Universidade Federal da Paraíba
Centro de Ciências Exatas e da Natureza
Programa de Pós-Graduação em Matemática
Mestrado em Matemática

Limites de Escala em Modelos de Armadilhas

Lucas Araújo Santos

João Pessoa - PB
Dezembro de 2015

Universidade Federal da Paraíba
Centro de Ciências Exatas e da Natureza
Programa de Pós-Graduação em Matemática
Mestrado em Matemática

Limites de Escala em Modelos de Armadilhas

por

Lucas Araújo Santos

sob a orientação do

Prof. Dr. Alexandre de Bustamante Simas

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática da Universidade Federal da Paraíba como requisito parcial para a obtenção do título de Mestre em Matemática.

João Pessoa - PB
Dezembro de 2015

S237l Santos, Lucas Araújo.

Limites de escala em modelos de armadilhas / Lucas Araújo Santos.- João Pessoa, 2015.

77f.

Orientador: Alexandre de Bustamante Simas

Dissertação (Mestrado) - UFPB/CCEN

1. Matemática. 2. Modelos de armadilhas. 3. Limite de escala. 4. Leis estáveis. 5. Processos de Lévy.

UFPB/BC CDU: 51(043)

Limites de Escala em Modelos de Armadilhas

por

Lucas Araújo Santos [1]

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática da Universidade Federal da Paraíba como requisito parcial para a obtenção do título de Mestre em Matemática.

Área de Concentração: Análise

Aprovada em 11 de Dezembro de 2015.

Banca Examinadora:

Prof. Dr. Alexandre de Bustamante Simas – UFPB
(Orientador)

Prof. Dr. Wagner Barreto de Souza – UFMG
(Examinador Externo)

Profa. Dra. Evelina Shamarova – UFPB
(Examinador Interno)

[1] O autor foi bolsista PICME da Capes durante a elaboração desta dissertação.

À minha família

Agradecimentos

Agradeço, primeiramente, a Deus, meu Pai, a quem confio minha existência. Que todos os frutos de minha vida honrem e glorifiquem o nome dEle.

À minha família, pelo apoio em todas as minhas realizações.

Aos professores do Departamento de Matemática da UFPB, em especial aos professores Carlos Bocker, pelo acompanhamento durante a graduação, e Alexandre Simas pelo acompanhamento e orientação no mestrado.

Aos meus amigos que me acompanharam nesta jornada.

À todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

Resumo

Seja $\mathcal{X} = \{\mathcal{X} \geq 0, \mathcal{X}_0 = 0\}$ um passeio aleatório de média zero β-estável sobre $\mathbb{Z}$ com taxas de saltos não homogêneas $\{\tau_i^{-1}, i \in \mathbb{Z}\}$, com $\beta \in (1, 2]$ e $\{\tau_i : i \in \mathbb{Z}\}$ uma família de variáveis aleatórias independentes com distribuição marginal comum na bacia de atração de uma lei α-estável com $\alpha \in (0, 2]$. Neste trabalho, obtemos resultados sobre o comportamento a longo prazo deste processo obtendo seu limite de escala. Para isso, faremos previamente um estudo sobre probabilidade em espaços métricos, mais especificamente sobre o espaço D das funções contínuas à direita com limite à esquerda. Também iremos expor alguns resultados que tratam de leis estáveis que estão relacionadas diretamente ao problema supracitado.

Palavras-chave: Modelos de armadilhas, Limite de escala, Leis estáveis, Processos de Lévy.

Abstract

Let $\mathcal{X} = \{\mathcal{X} \geq 0, \mathcal{X}_0 = 0\}$ be a mean zero β-stable random walk on $\mathbb{Z}$ with inhomogeneous jump rates $\{\tau_i^{-1}, i \in \mathbb{Z}\}$, with $\beta \in (1, 2]$ and $\{\tau_i : i \in \mathbb{Z}\}$ is a family of independent random walk variables with common marginal distribution in the basis of attraction of an α-stable law with $\alpha \in (0, 2]$. In this paper we derive results about the long time behavior of this process, we obtain the scaling limit. To this end, first we will approach probability on metric spaces, specifically treat the D space of the functions that are right-continuous and have left-hand limits. We will also expose some results dealing with stable laws that are directly related to the above problem.

Keywords: Trap models, Scaling limit, Stable laws, Lévy processes.

Sumário

Notações

A seguir, listamos algumas notações utilizadas neste trabalho.

• $[t]$ denota a parte inteira do número real t, isto é, $[t]$ é o maior número inteiro que não ultrapassa t;

• $f_1(x) \sim f_2(x)$ significa que $\lim_{x\to\infty} \frac{f_1(x)}{f_2(x)} = 1$;

• id denota a aplicação identidade;

• $X \stackrel{d}{=} Y$ significa que os elementos aleatórios X e Y possuem a mesma distribuição;

• $X \Rightarrow Y$ denota convergência fraca das distribuições correspondentes aos elementos aleatórios X e Y;

• A função $sgn : \mathbb{R} \to \mathbb{R}$ é definida da seguinte forma: $sgn(x) = 1$, se $x > 0$, $sgn(x) = -1$, se $x < 0$, e $sgn(x) = 0$, se $x = 0$;

• I_A denota a função indicadora do conjunto A. $I_A(x) = 1$, se $x \in A$ e $I_A(x) = 0$, se $x \notin A$;

• $\square$ denota o final de uma demonstração.

Introdução

Neste trabalho, com base no artigo [1] de W. Barreto-Souza e L.R.G. Fontes, estudamos o limite de escala de um modelo de armadilha. Grosseiramente falando, um modelo de armadilha é um passeio aleatório a tempo contínuo sobre algum grafo regular com taxas de transição aleatórias dadas em termos geralmente de variáveis aleatórias fortemente atadas, o ambiente aleatório. Os casos mais estudados na literatura matemática envolve um passeio aleatório esqueleto que é independente do ambiente aleatório, e o inverso das taxas de saltos dado por variáveis aleatórias $i.i.d$ fortemente atadas, vistas neste caso como armadilhas profundas. Nestes casos, o modelo de armadilha é, portanto, um passeio aleatório discreto com uma mudança de tempo.

Estamos interessados num modelo de armadilha em $\mathbb{Z}$, e iremos supor que o passeio esqueleto é de média zero, β-estável, com $\beta \in (1, 2]$. No caso em que $\alpha \in (0, 1)$, o limite de escala é um processo β-estável com uma mudança de tempo pelo inverso de outro processo, envolvendo o tempo local do processo β-estável e um processo α-estável independente subordinado; o processo resultante pode ser chamado de processo *quase-estável*.

Antes de abordarmos o limite de escala em nosso modelo de armadilha, feito no terceiro capítulo, introduzimos temas auxiliares que por si têm uma grande relevância na teoria de probabilidade. No primeiro capítulo explicamos um pouco da teoria das probabilidades em espaços métricos. Nosso principal interesse é estudar a convergência fraca de distribuições, que são medidas de probabilidade, e para isso desenvolvemos vários critérios para sua ocorrência. Outro tema importante neste contexto é a compacidade, caracterizado, via Teorema de Prohorov, sob certas hipóteses, por uma propriedade denominada *tightness*. A compacidade, em certo sentido, é importante para assegurar que, dada uma sequência de distribuições que converge fracamente para um determinado limite, tal limite seja de fato a distribuição de algum elemento aleatório. Ainda no primeiro capítulo, tratamos mais especificamente de convergência fraca e *tightness* no espaço métrico D das funções contínuas à direita com limite à esquerda munido da topologia de Skorohod.

No segundo capítulo apresentamos um teorema central do limite para leis estáveis que é válido até mesmo se tivermos média infinita. Além disso, introduzimos os processos de Lévy, a principal classe de processos de nosso interesse, e expomos algumas de suas principais propriedades como a propriedade de Markov e a propriedade forte de Markov.

Por fim, o apêndice é dedicado aos resultados mais elementares em probabilidade que decidimos, afim de obter uma leitura mais direta e objetiva, não deixá-los no texto.

Capítulo 1

Princípios de invariância

1.1 Probabilidades em espaços métricos

Nesta seção faremos uma breve revisão dos principais resultados que tratam de probabilidade em espaços métricos. Para um aprofundamento nesta teoria veja [2].

1.1.1 Convergência fraca em espaços métricos

Seja S um espaço métrico. Iremos estudar medidas de probabilidade sobre uma classe $\mathcal{S}$ de borelianos em S. Aqui $\mathcal{S}$ é a σ-álgebra gerada pelos conjuntos abertos $-$ a menor σ-álgebra contendo os conjuntos abertos $-$ e a medida de probabilidade P em $\mathcal{S}$ é não-negativa, contável aditiva com $P(S) = 1$.

Definição 1.1. Se as medidas de probabilidade P_n e P satisfazem

$$\int_S f \; dP_n \to \int_S f \; dP$$

para toda função real f em S limitada e contínua, dizemos que P_n **converge fracamente** para P e escrevemos $P_n \Rightarrow P$. Denotaremos por $C(S)$ a classe de tais funções f.

Teorema 1.1. *Toda medida de probabilidade sobre $(S, \mathcal{S})$ é regular, isto é, se $A \in \mathcal{S}$ e $\varepsilon > 0$, então existe um conjunto fechado F e um conjunto aberto G tais que $F \subset A \subset G$ e $P(G - F) < \varepsilon$*

O teorema acima implica que uma medida $P \in (S, \mathcal{S})$ é determinada pelos valores de $P(B)$ para os conjuntos fechados B. De fato, sejam $A \in \mathcal{S}$ e $Q \in (S, \mathcal{S})$ tal que $P(B) = Q(B)$ para todo conjunto fechado B. Dados $A \in \mathcal{S}$ e ε positivo, podemos obter um conjunto fechado F e um conjunto aberto G de forma que $F \subset A \subset G$,

$P(G - F) < \varepsilon$ e $Q(G - F) < \varepsilon$. Observe que

$$P(A) \leq P(G) = P(G - F) + P(F) = P(G - F) + Q(F) \leq \varepsilon + Q(A).$$

Como ε é arbitrário, segue que $P(A) \leq Q(A)$ e similarmente $Q(A) \leq P(A)$, portanto, $P(A) = Q(A)$

Teorema 1.2. *Se F é fechado e ε é positivo, existe uma função f em $C(S)$ tal que $f(x) = 1$, se $x \in F$, $f(x) = 0$, se $d(x, F) \geq \varepsilon$, e $0 \leq f(x) \leq 1$, para outro caso. Aqui $d(x, F)$ representa a distância de x ao conjunto F, isto é, $d(x, F) = \inf_{y \in F} d(x, y)$. A função f pode se obtida de forma que seja uniformemente contínua.*

Teorema 1.3. *Duas medidas de probabilidades P e Q em $(S, \mathcal{S})$ coincidem se*

$$\int f\ dP = \int f\ dQ \tag{1.1}$$

para toda $f \in C(S)$.

Demonstração. Seja F um conjunto fechado. Considere a função real φ dada por

$$\varphi(t) = \begin{cases} 1 & \text{se} \quad t \leq 0, \\ 1 - t & \text{se} \quad 0 \leq t \leq 1, \\ 0 & \text{se} \quad 1 \leq t. \end{cases}$$

Defina, para cada inteiro positivo u,

$$\varphi_u(t) = \varphi(ut) \tag{1.2}$$

e

$$f_u(x) = \varphi_u(d(x, F)).$$

Então $\{f_u\}$ é uma sequência não-crescente de elementos de $C(S)$ convergindo pontualmente para a função indicadora I_F do conjunto F. Pelo teorema da convergência dominada, $P(F) = \lim_u \int f_u dP$ e $Q(F) = \lim_u \int f_u dQ$, assim, se (1.1) vale para toda $f \in C(S)$, $P(F) = Q(F)$. Como P e Q coincidem nos conjuntos fechados, segue do teorema 1.1 que $P = Q$.

□

1.1.2 Tightness

Definição 1.2. Uma medida de probabilidade P em $(S, \mathcal{S})$ é ***tight*** se para cada ε positivo existe um conjunto compacto K tal que $P(K) > 1 - \varepsilon$.

Pelo teorema 1.1, P é *tight* se, e somente se, $P(A)$ é, para todo $A \in \mathcal{S}$, o supremo de $P(K)$ sobre os conjuntos compactos $K \subset A$.

Teorema 1.4. *Se S é separável e completo, então toda medida de probabilidade em $(S, \mathcal{S})$ é tight.*

1.2 Propriedades da convergência fraca

Observe que, como as integrais $\int f \, dP$ determinam P completamente (Teorema 1.3), a sequência $\{P_n\}$ não pode convergir fracamente para dois limites distintos ao mesmo tempo. Observe também que a convergência fraca depende somente da topologia de S, e não da métrica específica que a gera: Duas métricas que geram a mesma topologia dão origem à mesmas classes de $\mathcal{S}$ e $C(S)$, portanto, à mesma noção de convergência fraca.

Teorema Portemanteau

O teorema a seguir nos dá condições equivalentes para a convergência fraca.

Definição 1.3. Um conjunto $A \in \mathcal{S}$ cuja fronteira ∂A satisfaz $P(\partial A) = 0$ é dito um **conjunto P-contínuo**.

Teorema 1.5. *Sejam P_n, P medidas de probabilidade em $(S, \mathcal{S})$. As seguintes condições são equivalentes:*

(i) $P_n \Rightarrow P$

(ii) $\lim_{n\to\infty} \int f \, dP_n = \int f \, dP$, *para toda função f limitada e uniformemente contínua.*

(iii) $\limsup_{n\to\infty} P_n(F) \leq P(F)$, *para todo fechado F.*

(iv) $\liminf_{n\to\infty} P_n(G) \geq P(G)$, *para todo aberto G.*

(v) $\lim_{n\to\infty} P_n(A) = P(A)$, *para todo conjunto A P-contínuo.*

Outro Critério

As vezes é conveniente provar convergência fraca mostrando que $P_n(A) \to P(A)$ para alguma classe especial de conjuntos A.

Teorema 1.6. *Seja $\mathcal{U}$ uma subclasse de $\mathcal{S}$ tal que (i) $\mathcal{U}$ é fechado sob interseções finitas e (ii) cada conjunto aberto em S é uma união finita ou contável dos elementos de $\mathcal{U}$. Se $P_n(A) \to P(A)$ para todo $A \in \mathcal{U}$, então, $P_n \Rightarrow P$.*

Demonstração. Seja $A_i \in \mathcal{U}$ para $i = 1, \ldots, m$. Por hipótese, as interseções de tais elementos também estão em $\mathcal{U}$. Pela fórmula de inclusão-exclusão,

$$\begin{aligned} P_n\left(\bigcup_{i=1}^m A_i\right) &= \sum_i P_n(A_i) - \sum_{ij} P_n(A_i \cap A_j) + \sum_{ijk} P_n(A_i \cap A_j \cap A_k) - \cdots \\ &\to \sum_i P(A_i) - \sum_{ij} P(A_i \cap A_j) + \sum_{ijk} P(A_i \cap A_j \cap A_k) - \cdots \\ &= P\left(\bigcup_{i=1}^m A_i\right). \end{aligned}$$

Se G é aberto, então $G = \cup_i A_i$ para alguma sequência $\{A_i\}$ de elementos de $\mathcal{U}$. Dado $\varepsilon > 0$, escolha m tal que $P(\cup_{i \le m} A_i) > P(G) - \varepsilon$. Pela relação que acabamos de provar, $P(G) - \varepsilon < P(\cup_{i \le m} A_i) = \lim_n P_n(\cup_{i \le m} A_i) \le \liminf_n P_n(G)$. Como ε é arbitrário, vale o item (iv) do teorema anterior. □

Denotamos a bola aberta de raio ε em torno de x por $S(x, \varepsilon)$.

Corolário 1.7. Seja $\mathcal{U}$ uma classe de conjuntos tal que (i) $\mathcal{U}$ é fechado sob interseções finitas e (ii) para cada $x \in S$ e todo ε positivo existe um $A \in \mathcal{U}$ com $x \in int(A) \subset A \subset S(x, \varepsilon)$. Se S é separável e $P_n(A) \to P(A)$ para todo $A \in \mathcal{U}$, então $P_n \Rightarrow P$.

Demonstração. A condição (ii) implica que, para cada ponto x de um conjunto aberto G, $x \in int(A) \subset A \subset G$ para algum $A \in \mathcal{U}$. Como S é separável, existe em $\mathcal{U}$ uma sequência $\{A_i\}$ tal que $G \subset \cup_i int(A_i)$ e $A_i \subset G$, que implica que $G = \cup_i A_i$. Portanto, $\mathcal{U}$ satisfaz as hipóteses do teorema 1.6. □

Iremos ver mais uma condição para a convergência fraca. Uma sequência $\{x_n\}$ de números reais converge para um limite x se, e somente se, toda subsequência $\{x_{n'}\}$ possui uma subsequência $\{x_{n''}\}$ que converge para x. Para este fato é fácil deduzir um análogo para convergência fraca.

Teorema 1.8. *$P_n \Rightarrow P$ se, e somente se, toda subsequência $\{P_{n'}\}$ contém uma subsequência $\{P_{n''}\}$ tal que $\{P_{n''}\} \Rightarrow P$.*

1.3 Alguns Casos Especiais

O Espaço Euclideano

Seja $\mathbb{R}^k$ o espaço k-dimensional, o qual sempre consideraremos munido da métrica $d(x, y) = |x - y| = [\sum_{i=1}^k (x_i - y_i)^2]^{1/2}$. Iremos relembrar a relação entre a convergência fraca de medidas de probabilidade e a noção usual de convergência para as funções de distribuição correspondentes. Considere as medidas de probabilidade P_n, P e suas respectivas funções de distribuição F_n e F, para $n = 1, 2, \ldots$. Então, $P_n \Rightarrow P$ se, e somente se, $F_n(x) \to F(x)$ para todo ponto de continuidade x de F. Em outras palavras, os conjuntos $\{y : y \leq x\}$ formam uma classe determinante de convergência.

O Espaço $\mathbb{R}^\infty$

Os resultados obtidos para $\mathbb{R}^k$ são transferidos em todos os aspectos essenciais para o espaço $\mathbb{R}^\infty$ das sequências $x = (x_1, x_2, \ldots)$ de números reais munido da topologia produto. Essa topologia possui como vizinhanças básicas de um ponto x os conjuntos da forma

$$N_{k,\varepsilon}(x) = \{y : |x_i - y_i| < \varepsilon,\ i = 1, \ldots, k\} \tag{1.3}$$

com $\varepsilon > 0$ e $k = 1, 2, \ldots$. Com esta topologia, $\mathbb{R}^\infty$ é um espaço métrico completo e separável.

Seja π_k a projeção natural de $\mathbb{R}^\infty$ para $\mathbb{R}^k$, definida por $\pi_k(x) = (x_1, \ldots, x_k)$. Um conjunto finito dimensional, ou cilindro, é, por definição, um conjunto da forma $\pi_k^{-1}(H)$ com $k \geq 1$ e $H \in \mathcal{R}^k$. Como cada π_k é contínuo e, portanto, mensurável, os conjuntos finito-dimensionais pertencem à σ-álgebra $\mathcal{R}^\infty$ de borelianos em $\mathbb{R}^\infty$. Seja $\mathcal{F}$ a classe dos conjuntos finito-dimensionais. Como cada conjunto em (1.3) pertence à $\mathcal{F}$ e $\mathbb{R}^\infty$ é separável, $\mathcal{F}$ gera $\mathcal{R}^\infty$. Como $\mathcal{F}$ é uma álgebra, segue que $\mathcal{F}$ é uma classe determinante.

Para k e x fixos, os conjuntos em (1.3), para diferentes valores de ε, possuem fronteiras disjuntas ($\varepsilon < \delta$ implica $\overline{N_{k,\varepsilon}} \subset int(N_{k,\delta})$). Aplicando o corolário 1.7 do teorema 1.6 para a classe $\mathcal{U}$ de conjuntos P-contínuos em $\mathcal{F}$, vemos que $\mathcal{F}$ é uma classe determinante de convergência. Assim $P_n \Rightarrow P$ se, e somente se $P_n(A) \to P(A)$ ocorre para todos os conjuntos A finito-dimensionais P-contínuos.

O Espaço C

Em $\mathbf{C} = \mathbf{C}[0, 1]$, o espaço das funções contínuas em $[0, 1]$ com a métrica uniforme $d(x, y) = \sup_t |x(t) - y(t)|$, a situação difere de $\mathbb{R}^\infty$. Para os pontos $t_1, \ldots, t_k$ em $[0, 1]$, seja $\pi_{t_1 \cdots t_k}$ a função que leva o ponto x de $\mathbf{C}$ ao ponto $(x(t_1), \ldots, x(t_k))$ de $\mathbb{R}^k$. Os conjuntos finito-dimensionais são agora definidos como conjuntos da forma $\pi_{t_1 \cdots t_k}^{-1}(H)$ com $H \in \mathcal{R}^k$. Como $\pi_{t_1 \cdots t_k}$ é contínuo, estes conjuntos pertencem à classe $\mathcal{C}$ de borelianos em $\mathbf{C}$. Por outro lado, a bola fechada $\{y : d(x, y) \leq \varepsilon\}$ é o limite dos conjuntos finito-dimensionais $\{y : |x(i/n) - y(i/n)| \leq \varepsilon,\ i = 1, \ldots, n\}$; como $\mathbf{C}$ é separável, cada conjunto aberto é uma união contável de bolas abertas e, portanto, de bolas fechadas, assim os conjuntos finito-dimensionais geram $\mathcal{C}$. Como eles formam uma álgebra, os conjuntos finito-dimensionais são uma classe determinante.

