [1 - \cos(t_0(x-\lambda))]dF(x) = 0.$$

Como $1-\cos(t_0(x-\lambda)) \geq 0$ para todo $x \in \mathbb{R}$, segue da igualdade acima que $1-\cos(t_0(X-\lambda)) = 0$ em quase toda parte, isto é, exceto num conjunto de medida nula. Segue daí que $t_0(X-\lambda) = 2\pi n$, $n \in \mathbb{Z}$, e consequentemente, $X = \lambda + \frac{2\pi n}{t_0}$ em quase toda parte, e daí,

$$\sum_{n\in\mathbb{Z}} P\left(X = \lambda + \frac{2\pi n}{t_0}\right) = 1.$$

Quanto à segunda afirmação, sendo $|\varphi(t_0)| = |\varphi(\alpha t_0)| = 1$, temos pela primeira afirmação que existem $\lambda_1, \lambda_2 \in \mathbb{R}$, tais que

$$\sum_{n\in\mathbb{Z}} P\left(X = \lambda_1 + \frac{2\pi n}{t_0}\right) = \sum_{n\in\mathbb{Z}} P\left(X = \lambda_2 + \frac{2\pi n}{\alpha t_0}\right) = 1.$$

Daí, supondo por absurdo que a distribuição de X não é degenerada, existirão $n_1 \neq n_2$ e $m_1 \neq m_2$, tais que

$$\lambda_1 + \frac{2\pi n_1}{t_0} = \lambda_2 + \frac{2\pi m_1}{\alpha t_0}, \quad \text{e,} \quad \lambda_1 + \frac{2\pi n_2}{t_0} = \lambda_2 + \frac{2\pi m_2}{\alpha t_0},$$

onde a probabilidade de ambos não é nula. Segue que

$$\frac{2\pi}{t_0}(n_1 - n_2) = \frac{2\pi}{\alpha t_0}(m_1 - m_2),$$

e consequentemente,

$$\alpha = \frac{m_1 - m_2}{n_1 - n_2}.$$

Chegamos assim a uma contradição, já que α é um número irracional, e portanto, não pode ser escrito como fração de números inteiros. □

2.5.2 Distribuição Normal

Definição 2.5.2. *A variável aleatória X tem distribuição normal padrão se ela tiver função densidade dada por*

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

isto é, se a sua função de distribuição for dada por

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt, \quad x \in \mathbb{R}.$$

Nesse caso, escrevemos $X \sim N(0,1)$.

Para verificar que f é de fato uma função de densidade, note que $f(x) > 0$ para todo $x \in \mathbb{R}$, e logo basta mostrar que $\int_{-\infty}^{\infty} f(x)dx = 1$. Para tanto, observemos primeiro que

$$\left(\int_{-\infty}^{\infty} f(x)dx\right)^2 = \int_{-\infty}^{\infty} f(x)dx \int_{-\infty}^{\infty} f(y)dy = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\frac{-(x^2+y^2)}{2}} dxdy.$$

Em coordenadas polares temos $x = \rho\cos(\theta)$ e $y = \rho\sin(\theta)$, com $\rho \in [0, \infty)$ e $\theta \in [0, 2\pi]$. Logo,

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\frac{-(x^2+y^2)}{2}}dxdy &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{\frac{-\rho^2}{2}}\rho d\rho d\theta = \frac{1}{2\pi}\int_0^{2\pi} -e^{\frac{-\rho^2}{2}}\Big|_{\rho=0}^{\rho=\infty} d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} d\theta = 1.
\end{aligned}$$

Assim, como $\left(\int_{-\infty}^{\infty} f(x)dx\right)^2 = 1$ e $f(x) > 0$ para todo $x \in \mathbb{R}$, segue que $\int_{-\infty}^{\infty} f(x)dx = 1$.

Proposição 2.5.2. *A função característica correspondente à distribuição normal padrão é*

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

Demonstração. Pela Proposição 2.3.1 temos

$$\begin{aligned}
\varphi(t) &= \int_{-\infty}^{\infty} e^{itx} f(x)dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{itx}e^{-\frac{x^2}{2}}dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{(itx-\frac{x^2}{2})}dx = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}}dx \\
&= \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\int_{\gamma} e^{-\frac{z^2}{2}}dz,
\end{aligned}$$

onde $z = x - it$ e $\gamma : \mathbb{R} \to \mathbb{C}$ é dada por $\gamma(s) = s - it$.

Note que $\int_{\gamma} e^{-\frac{z^2}{2}}dz = \lim_{n\to\infty}\int_{[-n-it,n-it]} e^{-\frac{z^2}{2}}dz$, e pela teoria da integração complexa de Cauchy, temos para todo $n \in \mathbb{R}$ que

$$\int_{[-n-it,n-it]} e^{-\frac{z^2}{2}}dz + \int_{[n-it,n]} e^{-\frac{z^2}{2}}dz + \int_{n}^{-n} e^{-\frac{z^2}{2}}dz + \int_{[-n,-n-it]} e^{-\frac{z^2}{2}}dz = 0.$$

Logo,

$$\int_{\gamma} e^{-\frac{z^2}{2}}dz + \lim_{n\to\infty}\int_{[n-it,n]} e^{-\frac{z^2}{2}}dz - \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}dz + \lim_{n\to\infty}\int_{[-n,-n-it]} e^{-\frac{z^2}{2}}dz = 0.$$

Agora, como $[n - it, n]$ é determinado por $\lambda(s) = n + is$, $s \in [-t, 0]$, temos

$$\begin{aligned}
\left|\int_{[n-it,n]} e^{-\frac{z^2}{2}}dz\right| &= \left|\int_{-t}^{0} ie^{-\frac{(n+is)^2}{2}}ds\right| \leq \int_{-t}^{0}\left|ie^{\frac{-n^2-2ins+s^2}{2}}\right|ds \\
&= \int_{-t}^{0}|i|\left|e^{\frac{s^2-n^2}{2}}\right||e^{-ins}|ds = \int_{-t}^{0} e^{\frac{s^2-n^2}{2}}ds \\
&= e^{\frac{-n^2}{2}}\int_{-t}^{0} e^{\frac{s^2}{2}}ds \overset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Analogamente temos também que $|\int_{[-n,-n-it]} e^{-\frac{z^2}{2}}dz| \overset{n\to\infty}{\longrightarrow} 0$, e consequentemente,

$$\int_{\gamma} e^{-\frac{z^2}{2}}dz = \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}dz = \sqrt{2\pi}.$$

Segue daí que

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

□

Definição 2.5.3. *Uma variável aleatória Y tem distribuição normal com parâmetros μ e σ^2 se tivermos $Y = \sigma X + \mu$, onde X tem distribuição normal padrão. Nesse caso escrevemos $Y \sim N(\mu, \sigma^2)$.*

Como veremos no Corolário 2.5.2, os parâmetros μ e σ^2, são a esperança e variância de Y, respectivamente.

Corolário 2.5.1. *A função característica correspondente à distribuição normal com parâmetros μ e σ^2 é*

$$\varphi_Y(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}.$$

Demonstração. Como $Y = \sigma X + \mu$, onde a função característica de X é $\varphi_X(t) = e^{-\frac{t^2}{2}}$, temos da Proposição 2.4.1 que

$$\varphi_Y(t) = e^{it\mu}\varphi_X(\sigma t) = e^{it\mu}e^{-\frac{\sigma^2 t^2}{2}} = e^{it\mu - \frac{\sigma^2 t^2}{2}}.$$

□

Corolário 2.5.2. *Se Y tem distribuição normal com parâmetros μ e σ^2, então $\mathbb{E}Y = \mu$ e $VarY = \sigma^2$.*

Demonstração. Sendo $\varphi_Y(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}$ a função característica de Y, temos $\varphi_Y'(t) = (i\mu - \sigma^2 t)e^{it\mu - \frac{\sigma^2 t^2}{2}}$ e $\varphi_Y''(t) = -\sigma^2 e^{it\mu - \frac{\sigma^2 t^2}{2}} + (i\mu - \sigma^2 t)^2 e^{it\mu - \frac{\sigma^2 t^2}{2}}$, que são funções contínuas em $\mathbb{R}$ com $\varphi_Y'(0) = i\mu$ e $\varphi_Y''(0) = -\sigma^2 - \mu^2$. Segue do Corolário 2.4.4 que

$$\mathbb{E}X = -i\varphi_Y'(0) = \mu \quad \text{e}$$

$$VarX = -\varphi_Y''(0) + \varphi_Y'(0)^2 = \sigma^2.$$

□

2.5.3 Distribuição de Cauchy

Definição 2.5.4. *A variável aleatória X tem distribuição de Cauchy padrão se ela tiver função densidade dada por*

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R},$$

isto é, se a sua função de distribuição for dada por

$$F(x) = \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{1+t^2}dt, \quad x \in \mathbb{R}.$$

Note que f é de fato uma função de densidade, já que $f(x) > 0$ para todo $x \in \mathbb{R}$, e,

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1}{1+x^2}dx = \frac{1}{\pi}\arctan(x)\Big|_{x=-\infty}^{x=\infty} = \frac{1}{\pi}\left(\frac{\pi}{2} - (-\frac{\pi}{2})\right) = 1.$$

Proposição 2.5.3. *A função característica correspondente à distribuição de Cauchy padrão é*

$$\varphi(t) = e^{-|t|}.$$

Demonstração. Pela Proposição 2.3.1 temos

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

Note que $g(x) = \frac{e^{itx}}{1+x^2}$ tem polo simples em $x = \pm i$. Logo, o resíduo de g calculado no ponto i será

$$res(g, i) = \lim_{x\to i} \frac{(x-i)e^{itx}}{1+x^2} = \frac{e^{-t}}{2i}.$$

Sendo C_r o semicírculo de raio r centrado na origem e percorrido no sentido anti-horário, e $\gamma_r = C_r \cup [-r, r]$, temos pelo Teorema dos Resíduos (SOARES, 1999, Capítulo 6) que para todo $r > 1$

$$\frac{1}{\pi}\int_{\gamma_r} \frac{e^{itx}}{1+x^2}dx = \frac{2\pi i}{\pi} res(g, i) = e^{-t}.$$

Por outro lado, como $x = r(\cos(s) + i\sin(s)), s \in [0, \pi]$, é uma parametrização de C_r, temos que dado $t \geq 0$,

$$\begin{aligned} \left|\int_{C_r} \frac{e^{itx}}{1+x^2}dx\right| &= \left|\int_0^{\pi} \frac{e^{itr(\cos(s)+i\sin(s))} r(-\sin(s) + i\cos(s))}{r^2(\cos(s) + i\sin(s))^2 + 1}ds\right| \\ &\leq \int_0^{\pi} \frac{|e^{-tr\sin(s)}|}{r|(\cos(s) + i\sin(s))^2 + \frac{1}{r^2}|}ds \\ &\leq \int_0^{\pi} \frac{e^{-tr\sin(s)}}{r(\cos(s) + i\sin(s))^2 + \frac{1}{r}}ds \xrightarrow{r\to\infty} 0. \end{aligned}$$

Segue portanto, que para todo $t \geq 0$,

$$\varphi(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{e^{itx}}{1+x^2}dx = \lim_{r\to\infty} \frac{1}{\pi}\int_{\gamma_r} \frac{e^{itx}}{1+x^2}dx - \lim_{r\to\infty} \frac{1}{\pi}\int_{C_r} \frac{e^{itx}}{1+x^2}dx = e^{-t}.$$

Para $t < 0$ temos da Proposição 2.4.1 que

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

e portanto, para todo $t \in \mathbb{R}$,

$$\varphi(t) = e^{-|t|}.$$

□

Definição 2.5.5. *Uma variável aleatória Y tem distribuição de Cauchy com parâmetro $\theta > 0$ se tivermos $Y = \theta X$, onde X tem distribuição de Cauchy padrão.*

Corolário 2.5.3. *A função característica correspondente à distribuição de Cauchy com parâmetro $\theta > 0$ é*

$$\varphi_Y(t) = e^{-\theta|t|}.$$

Demonstração. Como $Y = \theta X$, onde a função característica de X é $\varphi_X(t) = e^{-|t|}$, temos da Proposição 2.4.1 que

$$\varphi_Y(t) = \varphi_X(\theta t) = e^{-\theta|t|}.$$

$\square$

Corolário 2.5.4. *Se Y tem distribuição de Cauchy com parâmetro $\theta > 0$, então Y não é integrável.*

Demonstração. Sendo $\varphi_Y(t) = e^{-\theta|t|}$ a função característica de Y, temos

$$\varphi_Y'(t) = \begin{cases} -\theta e^{-\theta t}, & \text{se } t > 0 \\ \theta e^{\theta t}, & \text{se } t < 0, \end{cases}$$

que é uma função descontínua em $t = 0$. Segue do Corolário 2.4.4 que Y não é integrável.

$\square$

2.5.4 Distribuição de Poisson

Definição 2.5.6. *A variável aleatória discreta X, que assume apenas valores inteiros positivos, tem distribuição de Poisson com parâmetro $\lambda > 0$ se tivermos*

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \text{ para todo } k \in \mathbb{N}.$$

Nesse caso, escrevemos $X \sim Poisson(\lambda)$.

Note que a função de distribuição de X será

$$F(x) = P(X \leq x) = \sum_{0 \leq k \leq x} P(X = k),$$

isto é,

$$F(x) = \begin{cases} 0, & \text{se } x < 0 \\ \sum_{k=0}^{\lfloor x \rfloor} \frac{e^{-\lambda}\lambda^k}{k!}, & \text{se } x \geq 0, \end{cases}$$

onde $\lfloor x \rfloor$ denota a parte inteira de x.

Proposição 2.5.4. *A função característica correspondente à distribuição de Poisson com parâmetro $\lambda > 0$ é*

$$\varphi(t) = e^{\lambda(e^{it}-1)}.$$

Demonstração. Pela Proposição 2.3.1 segue que

$$\begin{aligned} \varphi(t) &= \mathbb{E}e^{itX} = \sum_{n=0}^{\infty} e^{itn} P(X = n) \\ &= \sum_{n=0}^{\infty} e^{itn} \frac{e^{-\lambda}\lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(e^{it}\lambda)^n}{n!} \\ &= e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}. \end{aligned}$$

□

Corolário 2.5.5. *Se X tem distribuição de Poisson com parâmetro $\lambda > 0$, então $\mathbb{E}X = VarX = \lambda$.*

Demonstração. Sendo $\varphi(t) = e^{\lambda(e^{it}-1)}$ a função característica de X, temos imediatamente $\varphi'(t) = \lambda i e^{it} e^{\lambda(e^{it}-1)}$ e $\varphi''(t) = (-\lambda e^{it} - \lambda^2 e^{2it})e^{\lambda(e^{it}-1)}$, que são funções contínuas em $\mathbb{R}$ com $\varphi'(0) = i\lambda$ e $\varphi''(0) = -\lambda^2 - \lambda$. Segue do Corolário 2.4.4 que

$$\mathbb{E}X = -i\varphi'(0) = \lambda \quad \text{e}$$

$$VarX = -\varphi''(0) + \varphi'(0)^2 = \lambda.$$

□

2.5.5 Distribuição Uniforme

Definição 2.5.7. *A variável aleatória X tem distribuição uniforme em $[a, b]$ se a sua função de distribuição for dada por*

$$F(x) = \begin{cases} 0, & se\ x \leq a \\ \frac{x-a}{b-a}, & se\ a \leq x \leq b \\ 1, & se\ x > b \end{cases}.$$

Nesse caso, escrevemos $X \sim U[a, b]$.

Note que tal variável aleatória tem função densidade dada por $f(x) = \frac{1}{b-a}$ para $a \leq x \leq b$ e $f(x) = 0$ para $x < a$ e $x > b$.

Proposição 2.5.5. *A função característica correspondente à distribuição uniforme em $[a, b]$ é*

$$\varphi(t) = \begin{cases} \frac{e^{it\frac{b+a}{2}}\sin(t\frac{b-a}{2})}{t\frac{b-a}{2}}, & se\ t \neq 0 \\ 1, & se\ t = 0 \end{cases}.$$

Demonstração. Vamos começar pelo caso em que $X \sim U[-1, 1]$, onde temos pela Proposição 2.3.1 que, para $t \neq 0$

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx = \int_{-1}^{1} \frac{e^{itx}}{2}dx = \left.\frac{e^{itx}}{2it}\right|_{x=-1}^{x=1} = \frac{e^{it} - e^{-it}}{2it} = \frac{\sin(t)}{t}.$$

Para $t = 0$ é imediato que

$$\varphi(0) = \int_{-\infty}^{\infty} f(x)dx = \int_{-1}^{1} \frac{1}{2}dx = 1.$$

Por outro lado, se $Y \sim U[a,b]$, temos que $Y = \frac{b-a}{2}X + \frac{b+a}{2}$, onde $X \sim U[-1,1]$. Logo, pela Proposição 2.4.1, a função característica de Y será

$$\varphi(t) = \begin{cases} \frac{e^{it\frac{b+a}{2}}\sin(t\frac{b-a}{2})}{t\frac{b-a}{2}}, & \text{se } t \neq 0 \\ 1, & \text{se } t = 0 \end{cases}.$$

□

Corolário 2.5.6. *Se Y tem distribuição uniforme em $[a,b]$, então $\mathbb{E}Y = \frac{b+a}{2}$ e $VarY = \frac{(b-a)^2}{12}$.*

Demonstração. Se $X \sim U[-1,1]$, então, para $t \neq 0$, a função característica de X será $\varphi_X(t) = \frac{\sin(t)}{t}$ e logo, $\varphi_X'(t) = \frac{t\cos(t)-\sin(t)}{t^2}$ e $\varphi_X''(t) = \frac{(2-t^2)\sin(t)-2t\cos(t)}{t^3}$, que são funções possivelmente descontínuas em $t = 0$. Mas, pela regra de L'Hôpital temos

$$\lim_{t\to 0}\varphi_X'(t) = \lim_{t\to 0}\frac{t\cos(t)-\sin(t)}{t^2} = \lim_{t\to 0}\frac{-t\sin(t)}{2t} = 0 \quad \text{e}$$

$$\lim_{t\to 0}\varphi_X''(t) = \lim_{t\to 0}\frac{(2-t^2)\sin(t)-2t\cos(t)}{t^3} = \lim_{t\to 0}\frac{-t^2\cos(t)}{3t^2} = \frac{-1}{3}.$$

Assim $\varphi'(t)$ e $\varphi''(t)$ são funções contínuas em $\mathbb{R}$ com $\varphi'(0) = 0$ e $\varphi''(0) = \frac{-1}{3}$. Segue do Corolário 2.4.4 que

$$\mathbb{E}X = -i\varphi_Y'(0) = 0 \quad \text{e}$$

$$VarX = -\varphi_X''(0) + \varphi_X'(0)^2 = \frac{1}{3}.$$

Agora, se $Y \sim U[a,b]$, $Y = \frac{b-a}{2}X + \frac{b+a}{2}$, onde $X \sim U[-1,1]$. Logo, pelas Proposições 2.3.3 e 2.3.4 temos

$$\mathbb{E}Y = \frac{b-a}{2}\mathbb{E}X + \frac{b+a}{2} = \frac{b+a}{2} \quad \text{e}$$

$$VarY = (\frac{b-a}{2})^2 VarY = \frac{(b-a)^2}{12}.$$

□

2.5.6 Distribuição Exponencial

Definição 2.5.8. *A variável aleatória X tem distribuição exponencial com parâmetro λ se $P(X > x) = e^{-\lambda x}$ para $x \geq 0$ e $P(X < 0) = 0$. Nesse caso, escrevemos $X \sim \exp(\lambda)$.*

Note que a função de distribuição de X será

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

e que sua função densidade é dada por $f(x) = \lambda e^{-\lambda x}$ para $x \geq 0$ e $f(x) = 0$ para $x < 0$.

Proposição 2.5.6. *A função característica correspondente à distribuição exponencial com parâmetro λ é*

$$\varphi(t) = \frac{-\lambda}{it - \lambda} = \frac{1}{1 - \frac{it}{\lambda}}.$$

Demonstração. Se $X \sim \exp(\lambda)$, temos pela Proposição 2.3.1 que

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx = \int_{0}^{\infty} \lambda e^{(it-\lambda)x}dx = \left.\frac{\lambda e^{(it-\lambda)x}}{it - \lambda}\right|_{x=0}^{x=\infty} = \frac{-\lambda}{it - \lambda} = \frac{1}{1 - \frac{it}{\lambda}}.$$

□

Corolário 2.5.7. *Se X tem distribuição exponencial com parâmetro λ, então $\mathbb{E}X = \frac{-1}{\lambda}$ e $VarX = \frac{1}{\lambda^2}$.*

Demonstração. Sendo $\varphi(t) = \frac{1}{1-\frac{it}{\lambda}}$ a função característica de X, temos, $\varphi'(t) = \frac{\frac{i}{\lambda}}{(1-\frac{it}{\lambda})^2}$ e $\varphi''(t) = \frac{\frac{-2}{\lambda^2}}{(1-\frac{it}{\lambda})^3}$, que são funções contínuas em $\mathbb{R}$ com $\varphi'(0) = \frac{i}{\lambda}$ e $\varphi''(0) = \frac{-2}{\lambda^2}$. Segue do Corolário 2.4.4 que

$$\mathbb{E}X = -i\varphi'(0) = \frac{-1}{\lambda} \quad \text{e}$$

$$VarX = -\varphi''(0) + \varphi'(0)^2 = \frac{1}{\lambda^2}.$$

□

2.5.7 Distribuição Binomial

Definição 2.5.9. *A variável aleatória discreta X, que assume apenas um número finito de valores inteiros positivos, tem distribuição binomial com parâmetros n e p, onde $n \in \mathbb{N}$ e $0 < p < 1$, se tivermos*

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \;\; k = 0, 1, \cdots, n.$$

Nesse caso, escrevemos $X \sim b(n, p)$.

Note que a função de distribuição de X é dada por

$$F(x) = P(X \le x) = \sum_{0 \le k \le x} P(X = k) = \begin{cases} 0, & \text{se } x < 0 \\ \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} p^k (1-p)^{n-k}, & \text{se } x \ge 0 \end{cases}.$$

Proposição 2.5.7. *A função característica correspondente à distribuição binomial com parâmetros n e p é dada por*

$$\varphi(t) = (1 + p(e^{it} - 1))^n.$$

Demonstração. Pela Proposição 2.3.1 segue que

$$\begin{aligned} \varphi(t) &= \mathbb{E}e^{itX} = \sum_{k=0}^{n} e^{itk} P(X = k) \\ &= \sum_{k=0}^{n} \binom{n}{k} e^{itk} p^k (1-p)^{n-k} = (pe^{it} + 1 - p)^n \\ &= (1 + p(e^{it} - 1))^n, \end{aligned}$$

onde a penúltima igualdade vem do binômio de Newton. □

Corolário 2.5.8. *Se X tem distribuição binomial com parâmetros n e p, então $\mathbb{E}X = np$ e $VarX = np(1-p)$.*

Demonstração. Sendo $\varphi(t) = (1+p(e^{it}-1))^n$ a função característica de X, temos $\varphi'(t) = ipne^{it}(1+p(e^{it}-1))^{n-1}$ e $\varphi''(t) = -npe^{it}(1+p(e^{it}-1))^{n-1} - n(n-1)p^2e^{2it}(1+p(e^{it}-1))^{n-2}$, que são funções contínuas em $\mathbb{R}$ com $\varphi'(0) = ipn$ e $\varphi''(0) = -pn(1+pn-p)$. Segue do Corolário 2.4.4 que

$$\mathbb{E}X = -i\varphi'(0) = np \quad \text{e}$$

$$VarX = -\varphi''(0) + \varphi'(0)^2 = np(1-p).$$

□

Para encerrar este capítulo, destaco que a esperança e a variância de uma variável aleatória podem ser obtidos sem determinar a sua função característica. No entanto, uma vez que temos a função característica, é muito mais fácil obter a esperança e a variância a partir dela, como fizemos aqui.

3 Convergência de Sequências de Variáveis Aleatórias

De um modo geral, a convergência de uma sequência está associada à noção de aproximação. No nosso caso, estamos preocupados com a convergência de uma sequência de variáveis aleatórias. Especificamente, vamos considerar três tipos de convergência: a convergência em quase toda parte, em probabilidade e em distribuição. Vamos defini-las e mostrar que ocorrendo a primeira, ocorrerá também a segunda, e consequentemente a terceira. Em seguida vamos apresentar alguns teoremas envolvendo o limite de sequências de variáveis aleatórias. Especificando, vamos enunciar e demonstrar o Teorema de Poisson, a Lei Fraca dos Grandes Números de Khintchin e algumas versões do Teorema Central do Limite. Outro assunto que abordaremos neste capítulo são as famílias de medidas de probabilidade rígidas e relativamente compactas. Tal abordagem será fundamental na demonstração do Teorema 4.2.1, o qual é um dos principais desta dissertação. Os resultados aqui obtidos têm como referência (JAMES, 2006) e (SHIRYAEV, 1996).

3.1 Tipos de Convergência

Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias definidas no mesmo espaço de probabilidade $(\Omega, \mathcal{F}, P)$. Naturalmente, surgem as sequências $\{F_n\}_{n\geq 1}$ e $\{\varphi_n\}_{n\geq 1}$ formadas respectivamente pelas funções de distribuição e funções características correspondentes às variáveis aleatórias, e fica a seguinte questão: se uma das três sequências converge de alguma forma, o que acontece com as outras duas? Na resposta de tal questão é que estaremos preocupados no que segue.

Definição 3.1.1. *Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias definidas no mesmo espaço de probabilidade $(\Omega, \mathcal{F}, P)$, e $\{F_n\}_{n\geq 1}$ a sequência formada pelas respectivas funções de distribuição. Então, dizemos que:*

1. *X_n converge para X em* ***quase toda parte*** *se $P(X_n \overset{n\to\infty}{\longrightarrow} X) = 1$, isto é, se X_n converge pontualmente para X em quase toda parte (exceto num conjunto de medida nula).*

2. *X_n converge para X em* ***probabilidade*** *se para todo $\epsilon > 0$, $P(|X_n - X| \geq \epsilon) \overset{n\to\infty}{\longrightarrow} 0$. Nesse caso, escrevemos $X_n \overset{P}{\longrightarrow} X$.*

3. *X_n converge para X em* ***distribuição*** *se $\{F_n\}_{n\geq 1}$ converge fracamente para F, isto é, se $F_n(x) \overset{n\to\infty}{\longrightarrow} F(x)$, para todo x que é ponto de continuidade de F. Nesse caso, escrevemos $X_n \overset{D}{\longrightarrow} X$.*

O exemplo a seguir mostra que a convergência em probabilidade não garante a convergência em quase toda parte. No entanto, a convergência em quase toda parte garante a convergência em probabilidade, como veremos em seguida.

Exemplo 3.1.1. *Seja X uma variável aleatória com distribuição uniforme no intervalo $[0,1]$. Considere os intervalos*

$$\begin{aligned}
&I_1 = [0,1],\\
&I_2 = \left[0, \frac{1}{2}\right], I_3 = \left[\frac{1}{2}, 1\right],\\
&I_4 = \left[0, \frac{1}{4}\right], I_5 = \left[\frac{1}{4}, \frac{1}{2}\right], I_6 = \left[\frac{1}{2}, \frac{3}{4}\right], I_7 = \left[\frac{3}{4}, 1\right],\\
&I_8 = \left[0, \frac{1}{8}\right], \cdots,
\end{aligned}$$

onde para todo $k \in \mathbb{N}$, $I_{2^k} \cup \cdots \cup I_{2^{k+1}-1} = [0,1]$. Seja para todo $n \in \mathbb{N}$, a variável aleatória Y_n dada por

$$Y_n(\omega) = \begin{cases} 1, & \textit{se } X(\omega) \in I_n \\ 0, & \textit{se } X(\omega) \notin I_n \end{cases}.$$

Então, Y_n converge a zero em probabilidade, mas não converge para 0 em quase toda parte.

De fato, dado $0 < \epsilon \leq 1$, temos

$$P(|Y_n - 0| \geq \epsilon) = P(Y_n = 1) = P(X \in I_n) \overset{n\to\infty}{\longrightarrow} 0.$$

Para $\epsilon > 1$, é imediato que $P(|Y_n - 0| \geq \epsilon) = P(\emptyset) = 0$, e portanto, $Y_n \xrightarrow{P} 0$. Por outro lado, dado $\omega \in \Omega$ arbitrário, temos $X(\omega) \in I_k$ para algum k tal que $2^n \leq k < 2^{n+1} - 1$, qualquer que seja $n \in \mathbb{N}$. Daí, existe uma subsequência $\{Y_{n_j}\}_{j\geq 1}$ infinita tal que $Y_{n_j}(\omega) = 1$ para todo $j \in \mathbb{N}$. De modo análogo, existe também uma subsequência infinita $\{Y_{n_l}\}_{l\geq 1}$ tal que $Y_{n_l}(\omega) = 0$ para todo $l \in \mathbb{N}$, e logo $Y_n(\omega)$ não converge. Portanto, Y_n não converge para 0 em quase toda parte. Aliás, não converge em ponto algum.

Proposição 3.1.1. *Se X_n converge para X em quase toda parte, então $X_n \xrightarrow{P} X$.*

Demonstração. Dado $\epsilon > 0$ arbitrário, precisamos mostrar que $P(|X_n - X| \geq \epsilon) \overset{n\to\infty}{\longrightarrow} 0$. Para tanto, considere os seguintes eventos aleatórios:

$$\begin{aligned}
A_0 &= \{\omega \in \Omega : X_n(\omega) \overset{n\to\infty}{\longrightarrow} X(\omega)\}, \quad \text{e}\\
A_n &= \{\omega \in \Omega : \text{ para todo } k \geq n, |X_k(\omega) - X(\omega)| < \epsilon\}.
\end{aligned}$$

Pela definição da convergência em quase toda parte temos $P(A_0) = 1$. Além disso, se $\omega \in A_0$, então para n suficientemente grande temos $|X_k(\omega) - X(\omega)| < \epsilon$ para todo $k \geq n$, e logo, $\omega \in A_n$. Daí, temos

$$A_0 \subset \bigcup_{n\geq 1} A_n.$$

Por outro lado, se $|X_k(\omega) - X(\omega)| < \epsilon$ para todo $k \leq n$, obviamente temos $|X_k(\omega) - X(\omega)| < \epsilon$ para todo $k \leq n+1$. Logo $A_n \subset A_{n+1}$ para todo n, e portanto,

$$A_0 \subset \bigcup_{n\geq 1} A_n = \lim_{n\to\infty} A_n.$$

Segue portanto que $1 = P(A_0) \leq P(\bigcup_{n\geq 1} A_n)$. Pela continuidade da probabilidade, temos $P(A_n) \nearrow 1$. Para finalizar, observe que para todo n temos

$$A_n \subset \{\omega \in \Omega : |X_n(\omega) - X(\omega)| < \epsilon\},$$

e logo, $P(A_n) \leq P(|X_n - X| < \epsilon)$ qualquer que seja n. Segue daí que

$$P(|X_n - X| < \epsilon) \stackrel{n\to\infty}{\longrightarrow} 1,$$

e logo,

$$P(|X_n - X| \geq \epsilon) = 1 - P(|X_n - X| < \epsilon) \stackrel{n\to\infty}{\longrightarrow} 0.$$

□

Proposição 3.1.2. *Se $X_n \stackrel{P}{\longrightarrow} X$, então $X_n \stackrel{D}{\longrightarrow} X$.*

Demonstração. Sendo F_n a função de distribuição de X_n, $n \geq 1$, F a função de distribuição de X, e x um ponto de continuidade de F, precisamos mostrar que $F_n(x) \stackrel{n\to\infty}{\longrightarrow} F(x)$. Para tanto, observe que dado $\epsilon > 0$ arbitrário e $\omega \in \Omega$, temos $X(\omega) \leq x + \epsilon$ ou $X(\omega) > x + \epsilon$, e daí, $X_n(\omega) \leq x$ implica $X(\omega) \leq x + \epsilon$ ou $X(\omega) - X_n(\omega) > \epsilon$. Assim, temos $[X_n \leq x] \subset [X \leq x + \epsilon] \cup [|X - X_n| > \epsilon]$, e logo,

$$\begin{aligned} F_n(x) &= P(X_n \leq x) \leq P(X \leq x + \epsilon) + P(|X - X_n| > \epsilon) \\ &= F(x+\epsilon) + P(|X - X_n| > \epsilon). \end{aligned}$$

De modo análogo, temos que

$$X(\omega) \leq x - \epsilon \Rightarrow X_n(\omega) \leq x \text{ ou } X_n(\omega) - X(\omega) > \epsilon.$$

Desta forma, $[X \leq x - \epsilon] \subset [X_n \leq x] \cup [|X_n - X| > \epsilon]$, e assim, $F(x - \epsilon) \leq F_n(x) + P(|X_n - X| > \epsilon)$. Juntando as duas desigualdades obtemos

$$F(x-\epsilon) - P(|X_n - X| > \epsilon) \leq F_n(x) \leq F(x+\epsilon) + P(|X - X_n| > \epsilon).$$

Como $X_n \stackrel{P}{\longrightarrow} X$, obtemos, fazendo n tender ao infinito, que

$$F(x-\epsilon) \leq \lim_{n\to\infty} \inf F_n(x) \leq \lim_{n\to\infty} \sup F_n(x) \leq F(x+\epsilon),$$

e como x é um ponto de continuidade de F, fazendo ϵ tender a zero obtemos

$$F(x) = \lim_{n\to\infty} F_n(x),$$

o que conclui a prova da proposição. □

Exemplo 3.1.2. *Sejam $X, X_1, X_2, \cdots$ variáveis aleatórias independentes com distribuição normal com parâmetros* 0 *e* $\frac{1}{2}$. *Então $X_n \xrightarrow{D} X$, mas X_n não converge para X em probabilidade.*

De fato, como as variáveis aleatórias são identicamente distribuídas, é óbvio que $X_n \xrightarrow{D} X$. Por outro lado, sendo φ e φ_n as funções características de X e X_n respectivamente, temos da independência de tais e da Proposição 2.4.1 que a função característica de $X_n - X$ será dada por

$$\tilde{\varphi}(t) = \varphi_n(t)\overline{\varphi(t)} = e^{-\frac{t^2}{4}} e^{-\frac{t^2}{4}} = e^{-\frac{t^2}{2}},$$

e assim, $X_n - X$ tem distribuição normal padrão, qualquer que seja n. Assim, sendo $\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{s^2}{2}} ds$ a função de distribuição normal (que é simétrica em torno de zero), temos

$$\begin{aligned} P(|X_n - X| \geq \epsilon) &= P(X_n - X \geq \epsilon) + P(X_n - X \leq -\epsilon) \\ &= P(X_n - X \geq \epsilon) + P(X_n - X \geq \epsilon) \\ &= 2P(X_n - X \geq \epsilon) \\ &= 2(1 - \phi(\epsilon)). \end{aligned}$$

Assim temos que $P(|X_n - X| \geq \epsilon) \xrightarrow{n\to\infty} 2(1 - \phi(\epsilon)) \neq 0$ se $\epsilon \neq 0$, e portanto, não temos $X_n \xrightarrow{P} X$.

Proposição 3.1.3. *Se $X_n \xrightarrow{D} c$, onde c é uma constante real, então $X_n \xrightarrow{P} c$.*

Demonstração. Sejam $F, F_1, F_2, \cdots$ as funções de distribuição de $X, X_1, X_2, \cdots$ respectivamente, onde $X \equiv c$. Temos que

$$F(x) = \begin{cases} 1, & \text{se } x \geq c \\ 0, & \text{se } x < c \end{cases}.$$

Assim, todo $x \neq c$ é ponto de continuidade de F, e como $X_n \xrightarrow{D} c$, temos $F_n(x) \xrightarrow{n\to\infty} F(x)$ para todo $x \neq c$, isto é, $F_n(x) \xrightarrow{n\to\infty} 1$ se $x > c$ e $F_n(x) \xrightarrow{n\to\infty} 0$ se $x < c$. Logo, dado $\epsilon > 0$, temos

$$\begin{aligned} P(|X_n - c| \leq \epsilon) &= P(c - \epsilon \leq X_n \leq c + \epsilon) \\ &\geq P(c - \epsilon < X_n \leq c + \epsilon) \\ &= P(X_n \leq c + \epsilon) - P(X_n \leq c - \epsilon) \\ &= F_n(c + \epsilon) - F_n(c - \epsilon) \xrightarrow{n\to\infty} 1. \end{aligned}$$

Assim, $P(|X_n - c| > \epsilon) = 1 - P(|x_n - c| \leq \epsilon) \xrightarrow{n\to\infty} 0$, e portanto, $X_n \xrightarrow{P} c$. □

Pelo Exemplo 3.1.2 fica claro que a recíproca da Proposição 3.1.2 não vale em geral, mas como vimos na Proposição 3.1.3, temos que vale no caso em que X degenera em um ponto, e isso será de grande utilidade na demonstração da Lei dos Grande Números.

Recorde que no início da seção fizemos a seguinte pergunta: qual é a relação entre a convergência da sequência de variáveis aleatórias e a convergência das sequências de funções de distribuição e funções características que surgem? Pelo que vimos até agora temos que a convergência em quase toda parte implica a convergência em probabilidade, que por sua vez implica a convergência em distribuição, enquanto que a recíproca não vale em geral em nenhum dos dois casos. Falta então apenas analisar como estas convergências se relacionam com as funções características. Como vimos no Corolário 2.4.1, há uma relação biunívoca entre função de distribuição e função característica, e portanto é natural pensarmos que a convergência da sequência de funções características também estabeleça tal relação com a convergência em distribuição. De fato isso ocorre, como veremos nos resultados seguintes, os quais podem ser encontrados em (JAMES, 2006, Capítulo 6).

Teorema 3.1.1 (Teorema de Helly-Bray)**.** *Sejam $F, F_1, F_2, \cdots$ funções de distribuição com $\{F_n\}_{n\geq 1}$ convergindo fracamente para F, isto é, $F_n(x) \overset{n\to\infty}{\longrightarrow} F(x)$, para todo x que é ponto de continuidade de F. Então,*

$$\int_{-\infty}^{\infty} g(x)dF_n(x) \overset{n\to\infty}{\longrightarrow} \int_{-\infty}^{\infty} g(x)dF(x)$$

para toda função $g : \mathbb{R} \to \mathbb{R}$ contínua e limitada. Em particular, sendo $\varphi, \varphi_1, \varphi_2, \cdots$ as respectivas funções características de $F, F_1, F_2, \cdots$, temos que $F_n \overset{D}{\longrightarrow} F$ implica $\varphi_n(t) \overset{n\to\infty}{\longrightarrow} \varphi(t)$, para todo $t \in \mathbb{R}$.

Demonstração. Dados a e b tais que $-\infty < a < b < \infty$, temos

$$\left| \int_{-\infty}^{\infty} g(x)dF_n(x) - \int_{-\infty}^{\infty} g(x)dF(x) \right| \leq I + II + III,$$

onde

$$\begin{aligned} I &= \left| \int_{-\infty}^{\infty} g(x)dF_n(x) - \int_a^b g(x)dF_n(x) \right|, \\ II &= \left| \int_a^b g(x)dF_n(x) - \int_a^b g(x)dF(x) \right|, \\ III &= \left| \int_a^b g(x)dF(x) - \int_{-\infty}^{\infty} g(x)dF(x) \right|. \end{aligned}$$

Como g é limitada, temos $\sup\limits_{x\in\mathbb{R}} |g(x)| = c < \infty$, e logo,

$$\begin{aligned} III &= \left| \int_{-\infty}^{a} g(x)dF(x) + \int_b^{\infty} g(x)dF(x) \right| \leq \left| \int_{-\infty}^{a} g(x)dF(x) \right| + \left| \int_b^{\infty} g(x)dF(x) \right| \\ &\leq \int_{-\infty}^{a} |g(x)|dF(x) + \int_b^{\infty} |g(x)|dF(x) \leq c\left(\int_{-\infty}^{a} dF(x) + \int_b^{\infty} dF(x) \right) \\ &= c(F(a) + 1 - F(b)). \end{aligned}$$

Uma vez que, pela Proposição 2.1.2, os pontos de continuidade de F são densos em $\mathbb{R}$, e $\lim_{b\to\infty}\lim_{a\to-\infty}[F(a)+1-F(b)]=0$, dado $\epsilon>0$, podemos escolher a suficientemente pequeno e b suficientemente grande entre os pontos de continuidade de F, de modo que

$$III \leq c(F(a)+1-F(b)) < \epsilon.$$

Como F_n converge fracamente para F, temos para esses mesmos a e b que

$$\begin{aligned}
I &= \left|\int_{-\infty}^{a} g(x)dF_n(x) + \int_{b}^{\infty} g(x)dF_n(x)\right| \leq \left|\int_{-\infty}^{a} g(x)dF_n(x)\right| + \left|\int_{b}^{\infty} g(x)dF_n(x)\right| \\
&\leq \int_{-\infty}^{a} |g(x)|dF_n(x) + \int_{b}^{\infty} |g(x)|dF_n(x) \leq c\left(\int_{-\infty}^{a} dF_n(x) + \int_{b}^{\infty} dF_n(x)\right) \\
&= c(F_n(a)+1-F_n(b)) \stackrel{n\to\infty}{\longrightarrow} c(F(a)+1-F(b)).
\end{aligned}$$

Daí, para n suficientemente grande temos $I<\epsilon$, e logo, $I+III<2\epsilon$. Agora, como g é uma função contínua em $\mathbb{R}$, temos que g é uniformemente contínua no intervalo compacto $[a,b]$, e logo podemos escolher $x_0, x_1, \cdots, x_N$ tais que:

$$\begin{aligned}
&a = x_0 < x_1 < \cdots < x_N = b, \\
&x_0, x_1, \cdots, x_N \text{ são pontos de continuidade de F, e} \\
&|g(x)-g(x_i)| < \epsilon \text{ para todo } x \in [x_i, x_{i+1}],\ i = 0, 1, \cdots, N-1.
\end{aligned}$$

Segue agora que

$$\begin{aligned}
m_{n_i} &\stackrel{def}{=} (g(x_i)-\epsilon)[F_n(x_{i+1})-F_n(x_i)] \leq \int_{x_i}^{x_{i+1}} g(x)dF_n(x) \\
&\leq (g(x_i)+\epsilon)[F_n(x_{i+1})-F_n(x_i)] \stackrel{def}{=} M_{n_i}
\end{aligned}$$

e

$$\begin{aligned}
m_i &\stackrel{def}{=} (g(x_i)-\epsilon)[F(x_{i+1})-F(x_i)] \leq \int_{x_i}^{x_{i+1}} g(x)dF(x) \\
&\leq (g(x_i)+\epsilon)[F(x_{i+1})-F(x_i)] \stackrel{def}{=} M_i,
\end{aligned}$$

e consequentemente,

$$m_{n_i} - M_i \leq \int_{x_i}^{x_{i+1}} g(x)dF_n(x) - \int_{x_i}^{x_{i+1}} g(x)dF(x) \leq M_{n_i} - m_i,$$

qualquer que seja $i = 0, 1, \cdots, N-1$. Somando termo a termo segue que

$$\sum_{i=0}^{N-1}(m_{n_i} - M_i) \leq \int_a^b g(x)dF_n(x) - \int_a^b g(x)dF(x) \leq \sum_{i=0}^{N-1}(M_{n_i} - m_i).$$

Como os x_i são pontos de continuidade de F e $\{F_n\}_{n\geq 1}$ converge fracamente para F, temos que $m_{n_i} \stackrel{n\to\infty}{\longrightarrow} m_i$ e $M_{n_i} \stackrel{n\to\infty}{\longrightarrow} M_i$. Logo,

$$\begin{aligned}
\sum_{i=0}^{N-1}(m_{n_i} - M_i) &\stackrel{n\to\infty}{\longrightarrow} \sum_{i=0}^{N-1}(m_i - M_i) \\
&= \sum_{i=0}^{N-1}[(g(x_i)-\epsilon)(F(x_{i+1})-F(x_i)) - (g(x_i)+\epsilon)(F(x_{i+1})-F(x_i))] \\
&= -2\epsilon\sum_{i=0}^{N-1}(F(x_{i+1})-F(x_i)) \\
&= -2\epsilon(F(b)-F(a)) \geq -2\epsilon,
\end{aligned}$$

e analogamente,

$$\sum_{i=0}^{N-1}(M_{n_i} - m_i) \stackrel{n\to\infty}{\longrightarrow} \sum_{i=0}^{N-1}(M_i - m_i) - 2\epsilon(F(b)-F(a)) \leq 2\epsilon.$$

Assim, para n suficientemente grande temos

$$-3\epsilon \leq \int_a^b g(x)dF_n(x) - \int_a^b g(x)dF(x) \leq 3\epsilon,$$

e consequentemente,

$$\left|\int_{-\infty}^{\infty} g(x)dF_n(x) - \int_{-\infty}^{\infty} g(x)dF(x)\right| = I + II + III \leq 5\epsilon,$$

donde segue que

$$\int_{-\infty}^{\infty} g(x)dF_n(x) \stackrel{n\to\infty}{\longrightarrow} \int_{-\infty}^{\infty} g(x)dF(x).$$

□

Lema 3.1.1. *Suponha que para toda sequência de funções de distribuição $\{F_n\}_{n\geq 1}$, cujas funções características correspondentes convergem pontualmente para φ, existe uma subsequência $\{F_{n_j}\}_{j\geq 1}$ e uma função de distribuição F de forma que $\{F_{n_j}\}_{j\geq 1}$ converge fracamente para F. Então a sequência $\{F_n\}_{n\geq 1}$ de funções de distribuição converge fracamente para F. Note que ainda não podemos dizer que φ é uma função característica.*

Demonstração. Suponha por absurdo que $\{F_{n_j}\}_{j\geq 1}$ converge fracamente para F, mas que o mesmo não ocorre com $\{F_n\}_{n\geq 1}$. Então, existirão um ponto x no qual F é contínua e uma subsequência $\{F_{n'}\}_{n'\geq 1}$ de forma que $F_{n'}(x) \stackrel{n'\to\infty}{\longrightarrow} a \neq F(x)$. Como as funções características correspondentes a $\{F_n\}_{n\geq 1}$ convergem para φ, o mesmo ocorre com as funções características de $\{F_{n'}\}_{n'\geq 1}$, e portanto, pelas hipóteses do lema, $\{F_{n'}\}_{n'\geq 1}$ possui uma subsequência $\{F_{\tilde{n}}\}_{\tilde{n}\geq 1}$ que converge fracamente para G. Como F e G têm a mesma função característica, temos pelo Corolário 2.4.1 que $F = G$. Desta forma temos que $F_{\tilde{n}}(x) \stackrel{\tilde{n}\to\infty}{\longrightarrow} G(x) = F(x)$, ao mesmo tempo que $F_{\tilde{n}}(x) \stackrel{\tilde{n}\to\infty}{\longrightarrow} a \neq F(x)$. Chegamos assim a uma contradição, e portanto $\{F_n\}_{n\geq 1}$ converge fracamente para F. □

Teorema 3.1.2 (Teorema da Continuidade de Paul Lévy)**.** *Sejam $F_1, F_2, \cdots$ funções de distribuição cujas funções características correspondentes são $\varphi_1, \varphi_2, \cdots$, respectivamente. Se φ_n converge pontualmente para φ que é contínua em $t = 0$, então temos que*

a) existe uma função de distribuição F tal que F_n converge fracamente para F, e

b) φ é a função característica correspondente a F.

Demonstração. Note inicialmente que é suficiente provar o primeiro item, pois a partir daí o Teorema de Helly-Bray garante a validade do segundo. A demonstração será feita a partir das seguintes afirmações:

Afirmação 1. Existem uma subsequência $F_{n_1}, F_{n_2}, \cdots$ e uma função $F : \mathbb{R} \to [0, 1]$ não decrescente e contínua à direita onde $F_{n_j}(x) \stackrel{j\to\infty}{\longrightarrow} F(x)$ para todo x que é ponto de continuidade de F.

Afirmação 2. A função F obtida é uma função de distribuição.

Como a sequência $F_1, F_2, \cdots$, tomada é arbitrária, a existência de F no primeiro item segue do Lema 3.1.1.

Para provar a afirmação 1 observe que a sequência $\{F_n\}_{n\geq 1}$ é limitada, e logo, sendo $\mathbb{Q} = \{r_1, r_2, \cdots\}$ o conjunto (enumerável) dos números racionais, temos, pelo método da diagonal de Cantor (LIMA, 2009, Capítulo 10, Teorema 21), que existe uma subsequência $\{F_{n_j}\}_{j\geq 1}$ que converge pontualmente em $\mathbb{Q}$. Vamos definir F então no conjunto dos racionais por

$$F(x) = \lim_{j\to\infty} F_{n_j}(x),\ x \in \mathbb{Q}.$$

Como temos para todo n que $F_n(r_i) \in [0, 1]$ e F_n é não decrescente, o mesmo acontecerá com F no conjunto dos racionais.

Para x irracional vamos definir $F(x)$ como $\lim\limits_{\substack{r\searrow x \\ r\in\mathbb{Q}}} F(r)$, e teremos F não decrescente em $\mathbb{R}$. Além disso, $F_{n_j}(x) \stackrel{j\to\infty}{\longrightarrow} F(x)$ para todo x que é ponto de continuidade de F, pois dado $\epsilon > 0$, existem $\hat{r}$ e $\tilde{r}$ racionais tais que $\hat{r} < x < \tilde{r}$ e $F(\tilde{r}) - \epsilon < F(x) < F(\hat{r}) + \epsilon$, e logo

$$\begin{aligned} F(x) - \epsilon &< F(\hat{r}) = \lim_{j\to\infty} F_{n_j}(\hat{r}) \leq \lim_{j\to\infty} \inf F_{n_j}(x) \\ &\leq \lim_{j\to\infty} \sup F_{n_j}(x) \leq \lim_{j\to\infty} F_{n_j}(\tilde{r}) = F(\tilde{r}) < F(x) + \epsilon. \end{aligned}$$

Como a exigência é que $F_{n_j}(x) \stackrel{j\to\infty}{\longrightarrow} F(x)$ apenas nos pontos de continuidade de F, podemos redefinir F nos pontos de descontinuidade, de forma que F seja contínua à direita.

Para provar a afirmação 2, basta provarmos ainda que $F(\infty) = 1$ e $F(-\infty) = 0$. Para tanto, dada a função característica g correspondente a uma função de distribuição G, vamos definir a **função característica integrada** $\hat{g}$ por

$$\begin{aligned}\hat{g}(t) &= \int_0^t g(s)ds = \int_0^t \int_{-\infty}^{\infty} e^{isx} dG(x)ds \\ &= \int_{-\infty}^{\infty} \int_0^t e^{isx} ds dG(x) = \int_{-\infty}^{\infty} \frac{e^{itx} - 1}{ix} dG(x),\end{aligned}$$

onde a troca na ordem de integração se justifica no fato de que

$$\left| \int_0^t g(s)ds \right| \leq \int_0^t |g(s)|ds \leq \int_0^t ds \leq t.$$

Para t fixo, $\frac{e^{itx}-1}{ix}$ é uma função limitada e contínua, já que $\lim\limits_{x\to 0} \frac{e^{itx}-1}{ix} = t$ e $\lim\limits_{x\to\infty} \frac{e^{itx}-1}{ix} = \lim\limits_{x\to-\infty} \frac{e^{itx}-1}{ix} = 0$. Assim, de forma análoga à prova do Teorema de Helly-Bray (CHUNG, 2001, Teorema 4.4.1), temos que

$$\int_{-\infty}^{\infty} \frac{e^{itx} - 1}{ix} dF_{n_j}(x) \overset{j\to\infty}{\longrightarrow} \int_{-\infty}^{\infty} \frac{e^{itx} - 1}{ix} dF(x),$$

isto é,

$$\hat{\varphi}_{n_j}(t) = \int_0^t \varphi_{n_j}(s)ds \overset{j\to\infty}{\longrightarrow} \int_0^t \int_{-\infty}^{\infty} e^{isx} dF(x)ds. \quad (3.1)$$

Como φ_{n_j} é uma função contínua (pois é uma função característica), temos φ_{n_j} mensurável para todo j, e como $\varphi_{n_j} \overset{j\to\infty}{\longrightarrow} \varphi$, φ também é mensurável. Assim, sendo $I_{[0,t]}(s)$ dada por

$$I_{[0,t]}(s) = \begin{cases} 1, & \text{se } 0 \leq s \leq t \\ 0, & \text{caso contrário,} \end{cases}$$

temos do Teorema da Convergência Dominada (aplicado à parte real e à parte imaginária de $I_{[0,t]}(s)\varphi_{n_j}$) que

$$\hat{\varphi}_{n_j}(t) = \int_0^t \varphi_{n_j}(s)ds = \int_{-\infty}^{\infty} I_{[0,t]}(s)\varphi_{n_j}(s)ds \overset{j\to\infty}{\longrightarrow} \int_{-\infty}^{\infty} I_{[0,t]}(s)\varphi(s)ds = \int_0^t \varphi(s)ds. \quad (3.2)$$

Das equações (3.1) e (3.2) segue que

$$\int_0^t \varphi(s)ds = \int_0^t \int_{-\infty}^{\infty} e^{isx} dF(x)ds,$$

e logo, para todo $t \neq 0$,

$$\frac{1}{t} \int_0^t \varphi(s)ds = \frac{1}{t} \int_0^t \int_{-\infty}^{\infty} e^{isx} dF(x)ds. \quad (3.3)$$

Além disso, como $e^{itx} \overset{t\to s}{\longrightarrow} e^{isx}$ e $|e^{itx}| \leq 1$, temos novamente do Teorema da Convergência Dominada que $h(s) = \int_{-\infty}^{\infty} e^{isx}$ é contínua em $s = 0$. Como φ também é contínua em $s = 0$, podemos derivar $\int_0^t \varphi(s)ds$ e $\int_0^t \int_{-\infty}^{\infty} e^{isx} dF(x)ds$ numa vizinhança de $s = 0$, e logo, pela regra de L'Hôpital temos

$$\lim_{t\to 0} \frac{1}{t} \int_0^t \varphi(s)ds = \lim_{t\to 0} \varphi(t) = \varphi(0) \quad \text{e} \quad (3.4)$$

$$\lim_{t\to 0}\frac{1}{t}\int_0^t\int_{-\infty}^{\infty} e^{isx}dF(x)ds = \lim_{t\to 0}\int_{-\infty}^{\infty} e^{itx}dF(x) = \int_{-\infty}^{\infty} dF(x), \qquad (3.5)$$

onde a última igualdade decorre do Teorema da Convergência Dominada aplicado a e^{itx}. Como $\varphi(0) = \lim_{n\to\infty}\varphi_n(0) = 1$, temos das equações (3.3), (3.4) e (3.5) que

$$F(\infty) - F(-\infty) = \int_{-\infty}^{\infty} dF(x) = \varphi(0) = 1,$$

e como $0 \leq F(x) \leq 1$, temos $F(\infty) = 1$ e $F(-\infty) = 0$. $\square$

Corolário 3.1.1. *$X_n \xrightarrow{D} X$ se, e somente se, φ_{X_n} converge pontualmente para φ_X.*

Demonstração. Se $X_n \xrightarrow{D} X$, temos pelo Teorema de Helly-Bray que para $t \in \mathbb{R}$ fixo,

$$\mathbb{E}\cos(tX_n) = \int_{-\infty}^{\infty}\cos(tx)dF_{X_n}(x) \overset{n\to\infty}{\longrightarrow} \int_{-\infty}^{\infty}\cos(tx)dF_X(x) = \mathbb{E}\cos(tX) \quad \text{e}$$

$$\mathbb{E}\sin(tX_n) = \int_{-\infty}^{\infty}\sin(tx)dF_{X_n}(x) \overset{n\to\infty}{\longrightarrow} \int_{-\infty}^{\infty}\sin(tx)dF_X(x) = \mathbb{E}\sin(tX),$$

uma vez que $\cos(tx)$ e $\sin(tx)$ são funções contínuas e limitadas. Segue que

$$\varphi_{X_n}(t) = \mathbb{E}\cos(tX_n) + i\mathbb{E}\sin(tX_n) \overset{n\to\infty}{\longrightarrow} \mathbb{E}\cos(tX) + i\mathbb{E}\sin(tX) = \varphi_X(t),$$

qualquer que seja $t \in \mathbb{R}$, isto é, φ_{X_n} converge pontualmente para φ_X.

Reciprocamente, suponha que φ_{X_n} converge pontualmente para φ_X. Como φ_X é a função característica de X, temos que φ_X é contínua em $t = 0$. Logo, pelo Teorema da Continuidade de Paul-Lévy, temos $X_n \xrightarrow{D} X$. $\square$

No corolário acima temos, como aplicação do Teorema de Helly-Bray, que a convergência em distribuição implica a convergência pontual das funções características. No teorema que segue vamos mostrar, que nesse caso, a convergência das funções característica é uniforme em cada intervalo limitado.

Teorema 3.1.3. *Sejam $F, F_1, F_2, \cdots$ funções de distribuição, com $\{F_n\}_{n\geq 1}$ convergindo fracamente para F, e $\varphi, \varphi_1, \varphi_2, \cdots$, as respectivas funções características. Então, $\{\varphi_n\}_{n\geq 1}$ converge uniformemente para φ em todo intervalo limitado.*

Demonstração. Acabamos de ver, no corolário acima, que $\{\varphi_n\}_{n\geq 1}$ converge pontualmente para φ. Para mostrar que a convergência é uniforme, dado $\epsilon > 0$, observe que para quaisquer t e h em $\mathbb{R}$, temos

$$\begin{aligned}|\varphi_n(t+h) - \varphi_n(t)| &= \left|\int_{-\infty}^{\infty} e^{i(t+h)x}dF_n(x) - \int_{-\infty}^{\infty} e^{itx}dF_n(x)\right| \\ &\leq \int_{-\infty}^{\infty} |e^{ihx} - 1|dF_n(x).\end{aligned}$$

Daí temos, para A grande, que

$$\begin{aligned}
|\varphi_n(t+h)-\varphi_n(t)| &\leq \int_{|x|\leq A}|e^{ihx}-1|dF_n(x)+\int_{|x|>A}|e^{ihx}-1|dF_n(x)\\
&\leq \int_{|x|\leq A}|hx|dF_n(x)+2\int_{|x|>A}dF_n(x)\\
&\leq |h|A+2\int_{|x|>A}dF_n(x).
\end{aligned}$$

Por outro lado, como $\{F_n\}_{n\geq 1}$ converge fracamente para F, existe $n_0>0$ tal que,

$$\begin{aligned}
n\geq n_0 \quad &\Rightarrow \quad \left|\int_{|x|>A}dF_n(x)-\int_{|x|>A}dF(x)\right|<\frac{\epsilon}{6}\\
&\Rightarrow \quad \left|\int_{|x|>A}dF_n(x)\right|<\left|\int_{|x|>A}dF(x)\right|+\frac{\epsilon}{6},
\end{aligned}$$

e logo,

$$|\varphi_n(t+h)-\varphi_n(t)|\leq |h|A+2\int_{|x|>A}dF(x)+\frac{\epsilon}{3}.$$

Agora, tomando A suficientemente grande, temos

$$|\varphi_n(t+h)-\varphi_n(t)|\leq |h|A+\frac{2\epsilon}{3},$$

e daí, tomando h suficientemente pequeno, teremos

$$|\varphi_n(t+h)-\varphi_n(t)|\leq \epsilon.$$

Assim, temos que $\{\varphi_n\}_{n\geq 1}$ é uma família de funções equicontínua e limitada, e logo, pelo Teorema de Arzelá-Ascoli (LIMA, 2009, Capítulo 10, Teorema 22), converge uniformemente em cada intervalo limitado. □

3.2 Teoremas Limites

Nesta seção vamos começar a falar sobre a convergência das sequências do tipo $\frac{S_n+b_n}{a_n}$. Antes de enunciarmos a Lei dos Grandes Números e o Teorema Central do Limite, vamos expor alguns exemplos que deixam claro que tais sequências nem sempre convergem para uma variável aleatória que tenha distribuição normal ou degenerada. Tais exemplos serão a motivação para estudarmos as distribuições infinitamente divisíveis e estáveis no próximo capítulo. Outro fato que veremos adiante é que as sequências mencionadas podem ser escritas sob a forma de um arranjo triangular, e nesse contexto surge o Teorema de Poisson, o qual também estaremos abordando aqui.