1.4 Convergência em Distribuição

A teoria da convergência fraca pode ser parafraseada como a teoria da convergência em distribuição. Quando estabelecemos a terminologia da última teoria, que não envolve novas ideias, muitos resultados assumem uma forma compacta e tornam-se mais claros.

Elementos Aleatórios

Definição 1.4. Seja X uma função de um espaço de probabilidade $(\Omega, \mathcal{B}, \mathbf{P})$ num espaço métrico S. Se X é mensurável, o chamaremos de **elemento aleatório** de S. Se $S = \mathbb{R}$, chamamos X de **variável aleatória**; se $S = \mathbb{R}^k$, chamamos X de **vetor aleatório**; se $S = C$, chamamos X de **função aleatória**.

A distribuição de X é a medida de probabilidade $P = \mathbf{P}X^{-1}$ sobre $(S, \mathcal{S})$:

$$P(A) = \mathbf{P}(X^{-1}A) = \mathbf{P}\{\omega : X(\omega) \in A\} = \mathbf{P}\{X \in A\}, A \in \mathcal{S}. \tag{1.4}$$

No caso $S = \mathbb{R}^k$, temos também a **função de distribuição** associada à X, definida por

$$F(x) = P\{y : y \leq x\} = P\{X \leq x\}, x \in \mathbb{R}^k.$$

Note que $\mathbf{P}$ é uma medida de probabilidade sobre um espaço arbitrário, enquanto P é definido sempre sobre um espaço métrico. Por várias razões, a distribuição P contém todas as informações relevantes sobre o elemento aleatório X. Se h é uma função mensurável de $\mathbb{R}$ em S, então, pela fórmula de mudança de variável (teorema A.8),

$$\int h(X)d\mathbf{P} = \int hdP.$$

Cada medida de probabilidade num espaço métrico é a distribuição de algum elemento aleatório sobre algum espaço de probabilidade: Dado P em $(S, \mathcal{S})$, se considerarmos $(\Omega, \mathcal{B}, \mathbf{P}) = (S, \mathcal{S}, P)$, com X sendo a identidade,

$$X(\omega) = \omega,\ \omega \in \Omega = S,$$

então X é um elemento aleatório sobre Ω com valores em S e possui P como distribuição. Embora a classe de distribuição coincida com a classe de medidas de probabilidade sobre espaços métricos, geralmente chamamos uma medida sobre um espaço métrico de distribuição apenas quando ela é de fato a distribuição de algum elemento aleatório em discussão.

Convergência em Distribuição

Definição 1.5. Dizemos que uma sequência $\{X_n\}$ de elementos aleatórios **converge em distribuição** para o elemento aleatório X, e escrevemos

$$X_n \xrightarrow{\mathcal{D}} X,$$

se as distribuições P_n de X_n converge fracamente para a distribuição P de X:

$$P_n \Rightarrow P.$$

É claro que para ter sentido, o contradomínio S e sua topologia são os mesmos para todos os elementos aleatórios $X, X_1, X_2, \ldots$. Como $\int_S f(x)P(dx) = \int_\Omega f(X)d\mathbf{P}$, pela fórmula de mudança de variável, e analogamente para $\int f dP_n$, temos $X_n \xrightarrow{\mathcal{D}} X$ se, e somente se, $E\{f(X_n)\} \to E\{f(X)\}$ para toda $f \in C(S)$.

1.4.1 Convergência em Probabilidade

Muitos conceitos padrões e resultados para convergência em distribuição de variáveis aleatórias se generalizam para elementos aleatórios.

Definição 1.6. Se, para um elemento $a \in S$,

$$\mathbf{P}\{d(X_n, a) \geq \varepsilon\} \to 0$$

para cada ε positivo, dizemos que X_n **converge em probabilidade** para a e escrevemos

$$X_n \xrightarrow{\mathcal{P}} a.$$

Se a é considerado como um elemento aleatório constante, então, $X_n \xrightarrow{\mathcal{P}} a$ se, e somente se, $X_n \xrightarrow{\mathcal{D}} a$. Alternativamente, $X_n \xrightarrow{\mathcal{P}} a$ se, e somente se, a distribuição de X_n converge fracamente para a medida de probabilidade que atribui massa um ao ponto a.

Se X_n e Y_n possuem o mesmo domínio, faz sentido falar da distância $d(X_n, Y_n)-$ a função com valor $d(X_n(\omega), Y_n(\omega))$ em ω. Se S é separável, $d(X_n, Y_n)$ é uma variável aleatória. No teorema a seguir, assumimos que, para cada n, X_n e Y_n possuem o mesmo domínio e S é separável.

Teorema 1.9. *Se $X_n \xrightarrow{\mathcal{D}} X$ e $d(X_n, Y_n) \xrightarrow{\mathcal{P}} 0$, então $Y_n \xrightarrow{\mathcal{D}} X$.*

Demonstração. Se $F_\varepsilon = \{x : d(x, F) \leq \varepsilon\}$, então

$$\mathbf{P}\{Y_n \in F\} \leq \mathbf{P}\{d(X_n, Y_n) \geq \varepsilon\} + \mathbf{P}\{X_n \in F_\varepsilon\}.$$

Como F_ε é fechado a hipótese implica

$$\limsup_n \mathbf{P}\{Y_n \in F\} \leq \limsup_n \mathbf{P}\{X_n \in F_\varepsilon\} \leq \mathbf{P}\{X \in F_\varepsilon\}.$$

Se F é fechado, então $F_\varepsilon \downarrow F$ quando $\varepsilon \downarrow 0$ e o resultado segue do teorema 1.5. $\square$

1.5 Teorema de Prohorov

Compacidade Relativa

Definição 1.7. Seja Π uma família de medidas de probabilidade em $(S, \mathcal{S})$. Dizemos que Π é **relativamente compacta** se toda sequência de elementos de Π contém uma subsequência fracamente convergente, isto é, se para toda sequência $\{P_n\}$ em Π, existe uma subsequência $\{P_{n'}\}$ e uma medida de probabilidade Q (definida em $(S, \mathcal{S})$, mas não necessariamente um elemento de Π) tal que $P_{n'} \Rightarrow Q$.

Uma família Π de medidas de probabilidade sobre um espaço métrico S é dita *tight* se para cada ε positivo existe um conjunto compacto K tal que $P(K) > 1 - \varepsilon$, para todo $P \in \Pi$.

Teorema 1.10. *Se Π é tight, então Π é relativamente compacto.*

Teorema 1.11. *Suponha que S é separável e completo. Se Π é relativamente compacto, então Π é tight.*

Nos referimos aos dois teoremas acima conjuntamente como o Teorema de Prohorov.

Observação 1.1. A compacidade relativa é importante, dentre outros motivos, para assegurar que o limite fraco Q de uma sequência de distribuições seja, de fato, uma distribuição, no sentido que não ocorre $Q(S) < 1$. Vejamos um exemplo em $(\mathbb{R}, \mathcal{R})$.

Seja $a + b + c = 1$ e $F_n(x) = aI_{(x \geq n)} + bI_{(x \geq -n)} + cG(x)$. Onde G é uma função de distribuição. Assim, $F_n(x) \to F(x) = b + cG(x)$,

$$\lim_{x \downarrow -\infty} F(x) = b \quad \text{e} \quad \lim_{x \uparrow +\infty} F(x) = b + c = 1 - a.$$

Ou seja, um quantidade de massa a está escapando para $+\infty$ e uma quantidade de massa b para $-\infty$.

Isto não ocorre, se a sequência F_n é tight. De fato, suponha que para todo $\varepsilon > 0$, existe M_ε tal que $\limsup_{n\to\infty} 1 - F_n(M_\varepsilon) + F_n(-M_\varepsilon) \leq \varepsilon$ e que $F_n(x) \to F(x)$ para todo x ponto de continuidade de F. Sejam $r < M_\varepsilon$ e $s > M_\varepsilon$ pontos de continuidade de F. Como $F_n(r) \to F(r)$ e $F_n(s) \to F(s)$, temos

$$\begin{aligned} 1 - F(s) - F(r) &= \lim_{n\to\infty} 1 - F_n(s) - F_n(r) \\ &\leq \limsup_{n\to\infty} 1 - F_n(M_\varepsilon) + F_n(-M_\varepsilon) \leq \varepsilon. \end{aligned}$$

O último resultado implica que

$$\limsup_{x\to\infty} 1 - F(x) + F(-x) \leq \epsilon.$$

Como ε é arbitrário, F é uma função de distribuição.

Funções Características

Se P é uma medida de probabilidade sobre $(\mathbb{R}^k, \mathcal{R}^k)$, sua função característica p é definida como

$$p(t) = \int e^{it\cdot x} P(dx),\ \ t \in \mathbb{R}^k,$$

onde $t \cdot x = \sum_{u=1}^k t_u x_u$ denota o produto interno. Seja Q uma segunda medida de probabilidade sobre $(\mathbb{R}^k, \mathcal{R}^k)$, e seja q sua função característica. Provaremos o teorema de unicidade:

Teorema 1.12. *Se $p(t) = q(t)$ para todo t, então $P = Q$.*

Demonstração. Para cada t, $e^{it\cdot x}$ é, como uma função de x, um elemento de $C(\mathbb{R}^k)$ (ou sua parte real e sua parte imaginária são). Assim, no caso de $\mathbb{R}^k$, o teorema de unicidade refina o teorema 1.3, que afirma que P é determindo pelos valores de $\int f dP$ para $f \in C(S)$. Não iremos provar a unicidade obtendo uma formula de inversão explícita, usaremos o teorema de aproximação de Weierstrass.

Sabemos que os retângulos $(a, b]$ formam um classe determinante; e como tal retângulo é uma união crescente de retângulos fechados, é suficiente checar que P e Q coincidem para os conjuntos

$$A = [a_1, b_1] \times \cdots \times [a_k, b_k].$$

Pelo argumento do teorema 1.3, isto segue se, pra cada integral,

$$\int f dP = \int f dQ \tag{1.5}$$

quando f for da forma $f(x_1, \ldots, x_k) = f_1(x_1) \cdots f_k(x_k)$ com $f_j(s) = \varphi_u(d(s, [a_j, b_j]))$, onde d denota a distância linear e φ_u é definida como em (1.2).

Fixado u. Dado ε, escolha r suficientemente grande tal que f_j se anula fora do intervalo $[-r,r]$ e, ao mesmo tempo, se I_r denota o cubo $\{x : |x_m| \leq r, m = 1,\dots,k\}$, então $P(I_r^C) < \varepsilon$ e $Q(I_r^C) < \varepsilon$. Como $f_j(-r) = f_j(r), f_j(s)$ podem, pelo teorema de Weierstraiss, serem uniformemente aproximados em $[-r,r]$ por uma soma trigonométrica finita $\sum_l \alpha_l e^{il\pi s/r}$ de período $2r$. Multiplicando em conjunto estas somas para diferentes valores de j, vemos que f pode ser uniformemente aproximada em I_r por uma soma trigonométrica finita

$$g(x) = \sum_l \gamma_l e^{it^{(l)}\cdot x} \tag{1.6}$$

de período $2r$ em cada variável. Escolha (1.6) de forma que $|f(x) - g(x)| < \varepsilon$ para $x \in I_r$. Como f é limitada por um, g é limitada por $1+\varepsilon$ em I_r e, portanto, por periodicidade, em todo $\mathbb{R}^k$. Assim, $|f-g|$ é limitado por ε em I_r e por $2+\varepsilon$ globalmente. Assumindo $0 < \varepsilon < 1$, temos

$$\int |f-g| dP \leq \varepsilon + (2+\varepsilon)P(I_r^C) < 4\varepsilon.$$

De modo similar, $\int |f-g|dQ < 4\varepsilon$, assim $|\int f dP - \int f dQ| < |\int g dP - \int g dQ| + 8\varepsilon$. Mas $\int g dP = \int g dQ$ é uma consequência imediata de $p = q$ e de (1.6), e como ε é arbitrário, (1.5) é satisfeita. □

Dispositivo de Cramer-Wold

Por meio do dispositivo a seguir, devido a Cramer e Wold, problemas envolvendo vetores aleatórios em $\mathbb{R}^k$ podem ser reduzids a problemas envolvendo apenas variáveis aleatórias em $\mathbb{R}$. Suponha que os vetores aleatórios $X_n = (X_{n1},\dots,X_{nk})$ e $X = (X_1,\dots,X_k)$ k-dimensionais satisfazem

$$\sum_{j=1}^k t_j X_{nj} \xrightarrow{\mathcal{D}} \sum_{j=1}^k t_j X_j$$

para cada ponto $t = (t_1,\dots,t_k)$ de $\mathbb{R}^k$. Então as funções características $p_n(s) = E\{\exp(is\sum_{j=1}^k t_j X_{nj})\}$ dessas variáveis aleatórias converge para $p(s) = E\{\exp(is\sum_{j=1}^k t_j X_j)\}$ para cada real s. Tomando $s = 1$, obtemos

$$E\{e^{it\cdot X_n}\} \to E\{e^{it\cdot X}\}.$$

Como t foi arbitrário, $X_n \xrightarrow{\mathcal{D}} X$ segue do teorema de continuidade para funções características.

1.6 O espaço D

Seja $D = D[0,1]$ o espaço das funções x em $[0,1]$ que são contínuas à direita e possuem limite à esquerda:

(i) Para $0 \leq t < 1$, $x(t^+) = \lim_{s\downarrow t} x(s)$ existe e $x(t^+) = x(t)$.

(ii) Para $0 < t \leq 1$, $x(t^-) = \lim_{s\uparrow t} x(s)$ existe.

Dizemos que a função x possui uma *descontinuidade do primeiro tipo* em t se $x(t^-)$ e $x(t^+)$ existem mas são diferentes e $x(t)$ está entre $x(t^-)$ e $x(t^+)$. Toda descontinuidade de um elemento de D são do primeiro tipo. É claro que C é um subconjunto de D.

Definição 1.8. O **módulo de continuidade** de um elemento x de C é definido por

$$w_x(\delta) = w(x,\delta) = \sup_{|s-t|<\delta} |x(s) - x(t)|, \ 0 < \delta \leq 1. \tag{1.7}$$

Para $x \in D$ e $T_0 \subset [0,1]$, ponha

$$w_x(T_0) = sup\{|x(s) - x(t)| : s,t \in T_0\}. \tag{1.8}$$

Assim, o módulo de continuidade de x, definido por (1.7), pode ser expresso como

$$w_x(\delta) = \sup w_x[t, t+\delta]. \tag{1.9}$$

Uma função continua sobre $[0,1]$ é uniformemente contínua. O próximo lema nos dá a ideia correspondente de uniformidade para elementos de D.

Lema 1.13. Para cada x em D e cada ε positivo, existem pontos $t_0, t_1, \ldots, t_r$ tais que

$$0 = t_0 < t_1 < \cdots < t_r = 1 \tag{1.10}$$

e

$$w_x[t_{i-1}, t_i) < \varepsilon, \ i = 1, 2, \ldots, r. \tag{1.11}$$

Demonstração. Seja τ o supremo dos t em $[0,1]$ para os quais $[0,t)$ pode ser decomposto em finitos subintervalos $[t_{i-1}, t_i)$ satisfazendo (1.11). Como $x(0) = x(0^+)$, temos $\tau > 0$; como $x(\tau^-)$ existe, $[0,\tau)$ pode então se decompor; $\tau < 1$ não ocorre pois $x(\tau) = x(\tau^+)$ neste caso. $\square$

Iremos precisar de um módulo que desempenhe em D o mesmo papel que o módulo de continuidade desempenha em C. Para $0 < \delta < 1$, ponha

$$w'_x(\delta) = \inf_{\{t_i\}} \max_{0<i\leq r} w_x[t_{i-1}, t_i), \tag{1.12}$$

onde o ínfimo se estende sobre os conjuntos finitos $\{t_i\}$, satisfazendo

$$\begin{cases} 0 = t_0 < t_1 < \ldots < t_r = 1 \\ t_i - t_{i-1} > \delta, \ i = 1, 2, \ldots, r. \end{cases} \tag{1.13}$$

O lema 1.13 é equivalente a afirmação que

$$\lim_{\delta \to 0} w'_x(\delta) = 0 \tag{1.14}$$

para todo x em D.

A topologia de Skorohod

Duas funções x e y estão próximas, na topologia usada por C, se o gráfico de $x(t)$ pode ser transportado para o gráfico de $y(t)$ por uma pequena pertubação uniforme das ordenadas, com as abscissas mantidas fixas. Em D, iremos permitir uma pequena deformação uniforme da escala do tempo. Fisicamente, isto equivale a admitir que não conseguimos medir o tempo com uma precisão maior do que podemos posicionar. A topologia a seguir, idealizada por Skorohod, incorpora esta ideia.

Seja Λ a classe de aplicações estritamente crescente e contínuas de $[0, 1]$ em $[0, 1]$. Se $\lambda \in \Lambda$, então $\lambda(0) = 0$ e $\lambda(1) = 1$. Para x e y em D, defina $d(x, y)$ como o ínfimo de todos os números positivos ε para os quais existe um λ em Λ tal que

$$\sup_t |\lambda t - t| \leq \varepsilon \tag{1.15}$$

e

$$\sup_t |x(t) - y(\lambda t)| \leq \varepsilon. \tag{1.16}$$

Pode-se mostrar que d é de fato uma métrica, à sua topologia associada denominamos topologia de Skorohod. A distância uniforme entre x e y pode ser definida como o ínfimo dos números positivos para os quais $\sup_t |x(t) - y(t)| \leq \varepsilon$. O λ em (1.15) e (1.16) representa a pequena deformação uniforme da escala do tempo mencionado acima.

Iremos introduzir em D uma outra métrica d_0 - uma métrica equivalente à d, mas sob a qual D é completo. A completude facilita a caracterização de compacidade.

A ideia de definir d_0 é requerer que o tempo de deformação λ, que intervém na definição de d, esteja próximo da função identidade num certo sentido que a princípio parece mais rigoroso que (1.15); nominalmente, requeremos que a inclinação $(\lambda t - \lambda s)/t - s$ de cada corda esteja perto de um ou, o que é a mesma coisa e analiticamente mais coveniente, que seu logaritmo esteja perto de zero.

Se λ é uma função não decrescente em $[0,1]$ com $\lambda(0) = 0$ e $\lambda(1) = 1$, defina

$$\|\lambda\| = sup_{s \neq t} \left| log \frac{\lambda t - \lambda s}{t - s} \right| \tag{1.17}$$

Se $\|\lambda\|$ é finito, então as inclinações das cordas de λ são limitadas longe de zero e no infinito, assim, λ é contínuo e estritamente crescente e, portanto, é um elemento de Λ. Embora $\|\lambda\|$ possa ser infinito até mesmo se λ pertence a Λ, tais elementos de Λ não entram na definição a seguir.

Definimos $d_0(x, y)$ como o ínfimo de todos os números positivos para os quais Λ contém algum λ com

$$\|\lambda\| \leq \varepsilon \tag{1.18}$$

e

$$\sup_t |x(t) - y(\lambda t)| \leq \varepsilon \tag{1.19}$$

Lema 1.14. Se $d(x, y) < \delta^2$, onde $0 < \delta < 1/4$, então $d_0(x, y) \leq 4\delta + w'_x(\delta)$.

Demonstração. Escolha pontos t_i satisfazendo (1.13) e

$$w_x[t_{i-1}, t_i) < w'_x(\delta) + \delta, \ i = 1, 2, \ldots, r. \tag{1.20}$$

Agora escolha um μ de Λ tal que

$$\sup_t |x(t) - y(\mu t)| = \sup_t |x(\mu^{-1} t) - y(t)| < \delta^2 \tag{1.21}$$

e

$$\sup_t |\mu t - t| < \delta^2. \tag{1.22}$$

Queremos definir um λ em Λ que ficará próximo de μ mas não terá, como μ pode ter, cordas com inclinação muito longe de um. Tome λ coincidindo com μ nos pontos t_i e linear entre eles. Como a composição $\mu^{-1}\lambda$ deixa fixos os t_i e é crescente, t e $\mu^{-1}\lambda t$ sempre estão no mesmo subintervalo $[t_{i-1}, t_i)$. Por (1.20)e (1.21), portanto:

$$\begin{aligned} |x(t) - y(\lambda t)| &\leq |x(t) - x(\mu^{-1}\lambda t)| + |x(\mu^{-1}\lambda t) - y(\lambda t)| \\ &< w'_x(\delta) + \delta + \delta^2 < 4\delta + w'_x(\delta). \end{aligned}$$

Agora é suficiente provar que $\|\lambda\| \leq 4\delta$. Como λ coincide com μ em t_i, segue de (1.22) e da desigualdade $t_i - t_{i-1} > \delta$ que

$$|(\lambda t_i - \lambda t_{i-1})| < |2\delta^2| < 2\delta(t_i - t_{i-1}).$$

Pela característica poligonal de λ, segue que

$$|(\lambda t - \lambda s) - (t - s)| \leq 2\delta|t - s|$$

é satisfeito para todo s e t. Portanto,

$$\log(1 - 2\delta) \log\left(\frac{\lambda t - \lambda s}{t - s}\right) \leq (1 + 2\delta);$$

como $\delta < 1/4$, segue que $\|\lambda\| \leq 4\lambda$. □

Teorema 1.15. *As métricas d e d_0 são equivalentes*

Demonstração. Vamos denotar a bola aberta de centro x e raio ε com a métrica d e d_0 por $S_d(x, \varepsilon)$ e $S_{d_0}(x, \varepsilon)$, respectivamente. Se $d_0(x, y) < \varepsilon$, (1.18) e (1.19) são satisfeitas para algum $\lambda \in \Lambda$. Se $\varepsilon < 1/4$, então, como $\lambda(0) = 0$,

$$\log(1 - 2\varepsilon) < -\varepsilon \leq \log\frac{\lambda t}{t} \leq \varepsilon < \log(1 + 2\varepsilon).$$

Assim, temos $|\lambda t - t| \leq 2\varepsilon$. Portanto, $d(x, y) \leq 2d_0(x, y)$ se $d_0(x, y) < 1/4$. Isto mostra que dentro de uma bola arbitrária $S_d(x, \varepsilon)$ podemos encontrar uma bola $S_{d_0}(x, \delta)$.

O lema 1.14 implica que, se

$$\delta < 1/4, \ 4\delta + w'_x(\delta) < \varepsilon, \tag{1.23}$$

então $S_d(x, \delta^2) \subset S_{d_0}(x, \varepsilon)$. Dados x e ε, podemos, por (1.14), encontrar um δ satisfazendo (1.23). Dentro de uma bola na métrica d_0 podemos encontrar então uma bola da métrica d com o mesmo centro. Portanto d e d_0 são métricas equivalentes. □

Teorema 1.16. *O espaço D é completo na métrica d_0.*

Demonstração. É suficiente mostrar que cada sequência de Cauchy contém uma subsequência convergente. Se $\{x_k\}$ é uma sequência de Cauchy, ela contém uma subsequência $\{y_n\} = \{x_{k_n}\}$ tal que

$$d_0(y_n, y_{n+1}) < 1/2^n. \tag{1.24}$$

Iremos provar que $\{y_n\}$ é convergente. Por (1.24), Λ contém μ_n tal que

$$\sup_t |y_n(t) - y_{n+1}(\mu_n t)| < 1/2^n \tag{1.25}$$

e

$$\|\mu_n\| < 1/2^n. \tag{1.26}$$

Isto implica, para $m \geq 1$,

$$\sup_t |\mu_{n+m+1}\mu_{n+m} \cdots \mu_{n+1}\mu_n t - \mu_{n+m} \cdots \mu_{n+1}\mu_n t| =$$

$$= \sup_s |\mu_{n+m+1}s - s| \leq 1/2^{n+m}.$$

(Aqui, $\mu_{n+m}\mu_{n+1}\mu_n$ denota composições iteradas.) Para n fixado, as funções

$$\mu_{n+m} \cdots \mu_{n+1}\mu_n t \tag{1.27}$$

são, portanto, uniformemente de Cauchy quando $m \to \infty$. Assim, (1.27) converge uniformemente para um limite

$$\lambda_n t = \lim_m \mu_{n+m} \cdots \mu_{n+1}\mu_n t. \tag{1.28}$$

A função λ_n deve ser contínua e não-decrescente e deve satisfazer $\lambda_n(0) = 0$, $\lambda_n(1) = 1$. Se provarmos que $\|\lambda_n\|$ é finito, segue que λ_n é estritamente crescente e, portanto, um elemento de Λ. Observe que para $\lambda_1,\ \lambda_2 \in \Lambda$, temos $\|\lambda_1\lambda_2\| \leq \|\lambda_1\| + \|\lambda_2\|$, disto segue que

$$\left|\log \frac{\mu_{n+m} \cdots \mu_{n+1}\mu_t - \mu_{n+m} \cdots \mu_{n+1}\mu_n s}{t-s}\right|$$

$$\leq \|\mu_{n+m} \cdots \mu_{n+1}\mu_n\| \leq \|\mu_n\| \leq \|\mu_{n+1}\| \leq \cdots \|\mu_{n+m}\| \leq 1/2^{n-1}.$$

Fazendo $m \to \infty$ no primeiro membro desta desigualdade, obtemos $\|\lambda_n\| \leq 1/2^{n-1}$, em particular, $\|\lambda_n\|$ é finito e $\lambda_n \in \Lambda$.

Por (1.28), $\lambda_n = \lambda_{n+1}\mu_n$. Portanto, por (1.25),

$$\sup_t |y_n(\lambda_n^{-1}t) - y_{n+1}(\lambda_{n+1}^{-1})| = \sup_s |y_n(s) - y_{n+1}(\mu_n s)| < 1/2^n.$$

Em consequência, as funções $y_n(\lambda_n^{-1}t)$, que são elementos de D, são uniformemente de Cauchy e, portanto, converge uniformemente para uma função limite $x(t)$. É fácil mostrar que x deve também ser um elemento de D. Como $\sup_t |y_n(\lambda^{-1}t) - x(t)| \to 0$

e $\|\lambda_n\| \to 0$, temos $d_0(y_n, x) \to 0$, que completa a prova.

□

Caracterização de compacidade

Considere o módulo $w''(\delta)$ dado por

$$w''(\delta) = \sup \min\{|x(t) - x(t_1)|, |x(t_2) - x(t)|\}, \tag{1.29}$$

onde o supremo se estende sobre os números t_1, t e t_2, que satisfaz

$$t_1 \leq t \leq t_2,\ t_2 - t_1 < \delta. \tag{1.30}$$

Dado δ e ε, decomponha $[0,1)$ em subintervalos $[s_{i-1}, s_i)$ tal que $s_i - s_{i-1} > \delta$ e $w_x[s_{i-1}, s_i) < w'_x(\delta) + \varepsilon$. Se (1.30)é satisfeita, então, t_1 e t_2 estão no mesmo subintervalo $[s_{i-1}, s_i)$, neste caso $|x(t) - x(t_1)| < w'_x(\delta) + \varepsilon$ e $|x(t_2) - x(t)| < w'_x(\delta) + \varepsilon$, ou então eles estão em intervalos adjacentes $[s_{i-1}, s_i)$ e $[s_i, s_{i+1})$, neste caso, $|x(t) - x(t_1)| < w'_x(\delta) + \varepsilon$ para $t_1 \leq t < s_i$ e $|x(t_2) - x(t)| < w'_x(\delta) + \varepsilon$ para $s_i \leq t \leq t_2$. Portanto,

$$w'_x(\delta) \leq w'_x(\delta) \tag{1.31}$$

É possível formular uma condição de compacidade em termos de $w''_x(\delta)$ e do comportamento de x perto de 0 e de 1.

Teorema 1.17. *Um conjunto A possui o fecho compacto na topologia de Skorohod se, e somente se,*

$$\sup_{x \in A} \sup_t |x(t)| < \infty \tag{1.32}$$

e

$$\begin{cases} \lim_{\delta \to 0} \sup_{x \in A} w''_x(\delta) = 0, \\ \lim_{\delta \to 0} \sup_{x \in A} w_x[0, \delta) = 0, \\ \lim_{\delta \to 0} \sup_{x \in A} w_x[1-\delta, 1) = 0 \end{cases} \tag{1.33}$$

Demonstração. Pelo teorema 1.16, é suficiente mostrar que (1.33) é equivalente a

$$\lim_{\delta \to 0} \sup_{x \in A} w'(\delta) = 0 \tag{1.34}$$

Note que (1.34) implica (1.33) por (1.31) e da definição de $w'(\delta)$. Iremos provar a recíproca.

Dado um ε positivo, escolha um δ positivo tal que, para todo $x \in A$,

$$w_x''(\delta) < \varepsilon,\ w_x[0,\delta) < \varepsilon \text{ e } w_x[1-\delta,1) < \varepsilon. \tag{1.35}$$

Assuma que $x \in A$, iremos mostrar que

$$w'_{(\frac{1}{2}\delta)} \leq 6\varepsilon, \tag{1.36}$$

que é suficiente para provar o teorema.

Vamos provar primeiro que

$$t_1 \leq s \leq t \leq t_2,\ t_2 - t_1 \leq \delta \tag{1.37}$$

implica

$$\min\{|x(s) - x(t_1)|,\ |x(t_2) - x(t)|\} < 2\varepsilon. \tag{1.38}$$

De fato, se $|x(s) - x(t_1)| > \varepsilon$, então, por (1.35), $|x(t) - x(s)| < \varepsilon$ e $|x(t_2) - x(s)| < \varepsilon$, assim, $|x(t_2) - x(t)| < 2\varepsilon$.

Suponha que x possui um salto excedendo 2ε nos pontos τ_1 e τ_2. Se $0 < \tau_2 - \tau_1 < \delta$, então existe pontos t_1, s, t, t_2 satisfazendo (1.37), $t_1 < \tau_1 = s$, e $t < \tau_2 = t_2$, se t_1 está próximo suficiente de τ_1, e se t está próximo suficiente de τ_2, então (1.38) não é satisfeito, portanto $[0, 1]$ não pode conter dois pontos cuja distância seja menor que δ em cada x que possui salto maior que 2ε. Por (1.35), nem $[0, \delta)$ ou $[1 - \delta, 1)$ podem conter um ponto no qual x possui um salto maior que 2ε.

Desta forma, existem pontos s_i, com $0 = s_0 < s_1 < \cdots < s_r = 1$, tais que $s_i - s_{i-1} \geq \delta$ e tal que qualquer ponto no qual x salta mais que 2ε é um dos s_i. Se $s_j - s_{j-1} > \delta$ para um par de pontos adjacentes, aumente o sistema $\{s_i\}$ incluindo seus pontos médios. Continuando desta forma, chegamos a um novo sistema aumentado $s_0, \ldots, s_r$ (com um novo r) que satisfaz

$$\frac{1}{2}\delta < s_i - s_{i-1} \leq \delta,\ i = 1, 2, \ldots, r.$$

Agora, (1.36) será válido se mostrarmos que

$$w_x[s_{i-1}, s_i) \leq 6\varepsilon \tag{1.39}$$

para cada i. Suponha $s_{i-1} \leq t_1 < t_2 < s_i$. Então $t_2 - t_1 < \delta$. Seja σ_1 o supremo dos $\sigma \in [t_1, t_2]$ para os quais $\sup_{t_1 \leq u \leq \sigma} |x(u) - x(t_1)| \leq \varepsilon$; seja σ_2 o ínfimo dos $\sigma \in [t_1, t_2]$ para os quais $\sup_{\sigma \leq u \leq t_2} |x(t_2) - x(u)| \leq 2\varepsilon$. Se $\sigma_1 < \sigma_2$, existem pontos s à direita de σ_1 com $|x(s) - x(t_1)| > 2\varepsilon$ e existem pontos t à esquerda de σ_2 com $|x(t_2) - x(t)| > 2\varepsilon$, como podemos afirmar que $s < t$, isto contradiz o fato que (1.37) implica (1.38). Assim,

$\sigma_2 \leq \sigma_1$ e segue que $|x(\sigma_1^-) - x(t_1)| \leq 2\varepsilon$ e $|x(t_2) - x(\sigma_1)| \leq 2\varepsilon$. Como $t_1 < \sigma_1 \leq t_2$, $\sigma_1 \in (s_{i-1}, s_i)$, logo o salto em σ_1 é pelo menos de 2ε. Desta forma $|x(t_2) - x(t_1)| \leq 6\varepsilon$. Isto estabelece (1.39), que prova (1.36) e o teorema. □

Conjuntos finito-dimensionais

Conjuntos finito-dimensionais exercem em D o mesmo papel que exercem em C. Para pontos $t_1, \ldots, t_k \in [0,1]$, defina a projeção natural $\pi_{t_1 \cdots t_k}$ de D para $\mathbb{R}^k$ como a usual:

$$\pi_{t_1 \ldots t_k}(x) = (x(t_1), \ldots, x(t_k)). \tag{1.40}$$

Observe que π_0 e π_1 são contínuas em todo ponto. Suponha $0 < t < 1$. Se x_n converge para x na topologia de Skorohod e x é contínuo em t, então $x_n(t) \to x(t)$. Suponha por outro lado que x é descontínuo em t. Se $\lambda_n \in \Lambda$ é tal que é linear em $[0,t]$ e em $[t,1]$ e satisfaz $\lambda_n t = t - 1/n$, e se pusermos $x_n(s) = x(\lambda_n s)$, então x_n converge para x na topologia de Skorohod, mas $x_n(t)$ não converge para $x(t)$. Portanto, se $0 < t < 1$, então π_t é contínuo em x se, e somente se, x é contínuo em t.

Vamos provar que $\pi_{t_1 \cdots t_k}$ é mensurável com respeito a σ-álgebra $\mathcal{D}$ dos conjuntos borelianos na topologia de Skorohod. Precisamos considerar apenas um único ponto t (já que uma função cujo contradomínio é $\mathbb{R}^k$ é mensurável se cada componente for mensurável), e iremos assumir que $t < 1$ (pois π_1 é contínua). Se x_n converge para x na topologia de Skorohod, então $x_n(s) \to x(s)$ para os pontos de continuidade s de x e, portanto, para pontos s fora de um conjunto de medida de Lebesgue nula. Como x_n é uniformemente limitada,

$$\frac{1}{\varepsilon}\int_t^{t+\varepsilon} x_n(s)ds \longrightarrow \frac{1}{\varepsilon}\int_t^{t+\varepsilon} x(s)ds \quad (n \to \infty) \tag{1.41}$$

para cada ε positivo. Assim, $h_\varepsilon(x) = \varepsilon^{-1}\int_t^{t+\varepsilon} x(s)ds$ é contínuo na topologia de Skorohod. Por continuidade à direita, $h_\varepsilon(x) \to \pi_t(x)$ para cada x quando $\varepsilon \to 0$. Portanto, $\pi_t(x)$ é mensurável. Assim, podemos, como em C, definir os conjuntos finito-dimensionais como conjuntos da forma $\pi^{-1}_{t_1 \cdots t_k} H$ com $k \geq 1$ e $H \in \mathcal{R}^k$.

Se T_0 é um subconjunto de $[0,1]$, seja $\mathcal{F}_{T_0}$ a classe dos conjuntos $\pi^{-1}_{t_1 \cdots t_k} H$, onde k é arbitrário, os t_i são pontos arbitrários de T_0 e $H \in \mathcal{R}^k$. Então, $\mathcal{F}_{T_0}$ é uma álgebra. Neste caso, $\mathcal{F}_{[0,1]}$ é a classe dos conjuntos finito-dimensionais.

Teorema 1.18. *Se T_0 contém 1 e é denso em $[0,1]$, então $\mathcal{F}_{T_0}$ gera $\mathcal{D}$.*

Demonstração. Como D é separável, é suficiente mostrar que cada bola aberta $B_{d_0}(x,r)$ pertence à σ-álgebra gerada por $\mathcal{F}_{T_0}$. Fixe o centro x e o raio r. Escolha em T_0 uma sequência $t_1, t_2, \ldots$ que seja densa em $[0,1]$ com $t_1 = 1$. Para $0 < \varepsilon < r$ e $k \geq 1$, seja $A_k(\varepsilon)$ o conjunto dos y para os quais existe um $\lambda \in \Lambda$ satisfazendo

$$\|\lambda\| < r - \varepsilon$$

e

$$\max_{1 \leq i \leq k} |y(t_i) - x(\lambda t_i)| < r - \varepsilon.$$

É suficiente provar que

$$S_{d_0}(x, r) = \bigcup_{\varepsilon} \bigcap_{k=1}^{\infty} A_k(\varepsilon), \tag{1.42}$$

onde a união é tomada sobre os ε racionais em $(0, r)$ e provar que cada $A_k(\varepsilon)$ pertence à $\mathcal{F}_{T_0}$.

Iremos provar a segunda afirmação primeiro. Para ε e k fixos, seja H_1 o conjunto dos pontos $(x(\lambda t_1), \ldots, x(\lambda t_k)) \in \mathbb{R}^k$, onde λ varia sobre essas funções em Λ satisfazendo $\|\lambda\| < r - \varepsilon$. Seja H_2 o conjunto dos pontos $(\alpha_1, \ldots, \alpha_k) \in \mathbf{R}^k$ tal que $|\alpha_i - \beta_i| < r - \varepsilon$, $i = 1, \ldots, k$, para algum $(\beta_1, \ldots, \beta_k) \in H_1$. Então H_2 é aberto, e portanto, pertence a $\mathcal{R}^k$, e $A_k(\varepsilon) = \pi^{-1}_{t_1 \cdots t_k} H_2$. Logo $A_k(\varepsilon) \in \mathcal{F}_{T_0}$.

É fácil ver que o lado esquerdo de (1.42) está contido no lado direito. Completamos a prova se mostrarmos que

$$\bigcap_{k=1}^{\infty} A_k(\varepsilon) \subset S_{d_0}(x, r). \tag{1.43}$$

Se y pertence à interseção acima, escolha, para cada k, um λ_k em Λ com

$$\|\lambda_k\| < r - \varepsilon \tag{1.44}$$

e

$$\max_{1 \leq i \leq k} |y(t_i) - x(\lambda_k t_i)| < r - \varepsilon. \tag{1.45}$$

Pelo teorema de seleção de Helly A.5, existe uma subsequência $\{\lambda_{k'}\}$ e uma função não-decrescente λ tal que

$$\lim_{k'} \lambda_{k'} t = \lambda t \tag{1.46}$$

é satisfeito para os pontos de continuidade t de λ. Iremos mostrar que $\lambda \in \Lambda$, que $\|\lambda\| \leq r - \varepsilon$, e que $\sup_t |y(t) - x(\lambda t)| \leq r - \varepsilon$. Isto implica que $d_0(x, y) < r$ e prova (1.43).

Se s e t são pontos de continuidade distintos de λ, então, por (1.44),

$$\left|log\frac{\lambda t-\lambda s}{t-s}\right|=\lim_{k'}\left|log\frac{\lambda_{k'}t-\lambda_{k'}s}{t-s}\right|\leq r-\varepsilon. \tag{1.47}$$

Esta relação mostra que é impossível ocorrer um salto em λ (em particular, (1.46) vale para todo t) e isto implica que λ é estritamente crescente, logo $\lambda \in \Lambda$. A desigualdade (1.47) também implica que $\|\lambda\| \leq r-\varepsilon$. Por (1.45), $|y(t_1)-x(t_1)|<r-\varepsilon$. Se $i>1$, então, por (1.45), $|y(t_i)-x(\lambda_{k'}t_i)|<r-\varepsilon$ para $k'\geq i$. Como $\lambda_{k'}t\to\lambda t$, temos que $|y(t_i)-x(\lambda t_i)|\leq r-\varepsilon$ ou $|y(t_i)-x((\lambda t_i)^-)|\leq r-\varepsilon$ e, como $\{t_i\}$ é denso, temos $\sup_t|y(t)-x(\lambda t)|\leq r-\varepsilon$. Isto prova o teorema 1.18. □

1.6.1 Convergência fraca e tightness em D

Uma técnica muito útil em C para provar convergência fraca consiste em provar convergência fraca das distribuições finito-dimensionais junto com o fato da sequência em questão ser tight. Queremos adaptar isto para D. Como D é separável e completo sob a métrica d_0, uma família de medidas de probabilidade em $(D,\mathcal{D})$ é relativamente compacta se, e somente se, ela é tight, assim, não há dificuldade neste quesito. Por outro lado, o fato que as projeções naturais não são contínuas traz uma certa dificuldade.

Distribuições finito-dimensionais

Para uma medida de probabilidade P em $(D,\mathcal{D})$, seja T_P o conjunto dos $t\in[0,1]$ para os quais a projeção π_t é contínua exceto nos pontos que formam um conjunto de medida P nula. Os pontos 0 e 1 sempre estão em T_P. Se $0<t<1$, então $t\in T_P$ se, e somente se, $P(J_t)=0$, onde

$$J_t=\{x:x(t)\neq x(t^-)\} \tag{1.48}$$

é o conjunto dos x que são descontínuos em t, para $0<t<1$, π_t é contínua em x se, e somente se, x é contínua em t.

Um elemento de D possui no máximo uma quantidade enumerável de saltos. Vamos provar o fato correspondente que $P(J_t)>0$ é possível para no máximo uma quantidade enumerável de pontos t. Seja $J_t(\varepsilon)$ o conjunto dos x que possuem em t um salto $|x(t)-x(t^-)|$ excedendo ε. Para ε e δ fixos, pode haver no máximo um número finito de pontos para os quais $P(J_t(\varepsilon))\geq\delta$, se esta desigualdade fosse válida para uma sequência de pontos distintos $t_1,t_2,\ldots$, então o conjunto $\limsup_n J_{t_n}(\varepsilon)$ teria medida no mínimo δ e, portanto, seria não-vazio, contradizendo o fato que para cada x o salto pode exceder ε apenas numa quantidade finita de pontos. Para um ε positivo fixado, portanto, $P(J_t(\varepsilon))$ pode exceder zero para no máximo uma quantidade enumerável de pontos t. Como $P(J_t(\varepsilon)) \uparrow P(J_t)$, quando $\varepsilon \downarrow 0$, segue o resultado.