Exemplo 3.2.1. *Sejam $X_1, X_2, \cdots$ variáveis aleatórias independentes onde X_n tem distribuição de Poisson com parâmetro $\lambda_n>0$ e $\sum\limits_{n=1}^{\infty}\lambda_n=\lambda$. Então, sendo $S_n=X_1+\cdots+X_n$, temos que $\frac{S_n-\mathbb{E}S_n}{\sqrt{VarS_n}}\xrightarrow{D} T=\frac{X-\lambda}{\sqrt{\lambda}}$, onde X tem distribuição de Poisson com parâmetro λ.*

Demonstração. De acordo com a Proposição 2.5.4, para cada n, a função característica de X_n será $\varphi_{X_n}(t) = e^{\lambda_n(e^{it}-1)}$, e como, pelas Proposições 2.3.3 e 2.3.4,

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}} = \frac{S_n - (\lambda_1 + \cdots + \lambda_n)}{\sqrt{\lambda_1 + \cdots + \lambda_n}} = \frac{S_n}{\sqrt{\lambda_1 + \cdots + \lambda_n}} - \sqrt{\lambda_1 + \cdots + \lambda_n},$$

de acordo com a Proposição 2.4.1, a função característica de $\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}}$ será

$$\begin{aligned}
\varphi_n(t) &= e^{-it\sqrt{\lambda_1+\cdots+\lambda_n}}\varphi_{S_n}\left(\frac{t}{\sqrt{\lambda_1 + \cdots + \lambda_n}}\right) \\
&= e^{-it\sqrt{\lambda_1+\cdots+\lambda_n}}\varphi_{X_1}\left(\frac{t}{\sqrt{\lambda_1 + \cdots + \lambda_n}}\right)\cdots\varphi_{X_n}\left(\frac{t}{\sqrt{\lambda_1 + \cdots + \lambda_n}}\right) \\
&= e^{-it\sqrt{\lambda_1+\cdots+\lambda_n}}\exp\left(\lambda_1(e^{it\backslash\sqrt{\lambda_1+\cdots+\lambda_n}} - 1)\right)\cdots\exp\left(\lambda_n(e^{it\backslash\sqrt{\lambda_1+\cdots+\lambda_n}} - 1)\right) \\
&= e^{-it\sqrt{\lambda_1+\cdots+\lambda_n}}\exp\left((\lambda_1 + \cdots + \lambda_n)(e^{it\backslash\sqrt{\lambda_1+\cdots+\lambda_n}} - 1)\right).
\end{aligned}$$

Segue que $\varphi_n(t) \overset{n\to\infty}{\longrightarrow} e^{-it\sqrt{\lambda}}e^{\lambda(e^{i\frac{t}{\sqrt{\lambda}}}-1)}$, que é uma função contínua em $t = 0$, e logo, pelo Teorema da Continuidade de Paul-Lévy, é função característica da variável aleatória T tal que $\frac{S_n - \mathbb{E}S_n}{s_n} \overset{D}{\longrightarrow} T$. Para finalizar, das Proposições 2.4.1 e 2.5.4, temos $T = \frac{X-\lambda}{\sqrt{\lambda}}$ onde X tem distribuição de Poisson com parâmetro λ. □

Exemplo 3.2.2. *Sejam $X_1, X_2, \cdots$ variáveis aleatórias independentes com distribuição de Cauchy com parâmetro $\theta > 0$ e $S_n = X_1 + \cdots + X_n$. Então $\frac{S_n}{n} \overset{D}{\longrightarrow} T$, onde T também tem distribuição de Cauchy com parâmetro θ.*

Demonstração. De acordo com o Corolário 2.5.3 as funções características de $X_1, X_2, \cdots$ serão todas iguais a $\varphi(t) = e^{-\theta|t|}$, e de acordo com a Proposição 2.4.1 a função característica de $\frac{S_n}{n}$ será

$$\varphi_n(t) = \varphi_{\frac{X_1}{n}}(t)\cdots\varphi_{\frac{X_n}{n}}(t) = \varphi\left(\frac{t}{n}\right)\cdots\varphi\left(\frac{t}{n}\right) = \left(e^{-\frac{\theta|t|}{n}}\right)^n = e^{-\theta|t|},$$

qualquer que seja n. Assim, $\varphi_n(t) \overset{n\to\infty}{\longrightarrow} e^{-\theta|t|}$, e pelo Teorema da Continuidade de Paul-Lévy $\frac{S_n}{n} \overset{D}{\longrightarrow} T$, onde T tem distribuição de Cauchy com parâmetro θ. □

Lema 3.2.1. *Seja $c_{n,k}$, $n \geq 1$ e $1 \leq k \leq n$, uma sequência de números complexos tais que $\sum\limits_{k=1}^{n} c_{n,k} \overset{n\to\infty}{\longrightarrow} c$, $\max\limits_{1\leq k\leq n}|c_{n,k}| \overset{n\to\infty}{\longrightarrow} 0$ e $\sum\limits_{k=1}^{n}|c_{n,k}| \leq M < \infty$, onde M é uma constante que não depende de n. Então,*

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

Demonstração. De $\max\limits_{1\leq k\leq n}|c_{n,k}| \overset{n\to\infty}{\longrightarrow} 0$ temos que existe n_0 tal que $|c_{n,k}| \leq \frac{1}{2}$ para todo $n \geq n_0$, e logo $1 + c_{n,k} \in B = \{z \in \mathbb{C}; |1 - z| \leq \frac{1}{2}\}$. Logo podemos considerar o logaritmo principal em $\mathbb{C} - \mathbb{R}^-$, o qual em B assume a forma

$$\ln(1 + z) = \sum_{j=1}^{\infty}\frac{(-1)^{j-1}z^j}{j} = z + \sum_{j=2}^{\infty}\frac{(-1)^{j-1}z^j}{j},\ |z| \leq \frac{1}{2}.$$

Segue que para $n \geq n_0$,

$$\sum_{k=1}^{n} \ln(1 + c_{n,k}) = \sum_{k=1}^{n} c_{n,k} + \sum_{k=1}^{n} \sum_{j=2}^{\infty} \frac{(-1)^{j-1} c_{n,k}^j}{j}.$$

Agora, observe que

$$\begin{aligned}
\lim_{n\to\infty} \left| \sum_{k=1}^{n} \sum_{j=2}^{\infty} \frac{(-1)^{j-1} c_{n,k}^j}{j} \right| &\leq \lim_{n\to\infty} \sum_{k=1}^{n} \left| \sum_{j=2}^{\infty} \frac{(-1)^{j-1} c_{n,k}^j}{j} \right| \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \left(|c_{n,k}|^2 \left| \sum_{j=0}^{\infty} \frac{(-1)^{j+1} c_{n,k}^j}{j+2} \right| \right) \\
&\leq \lim_{n\to\infty} \max_{1\leq k\leq n} |c_{n,k}| \sum_{k=1}^{n} \left(|c_{n,k}| \left| \sum_{j=0}^{\infty} \frac{(-1)^{j+1} c_{n,k}^j}{j+2} \right| \right) \\
&\leq \lim_{n\to\infty} \max_{1\leq k\leq n} |c_{n,k}| \sum_{k=1}^{n} \left(|c_{n,k}| \left| \sum_{j=0}^{\infty} \frac{(-1)^{j+1} \max_{1\leq k\leq n} c_{n,k}^j}{j+2} \right| \right) \\
&\leq \lim_{n\to\infty} \left(\max_{1\leq k\leq n} |c_{n,k}| \right) \left(\sum_{k=1}^{n} |c_{n,k}| \right) \left(\left| \sum_{j=0}^{\infty} \frac{(-1)^{j+1} \max_{1\leq k\leq n} c_{n,k}^j}{j+2} \right| \right) \\
&= 0,
\end{aligned}$$

uma vez que $\max_{1\leq k\leq n} |c_{n,k}| \stackrel{n\to\infty}{\longrightarrow} 0$, $\sum_{k=1}^{n} |c_{n,k}| \leq M < \infty$ para todo n e $\sum_{j=0}^{\infty} \frac{(-1)^{j+1} \max_{1\leq k\leq n} c_{n,k}^j}{j+2} \stackrel{n\to\infty}{\longrightarrow} -\frac{1}{2}$. Segue portanto que

$$\lim_{n\to\infty} \sum_{k=1}^{n} \ln(1 + c_{n,k}) = \lim_{n\to\infty} \sum_{k=1}^{n} c_{n,k} = c,$$

e consequentemente,

$$\prod_{k=1}^{n} (1 + c_{n,k}) = \exp\left(\sum_{k=1}^{n} \ln(1 + c_{n,k}) \right) \stackrel{n\to\infty}{\longrightarrow} e^c.$$

□

Teorema 3.2.1 (Teorema de Poisson)**.** *Para todo $n \geq 1$ considere as variáveis aleatórias independentes $X_{n,1}, \cdots, X_{n,n}$ de modo que se forma a seguinte sequência de variáveis aleatórias, a qual chamaremos de Arranjo Triangular:*

$$\begin{aligned}
S_1 &= X_{1,1} \\
S_2 &= X_{2,1} + X_{2,2} \\
S_3 &= X_{3,1} + X_{3,2} + X_{3,3} \\
&\vdots \\
S_n &= X_{n,1} + X_{n,2} + X_{n,3} + \cdots + X_{n,n} \\
&\vdots
\end{aligned}$$

Suponha que para todo $1 \leq k \leq n$, *tenhamos* $P(X_{n,k} = 1) = p_{n,k}$ *e* $P(X_{n,k} = 0) = q_{n,k}$, *com* $p_{n,k}+q_{n,k} = 1$, $\max\limits_{1\leq k\leq n} p_{n,k} \overset{n\to\infty}{\longrightarrow} 0$ *e* $\sum\limits_{k=1}^{n} p_{n,k} \overset{n\to\infty}{\longrightarrow} \lambda > 0$. *Então, para todo* $m = 0, 1, 2, \cdots$ *vale que*

$$P(S_n = m) \overset{n\to\infty}{\longrightarrow} \frac{e^{-\lambda}\lambda^m}{m!},$$

isto é, $S_n \xrightarrow{D}$ *Poisson*(λ).

Demonstração. Para todo $1 \leq k \leq n$ temos da Proposição 2.3.1 que a função característica de $X_{n,k}$ será

$$\varphi_{X_{n,k}}(t) = \mathbb{E}e^{itX_{n,k}} = \sum_{i\in\{0,1\}} e^{itx_i} P(X_{n,k} = x_i) = e^{it}p_{n,k} + q_{n,k}.$$

Como as variáveis aleatórias são independentes, temos da Proposição 2.4.1 que a função característica de S_n será

$$\begin{aligned} \varphi_n(t) &= \varphi_{X_{n,1}}(t)\cdots\varphi_{X_{n,n}}(t) = \prod_{k=1}^{n}(e^{it}p_{n,k} + q_{n,k}) \\ &= \prod_{k=1}^{n}(e^{it}p_{n,k} + 1 - p_{n,k}) = \prod_{k=1}^{n}(1 + p_{n,k}(e^{it} - 1)). \end{aligned}$$

Agora, tomando $c_{n,k} = p_{n,k}(e^{it} - 1)$, temos claramente do lema anterior que

$$\sum_{k=1}^{n} c_{n,k} \overset{n\to\infty}{\longrightarrow} \lambda(e^{it} - 1). \tag{3.6}$$

Além disso, como $|c_{n,k}| = p_{n,k}|(e^{it} - 1)| \leq 2p_{n,k}$ para quaisquer k e n, temos

$$\max_{1`k\leq n} |c_{n,k}| \overset{n\to\infty}{\longrightarrow} 0, \tag{3.7}$$

ao mesmo tempo que

$$\sum_{k=1}^{n} |c_{n,k}| \leq 2\sum_{k=1}^{n} p_{n,k} \overset{n\to\infty}{\longrightarrow} 2\lambda < \infty. \tag{3.8}$$

Por (3.6), (3.7) e (3.8), segue do Lema 3.2.1 que

$$\varphi_n(t) \overset{n\to\infty}{\longrightarrow} e^{\lambda(e^{it}-1)},$$

e portanto, do Teorema da Continuidade de Paul-Lévy, $S_n \xrightarrow{D}$ *Poisson*(λ). □

Uma sequência de variáveis aleatórias $X_1, X_2, \cdots$ satisfaz a Lei Fraca dos Grandes Números quando $\frac{S_n - \mathbb{E}S_n}{n} \xrightarrow{P} 0$, onde $S_n = X_1 + \cdots + X_n$. Analogamente, dizer que $X_1, X_2, \cdots$ satisfazem a Lei Forte dos Grandes Números significa que $\frac{S_n - \mathbb{E}S_n}{n}$ converge em quase toda parte para 0. Note que o Exemplo 3.1.1 e a Proposição 3.1.1 deixam claro que a Lei Forte implica a Lei Fraca, enquanto a recíproca não vale em geral, justificando assim o uso dos termos Forte e Fraca. No caso em que as variáveis aleatórias são identicamente distribuídas com média $m < \infty$, temos $\frac{S_n - \mathbb{E}S_n}{n} = \frac{S_n}{n} - m$, e logo, $\frac{S_n - \mathbb{E}S_n}{n} \xrightarrow{P} 0$ é equivalente a $\frac{S_n}{n} \xrightarrow{P} m$.

Teorema 3.2.2 (A Lei Fraca dos Grandes Números de Khintchin)**.** *Sejam $X_1, X_2, \cdots$ variáveis aleatórias independentes e identicamente distribuídas com $\mathbb{E}X_i = m < \infty$ para todo $i = 1, 2, \cdots$, e $S_n = X_1 + \cdots + X_n$. Então,*

$$\frac{S_n}{n} \xrightarrow{P} m.$$

Noutros termos, $X_1, X_2, \cdots$ satisfazem a Lei Fraca dos Grandes Números.

Demonstração. Seja $\varphi(t)$ a função característica de $X_1, X_2, \cdots$, e $\varphi_n(t)$ a função característica de $\frac{S_n}{n} = \frac{X_1+\cdots+X_n}{n}$. Como as variáveis aleatórias são independentes, temos da Proposição 2.4.1 que

$$\varphi_n(t) = \varphi_{\frac{X_1}{n}}(t) \cdots \varphi_{\frac{X_n}{n}}(t) = \left(\varphi\left(\frac{t}{n}\right)\right)^n.$$

Por outro lado, como $m < \infty$, temos do Teorema 2.4.2 que φ possui pelo menos uma derivada contínua, a qual no ponto $t = 0$ é dada por $\varphi'(0) = im$. Além disso, pela expansão de Taylor temos, para t próximo de zero, que

$$\varphi(t) = \varphi(0) + \varphi'(0)t + O(t) = 1 + imt + O(t).$$

Daí, para todo $t \in \mathbb{R}$ vale que, para n suficientemente grande,

$$\varphi\left(\frac{t}{n}\right) = 1 + \frac{itm}{n} + O\left(\frac{1}{n}\right),$$

e consequentemente, tomando $c_{n,k} = \frac{itm}{n} + O\left(\frac{1}{n}\right)$, $k = 1, \cdots, n$, temos do Lema 3.2.1 que

$$\varphi_n(t) = \left(1 + \frac{itm}{n} + O\left(\frac{1}{n}\right)\right)^n \xrightarrow{n\to\infty} e^{itm}.$$

Como $\psi(t) = e^{itm}$ é a função característica da variável aleatória X degenerada em m, segue do Teorema da Continuidade de Paul-Lévy que $\frac{S_n}{n} \xrightarrow{D} m$, e consequentemente, pela Proposição 3.1.3, $\frac{S_n}{n} \xrightarrow{P} m$. □

Nas condições do Teorema 3.2.2 vale também a Lei Forte dos Grandes Números, isto é, $\frac{S_n}{n}$ converge em quase toda parte para m, o que é conhecido como a *Lei Forte de Kolmogorov* (JAMES, 2006, Capítulo 5).

Teorema 3.2.3 (Teorema Central do Limite para Variáveis Aleatórias Independentes e Identicamente Distribuídas)**.** *Sejam $X_1, X_2, \cdots$ variáveis aleatórias independentes e identicamente distribuídas com $\mathbb{E}X_i = m < \infty$ e $0 < VarX_i = \sigma^2 < \infty$ para todo $i = 1, 2, \cdots$, e $S_n = X_1 + \cdots + X_n$. Então, quando $n \to \infty$,*

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}} = \frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow{D} N(0,1).$$

Demonstração. Sem perda de generalidade vamos supor que $m = 0$, pois considerando a sequência formada por $Y_i = X_i - m$ teremos que o valor $\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}}$ não se altera quando trocamos X_i por Y_i. Seja então $\varphi(t)$ a função característica de $X_1, X_2, \cdots$, e $\varphi_n(t)$ a função característica de $\frac{S_n}{\sigma\sqrt{n}}$. Como as variáveis aleatórias são independentes, temos da Proposição 2.4.1 que

$$\varphi_n(t) = \varphi_{\frac{X_1}{\sigma\sqrt{n}}}(t) \cdots \varphi_{\frac{X_n}{\sigma\sqrt{n}}}(t) = \left(\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n.$$

Por outro lado, como $\mathbb{E}|X_i|^2 < \infty$, temos do Teorema 2.4.2 e do Corolário 2.4.4 que φ possui pelo menos duas derivadas contínuas, as quais no ponto $t = 0$ são dadas por $\varphi'(0) = im = 0$ e $\varphi''(0) = i^2\mathbb{E}X^2 = -\sigma^2$. Além disso, pela expansão de Taylor teremos, para t próximo zero, que

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)t^2}{2} + O(t^2) = 1 - \frac{\sigma^2 t^2}{2} + O(t^2).$$

Daí, segue que para todo $t \in \mathbb{R}$, quando n é suficientemente grande,

$$\varphi_n(t) = \left(\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n = \left(1 - \frac{\sigma^2 t^2}{2\sigma^2 n} + O\left(\frac{1}{n}\right)\right)^n = \left(1 - \frac{t^2}{2n} + O\left(\frac{1}{n}\right)\right)^n,$$

e consequentemente, tomando $c_{n,k} = -\frac{t^2}{2n} + O\left(\frac{1}{n}\right)$, $k = 1, \cdots, n$, temos do Lema 3.2.1 que

$$\varphi_n(t) \stackrel{n\to\infty}{\longrightarrow} e^{-\frac{t^2}{2}}.$$

Como $\psi(t) = e^{-\frac{t^2}{2}}$ é a função característica correspondente a distribuição normal padrão, segue do Teorema da Continuidade de Paul-Lévy que

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}} = \frac{S_n - nm}{\sigma\sqrt{n}} \stackrel{D}{\longrightarrow} N(0,1).$$

□

Para enunciarmos uma versão mais geral do teorema anterior, observe que para a expressão $\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}}$ fazer sentido, não é necessário que as variáveis aleatórias sejam identicamente distribuídas. Basta que as variâncias sejam todas finitas com pelo menos uma diferente de zero. Um fato que ocorre no entanto é que, sendo $s_n = \sqrt{VarS_n}$, as parcelas $\frac{X_k - \mathbb{E}X_K}{s_n}$ da soma $\frac{S_n - \mathbb{E}S_n}{s_n}$ são uniformemente pequenas para n grande. Vamos então tentar estabelecer condições sob as quais $\max\limits_{1\leq k\leq n} \frac{X_k - \mathbb{E}X_K}{s_n} \stackrel{n\to\infty}{\longrightarrow} 0$, ou equivalentemente, sendo μ_k a média e σ_k a variância de X_k, estabelecer condições sob as quais

$$\max_{1\leq k\leq n} \frac{\sigma_k^2}{s_n^2} = \max_{1\leq k\leq n} \frac{\mathbb{E}(X_k - \mu_k)^2}{s_n^2} = \max_{1\leq k\leq n} \frac{1}{s_n^2} \int_{-\infty}^{\infty} (x - \mu_k)^2 dF_k(x) \stackrel{n\to\infty}{\longrightarrow} 0.$$

Para tanto, observe que dado $\epsilon > 0$ arbitrário, temos

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

Como o resultado obtido não depende do índice k escolhido, segue que

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

Como $\epsilon > 0$ é arbitrário, para que $\max\limits_{1\leq k\leq n}\frac{\sigma_k^2}{s_n^2} \stackrel{n\to\infty}{\longrightarrow} 0$ é suficiente que tenhamos para todo $\epsilon > 0$

$$\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon S_n}(x-\mu_k)^2 dF_k(x) \stackrel{n\to\infty}{\longrightarrow} 0,$$

o que é conhecido como *condição de Lindeberg*. Daí temos o seguinte resultado:

Teorema 3.2.4 (Teorema Central do Limite de Lindeberg)**.** *Sejam $X_1, X_2, \cdots$ variáveis aleatórias independentes com $\mathbb{E}X_i = \mu_i < \infty$ e $VarX_i = \sigma_i^2 < \infty$ para todo $i = 1, 2, \cdots$, $\sigma_i > 0$ para pelo menos um índice i. Sendo $S_n = X_1 + \cdots + X_n$ e $s_n = \sqrt{VarS_n} = \sqrt{\sigma_1^2 + \cdots + \sigma_n^2}$, então, para que $\frac{S_n - \mathbb{E}S_n}{s_n} \stackrel{D}{\longrightarrow} N(0,1)$ quando $n \to \infty$, é suficiente que seja satisfeita a condição de Lindeberg.*

Demonstração. Sejam $\varphi_1, \varphi_2, \cdots$ as funções características e $F_1, F_2, \cdots$ as funções de distribuição de $X_1, X_2, \cdots$, respectivamente, e $\varphi_n(t)$ a função característica de $\frac{S_n - \mathbb{E}S_n}{s_n}$. Como as variáveis aleatórias são independentes, temos da Proposição 2.3.3 que

$$\varphi_n(t) = \mathbb{E}\exp\left(it\frac{S_n - \mathbb{E}S_n}{s_n}\right) = \prod_{k=1}^{n}\mathbb{E}\exp\left(it\frac{X_k - \mu_k}{s_n}\right), \tag{3.9}$$

e pelo Teorema da Continuidade de Paul-Lévy é suficiente mostrarmos que para todo $t \in \mathbb{R}$, $\varphi_n(t) \stackrel{n\to\infty}{\longrightarrow} e^{-\frac{t^2}{2}}$. Para tanto, fixado $t \in \mathbb{R}$, vamos usar as seguintes versões da fórmula de Taylor para e^{itx}:

$$e^{itx} = 1 + itx + \theta_1(x)\frac{t^2x^2}{2}, \text{ onde } |\theta_1(x)| \leq 1, \quad \text{e,}$$

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

Daí, dado $\epsilon > 0$ arbitrário, usando a primeira fórmula para $|x| > \epsilon$ e a segunda para $|x| \leq \epsilon$, teremos a seguinte representação para e^{itx}:

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

onde

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

e consequentemente,

$$\begin{aligned} \mathbb{E}\exp\left(it\frac{X_k-\mu_k}{s_n}\right) &= \int_{-\infty}^{\infty} \exp\left(it\frac{x-\mu_k}{s_n}\right) dF_k(x) \\ &= \int_{-\infty}^{\infty} \left[1 + it\left(\frac{x-\mu_k}{s_n}\right) - \frac{t^2}{2}\left(\frac{x-\mu_k}{s_n}\right)^2 + r_\epsilon\left(\frac{x-\mu_k}{s_n}\right)\right] dF_k(x) \\ &= 1 + it\mathbb{E}\left(\frac{X_k-\mu_k}{s_n}\right) - \frac{t^2}{2}\mathbb{E}\left(\frac{X_k-\mu_k}{s_n}\right)^2 \\ &\quad + \frac{t^2}{2}\int_{|x-\mu_k|>\epsilon s_n} \left[1+\theta_1\left(\frac{x-\mu_k}{s_n}\right)\right]\left(\frac{x-\mu_k}{s_n}\right)^2 dF_k(x) \\ &\quad + \frac{t^3}{6}\int_{|x-\mu_k|\le\epsilon s_n} \theta_2\left(\frac{x-\mu_k}{s_n}\right)\left(\frac{x-\mu_k}{s_n}\right)^3 dF_k(x). \end{aligned}$$

Agora, como $\mathbb{E}X_k = \mu_k$ e $VarX_k = \sigma_k^2$, segue que para todo $k = 1, \cdots, n$, $\mathbb{E}\left(\frac{X-\mu_k}{s_n}\right) = 0$ e $\mathbb{E}\left(\frac{X-\mu_k}{s_n}\right)^2 = \frac{VarX_k}{s_n^2} = \frac{\sigma_k^2}{s_n^2}$, e logo,

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

onde

$$\begin{aligned} \xi_{n,k} &= \frac{t^2}{2}\int_{|x-\mu_k|>\epsilon s_n} \left[1+\theta_1\left(\frac{x-\mu_k}{s_n}\right)\right]\left(\frac{x-\mu_k}{s_n}\right)^2 dF_k(x) \\ &\quad + \frac{t^3}{6}\int_{|x-\mu_k|\le\epsilon s_n} \theta_2\left(\frac{x-\mu_k}{s_n}\right)\left(\frac{x-\mu_k}{s_n}\right)^3 dF_k(x). \end{aligned}$$

Note que

$$\begin{aligned} |\xi_{n,k}| &\le \frac{t^2}{2}\int_{|x-\mu_k|>\epsilon s_n} \left|1+\theta_1\left(\frac{x-\mu_k}{s_n}\right)\right|\left(\frac{x-\mu_k}{s_n}\right)^2 dF_k(x) \\ &\quad + \frac{|t^3|}{6}\int_{|x-\mu_k|\le\epsilon s_n} \left|\theta_2\left(\frac{x-\mu_k}{s_n}\right)\right|\left|\frac{x-\mu_k}{s_n}\right|\left(\frac{x-\mu_k}{s_n}\right)^2 dF_k(x) \\ &\le t^2\int_{|x-\mu_k|>\epsilon s_n} \left(\frac{x-\mu_k}{s_n}\right)^2 dF_k(x) + \frac{|t^3|}{6}\int_{|x-\mu_k|\le\epsilon s_n} \epsilon.\left(\frac{x-\mu_k}{s_n}\right)^2 dF_k(x) \\ &\le \frac{t^2}{s_n^2}\int_{|x-\mu_k|>\epsilon s_n} (x-\mu_k)^2 dF_k(x) + \frac{\epsilon|t^3|}{6s_n^2}\int_{-\infty}^{\infty}(x-\mu_k)^2 dF_k(x) \\ &\le \frac{t^2}{s_n^2}\int_{|x-\mu_k|>\epsilon s_n} (x-\mu_k)^2 dF_k(x) + \frac{\epsilon|t^3|\sigma_k^2}{6s_n^2}, \end{aligned}$$

e logo temos que

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

Pela condição de Lindeberg a primeira parcela do termo à direita tende a zero quando $n \to \infty$. Logo, para n suficientemente grande, temos $\sum\limits_{k=1}^{n} |\xi_{n,k}| \leq \frac{\epsilon|t^3|}{3}$, e daí, tomando uma sequência de $\epsilon's$ tendendo a zero, temos que a sequência $\xi_{n,k}$ correspondente satisfaz

$$\sum_{k=1}^{n} |\xi_{n,k}| \overset{n\to\infty}{\longrightarrow} 0. \tag{3.10}$$

Assim, substituindo $\mathbb{E}\exp\left(it\frac{X_k-\mu_k}{s_n}\right)$ por $1 - \frac{t^2\sigma_k^2}{2s_n^2} + \xi_{n,k}$ em (3.9), temos

$$\varphi_n(t) = \prod_{k=1}^{n}\left(1 - \frac{t^2\sigma_k^2}{2s_n^2} + \xi_{n,k}\right),$$

e tomando os $\xi_{n,k}$ como em (3.10), temos do Lema 3.2.1 que

$$\varphi_n(t) \overset{n\to\infty}{\longrightarrow} e^{-\frac{t^2}{2}},$$

uma vez que, tomando $c_{n,k} = -\frac{t^2\sigma_k^2}{2s_n^2} + \xi_{n,k}$, temos

- $\sum\limits_{k=1}^{n} c_{n,k} \overset{n\to\infty}{\longrightarrow} -\frac{t^2}{2}$;
- $\sum\limits_{k=1}^{n} |c_{n,k}| \leq \frac{t^2\sum\limits_{k=1}^{n}\sigma_k^2}{2s_n^2} + \sum\limits_{k=1}^{n}\xi_{n,k} = \frac{t^2}{2} + \sum\limits_{k=1}^{n}\xi_{n,k} \overset{n\to\infty}{\longrightarrow} \frac{t^2}{2}$, e logo, existe $M > 0$ tal que $\sum\limits_{k=1}^{n}\xi_{n,k} \leq M$;
- $\max\limits_{1\leq k\leq n} |c_{n,k}| \leq \max\limits_{1\leq k\leq n} \frac{t^2\sigma_k^2}{2s_n^2} + \max\limits_{1\leq k\leq n} |\xi_{n,k}| \leq \frac{t^2}{2s_n^2}\max\limits_{1\leq k\leq n}\sigma_k^2 + \sum\limits_{k=1}^{n}|\xi_{n,k}| \overset{n\to\infty}{\longrightarrow} 0$ por (3.10) e pela condição de Lindeberg.