Desta forma, T_P contém 0 e 1 e seu complemento em $[0,1]$ é no máximo uma quantidade enumerável. Se $t_1, \ldots, t_k$ estão em T_P, então $\pi_{t_1 \cdots t_k}$ é continua exceto num conjunto de medida P nula.

Lema 1.19. Seja h uma função mensurável e D_h o conjunto dos pontos de descontinuidade de h. Se $P_n \Rightarrow P$ e $P(D_h) = 0$, então $P_n h^{-1} \Rightarrow P h^{-1}$

Demonstração. Iremos mostrar que se F é uma fechado, então,

$$\limsup_{n\to\infty} P_n h^{-1}(F) \leq P h^{-1}(F).$$

Como $P_n \Rightarrow P$, temos

$$\limsup_n P_n(h^{-1}F) \leq \limsup_n P_n((h^{-1}F)^-) \leq P((h^{-1}F)^-).$$

Assim, é suficiente provar que $P((h^{-1}F)^-) = P(h^{-1}F)$, que segue da hipótese $P(D_h) = 0$ e do fato que $(h^{-1}F)^- \subset D_h \cup (h^{-1}F)$. □

Suponha

$$P_n \Rightarrow P, \tag{1.49}$$

onde P_n e P são medidas de probabilidade em $(D, \mathcal{D})$, pelo lema 1.19,

$$P_n \pi^{-1}_{t_1 \cdots t_k} \Rightarrow P \pi^{-1}_{t_1 \cdots t_k} \tag{1.50}$$

vale se todos os t_i estão em T_P. Em geral, (1.50) não segue de (1.49) se algum t_i está fora de T_P: Se P é uma unidade de massa na função $I[0, 1/2)$ e P_n uma unidade de massa em $I[0, 1/2 + n^{-1})$, então $P_n \Rightarrow P$ vale enquanto $P_n \pi^{-1} \Rightarrow P \pi^{-1}_{1/2}$ não vale.

Teorema 1.20. *Se $\{P_n\}$ é tight e se $P_n \pi^{-1}_{t_1 \cdots t_k} \Rightarrow P \pi^{-1}_{t_1 \cdots t_k}$ para todos $t_1, \ldots, t_k$ em T_P, então $P_n \Rightarrow P$.*

Demonstração. Por tightness, cada subsequência $\{P_{n'}\}$ contém uma subsequência $\{P_{n''}\}$ convergindo fracamente para algum limite Q. É suficiente mostrar que Q sempre coincide com P.

Se $t_1, \ldots, t_k$ pertencem à $T_P \cap T_Q$, temos $P_{n''} \pi^{-1}_{t_1 \cdots t_k} \Rightarrow P \pi^{-1}_{t_1 \cdots t_k}$ por hipótese e

$$P_{n''} \pi^{-1}_{t_1 \cdots t_k} \Rightarrow Q \pi^{-1}_{t_1 \cdots t_k},\ t_1, \ldots, t_k \in T_P \cap T_Q. \tag{1.51}$$

Como T_P e T_Q contém $[0,1]$ exceto uma quantidade enumerável, o mesmo é válido para $T_P \cap T_Q$; em particular, esta interseção é densa. Como $T_P \cap T_Q$ contém 1, segue do teorema 1.18 que $\mathcal{D}$ é gerado por conjuntos finito-dimensionais com base nos pontos de $T_P \cap T_Q$. Portanto, $P = Q$ segue de (1.51), e a prova está completa. $\square$

Teorema 1.21. *Uma sequência $\{P_n\}$ é tight se, e somente se, as seguintes condições são satisfeitas:*

(i) Para cada η positivo, existe um a tal que

$$P_n\{x : \sup_t |x(t)| > a\} \leq \eta,\ n \geq 1.$$

(ii) Para cada η e ε positivos, existem um δ, $0 < \delta < 1$, e um inteiro n_0 tal que

$$P_n\{x : w''_x(\delta) \geq \varepsilon\} \leq \eta,\ n \geq n_0, \tag{1.52}$$

$$P_n\{x : w_x[0,\delta) \geq \varepsilon\} \leq \eta,\ n \geq n_0, \tag{1.53}$$

$$P_n\{x : w_x[1-\delta,1) \geq \varepsilon\} \leq \eta,\ n \geq n_0. \tag{1.54}$$

Demonstração. Ver [2] p. 125. $\square$

Teorema 1.22. *Suponha que*

$$P_n\pi^{-1}_{t_1\cdots t_k} \Rightarrow P\pi^{-1}_{t_1\cdots t_k} \tag{1.55}$$

para todos $t_1, \cdots, t_k$ em T_P. Suponha ainda que $P(J_1) = 0$ e que, para cada ε e η positivos, existe um δ, $0 < \delta < 1$, e um inteiro n_0 tal que

$$P_n\{x : w''_x(\delta) \geq \varepsilon\} \leq \eta,\ n \geq n_0. \tag{1.56}$$

Então $P_n \Rightarrow P$.

Demonstração. Pelo teorema 1.20, é suficiente mostrar que $\{P_n\}$ é tight. Iremos primeiro verificar a condição (i) do teorema anterior. Dado um η positivo, escolha δ e n_0 satisfazendo (1.56) com $\varepsilon = 1$, digamos. Por ser denso, T_P contém pontos $t_1, \ldots, t_k$ com $0 = t_1 < \cdots < t_k = 1$, tal que $t_i - t_{i-1} < \delta$. De (1.55) segue que $\{P_n\pi^{-1}_{t_1\cdots t_k}\}$ é uma sequência tight em $\mathbb{R}^k$ e, portanto, que

$$P_n\{x : \max_{1\leq i\leq k} |x(t_i)| > a_0\} \leq \eta,\ n \geq 1, \tag{1.57}$$

para uma escolha apropriada de a_0. Se $|x(t_i)| \leq a_0$ $i = 1, \ldots, k$, e se $w''_x(\delta) < 1$, então $|x(t)| < a_0 + 1$, para todo t. Se $a = a_0 + 1$, então por (1.57) e (1.56) (com $\varepsilon = 1$)

temos $P_n\{x : \sup_t |x(t)| > a\} \leq 2\eta$ para $n \geq n_0$. Para uma quantidade finita de n precedendo n_0, temos certeza que esta desigualdade é válida aumentando-se a. Cada P_n sendo tight. Isto estabelece a condição (i).

Como para a condição (ii) do teorema anterior, devemos encontrar, dados ε e η, um δ e n_0 satisfazendo (1.53) e (1.54). Como cada x em D e contínuo à direita, temos

$$P\{x : |x(\delta) - x(0)| \geq \varepsilon\} < \eta \tag{1.58}$$

para todo δ suficientemente pequeno. Escolha em T_P um δ pequeno o suficiente que satisfaça (1.58) e também satisfaça (1.56) para um n_0 apropriado. Como 0 e δ estão em T_P, temos $P_n\pi_{0,\delta}^{-1} \Rightarrow P\pi_{0,\delta}^{-1}$ como um caso especial de (1.55), e segue de (1.58) que

$$P_n\{x : |x(\delta) - x(0)| \geq \varepsilon\} < \eta \tag{1.59}$$

para todo n suficientemente grande. Se $w''_x(\delta) < \varepsilon$ e $|x(\delta) - x(0)| < \varepsilon$, então $|x(s) - x(0)| < 2\varepsilon$, para todo s em $[0, \delta]$, e assim $w_x[0, \delta] < 4\varepsilon$. Desta forma, (1.56) e (1.59) implica $P\{x : w_x[0, \delta) \geq 4\varepsilon\} \leq 2\eta$. Isto mostra que (1.53) é satisfeito.

Com uma mudança, o argumento simétrico funciona para (1.54). No lugar de (1.58), precisamos de

$$P\{x : |x(1) - x(1 - \delta)| \geq \varepsilon\} < \eta, \tag{1.60}$$

para δ pequeno. Como um elemento de D não precisa ser contínuo à esquerda, (1.60) não é válido em geral, já que assumimos $P(J_1) = 0$, x é contínuo à esquerda em $t = 1$ exceto para x num conjunto de medida P nula, e (1.60) vale para todo δ suficientemente pequeno. Isto completa a prova do teorema. $\square$

Elementos aleatórios de D

Iremos chamar um elemento aleatório de D de função aleatória. Se X é uma função de $(\Omega, \mathcal{B}, \mathbf{P})$ em D, então, para cada ω, $X(\omega)$ é um elemento de D cujo valor em t denotaremos por $X(t, \omega)$. Para cada t, $X(t)$ denota a função real $\pi_t X$ sobre $\Omega-$ seu valor em ω é $X(t, \omega)$. Assim como no espaço C, X é um elemento aleatório de D ($X^{-1}\mathcal{D} \subset \mathcal{B}$) se, e somente se, para cada t, $X(t)$ é uma variável aleatória (use o teorema 1.18).

Uma sequência X_n de elementos aleatórios de D é por definição tight quando a sequência de suas distribuições correspondentes é tight.

Um critério para convergência

Sejam X_n e X elementos aleatórios de D. Escreva T_X para T_P, onde P é a distribuição de X. Então T_X contém 0 e 1, e, se $0 < t < 1$, $t \in T_X$ se, e somente se, $\mathbf{P}\{X(t) \neq X(t^-)\} = 0$. Iremos escrever $w''(X, \delta)$ no lugar de $w''_X(\delta)$.

Lema 1.23. Se $\gamma \geq 0$ e $\alpha > 1/2$, e se

$$\mathbf{P}\{|S_j - S_i| \geq \lambda,\ |S_k - S_j| \geq \lambda\} \leq \frac{1}{\lambda^{2\gamma}} \left(\sum_{i<l\leq k} u_l \right)^{2\alpha},\ 0 \leq i \leq l \leq k \leq m, \tag{1.61}$$

é satisfeito para todo λ positivo, então, para todo λ positivo,

$$\mathbf{P}\{M''_m \geq \lambda\} \leq \frac{K_{\gamma,\alpha}}{\lambda^2 \gamma} (u_1 + \cdots + u_m)^{2\alpha}, \tag{1.62}$$

onde $K''_{\gamma,\alpha}$ é uma constante que depende somente de γ e α.

Demonstração. Ver [2] p. 98. $\square$

Teorema 1.24. *Suponha que*

$$(X_n(t_1), \ldots, X_n(t_k)) \xrightarrow{\mathcal{D}} (X(t_1), \ldots, X(t_k)) \tag{1.63}$$

para todos $t_1, \ldots, t_k$ em T_X, que $\mathbf{P}\{X(1) \neq X(1^-)\} = 0$, e que

$$\mathcal{P}\{|X_n(t) - X_n(t_1)| \geq \lambda, |X_n(t_2) - X_n(t)| \geq \lambda\} \leq \frac{1}{\lambda^{2\gamma}} [F(t_2) - F(t_1)]^{2\alpha} \tag{1.64}$$

para $t_1 \leq t \leq t_2$ e $n \geq 1$, onde $\gamma \geq 0$, $\alpha > 1/2$, e F é uma função não decrescente e contínua em $[0, 1]$. Então $X_n \xrightarrow{\mathcal{D}} X$.

Demonstração. Pelo teorema 1.22, é suficiente mostrar que para cada ε e η positivos existe um δ positivo tal que

$$\mathbf{P}\{w''(X_n, \delta) \geq \varepsilon\} \leq \eta \tag{1.65}$$

vale para todo n. Fixados τ, δ e n, e para m inteiro positivo, considere as variáveis aleatórias

$$\xi_i = X_n \left(\tau + \frac{1}{m} \delta \right) - X_n \left(\tau + \frac{i-1}{m} \delta \right),\ i = 1, 2, \ldots, m. \tag{1.66}$$

Seja

$$M''_m = \max \min \left\{ \left| X_n \left(\tau + \frac{j}{m} \delta \right) - X_n \left(\tau + \frac{i}{m} \delta \right) \right|, \left| X_n \left(\tau + \frac{k}{m} \delta \right) - X_n \left(\tau + \frac{j}{m} \delta \right) \right| \right\},$$

onde o máximo se estende sobre $0 \le i \le j \le k \le m$. Por (1.64) e pelo lema 1.23,

$$\mathbf{P}\{M''_m \ge \varepsilon\} \le \frac{K}{\varepsilon^{2\gamma}}[F(\tau+\delta) - F(\tau)]^{2\alpha}, \tag{1.67}$$

onde $K = K''_{\gamma,\alpha}$ depende apenas de γ e α.

Ponha

$$w''(X_n, [\tau, \tau+\delta]) = \sup \min\{|X_n(t) - X_n(t_1)|, |X_n(t_2) - X_n(t)|\}, \tag{1.68}$$

onde o supremo se estende sobre t_1, t, t_2, satisfazendo $\tau \le t_1 \le t \le t_2 \le \tau + \delta$. Como X_n é uma função contínua à direita sobre $[0,1]$, fazendo $m \to \infty$ em (1.67), obtemos

$$\mathbf{P}\{w''(X_n, [\tau, \tau+\delta]) > \varepsilon\} \le \frac{K}{\varepsilon^{2\gamma}}[F(\tau+\delta) - F(\tau)]^{2\alpha}. \tag{1.69}$$

Suponha agora que $\delta = 1/(2u)$ é o inverso de um inteiro positivo, e suponha que

$$w''(X_n, [2i\delta, (2i+2)\delta]) \le \varepsilon, \ 0 \le i \le u-1, \tag{1.70}$$

e

$$w''(X_n, [(2_i+1)\delta, (2_i+3)\delta]) \le \varepsilon, \ 0 \le i \le u-2. \tag{1.71}$$

Se $t_1 \le t \le t_2$ e $t_2 - t_1 \le \delta$, então t_1 e t_2 estão em algum dos $2_u - 1$ intervalos envolvidos em (1.70) e (1.71), assim, $\min\{|X_n(t) - X_n(t_1)|, |X_n(t_2) - X_n(t)|\} \le \varepsilon$. Portanto, (1.70) e (1.71) implica $w''(X_n, \delta) \le \varepsilon$. Segue, agora, por (1.69) que

$$\mathbf{P}\{w''(X_n, \delta) \ge \varepsilon\} \le \frac{K}{\varepsilon^{2\gamma}}\left(\sum{}' + \sum{}''\right), \tag{1.72}$$

onde $\sum'$ e $\sum''$ são somas da forma

$$\sum_{k=1}^{r}[F(t_k) - F(t_{k-1})]^{2\alpha}$$

com $0 \le t_1 \le \cdots \le t_r \le 1$ e $t_k - t_{k-1} \le 2\delta$. Assim,

$$\mathbf{P}\{w''(X_n, \delta) \ge \varepsilon\} \le \frac{2K}{\varepsilon^{2\gamma}}[F(1) - F(0)][w_F(2\delta)]^{2\alpha-1}. \tag{1.73}$$

Como $2\alpha > 1$ e F é contínuo, o lado direito da desigualdade acima vai para 0 com δ e isto prova (1.65). □

Existência do movimento browniano

A medida de Wiener, denotada por W, é a principal medida de probabilidade sobre $(C, \mathcal{C})$. Ela é caracterizada pelas seguintes propriedades:

(i) $W\{x : x(0) = 0\} = 1$

(ii) $W\pi_t^{-1}$ segue distribuição $\mathcal{N}(0, t)$

(iii) O processo estocástico $\{x_t : 0 \leq t \leq 1\}$ possui incrementos independente sob W, isto é, para todos $0 \leq t_1 < t_2 < \cdots < t_k \leq 1$, as variáveis aleatórias $t_2 - t_1$, $t_3 - t_2$, ..., $t_k - t_{k-1}$ são independentes.

Lema 1.25. Se $\gamma \geq 0$ e $\alpha > 1$, e se

$$\mathbf{P}\{|S_j - S_i| \geq \lambda\} \leq \frac{1}{\lambda^\gamma} \left(\sum_{i<l\leq j} u_l \right)^\alpha, \ 0 \leq i \leq j \leq m.$$

para todo λ positivo, então, para todo λ positivo,

$$\mathbf{P}\{M_m \geq \lambda\} \leq \frac{K'_{\gamma,\alpha}}{\lambda^\gamma}(u_1 + \cdots + u_m)^\alpha, \tag{1.74}$$

onde $K'_{\gamma,\alpha}$ depende apenas de γ e α.

Demonstração. ver [2] p. 94. □

Se X_n são elementos aleatórios de S, dizemos que $\{X_n\}$ é tight quando $\{P_n\}$ é tight, onde P_n é a distribuição de X_n.

O lema 1.25 nos direciona a uma condição de tightness para uma sequência $\{X_n\}$ de elementos de C.

Teorema 1.26. *A sequência $\{X_n\}$ é tight se ela satisfaz as seguintes condições:*

(i) A sequência $\{X_n(0)\}$ é tight.

(ii) Existem constantes $\gamma \geq 0$ e $\alpha > 1$ e uma função F sobre $[0,1]$ não decrescente e contínua tal que

$$\mathbf{P}\{|X_n(t_2) - X_n(t_1)| \geq \lambda\} \leq \frac{1}{\lambda^\gamma}|F(t_2) - F(t_1)|^\alpha \tag{1.75}$$

para todos t_1, t_2, n e todo λ positivo.

A condição

$$E\{|X_n(t_2) - X_n(t_1)|^\gamma\} \leq |F(t_2) - F(t_1)|^\alpha \tag{1.76}$$

implica (1.75).

Demonstração. Pelo teorema 1.21 é suficiente obtermos, dados ε e η, um $\delta(0 < \delta < 1)$ para os quais,

$$\mathbf{P}\{w(X_n, \delta) \geq 3\varepsilon\} \leq \eta \tag{1.77}$$

para todo n, e, pelo corolário do teorema (8.3), isto será válido se δ^{-1} é um inteiro e

$$\sum_{j<\delta^{-1}} \mathbf{P}\left\{\sup_{j\delta \leq s \leq (j+1)\delta} |X_n(s) - X_n(j\delta)| \geq \varepsilon\right\} \leq \eta. \tag{1.78}$$

Fixe n, δ e j "por um momento"e, para um inteiro positivo m, considere as variáveis aleatórias

$$\xi_i = X_n\left(j\delta + \frac{i}{m}\delta\right) - X_n\left(j\delta + \frac{i-1}{m}\delta\right),\ i = 1, 2, \ldots, m. \tag{1.79}$$

Por (1.75), estas variáveis aleatórias satisfazem (12.43) (ver isto) com $u_i = F(j\delta + i\delta m^{-1}) - F(j\delta + (i-1)\delta m^{-1})$. Pelo teorema 1.25, portanto,

$$\mathbf{P}\left\{\max_{0\leq i\leq m} |X_n\left(j\delta + \frac{i}{m}\delta\right) - X_n(j\delta)| \geq \varepsilon\right\} \leq \frac{K}{\varepsilon^\gamma}[F((j+1)\delta) - F(j\delta)]^\alpha, \tag{1.80}$$

onde $K = K'_{\gamma,\alpha}$. Como X_n pertence à C, fazendo $m \to \infty$, obtemos

$$\mathbf{P}\left\{\sup_{j\delta \leq s \leq (j+1)\delta} |X_n(s) - X_n(j\delta)| \geq \varepsilon\right\} \leq \frac{K}{\varepsilon^\gamma}[F((j+1)\delta) - F(j\delta)]^\alpha. \tag{1.81}$$

Portanto,

$$\mathbf{P}\left\{\sup_{j\delta \leq s \leq (j+1)\delta} |X_n(s) - X_n(j\delta)| \geq \varepsilon\right\} \leq \frac{K}{\varepsilon^\gamma}[F(1) - F(0)][\max_{j<\delta^{-1}} |F(j+1)\delta - F(j\delta)|]^{\alpha-1} \tag{1.82}$$

se δ^{-1} é inteiro. Como F é contínuo e $\alpha > 1$, podemos fazer esta última quantidade ficar pequena tomando δ como sendo o inverso de um número inteiro grande, e isto completa a prova. □

Teorema 1.27. *Existe em C um elemento aleatório com distribuição finito-dimensional $\mu_{t_1\cdots t_k}$, desde que tais distribuições sejam consistentes e desde que existam constantes*

$\gamma \geq 0$ e $\alpha > 1$ e uma função F sobre $[0,1]$ não decrescente e contínua tal que

$$\mu_{t_1,t_2}\{(\beta_1,\beta_2) : |\beta_1 - \beta_2| \geq \lambda\} \leq \frac{1}{\lambda^\gamma}|F(t_2) - F(t_1)|^\alpha, \tag{1.83}$$

para todos t_1, t_2 e λ positivo.

Se $\mu_{t_1 \cdots t_k}$ satisfaz

$$\int_{\mathbb{R}^2} |\beta_1 - \beta_2|^\gamma d\mu_{t_1,t_2}(\beta_1,\beta_2) \leq |F(t_2) - F(t_1)|^\alpha, \tag{1.84}$$

então (1.83) *segue.*

Demonstração. Para cada n, construa uma variável aleatória poligonal X_n que é linear sobre cada intervalo $[(i-1)2^{-n}, i2^{-n}]$ e para o qual a distribuição de

$$X_n(0), X_n\left(\frac{1}{2^n}\right), \ldots, X_n\left(\frac{2^n-1}{2^n}\right), X_n(1)$$

é $\mu_{t_0,\ldots,t_2^n}$ com $t_i = i/2^n$. Se t_1 e t_2 são múltiplos de 2^{-n}, então, por (1.83),

$$\mathbf{P}\{|X_n(t_2) - X_n(t_1)| \geq \lambda\} \leq \frac{1}{\lambda^\gamma}|F(t_2) - F(t_1)|^\alpha. \tag{1.85}$$

Como na prova anterior, considere fixos n, δ *e* j, onde δ^{-1} é inteiro por hipótese. Se os pontos

$$j\delta + i\delta m^{-1}, i = 0, 1, \ldots, m \tag{1.86}$$

envolvidos em (1.79) são todos múltiplos de 2^{-n}, então (1.80) segue como antes. Suponha agora que $\delta 2^n$ é um inteiro. Se temos $m = \delta 2^n$, então os pontos (1.86) são, de fato, múltiplos de 2^{-n}, assim (1.80) é válido. Além disso, os pontos (1.86) são, neste caso, exatamente os múltiplos de 2^{-n} no intervalo $[j\delta, (j+1)\delta]$, de modo que, por causa da característica poligonal de X_n, (1.81) e (1.82) são satisfeitos.

Portanto, (1.82) é válido se δ^{-1} e $\delta 2^n$ são ambos inteiros. Se temos $\delta = 2^{-v}$, onde v é um inteiro suficientemente grande de modo que o lado direito de (1.82) é menor que η, logo (1.77) é válido para todo $n \geq v$, o qual, como $X_n(0)$ possuem todos a mesma distribuição, é suficiente para ser tight.

Pelo teorema de Prohorov, portanto, existe um elemento aleatório X de C e uma subsequência $\{X_{n'}\}$ tal que $X_n \xrightarrow{\mathcal{D}} X$. Se $t_1, \ldots, t_k$ são racionais cujo denominadores são potências de dois, então, pela hipótese de consistência, a distribuição de $(X_n(t_1), \ldots, X_n(t_k))$ é exatamente $\mu_{t_1 \cdots t_k}$ para todo n suficientemente grande, e que $(X(t_1), \ldots, X(t_k))$ possui distribuição $\mu_{t_1 \cdots t_k}$. Para pontos genéricos de $[0,1]$, existem racionais cujos denominadores são potências de dois $s_1^{(v)}, \ldots, s_k^{(v)}$ com $\lim_v s_i^{(v)} =$

t_i, como $(X(s_1^{(v)}), \ldots, X(s_k^{(v)}))$ converge em distribuição para $(X(t_1), \ldots, X(t_k))$, e como, por (1.83) e pela continuidade de F, $\mu_{s_1^{(v)} \cdots s_k^{(v)}}$ converge fracamente para $\mu_{t_1 \cdots t_k}$, $(X(t_1), \ldots, X(t_k))$ possui $\mu_{t_1 \cdots t_k}$ como distribuição. Desta forma, X é a função aleatória requerida.

A existência de W e W° seguem do teorema 1.27. □

Teorema de Donsker

Em alguns aspectos, o espaço D é mais conveniente que C para a formulação do teorema de Donsker. Dadas variáveis aleatórias $\xi_1, \xi_2, \ldots$ em $(\Omega, \mathcal{B}, \mathbf{P})$ com somas parciais $S_n = \xi_1 + \cdots + \xi_n$, seja $X_n(\omega)$ a função em D com valor

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega) \tag{1.87}$$

em t. Como cada $X_n(t)$ é uma variável aleatória, X_n é uma função aleatória - um elemento aleatório de D.

Queremos provar, sob condições adequadas que a distribuição de X_n converge fracamente para a medida de Wiener W. W está definido sobre $(C, \mathcal{C})$, mas é facilmente estendido para $(D, \mathcal{D})$: Como $C \in D$, e como a topologia relativa de Skorohod em C coincide com a topologia uniforme, $A \in \mathcal{D}$ implica $A \cap C \in \mathcal{C}$. Podemos, portanto, estender W para $(D, \mathcal{D})$ dando o valor $W(A \cap C)$ para A em $\mathcal{D}$. De agora em diante, iremos interpretar W como esta medida de probabilidade sobre $(D, \mathcal{D})$ ou como um elemento aleatório de D com esta medida de probabilidade como sua distribuição.

Teorema 1.28. *Suponha que as variáveis aleatórias ξ_n são independentes e identicamente distribuidas com média zero, com variância positiva σ^2:*

$$E\{\xi_n\} = 0, E\{\xi_n^2\} = \sigma^2. \tag{1.88}$$

Então as funções aleatórias X_n definidas por (1.87) *satisfazem*

$$X_n \xrightarrow{\mathcal{D}} W. \tag{1.89}$$

Demonstração. A convergência das distribuições finito-dimensionais segue do teorema central do limite. Assim, pelo teorema 1.24, é suficiente verificar a desigualdade

$$E\{|X_n(t) - X_n(t_1)|^2 |X_n(t_2) - X_n(t)|^2\} \leq 4(t_2 - t_1)^2, \quad t_1 \leq t \leq t_2. \tag{1.90}$$

Por (1.87), (1.88), e pela independência de ξ_n, o lado esquerdo de (1.90) fica

$$\begin{aligned} \frac{1}{\sigma^2 n^2} & \quad E\{|S_{[nt]} - S_{[nt_1]}|^2 E\{|S_{[nt_2]} - S_{[nt]}|^2\}\} = \\ = & \quad \frac{1}{n^2}([nt] - [nt_1])([nt_2] - [nt]) \leq \left\{ \frac{[nt_2] - [nt_1]}{n} \right\}^2 \end{aligned}$$

Se $t_2 - t_1 \geq 1/n$, (1.90) segue. Se $t_2 - t_1 < 1/n$, então t_1 e t ou t_2 e t estão no mesmo subintervalo $[(i-1)/n, i/n)$, em ambos os casos, o lado esquerdo de (1.90) se anula. Isto estabelece (1.90) em geral e prova o teorema. $\square$

Capítulo 2

Leis Estáveis e Processos de Lévy

2.1 Leis Estáveis

Sejam $X_1, \ldots, X_n$ variáveis aleatórias *i.i.d.* e $S_n = X_1 + \cdots + X_n$. Lembremos do Teorema Central do limite que se $EX_i = \mu$ e $var(X_i) = \sigma^2 \in (0, \infty)$, então

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow \chi$$

Onde χ possui distribuição normal padrão. Nesta seção iremos investigar o caso em que ${EX_1}^2 = \infty$ e daremos condições necessárias e suficientes para a existência de constantes a_n e b_n tais que

$$\frac{S_n - b_n}{a_n} \Rightarrow Y, \quad \text{onde } Y \text{ é não degenerado.}$$

Começaremos com um exemplo.

Exemplo 2.1. Suponha que a distribuição de X_i é tal que

$$P(X_1 > x) = P(X_1 < -x) = \frac{x^{-\alpha}}{2} \quad \text{para} \quad x \geq 1. \tag{2.1}$$

Onde $0 < \alpha < 2$. Se $\varphi(t) = E\exp(itX_1)$, então

$$\begin{aligned} 1 - \varphi(t) &= \int_1^\infty (1 - e^{itx})\frac{\alpha}{2|x|^{\alpha+1}}dx + \int_{-\infty}^{-1}(1 - e^{itx})\frac{\alpha}{2|x|^{\alpha+1}}dx \\ &= \alpha\int_1^\infty \frac{1 - \cos(tx)}{x^{\alpha+1}}dx \end{aligned}$$

Fazendo a mudança $tx = u$, $dx = du/t$ a última integral se torna

$$= \alpha \int_t^\infty \frac{1-\cos u}{(u/t)^{\alpha+1}} \frac{du}{t} = t^\alpha \alpha \int_t^\infty \frac{1-\cos u}{u^{\alpha+1}} du$$

Quando $u \to 0$, $1 - \cos u \sim u^2/2$. Então $\dfrac{1-\cos u}{u^{\alpha+1}} \sim u^{-\alpha+1}/2$, que é integrável, pois $\alpha < 2$ implica $-\alpha + 1 > -1$. Se considerarmos

$$C = \alpha \int_0^\infty \frac{1-\cos u}{u^{\alpha+1}} du < \infty$$

e observarmos que (2.1) implica $\varphi(t) = \varphi(-t)$, então os resultados acima mostram que

$$1 - \varphi(t) \sim C|t|^\alpha \quad \text{quando} \quad t \to 0 \tag{2.2}$$

Sejam $X_1, \ldots, X_n$ *i.i.d.* com distribuição dada em (2.1) e seja $S_n = X_1 + \cdots + X_n$.

$$E \exp\left(\frac{itS_n}{n^{1/\alpha}}\right) = \varphi\left(\frac{t}{n^{1/\alpha}}\right)^n = \left(1 - \{1 - \varphi(t/n^{1/\alpha})\}\right)^n$$

Quando $n \to \infty$, $n(1 - \varphi(t/n^{1/\alpha})) \to C|t|^\alpha$, assim segue que

$$E \exp\left(\frac{itS_n}{n^{1/\alpha}}\right) \to \exp(-C|t|^\alpha)$$

Desta forma, o lado direito da igualdade acima é a função característica de alguma variável aleatória Y e

$$\frac{S_n}{n^{1/\alpha}} \Rightarrow Y \tag{2.3}$$

Lema 2.1. Se $h_n(\epsilon) \to g(\epsilon)$ para cada $\epsilon > 0$ e $g(\epsilon) \to g(0)$ quando $\epsilon \to 0$, então podemos escolher $\epsilon_n \to 0$ de forma que $h_n(\epsilon_n) \to g(0)$.

Demonstração. Escolha N_m de forma que $|h_n(1/m) - g(1/m)| \leq 1/m$ para $n \geq N_m$ e $m \to N_m$ é crescente. Seja $\epsilon_n = 1/m$ para $N_m \leq n < N_{m+1}$ e $\epsilon_n = 1$ para $n < N_1$. Quando $N_m \leq n < N_{m+1}$, $\epsilon_n = 1/m$, então segue da desigualdade triangular e da definição de ϵ_n que

$$\begin{aligned} |h_n(\epsilon_n) - g(0)| &\leq |h_n(1/m) - g(1/m)| + |g(1/m) - g(0)| \\ &\leq 1/m + |g(1/m) - g(0)| \end{aligned}$$

Quando $n \to \infty$, temos $m \to \infty$ e segue o resultado. □

Definição 2.1. Uma função L é de **variação lenta** se

$$\lim_{x\to\infty}\frac{L(tx)}{L(x)} = 1 \quad \text{para todo} \quad t > 0.$$

Teorema 2.2. *Sejam $X_1, \dots, X_n$ variáveis aleatórias i.i.d. com distribuição que satisfaz*

(i) $\lim_{x\to\infty}\dfrac{P(X_1 > x)}{P(|X_1| > x)} = \theta \in [0,1]$

(ii) $P(|X_1| > x) = x^{-\alpha}L(x)$

onde $\alpha < 2$ e L é de variação lenta. Sejam $S_n = X_1 + \cdots + X_n$,

$$a_n = \inf\{x : P(|X_1| > x) \le n^{-1}\} \quad e \quad b_n = nE(X_1 1_{(|X_1|\le a_n)})$$

Se $n \to \infty$, então $\dfrac{S_n - b_n}{a_n} \Rightarrow Y$ onde Y tem distribuição não degenerada.

Demonstração. Primeiro, veremos que (ii) implica

$$nP(|X_1| > a_n) \to 1 \tag{2.4}$$

Para provar isto, note que $nP(|X_1| > a_n) \le 1$. sejam $\epsilon > 0$, $x = a_n/(1+\epsilon)$ e $t = 1+2\epsilon$, (ii) implica

$$(1+2\epsilon)^{-\alpha} = \lim_{n\to\infty}\frac{P(|X_1| > (1+2\epsilon)a_n/(1+\epsilon))}{P(|X_1| > a_n/(1+\epsilon))} \le \liminf_{n\to\infty}\frac{P(|X_1| > a_n)}{1/n}$$

provando (2.4) já que ϵ é arbitrário. Combinando (2.4) com (i) e (ii), temos

$$\begin{aligned}
\lim_{n\to\infty} nP(X_1 > xa_n) &= \lim_{n\to\infty}\frac{nP(X_1 > xa_n)}{nP(|X_1| > xa_n)}nP(|X_1| > xa_n) \\
&= \theta\lim_{n\to\infty} nP(|X_1| > xa_n) = \theta\lim_{n\to\infty} nx^{-\alpha}a_n^{-\alpha}L(xa_n) \\
&= \theta x^{-\alpha}\lim_{n\to\infty} na_n^{-\alpha}L(a_n)\frac{L(xa_n)}{L(a_n)} \\
&= \theta x^{-\alpha}, \quad \text{para} \quad x > 0.
\end{aligned}$$

Portanto,

$$nP(X_1 > xa_n) \to \theta x^{-\alpha} \quad \text{para} \quad x > 0. \tag{2.5}$$

Se colocarmos $I_j^n = 1_{(X_j > xa_n)}$, então $|m \leq n : X_m > xa_n| = \sum_{j=1}^n I_j^n$. Observe que I_j^n possui distribuição de Bernoulli com parâmetro $p_n = P(X_1 > xa_n)$ e $I_1^n, \ldots, I_n^n$ são independentes, logo, pelo lema A.1, $\sum_{j=1}^n I_j^n$ possui distribuição Binomial(n, p_n). Além disso, (2.5) implica $np_n \to \theta x^{-\alpha}$, portanto, usando o lema A.2, concluímos que $|m \leq n : X_m > xa_n| \Rightarrow \text{Poisson}(\theta x^{-\alpha})$.

Para somar os pontos, considere $I_n(\epsilon) = \{m \leq n : |X_m| > \epsilon a_n\}$

$$\begin{aligned}
\hat{\mu}(\epsilon) &= EX_m 1_{(\epsilon a_n < |X_m| \leq a_n)}; \quad \hat{S}_n(\epsilon) = \sum_{m \in I_n(\epsilon)} X_m; \\
\overline{\mu}(\epsilon) &= EX_m 1_{(|X_m| \leq \epsilon a_n)};
\end{aligned}$$

$$\begin{aligned}
\overline{S}_n(\epsilon) &= (S_n - b_n) - (\hat{S}_n(\epsilon) - n\hat{\mu}(\epsilon)) \\
&= \sum_{m=1}^n \{X_m 1_{(|X_m| \leq \epsilon a_n)}\} + n(EX_m 1_{(\epsilon a_n \leq 1 \leq a_n)} - EX_m 1_{(|X_m| \leq a_n)}) \\
&= \sum_{m=1}^n \{X_m 1_{(|X_m| \leq \epsilon a_n)}\} - nEX_m 1_{(|X_m|)} \\
&= \sum_{m=1}^n \{X_m 1_{(|X_m| \leq \epsilon a_n)} - \overline{\mu}(\epsilon)\}
\end{aligned}$$

Se colocarmos $\overline{X}_m(\epsilon) = X_m 1_{(|X_m| \leq \epsilon a_n)}$, usando os lemas A.3 e (A.4), obtemos

$$\begin{aligned}
E(\overline{S}_n(\epsilon)/a_n)^2 &= n\ var(\overline{X}_1(\epsilon)/a_n) \leq nE(\overline{X}_1(\epsilon)/a_n)^2 = \int_0^\epsilon 2yP(|X_1| > ya_n)dy \\
E(\overline{S}_n(\epsilon)/a_n)^2 &\leq \int_0^\epsilon 2yP(|X_1| > ya_n)dy \\
&= P(|X_1| > a_n) \int_0^\epsilon 2y \frac{P(|X_1| > ya_n)}{P(|X_1| > a_n)} dy
\end{aligned}$$

Gostaríamos de usar (2.5) e (ii),trocando o limite com a integral, para concluir

$$E(\overline{S}_n(\epsilon)/a_n)^2 \to \int_0^\epsilon 2yy^{-\alpha}dy = \frac{2}{2-\alpha}\epsilon^{2-\alpha}$$

e, portanto,

$$\limsup_{n\to\infty} E(\overline{S}(\epsilon)/a_n)^2 \leq \frac{2\epsilon^{2-\alpha}}{2-\alpha} \tag{2.6}$$

Para justificar a troca do limite com a integral e completar a prova de (2.6), iremos mostrar o seguinte (tome $\delta < 2 - \alpha$):

Lema 2.3. Para qualquer $\delta > 0$ existe um C tal que para todo $t \geq t_0$ e $y \leq 1$

$$P(|X_1| > yt)/P(|X_1| > t) \leq Cy^{-\alpha-\delta}$$

Demonstração. (ii) implica que quando $t \to \infty$

$$P(|X_1| > t/2)/P(|X_1| > t) \to 2^{\alpha}$$

desta forma, para $t \geq t_0$, temos

$$P(|X_1| > t/2)/P(|X_1| > t) \leq 2^{\alpha+\delta}$$

iterando e parando na primeira vez que $t/2^m < t_0$ nós temos para todo $n \geq 1$

$$P(|X_1| > t/2^n)/P(|X_1| > t) \leq C2^{(\alpha+\delta)n}$$

Onde $C = 1/P(|X_1| > t_0)$. Aplicando o resultado para os primeiros n com $1/2^n < y$ e notando que $y \leq 1/2^{n-1}$, temos

$$P(|X_1| > yt)/P(|X_1| > t) \leq C2^{\alpha+\delta}y^{-\alpha-\delta}$$

que prova o lema. □

Para calcular o limite de $\hat{S}_n(\epsilon)$, observamos que, com um cálculo análogo ao qual obtemos (2.5), $nP(|X_1| > \epsilon a_n) \to \epsilon^{-\alpha}$, assim $|I_n(\epsilon)| \Rightarrow Poisson(\epsilon^{-\alpha})$. Dado $|I_n(\epsilon)| = m$, $\hat{S}_n(\epsilon)/a_n$ é a soma de m variáveis aleatórias independentes com distribuição F_n^{ϵ} tal que, para $x \geq \epsilon$

$$\begin{aligned}
1 - F_n^{\epsilon}(x) &= P(X_1/a_n) > x||X_1|/a_n \geq \epsilon) \\
&= \frac{P(X_1/a_n > x||X_1|/a_n > \epsilon)}{P(|X_1|/a_n > \epsilon)} = \frac{P(X_1/a_n > x)}{P(|X_1|/a_n > \epsilon)} \\
&= \frac{nP(X_1/a_n > x)}{nP(|X_1|/a_n > \epsilon)} \to \theta x^{-\alpha}/\epsilon^{-\alpha} \\
F_n^{\epsilon}(-x) &= P(X_1/a_n < -x||X_1|/a_n > \epsilon) \\
&= \frac{P(X_1/a_n < -x, |X_1|/a_n > \epsilon)}{P(|X_1|/a_n > \epsilon)} = \frac{P(X_1/a_n < -x)}{P(|X_1|/a_n > \epsilon)} \to (1-\theta)|x|^{-\alpha}/\epsilon^{-\alpha}
\end{aligned}$$

Portanto, $F_n^\epsilon \Rightarrow F^\epsilon$, onde

$$F^\epsilon(x) = \begin{cases} 1 - \theta x^{-\alpha}/\epsilon^{-\alpha}, \text{ se } x \geq \epsilon \\ (1-\theta)|x|^{-\alpha}/\epsilon^{-\alpha}, \text{ se } x \leq -\epsilon \end{cases}$$

Se denotarmos por $\psi_n^\epsilon(t)$ a função característica de F_n^ϵ então o Teorema de Continuidade de Lévy A.9 implica

$$\psi_n^\epsilon(t) \to \psi^\epsilon(t) = \int_\epsilon^\infty e^{itx}\theta\epsilon^\alpha\alpha x^{-(\alpha+1)}dx + \int_{-\infty}^{-\epsilon} e^{itx}(1-\theta)\epsilon^\alpha\alpha|x|^{-(\alpha+1)}dx$$

quando $n \to \infty$.

$$\begin{aligned} E\exp(it\hat{S}_n(\epsilon)/a_n) &= \sum_{k=0}^n E(e^{it\hat{S}_n(\epsilon)/a_n}|I_n(\epsilon)=k)P(I_n(\epsilon)=k) \\ &= \sum_{k=0}^n [E(e^{it\overline{X}_1(\epsilon)/a_n}|I_n(\epsilon)=k)]^n P(I_n(\epsilon)=k) \\ &= \sum_{k=0}^n (\psi_n^\epsilon(t))^k P(I_n(\epsilon)=k) \\ &= \sum_{k=0}^n \frac{(\psi_n^\epsilon(t))^k}{k!}(np_n)^k(1-p_n)^n \\ &= \sum_{k=0}^n \frac{(np_n\psi_n^\epsilon(t))^k}{k!}(1-p_n)^n \\ &\to \exp(\epsilon^{-\alpha}\psi^\epsilon)\exp(-\epsilon^\alpha) = \exp(-\epsilon^\alpha(1-\psi^\epsilon)) \end{aligned}$$

Note que $\epsilon^{-\alpha} = \int_\epsilon^\infty \alpha x^{-(\alpha+1)}dx$. Desta forma,

$$\exp(-\epsilon^\alpha(1-\psi^\epsilon)) = \exp\left(\int_\epsilon^\infty (e^{itx}-1)\theta\alpha x^{-(\alpha+1)}dx + \int_{-\infty}^{-\epsilon}(e^{itx}-1)(1-\theta)\alpha|x|^{-(\alpha+1)}dx\right)$$