Para finalizar, o Teorema da Continuidade de Paul-Lévy garante que $\frac{S_n-\mathbb{E}S_n}{s_n} \xrightarrow{D} N(0,1)$.

□

Verificar se uma sequência de variáveis aleatórias satisfaz a condição de Lindeberg costuma ser uma tarefa bem complicada. No entanto, existem propriedades mais simples de serem verificadas e que implicam tal condição. Uma delas é a condição de *Liapunov*, que estaremos definindo no corolário que segue.

Corolário 3.2.1 (Teorema Central do Limite de Liapunov). *Sejam $X_1, X_2, \cdots$ variáveis aleatórias independentes com $\mathbb{E}X_i = \mu_i < \infty$ e $VarX_i = \sigma_i^2 < \infty$ para todo $i = 1, 2, \cdots$, e, $\sigma_i > 0$ para pelo menos um i. Sendo $S_n = X_1 + \cdots + X_n$ e $s_n = \sqrt{VarS_n} = \sqrt{\sigma_1^2 + \cdots + \sigma_n^2}$, então, para que $\frac{S_n-\mathbb{E}S_n}{s_n} \xrightarrow{D} N(0,1)$ quando $n \to \infty$, é suficiente que seja satisfeita a condição de Liapunov, isto é, que exista $\delta > 0$ tal que*

$$\frac{1}{s_n^{2+\delta}}\sum_{k=1}^{n}\mathbb{E}|X_k - \mu_k|^{2+\delta} \overset{n\to\infty}{\longrightarrow} 0.$$

Demonstração. Basta mostrarmos que a condição de Liapunov implica a condição de Lindeberg. Seja então $\epsilon > 0$ arbitrário. Se $|x - \mu_k| > \epsilon s_n$, então $\frac{|x-\mu_k|^\delta}{\epsilon^\delta s_n^\delta} > 1$, de modo que teremos

$$\begin{aligned}
\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon s_n}(x-\mu_k)^2 dF_k(x) &\leq \frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon s_n}(x-\mu_k)^2\frac{|x-\mu_k|^\delta}{\epsilon^\delta s_n^\delta}dF_k(x)\\
&\leq \frac{1}{\epsilon^\delta s_n^{2+\delta}}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon s_n}|x-\mu_k|^{2+\delta}dF_k(x)\\
&\leq \frac{1}{\epsilon^\delta s_n^{2+\delta}}\sum_{k=1}^{n}\int_{-\infty}^{\infty}|x-\mu_k|^{2+\delta}dF_k(x)\\
&= \frac{1}{\epsilon^\delta s_n^{2+\delta}}\sum_{k=1}^{n}\mathbb{E}|X_k-\mu_k|^{2+\delta} \stackrel{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

onde a última passagem se deve à condição de Liapunov. Assim verifica-se a condição de Lindeberg, e logo, $\frac{S_n - \mathbb{E}S_n}{s_n} \xrightarrow{D} N(0,1)$. □

Para finalizar esta seção, observemos que da mesma forma que o Teorema Central do Limite de Liapunov, o Teorema 3.2.3 também pode ser demonstrado a partir da verificação da condição de Lindeberg. Basta lembrarmos que nesse caso temos $s_n^2 = n\sigma^2$ e $\mu_1 = \mu_2 = \cdots = \mu$, e daí, sendo F a distribuição comum de $X_1, X_2, \cdots$, temos

$$\begin{aligned}
\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|\leq\epsilon s_n}(x-\mu_k)^2 dF_k(x) &= \frac{1}{n\sigma^2}\sum_{k=1}^{n}\int_{|x-\mu|\leq\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF(x)\\
&= \frac{1}{\sigma^2}\int_{|x-\mu|\leq\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF(x)\\
&\stackrel{n\to\infty}{\longrightarrow} \frac{1}{\sigma^2}\int_{-\infty}^{\infty}(x-\mu)^2 dF(x) = 1,
\end{aligned}$$

donde decorre que

$$\begin{aligned}
\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon S_n}(x-\mu_k)^2 dF_k(x) &= \frac{1}{s_n^2}\sum_{k=1}^{n}\int_{-\infty}^{\infty}(x-\mu_k)^2 dF_k(x)\\
&\quad -\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|\leq\epsilon s_n}(x-\mu_k)^2 dF_k(x)\\
&= 1-\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|\leq\epsilon s_n}(x-\mu_k)^2 dF_k(x) \stackrel{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

como queríamos demonstrar.

3.3 Famílias Relativamente Compactas e Rígidas

Nesta seção vamos definir o que são famílias de medidas de probabilidade rígidas e relativamente compactas. Tal abordagem será fundamental na demonstração do Teorema 4.2.1, que é um dos principais resultados desta dissertação.

Definição 3.3.1 (Relativamente Compacto)**.** *Seja $\mathcal{P} = \{P_\alpha; \alpha \in \mathcal{L}\}$ uma família de medidas de probabilidade. Dizemos que $\mathcal{P}$ é relativamente compacta se toda sequência de $\mathcal{P}$ admite uma subsequência que converge fracamente a alguma medida de probabilidade.*

Definição 3.3.2 (Rígida)**.** *Seja $\mathcal{P} = \{P_\alpha; \alpha \in \mathcal{L}\}$ uma família de medidas de probabilidade. Dizemos que $\mathcal{P}$ é rígida se para todo $\epsilon > 0$ existe um conjunto compacto $K \subset \mathbb{R}$ tal que*

$$\sup_{\alpha \in \mathcal{L}} P_\alpha(\Omega - K) \leq \epsilon.$$

Definição 3.3.3. *Seja $\mathcal{F} = \{F_\alpha; \alpha \in \mathcal{L}\}$ uma família de funções de distribuição. Dizemos que $\mathcal{F}$ é relativamente compacta ou rígida se a família de medidas de probabilidade correspondente o for.*

Teorema 3.3.1 (Teorema de Helly)**.** *Seja $\mathcal{I} = \{G : \mathbb{R} \to \mathbb{R}\}$ a coleção de funções tais que*

a) G *é não decrescente;*

b) G *é contínua à direita;*

c) $0 \leq G(-\infty) \leq G(\infty) \leq 1$.

Então, a classe $\mathcal{I}$ é sequencialmente compacta, isto é, para toda sequência $\{G_n\}_{n\geq 1}$ de funções de $\mathcal{I}$, existe uma função $G \in \mathcal{I}$ e uma subsequência $\{G_{n_k}\}_{k\geq 1}$ tais que

$$G_{n_k}(x) \stackrel{k\to\infty}{\longrightarrow} G(x),$$

para todo x que é ponto de continuidade de G.

Demonstração. Seja $T = \{x_1, x_2, \cdots\}$ um subconjunto denso e enumerável de $\mathbb{R}$. Pelo método da Diagonal de Cantor (LIMA, 2009, Capítulo 10, Teorema 21), existe uma subsequência $\{G_{n_k}\}_{k\geq 1}$ tal que $\lim_k G_{n_k}(x_i) = g_i$, $i = 1, 2, \cdots$ e a partir daí vamos definir $G_T(x_i) = \lim_{k\to\infty} G_{n_k}(x_i) = g_i$ para $x_i \in T$. Agora vamos definir $G(x)$, $x \in \mathbb{R}$, pondo

$$G(x) = \inf\{G_T(y);\, y \in T \text{ e } y > x\}, \tag{3.11}$$

e vamos mostrar que $G(x)$ é a função tal que $G_{n_k}(x) \stackrel{k\to\infty}{\longrightarrow} G(x)$, para todo x que é ponto de continuidade de G e que satisfaz as hipóteses do teorema. Como toda função G_n em consideração é não decrescente temos $G_{n_k}(x) \leq G_{n_k}(y)$ para todo $x, y \in T$ tais que $x \leq y$ e para todo $k \geq 1$. Daí, para tais x e y temos $G_T(x) \leq G_T(y)$, e logo, por (3.11) temos que G é não decrescente. Para mostrar que G é contínua à direita, considere uma sequência $\{x_m\}_{m\geq 1}$ tal que $x_m \searrow x$ e $d = \lim_m G(x_m)$. Claramente temos $G(x) \leq d$ e precisamos mostrar que $G(x) = d$. Suponhamos então que $G(x) < d$. Por (3.11) temos que existe $y \in T$ tal que $x < y$ e $G_T(y) < d$. Mas para m suficientemente grande temos

$x < x_m < y$, e portanto, $G(x_m) \leq G(y) \leq d$. Por outro lado, como G é não decrescente, temos $G(x_m) \searrow d$. Em particular, $G(x_m) \geq d$, e logo, temos $G(x_m) = d$ para todo m suficientemente grande. Segue que $G(x) = \inf\{G_T(y);\ y \in T \text{ e } y > x\} = d$. Agora, como $0 \leq G_T(x_i) \leq 1$ para todo $x_i \in T$, teremos também $0 \leq G(x) \leq 1$ para todo $x \in \mathbb{R}$, e assim G pertencente a $\mathcal{I}$. Resta portanto provar que $G_{n_k}(x_0) \stackrel{k\to\infty}{\longrightarrow} G(x_0)$ para todo x_0 que é ponto de continuidade de G. Se $x_0 < y \in T$, então

$$\limsup_k G_{n_k}(x_0) \leq \limsup_k G_{n_k}(y) = G_T(y),$$

e consequentemente,

$$\limsup_k G_{n_k}(x_0) \leq \inf\{G_T(y);\ y \in T \text{ e } y > x_0\} = G(x_0). \tag{3.12}$$

Por outro lado, se $\tilde{x} < y < x_0$, $y \in T$, então

$$G(\tilde{x}) \leq G_T(y) = \lim_k G_{n_k}(y) = \liminf_k G_{n_k}(y) \leq \liminf_k G_{n_k}(x_0).$$

Tomando $\tilde{x} \nearrow x_0$, teremos

$$G(x_0-) \leq \liminf_k G_{n_k}(x_0). \tag{3.13}$$

Como x_0 é um ponto de continuidade de G, temos $G(x_0-) = G(x_0)$, e logo, por (3.12) e (3.13) temos que

$$G_{n_k}(x_0) \stackrel{k\to\infty}{\longrightarrow} G(x_0).$$

□

Teorema 3.3.2 (Teorema de Prohorov)**.** *Seja $\mathcal{P} = \{P_\alpha;\ \alpha \in \mathcal{L}\}$ uma família de medidas de probabilidade definidas no espaço $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Então $\mathcal{P}$ é Relativamente Compacta se, e somente se, é Rígida.*

Demonstração. Suponhamos inicialmente que a família $\mathcal{P}$ é relativamente compacta e não é rígida. Então existe um $\epsilon > 0$ tal que para todo compacto $K \subset \mathbb{R}$ temos $\sup_\alpha P_\alpha(\mathbb{R}-K) > \epsilon$, e portanto, para todo intervalo $I = (a, b)$ teremos $\sup_\alpha P_\alpha(\mathbb{R} - I) > \epsilon$. Segue que para todo intervalo $I_n = (-n, n)$, $n \geq 1$, existe uma medida de probabilidade P_{α_n} tal que $P_{\alpha_n}(\mathbb{R}-I_n) > \epsilon$. Como a família $\mathcal{P}$ é relativamente compacta, a sequência $\{P_{\alpha_n}\}_{n\geq 1}$ possui uma subsequência $\{P_{\alpha_{n_k}}\}_{k\geq 1}$ que converge fracamente para uma medida de probabilidade Q. Logo, para todo $n \geq 1$,

$$\limsup_{k\to\infty} P_{\alpha_{n_k}}(\mathbb{R} - I_n) \leq Q(\mathbb{R} - I_n).$$

Como $Q(\mathbb{R} - I_n) \searrow 0$ quando $n \to \infty$, segue que existe n_k tal que $P_{\alpha_{n_k}}(\mathbb{R} - I_n) < \epsilon$, contradizendo assim $\sup_\alpha P_\alpha(\mathbb{R} - I) > \epsilon$. Assim, toda família de medidas de probabilidade relativamente compacta será rígida.

Reciprocamente, suponhamos que a família $\mathcal{P}$ é rígida, e considere a sequência $\{P_n\}_{n\geq 1}$ em $\mathcal{P}$, onde $\{F_n\}_{n\geq 1}$ é a sequência de funções de distribuição correspondente. Pelo Teorema 3.3.1 existe uma subsequência F_{n_k} e uma função de distribuição generalizada G, com $0 \leq G(-\infty) \leq G(\infty) \leq 1$, de forma que $F_{n_k}(x) \overset{k\to\infty}{\longrightarrow} G(x)$, para todo x que é ponto de continuidade de G. Para provar que a família $\mathcal{P}$ é relativamente compacta basta mostrar que G é uma função de distribuição, e para tanto resta mostrar que $G(-\infty) = 0$ e $G(\infty) = 1$. Como a família $\mathcal{P}$ é rígida, dado $\epsilon > 0$, existe um intervalo $I = (a, b]$ para o qual $\sup_n P_n(\mathbb{R} - I) < \epsilon$, ou equivalentemente, $1 - \epsilon \leq P_n(a, b]$ para todo $n \geq 1$. Sejam então a' e b' pontos de continuidade de G tais que $a' < a$ e $b' > b$. Logo,

$$1 - \epsilon \leq P_{n_k}(a, b] \leq P_{n_k}(a', b'] = F_{n_k}(b') - F_{n_k}(a') \overset{k\to\infty}{\longrightarrow} G(b') - G(a').$$

Como $\epsilon > 0$ é arbitrário, segue que $G(\infty) - G(-\infty) = 1$, e como $0 \leq G(-\infty) \leq G(\infty) \leq 1$, temos $G(-\infty) = 0$ e $G(\infty) = 1$. $\square$

O Teorema de Prohorov na verdade é mais geral do que enunciamos. Na realidade a demonstração pode ser adaptada para o espaço euclidiano $\mathbb{R}^n$, $n \geq 2$ arbitrário, e daí podemos estender o teorema sucessivamente para $\mathbb{R}^\infty$, os espaços σ-compactos, e finalmente, para os espaços métricos separáveis e completos. Veja (SHIRYAEV, 1996, Capítulo 3).

4 Distribuições Infinitamente Divisíveis e Estáveis

Como já dissemos na introdução, o objetivo principal desta dissertação é estudar as classes das distribuições infinitamente divisíveis e estáveis, bem como os correspondentes teoremas limites, e isto será feito neste capítulo. Para ajudar a identificar tais distribuições faremos primeiro um estudo detalhado sobre as funções características. Tal abordagem passará pelas funções definidas positivas, pelo Teorema de Pólya, as distribuições de Poisson generalizadas e as distribuições α-estáveis. Depois mostraremos a relação entre as distribuições infinitamente divisíveis e os arranjos triangulares no Teorema 4.2.1, e a relação entre as distribuições estáveis e as sequências do tipo $\frac{S_n+b_n}{a_n}$ no Teorema 4.3.1, bem como as principais propriedades de tais distribuições. As demonstrações feitas nesse capítulo têm como referência principal (CHUNG, 2001) e (SHIRYAEV, 1996).

4.1 Caracterização das Funções Características

Para caracterizar os possíveis limites das sequências de variáveis aleatórias que formam um arranjo triangular como no Teorema de Poisson, bem como as sequências do tipo $\frac{S_n}{n}$ da Lei dos Grandes Números ou as somas normalizadas do Teorema Central do Limite, em especial as identicamente distribuídas, torna-se necessário conhecer as distribuições infinitamente divisíveis e estáveis, e para tanto torna-se imprescindível saber julgar quando uma função $\varphi : \mathbb{R} \to \mathbb{C}$ é uma função característica. Já vimos no capítulo anterior que uma função característica sempre satisfaz as propriedades da Proposição 2.4.1. Vimos também, pela mesma proposição, que o produto de funções características continua sendo uma função característica e pelo Corolário 2.4.3 o mesmo se dá com o quadrado do seu módulo. Mas esses critérios não serão suficientes para caracterizar as funções características. Vamos portanto estabelecer a partir de agora, critérios mais específicos em relação ao que vamos precisar.

Proposição 4.1.1. *Seja $\varphi_1, \varphi_2, \cdots$ uma sequência de funções características e $\lambda_1, \lambda_2, \cdots$ uma sequência de números reais tais que $\lambda_k \geq 0$ para todo k e $\sum\limits_{n=1}^{\infty} \lambda_n = 1$. Então $\varphi = \sum\limits_{n=1}^{\infty} \lambda_n \varphi_n$ também é uma função característica, isto é, uma combinação convexa de funções características é também uma função característica.*

Demonstração. Note que no caso de termos uma combinação convexa finita, o resultado segue da linearidade da Integral de Stieltjes, segundo a qual temos

$$\varphi(t) = \sum_{k=1}^{n} \lambda_k \varphi_k(t) = \sum_{k=1}^{n} \lambda_k \int_{-\infty}^{\infty} e^{itx} dF_k(x) = \int_{-\infty}^{\infty} e^{itx} d\left(\sum_{k=1}^{n} \lambda_k F_k(x)\right) = \int_{-\infty}^{\infty} e^{itx} dF(x),$$

onde F_j é a função de distribuição correspondente a φ_j, $j = 1, \cdots, n$, e $F = \sum_{k=1}^{n} \lambda_n F_n$. Daí, para mostrar que φ é uma função característica, basta mostrar que F é uma função de distribuição. Para tanto, basta observarmos que

- F é não decrescente e contínua à direita, já que é uma composição finita de funções com estas propriedades,
- $F(-\infty) = \lim_{x \to -\infty} F(x) = \sum_{k=1}^{n} \lambda_k \lim_{x \to -\infty} F_k(x) = 0$, e,
- $F(\infty) = \lim_{x \to \infty} F(x) = \sum_{k=1}^{n} \lambda_k \lim_{x \to \infty} F_k(x) = \sum_{k=1}^{n} \lambda_k = 1$.

Para o caso geral, vamos definir para todo n,

$$\tilde{\varphi}_n(t) = \sum_{k=1}^{n} \lambda_k \varphi_k(t) + \left(1 - \sum_{k=1}^{n} \lambda_k\right) \varphi_{n+1}(t).$$

Note que $\tilde{\varphi}_n$ é uma combinação convexa finita de $\varphi_1, \cdots, \varphi_{n+1}$, e portanto é uma função característica, como acabamos de provar. Além disso, $\tilde{\varphi}_n(t)$ converge pontualmente para $\varphi(t)$ e $\varphi(0) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(0) = \sum_{n=1}^{\infty} \lambda_n = 1$. Logo, pelo Teorema da Continuidade de Paul-Lévy, resta mostrar que φ é contínua em $t = 0$. Seja então $\epsilon > 0$ arbitrário. Como $\sum_{n=1}^{\infty} \lambda_n = 1$, existe $n_0 = n_0(\epsilon)$ tal que $\sum_{n=1}^{n_0} \lambda_n > 1 - \frac{\epsilon}{4}$, donde segue que

$$\begin{aligned}
\left| \sum_{n=n_0}^{\infty} \lambda_n \varphi_n(t) - \sum_{n=n_0}^{\infty} \lambda_n \varphi_n(0) \right| &\leq \sum_{n=n_0}^{\infty} \lambda_n |\varphi_n(t) - 1| \leq \sum_{n=n_0}^{\infty} \lambda_n (|\varphi_n(t)| + 1) \\
&\leq 2 \sum_{n=n_0}^{\infty} \lambda_n = 2\left(\sum_{n=1}^{\infty} \lambda_n - \sum_{n=1}^{n_0} \lambda_n\right) \\
&< 2\left(1 - \left(1 - \frac{\epsilon}{4}\right)\right) = \frac{\epsilon}{2}.
\end{aligned}$$

Por outro lado, como as φ_n são uniformemente contínuas, temos que para todo $n = 1, \cdots, n_0$ existe $\delta_n = \delta(n)$ tal que $|t| < \delta_n \Rightarrow |\varphi_n(t) - \varphi_n(0)| = |\varphi_n(t) - 1| < \frac{\epsilon}{2}$. Tomando $\delta = \delta(n_0) = \min_{n \leq n_0}\{\delta_n\}$, teremos que para $|t| < \delta$,

$$\left| \sum_{n=1}^{n_0} \lambda_n \varphi_n(t) - \sum_{n=1}^{n_0} \lambda_n \varphi_n(0) \right| \leq \sum_{n=1}^{n_0} \lambda_n |\varphi_n(t) - 1| \leq \frac{\epsilon}{2} \sum_{n=1}^{n_0} \lambda_n < \frac{\epsilon}{2}.$$

Daí segue que para $|t| < \delta$,

$$\begin{aligned}
|\varphi_n(t) - \varphi_n(0)| &\leq \left| \sum_{n=1}^{n_0} \lambda_n \varphi_n(t) - \sum_{n=1}^{n_0} \lambda_n \varphi_n(0) \right| + \left| \sum_{n=n_0}^{\infty} \lambda_n \varphi_n(t) - \sum_{n=n_0}^{\infty} \lambda_n \varphi_n(0) \right| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Assim φ é contínua em $t = 0$, e portanto, é uma função característica. □

À primeira vista, a demonstração da proposição para o caso da sequência finita nos leva a pensar que o caso da sequência infinita é uma aplicação direta do Teorema de Helly-Bray à sequência $\hat{F}_k = \sum_{k=1}^{n} \lambda_k F_k$. No entanto isso não pode ser feito, uma vez que as $\hat{F}_n$ não são funções de distribuição. Por outro lado, é possível obter o resultado dessa maneira se considerarmos a sequência $\tilde{F}_n = \sum_{k=1}^{n} \lambda_k F_k + \left(1 - \sum_{k=1}^{n} \lambda_k\right) F_{k+1}$ (como fizemos com as funções características), que é formada por funções de distribuição, e daí, por Helly-Bray teríamos

$$\int_{-\infty}^{\infty} e^{itx} d\tilde{F}_n(x) \stackrel{n\to\infty}{\longrightarrow} \int_{-\infty}^{\infty} e^{itx} dF(x) = \varphi(t).$$

Mas, seria preciso ainda provar que F é uma função de distribuição, o que seria muito mais trabalhoso do que aplicar o Teorema de Paul-Lévy às funções características.

Definição 4.1.1 (Função Definida Positiva)**.** *Dizemos que uma função $\varphi : \mathbb{R} \to \mathbb{C}$ é definida positiva se para quaisquer conjuntos $\{t_1, \cdots, t_n\} \subset \mathbb{R}$ e $\{z_1, \cdots, z_n\} \subset \mathbb{C}$ tivermos*

$$\sum_{j=1}^{n} \sum_{k=1}^{n} \varphi(t_j - t_k) z_j \overline{z_k} \geq 0.$$

Teorema 4.1.1. *Seja $\varphi : \mathbb{R} \to \mathbb{C}$ uma função definida positiva. Então, para todo $t \in \mathbb{R}$ temos $\varphi(-t) = \overline{\varphi(t)}$ e $|\varphi(t)| \leq \varphi(0)$. Além disso, se φ for contínua em $t = 0$, então será uniformemente contínua em $\mathbb{R}$, e nesse caso, teremos para toda função contínua $\xi : \mathbb{R} \to \mathbb{C}$ e para todo $T > 0$,*

$$\int_0^T \int_0^T \varphi(s - t)\xi(s)\overline{\xi(t)} ds dt \geq 0.$$

Demonstração. Considerando $n = 1$ e tomando $t_1 = 0$ e $z_1 = 1$, pela Definição 4.1.1 temos

$$0 \leq \sum_{j=1}^{1} \sum_{k=1}^{1} \varphi(t_j - t_k) z_j \overline{z_k} = \varphi(t_1 - t_1) z_1 \overline{z_1} = \varphi(0).$$

Considerando agora $n = 2$, novamente da Definição 4.1.1 temos que

$$\begin{aligned} 0 &\leq \sum_{j=1}^{2} \sum_{k=1}^{2} \varphi(t_j - t_k) z_j \overline{z_k} \\ &= \varphi(t_1 - t_1) z_1 \overline{z_1} + \varphi(t_1 - t_2) z_1 \overline{z_2} + \varphi(t_2 - t_1) z_2 \overline{z_1} + \varphi(t_2 - t_2) z_2 \overline{z_2} \\ &= (z_1 \overline{z_1} + z_2 \overline{z_2}) \varphi(0) + z_1 \overline{z_2} \varphi(t_1 - t_2) + z_2 \overline{z_1} \varphi(t_2 - t_1). \end{aligned}$$

Daí, tomando $t_1 = 0$, $t_2 = t$ e $z_1 = z_2 = 1$, teremos

$$0 \leq 2\varphi(0) + \varphi(-t) + \varphi(t),$$

donde segue que $\varphi(-t) + \varphi(t)$ é um número real, e logo, as partes imaginárias de $\varphi(t)$ e $\varphi(-t)$ são opostas. Por outro lado, tomando $t_1 = 0$, $t_2 = t$, $z_1 = 1$ e $z_2 = i$, teremos

$$0 \leq 2\varphi(0) + i\varphi(t) - i\varphi(-t) = 2\varphi(0) + i(\varphi(t) - \varphi(-t)),$$

donde segue que $\varphi(t) - \varphi(-t)$ é um número imaginário puro, e logo, as partes reais de $\varphi(t)$ e $\varphi(-t)$ são iguais, e portanto temos $\varphi(-t) = \overline{\varphi(t)}$. Ainda no caso $n = 2$, tomando $z_1 = \varphi(t)$, $z_2 = -|\varphi(t)|$, $t_1 = 0$ e $t_2 = t$, teremos

$$\begin{aligned} 0 &\leq (\varphi(t)\overline{\varphi(t)} + |\varphi(t)|^2)\varphi(0) - \varphi(t)|\varphi(t)|^2\varphi(-t) - \overline{\varphi(t)}|\varphi(t)|\varphi(t) \\ &= 2|\varphi(t)|^2\varphi(0) - 2|\varphi(t)|^3 \\ &= 2|\varphi(t)|^2(\varphi(0) - |\varphi(t)|), \end{aligned}$$

e daí segue que $|\varphi(t)| \leq \varphi(0)$.

Agora, supondo φ contínua em $t = 0$, vamos mostrar que φ é uniformemente contínua em $\mathbb{R}$. Se $\varphi(0) = 0$, teremos $\varphi \equiv 0$, e não há mais o que fazer. Se $\varphi(0) \neq 0$, vamos supor, sem perda de generalidade, que $\varphi(0) = 1$ (basta observar que $\psi = \frac{\varphi}{\varphi(0)}$ também é definida positiva). Considerando então $n = 3$, e tomando $t_1 = 0$, $t_2 = t$ e $t_3 = t + h$, teremos da Definição 4.1.1 que

$$\begin{aligned} 0 &\leq \sum_{j=1}^{3}\sum_{k=1}^{3} \varphi(t_j - t_k) z_j \overline{z_k} \\ &= \varphi(0)(|z_1|^2 + |z_2|^2 + |z_3|^2) + \varphi(-t)z_1\overline{z_2} + \varphi(-t-h)z_1\overline{z_3} \\ &\quad + \varphi(t)\overline{z_1}z_2 + \varphi(-h)z_2\overline{z_3} + \varphi(t+h)\overline{z_1}z_3 + \varphi(h)\overline{z_2}z_3. \end{aligned}$$

Daí, sendo $H = \begin{bmatrix} \varphi(0) & \varphi(-t) & \varphi(-t-h) \\ \varphi(t) & \varphi(0) & \varphi(-h) \\ \varphi(t+h) & \varphi(h) & \varphi(0) \end{bmatrix}$ e $(z_1, z_2, z_3) = v \in \mathbb{R}^3$, teremos

$$\begin{aligned} \langle Hv, v \rangle &= \varphi(0)(z_1^2 + z_2^2 + z_3^2) + \varphi(-t)z_1z_2 + \varphi(-t-h)z_1z_3 + \varphi(t)z_1z_2 \\ &\quad + \varphi(-h)z_2z_3 + \varphi(t+h)z_1z_3 + \varphi(h)z_2z_3 \geq 0. \end{aligned}$$

Segue que $\langle Hv, v \rangle$ é uma forma quadrática não negativa, e logo, $\det H \geq 0$, isto é,

$$1 - |\varphi(t)|^2 - |\varphi(t+h)|^2 - |\varphi(h)|^2 + 2\Re(\varphi(t)\varphi(h)\overline{\varphi(t+h)}) \geq 0,$$

ou ainda,

$$|\varphi(t)|^2 + |\varphi(t+h)|^2 \leq 1 - |\varphi(h)|^2 + 2\Re(\varphi(t)\varphi(h)\overline{\varphi(t+h)}),$$

onde $\Re(z)$ denota a parte real do complexo z. Note que sendo $\varphi(t) = a + bi$ e $\varphi(t+h) = c + di$, temos $\varphi(t)\overline{\varphi(t+h)} = ac + bd + i(-ad + bc)$, e daí, usando a desigualdade acima, temos

$$\begin{aligned}
|\varphi(t)-\varphi(t+h)|^2 &= (a-c)^2+(b-d)^2 = a^2+b^2+c^2+d^2-2(ac+bd)\\
&= |\varphi(t)|^2+|\varphi(t+h)|^2-2\Re(\varphi(t)\overline{\varphi(t+h)})\\
&\leq 1-|\varphi(h)|^2+2\Re(\varphi(t)\varphi(h)\overline{\varphi(t+h)}-2\Re(\varphi(t)\overline{\varphi(t+h)})\\
&= 1-|\varphi(h)|^2+2\Re[\varphi(t)\overline{\varphi(t+h)}(\varphi(h)-1)]\\
&\leq 1-|\varphi(h)|^2+2|\varphi(t)\overline{\varphi(t+h)}(\varphi(h)-1)|\\
&\leq 1-|\varphi(h)|^2+2|1-\varphi(h)|\\
&= (1+|\varphi(h)|)(1-|\varphi(h)|)+2|1-\varphi(h)|\\
&\leq 2(1-|\varphi(h)|)+2|1-\varphi(h)|\\
&\leq 2|1-\varphi(h)|+2|1-\varphi(h)|\\
&= 4|1-\varphi(h)|.
\end{aligned}$$

Assim, como φ é contínua em $t=0$, temos que dado $\epsilon>0$, existe $\delta>0$ tal que $|h|<\delta$ implica

$$|1-\varphi(h)| = |\varphi(0)-\varphi(h)| < \frac{\epsilon^2}{4},$$

e assim,

$$|\varphi(t)-\varphi(t+h)| < 2\sqrt{|1-\varphi(h)|} < \epsilon.$$

Como o resultado não depende de t, segue que φ é uniformemente contínua em $\mathbb{R}$. Por fim, para mostrar que $\int_0^T\int_0^T \varphi(s-t)\xi(s)\overline{\xi(t)}dsdt \geq 0$, considere as partições de $[0,T]$ dadas por $P_n=\{v_0=0, v_1=\frac{T}{n}, v_2=\frac{2T}{n}, \cdots, v_n=T\}$. Logo,

$$\begin{aligned}
\int_0^T \varphi(s-t)\xi(s)\overline{\xi(t)}ds &= \lim_{n\to\infty}\sum_{k=1}^{n}(v_k-v_{k-1})\varphi(s_k-t)\xi(s_k)\overline{\xi(t)}\\
&= \lim_{n\to\infty}\sum_{k=1}^{n}\frac{T}{n}\varphi(s_k-t)\xi(s_k)\overline{\xi(t)},
\end{aligned}$$

onde $v_{k-1}\leq s_k\leq v_k$. Tomando as mesmas partições para a outra integral, segue da Definição 4.1.1 que

$$\begin{aligned}
\int_0^T\int_0^T \varphi(s-t)\xi(s)\overline{\xi(t)}dsdt &= \lim_{n\to\infty}\sum_{j=1}^{n}\left[(v_j-v_{j-1})\lim_{n\to\infty}\sum_{k=1}^{n}\frac{T}{n}\varphi(s_k-t_j)\xi(s_k)\overline{\xi(t)}\right]\\
&= \lim_{n\to\infty}\sum_{j=1}^{n}\sum_{k=1}^{n}\left(\frac{T}{n}\right)^2\varphi(s_k-t_j)\xi(s_k)\overline{\xi(t_j)} \geq 0.
\end{aligned}$$

□

Lema 4.1.1. *Dada uma função contínua $\varphi:\mathbb{R}\to\mathbb{C}$ com $|\varphi(t)|\leq\varphi(0)=1$, defina*

$$\rho_T(x) = \frac{1}{2\pi T}\int_0^T\int_0^T \varphi(s-t)e^{-i(s-t)x}dsdt.$$

Se $\rho_T(x)\geq 0$ para todo $T>0$ e $x\in\mathbb{R}$, então φ é uma função característica.