Denotando $F_Y \equiv nX_m/a_n$, observamos que (2.5) implica $nP(xa_n < X_m \leq ya_n) = F_Y(y) - F_Y(x) \to \theta(x^{-\alpha} - y^{-\alpha})$, logo, $d_{F_Y}(y) = \theta\alpha y^{-\alpha-1}dy$, além disso, como $nP(X_1 < -a_ny) \to (1-\theta)y^{-\alpha}$, $d_{F_Y}(y) = -\alpha(1-\theta)|y|^{-\alpha-1}dy$, para $y < 0$. Lembrando que

$$\hat{\mu}(\epsilon) = EX_m 1_{(\epsilon a_n < |X_m| \leq a_n)}$$

Concluímos

$$\begin{aligned} n\hat{\mu}(\epsilon)/a_n &= \int_\epsilon^1 x d_{F_Y}(x) + \int_{-1}^{-\epsilon} x d_{F_Y}(x) \\ &\to \int_\epsilon^1 x\theta\alpha x^{-\alpha-1}dx + \int_{-1}^{-\epsilon} x\theta\alpha|x|^{-\alpha-1}dx \end{aligned} \tag{2.7}$$

Disto, segue que $E\exp(it\{\hat{S}_n(\epsilon) - n\hat{\mu}(\epsilon)\}/a_n) \to$

$$\exp(\int_1^\infty (e^{itx}-1)\theta\alpha x^{-(\alpha+1)}dx$$

$$\begin{aligned} &+ \int_\epsilon^1 (e^{itx}-1-itx)\theta\alpha x^{-(\alpha+1)}dx \\ &+ \int_{-1}^{-\epsilon} (e^{itx}-1-itx)(1-\theta)\alpha|x|^{-(\alpha+1)}dx \\ &+ \int_{-\infty}^{-1} (e^{itx}-1)(1-\theta)\alpha|x|^{-(\alpha+1)}dx) \end{aligned} \tag{2.8}$$

A última expressão é confusa, mas $e^{itx}-1-itx \sim -t^2x^2/2$ quando $t \to 0$, assim, precisamos subtrair itx para obter

$$\int_0^1 (e^{itx}-1-itx)x^{-(\alpha+1)}dx \ \text{ converge quando } \ \alpha \geq 1$$

Para reduzir o número de integrais de quatro para duas, podemos escrever o limite quando $\epsilon \to 0$ no lado direito de (2.8) como

$$\exp\left(itc + \int_0^\infty (e^{itx}-1-\frac{itx}{1+x^2})\theta\alpha x^{-(\alpha+1)}dx + \int_{-\infty}^0 (e^{itx}-1-\frac{itx}{1+x^2})(1-\theta)\alpha|x|^{-(\alpha+1)}\right)dx \tag{2.9}$$

onde c é uma constante. Combinando (2.4) e (2.8), usando o lema 2.1, segue que $(S_n - b_n)/a_n \Rightarrow Y$, onde Ee^{itY} é dada em (2.9). □

Fazendo alguns cálculos (veja [4](1968), p. 204-206) podemos reescrever (2.9) como

$$\exp\left(itc - b|t|^\alpha\{1+\kappa sgn(t) w_\alpha(t)\}\right) \tag{2.10}$$

onde $-1 \leq \kappa \leq 1$, $(\kappa = 2\theta - 1)$ e

$$w_\alpha(t) = \begin{cases} \tan(\pi\alpha/2), & \text{se } \alpha \neq 1 \\ (2/\pi)\log|t|, & \text{se } \alpha = 1. \end{cases}$$

Note que até agora assumimos $0 < \alpha < 2$. Se considerarmos $\alpha = 2$, então o termo com κ se anula e (2.10) se reduz a função característica da distribuição normal com média c e variância $2b$.

Definição 2.2. As distribuições cujas as funções características são dadas como em (2.10) são chamadas de **leis estáveis**. α é geralmente chamado de **índice**.

Quando $\alpha = 1$ e $\kappa = 0$, obtemos a distribuição de Cauchy. Além da Cauchy e da normal, existe apenas um caso em que a densidade é conhecida: Quando $\alpha = 1/2, \kappa = 1, c = 0$, e $b = 1$, a densidade é

$$(2\pi y^3)^{-1/2} \exp(-1/2y) \quad \text{para} \quad y \geq 0 \tag{2.11}$$

Para confirmação basta calcular a função característica.

Exemplo 2.2. Sejam $X_1, X_2, \ldots$ variáveis aleatórias *i.i.d.* com densidade simétria em torno de zero e contínua e positiva em zero. Afirmamos que

$$\frac{1}{n}\left(\frac{1}{X_1} + \cdots + \frac{1}{X_n}\right) \Rightarrow \textit{Cauchy } (\alpha = 1, \kappa = 0)$$

Para verificar isto, note que

$$P(1/X_i > x) = P(0 < X_i < x^{-1}) = \int_0^{x^{-1}} f(y)dy \sim f(0)/x$$

quando $x \to \infty$. Um cálculo análogo mostra que $P(1/X_i < -x) \sim f(0)/x$, assim (i) do teorema 2.2 vale com $\theta = 1/2$, e (ii) é satisfeito com $\alpha = 1$. Além disso $a_n \sim 2f(0)n$, enquanto os coeficientes b_n se anulam, já que supomos que a distribuição de X_1 é simétrica em torno de zero.

Nosso próximo resultado explica o nome **leis estáveis**.

Definição 2.3. Dizemos que uma variável aleatória Y possui uma **lei estável** se para todo inteiro $k > 0$ existem constantes a_k e b_k tais que se $Y_1, \ldots, Y_k$ são *i.i.d.* e possuem a mesma distribuição de Y, então $(Y_1 + \cdots + Y_k - b_k)/a_k =_d Y$.

Teorema 2.4. *Y é o limite de $(X_1 + \cdots + X_k - b_k)/a_k$ para alguma sequência i.i.d. X_i se, e somente se, Y possui uma lei estável.*

Demonstração. Se Y possui uma lei estável, podemos tomar $X_1, X_2, \ldots$ *i.i.d.* com distribuição Y. Reciprocamente, seja

$$Z_n = (X_1 + \ldots + X_n - b_n)/a_n$$

e $S_n^j = X_{(j-1)n+1} + \cdot + X_{jn}$. Alguns cálculos nos dá

$$
\begin{aligned}
Z_{nk} &= (S_n^1 + \ldots + S_n^k - b_{nk})/a_{nk} \\
a_{nk} Z_{nk} &= (S_n^1 - b_n) + \cdots + (S_n^k - b_n) + (kb_n - b_{nk}) \\
a_{nk} Z_{nk}/a_n &= (S_n^1 - b_n)/a_n + \cdots + (S_n^k - b_n)/a_n + (kb_n - b_{nk})/a_n
\end{aligned}
$$

Os primeiros k termos do lado direito $\Rightarrow Y_1 + \cdots + Y_k$ quando $n \to \infty$, onde $Y_1, \cdots, Y_k$ são independentes e têm a mesma distribuição de Y, e $Z_{nk} \Rightarrow Y$. Tomando $W_n = Z_{nk}$ e

$$W_n' = \frac{a_{kn}}{a_n} Z_{nk} - \frac{kb_n - b_{nk}}{a_n}$$

nos dá o resultado desejado.

□

2.1.1 Distribuições Infinitamente Divisíveis

Na seção passada, indentificamos as distribuições que podem aparecer como o limite de somas normalizadas de variáveis aleatórias *i.i.d.* Nesta seção, iremos identificar aquelas que são limites de somas

$$(*)\ S_n = X_{n,1} + \ldots + X_{n,n} \tag{2.12}$$

onde $X_{m,n}$ são *i.i.d.*

Definição 2.4. Dizemos que Z possui **distribuição infinitamente divisível** se para cada n existe uma sequência *i.i.d.* $Y_{n,1}, \cdots, Y_{n,n}$ tal que

$$Z \stackrel{d}{=} Y_{n,1} + \ldots + Y_{n,n}$$

Uma condição suficiente para Z ser o limite de somas da forma $(*)$ é que Z possua uma distribuição infinitamente divisível. Nosso primeiro resultado mostra que esta condição é também necessária.

Teorema 2.5. *Z é o limite de somas do tipo $(*)$ se, e somente se, Z possui distribuição infinitamente divisível.*

Demonstração. Como observado acima, precisamos provar apenas a necessidade. Escreva

$$S_{2n} = (X_{2n,1} + \cdots + X_{2n,n}) + (X_{2n,n+1} + \cdots + X_{2n,2n}) \equiv Y_n + Y_n'$$

As variáveis aleatórias Y_n e Y'_n são independentes e possuem a mesma distribuição. Se $S_n \Rightarrow Z$, então as distribuições de Y_n formam uma sequência tight, pois

$$P(Y_n > y)^2 = P(Y_n > Y)P(Y'_n > Y) \leq P(S_{2n>2y})$$

e analogamente $P(Y_n < -y)^2 \leq P(S_{2n} < -2y)$. Se tomarmos uma subsequência n_k tal que $Y_{n_k} \Rightarrow Y$ (e portanto $Y'_{n_k} \Rightarrow Y'$), então $Z =_d Y + Y'$. Um argumento similar mostra que Z pode ser dividido em $n > 2$ pedaços e a prova está completa. □

Teorema 2.6. *(Lévy-Khinchin) Z possui distribuição infinitamente divisível se, e somente se, sua função característica satisfaz*

$$log\varphi(t) = ict - \frac{\sigma^2 t^2}{2} + \int \left(e^{itx} - 1 - \frac{itx}{1+x^2} \right) \mu d(x)$$

onde μ é uma medida tal que $\mu(\{0\}) = 0$ e $\int \frac{x^2}{1+x^2}\mu(dx) < \infty$.

Observação 2.1. Perceba a diferença entre uma lei estável e uma lei infinitamente divisível. A primeira aparece como um limite de uma soma $S_n = X_1 + \cdots + X_n$ normalizada, enquanto a outra, pelo teorema 2.5, é o limite de somas da forma $S_n = X_{n,1} + \cdots + X_{n,n}$, onde para cada $i = 1, \ldots, n$ a parcela $X_{n,i}$ depende de n.

2.1.2 Processos de Lévy

Um processo indexado nos reais positivos, $B = \{B_t : t > 0\}$, definido num espaço de probabilidade $(\Omega, \mathcal{F}, \mathbf{P})$ é dito um **movimento Browniano** se os seguintes itens são satisfeitos:

(i) Os caminhos de B são $\mathbf{P}-$contínuos quase certamente.

(ii) $P(B_0 = 0) = 1$.

(iii) Para $0 \leq s \leq t$, $B_t - B_s$ é igual em distribuição a B_{t-s}.

(iv) Para $0 \leq s \leq t$, $B_t - B_s$ é independente de $\{B_u : u \leq s\}$.

(v) Para cada $t > 0$, B_t tem distribuição normal com média zero e variância t.

Um processo indexado nos inteiros não negativos, $N = \{N_t : t \geq 0\}$, definido num espaço de probabilidade $(\Omega, \mathcal{F}, \mathbf{P})$ é dito um **processo de Poisson** com intensidade $\lambda > 0$ se forem satisfeitos os seguintes itens:

(i) Os caminhos de N são $\mathbf{P}-$contínuos à direita com limite à esquerda quase certamente.

(ii) $P(N_0 = 0) = 1$.

(iii) Para $0 \leq s \leq t,\ N_t - N_s$ é igual em distribuição a N_{t-s}.

(iv) Para $0 \leq s \leq t,\ N_t - N_s$ é independente de $\{N_u : u \leq s\}$.

(v) Para cada $t > 0,\ N_t$ tem distribuição de Poisson com parâmetro λt.

Definição 2.5. Um processo $X = \{X_t : t \geq 0\}$, definido num espaço de probabilidade $(\Omega, \mathcal{F}, P)$ é dito um **processo de Lévy** se satisfaz as seguintes propriedades:

(i) Os caminhos de X são P-quase certamente contínuos à direita com limite à esquerda.

(ii) $P(X_0 = 0) = 1$

iii Para $0 \leq s \leq t,\ X_t - X_s$ é igual em distribuição a X_{t-s}.

(iv) Para $0 \leq s \leq t,\ X_t - X_s$ é independente de $\{X_u : u \leq s\}$.

Iremos associar a X a filtração $F = \{\mathcal{F}_t : t \geq 0\}$, onde, para cada $t \geq 0$, $\mathcal{F}_t$ é a ampliação natural da σ-álgebra gerada por $\{X_s : s \leq t\}$. Em particular, essa suposição assegura que, para cada $t \geq 0$, $\mathcal{F}_t$ é completa com respeito aos conjuntos de medida nula de $P|_{\mathcal{F}_t}$.

Vamos discutir agora a relação existente entre distribuição infinitamente divisível e processos que possuem incrementos estacionários e independentes.

Da definição de processo de Lévy, afirmamos que, para qualquer $t > 0$, X_t é uma variável aleatória que possui distribuição infinitamente divisível. Isto segue do fato que, para qualquer $n = 1, 2, \ldots,$

$$X_t = X_{t/n} + (X_{2t/n - X_{t/n}}) + \cdots + (X_t - X_{(n-1)t/n}), \tag{2.13}$$

junto com o fato de que X_t possui incremento estacionário e independente e que $X_0 = 0$. Suponha, agora, que definimos, para todo $\theta \in \mathbb{R}$, $t \geq 0$,

$$\psi_t(\theta) \equiv -\log E(e^{i\theta X_t})$$

Usando (2.13) duas vezes, temos, para quaisquer inteiros positivos $m,\ n$, que

$$m\psi_1(\theta) = \psi_m(\theta) = n\psi_{m/n}(\theta)$$

Portanto, para qualquer $t > 0$ racional,

$$\psi_t(\theta) = t\psi_1(\theta) \tag{2.14}$$

Se t for um número irracional, então podemos escolher uma sequência decrescente de racionais $\{t_n : n \geq 1\}$ tal que $t_n \downarrow t$ quando $n \to \infty$. A continuidade à direita quase certamente de X implica a continuidade à direita de $\exp\{-\psi_t(\theta)\}$(pelo Teo. da Convergência Dominada) e portanto (2.14) vale para todo $t \geq 0$.

Pelo que vimos, todo processo de Lévy possui a propriedade que, para todo $t \geq 0$,

$$E(e^{i\theta X_t}) = e^{-t\psi(\theta)}$$

onde $\psi(\theta) = \psi_1(\theta)$. Ou seja, podemos estabelecer a função característica de X_t, para qualquer $t > 0$, conhecendo apenas a função característica de X_1. Além disso, como X_t possui distribuição infinitamente divisível para qualquer $t > 0$, ainda podemos caracterizar $E(e^{i\theta X_t}) = e^{-t\psi(\theta)}$ pela fórmula de Lévy-Khinchin do teorema 2.6.

2.1.3 Propriedade Forte de Markov

O processo $\{X_t : t \geq 0\}$ possui a propriedade de Markov se, para cada $B \in \mathcal{R}$ e $s, t \geq 0$,

$$P(X_{t+s} \in B|\mathcal{F}_t) = P(X_{s+t} \in B|\sigma(X_t)) \tag{2.15}$$

Vejamos que a propriedade de Markov é satisfeita por todos processos de Lévy.

Lema 2.7. Se $f : \mathbb{R}^2 \to \mathbb{R}$ é mensurável, X é independente de $\mathcal{F}$ e Y é $\mathcal{F}$-mensurável, então,

$$E[f(X,Y)|\mathcal{F}] = E[f(X,Y)|Y]$$

Demonstração. Por definição, $E[f(X,Y)]$ é $\sigma(Y)$-mensurável e, portanto, $\mathcal{F}$-mensurável. Basta, então, mostrarmos que para todo $A \in \mathcal{F}$,

$$E[1_A E[f(X,Y)|Y]] = E[1_A f(X,Y)].$$

Note que,

$$\begin{aligned} E[1_A E[f(X,Y)|Y]] &= E[E[1_A E[f(X,Y)|Y]|Y]] \\ &= E[E[E[1_A|Y]f(X,Y)|Y]] \\ &= E[E[1_A|Y]f(X,Y)] \end{aligned}$$

Afirmação: Se X é independente de $\mathcal{F}$, então,

$$E[f(X,Y)E[1_A|Y]] = E[1_A f(X,Y)].$$

De fato, segue dos seguintes casos para a função f,

$$\text{indicadora} \rightarrow \text{função simples} \rightarrow \text{não negativa} \rightarrow \text{integrável}$$

Seja $f(X,Y) = 1_D(X)1_B(Y)$. Assim,

$$\begin{aligned}
E[1_D(X)1_B(Y)E[1_A|Y]] &= E[1_D(X)E[1_A1_B(Y)|Y]] \\
&= E[1_D(X)]E[E[1_A1_B(Y)|Y]] \\
&= E[1_D(X)]E[1_A1_B(Y)] \\
&= E[1_D(X)1_B(Y)1_A] = E[f(X,Y)1_A]
\end{aligned}$$

Este resultado se estende para f simples por linearidade. Se f é não negativa, aproximamos por funções simples da forma

$$f(x,y) = \sum_{i=1}^{n} a_i 1_{\{A_i \times B_i\}}(x,y) = \sum_{i=1}^{n} a_i 1_{A_i}(x) 1_{B_i}(y)$$

Obtendo o resultado pelo teorema da convergência monótona. Para provar o resultado em geral, decompomos f em parte positiva e negativa. $\square$

Aplicando o lema 2.7 à função $f(X,Y) = 1_{\{X+Y \in B\}}$, obtemos

$$E[1_{\{X+Y\in B\}}|\mathcal{F}] = E[1_{\{X+Y\in B\}}|Y].$$

Ou seja,

$$P(X+Y \in B|\mathcal{F}) = P(X+Y \in B|Y) \tag{2.16}$$

Lema 2.8. Se X é independente de $\mathcal{F}$, então X é independente de $\mathcal{F}$-aumentada por P.

Demonstração. Relembre que, por definição, X é independente de $\mathcal{F}$ se para todo B $\mathcal{F}$-mensurável e todo $A \in \mathcal{B}(\mathbb{R})$,

$$P((X \in A) \cap B) = P(X \in A)P(B).$$

Seja $\overline{\mathcal{F}}$ a σ-álgebra aumentada. Se $C \in \overline{\mathcal{F}}$, então $C = B \cup D$, onde $P(D) = 0$ e $B \in \mathcal{F}$.

Observe que

$$\begin{aligned}
P((X \in A) \cap C) &= P((X \in A) \cap (B \cup D)) \\
&= P(((X \in A) \cap B) \cup ((X \in A) \cap D)) \\
&= P((X \in A) \cap B) \\
&= P(X \in A)P(B) \\
&= P(X \in A)P(B \cup D) = P(X \in A)P(C).
\end{aligned}$$

□

Pelo lema 2.8, dado um processo de Lévy $X = \{X_t : t \geq 0\}$, $X_{t+s} - X_t$ é independente de $\mathcal{F}_t$, a σ-álgebra aumentada de $\sigma(X_u, u \leq t)$, para todos $s, t \geq 0$.

Teorema 2.9. *Se X é um processo de Lévy, então X possui a propriedade de Markov.*

Demonstração. Sejam $s, t \geq 0$. Então,

$$X_{t+s} = X_{t+s} - X_t + X_t$$

Faça $Z = X_{t+s} - X_t$ e $W = X_t$. Pelo que vimos, Z é independente de $\mathcal{F}_t$ e W é $\mathcal{F}_t$-mensurável. Pela equação (2.16),

$$P(Z + W \in B|\mathcal{F}_t) = P(Z + W \in B|\sigma(W))$$

isto é,

$$P(X_{t+s} \in B|\mathcal{F}_t) = P(X_{t+s} \in B|\sigma(X_t))$$

□

Definição 2.6. Uma variável aleatória τ é dita um **tempo de parada** se

$$\{\tau \leq t\} \in \mathcal{F}_t,$$

para todo $t \geq 0$. Para qualquer tempo de parada τ,

$$\{\tau < t\} = \bigcup_{n \geq 1} \{\tau \leq t - 1/n\} \in \bigcup_{n \geq 1} \mathcal{F}_{t-1/n} \subset \mathcal{F}_t$$

Associada a um tempo de parada τ está a sigma-álgebra

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t, \text{para todo } t \geq 0\}.$$

Definição 2.7. Dizemos que o processo X satisfaz a propriedade **forte de Markov** se, para cada tempo de parada τ,

$$P(X_{\tau+s} \in B|\mathcal{F}_\tau) = P(X_{\tau+s} \in B|\sigma(X_\tau))$$

em $\{\tau < \infty\}$.

Nosso próximo objetivo é mostrar que todo processo de Lévy satisfaz a propriedade forte de Markov, para isso, utilizaremos três lemas auxiliares.

Lema 2.10. (Dispositivo de Cramer-Wold). Sejam X e Y vetores aleatórios em $\mathbb{R}^n$. Se definirmos a função característica de X como

$$\varphi_X(t) = E[e^{i\langle t,X\rangle}],\ t \in \mathbb{R}^n$$

Então, $X =_d Y$ se, e somente, se $\varphi_X(t) = \varphi_Y(t)$, para todo $t \in \mathbb{R}^n$.