Demonstração. Para mostrar que φ é uma função característica vamos mostrar primeiro que ρ_T é uma função de densidade. Depois mostraremos que a função característica correspondente a ρ_T é

$$\varphi_T(t) = \begin{cases} \left(1 - \frac{|t|}{T}\right)\varphi(t) & \text{se } |t| \leq T \\ 0 & \text{se } |t| > T \end{cases},$$

e como φ_T converge para φ quando T tende ao infinito, bastará então aplicar o Teorema da continuidade de Paul-Lévy. Para começar, sendo $u = s - t$, temos $du = ds$, $u = -t$ para $s = 0$ e $u = T - t$ para $s = T$. Logo temos

$$\rho_T(x) = \frac{1}{2\pi T}\int_0^T \int_{-t}^{T-t} \varphi(u)e^{-iux}dudt = \frac{1}{2\pi T}\int_0^T g(t)h'(t)dt,$$

onde $g(t) = \int_{-t}^{T-t}\varphi(u)e^{-iux}du$ e $h(t) = t$. Integrando por partes obtemos

$$\begin{aligned} \rho_T(x) &= \frac{1}{2\pi T}\left[t\int_{-t}^{T-t}\varphi(u)e^{-iux}du\Big|_{t=0}^{t=T} - \int_0^T tg'(t)dt\right] \\ &= \frac{1}{2\pi T}\left[T\int_{-T}^{0}\varphi(u)e^{-iux}du - \int_0^T tg'(t)dt\right]. \end{aligned}$$

Pelo Teorema Fundamental do Cálculo temos que $g(t) = H(T-t) - H(-t)$, onde $H'(s) = \varphi(s)e^{-isx}ds$. Daí temos

$$g'(t) = -H'(T-t) + H'(-t) = -\varphi(T-t)e^{-i(T-t)x} + \varphi(-t)e^{itx},$$

e logo,

$$\begin{aligned} \rho_T(x) &= \frac{1}{2\pi}\left[\int_{-T}^{0}\varphi(t)e^{-itx}dt - \int_0^T \frac{t}{T}\left(\varphi(-t)e^{itx} - \varphi(T-t)e^{-i(T-t)x}\right)dt\right] \\ &= \frac{1}{2\pi}\left[\int_{-T}^{0}\varphi(t)e^{-itx}dt - \int_0^{-T} \frac{-t}{T}\varphi(t)e^{-itx}dt - \int_T^0 \frac{T-t}{T}\varphi(t)e^{-itx}dt\right] \\ &= \frac{1}{2\pi}\left[\int_{-T}^{0}\varphi(t)e^{-itx}dt + \int_{-T}^{0} \frac{t}{T}\varphi(t)e^{-itx}dt + \int_0^T \frac{T-t}{T}\varphi(t)e^{-itx}dt\right] \\ &= \frac{1}{2\pi}\left[\int_{-T}^{0}\left(1 - \frac{|t|}{T}\right)\varphi(t)e^{-itx}dt + \int_0^T\left(1 - \frac{|t|}{T}\right)\varphi(t)e^{-itx}dt\right] \\ &= \frac{1}{2\pi}\int_{-T}^{T}\left(1 - \frac{|t|}{T}\right)\varphi(t)e^{-itx}dt. \end{aligned}$$

Agora, dados $\alpha > 0$ e $t \neq 0$ arbitrários, temos que

$$\begin{aligned} \frac{1}{\alpha}\int_0^\alpha d\beta\int_{-\beta}^{\beta}e^{-itx}dx &= \frac{1}{\alpha}\int_0^\alpha \frac{e^{-itx}}{-it}\Big|_{x=-\beta}^{x=\beta}d\beta = \frac{1}{\alpha}\int_0^\alpha \frac{e^{-it\beta} - e^{it\beta}}{-it}d\beta \\ &= \frac{1}{\alpha}\int_0^\alpha \frac{\cos(-t\beta) + i\sin(-t\beta) - \cos(t\beta) - i\sin(t\beta)}{-it}d\beta \\ &= \frac{1}{\alpha}\int_0^\alpha \frac{2\sin(t\beta)}{t}d\beta = \frac{-2\cos(t\beta)}{\alpha t^2}\Big|_{\beta=0}^{\beta=\alpha} = \frac{2(1-\cos(t\alpha))}{\alpha t^2}. \end{aligned}$$

Para $t = 0$,

$$\frac{1}{\alpha}\int_0^\alpha d\beta \int_{-\beta}^{\beta} e^{-itx}dx = \frac{1}{\alpha}\int_0^\alpha \int_{-\beta}^{\beta} dxd\beta = \frac{1}{\alpha}\int_0^\alpha 2\beta d\beta = \frac{1}{\alpha}\beta^2\Big|_{\beta=0}^{\beta=\alpha} = \alpha = \lim_{t\to 0}\frac{2(1-\cos(t\alpha))}{\alpha t^2}.$$

Segue daí que para todo $\alpha > 0$ temos

$$\begin{aligned}
\frac{1}{\alpha}\int_0^\alpha d\beta \int_{-\beta}^{\beta} \rho_T(x)dx &= \frac{1}{2\pi\alpha}\int_0^\alpha \int_{-\beta}^{\beta}\int_{-T}^{T}\left(1-\frac{|t|}{T}\right)\varphi(t)e^{-itx}dtdxd\beta \\
&= \frac{1}{2\pi\alpha}\int_{-T}^{T}\int_0^\alpha \int_{-\beta}^{\beta}\left(1-\frac{|t|}{T}\right)\varphi(t)e^{-itx}dxd\beta dt \\
&= \frac{1}{2\pi}\int_{-T}^{T}\left(1-\frac{|t|}{T}\right)\varphi(t)\frac{1}{\alpha}\int_0^\alpha \int_{-\beta}^{\beta} e^{-itx}dxd\beta dt \\
&= \frac{1}{2\pi}\int_{-T}^{T}\left(1-\frac{|t|}{T}\right)\varphi(t)\frac{2[1-\cos(t\alpha)]}{\alpha t^2}dt \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty}\varphi_T(t)\frac{1-\cos(t\alpha)}{\alpha t^2}dt \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty}\varphi_T\left(\frac{t}{\alpha}\right)\frac{1-\cos(t)}{t^2}dt,
\end{aligned}$$

onde a troca na ordem de integração se justifica na continuidade do integrando e no fato dos intervalos de integração serem limitados. Prosseguindo, como $|\varphi_T(t)| \le |\varphi(t)| \le 1$ e $\frac{1-\cos(t)}{t^2}$ é limitado em $\mathbb{R}$ podemos majorar o integrando por $\frac{|1-\cos(t)|}{t^2}$ que é integrável em $\mathbb{R}$. Daí, usando o Teorema da Convergência Dominada temos

$$\lim_{\alpha\to\infty}\frac{1}{\alpha}\int_0^\alpha d\beta \int_{-\beta}^{\beta} \rho_T(x)dx = \frac{1}{\pi}\int_{-\infty}^{\infty}\lim_{\alpha\to\infty}\varphi_T\left(\frac{t}{\alpha}\right)\frac{1-\cos(t)}{t^2}dt, \tag{4.1}$$

e como $\varphi_T(0) = \varphi(0) = 1$, segue que

$$\begin{aligned}
\lim_{\alpha\to\infty}\frac{1}{\alpha}\int_0^\alpha d\beta \int_{-\beta}^{\beta} \rho_T(x)dx &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1-\cos(t)}{t^2}dt = \frac{2}{\pi}\int_0^{\infty}\frac{1-\cos(t)}{t^2}dt \\
&= \frac{-2(1-\cos(t))}{\pi t}\Big|_{t=0}^{t=\infty} + \frac{2}{\pi}\int_0^\infty \frac{\sin(t)}{t}dt = 1,
\end{aligned}$$

onde as três últimas igualdades decorrem do fato do integrando ser uma função par, da integração parcial e da Integral de Dirichlet (Lema 2.4.1), respectivamente. Agora, como $\rho_T(x) \ge 0$, temos que $\int_{-\beta}^{\beta}\rho_T(x)dx$ não decresce em β. Além disso, como

$$\lim_{\alpha\to\infty}\frac{1}{\alpha}\int_0^\alpha d\beta \int_{-\beta}^{\beta} \rho_T(x)dx = 1,$$

temos que $\int_{-\beta}^{\beta}\rho_T(x)dx$ é limitado para todo $\beta > 0$. Segue portanto que existe $M < \infty$ tal que

$$\lim_{\beta\to\infty}\int_{-\beta}^{\beta}\rho_T(x)dx = M.$$

Como $\rho_T(x) \ge 0$, temos $\int_{-\infty}^{\infty}\rho_T(x)dx = M$. De fato, para quaisquer $A_n, B_n \nearrow \infty$, denote $C_n = \max\{A_n, B_n\}$ e $c_n = \min\{A_n, B_n\}$, e teremos

$$\int_{-c_n}^{c_n}\rho_T(x)dx \le \int_{-A_n}^{B_n}\rho_T(x)dx \le \int_{-C_n}^{C_n}\rho_T(x)dx,$$

e logo tem-se

$$\lim_n \int_{-A_n}^{B_n} \rho_T(x)dx = M.$$

Assim, como $\lim_{\alpha\to\infty} \frac{1}{\alpha}\int_0^\alpha d\beta \int_{-\beta}^{\beta} \rho_T(x)dx = 1$ para todo $\beta > 0$, temos

$$\begin{aligned} 1 &= \lim_{\alpha\to\infty}\lim_{\beta\to\infty} \frac{1}{\alpha}\int_0^\alpha d\beta \int_{-\beta}^{\beta} \rho_T(x)dx = \lim_{\alpha\to\infty} \frac{1}{\alpha}\int_0^\alpha d\beta \int_{-\infty}^{\infty} \rho_T(x)dx \\ &= \int_{-\infty}^{\infty} \rho_T(x)dx \lim_{\alpha\to\infty} \frac{1}{\alpha}\int_0^\alpha d\beta = \int_{-\infty}^{\infty} \rho_T(x)dx. \end{aligned}$$

Daí segue que para todo $T > 0$, ρ_T é uma função de densidade de probabilidade. Para calcular a função característica correspondente a ρ_T, observemos primeiramente que

$$\begin{aligned} \int_{-\beta}^{\beta} e^{i\tau x}e^{-itx}dx &= \int_{-\beta}^{\beta} e^{i(\tau-t)x}dx = \left.\frac{e^{i(\tau-t)x}}{i(\tau-t)}\right|_{x=-\beta}^{\beta} \\ &= \frac{\cos(\beta(\tau-t)) + i\sin(\beta(\tau-t)) - \cos(-\beta(\tau-t)) - i\sin(-\beta(\tau-t))}{i(\tau-t)} \\ &= \frac{2\sin(\beta(\tau-t))}{\tau-t}. \end{aligned}$$

Assim, para todo $\alpha > 0$,

$$\begin{aligned} \frac{1}{\alpha}\int_0^\alpha d\beta \int_{-\beta}^{\beta} e^{i\tau x}\rho_T(x)dx &= \frac{1}{\alpha}\int_0^\alpha \int_{-\beta}^{\beta} e^{i\tau x}\frac{1}{2\pi}\int_{-T}^{T}\left(1-\frac{|t|}{T}\right)\varphi(t)e^{-itx}dtdxd\beta \\ &= \frac{1}{2\pi}\int_{-T}^{T}\left(1-\frac{|t|}{T}\right)\varphi(t)\frac{1}{\alpha}\int_0^\alpha\int_{-\beta}^{\beta} e^{i\tau x}e^{-itx}dxd\beta dt \\ &= \frac{1}{2\pi}\int_{-T}^{T}\left(1-\frac{|t|}{T}\right)\varphi(t)\frac{1}{\alpha}\int_0^\alpha \frac{2\sin(\beta(\tau-t))}{\tau-t}d\beta dt \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi_T(t)\left.\frac{-2\cos(\beta(\tau-t))}{\alpha(\tau-t)^2}\right|_{\beta=0}^{\beta=\alpha}dt \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty}\varphi_T(t)\frac{1-\cos(\alpha(\tau-t))}{\alpha(\tau-t)^2}dt \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty}\varphi_T\left(\tau-\frac{t}{\alpha}\right)\frac{1-\cos(t)}{t^2}dt. \end{aligned}$$

Pelo mesmo argumento que usamos para $\rho_T(x)$ em (4.1), temos de Teorema da Convergência Dominada que

$$\begin{aligned} \lim_{\alpha\to\infty}\frac{1}{\alpha}\int_0^\alpha d\beta\int_{-\beta}^{\beta} e^{i\tau x}\rho_T(x)dx &= \lim_{\alpha\to\infty}\frac{1}{\pi}\int_{-\infty}^{\infty}\varphi_T\left(\tau-\frac{t}{\alpha}\right)\frac{1-\cos(t)}{t^2}dt \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty}\lim_{\alpha\to\infty}\varphi_T\left(\tau-\frac{t}{\alpha}\right)\frac{1-\cos(t)}{t^2}dt \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty}\varphi_T(\tau)\frac{1-\cos(t)}{t^2}dt \\ &= \frac{\varphi_T(\tau)}{\pi}\int_{-\infty}^{\infty}\frac{1-\cos(t)}{t^2}dt \\ &= \varphi_T(\tau). \end{aligned}$$

Uma vez que $\rho_T(x)$ é uma densidade de probabilidade e $|e^{i\tau x}\rho_T(x)| \leq \rho_T(x)$, podemos usar o Teorema da Convergência Dominada para obtermos que

$$\begin{aligned}
\varphi_T(\tau) &= \lim_{\alpha\to\infty}\lim_{\beta\to\infty}\frac{1}{\alpha}\int_0^\alpha d\beta\int_{-\beta}^{\beta} e^{i\tau x}\rho_T(x)dx \\
&= \lim_{\alpha\to\infty}\frac{1}{\alpha}\int_0^\alpha d\beta\int_{-\infty}^{\infty} e^{i\tau x}\rho_T(x)dx \\
&= \int_{-\infty}^{\infty} e^{i\tau x}\rho_T(x)dx \lim_{\alpha\to\infty}\frac{1}{\alpha}\int_0^\alpha d\beta \\
&= \int_{-\infty}^{\infty} e^{i\tau x}\rho_T(x)dx.
\end{aligned}$$

Assim, φ_T é a função característica correspondente à densidade $\rho_T(x)$, qualquer que seja $T \in \mathbb{R}$. Além disso, $\varphi_T(t) \stackrel{T\to\infty}{\longrightarrow} \varphi(t)$ para todo $t \in \mathbb{R}$ e φ é contínua em $t = 0$. Logo, do Teorema da Continuidade de Paul-Lévy segue que φ é uma função característica. □

Teorema 4.1.2. *Seja $\varphi : \mathbb{R} \to \mathbb{C}$ uma função contínua tal que $\varphi(0) = 1$. Então φ é uma função característica se, e somente se, é definida positiva.*

Demonstração. Se φ é uma função característica, então temos

$$\begin{aligned}
\sum_{j=1}^{n}\sum_{k=1}^{n}\varphi(t_j-t_k)z_j\overline{z_k} &= \sum_{j=1}^{n}\sum_{k=1}^{n}\int_{-\infty}^{\infty} e^{i(t_j-t_k)x}z_j\overline{z_k}dF(x) \\
&= \int_{-\infty}^{\infty}\sum_{j=1}^{n}\sum_{k=1}^{n} e^{i(t_j-t_k)x}z_j\overline{z_k}dF(x) \\
&= \int_{-\infty}^{\infty}\sum_{j=1}^{n}\sum_{k=1}^{n} e^{it_jx}z_j\overline{e^{it_kx}z_k}dF(x) \\
&= \int_{-\infty}^{\infty}\left|\sum_{j=1}^{n} e^{it_jx}z_j\right|^2 dF(x) \geq 0,
\end{aligned}$$

quaisquer que sejam $t_1, \cdots, t_n \in \mathbb{R}$ e $z_1, \cdots, z_n \in \mathbb{C}$. Daí, φ é definida positiva.

Reciprocamente, sendo φ definida positiva, defina $\rho_T(x)$ como no Lema 4.1.1. Observe que tomando $\xi(t) = e^{-itx}$ no Teorema 4.1.1, temos para todo $T > 0$ que

$$\rho_T(x) = \frac{1}{2\pi T}\int_0^T\int_0^T \varphi(s-t)e^{-i(s-t)x}dsdt = \frac{1}{2\pi T}\int_0^T\int_0^T \varphi(s-t)\xi(s)\overline{\xi(t)}dsdt \geq 0.$$

Além disso, ainda do Teorema 4.1.1 temos $|\varphi(t)| \leq \varphi(0) = 1$, e daí, segue do Lema 4.1.1 que φ é uma função característica. □

Observe que a partir do Lema 4.1.1 a demonstração do teorema acima se torna razoavelmente simples, e esse é o único motivo pelo qual está sendo enunciado aqui. Na realidade ele não desempenha nenhum papel em relação aos objetivos da dissertação, e portanto não justifica a energia gasta na demonstração do Lema. No entanto, o próximo teorema será de grande utilidade e também será demonstrado a partir dele.

Teorema 4.1.3 (Teorema de Pólya)**.** *Seja $\varphi : \mathbb{R} \to \mathbb{R}$ uma função não crescente, contínua e convexa em $\mathbb{R}^+ = [0, \infty)$. Se $\varphi(0) = 1$, $\varphi(t) = \varphi(-t)$ e $\varphi(t) \geq 0$ para todo $t \in \mathbb{R}$, então φ é uma função característica.*

Demonstração. Como φ é convexa em $\mathbb{R}^+$, temos que existe $\varphi'_+(t)$ e $\varphi'_-(t)$ com $\varphi'_+(t) = \varphi'_-(t)$ em quase toda parte, isto é, exceto num conjunto enumerável, de modo que φ é a integral de uma dessas derivadas, que vamos denotar por φ'. Note que como φ é não crescente em $\mathbb{R}^+$, temos $\varphi'(t) \leq 0$ para $t \geq 0$. Além disso, como φ é convexa, temos $\varphi''(t) \geq 0$, e consequentemente, φ' crescente em $\mathbb{R}^+$. Vamos definir então $\rho_T(x)$ e $\varphi_T(x)$ como no Lema 4.1.1. Observe que

$$-\varphi'_T(t) = \begin{cases} \frac{1}{T}\varphi(t) - \left(1 - \frac{t}{T}\right)\varphi'(t), & \text{se } 0 \leq t < T \\ 0 & \text{se } t \geq T, \end{cases}$$

e logo, como $\varphi(t) \geq 0$, $1 - \frac{t}{T} > 0$ e $\varphi'(t) \leq 0$ para $0 \leq t \leq T$, temos $-\varphi'_T(t) \geq 0$ em $\mathbb{R}^+$. Do mesmo modo, como φ, $1 - \frac{t}{T}$ e $-\varphi'$ são não crescentes para $0 \leq t \leq T$, temos $-\varphi'_T$ não crescente em $\mathbb{R}^+$.

Agora observe que para todo $x \neq 0$,

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-itx}\varphi_T(t)dt &= \int_{-\infty}^{0} e^{-itx}\varphi_T(t)dt + \int_{0}^{\infty} e^{-itx}\varphi_T(t)dt \\ &= \int_{0}^{\infty} e^{itx}\varphi_T(t)dt + \int_{0}^{\infty} e^{-itx}\varphi_T(t)dt \\ &= 2\int_{0}^{\infty} \cos(tx)\varphi_T(t)dt \\ &= 2\left[\sin(tx)x\varphi_T(t)\Big|_{t=0}^{t=\infty} - \int_{0}^{\infty} \frac{\sin(tx)}{x}\varphi'_T(t)dt\right]. \end{aligned}$$

Supondo $\varphi(\infty) = \lim_{t\to\infty} \varphi(t) = 0$, segue que para todo $x \neq 0$,

$$\int_{-\infty}^{\infty} e^{-itx}\varphi_T(t)dt = \frac{2}{x}\int_{0}^{\infty} \sin(tx)\varphi'_T(t)dt = \frac{2}{x}\sum_{k=0}^{\infty}\int_{\frac{k\pi}{x}}^{\frac{(k+1)\pi}{x}} \sin(tx)\varphi'_T(t)dt.$$

Como $-\varphi'_T(t) \geq 0$ em $\mathbb{R}^+$ e $\sin(tx)$ troca de sinal em $t = \frac{\lambda\pi}{x}$ para todo inteiro positivo λ, temos que a série acima é alternada e tem o primeiro termo positivo. Além disso, como $-\varphi'$ é não crescente e positiva em $\mathbb{R}^+$, temos que os termos da série decrescem em módulo convergindo para zero, e daí temos

$$\int_{-\infty}^{\infty} e^{-itx}\varphi_T(t)dt \geq 0.$$

Para $x = 0$ é trivial que

$$\int_{-\infty}^{\infty} e^{-it0}\varphi_T(t)dt = \int_{-\infty}^{\infty} \varphi_T(t)dt \geq 0.$$

Assim, temos $\rho_T(x) \geq 0$ para todo $x \in \mathbb{R}$, e a partir daí, como φ é contínua com $|\varphi(t)| \leq \varphi(0) = 1$, temos do Lema 4.1.1 que φ é uma função característica.

Se $\varphi(\infty) \neq 0$, então $\varphi(\infty) = 1$ ou $\varphi(\infty) = \lambda \in (0,1)$. No caso em que $\varphi(\infty) = 1$, temos $\varphi \equiv 1$, e logo φ é a função característica da variável aleatória $X \equiv 0$.

No caso em que $\varphi(\infty) = \lambda \in (0,1)$, vamos considerar a função

$$\psi(t) = \frac{\varphi(t) - \lambda}{1 - \lambda}.$$

Note que ψ satisfaz todas as hipóteses que utilizamos na demonstração do teorema para o caso em que $\varphi(\infty) = 0$, isto é:

- $\psi(0) = \frac{\varphi(0)-\lambda}{1-\lambda} = 1$,
- $\psi(t) = \frac{\varphi(t)-\lambda}{1-\lambda} \geq 0$,
- $\psi(-t) = \frac{\varphi(-t)-\lambda}{1-\lambda} = \frac{\varphi(t)-\lambda}{1-\lambda} = \psi(t)$,
- ψ é contínua e convexa em $\mathbb{R}^+$, pois é a composição das funções φ e $\tilde{\lambda}(x) = \frac{x-\lambda}{1-\lambda}$, ambas com tal propriedade, e
- $\psi(\infty) = \frac{\varphi(\infty)-\lambda}{1-\lambda} = 0$.

Assim temos que ψ é uma função característica, e para finalizar, como $\varphi(t) = (1-\lambda)\psi(t) + \lambda$, temos que φ é uma combinação convexa das funções características ψ e 1, e logo, pela Proposição 4.1.1 φ é uma função característica. □

Teorema 4.1.4. *Se $\varphi : \mathbb{R} \to \mathbb{C}$ é uma função característica, então para todo $\lambda \geq 0$,*

$$\varphi_\lambda(t) = e^{\lambda(\varphi(t)-1)}$$

será também uma função característica.

Demonstração. Fixado $\lambda \geq 0$, para $n > \lambda$ temos que

$$1 - \frac{\lambda}{n} + \frac{\lambda}{n}\varphi(t) = 1 + \frac{\lambda}{n}(\varphi(t) - 1)$$

é uma combinação convexa das funções características φ e 1, e portanto, pela Proposição 4.1.1 será uma função característica, e pela Proposição 2.4.1, o mesmo ocorre com

$$\left(1 + \frac{\lambda}{n}(\varphi(t) - 1)\right)^n,$$

qualquer que seja o inteiro $n > \lambda$. Daí, como

$$\left(1 + \frac{\lambda}{n}(\varphi(t) - 1)\right)^n \overset{n\to\infty}{\longrightarrow} e^{\lambda(\varphi(t)-1)} = \varphi_\lambda(t),$$

que é uma função contínua em $t = 0$, temos do Teorema da Continuidade de Paul-Lévy que φ_λ é uma função caraterística. □

As distribuições correspondentes às funções características $\varphi_\lambda(t) = e^{\lambda(\varphi(t)-1)}$, onde $\varphi(t)$ é uma função característica, são chamadas de **Distribuições de Poisson Generalizadas**. No caso particular em que $\varphi(t) = e^{it}$ (que obtemos de $X \equiv 1$), temos

$$\varphi_\lambda(t) = e^{\lambda(e^{it}-1)},$$

que é a função característica correspondente a distribuição de Poisson com parâmetro λ usual.

Teorema 4.1.5. *Para todo $\alpha \in (0, 2]$,*

$$\varphi_\alpha(t) = e^{-|t|^\alpha}$$

é uma função característica.

Demonstração. A prova é dividida em dois casos: $0 < \alpha \leq 1$ e $1 < \alpha < 2$. A prova no primeiro caso consiste em verificar as hipóteses do Teorema de Pólya (Teorema 4.1.3). O segundo caso é uma aplicação do Teorema da Continuidade de Paul-Lévy.

Para $0 < \alpha \leq 1$, temos claramente que $\varphi_\alpha(0) = 1$, $\varphi_\alpha(-t) = \varphi_\alpha(t)$, $\varphi_\alpha(t) \geq 0$, φ_α é contínua e decrescente em $\mathbb{R}^+$. Além disso, para $t \in \mathbb{R}^+$ temos $\varphi'_\alpha(t) = -\alpha t^{\alpha-1} e^{-t^\alpha}$, e logo,

$$\varphi''_\alpha(t) = e^{-t^\alpha}\left(\alpha^2 t^{2\alpha-2} - \alpha(\alpha-1)t^{\alpha-2}\right) \geq 0,$$

uma vez que $\alpha^2 > 0$ e $-\alpha(\alpha-1) \geq 0$ para $0 < \alpha \leq 1$. Daí, φ_α é convexa em $\mathbb{R}^+$, e logo, pelo Teorema de Pólya, φ_α é uma função característica.

Para $1 < \alpha < 2$, considere a função $\rho_\alpha : \mathbb{R} \to \mathbb{R}$ dada por

$$\rho_\alpha(x) = \begin{cases} \frac{\alpha}{2|x|^{\alpha+1}}, & \text{se } |x| > 1 \\ 0, & \text{se } |x| \leq 1 \end{cases}.$$

Observe que ρ_α é uma função par, e logo,

$$\int_{-\infty}^{\infty} \rho_\alpha(x)dx = 2\int_0^{\infty} \rho_\alpha(x)dx = \int_1^{\infty} \frac{\alpha}{x^{\alpha+1}}dx = \left.\frac{-1}{x^\alpha}\right|_{x=1}^{x=\infty} = 1.$$

Daí temos que ρ_α é uma função densidade de probabilidade. Denotando por $\tilde{\varphi}_\alpha$ a função característica correspondente, $\tilde{\varphi}_\alpha(t) = \int_{-\infty}^{\infty} e^{itx}\rho_\alpha(x)dx$, segue que

$$\begin{aligned} 1 - \tilde{\varphi}_\alpha(t) &= \int_{-\infty}^{\infty} \rho_\alpha(x)dx - \int_{-\infty}^{\infty} e^{itx}\rho_\alpha(x)dx \\ &= \int_{-\infty}^{\infty} (1 - e^{itx})\rho_\alpha(x)dx \\ &= \int_{|x|>1} (1 - \cos(tx) - i\sin(tx))\frac{\alpha}{2|x|^{\alpha+1}}dx \\ &= \frac{\alpha}{2}\int_{|x|>1} \frac{(1-\cos(tx))}{|x|^{\alpha+1}}dx - \frac{i\alpha}{2}\int_{|x|>1} \frac{\sin(tx)}{|x|^{\alpha+1}}dx. \end{aligned}$$

Como $\frac{\sin(tx)}{|x|^{\alpha+1}}$ é uma função ímpar e $\frac{1-\cos(tx)}{|x|^{\alpha+1}}$ é uma função par, segue que

$$1 - \tilde{\varphi}_\alpha(t) = \frac{\alpha}{2}\int_{|x|>1} \frac{1-\cos(tx)}{|x|^{\alpha+1}}dx = \alpha\int_1^\infty \frac{1-\cos(tx)}{x^{\alpha+1}}dx.$$

Suponha que $t \geq 0$ e tome $u = tx$. Logo, $x = \frac{u}{t}$, $dx = \frac{du}{t}$, $u = t$ para $x = 1$ e $u = \infty$ para $x = \infty$, e daí,

$$\begin{aligned} 1 - \tilde{\varphi}_\alpha(t) &= \alpha\int_t^\infty \frac{(1-\cos(u))}{\left(\frac{u}{t}\right)^{\alpha+1}}\frac{du}{t} = \alpha t^\alpha \int_t^\infty \frac{1-\cos(u)}{u^{\alpha+1}}du \\ &= \alpha t^\alpha \left(\int_0^\infty \frac{1-\cos(u)}{u^{\alpha+1}}du - \int_0^t \frac{1-\cos(u)}{u^{\alpha+1}}du\right). \end{aligned}$$

Para $t < 0$, como $\tilde{\varphi}_\alpha(t) \in \mathbb{R}$ para $t \geq 0$, temos $\tilde{\varphi}_\alpha(-t) = \overline{\tilde{\varphi}_\alpha(t)} = \tilde{\varphi}_\alpha(t)$, e assim,

$$1 - \tilde{\varphi}_\alpha(t) = \alpha|t|^\alpha \left(\int_0^\infty \frac{1-\cos(u)}{u^{\alpha+1}}du - \int_0^t \frac{1-\cos(u)}{u^{\alpha+1}}du\right).$$

Como $1 - \cos(u)$ é assintoticamente equivalente a $\frac{u^2}{2}$ em uma vizinhança $u = 0$, temos que quando t tende a zero, $\int_0^t \frac{1-\cos(u)}{u^{\alpha+1}}du$ é assintoticamente equivalente a

$$\int_0^t \frac{u^{1-\alpha}}{2}du = \left.\frac{u^{2-\alpha}}{2(2-\alpha)}\right|_{u=0}^{u=t} = \frac{t^{2-\alpha}}{2(2-\alpha)}.$$

Daí, como $\int_t^\infty \frac{1-\cos(u)}{u^{\alpha+1}}du < \infty$ para todo $t > 0$, temos

$$\int_0^\infty \frac{1-\cos(u)}{u^{\alpha+1}}du = \int_0^t \frac{1-\cos(u)}{u^{\alpha+1}}du + \int_t^\infty \frac{1-\cos(u)}{u^{\alpha+1}}du = M < \infty.$$

Segue que

$$1 - \tilde{\varphi}_\alpha(t) = \alpha|t|^\alpha \left(M - \int_0^t \frac{1-\cos(u)}{u^{\alpha+1}}du\right),$$

e logo,

$$\tilde{\varphi}_\alpha(t) = 1 - C_\alpha|t|^\alpha + \psi_\alpha(t),$$

onde $C_\alpha = \alpha M$ é uma constante positiva que depende de α e $\psi_\alpha(t) = \alpha|t| \int_0^t \frac{1-\cos(u)}{u^{\alpha+1}}du$ é uma função de ordem quadrática quando t tende a zero, já que nesse caso é assintoticamente equivalente a $\frac{\alpha t^2}{2(2-\alpha)}$. Agora, pela Proposição 2.4.1 temos que para todo n,

$$\left(\tilde{\varphi}_\alpha\left(\frac{t}{n^{\frac{1}{\alpha}}}\right)\right)^n = \left(1 - \frac{C_\alpha|t|^\alpha}{n} + \psi_\alpha\left(\frac{t}{n^{\frac{1}{\alpha}}}\right)\right)^n$$

é uma função característica, e como $\lim\limits_{n\to\infty} \frac{t}{n^{\frac{1}{\alpha}}} = 0$, fixando $t \in \mathbb{R}$ temos

$$\lim_{n\to\infty} n\psi_\alpha\left(\frac{t}{n^{\frac{1}{\alpha}}}\right) = \lim_{n\to\infty} \frac{n\alpha t^2}{2(2-\alpha)n^{\frac{2}{\alpha}}} = \lim_{n\to\infty} \frac{\alpha t^2}{2(2-\alpha)n^{\frac{2-\alpha}{\alpha}}} = 0,$$

e daí, tomando $c_{n,k} = \frac{-C_\alpha|t|^\alpha}{n} + \psi_\alpha\left(\frac{t}{n^{\frac{1}{\alpha}}}\right)$ para $k = 1, \cdots, n$, temos

- $\sum\limits_{k=1}^{n} c_{n,k} = -C_\alpha|t|^\alpha + n\psi_\alpha\left(\frac{t}{n^{\frac{1}{\alpha}}}\right) \overset{n\to\infty}{\longrightarrow} -C_\alpha|t|^\alpha$;
- $\max\limits_{1\leq k\leq n} |c_{n,k}| = \frac{C_\alpha|t|^\alpha}{n} + \psi_\alpha\left(\frac{t}{n^{\frac{1}{\alpha}}}\right) \overset{n\to\infty}{\longrightarrow} 0$;
- $\sum\limits_{k=1}^{n} |c_{n,k}| \leq C_\alpha|t|^\alpha + \left|n\psi_\alpha\left(\frac{t}{n^{\frac{1}{\alpha}}}\right)\right| \leq \tilde{M} < \infty$, onde $\tilde{M}$ é uma constante que não depende de n.

Segue do Lema 3.2.1 que

$$\left(\tilde{\varphi}_\alpha\left(\frac{t}{n^{\frac{1}{\alpha}}}\right)\right)^n \overset{n\to\infty}{\longrightarrow} e^{-C_\alpha|t|^\alpha},$$

que é uma função contínua em $t = 0$, e consequentemente, pelo Teorema da Continuidade de Paul-Lévy é uma função característica. Como $\varphi_\alpha(t) = \exp\left(-C_\alpha|\frac{t}{C_\alpha^{\frac{1}{\alpha}}}|^\alpha\right)$, segue da Proposição 2.4.1 que φ_α também é uma função característica. □

A mudança de parâmetro que fizemos no final da demonstração deixa claro que para qualquer número real positivo γ, $\varphi_{\gamma,\alpha}(t) = e^{-\gamma|t|^\alpha}$ também é uma função característica. Na verdade, Lévy mostrou em 1923 que existem constantes γ (que dependem de α), complexas e não reais, para as quais $\varphi_{\gamma,\alpha}(t) = e^{-\gamma|t|^\alpha}$ é uma função característica, e determinou a forma exata de tais constantes. A partir daí obteve-se a forma canônica das funções características correspondentes às distribuições estáveis que veremos no Teorema 4.3.2.

4.2 Distribuições Infinitamente Divisíveis

Pela Lei dos Grandes Números de Khintchin (Teorema 3.2.2) temos que se as variáveis aleatórias $X_1, X_2, \cdots$ são independentes e identicamente distribuídas, com $\mathbb{E}X_i = m < \infty$ para todo $i = 1, 2, \cdots$, e $S_n = X_1 + \cdots + X_n$, então $\frac{S_n}{n} \xrightarrow{P} m$. Se tirarmos a hipótese de que a esperança das variáveis aleatórias é finita e bem definida, temos do Exemplo 3.2.2 que $\frac{S_n}{n}$ converge para a distribuição de Cauchy com parâmetro θ se $X_1, X_2, \cdots$ tiverem tal distribuição, e fica seguinte questão: o que podemos dizer do limite T se $\frac{S_n}{n} \xrightarrow{D} T$?

A resposta para a pergunta acima é que T é uma variável aleatória com distribuição estável, como veremos na próxima seção. Mas antes vamos mostrar que existe uma relação entre as sequências acima e os arranjos triangulares, e analisar o limite de tal sequência. Dadas as sequências $\{a_n\}_{n\geq 1}$ e $\{b_n\}_{n\geq 1}$ de números reais com $b_n < \infty$ e $0 < a_n < \infty$ para todo n, defina

$$X_{n,k} = \frac{X_k}{a_n} - \frac{b_n}{na_n}, \ 1 \leq k \leq n. \tag{4.2}$$

Logo, temos

$$X_{n,1} + \cdots + X_{n,n} = \frac{X_1 + \cdots + X_n}{a_n} - \frac{nb_n}{na_n} = \frac{S_n - b_n}{a_n},$$

formando-se assim a seguinte sequência:

$$\begin{aligned}
\frac{S_1 - b_1}{a_1} &= X_{1,1} \\
\frac{S_2 - b_2}{a_2} &= X_{2,1} + X_{2,2} \\
&\vdots \\
\frac{S_n - b_n}{a_n} &= X_{n,1} + X_{n,2} + \cdots + X_{n,n} \\
&\vdots
\end{aligned}$$

Assim, os possíveis limites das sequências $\frac{S_n - \mathbb{E}S_n}{n}$ da Lei dos Grandes Números (Teorema 3.2.2) e $\frac{S_n - \mathbb{E}S_n}{s_n}$ do Teorema Central do Limite (Teorema 3.2.3) são os mesmos que podem aparecer no arranjo triangular do Teorema de Poisson (Teorema 3.2.1).

Definição 4.2.1 (Distribuição Infinitamente Divisível)**.** *Seja X uma variável aleatória cuja função de distribuição é F. Dizemos que X tem distribuição infinitamente divisível se para todo $n \geq 1$ existe uma função de distribuição F_n tal que*

$$F = \underbrace{F_n * \cdots * F_n}_{n-fatores}.$$

Proposição 4.2.1. *Seja X uma variável aleatória cuja função de distribuição é F e a função característica é φ. Então as seguintes afirmações são equivalentes:*

1. *X tem distribuição infinitamente divisível;*

2. *Para todo $n \geq 1$ existem variáveis aleatórias $X_{n,1}, \cdots, X_{n,n}$ independentes e identicamente distribuídas tais que $X \stackrel{D}{=} X_{n,1} + \cdots + X_{n,n}$;*

3. *Para todo $n \geq 1$ existe uma função caraterística φ_n tal que $\varphi = (\varphi_n)^n$.*

Assim, fará sentido nos referirmos também à função característica φ como infinitamente divisível.

Demonstração. No Teorema 2.4.1 e Corolário 2.4.2 já mostramos que a validade da segunda afirmação implica a validade da primeira e da terceira. Para mostrar que a primeira implica a segunda, observe que se X tem distribuição infinitamente divisível, então para todo n existe uma função de distribuição F_n tal que $F = \underbrace{F_n * \cdots * F_n}_{n-fatores}$. Por outro lado, pelo Teorema 2.2.1 existem variáveis aleatórias $X_{n,1}, \cdots, X_{n,n}$ independentes com função de distribuição comum e igual a F_n, e logo, pelo Corolário 2.4.2, a distribuição de $X_{n,1} + \cdots + X_{n,n}$ será F, isto é, $X \stackrel{D}{=} X_{n,1} + \cdots + X_{n,n}$. A prova de que a terceira implica a segunda é análoga. $\square$

Exemplo 4.2.1. *A distribuição normal com parâmetros μ e σ^2 é infinitamente divisível.*

Demonstração. Pelo Corolário 2.5.1 temos que a função característica correspondente é $\varphi(t) = \exp(it\mu - \frac{\sigma^2 t^2}{2})$. Assim, dado $n \geq 1$, temos

$$\varphi(t) = \left(\exp\left(\frac{it\mu}{n} - \frac{\sigma^2 t^2}{2n}\right)\right)^n = (\varphi_n(t))^n.$$

Para n fixo, novamente pelo Corolário 2.5.1 temos que $\varphi_n(t) = \exp(\frac{it\mu}{n} - \frac{\sigma^2 t^2}{2n})$ é a função característica correspondente à variável aleatória Y_n tal que $Y_n \sim N(\frac{\mu}{n}, \frac{\sigma^2}{n})$. Assim a distribuição normal é infinitamente divisível, quaisquer que sejam os parâmetros μ e σ^2. □

Exemplo 4.2.2. *A distribuição de Poisson com parâmetros $\lambda > 0$ é infinitamente divisível.*

Demonstração. Pela Proposição 2.5.4 temos que a função característica correspondente é $\varphi(t) = e^{\lambda(e^{it}-1)}$. Assim, dado $n \geq 1$, temos

$$\varphi(t) = \left(e^{\frac{\lambda}{n}(e^{it}-1)}\right)^n = (\varphi_n(t))^n.$$

Para n fixo, novamente pela Proposição 2.5.4 temos que $\varphi_n(t) = e^{\frac{\lambda}{n}(e^{it}-1)}$ é a função característica correspondente à variável aleatória Y_n tal que $Y_n \sim Poisson\left(\frac{\lambda}{n}\right)$. Assim a distribuição de Poisson é infinitamente divisível, qualquer que seja o parâmetro $\lambda > 0$. □

Exemplo 4.2.3. *A distribuição de Cauchy com parâmetro $\theta > 0$ é infinitamente divisível.*

Demonstração. Pelo Corolário 2.5.3 temos que a função característica correspondente é $\varphi(t) = e^{\theta|t|}$. Assim, dado $n \geq 1$, temos

$$\varphi(t) = \left(e^{\frac{\theta|t|}{n}}\right)^n = (\varphi_n(t))^n.$$

Para n fixo, novamente pelo Corolário 2.5.3 temos que $\varphi_n(t) = e^{\frac{\theta|t|}{n}}$ é a função característica correspondente à distribuição de Cauchy com parâmetro $\frac{\theta}{n}$. Assim a distribuição de Cauchy é infinitamente divisível, qualquer que seja o parâmetro $\theta > 0$. □

Lema 4.2.1. *Sejam X e Y variáveis aleatórias discretas independentes cujas distribuições F_X e F_Y têm m e n átomos respectivamente. Então $F_X * F_Y$ terá no mínimo $m+n-1$ e no máximo mn átomos.*

Demonstração. Como F_X e F_Y têm m e n átomos, respectivamente, temos $X(\Omega) = \{x_1, \cdots, x_m\}$ e $Y(\Omega) = \{y_1, \cdots, y_n\}$, e logo, $(X+Y)(\Omega) = \{x_i + y_j;\ i = 1, \cdots, m \text{ e } j = 1, \cdots, n\}$. Como X e Y são independentes, temos que $F_X * F_Y$ é a distribuição de $X+Y$, e portanto tem uma quantidade de átomos limitada por $\#(X+Y)(\Omega) \leq mn$, o que prova que a quantidade máxima de átomos de $F_X * F_Y$ é mn. Por outro lado, supondo sem perda de generalidade que $x_1 < x_2 < \cdots < x_m$ e $y_1 < y_2 < \cdots < y_n$, temos que

$$x_1 + y_1 < x_2 + y_1 < \cdots < x_m + y_1 < x_m + y_2 < \cdots < x_m + y_n$$

são $m + n - 1$ elementos distintos de $(X + Y)(\Omega)$, o que prova que $m + n - 1$ é uma cota inferior para a quantidade de átomos de $F_X * F_Y$. É fácil verificar que o mínimo é atingido quando $X(\Omega) = \{x, x+\delta, \cdots, x+(m-1)\delta\}$ e $Y(\Omega) = \{y, y+\delta, \cdots, y+(n-1)\delta\}$, com $\delta > 0$, pois nesse caso, $(X+Y)(\Omega) = \{x+y, x+y+\delta, \cdots, x+y+(m+n-2)\delta\}$. O máximo, por sua vez, é atingido quando $X(\Omega) = \{x, x+\delta, \cdots, x+(m-1)\delta\}$ e $Y(\Omega) = \{y, y+m\delta, \cdots, y+(n-1)m\delta\}$, com $\delta > 0$, pois nesse caso, $(X+Y)(\Omega) = \{x+y, x+y+\delta, \cdots, x+y+(mn-1)\delta\}$. □

Proposição 4.2.2. *Seja X uma variável aleatória discreta com imagem no conjunto finito $\{x_1, \cdots, x_k\}$. Então X tem distribuição infinitamente divisível se, e somente se, é degenerada num ponto $\lambda \in \mathbb{R}$.*

Demonstração. Se X é degenerada em $\lambda \in \mathbb{R}$, então a função característica de X é $\varphi(t) = e^{it\lambda}$, $t \in \mathbb{R}$. Logo, para todo $n \geq 1$ temos $\varphi(t) = \left(e^{it\frac{\lambda}{n}}\right)^n$, onde $e^{it\frac{\lambda}{n}}$ é a função característica das variáveis aleatórias degeneradas em $\frac{\lambda}{n}$, e portanto, X tem distribuição infinitamente divisível.

Reciprocamente, se $X(\Omega) = \{x_1, \cdots, x_n\}$ com $n > 1$, então X não tem distribuição infinitamente divisível, pois dado $k > n$, temos pelo Lema 4.2.1 (aplicado $k - 1$ vezes) que a soma de k variáveis aleatórias independentes e identicamente distribuídas ou é degenerada num ponto (no caso em que as variáveis aleatórias o são), ou então possui no mínimo $k + 1$ átomos (no caso em que as variáveis aleatórias possuem dois átomos). Logo, não temos como decompor X como a soma de k variáveis aleatórias independentes e identicamente distribuídas, e portanto, X não tem distribuição infinitamente divisível. Para verificar que o número mínimo de átomos de $X_1 + \cdots + X_k$ é realmente $k + 1$, onde $X_1, \cdots, X_k$ são variáveis aleatórias independentes não degeneradas com n átomos cada, basta observar que o mínimo é atingido para $n = 2$, onde temos que

- $X_1 + X_2$ têm no mínimo $2 + 2 - 1 = 3$ átomos,
- $X_1 + X_2 + X_3$ têm no mínimo $3 + 2 - 1 = 4$ átomos,

e por recorrência,

- $X_1 + X_2 + \cdots + X_k$ têm no mínimo $k + 2 - 1 = k + 1$ átomos.

□

Exemplo 4.2.4. *Como aplicação direta da Proposição 4.2.2 temos que se $X \sim b(n, p)$, então X não tem distribuição infinitamente divisível, a não ser que $n = p = 1$.*

Teorema 4.2.1. *Considere a sequência de variáveis aleatórias que formam um arranjo triangular, isto é,*

$$\begin{aligned} T_1 &= X_{1,1} \\ T_2 &= X_{2,1} + X_{2,2} \\ T_3 &= X_{3,1} + X_{3,2} + X_{3,3} \\ &\vdots \\ T_n &= X_{n,1} + \cdots + X_{n,n} \qquad (4.3) \\ &\vdots \end{aligned}$$

Suponha que $X_{n,1}, \cdots, X_{n,n}$ são variáveis aleatórias independentes e identicamente distribuídas para todo $n \geq 1$. Então $T_n \xrightarrow{D} T$ se, e somente se, T tem distribuição infinitamente divisível.

Demonstração. Suponhamos inicialmente que T tem distribuição infinitamente divisível. Logo, para todo $n \geq 1$ existem variáveis aleatórias e identicamente distribuídas $X_{n,1}, \cdots, X_{n,n}$ tais que

$$T \overset{D}{=} X_{n,1} + \cdots + X_{n,n} \overset{Def}{=} T_n,$$

e trivialmente $T_n \xrightarrow{D} T$.

Reciprocamente, suponha que $T_n \xrightarrow{D} T$. Precisamos mostrar que para todo $k \geq 1$ existem variáveis aleatórias $\eta_1, \cdots, \eta_n$ independentes e identicamente distribuídas tais que $T \overset{D}{=} \eta_1 + \cdots + \eta_k$. Para tanto, dado $k \geq 1$, vamos definir $X_n^{(k)}$ por

$$X_n^{(1)} = X_{nk,1} + \cdots + X_{nk,n}, \quad \cdots, \quad X_n^{(k)} = X_{nk,n(k-1)+1} + \cdots + X_{nk,nk},$$

e logo, por (4.3)

$$\begin{aligned} T_1 &= X_{1,1} \\ T_2 &= X_{2,1} + X_{2,2} \\ &\vdots \\ T_k &= \underbrace{X_{k,1}}_{=X_1^{(1)}} + \underbrace{X_{k,2}}_{=X_1^{(2)}} + \cdots + \underbrace{X_{k,k}}_{=X_1^{(k)}} \\ &\vdots \\ T_{2k} &= \underbrace{X_{2k,1} + X_{2k,2}}_{=X_2^{(1)}} + \underbrace{X_{2k,3} + X_{2k,4}}_{=X_2^{(2)}} + \cdots + \underbrace{X_{2k,2k-1} + X_{2k,2k}}_{=X_2^{(k)}} \\ &\vdots \\ T_{nk} &= \underbrace{X_{nk,1} + \cdots + X_{nk,n}}_{=X_n^{(1)}} + \underbrace{X_{nk,n+1} + \cdots + X_{nk,2n}}_{=X_n^{(2)}} + \cdots + \underbrace{X_{nk,(k-1)n+1} + \cdots + X_{nk,nk}}_{=X_n^{(k)}} \\ &\vdots \end{aligned}$$

ou seja, $T_{nk} = X_n^{(1)} + \cdots + X_n^{(k)}$.

Como $T_{nk} \xrightarrow{D} T$, a sequência de funções de distribuição correspondente às variáveis aleatórias T_{nk} é relativamente compacta, e logo, pelo Teorema de Prohorov (Teorema 3.3.2) é rígida. Além disso, como $X_n^{(1)}, \cdots, X_n^{(k)}$ são independentes e identicamente distribuídas, temos

$$[P(X_n^{(1)} > z)]^k = P(X_n^{(1)} > z, \cdots, X_n^{(k)} > z) \leq P(T_{nk} > kz),$$

ao mesmo tempo que

$$[P(X_n^{(1)} < z)]^k = P(X_n^{(1)} < z, \cdots, X_n^{(k)} < z) \leq P(T_{nk} < kz),$$

onde as desigualdades se justificam no fato de termos $\{\omega; X_n^{(j)}(\omega) > z, j = 1, \cdots, k\} \subset \{\omega; T_n(\omega) = X_n^{(1)}(\omega) + \cdots, X_n^{(k)}(\omega) > kz\}$ e $\{\omega; X_n^{(j)}(\omega) < z, j = 1, \cdots, k\} \subset \{\omega; T_n(\omega) = X_n^{(1)}(\omega) + \cdots, X_n^{(k)}(\omega) < kz\}$. Daí temos que a sequência de distribuições de $\{X_n^{(1)}\}_{n\geq 1}$ também é rígida, e novamente pelo Teorema de Prohorov, é relativamente compacta. Assim, tomando uma subsequência se necessário, temos que existe uma variável aleatória η_1 tal que $X_n^{(1)} \xrightarrow{D} \eta_1$. Além disso, como $X_n^{(1)}, \cdots, X_n^{(k)}$ são identicamente distribuídas, temos também, $X_n^{(2)} \xrightarrow{D} \eta_2, \cdots, X_n^{(k)} \xrightarrow{D} \eta_k$, onde $\eta_1, \cdots, \eta_k$ são identicamente distribuídas. Dessa forma

$$T_{nk} = X_n^{(1)} + \cdots + X_n^{(k)} \xrightarrow{D} \eta_1 + \cdots + \eta_k,$$

ao mesmo tempo que $T_{nk} \xrightarrow{D} T$, e portanto

$$T \stackrel{D}{=} \eta_1 + \cdots + \eta_k,$$

qualquer que seja $k \geq 1$. Para provar que a distribuição de T é infinitamente divisível, falta apenas verificar que $\eta_1, \cdots, \eta_k$ são independentes, o que decorre da Proposição 2.1.3, já que, pela independência de $X_n^{(1)}, \cdots, X_n^{(k)}$ para todo $n \geq 1$, temos

$$\begin{aligned} F_{\eta_1,\cdots,\eta_k}(x_1, \cdots, x_k) &= P(\eta_1 \leq x_1, \cdots, \eta_k \leq x_k) \\ &= \lim_{n\to\infty} P(X_n^{(1)} \leq x_1, \cdots, X_n^{(k)} \leq x_k) \\ &= \lim_{n\to\infty} \left[P(X_n^{(1)} \leq x_1) \cdots P(X_n^{(k)} \leq x_k)\right] \\ &= P(\eta_1 \leq x_1) \cdots P(\eta_k \leq x_k), \end{aligned}$$

ao mesmo tempo que $\lim_{x\to\infty} P(\eta_j \leq x) = 1$ para todo $j = 1, \cdots, k$, uma vez que $P(\eta_j \leq x)$ é a função de distribuição de η_j. □

Uma vez provado o teorema, fica clara a importância das distribuições infinitamente divisíveis, visto que, tratando-se de variáveis aleatórias independentes e identicamente distribuídas, tanto os arranjos triangulares quanto as sequências do tipo $\frac{S_n + b_n}{a_n}$ convergem exclusivamente para tais. Precisamos portanto, conseguir propriedades que ajudem a identificar as distribuições infinitamente divisíveis.

Proposição 4.2.3. *Uma função característica infinitamente divisível nunca se anula.*

Demonstração. Seja $\varphi : \mathbb{R} \to \mathbb{C}$ uma função característica infinitamente divisível. Logo, para todo $n \geq 1$ existe uma função característica $\varphi_n : \mathbb{R} \to \mathbb{C}$ tal que

$$\varphi(t) = [\varphi_n(t)]^n.$$

Sejam $g : \mathbb{R} \to \mathbb{R}^+$ e $g_n : \mathbb{R} \to \mathbb{R}^+$ definidas por

$$g(t) = |\varphi(t)|^2 \qquad \text{e} \qquad g_n(t) = |\varphi_n(t)|^2.$$

Note que pelo Corolário 2.4.3, g e g_n também são funções características. Além disso, como $g(t) \geq 0$ para todo $t \in \mathbb{R}$, sua n-ésima raiz, real e positiva, é unicamente determinada. Vamos denotá-la por $[g(t)]^{\frac{1}{n}}$. Observe que

$$g(t) = |\varphi(t)|^2 = |[\varphi_n(t)]^n|^2 = |\varphi_n(t)|^{2n} = [g_n(t)]^n.$$

Logo, como $g_n(t) \geq 0$, temos

$$g_n(t) = [g(t)]^{\frac{1}{n}}, \text{ para todo } t \in \mathbb{R}.$$

Por outro lado, como $0 \leq g(t) \leq 1$, temos que $[g(t)]^{\frac{1}{n}} \overset{n\to\infty}{\longrightarrow} 0$ se $g(t) = 0$ e $[g(t)]^{\frac{1}{n}} \overset{n\to\infty}{\longrightarrow} 1$ se $g(t) \neq 0$. Assim $g_n(t)$ converge pontualmente para a função $h(t)$ definida por

$$h(t) = \begin{cases} 0, & \text{se } g(t) = 0 \\ 1, & \text{se } g(t) \neq 0 \end{cases}.$$

Além disso, como g é contínua em $t = 0$ com $g(0) = 1$, existe $t_0 > 0$ tal que $g(t) \neq 0$ para todo $t \in (-t_0, t_0)$, e logo temos $h(t) = 1$ para $t \in (-t_0, t_0)$. Assim h é uma função contínua em $t = 0$ com $h(0) = 1$, e assim, pelo Teorema da Continuidade de Paul-Lévy, $h(t)$ é uma função característica. Segue que h é contínua em $\mathbb{R}$, e portanto, $h(t) = 1$ para todo $t \in \mathbb{R}$, o que acarreta em $g(t) \neq 0$, e consequentemente, $\varphi(t) \neq 0$ para todo $t \in \mathbb{R}$. $\square$

Uma consequência imediata da Proposição 4.2.3 é que a distribuição uniforme não é infinitamente divisível, já que para $t \neq 0$, a função característica correspondente é da forma

$$\varphi(t) = \frac{\sin(\delta t)}{\delta t} e^{it\mu},$$

onde $\delta > 0$ e $\mu \geq 0$, e portanto se anula infinitas vezes. Nesse sentido, tal proposição torna-se uma ferramenta eficaz em identificar distribuições que não são infinitamente divisíveis. No entanto, a recíproca não é verdadeira, isto é, uma função característica que nunca se anula não é necessariamente infinitamente divisível. Um exemplo nesse sentido é $\varphi(t) = (1-p)\cos(\delta t) + p$, $0 \leq p < 1$ e $\delta > 0$, que é a função característica da variável aleatória X tal que $P(X = \delta) = P(X = -\delta) = \frac{1-p}{2}$ e $P(X = 0) = p$. Vimos na Proposição 4.2.2 que tais funções características não são infinitamente divisíveis para $0 \leq p < 1$, e ao mesmo tempo, nunca se anulam para $0 < p < 1$.