Lema 2.11. Seja X uma variável aleatória e $\mathcal{F}$ uma sigma-álgebra. X é independente de $\mathcal{F}$ se, e somente, se para todo $H \in \mathcal{F}$, para todo $t \in \mathbb{R}$,

$$E[e^{itX}; H] = E[e^{itX}]P(H) \tag{2.17}$$

Demonstração. É claro que se X é independente de $\mathcal{F}$, vale a última igualdade. Reciprocamente, suponha que (2.17) é satisfeita. Seja $P_1, P_2, \ldots$ uma sequência de polinômios trigonométricos convergindo pontualmente para 1_A (Obtemos tal sequência usando o teorema de Stone-Weierstrass nos compactos $I_n = [-n, n]$, observando que $\mathbb{R} = \cup_{n=1}^{\infty} I_n$ junto com o fato de que as funções contínuas são densas no conjuntos das funções integráveis). Daí, como 1_A é limitada, (P_n) é limitada. Além disso, $P_n \circ X \to 1_{\{X\in A\}}$, pontualmente. Pelo teorema da convergência dominada,

$$E[P_n \circ X; H] \to E[1_{\{X\in A\}}; H].$$

Por outro lado, temos para todo n,

$$E[P_n \circ X; H] = E[P_n \circ X]P(H).$$

Fazendo $n \to \infty$, obtemos

$$E[1_{\{X\in A\}}; H] = E[1_{\{X\in A\}}]P(H).$$

Ou seja,

$$P((X \in A) \cap H) = P(X \in A)P(H)$$

$\square$

Lema 2.12. Sejam X_1 e X_2 variáveis aleatórias com funções características φ_1 e φ_2, respectivamente. X_1 e X_2 são independentes se, e somente se, $X_1 + X_2$ possui função característica $\varphi_1(t)\varphi_2(t)$.

Demonstração.

$$E[e^{it(X_1+X_2)}] = E[e^{itX_1}e^{itX_2}] = E[e^{itX_1}]E[e^{itX_2}]$$

pois e^{itX_1} e e^{itX_2} são independentes. Para a recíproca, utilizamos o mesmo método do lema 2.11, usando polinômios trigonométricos para aproximar funções indicadoras. $\square$

Teorema 2.13. *Suponha que τ é um tempo de parada. Defina em $\{\tau < \infty\}$ o processo $\widetilde{X} = \{\widetilde{X}_t : t \geq 0\}$ onde*

$$\widetilde{X}_t = X_{\tau+t} - X_\tau,\ t \geq 0.$$

Então, sob o evento $\tau < \infty$, o processo $\widetilde{X}$ é independente de $\mathcal{F}_\tau$, possui a mesma lei de X e, portanto, em particular, é um processo de Lévy.

Demonstração. $\widetilde{X}$ é claramente contínua à direita com limite à esquerda. Precisamos mostrar que $\widetilde{X}$ possui incrementos estacionários e independentes. Se provarmos que para todo $n \in \mathbb{N}$, $0 \leq s_1 \leq t_1 \leq \cdots \leq s_n \leq t_n < \infty$, $H \in \mathcal{F}_\tau$ e $\theta_1, \ldots, \theta_n \in \mathbb{R}$,

$$E\left(\prod_{i=1}^{n} e^{i\theta_i(X_{\tau+t_i} - X_{\tau+s_i})}; H \cap \{\tau < \infty\}\right) = \prod_{i=1}^{n} e^{-\psi(\theta_i)(t_i - s_i)} P(H \cap \{\tau < \infty\}) \tag{2.18}$$

onde ψ é o expoente característico de X. Então, o lema 2.12 implica que $\widetilde{X}$ possui incrementos independentes, o lema 2.10 implica $\widetilde{X}$ com incrementos estacionários e mesma lei de X e, por fim, o lema 2.11 garante que $\widetilde{X}$ é independente de $\mathcal{F}_\tau$. Nos resta verificar (2.18). Para isso, defina a sequência de tempos de parada $\{\tau^{(n)} : n \geq 1\}$ por

$$\tau^{(n)} = \begin{cases} k2^{-n}, & \text{se } (k-1)2^{-n} < \tau \leq k2^{-n}, k = 1, 2, \ldots \\ 0, & \text{se } \tau = 0 \\ \infty, & \text{se } \tau = \infty \end{cases}$$

Incrementos estacionários e independentes junto com o fato que $H \cap \{\tau^n = k2^{-n}\} \in$

$\mathcal{F}_{k2^{-n}}$ nos permite escrever

$$\begin{aligned}
& E\left(\prod_{i=1}^{n} e^{i\theta_i(X_{\tau^{(n)}-t_i}-X_{\tau^{(n)}+s_i})}; H\cap\{\tau^{(n)}<\infty\}\right)\\
= & \sum_{k\geq 0} E\left(\prod_{i=1}^{n} e^{i\theta_i(X_{\tau^{(n)}-t_i}-X_{\tau^{(n)}+s_i})}; H\cap\{\tau^{(n)}=k2^{-n}\}\right)\\
= & \sum_{k\geq 0} E\left[E\left(\prod_{i=1}^{n} e^{i\theta_i(X_{k2^{-n}+t_i}-X_{k2^{-n}+s_i})}|\mathcal{F}_{k2^{-n}}\right); H\cap\{\tau^{(n)}=k2^{-n}\}\right]\\
= & \sum_{k\geq 0}\prod_{i=1}^{n} E(e^{i\theta_i X_{t_i-s_i}})P(H\cap\{\tau^{(n)}=k2^{-n}\})\\
= & \prod_{i=1}^{n} e^{-\psi(\theta_i)(t_i-s_i)}P(H\cap\{\tau^{n}<\infty\})
\end{aligned}$$

A continuidade à direita de X nos garante que $\tau^n \downarrow \tau$ em $\{\tau<\infty\}$, quando $n\to\infty$. Portanto, $X_{\tau^{(n)}+s}\to X_{\tau+s}$ quase certamente em $\tau<\infty$, para todo $s\geq 0$, quando $n\to\infty$. Segue do teorema da convergência dominada que

$$\begin{aligned}
& E\left(\prod_{i=1}^{n} e^{i\theta_i(X_{\tau-t_i}-X_{\tau}+s_i)}; H\cap\{\tau<\infty\}\right)\\
= & \lim_{n\to\infty} E\left(\prod_{i=1}^{n} e^{i\theta_i(X_{\tau^{(n)}-t_i}-X_{\tau^{(n)}+s_i})}; H\cap\{\tau^{(n)}<\infty\}\right)\\
= & \lim_{n\to\infty}\prod_{i=1}^{n} e^{-\psi(\theta_i)(t_i-s_i)}P(H\cap\{\tau^{n}<\infty\})\\
= & \prod_{i=1}^{n} e^{-\psi(\theta_i)(t_i-s_i)}P(H\cap\{\tau<\infty\}).
\end{aligned}$$

Como queríamos. □

Corolário 2.14. Todo processo de Lévy é um processo forte de Markov.

Demonstração. Análoga à demontração do teorema 2.9. □

Corolário 2.15. Em todo processo de Lévy, a filtração aumentada $\mathcal{F}_{t^+}$, isto é, a filtração que torna $\mathcal{F}_t$ completa com relação aos conjuntos de medida nula de $P|\mathcal{F}_t$, é uma filtração contínua à direita, ou seja,

$$\mathcal{F}_{t^+}=\bigcap_{s>t}\mathcal{F}_s,\ \text{ e vale }\ \mathcal{F}_t=\mathcal{F}_{t^+}.$$

Demonstração. A proposição 7.7 de [7] (página 90) nos diz que todo processo forte de Markov é tal que a filtração aumentada é contínua à direita. O resultado segue do corolário 2.14. □

Definição 2.8. Uma variável aleatória τ é dita um *tempo de opcional* se

$$\{\tau < t\} \in \mathcal{F}_t.$$

Associada a um tempo opcional τ está a sigma-álgebra

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau < t\} \in \mathcal{F}_t, \text{para todo } t \geq 0\}.$$

Proposição 2.16. Numa filtração $\mathcal{F}_t$ contínua à direita, uma variável aleatória τ é tempo de parada se, e somente se, é tempo opcional.

Demonstração. Se τ é tempo de parada, para todo $t \geq 0$

$$\{\tau < t\} = \bigcup_{n \geq 1} \{\tau \leq t - 1/n\} \in \bigcup_{n \geq 1} \mathcal{F}_{t-1/n} \subset \mathcal{F}_t.$$

Reciprocamente, se τ é tempo opcional, isto é, $\{\tau < t\} \in \mathcal{F}_t$, para todo $t \geq 0$, usando a continuidade à direita da filtração, obtemos

$$\{\tau \leq t\} = \bigcap_{n \geq 1} \{\tau < t + 1/n\} \in \bigcap_{n \geq 1} \mathcal{F}_{t+1/n} = \mathcal{F}_{t^+} = \mathcal{F}_t.$$

□

Capítulo 3

Modelo de Armadilha

3.1 Hipóteses e processo relógio

Seja $\mathcal{X} = \{\mathcal{X}_t : t \geq 0, \mathcal{X}_0 = 0\}$ um passeio aleatório a tempo contínuo com o passeio esqueleto $X = \{X_i, i \in \mathbb{N} \cup \{0\}, X_0 = 0\}$ satisfazendo a seguinte representação estocástica:

$$X_n = \sum_{j=1}^{n} \epsilon_j \tag{3.1}$$

para $n > 0$, onde $\{\epsilon_j, j \in \mathbb{N}\}$ é uma sequência *i.i.d* de variáveis aletórias discretas na bacia de atração de uma lei estável com índice $\beta \in (1, 2]$. A sequência em (3.1) é chamado um passeio aleatório $\beta-$estável; este processo também é conhecido como passeio aleatório de longo alcance. Iremos asssumir que $E(\epsilon_1) = 0$, e $E(e^{(it\epsilon_1)}) = 1$ se, e somente se t é mutiplo de 2π.

A hipótese de estabilidade implica que existe $c^- > 0$ e $c^+ > 0$ e uma função de variação lenta no infinito $h(\cdot)$ tal que

$$P(\epsilon_1 < -x) \sim x^{-\beta}(c^- + o(1))h(x)$$

e

$$P(\epsilon_1 > x) \sim x^{-\beta}(c^+ + o(1))h(x),$$

lembre que $f_1(x) \sim f_2(x)$ significa $\lim_{x\to\infty} \frac{f_1(x)}{f_2(x)} = 1$. Segue que existe um função de variação lenta positiva $v(\cdot)$ tal que

$$h(n^{1/\beta}v(n))v^{-\beta}(n) \to 1, \quad \text{para } \beta \in (1, 2)$$

e

$$H(n^{1/2}v(n))v^{-2}(n) \to 2(c^+ + c^-), \quad \text{para } \beta = 2,$$

quando $n \to \infty$, onde $H(z) = \int_{-z}^{z} x^2 dF(x)$.

Seja $L(n,x) = \sum_{i=0}^{n} 1\{X_i = x\}$ o tempo local (tempo de ocupação) do passeio aleatório X, isto é, o número de vezes que X visita o ponto x até o tempo $n \in \mathbb{N} \cup \{0\}$. Defina as sequências

$$d_n = n^{1/\beta}v(n), \;\; r_n = d_n^{-1}, \tag{3.2}$$

e o tempo local reescalado e o passeio aleatório esqueleto reescalado

$$\phi_n(t,x) = r_n^{-1}L([nt],[xd_n]), \;\; Z_t^{(n)} = d_n^{-1}X_{[nt]}, \tag{3.3}$$

para $t \in [0,\infty)$ e $x \in \mathbb{R}$. É conhecido que o processo $Z_t^{(n)}$ converge fracamente, no espaço $D([0,\infty),\mathbb{R})$ munido da topologia J_1, para Z_t, que é um processo estável contínuo à direita com função característica dada por

$$E(e^{isZ_t}) = \exp\{-ct|s|^{-\beta}[1 + iqsgn(s)]\}, \tag{3.4}$$

onde $c = -\Gamma(2-\beta)(\beta-1)^{-1}\cos(\pi\beta/2)$ e $q = (c^- - c^+)\tan(\pi\beta/2)/(c^+ + c^-)$. Para o processo Z_t existe o tempo local, isto é, a função $\phi(t,x)$ que é conjuntamente contínua com probabilidade um, que satisfaz

$$\mu_L(s : Z_s \in A, 0 \le s \le t) = \int_A \phi(t,x)dx,$$

para qualquer boreliano A, onde μ_L denota a medida de Lebesgue.

Para definir completamente nosso modelo de armadilha $\mathcal{X}_t$, precisamos definir suas taxas de saltos associadas. Seja $\tau = \{\tau_i : i \in \mathbb{Z}\}$ uma família de variáveis aleatórias $i.i.d$ positivas na bacia de atração de uma lei estável com índice $\alpha \in (0,2]$. Mais especificamente, para $\alpha \in (0,1)$, assumimos que a distribuição de τ_0 satisfaz a condição

$$P(\tau_0 > x) \sim x^{-\alpha}(\kappa^+ + o(1))s(x), \tag{3.5}$$

com $x > 0$, onde $\kappa^+ > 0$ e $s(\cdot)$ é uma função positiva de variação lenta. Segue que existe uma função positiva de variação lenta $w(\cdot)$ tal que

$$s(d_n^{1/\alpha}w(n))w^{-\alpha}(n) \to 1, \tag{3.6}$$

quando $n \to \infty$. Para $\alpha \in [1,2]$, supomos por simplicidade que a distribuição de τ_0 é dada por

$$P(\tau_0 > x) = x^{-\alpha}, \tag{3.7}$$

para $x \geq 1$. Supomos que τ é independente de X. Tomamos a taxa de salto de $\mathcal{X}$ em $i \in \mathbb{Z}$ dada por τ_i^{-1}. Desta forma, o passeio aleatório $\beta-$estável X (de tempo discreto) dado em (3.1) e a família τ satisfazendo (3.5) e (3.7) para α pertencendo aos intervalos $(0,1)$ e $[1,2]$, respectivamente, definem completamente o modelo de armadilha $\mathcal{X}_t$ que iremos considerar neste trabalho.

Com as definições acima, estamos prontos para intoduzir o processo relógio. Tal processo é definido por

$$C_t = \sum_{i=0}^{[t]} \tau_{X_i} T_i, \tag{3.8}$$

onde $\{T_i : i \in \mathbb{N} \cup \{0\}\}$ é uma sequência de variáveis exponenciais $i.i.d$ com média 1 independente de X e τ.

Temos que o processo relógio (3.8) é igual em distribuição ao processo

$$\bar{C}_t = \sum_{i \in \mathbb{Z}} \tau_i \sum_{j=1}^{L([t],i)} E_{ij}, \tag{3.9}$$

onde $\{E_{ij} : i \in \mathbb{Z}, j \in \mathbb{N}\}$ é um família de variáveis aleatórias exponenciais com média 1 e independente de todas variáveis aleatórias definidas anteriormente; definimos $\sum_{j-1}^{0} E_{ij} = 0$.

Para nosso propósito neste trabalho, o limite de escala do processo relógio desempenha um papel importante. Portanto, apresentaremos a ordem correta de magnitude deste processo, que depende de β e α. Vamos começar com o caso $\alpha \in [1,2]$. Defina uma sequência de números reais positivos $\{c_n : n \in \mathbb{N}\}$ por

$$c_n = \begin{cases} n \log n & se \quad \alpha = 1, \\ n & se \quad \alpha \in (1,2]. \end{cases}$$

Para $\alpha \in [1,2]$, definimos o processo relógio reescalado $\tilde{C}_t^{(n)}$ como

$$\tilde{C}_t^{(n)} = c_n^{-1} \bar{C}_{nt}, \ t \geq 0, \ n \in \mathbb{N}. \tag{3.10}$$

Iremos mostrar mais adiante que $\tilde{C}_t^{(n)}$ converge em probabiidade (na topologia uniforme) para um determinado limite quando $n \to \infty$. Mais especificamente, o limite é igual a $tE(\tau_0)$ e t para os casos $\alpha \in (1,2]$ e $\alpha = 1$, respectivamente. Para o caso $\alpha \in (0,1)$, é conveniente encontrar uma versão do processo relógio baseado num acoplamento de taxas aleatórias reescaladas, que discutiremos a seguir.

3.2 O acoplamento para o caso $\alpha \in (0, 1)$

Nesta seção apresentaremos uma versão útil para o ambiente aleatório τ onde a reescalação converge fortemente. Com esta versão de τ, definimos uma outra versão para o processo relógio. Nossa prova do limite de escala de $\mathcal{X}_t$ será baseada nesta versão particular do processo relógio. Primeiro, apresentamos a versão reescalada do processo (3.9), que é dada por

$$\bar{C}_t^{(n)} \equiv a_n^{-1} C_n^t, \tag{3.11}$$

onde, para cada $n \in \mathbb{N}$, definimos

$$a_n = \tau_n b_n \tag{3.12}$$

e

$$b_n = d_n^{1/\alpha} w(n) \tag{3.13}$$

Iremos agora mudar o processo acima para uma versão mais conveniente como feito em Fontes, Isopi e Newman. Seja $\{V_x : x \in \mathbb{R}\}$ um processo $\alpha-$estável bilateral independente de $\{E_{ij} : i \in \mathbb{Z}, j \in \mathbb{N}\}$ e $\{L([t], i) : i \in \mathbb{Z}, t \geq 0\}$. Considere a função $G : [0, \infty) \to [0, \infty)$ satisfazendo $P(V_1 > G(y)) = P(\tau_0 > y)$, para todo $y > 0$. Assim, defina a função $g_n : [0, \infty) \to [0, \infty)$ indexada em $n \in \mathbb{N}$ por $g_n(y) = b_n^{-1} G^{-1}(d_n^{1/\alpha} y)$.

Definiremos agora nossa versão conveniente das taxas aleatórias reescaladas $\{\tau_i : i \in \mathbb{Z}\}$ por

$$\tau_x^{(n)} \equiv b_n g_n(V_{x+d_n^{-1}} - V_x), x \in d_n^{-1}\mathbb{Z}$$

Observe que $\{\tau_x^{(n)} : x \in d_n^{-1}\mathbb{Z}\}$ possui a mesma distribuição de $\{\tau_i : i \in \mathbb{Z}\}$. Portanto, segue que o processo (3.11) segue a mesma lei que o seguinte processo:

$$\bar{C}_t^{(n)} \equiv \sum_{x \in d_n^{-1} \in \mathbb{Z}} g_n(V_{x+d_n^{-1}} - V_x) \phi_n(t, x) \bar{T}_{xd_n}(nt), \ t \geq 0, \tag{3.14}$$

onde $\bar{T}_i(t) = \sum_{j=1}^{L([t],i)} E_{ij}/L([t], i)$, se o sítio i é visitado pelo passeio aleatório X até o tempo $[t]$ e igual a 1 em outro caso, e $\phi_n(t, x)$ é o tempo local reescalado definido em (3.3). A versão (3.14) do processo relógio irá desempenhar um importante papel neste trabalho, como iremos ver logo mais. Na seção seguinte iremos mostrar que quando $\alpha < 1$ o processo relógio reescalado converge fracamente para um processo não trivial.

3.3 Limite de escala

Começaremos esta seção introduzindo o limite de escala do nosso modelo de armadilha $\mathcal{X}_t$ quando $\alpha \in (0,1)$. Chamaremos este limite de escala de processo quase-estável.

Definição 3.1. (Processo quase-estável) Seja $\mu : \mathbb{R} \to \mathbb{R}^+$ uma função não-decrescente e $\phi(x,t)$ o tempo local de um processo $\beta-$estável Z_t (com $\beta \in (1,2]$) como definido na seção (3.2). Definimos o processo quase-estável por $Z_{S_t^{-1}}$, onde S_t^{-1} é a função inversa de $S_t = \int_{-\infty}^{\infty} \phi(t,x)d\mu(x)$ (S_t é uma função estritamente crescente e contínua com probabilidade 1).

Observação 3.1. De acordo com a definição acima, os processos quase-estáveis são processos $\beta-$estáveis de mudança de tempo pelo inverso do processo relógio, que envolve o tempo local do processo $\beta-$estável e uma medida induzida por μ. Estes processos foram introduzidos por Stone (1963) e são processos estacionários forte de Markov.

Observação 3.2. Seja $V = \{V_x : x \in \mathbb{R}, V_0 = 0\}$ um processo $\alpha-$estável bilateral com $\alpha \in (0,1)$. Neste trabalho, processos estáveis de nosso interesse será como o da definição (3.1) com V no lugar de μ. A menos que mencionemos o contrário, de agora em diante consideraremos esta substituição.

Denote por $D([0,T],\mathbb{R})$ e $D([0,T],\mathbb{R}^+)$ o espaço das funções reais e reais não negativas, contínuas à direita com limite à esquerda em $[0,T]$, onde $T > 0$. Além disso, sejam d_T e u_T^+ as métricas J_1-Skorohod em $D([0,T],\mathbb{R})$ e a uniforme em $D([0,T],\mathbb{R}^+)$, respectivamente, $d = \sum_{n=1}^{\infty} 2^{-n}\min(d_n,1)$ e $u^+ = \sum_{n=1}^{\infty} 2^{-n}\min(u_n^+,1)$. Denotamos a topologia uniforme por U. O próximo teorema nos dá um limite de escala de $\mathcal{X}_t$ onde $\alpha \in (0,1)$ e é um dos principais resultados neste trabalho.