Lema 4.2.2. *Seja $\varphi : \mathbb{R} \to \mathbb{C}$ uma função característica que nunca se anula. Então existe uma única função contínua $\lambda : \mathbb{R} \to \mathbb{C}$ tal que $\lambda(0) = 0$ e $\varphi(t) = e^{\lambda(t)}$, $t \in \mathbb{R}$.*

Demonstração. Fixemos um inteiro positivo n. Como $[-n, n]$ é compacto e φ nunca se anula, existe $\rho = \rho(n) > 0$ tal que $|\varphi(t)| \geq \rho$ para todo $t \in [-n, n]$. Além disso, como φ é uniformemente contínua, existe $\delta = \delta(\rho(n))$ tal que para quaisquer t e t' em $[-n, n]$ com $|t - t'| < \delta$ tem-se $|\varphi(t) - \varphi(t')| \leq \frac{\rho}{2} \leq \frac{1}{2}$. Consideremos então uma partição de $[-n, n]$ dada por

$$-n = t_{-l} < \cdots < t_{-1} < t_0 = 0 < t_1 < \cdots < t_l = n,$$

onde $|t_k - t_{k-1}| < \delta$ para $k = -l + 1, \cdots, l$. Consideremos também a função $L : \mathbb{C} \to \mathbb{C}$ dada por

$$L(z) = \sum_{j=1}^{\infty} \frac{(-1)^j}{j!}(z-1)^j,$$

que é a única determinação de $\ln(z)$ em $|z - 1| \leq \frac{1}{2}$ que se anula em $z = 1$. Para $t_{-1} \leq t \leq t_1$, vamos definir $\lambda(t)$ por

$$\lambda(t) = L(\varphi(t)) = \sum_{j=1}^{\infty} \frac{(-1)^j}{j!}(\varphi(t) - 1)^j.$$

Nesse caso, temos $|t-0| = |t| \leq \delta$, e logo, $|\varphi(t)-1| = |\varphi(t)-\varphi(0)| \leq \frac{1}{2}$. Daí, $\sum\limits_{j=1}^{\infty} \frac{(-1)^j}{j!}(\varphi(t)-1)^j$ converge uniformemente a uma função contínua, no caso, $\lambda(t)$. Além disso, temos trivialmente que $\lambda(0) = 0$.

Agora, supondo λ bem definida em $[t_{-k}, t_k]$, vamos definir $\lambda(t)$ em $[t_k, t_{k+1}]$ por

$$\lambda(t) = \lambda(t_k) + \sum_{j=1}^{\infty} \frac{(-1)^j}{j!} \left(\frac{\varphi(t)}{\varphi(t_k)} - 1 \right)^j.$$

Nesse caso, como $|t - t_k| \leq \delta$, temos $|\varphi(t) - \varphi(t_k)| \leq \frac{\rho}{2}$, e como $|\varphi(t)| \geq \rho$, segue que

$$\left| \frac{\varphi(t)}{\varphi(t_k)} - 1 \right| = \frac{|\varphi(t) - \varphi(t_k)|}{|\varphi(t_k)|} \leq \frac{\frac{\rho}{2}}{\rho} = \frac{1}{2}.$$

Daí, $\sum\limits_{j=1}^{\infty} \frac{(-1)^j}{j!} \left(\frac{\varphi(t)}{\varphi(t_k)} - 1 \right)^j$ converge uniformemente a uma função contínua, no caso, $L\left(\frac{\varphi(t)}{\varphi(t_k)}\right)$. Logo λ é uma função contínua em $[t_{-k}, t_{k+1}]$ com

$$\lambda(t) = \begin{cases} L(\varphi(t)), & \text{se } t_{-1} \leq t \leq t_1 \\ \lambda(t_k) + L\left(\frac{\varphi(t)}{\varphi(t_k)}\right), & \text{se } t_k \leq t \leq t_{k+1} \end{cases}.$$

De modo análogo podemos estender $\lambda(t)$ continuamente a $[t_{-k-1}, t_{-k}]$ pondo

$$\lambda(t) = \lambda(t_{-k}) + L\left(\frac{\varphi(t)}{\varphi(t_{-k})} \right), \text{ se } t_{-k-1} \leq t \leq t_{-k},$$

e assim teremos $\lambda(t)$ bem definida e contínua em $[t_{-k-1}, t_{k+1}]$. Por recorrência conseguimos estender $\lambda(t)$ continuamente para $[-n, n]$ com $\lambda(0) = 0$. Além disso, da forma como $\lambda(t)$ foi construída, teremos também

$$e^{\lambda(t)} = \exp\left(\lambda(t_k) + L\left(\frac{\varphi(t)}{\varphi(t_k)}\right)\right) = e^{\lambda(t_k)} \exp\left(L\left(\frac{\varphi(t)}{\varphi(t_k)}\right)\right) = \varphi(t_k)\frac{\varphi(t)}{\varphi(t_k)} = \varphi(t).$$

Agora, tendo $\lambda(t)$ bem definida em $[-n, n]$, podemos estender $\lambda(t)$ a $[-n-1, n+1]$ da seguinte maneira: obtemos $\rho = \rho(n+1) > 0$ tal que $|\varphi(t)| \geq \rho$ para $t \in [-n-1, -n]$ e $t \in [n, n+1]$, e $\delta = \delta(\rho(n+1)) > 0$ tal que, para quaisquer t e t' em $[-n-1, -n]$ e $[n, n+1]$ com $|t - t'| \leq \delta$, tem-se $|\varphi(t) - \varphi(t')| \leq \frac{\rho}{2}$; depois estendemos a partição de $[-n, n]$ para $[-n-1, n+1]$ pondo

$$n = t_l < t_{l+1} < \cdots < t_{l+j} = n+1,$$

e

$$-n-1 = t_{-l-j} < \cdots < t_{-l-1} < t_{-l} = -n,$$

onde $|t_{k+1} - t_k| < \delta$ para $k = -l-j, \cdots, -l-1, l, \cdots, l+j-1$. A partir daí basta procedermos de modo análogo à construção que fizemos para $t \in [-n, n]$ e teremos $\lambda(t)$ bem definida em $[-n-1, n+1]$. Repetindo sucessivamente o processo, obteremos $\lambda(t)$ bem definida em $[-\infty, \infty]$. Por fim, para provar que $\lambda(t)$ é única, basta observar que se $\tilde{\lambda}(t)$ é contínua em $\mathbb{R}$ com $\tilde{\lambda}(0) = 0$ e $e^{\tilde{\lambda}(t)} = \varphi(t)$ para todo $t \in \mathbb{R}$, então $\tilde{\lambda}(t) - \lambda(t) = 2\pi i m(t)$, onde $m(t) \in \mathbb{Z}$ para todo $t \in \mathbb{R}$. Como $\lambda(t)$ e $\tilde{\lambda}(t)$ são contínuas, segue que $m : \mathbb{R} \to \mathbb{Z}$ é contínua, e portanto, constante. Além disso, temos $m(0) = \frac{\tilde{\lambda}(0) - \lambda(0)}{2\pi i} = 0$, e logo, $m \equiv 0$, donde segue que $\tilde{\lambda}(t) = \lambda(t)$ para todo $t \in \mathbb{R}$. □

Teorema 4.2.2. *Seja $\varphi : \mathbb{R} \to \mathbb{C}$ uma função característica infinitamente divisível. Então, para todo $n \geq 1$, a função característica φ_n tal que $\varphi = (\varphi_n)^n$ é o ramo da n-ésima raiz de φ obtido pelo ramo de logaritmo λ do Lema 4.2.2.*

Demonstração. Pela Proposição 4.2.3 temos que $\varphi(t) \neq 0$ para todo $t \in \mathbb{R}$. Logo, fixado n, teremos também $\varphi_n(t) \neq 0$ para todo $t \in \mathbb{R}$, já que $\varphi(t) = (\varphi_n(t))^n$. Assim, pelo Lema 4.2.2 existem $\lambda(t)$ e $\lambda_n(t)$ contínuas tais que $\lambda(0) = \lambda_n(0) = 1$, $e^{\lambda(t)} = \varphi(t)$ e $e^{\lambda_n(t)} = \varphi_n(t)$. Logo, como $e^{n\lambda_n(t)} = (\varphi_n(t))^n = \varphi(t) = e^{\lambda(t)}$, temos $\lambda(t) - n\lambda_n(t) = 2\pi i m_n(t)$, onde $m_n(t) \in \mathbb{Z}$ para todo $t \in \mathbb{R}$. Por outro lado, λ e λ_n são contínuas em $\mathbb{R}$, e logo, $m_n : \mathbb{R} \to \mathbb{Z}$ é contínua, e portanto, constante. Logo, como $m(0) = \frac{\lambda(0) - \lambda_n(0)}{2\pi i} = 0$, temos que $\lambda(t) = n\lambda_n(t)$, para todo $t \in \mathbb{R}$. Segue portanto que

$$\varphi_n(t) = e^{\lambda_n(t)} = e^{\frac{\lambda(t)}{n}},$$

e como λ é um ramo do logaritmo de φ, segue que φ_n é o ramo da n-ésima raiz correspondente. □

Corolário 4.2.1. *Uma função característica φ é infinitamente divisível se, e somente se, para todo $n \geq 1$, o ramo da n-ésima raiz de φ, φ_n tal que $\varphi_n(0) = 1$, for uma função característica. Em particular, se a função característica for real, então φ é infinitamente divisível se, e somente se, para todo $n \geq 1$, $\varphi_n = \sqrt[n]{\varphi}$ for uma função característica.*

Corolário 4.2.2. *Para todo $\alpha \in (0, 2]$, a função*

$$\varphi_\alpha(t) = e^{-|t|^\alpha},\ t \in \mathbb{R},$$

é uma função característica infinitamente divisível.

Demonstração. Já vimos no Teorema 4.1.5 que φ_α é uma função característica. Para mostrar que φ_α é infinitamente divisível, pelo Corolário 4.2.1 basta mostrar que para todo $n \geq 1$, $\sqrt[n]{e^{-|t|^\alpha}} = e^{\frac{-|t|^\alpha}{n}}$ é também uma função característica. Para tanto, observe que sendo X a variável aleatória cuja função característica é φ_α, então, pela Proposição 2.4.1, a função característica da variável aleatória $\frac{1}{n^{\frac{1}{\alpha}}}X$ será

$$\varphi_\alpha\left(\frac{t}{n^{\frac{1}{\alpha}}}\right) = e^{\frac{-|t|^\alpha}{n}}.$$

Isto conclui a prova. □

Corolário 4.2.3. *Para todo $\lambda \geq 0$, e para toda função característica φ, a função*

$$\varphi_\lambda(t) = e^{\lambda(\varphi(t)-1)},\ t \in \mathbb{R},$$

é uma função característica infinitamente divisível.

Demonstração. Já vimos no Teorema 4.1.4 que φ_λ é uma função característica. Para mostrar que φ_λ é infinitamente divisível, pelo Corolário 4.2.1 basta mostrar que para todo $n \geq 1$, $e^{\frac{\lambda(\varphi(t)-1)}{n}}$ é também uma função característica, e isso decorre imediatamente do Teorema 4.1.4, visto que $\frac{\lambda}{n} \geq 0$. □

Proposição 4.2.4. *A convolução de duas distribuições infinitamente divisíveis é infinitamente divisível.*

Demonstração. Sendo $\tilde{\varphi}$ e $\hat{\varphi}$ as funções características correspondentes, pela Proposição 4.2.1, basta mostrar que $\varphi(t) = \tilde{\varphi}(t)\hat{\varphi}(t)$ é uma função característica infinitamente divisível, e isso é imediato, pois dado $n \geq 1$, como $\tilde{\varphi}$ e $\hat{\varphi}$ são infinitamente divisíveis, existem $\tilde{\varphi}_n$ e $\hat{\varphi}_n$ tais que

$$\varphi(t) = \tilde{\varphi}(t)\hat{\varphi}(t) = (\tilde{\varphi}_n(t))^n(\hat{\varphi}_n(t))^n = (\tilde{\varphi}_n(t)\hat{\varphi}_n(t))^n,$$

onde $\tilde{\varphi}_n(t)\hat{\varphi}_n(t)$ é uma função característica, já que é o produto de tais. □

Teorema 4.2.3. *Seja $\{\varphi_k\}_{k\geq 1}$ uma sequência de funções características infinitamente divisíveis tais que φ_k converge pontualmente para uma função característica φ. Então, φ também é infinitamente divisível.*

Demonstração. Por definição, para todo $k \geq 1$ existem variáveis aleatórias independentes e identicamente distribuídas $X_{k,1}, \cdots, X_{k,k}$ tais que φ_k é a função característica de $T_k = X_{k,1} + \cdots + X_{k,k}$. Como φ_k converge pontualmente para a função característica φ, temos do Teorema da Continuidade de Paul-Lévy que φ corresponde a variável aleatória T tal que $T_n \xrightarrow{D} T$, a qual pelo Teorema 4.2.1 tem distribuição infinitamente divisível. $\square$

Vamos encerrar esta seção mostrando que a classe das distribuições infinitamente divisíveis coincide com o fecho (no sentido da convergência vaga) do conjunto gerado pela convolução de um número finito de distribuições de Poisson generalizadas. Para obter tal resultado vamos precisar do seguinte lema:

Lema 4.2.3. *Seja F uma função de distribuição. Então, para todo $n \geq 1$ existe $k \geq 1$ e uma partição $\mathcal{P} = \{-\infty < u_1 < \cdots < u_k < \infty\}$ de $\mathbb{R}$, tal que*

$$\sup_{|t|\leq n}\left|\sum_{j=1}^{k}(e^{itu_j}-1)n[F(u_j)-F(u_{j-1})] - \int_{-\infty}^{\infty}(e^{itu}-1)ndF(u)\right| \leq \frac{1}{n}.$$

Demonstração. Dado $n \geq 1$, temos imediatamente que existem $u_1 < \tilde{u}$ tais que $F(u_1) < \frac{1}{8n^2}$ e $F(\tilde{u}) > 1 - \frac{1}{8n^2}$. Por outro lado, sendo $\varphi(u) = e^{iu} - 1$, temos que φ é uniformemente contínua em $\mathbb{R}$, e daí, dado $\epsilon > 0$, existe $\delta > 0$ tal que $|u-u'| < \delta$ implica $|\varphi(u)-\varphi(u')| < \epsilon$, e logo,

$$|t| < n \text{ e } |u-u'| < \frac{\delta}{n} \Rightarrow |tu - tu'| < \delta \Rightarrow |\varphi(tu) - \varphi(tu')| = |(e^{itu}-1) - (e^{itu'}-1)| < \epsilon.$$

Assim, sendo $\epsilon = \frac{1}{2n^2}$, vamos considerar a partição

$$\mathcal{P} = \{-\infty < u_1 < u_2 < \cdots < u_k < \infty\},$$

onde $u_{j+1} - u_j = \frac{\delta}{n}$ para $j = 1, \cdots, k-1$ e $u_{k-1} \leq \tilde{u} < u_k$, e logo, teremos

$$|(e^{itu_j}-1) - (e^{itu}-1)| < \frac{1}{2n^2}$$

para todo $u_j \leq u \leq u_{j+1}$ e $|t| \leq n$. Agora, sendo

$$\varphi_t^k(u) = \begin{cases} e^{itu_j} - 1, & \text{se } u \in [u_j, u_{j+1}) \\ 0, & \text{se } u < u_1 \text{ ou } u \geq u_k, \end{cases}$$

segue que

$$\begin{aligned}
&\left|\sum_{j=1}^{k}(e^{itu_j}-1)[F(u_j)-F(u_{j-1})]-\int_{-\infty}^{\infty}(e^{itu}-1)dF(u)\right| \\
= &\left|\int_{-\infty}^{\infty}\varphi_t^k(u)dF(u)-\int_{-\infty}^{\infty}\varphi(tu)dF(u)\right| \\
= &\left|\int_{-\infty}^{\infty}[\varphi_t^k(u)-\varphi(tu)]dF(u)\right| \leq \int_{-\infty}^{\infty}|\varphi_t^k(u)-\varphi(tu)|dF(u) \\
= &\int_{-\infty}^{u_1}|\varphi(tu)|dF(u)+\int_{u_1}^{u_k}|(e^{itu_j}-1)-(e^{itu}-1)|dF(u)+\int_{u_k}^{\infty}|\varphi(tu)|dF(u) \\
\leq\; & 2\int_{-\infty}^{u_1}dF(u)+\frac{1}{2n^2}\int_{u_1}^{u_k}dF(u)+2\int_{u_k}^{\infty}dF(u) \\
\leq\; & 2F(u_1)+\frac{1}{2n^2}+2[1-F(u_k)]<\frac{1}{4n^2}+\frac{1}{2n^2}+\frac{1}{4n^2}=\frac{1}{n^2}.
\end{aligned}$$

Isso completa a prova do lema. □

Teorema 4.2.4. *Seja $\varphi:\mathbb{R}\to\mathbb{C}$ uma função característica. Então φ é infinitamente divisível se, e somente se, existem duas sequências de números reais, $a_{n,j}$ e $u_{n,j}$, com $1\leq j\leq k_n$, $n\geq 1$ e $a_{n,j}>0$, tais que*

$$\varphi(t)=\lim_{n\to\infty}\prod_{j=1}^{k_n}\exp\left(a_{n,j}(e^{itu_{n,j}}-1)\right).$$

Demonstração. Dadas as sequências $a_{n,j}$ e $u_{n,j}$, temos do Corolário 4.2.3 que para todo $n=1,\cdots,k_n$, $\exp\left(a_{n,j}(e^{itu_{n,j}}-1)\right)$ é uma função característica infinitamente divisível, e logo, pela Proposição 4.2.4, $\prod_{j=1}^{k_n}\exp\left(a_{n,j}(e^{itu_{n,j}}-1)\right)$ também o é. Segue portanto do Teorema 4.2.3 que $\varphi(t)=\lim_{n\to\infty}\prod_{j=1}^{k_n}\exp\left(a_{n,j}(e^{itu_{n,j}}-1)\right)$ é uma função característica infinitamente divisível.

Reciprocamente, seja φ uma função característica infinitamente divisível. Fixado $n\geq 1$, sejam φ_n a função característica tal que $\varphi(t)=(\varphi_n(t))^n$, F_n a função de distribuição correspondente a φ_n e λ o logaritmo de φ como no Lema 4.2.2. Daí, pela regra de L'Hôpital, para todo $t\in\mathbb{R}$ vale que

$$\begin{aligned}
\lim_{n\to\infty}n(\varphi_n(t)-1) &= \lim_{n\to\infty}n\left(e^{\frac{\lambda(t)}{n}}-1\right)=\lim_{n\to\infty}\frac{e^{\frac{\lambda(t)}{n}}-1}{\frac{1}{n}} \\
&= \lim_{n\to\infty}\frac{\frac{-\lambda(t)}{n^2}e^{\frac{\lambda(t)}{n}}}{\frac{-11}{n^2}}=\lim_{n\to\infty}\lambda(t)e^{\frac{\lambda(t)}{n}}=\lambda(t),
\end{aligned}$$

e consequentemente,

$$e^{n(\varphi_n(t)-1)}\xrightarrow{n\to\infty}e^{\lambda(t)}=\varphi(t). \tag{4.4}$$

Por outro lado, observe que pelo Lema 4.2.3, para todo $n\geq 1$ existe $k_n\geq 1$ e uma partição dada por $-\infty<u_{n,1}<\cdots<u_{n,k_n}<\infty$ tal que

$$\sup_{|t|\leq n}\left|\sum_{j=1}^{k_n}(e^{itu_{n,j}}-1)a_{n,j}-\int_{-\infty}^{\infty}(e^{itu}-1)ndF_n(u)\right|\leq\frac{1}{n},$$

onde $a_{n,j} = n[F_n(u_{n,j}) - F_n(u_{n,j-1})]$. Observe também que dados z e $z' \in \mathbb{C}$, temos

$$|e^z - e^{z'}| = |e^z|(e^{z-z'} - 1) \leq |e^z|(e^{|z-z'|} - 1).$$

Daí, tomando $z = \int_{-\infty}^{\infty}(e^{itu} - 1)ndF_n(u)$ e $z' = \sum_{j=1}^{k_n}(e^{itu_{n,j}} - 1)a_{n,j}$, segue que

$$\begin{aligned}
\sup_{|t|\leq n}\left|e^{n(\varphi_n(t)-1)} - \prod_{j=1}^{k_n}\exp(a_{n,j}(e^{itu_{n,j}} - 1))\right| &= \sup_{|t|\leq n}|e^z - e^{z'}| \\
&\leq \sup_{|t|\leq n}|e^z|(e^{|z-z'|} - 1) \\
&\leq e^{\frac{1}{n}} - 1 \overset{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

onde a última desigualdade se justifica no fato de termos $e^z \leq 1$ (pois $e^z = e^{n(\varphi_n(t)-1)}$ é uma função característica) e $|z - z'| \leq 1$. Segue daí e de (4.4) que

$$\lim_{n\to\infty}\prod_{j=1}^{k_n}\exp(a_{n,j}(e^{itu_{n,j}} - 1)) = \lim_{n\to\infty} e^{n(\varphi_n(t)-1)} = \varphi(t).$$

□

Como consequência do teorema acima temos que a classe das distribuições infinitamente divisíveis coincide com o fecho (no sentido da convergência fraca) do conjunto gerado pela convolução de um número finito de distribuições de Poisson generalizadas. A partir daí é possível obter também a forma canônica exata das funções características correspondentes.

Teorema 4.2.5 (Representação Canônica de Lévy-Khintchin)**.** *Toda função característica infinitamente divisível pode ser representada canonicamente por*

$$\varphi(t) = \exp\left[ait + \int_{-\infty}^{\infty}\left(e^{itu} - 1 - \frac{itu}{1+u^2}\right)\frac{1+u^2}{u^2}dG(u)\right],$$

onde $a \in \mathbb{R}$, G é uma função não decrescente e limitada em $(-\infty, \infty)$ e o integrando é contínuo em $u = 0$, sendo o seu valor dado por

$$\begin{aligned}
& \lim_{u\to\infty}\left(e^{itu} - 1 - \frac{itu}{1+u^2}\right)\frac{1+u^2}{u^2} \\
= \ & \lim_{u\to\infty}\frac{(e^{itu} - 1)(1+u^2) - itu}{u^2} \\
= \ & \lim_{u\to\infty}\frac{ite^{itu}(1+u^2) + 2u(e^{itu} - 1) - it}{2u} \\
= \ & \lim_{u\to\infty}\frac{-t^2e^{itu}(1+u^2) + 2itue^{itu} + 2(e^{itu} - 1) + 2itue^{itu}}{2} \\
= \ & \frac{-t^2}{2}.
\end{aligned}$$

Uma prova deste resultado pode ser visto em (GNEDENKO; KOLMOGOROV, 1954, Capítulo 3).

4.3 Distribuições Estáveis

O Exemplo 3.2.2 deixa claro que nem sempre as sequências do tipo $\frac{S_n}{n}$ (onde S_n é a soma das variáveis aleatórias independentes e identicamente distribuídas $X_1, \cdots, X_n$) convergem para uma variável aleatória com distribuição degenerada, como na Lei dos Grandes Números de Khintchin (Teorema 3.2.2). Assim somos levados a estudar os possíveis limites das sequências do tipo $\frac{S_n+b_n}{a_n}$ onde $a_n > 0$ e b_n são constantes reais. No Teorema 4.2.1 vimos que se $\frac{S_n+b_n}{a_n}$ converge para uma variável aleatória T, esta terá distribuição infinitamente divisível. Mas, nesta seção veremos que a classe das distribuições dos possíveis limites para tais sequências é um pouco mais restrita.

Definição 4.3.1 (Distribuição Estável)**.** *Seja X uma variável aleatória cuja função de distribuição é F. Dizemos que X tem distribuição estável se para todo $n \geq 1$ existem constantes $a_n > 0$ e b_n tais que*

$$F\left(\frac{x-b_n}{a_n}\right) = \underbrace{F_n(x) * \cdots * F_n(x)}_{n-fatores}.$$

Proposição 4.3.1. *Seja X uma variável aleatória cuja função de distribuição é F e a função característica é φ. Então as seguintes afirmações são equivalentes:*

1. *X tem distribuição estável;*
2. *Para todo $n \geq 1$ existem variáveis aleatórias $X_1, \cdots, X_n$ independentes e identicamente distribuídas a X, e constante $a_n > 0$ e b_n tais que $a_n X + b_n \stackrel{D}{=} X_1 + \cdots + X_n$;*
3. *Para todo $n \geq 1$ existem constantes $a_n > 0$ e b_n tais que $(\varphi(t))^n = \varphi(a_n t)e^{itb_n}$.*

Assim, fará sentido nos referirmos também à função característica φ como estável.

Demonstração. Para mostrar que a segunda afirmação implica a primeira e a terceira, basta observar que pelo Corolário 2.4.2, a função de distribuição e a função característica de $X_1 + \cdots + X_n$ são, respectivamente, $\underbrace{F_n(x) * \cdots * F_n(x)}_{n-fatores}$ e $(\varphi(t))^n$, ao mesmo tempo que a função de distribuição de $a_n X + b_n$ é $F_{a_n X + b_n}(x) = P(a_n X + b_n \leq x) = P\left(X \leq \frac{x-b_n}{a_n}\right) = F\left(\frac{x-b_n}{a_n}\right)$, e pela Proposição 2.4.1, a função característica de $a_n X + b_n$ é $\varphi(a_n t + b_n) = \varphi(a_n t)e^{itb_n}$.

Para mostrar que a primeira afirmação implica a segunda, observe que se X tem distribuição estável, então temos que para todo $n \geq 1$ existem constantes $a_n > 0$ e b_n tais que $F\left(\frac{x-b_n}{a_n}\right) = \underbrace{F_n(x) * \cdots * F_n(x)}_{n-fatores}$. Por outro lado, pelo Teorema 2.2.1 existem variáveis aleatórias $X_1, \cdots, X_n$ independentes e identicamente distribuídas a X, e pelo Corolário

2.4.2 a função de distribuição de $X_1 + \cdots + X_n$ será $\underbrace{F_n(x) * \cdots * F_n(x)}_{n-fatores}$, e portanto, $a_n X + b_n \stackrel{D}{=} X_1 + \cdots + X_n$. A prova de que a terceira implica a segunda é análoga. □

Exemplo 4.3.1. *A distribuição normal com parâmetros μ e σ^2 é estável.*

Demonstração. Pelo Corolário 2.5.1 temos que a função característica correspondente é $\varphi(t) = \exp(it\mu - \frac{\sigma^2 t^2}{2})$. Assim, dado $n \geq 1$, temos

$$(\varphi(t))^n = \exp\left(it\mu - \frac{\sigma^2 t^2}{2}\right)^n = \varphi(\sqrt{n}t)e^{it(n-\sqrt{n})\mu} = \varphi(a_n t)e^{itb_n},$$

onde $a_n = \sqrt{n}$ e $b_n = (n - \sqrt{n})\mu$. Temos portanto que a distribuição normal é estável. □

Exemplo 4.3.2. *A distribuição de Cauchy com parâmetro $\theta > 0$ é estável.*

Demonstração. Pelo Corolário 2.5.3 temos que a função característica correspondente é $\varphi(t) = e^{\theta|t|}$. Assim, dado $n \geq 1$, temos

$$(\varphi(t))^n = e^{-\theta|nt|} = \varphi(nt) = \varphi(a_n t)e^{itb_n},$$

onde $a_n = n$ e $b_n = 0$. Temos portanto que a distribuição de Cauchy é estável. □

Exemplo 4.3.3. *A distribuição de Poisson com parâmetro $\lambda > 0$ não é estável.*

Demonstração. Pela Proposição 2.5.4 temos que a função característica correspondente é $\varphi(t) = e^{\lambda(e^{it}-1)}$. Assim, dado $n \geq 1$, temos $(\varphi(t))^n = e^{n\lambda(e^{it}-1)}$, ao mesmo tempo que $\varphi(a_n t)e^{itb_n} = e^{\lambda(e^{ita_n}-1)}e^{itb_n}$. Assim, se a distribuição de Poisson fosse estável teríamos $e^{n\lambda(e^{it}-1)} = e^{\lambda(e^{ita_n}-1)}e^{itb_n} = e^{\lambda(e^{ita_n}-1+\frac{itb_n}{\lambda})}$ onde $a_n > 0$ e b_n são constantes. Nesse caso teríamos

$$e^{ita_n} - 1 + \frac{itb_n}{\lambda} - ne^{it} + n = \cos(ta_n) + i\sin(ta_n) - 1 + \frac{itb_n}{\lambda} - n\cos(t) - ni\sin(t) + n = 2k\pi,$$

onde $k \in \mathbb{Z}$. Segue que $\sin(ta_n) - n\sin(t) + \frac{tb_n}{\lambda} = 0$, e consequentemente,

$$b_n = \frac{\lambda(n\sin(t) - \sin(ta_n))}{t}. \tag{4.5}$$

Daí, tomando $t = \frac{\pi}{a_n}$ e $t = \frac{2\pi}{a_n}$, respectivamente, teremos

$$\begin{aligned}
b_n = \frac{na_n\lambda\sin\left(\frac{\pi}{a_n}\right)}{\pi} = \frac{na_n\lambda\sin\left(\frac{2\pi}{a_n}\right)}{2\pi} \quad &\Rightarrow \quad \sin\left(\frac{2\pi}{a_n}\right) = 2\sin\left(\frac{\pi}{a_n}\right) \\
&\Rightarrow \quad 2\sin\left(\frac{\pi}{a_n}\right)\cos\left(\frac{\pi}{a_n}\right) = 2\sin\left(\frac{\pi}{a_n}\right) \\
&\Rightarrow \quad 2\sin\left(\frac{\pi}{a_n}\right)\left[\cos\left(\frac{\pi}{a_n}\right) - 1\right] = 0 \\
&\Rightarrow \quad \frac{1}{a_n} \in \mathbb{Z} \\
&\Rightarrow \quad b_n = 0.
\end{aligned}$$

Agora, tomando $t = \frac{\pi}{2}$ teremos por (4.5) que $\sin\left(\frac{\pi a_n}{2}\right) = n$, o que é impossível para $n > 1$. $\square$

Proposição 4.3.2. *Toda distribuição estável é infinitamente divisível.*

Demonstração. Se X é uma variável aleatória com distribuição estável, então, para todo $n \geq 1$ existem constantes $a_n > 0$ e b_n, e variáveis aleatórias $X_1, \cdots, X_n$ independentes e identicamente distribuídas a X tais que $a_n X + b_n \stackrel{D}{=} X_1 + \cdots + X_n$. Segue que

$$X \stackrel{D}{=} \frac{X_1 + \cdots + X_n - b_n}{a_n} = \frac{X_1 - \frac{b_n}{n}}{a_n} + \cdots + \frac{X_n - \frac{b_n}{n}}{a_n},$$

onde $\frac{X_1 - \frac{b_n}{n}}{a_n}, \cdots, \frac{X_n - \frac{b_n}{n}}{a_n}$ são independentes e identicamente distribuídas. Assim, a distribuição de X é infinitamente divisível. $\square$

Proposição 4.3.3. *Seja X uma variável aleatória discreta com imagem no conjunto finito $\{x_1, \cdots, x_k\}$. Então X tem distribuição estável se, e somente se, é degenerada num ponto $\lambda \in \mathbb{R}$.*

Demonstração. Se a distribuição de X não for degenerada, então pela Proposição 4.2.2 não será infinitamente divisível, e logo, pela Proposição 4.3.2 também não será estável. Reciprocamente, se a distribuição de X é degenerada num ponto $\lambda \in \mathbb{R}$, então a função característica de X será $\varphi(t) = e^{it\lambda}$. Logo, dado $n \geq 1$ temos

$$(\varphi(t))^n = e^{it\lambda n} = \varphi(nt) = \varphi(a_n t) e^{itb_n},$$

onde $a_n = n$ e $b_n = 0$, e portanto, a distribuição de X é estável. $\square$

Proposição 4.3.4. *Para todo $\alpha \in (0, 2]$, a função $\varphi(t) = e^{-|t|^\alpha}$, $t \in \mathbb{R}$, é uma função característica estável.*

Demonstração. Já vimos no Corolário 4.2.2 que a função característica $\varphi(t) = e^{-|t|^\alpha}$, $t \in \mathbb{R}$, é infinitamente divisível. Para verificar que ela é estável basta observar que dado $n \geq 1$ temos

$$(\varphi(t))^n = e^{-n|t|^\alpha} = e^{-|n^{\frac{1}{\alpha}} t|^\alpha} = \varphi(a_n t) e^{itb_n},$$

onde $a_n = n^{\frac{1}{\alpha}}$ e $b_n = 0$. Segue que φ é estável. $\square$

Lema 4.3.1. *Seja $X_1, X_2, \cdots$ uma sequência de variáveis aleatórias e tal que $X_n \xrightarrow{D} X$, e sejam $a_n > 0$ e b_n duas sequências de números reais tais que $a_n X_n + b_n \xrightarrow{D} \tilde{X}$. Se X e $\tilde{X}$ são não degeneradas, então existem $a > 0$ e $b \in \mathbb{R}$ tais que $a_n \xrightarrow{n\to\infty} a$, $b_n \xrightarrow{n\to\infty} b$ e $\tilde{X} = aX + b$.*

Demonstração. Sejam $\varphi(t)$, $\varphi_n(t)$, $\tilde{\varphi}(t)$ e $\varphi_{a_nX_n+b_n}(t)$ as funções características de X, X_n, $\tilde{X}$ e $a_nX_n+b_n$, respectivamente. Logo, como $a_nX_n+b_n \xrightarrow{D} \tilde{X}$ e $X_n \xrightarrow{D} X$, temos pelo Corolário 3.1.1 que

$$e^{itb_n}\varphi_n(a_nt) = \varphi_{a_nX_n+b_n}(t) \xrightarrow{n\to\infty} \tilde{\varphi}(t), \quad \text{e} \tag{4.6}$$

$$\varphi_n(t) \xrightarrow{n\to\infty} \varphi(t), \tag{4.7}$$

onde, pelo Teorema 3.1.3, a convergência é uniforme em qualquer intervalo compacto. Seja agora $\{a_{n_k}\}_{k\geq 1}$ uma subsequência de $\{a_n\}_{n\geq 1}$ tal que $a_{n_k} \xrightarrow{k\to\infty} a$. Afirmamos que $a < \infty$ e a é único. Logo, todas as subsequências de $\{a_n\}_{n\geq 1}$ convergem para o mesmo valor $a < \infty$, e assim, $a_n \xrightarrow{n\to\infty} a$.

Para provar que $a < \infty$, suponha por absurdo que $a = \infty$. Logo, por (4.6) temos

$$|\varphi_n(a_nt)| = |e^{itb_n}\varphi_n(a_nt)| \xrightarrow{n\to\infty} |\tilde{\varphi}(t)|,$$

onde a convergência é uniforme em cada intervalo compacto, e logo,

$$\sup_{|t|\leq c}\Big||\varphi_n(a_nt)| - |\tilde{\varphi}(t)|\Big| \xrightarrow{n\to\infty} 0,$$

para todo $c > 0$. Daí, tomando $t = \frac{t_0}{a_{n_k}}$, e usando o fato de que $a_{n_k} \xrightarrow{n\to\infty} \infty$, temos que

$$\left|\varphi_{n_k}\left(a_{n_k}\frac{t_0}{a_{n_k}}\right)\right| - \left|\tilde{\varphi}\left(\frac{t_0}{a_{n_k}}\right)\right| \xrightarrow{k\to\infty} 0,$$

donde temos que

$$|\varphi_{n_k}(t_0)| \xrightarrow{k\to\infty} |\tilde{\varphi}(0)| = 1,$$

para todo $t_0 \in \mathbb{R}$. Por outro lado, de (4.7) temos $|\varphi_n(t_0)| \xrightarrow{n\to\infty} |\varphi(t_0)|$, donde segue que $|\varphi(t_0)| = 1$ para todo $t_0 \in \mathbb{R}$, e consequentemente, pela Proposição 2.5.1, a distribuição de X é degenerada. Chegamos assim a uma contradição, e portanto $a < \infty$.

Para provar que a é único, considere as subsequências $\{a_{n_j}\}_{j\geq 1}$ e $\{a_{n_{j'}}\}_{j'\geq 1}$ tais que $a_{n_j} \xrightarrow{j\to\infty} a$ e $a_{n_{j'}} \xrightarrow{j'\to\infty} a'$, supondo sem perda de generalidade que $0 < a' < a$. Por (4.6) e (4.7) temos

$$|\varphi_{n_j}(a_{n_j}t)| \xrightarrow{j\to\infty} |\varphi(at)| \quad \text{e} \quad |\varphi_{n_j}(a_{n_j}t)| \xrightarrow{j\to\infty} |\tilde{\varphi}(t)|,$$

donde segue que $|\varphi(at)| = |\tilde{\varphi}(t)|$. De modo análogo temos também que $|\varphi(a't)| = |\tilde{\varphi}(t)|$, e portanto, $|\varphi(at)| = |\varphi(a't)|$, para todo $t \in \mathbb{R}$. Segue que

$$|\varphi(t)| = \left|\varphi\left(a\frac{t}{a}\right)\right| = \left|\varphi\left(\frac{a'}{a}t\right)\right| = \cdots = \left|\varphi\left(\left(\frac{a'}{a}\right)^n t\right)\right| \xrightarrow{n\to\infty} |\varphi(0)| = 1,$$

e logo, $|\varphi(t)| = 1$ para todo $t \in \mathbb{R}$. Novamente chegamos à contradição de que X é degenerada, e portanto a é único.

Agora vamos provar que existe $b < \infty$ tal que $b_n \stackrel{n\to\infty}{\longrightarrow} b$ e que $a > 0$. Por (4.7) temos que $\varphi_n(a_n t) \stackrel{n\to\infty}{\longrightarrow} \varphi(at)$, e logo, por (4.6) segue que

$$e^{itb_n} \stackrel{n\to\infty}{\longrightarrow} \frac{\tilde{\varphi}(t)}{\varphi(at)}, \tag{4.8}$$

para todo t tal que $\varphi(at) \neq 0$. Note que como φ é contínua em $t = 0$ com $\varphi(0) = 1$, existe $\delta > 0$ tal que $\varphi(at) \neq 0$ para $|t| < \delta$, e logo existe $\lim_{n\to\infty} e^{itb_n}$ para tais $t's$. Daí segue que a sequência $\{b_n\}_{n\geq 1}$ é limitada, e portanto possui uma subsequência convergente. Sejam então $\{b_{n_j}\}_{j\geq 1}$ e $\{b_{n_{j'}}\}_{j'\geq 1}$ subsequências tais que $b_{n_j} \stackrel{j\to\infty}{\longrightarrow} b$ e $b_{n_{j'}} \stackrel{j'\to\infty}{\longrightarrow} b'$. Por (4.8) temos para $|t| < \delta$ que $e^{itb} = \lim_{j\to\infty} e^{itb_{n_j}} = \lim_{j'\to\infty} e^{itb_{n_{j'}}} = e^{itb'}$, e consequentemente, $b = b'$. Assim temos que existe $b < \infty$ tal que $b_n \stackrel{n\to\infty}{\longrightarrow} b$. Para finalizar, por (4.6) temos $\tilde{\varphi}(t) = e^{itb}\varphi(at)$, donde segue que $\tilde{X} \stackrel{D}{=} aX + b$, e como $\tilde{X}$ é não degenerada, segue que $a > 0$. □

Teorema 4.3.1. *Sejam $X_1, X_2, \cdots$ variáveis aleatórias independente e identicamente distribuídas, e $S_n = X_1 + \cdots + X_n$. Então, existem constantes reais $a_n > 0$ e b_n, e uma distribuição T tal que $\frac{S_n + b_n}{a_n} \stackrel{D}{\longrightarrow} T$ se, e somente se, T tem distribuição estável.*

Demonstração. Seja T uma variável aleatória com distribuição estável. Logo, para todo $n \geq 1$ existem constantes $a_n > 0$ e b_n tais que $a_n T + b_n \stackrel{D}{=} S_n$, e logo, $T \stackrel{D}{=} \frac{S_n - b_n}{a_n}$. Obviamente $\frac{S_n - b_n}{a_n} \stackrel{D}{\longrightarrow} T$.

Reciprocamente, suponhamos que $\frac{S_n - b_n}{a_n} \stackrel{D}{\longrightarrow} T$. Precisamos mostrar que T tem distribuição estável. Para tanto, vamos supor que T não é degenerada, pois caso contrário, já sabemos da Proposição 4.3.3 que ela tem distribuição estável. Daí, dado $n \geq 1$ e $1 \leq k \leq n$, façamos as seguintes definições:

$$S_n^{(k)} = X_{(k-1)n+1} + \cdots + X_{kn} \quad \text{e,}$$

$$T_n^{(k)} = \frac{S_n^{(k)} - b_n}{a_n}.$$

Note que para todo $k = 1, \cdots, n$, temos $S_n^{(k)} \stackrel{D}{=} S_n$, e logo,

$$T_n^{(k)} = \frac{S_n^{(k)} - b_n}{a_n} \stackrel{D}{=} \frac{S_n - b_n}{a_n} \stackrel{D}{\longrightarrow} T.$$

Agora, defina $U_n^{(k)}$ por

$$U_n^{(k)} = T_n^{(1)} + \cdots + T_n^{(k)}.$$

Imediatamente temos que $U_n^{(k)} \stackrel{D}{\longrightarrow} T^{(1)} + \cdots + T^{(k)}$, onde $T^{(1)} \stackrel{D}{=} \cdots \stackrel{D}{=} T^{(k)} \stackrel{D}{=} T$. Por outro lado, temos

$$U_n^{(k)} = T_n^{(1)} + \cdots + T_n^{(k)} = \frac{X_1 + \cdots + X_{kn} - kb_n}{a_n} = \alpha_n^{(k)} V_{kn} + \beta_n^{(k)},$$

onde $\alpha_n^{(k)} = \frac{a_{kn}}{a_n}$, $V_{kn} = \frac{X_1 + \cdots + X_{kn} - b_{kn}}{a_{kn}} = \frac{S_{kn} - b_{kn}}{a_{kn}}$ e $\beta_n^{(k)} = \frac{b_{kn} - kb_n}{a_n}$, e consequentemente,

$$V_{kn} = \frac{U_n^{(k)} - \beta_n^{(k)}}{\alpha_n^{(k)}},$$

onde $V_{kn} \xrightarrow{D} T$, $U_n^{(k)} \xrightarrow{D} T^{(1)} + \cdots + T^{(k)}$, e pelo Lema 4.3.1 existem $\alpha^{(k)} > 0$ e $\beta^{(k)}$ tais que $\alpha_n^{(k)} \overset{n\to\infty}{\longrightarrow} \alpha^{(k)}$ e $\beta_n^{(k)} \overset{n\to\infty}{\longrightarrow} \beta^{(k)}$. Segue que $T \overset{D}{=} \frac{T^{(1)}+\cdots+T^{(k)}-\beta^{(k)}}{\alpha^{(k)}}$, e consequentemente, $\alpha^{(k)}T + \beta^{(k)} \overset{D}{=} T^{(1)} + \cdots + T^{(k)}$. Como o resultado vale para qualquer $k \geq 1$, temos que a distribuição de T é estável. □

Da mesma forma que as funções características infinitamente divisíveis, as estáveis também possuem uma representação canônica determinada por Lévy-Khintchin.

Teorema 4.3.2 (Representação Canônica de Lévy-Khintchin)**.** *Toda função característica estável pode ser representada canonicamente por*

$$\varphi(t) = \exp\left[it\beta - d|t|^\alpha \left(1 + i\theta \frac{t}{|t|} G(t,\alpha)\right)\right], \tag{4.9}$$

onde $0 < \alpha \leq 2$, $\beta \in \mathbb{R}$, $d \geq 0$, $|\theta| \leq 1$, $\frac{t}{|t|} = 0$ *se* $t = 0$ *e*

$$G(t,\alpha) = \begin{cases} \tan(\frac{\pi\alpha}{2}), & se\ \alpha \neq 1 \\ \frac{2}{\pi}\ln|t|, & se\ \alpha = 1 \end{cases}.$$

Uma prova deste resultado pode ser visto em (GNEDENKO; KOLMOGOROV, 1954, Capítulo 7). Observe que as funções da Proposição 4.3.4 resultam do caso particular $\beta = 0$, $\theta = 0$ e $d = 1$. Uma consequência interessante da representação de Lévy-Khintchin é que ela permite calcular a média e a variância das variáveis aleatórias com distribuição estável.

Proposição 4.3.5. *Sendo X uma variável com distribuição estável e função característica dada em (4.9), temos que X não é integrável para $0 < \alpha \leq 1$, $\mathbb{E}X = \beta$ para $1 < \alpha \leq 2$, $VarX = 2d$ para $\alpha = 2$ e $VarX = \infty$ para $0 < \alpha < 2$.*

Demonstração. Sendo $\alpha \neq 1$, temos da representação de Lévy-Khintchin (Teorema 4.3.2) que para $t > 0$,

$$\varphi'(t) = \left[i\beta - \alpha d\left(1 + i\theta\tan\frac{\pi\alpha}{2}\right)t^{\alpha-1}\right]\varphi(t), \quad \text{e},$$

$$\varphi''(t) = \left[-\alpha(\alpha-1)d\left(1 + i\theta\tan\frac{\pi\alpha}{2}\right)t^{\alpha-2} + [i\beta - \alpha d\left(1 + i\theta\tan\frac{\pi\alpha}{2}\right)t^{\alpha-1}]^2\right]\varphi(t),$$

enquanto que para $t < 0$,

$$\varphi'(t) = \left[i\beta + \alpha d\left(1 - i\theta\tan\frac{\pi\alpha}{2}\right)(-t)^{\alpha-1}\right]\varphi(t), \quad \text{e},$$

$$\varphi''(t) = \left[-\alpha(\alpha-1)d\left(1 - i\theta\tan\frac{\pi\alpha}{2}\right)(-t)^{\alpha-2} + [i\beta + \alpha d\left(1 - i\theta\tan\frac{\pi\alpha}{2}\right)(-t)^{\alpha-1}]^2\right]\varphi(t).$$

A partir daí, para $0 < \alpha < 1$ temos $\lim_{t\to 0}|\varphi'(t)| = \infty$ e $\lim_{t\to 0}|\varphi''(t)| = \infty$, e logo, pelo Teorema 2.4.2 segue que X não é integrável e que $VarX = \infty$.

Para $1 < \alpha < 2$ temos $\varphi'(0) = i\beta$ e $\lim_{t\to 0} |\varphi''(t)| = \infty$, e logo, pelo Teorema 2.4.2 e pelo Corolário 2.4.4 segue que $\mathbb{E}X = \beta$ e $VarX = \infty$.

Agora, para $\alpha = 1$, temos

$$\varphi'(t) = \begin{cases} \left[i\beta - d - \frac{2id\theta}{\pi}\ln(t) - \frac{2id\theta}{\pi}\right], & \text{se } t > 0 \\ \left[i\beta + d - \frac{2id\theta}{\pi}\ln(-t) - \frac{2id\theta}{\pi}\right], & \text{se } t < 0 \end{cases}.$$

Se $\theta \neq 0$, então $\lim_{t\to 0} |\varphi'(t)| = \infty$, e se $\theta = 0$, então $\lim_{t\to 0^+} \varphi'(t) = i - d \neq i + d = \lim_{t\to 0^-} \varphi'(t)$, para $d \neq 0$. Logo X não é integrável e $VarX = \infty$.

Para $\alpha = 2$ temos $\varphi(t) = e^{it\beta - dt^2}$, e logo, pelo Corolário 2.5.1, $\mathbb{E}X = \beta$ e $VarX = 2d$. $\square$

5 Considerações Finais

Para finalizar esta dissertação, vamos neste capítulo fazer algumas considerações sobre a teoria descrita nos capítulos anteriores, em especial sobre os Teoremas 4.2.1 e 4.3.1. Salvo as citações que são feitas, as ideias aqui expressas são fruto do conhecimento adquirido durante o processo de elaboração da dissertação.

5.1 Diferenças entre os Teoremas 4.2.1 e 4.3.1

Sendo $a_n > 0$ e b_n constantes reais, e S_n a soma das variáveis aleatórias independentes e identicamente distribuídas $X_1, \cdots, X_n$, temos pelo Teorema 4.3.1 que os possíveis limites da sequência $\frac{S_n+b_n}{a_n}$ têm distribuição estável. Por outro lado, pela equação (4.2) temos que $\frac{S_n+b_n}{a_n}$ pode ser escrito como arranjo triangular, e logo, pelo Teorema 4.2.1 o limite, caso exista, tem distribuição infinitamente divisível. A diferença entre tais teoremas é que o Teorema 4.3.1 trata especificamente das sequências da forma $\frac{S_n+b_n}{a_n}$, enquanto que o Teorema 4.2.1 analisa o comportamento dos arranjos triangulares. Como já dissemos, $\frac{S_n+b_n}{a_n}$ sempre pode ser escrito como arranjo triangular, e portanto tem distribuição infinitamente divisível e estável, mas se tentarmos escrever um arranjo triangular na forma $\frac{S_n+b_n}{a_n}$, nem sempre conseguiremos fazê-lo de forma que S_n seja a soma de variáveis aleatórias independentes e identicamente distribuídas, e assim, o limite de um arranjo triangular, caso exista, será infinitamente divisível, mas não necessariamente estável. Um exemplo nesse sentido é o Teorema de Poisson (Teorema 3.2.1), onde temos um arranjo triangular convergindo para uma variável aleatória com distribuição de Poisson, a qual é infinitamente divisível, mas não é estável, conforme os Exemplos 4.2.2 e 4.3.3.

5.2 Consequências do Teorema 4.3.1

Na Seção 4.3 estudamos a classe das distribuições estáveis, e como motivação para tal abordagem usamos a Lei dos Grandes Números de Khintchin (Teorema 3.2.2) e o Exemplo 3.2.2, a partir dos quais obtemos que dada a sequência $\frac{S_n}{n}$, onde S_n é a soma das variáveis aleatórias independentes e identicamente distribuídas $X_1, X_2, \cdots$, a classe dos possíveis limites de $\frac{S_n}{n}$ contém pelo menos as distribuições degeneradas e de Cauchy. Posteriormente, no Teorema 4.3.1, vimos que se $\frac{S_n+b_n}{a_n} \xrightarrow{D} T$, onde $a_n > 0$ e b_n são constantes reais, então T tem distribuição estável. Mas no caso da sequência $\frac{S_n}{n}$ podemos restringir um pouco mais a classe dos possíveis limites. Observe que como as variáveis aleatórias são identicamente distribuídas, temos três possibilidades:

i. $\mathbb{E}X_i = \mu < \infty$,

ii. $\mathbb{E}X_i = \infty$,

iii. $\mathbb{E}X_i$ não está definida.

No caso de termos $\mathbb{E}X_i = \mu < \infty$, sabemos do Teorema 3.2.2 que o limite de $\frac{S_n}{n}$ tem distribuição degenerada em μ. Quando $\mathbb{E}X_i = \infty$, temos de (JAMES, 2006, Capítulo 5, Teorema 5.3) que $\frac{S_n}{n}$ diverge. Se $\mathbb{E}X_i$ não estiver definida, e $\frac{S_n}{n} \xrightarrow{D} T$, então $\mathbb{E}T$ também não estará definida. De fato basta observar que para todo $n \in \mathbb{N}$,

$$\mathbb{E}\left(\frac{S_n}{n}\right) = \frac{1}{n}\mathbb{E}(X_1 + \cdots + X_n) = \frac{1}{n}n\mathbb{E}X_i = \mathbb{E}X_i$$

Assim, a classe dos possíveis limites da sequência $\frac{S_n}{n}$ se reduz às distribuições estáveis degeneradas ou com esperança indefinida, excluindo-se assim o caso em que $1 < \alpha < 2$ na representação de Lévy-Khintchin do Teorema 4.3.2, já que neste caso temos $\mathbb{E}X_i = \beta < \infty$ pela Proposição 4.3.5.

Um raciocínio semelhante pode ser feito a respeito do Teorema Central do Limite para variáveis aleatórias independentes e identicamente distribuídas (Teorema 3.2.3), segundo o qual $\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}} = \frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow{D} N(0,1)$ se $\mathbb{E}X_i = m < \infty$ e $0 < VarX_i = \sigma^2 < \infty$. Observe que $\frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow{D} N(0,1)$ equivale a $\frac{S_n - nm}{\sqrt{n}} \xrightarrow{D} N(0,\sigma^2)$, e daí não precisamos da hipótese $0 < VarX_i = \sigma^2 < \infty$ para que a sequência exista. Nesse caso, temos três possibilidades a respeito da variância das variáveis aleatórias:

i. $0 < VarX_i = \sigma^2 < \infty$,

ii. $VarX_i = 0$,

iii. $VarX_i = \infty$.

No caso de termos $0 < VarX_i = \sigma^2 < \infty$, temos do Teorema 3.2.3 que $\frac{S_n - nm}{\sqrt{n}} \xrightarrow{D} N(0,\sigma^2)$. Se admitirmos $VarX_i = 0$ (o que ocorre se, e só se, as variáveis aleatórias têm distribuição degenerada), temos trivialmente que $\frac{S_n - nm}{\sqrt{n}} \xrightarrow{D} 0$. No último caso, se $VarX_i = \infty$, obtemos de modo análogo ao raciocínio desenvolvido no parágrafo anterior que se $\frac{S_n - nm}{\sqrt{n}} \xrightarrow{D} T$, então T tem distribuição estável com média zero e variância infinita. Assim, pela representação de Lévy-Khintchin do Teorema 4.3.2 e pela Proposição 4.3.5, se $\frac{S_n - nm}{\sqrt{n}} \xrightarrow{D} T$, a função característica de T terá a forma

$$\varphi(t) = \exp\left[-d|t|^{\alpha}\left(1 + i\theta\frac{t}{|t|}\tan\frac{\pi\alpha}{2}\right)\right],$$

onde $1 < \alpha \leq 2$, $d \geq 0$, $|\theta| \leq 1$ e $\frac{t}{|t|} = 0$ se $t = 0$.

5.3 Variáveis Aleatórias com Distribuições não Idênticas

Em vez de alterarmos as hipóteses sobre a esperança e a variância das variáveis aleatórias como fizemos acima, podemos retirar a condição delas serem identicamente distribuídas e analisar os possíveis limites das sequências do tipo $\frac{S_n+b_n}{a_n}$. Se considerarmos, por exemplo, que na sequência $X_1, X_2, \cdots$ temos $\mathbb{E}X_1 = \mu < \infty$, $0 < VarX_1 < \infty$ e que todas as outras variáveis aleatórias são degeneradas, temos que para todo $n \in \mathbb{N}$, $\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}} = \frac{X_1 - \mathbb{E}X_1}{\sqrt{VarX_1}}$, e logo, $\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}}$ converge em distribuição para $\frac{X_1 - \mathbb{E}X_1}{\sqrt{VarX_1}}$. Assim, qualquer variável aleatória com média igual a zero e variância igual a um é um possível limite das somas normalizadas $\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}}$. Uma condição que podemos impor para restringir a classe dos possíveis limites é a seguinte:

Definição 5.3.1 (Uniformidade Infinitesimal)**.** *Considere o arranjo triangular*

$$
\begin{aligned}
T_1 &= X_{1,1} \\
T_2 &= X_{2,1} + X_{2,2} \\
T_3 &= X_{3,1} + X_{3,2} + X_{3,3} \\
&\vdots
\end{aligned}
$$

Dizemos que a sequência $(X_{n,k})$ é uniformemente infinitesimal, se para todo $\delta > 0$ tivermos

$$\lim_{n \to \infty} \sup_{1 \le k \le n} P[|X_{n,k}| \ge \delta] = 0.$$

No Teorema 4.2.1 vimos que caso as variáveis aleatórias $X_{n,1}, \cdots, X_{n,n}$ sejam sempre identicamente distribuídas, então o limite da sequência T_n, caso exista, é infinitamente divisível. Ao mesmo tempo, o exemplo acima deixa claro que se tomarmos uma sequência arbitrária, não há restrição quanto aos limites que podem aparecer. No entanto, se a sequência for uniformemente infinitesimal, temos por (VARADHAN, 2001, Capítulo 3, Teorema 3.19) ou (BREIMAN, 1992, Capítulo 9, Teorema 9.19) que os possíveis limites da sequência $T_n + A_n$, onde $A_1, A_2, \cdots$ são constantes reais, continuam sendo infinitamente divisíveis, mesmo que as variáveis aleatórias não sejam identicamente distribuídas.

Quanto às sequências da forma $\frac{S_n+b_n}{a_n}$, recorde que por (4.2) podemos escrevê-la como arranjo triangular pondo $X_{n,k} = \frac{X_k}{a_n} + \frac{b_n}{na_n}$, $1 \le k \le n$. Assim, uma condição suficiente para que o limite da sequência seja infinitamente divisível é que tenhamos para todo $\delta > 0$,

$$\lim_{n \to \infty} \sup_{1 \le k \le n} P\left[\left|\frac{X_k}{a_n} - \frac{b_n}{na_n}\right| \ge \delta\right] = 0.$$

Vale ressaltar que o objetivo em (4.2) é deixar as variáveis aleatórias $X_{n,1}, \cdots, X_{n,n}$ identicamente distribuídas. Mas uma vez que tiramos esta exigência, podemos tomar $X_{n,k} = \frac{X_k}{a_n}$ e $A_n = \frac{b_n}{a_n}$, e logo, para que o limite da sequência $\frac{S_n+b_n}{a_n}$ seja infinitamente divisível é suficiente que para todo $\delta > 0$ tenhamos

$$\lim_{n \to \infty} \sup_{1 \le k \le n} P\left[\left|\frac{X_k}{a_n}\right| \ge \delta\right] = 0.$$

Assim, mesmo que as sequências $\frac{S_n - \mathbb{E}S_n}{n}$ e $\frac{S_n - \mathbb{E}S_n}{\sqrt{VarS_n}}$ não satisfaçam a Lei dos Grandes Números e o Teorema Central do Limite, respectivamente, temos que os limites, caso existam, serão infinitamente divisíveis, desde que tenhamos para todo $\delta > 0$,

$$\lim_{n\to\infty} \sup_{1\leq k\leq n} P\left[\left|\frac{X_k}{n}\right| \geq \delta\right] = 0 \quad \text{e} \quad \lim_{n\to\infty} \sup_{1\leq k\leq n} P\left[\left|\frac{X_k}{\sqrt{VarX_n}}\right| \geq \delta\right] = 0.$$

Referências

BREIMAN, L. *Probability.* University of California: Addison-Wesley Publishing Company, 1992. ISBN 0898712963. Citado 2 vezes nas páginas 11 e 102.

CHUNG, K. L. *A Course in Probability Theory - 3rd ed.* San Diego, Cailfornia: Academic Press, 2001. ISBN 9780121741518. Citado 7 vezes nas páginas 12, 13, 14, 18, 23, 52 e 67.

GNEDENKO, B. V.; KOLMOGOROV, A. N. *Limit Distributions for Sums of Independent Random Variables.* Cambridge: Addison-Wesley Publishing Company, 1954. ISBN 9780201024203. Citado 3 vezes nas páginas 13, 92 e 98.

JAMES, B. R. *Probabilidade: um Curso em Nível Intermediário - 3ª. ed.* Rio de Janeiro: IMPA, 2006. ISBN 9788524401015. Citado 6 vezes nas páginas 12, 14, 44, 48, 58 e 101.

LIMA, E. L. *Curso de Análise Volume I - 12ª. ed.* Rio de Janeiro: IMPA, 2009. ISBN 9788524401183. Citado 3 vezes nas páginas 51, 54 e 64.

SHIRYAEV, A. N. *Probability - 2nd ed.* New York: Springer, 1996. ISBN 9780387945491. Citado 7 vezes nas páginas 12, 13, 17, 22, 44, 66 e 67.

SOARES, M. G. *Cálculo em uma Variável Complexa - 4ª. ed.* Rio de Janeiro: IMPA, 1999. ISBN 9788524401442. Citado na página 38.

VARADHAN, S. R. S. *Probability Theory.* New York: American Mathematical Society, 2001. ISBN 9780821828256. Citado na página 102.