Teorema 3.1. *Para $\alpha \in (0,1)$ existem sequências de números reais positivos $\{a_n : n \in \mathbb{N}\}$ e $\{d_n : n \in \mathbb{N}\}$ tais que*

$$\mathcal{X}_t^{(n)} \equiv d_n^{-1}\mathcal{X}_{a_n t} \Rightarrow Z_{S_t^{-1}},$$

em $D([0,\infty),\mathbb{R})$ munido da topologia J_1-Skorohod, quando $n \to \infty$, onde o processo limite $Z_{S_t^{-1}}$ é o introduzido na definição na definição (3.1).

Observação 3.3. Uma expressão explícita para as sequências $\{a_n : n \in \mathbb{N}\}$ e $\{d_n : n \in \mathbb{N}\}$ são dadas em (3.12) e (3.2), respectivamente. Estas sequências são a ordem de magnitude do processo relógio e o passeio aleatório esqueleto, respectivamente.

Observação 3.4. Ao longo deste capítulo aparecem algumas quantidades aleatórias que dependem de V. Denotamos a probabilidade e a esperança condicionais (dado V) por $P(\cdot) = P(\cdot|V)$ e $E(\cdot) = E(\cdot|V)$, respectivamente.

Afim de provar o teorema (3.1), obteremos primeiro o limite de escala conjuntamente do processo relógio e do passeio aleatório esqueleto X, que é fornecido no próximo teorema. O maior trabalho está em mostrar a convergência conjunta das distribuições finito-dimensionais do processo relógio e do passeio aleatório esqueleto.

Teorema 3.2. *Para quase toda trajetória de V, $\{(Z_t^{(n)}, \tilde{C}_t^{(n)}) : t \geq 0\}$ converge para $\{(Z_t, S_t) : t \geq 0\}$ em distribuição quando $n \to \infty$ em $D([0,\infty),\mathbb{R}) \times D([0,\infty),\mathbb{R}^+)$ munido da topologia produto $J_1 \times U$.*

Observação 3.5. O processo limite do nosso processo relógio reescalado é quase certamente contínuo. A topologia adequada em que ocorre a convergência é a uniforme. Mesmo assim, nossa prova verifica o critério de convergência na topologia $J_1 \times M_1$ para o processo bivariado dado no teorema acima. Como S_t é quase certamente contínuo, a convergência na topologia $J_1 \times U$ segue.

Demonstração. Primeiro, troque $D([0,\infty),\mathbb{R})$ e $D([0,\infty),\mathbb{R}^+)$ por $D([0,T),\mathbb{R})$ e $D([0,T),\mathbb{R}^+)$ munidos das métricas d_T e u_T^+, respectivamente, com $T > 0$ arbitrário. Iremos provar a convergência das distribuições finito-dimensionais de $\{Z_t^{(n)}, \tilde{C}_t^{(n)}\}$ e mostrar que esta sequência é tight. Estes resultados nos levam à prova do teorema.

Convergência das distribuições finito-dimensionais

Para obter a convergência das distribuições finito-dimensionais do processo bivariado $\{(Z_t^{(n)}, \tilde{C}_t^{(n)}) : t \geq 0\}$, usaremos o teorema 1.1 de Borodin (1985) [3], que nos garante que é possível construir, sobre algum espaço de probabilidade, processos $\{Z_t'^{(n)} : t \in [0,T]\}$ e $\{Z_t' : t \in [0,T]\}$ tais que:

(i) Suas distribuições finito-dimensionais coincidem com as de $\{Z_t^{(n)} : t \in [0,T]\}$ e $\{Z_t : t \in [0,T]\}$, respectivamente;

(ii) $Z_t'^{(n)}$ converge quase certamente para Z_t' em $D([0,T],\mathbb{R})$ munido da topologia J_1-Skorohod;

(iii) O tempo local $\phi_n'(t,x)$ e $\phi(t,x)$ definidos com respeito aos processos $Z_t'^{(n)}$ e Z_t', respectivamente, para todo $T > 0$ e $\xi > 0$, satisfazem:

$$\lim_{n\to\infty} P\Big(\sup_{(t,x)\in[0,T]\times\mathbb{R}} |\phi_n'(t,x) - \phi'(t,x)| > \xi\Big) = 0. \tag{3.15}$$

Portanto, usando o teorema 1.1 de [3], tomamos o processo $Z_t^{(n)}$ como descrito acima e mostraremos convergência das distribuições finito-dimensionais de $\{(Z_t'^{(n)}, \tilde{C}_t^{(n)}) : t \geq 0\}$, onde $\tilde{C}_t'^{(n)}$ é definido como em (3.14), mas com as quantidades que dependem de $Z_t^{(n)}$ definidas com respeito a $Z_t'^{(n)}$. A quantidade $\mathcal{A}$ que depende de $Z_t^{(n)}$ será denotada por $\mathcal{A}'$ enquanto esta depende de $Z_t'^{(n)}$.

Agora definiremos o conjunto das armadilhas profundas. Para um $\delta > 0$ arbitrário, considere

$$\mathcal{T}_\delta = \{x \in \mathbb{R} : V_x - V_{x^-} > \delta\} = \{x_j : j \in \mathbb{Z}\} = \{\ldots < x_{-1} < x_0 < x_1 < \ldots\},$$

$x_j^{(n)} = d_n^{-1}[d_n x_j]$, para $i \in \mathbb{Z}$, e

$$\mathcal{T}_\delta^{(n)} = \{x_j^{(n)} : j \in \mathbb{Z}\}$$

Mostraremos primeiro que

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \sum_{x \in d_n^{-1}\mathbb{Z} \cap (\mathcal{T}_\delta^{(n)})^C} g_n(V_{x+d_n^{-1}} - V_x)\phi_n'(t,x)\bar{T}_{xd_n}'(nt) = 0 \tag{3.16}$$

em probabilidade, para todo $t \in [0,T]$, isto é, os termos do processo relógio reescalado que estão fora da armadilha profunda possuem uma contribuição despresível para o processo limite quando $n \to \infty$ e $\delta \downarrow 0$.

Observação 3.6. Seja $A_t^{(n,\delta)}$ uma quantidade aleatória dependendo de δ, n e t. Abaixo, quando dizemos que $\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \mathcal{A}_t^{(n,\delta)} = 0$ em probabilidade, significa que para todo $\epsilon > 0$ fixado temos que $\lim_{\delta \downarrow 0} \limsup_{n \to \infty} P(|A_t^{(n,\delta)}| > \epsilon) = 0$

O teorema (1.1) de [6] nos diz que para todo $\epsilon > 0$, existe $A_\epsilon > 0$ tal que

$$\sup_n P\left(\sup_{t \in [0,T]} |Z_t'^{(n)}| > A_\epsilon\right) < \epsilon \quad \text{e} \quad P\left(\sup_{t \in [0,T]} |Z_t'| > A_\epsilon\right) < \epsilon \tag{3.17}$$

Seja $I_\epsilon = (A_\epsilon, A_\epsilon)$. Então, para todo $\zeta > 0$, usando o resultado acima e a desigualdade de Markov, segue que

$$P\left(\sum_{x\in d_n^{-1}\mathbb{Z}\cap(\mathcal{T}_\delta^{(n)})^c} g_n(V_{x+d_n^{-1}}-V_x)\phi'_n(t,x)\bar{T}'_{xd_n}(nt)>\zeta\right)\leq$$
$$P\left(\sum_{x\in d_n^{-1}\mathbb{Z}\cap(\mathcal{T}_\delta^{(n)})^c} g_n(V_{x+d_n^{-1}}-V_x)\phi'_n(t,x)\bar{T}'_{xd_n}(nt)>\zeta\right)+\epsilon\leq$$
$$\zeta^{-1}\sum_{x\in d_n^{-1}\mathbb{Z}\cap(\mathcal{T}_\delta^{(n)})^c} g_n(V_{x+d_n^{-1}}-V_x)E(\phi'_n(t,x))+\epsilon, \tag{3.18}$$

para quase toda trajetória de V. Para todo $M>0$ inteiro temos que

$$E(\phi_n^M(t,x))\leq E(\phi_n^M(t,0))$$

e

$$\lim_{n\to\infty} E(\phi_n^M(t,0))=t^{M(1-1/\beta)}z\Gamma(1-1/\beta)/\Gamma(2-1/\beta), \tag{3.19}$$

onde z é o valor da densidade de Z_1 em zero; veja [6] p. 328. De (3.19) e usando a igualdade das distribuições finito-dimensionais de $\phi_n(t,x)$ e $\phi'_n(t,x)$, que é válido para todos $x\in\mathbb{R}$, obtemos que a soma em (3.18) é limitada superiormente pelas constantes de tempo

$$\sum_{x\in d_n^{-1}\mathbb{Z}\cap(\mathcal{T}_\delta^{(n)})^c} g_n(V_{x+d_n^{-1}}-V_x).$$

Agora, argumentando como em [6], parágrafos (3.25) à (3.28), segue que $\lim_{\delta\downarrow}\limsup_{n\to\infty}$ dos termos acima se anulam em probabilidade. Isto e a arbitrariedade de ϵ nos dá (3.16).

Com o resultado acima, definimos o processo relógio restrito às armadilhas profundas:

$$\tilde{C}_t'^{(n,\delta)}=\sum_{i\geq 1} g_n(V_{x_i^{(n)}+d_n^{-1}}-V_{x_i}^{(n)})\phi'_n(t,x_i^{(n)})\bar{T}_{x_i^{(n)}d_n}(nt) \tag{3.20}$$

Seja $\epsilon>0$ e tome A_ϵ satisfazendo (3.17). Defina o conjunto

$$\Omega_\epsilon=\{\sup_{t\in[0,T]}|Z_t^{(i)}|\leq A_\epsilon, 1\leq i\leq n\}\cap\{\sup_{t\in[0,T]}|Z_t|\leq A_\epsilon\}$$

Sobre Ω_ϵ, o processo dado em (3.20) fica

$$\tilde{C}_t'^{(n,\delta,\epsilon)}=\sum_{i=-N_{\delta,\epsilon}}^{N_{\delta,\epsilon}} g_n(V_{x_i+d_n^{-1}}-V_{x_i^{(n)}})\phi'_n(t,x_i^{(n)})\bar{T}'_{x_i^{(n)}d_n}(nt),$$

onde $N_{\delta,\epsilon} = \max\{|j| \in \mathbb{N} : x_j \in I_\epsilon\} < \infty$ para quase todo V.

O resultado (3.15) e a continuidade (com probabilidade 1) do tempo local $\phi'(t,x)$ (com respeito ao espaço variável) implicam que $\phi'_n(t, x_i^{(n)})$ converge em probabilidade para $\phi'(t, x_i)$ uniformemente em $i \in [-N_{\delta,\epsilon}, N_{\delta,\epsilon}]$ quando $n \to \infty$. Além disso, como $L'([nt], [x_i^{(n)} d_n]) \xrightarrow{P} \infty$ quando $n \to \infty$ (isto segue da convergência em probabilidade de $\phi'(t, x_i^{(n)})$). Segue da Lei dos grandes números que $\bar{T}'_{x_i^{(n)} d_n}(nt) \xrightarrow{P} 1$ quando $n \to \infty$, para todo i. Da proposição (3.1) de [6], segue que

$$g_n(V_{x_i^{(n)} + d_n^{-1}} - V_{x_i^{(n)}}) \to V_{x_i} - V_{x_i^-}, \tag{3.21}$$

quando $n \to \infty$, para quase todo V e para todo i.

A convergência em probabilidade de $\phi'_n(t, x_i^{(n)})$ e $\bar{T}'_{x_i^{(n)} d_n}(nt)$ (uniformemente em $i \in [-N_{\delta,\epsilon}, N_{\delta,\epsilon}]$) discutida acima, o resultado (3.21) e o teorema da função contínua A.6 implicam que para quase todo V

$$(Z_t'^{(n)}, C_t'^{(n,\delta,\epsilon)}) \xrightarrow{P} (Z'_t, S_t'^{(\delta,\epsilon)}), \tag{3.22}$$

quando $n \to \infty$, onde $S_t'^{(\delta,\epsilon)} = \sum_{x \in \mathcal{T}_\delta \cap I_\epsilon} (V_x - V_{x^-}) \phi'(t,x)$. Além disso, temos que

$$\lim_{\delta \downarrow 0} \lim_{\epsilon \downarrow 0} S_t'^{(\delta,\epsilon)} = \int_{-\infty}^{\infty} \phi'(t,x) dV_x, \tag{3.23}$$

em probabilidade para quase todo V. Assim, a convergência das distribuições finito-dimensionais do processo $\{(Z_t^{(n)}, \tilde{C}_t^{(n)}) : t \in [0,T]\}$ para as de $\{(Z_t, S_t) : t \in [0,T]\}$ segue de (3.22) e (3.23).

Tightness

Daqui em diante, iremos denominar a topologia de Skorohod associada ao espaço D, estudada no capítulo 1, de topologia J_1-Skorohod ou simplesmente topologia J_1. A seguir, mostraremos que a sequência $\{(Z_t^{(n)}, \tilde{C}_t^{(n)}) : t \in [0,T]\}$ é tight em $D([0,T], \mathbb{R}) \times D([0,T], \mathbb{R}^+)$ munido da topologia produto $J_1 \times M_1$. É suficiente mostrar isto para cada coordenada. A primeira coordenada converge em distribuição, logo é tight. Só nos resta mostrar que $\{\tilde{C}_t^{(n)} : t \in [0,T]\}$ é tight em $D([0,T], \mathbb{R}^+)$ munido da topologia M_1 (e consequentemente com a topologia U - veja a observação (3.5)).

Como o processo $\tilde{C}_t^{(n)}$ é não decrescente, podemos checar facilmente que a função w_s usada na condição (ii) do teorema 12.12.3 de [8] é igual a zero. Com isso, e usando o fato que o processo é não-negativo, obtemos um critério de tightness para o teorema numa maneira compacta, usando limites, dado por

(i) $\lim_{c \to \infty} \limsup_{n \to \infty} P(\tilde{C}_T^{(n)} > c) = 0$;

(ii) Para cada $\xi > 0$, $\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} P(\max\{\bar{v}(C_n^*, 0, \epsilon), \bar{v}(C_n^*, T, \epsilon)\} > \xi) = 0$,

onde, para $t \in [0, T]$, temos

$$\bar{v}(C_n^*, t, \epsilon) = \sup_{\max\{0, t-\epsilon\} \leq t_1 \leq t_2 \leq \min\{t+\epsilon, T\}} \{\tilde{C}_{t_2}^{(n)} - \tilde{C}_{t_1}^{(n)}\}.$$

Tomando $\epsilon < T$, temos que, para $t = 0$ e $t = T$, a quantidade acima se reduz a

$$\bar{v}(C_n^*, 0, \epsilon) = \tilde{C}_\epsilon^{(n)} \tag{3.24}$$

e

$$\bar{v}(C_n^*, T, \epsilon) = \tilde{C}_T^{(n)} - \tilde{C}_{T-\epsilon}^{(n)} \tag{3.25}$$

respectivamente. Mostraremos agora que de fato o processo relógio reescalado satisfaz as condições (i) e (ii). Daqui em diante, usaremos a restrição $\epsilon < T$. A condição (i) segue da convergência das distribuições finito-dimensionais, isto é,

$$\lim_{c\to\infty} \lim_{n\to\infty} P(\tilde{C}_T^{(n)} > c) = \lim_{c\to\infty} P(S_T > c) = 0$$

Usando (3.24) e a convergência das distribuições finito-dimensionais, também obtemos que para cada $\xi > 0$ fixo

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} P(\bar{v}(C_n^*, 0, \epsilon) > \xi) = \lim_{\epsilon\downarrow 0} P(S_\epsilon > \xi)$$

Como S_ϵ converge para 0 em probabilidade, fazendo $\epsilon \downarrow 0$, obtemos

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} P(\bar{v}(C_n^*, 0, \epsilon) > \xi) = 0 \tag{3.26}$$

Analogamente ao caso anterior e usando (3.25), para cada $\xi > 0$ fixo, segue que

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} P(\bar{v}(C_n^*, T, \epsilon) > \xi) = \lim_{\epsilon\downarrow 0} P(S_T - S_{T-\epsilon} > \xi).$$

Agora, usando o fato que S_t é quase certamente contínuo, obtemos que $S_T - S_{T-\epsilon} \xrightarrow{q.c} 0$ quando $\epsilon \to 0$. Com isso, obtemos

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} P(\bar{v}(C_n^*, T, \epsilon) > \xi) = 0 \tag{3.27}$$

Os resultados (3.26) e (3.27) implicam que a condição (ii) é satisfeita e, portanto, a sequência $\tilde{C}_t^{(n)}$ é tight.

Prova do teorema 3.1

Começamos a prova notando que a seguinte representação estocástica é válida:

$$\{\mathcal{X}_t^{(n)} : t \geq 0\} \stackrel{d}{=} \{Z_{\tilde{C}_t^{(n)-1}} : t \geq 0\}$$

Iremos definir alguns subconjuntos de $D([0,\infty),\mathbb{R})$ (a definição dos subconjuntos de $D([0,\infty)\mathbb{R}^+)$ são feitos de modo análogo).

Seja $D_\uparrow([0,\infty),\mathbb{R})$ o espaço das funções não negativas em $D([0,\infty),\mathbb{R})$ que são não-decrescente. Aqui denotamos o espaço das funções contínuas que são estritamente crescentes por $C_{\uparrow\uparrow}([0,\infty),\mathbb{R})$. Denote por $D_u([0,\infty),\mathbb{R})$ o espaço das funções em $D([0,\infty),\mathbb{R})$ que são ilimitadas. Consequentemente $D_{u,\uparrow}([0,\infty),\mathbb{R}) = D_u([0,\infty),\mathbb{R}) \cap D_\uparrow([0,\infty),\mathbb{R})$.

Seguindo as ideias usadas na prova do teorema (3.2) e o resultado acima, podemos ver que as distribuições finito-dimensionais de $(Z_t^{(n)}, \tilde{C}_t^{(n)^{-1}})$ converge fracamente para as de (Z_t, S_t^{-1}) quando $n \to \infty$. Além disso, temos que $D_{u,\uparrow}([0,\infty),\mathbb{R}^+)$ munido da topologia uniforme é separável e completo, assim, usando a convergência fraca de $\tilde{C}_t^{(n)^{-1}}$ para S_t^{-1} (que segue do teorema (3.2) e do teorema da função contínua A.6) e o teorema de Prohorov, temos que $\tilde{C}_t^{(n)^{-1}}$ é tight. Portanto, $(Z_t^{(n)}, \tilde{C}_t^{(n)^{-1}}) \Rightarrow (Z_t, S_t^{-1})$ quando $n \to \infty$ em $D([0,\infty),\mathbb{R} \times D_{u,\uparrow}([0,\infty),\mathbb{R}^+)$ munido com a topologia produto $J_1 \times U$.

Agora usamos o teorema 13.2.2 de whitt (2002) [8], que estabelece que a composição de $D([0,\infty),\mathbb{R} \times D_{u,\uparrow}([0,\infty),\mathbb{R}^+)$ para $D([0,\infty),\mathbb{R})$ é contínua em $D([0,\infty),\mathbb{R} \times C_{\uparrow,\uparrow}([0,\infty),\mathbb{R}^+)$ munido da topologia J_1-Skorohod. Temos que (Z_t, S_t^{-1}) pertence a $D([0,\infty),\mathbb{R} \times C_{\uparrow,\uparrow}([0,\infty),\mathbb{R}^+)$: O teorema da função contínua A.6 e os resultados acima implicam que $\mathcal{X}_t \Rightarrow Z_{S_t}^{-1}$ em $D([0,\infty),\mathbb{R})$ munido da topologia J_1-Skorohod, assim, provamos o teorema.

□

Apêndice A

Resultados Básicos

A.1 Conceitos Básicos em Probabilidade

Seja X uma variável aleatória. Dizemos que X possui distribuição de *Bernoulli* com parâmetro p se $P(X = 1) = p$ e $P(X = 0) = 1 - p$. X possui distribuição *Binomial*(n, p) se

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

X possui distribuição de *Poisson* com parâmetro λ se

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} \quad \text{para } k = 0, 1, 2, \ldots.$$

Lema A.1. A soma de n variáveis aleatórias Bernoulli(p) independentes é uma variável aleatória Binomial(n, p).

Lema A.2. Sejam $X_1, X_2, \ldots$ variáveis aleatórias independentes. Se X_n possui distribuição Binomial(n, p_n) e $\lim_{n\to\infty} np_n = \lambda$, então $X_n \Rightarrow$ Poisson(λ), quando $n \to \infty$.

Lema A.3. Se $X_1, \ldots, X_n$ são independentes com $EX_i^2 < \infty$. Então

$$var(X_1 + \cdots + X_n) = var(X_1) + \cdots + var(X_n)$$

Lema A.4. Se $X > 0$ é uma variável aleatória com função de distribuição F, então seu k-ésimo momento é dado por

$$EX^k = \int_0^\infty kt^{k-1} P(X > t)dt$$

Teorema A.5. *(Teorema de seleção de Helly) Para toda sequência F_n de funções de distribuição, existe uma subsequência $F_{n(k)}$ e uma função contínua à direita e não* *decrescente F tal que* $\lim_{k\to\infty} F_{n(k)}(y) = F(y)$ *em todos os pontos de continuidade y de F.*

Teorema A.6. *(Teorema da Função Contínua) Seja g uma função mensurável e $D_g = \{x : g$ é descontínua em $x\}$. Se $X_n \Rightarrow X$ e $P(X \in D_g) = 0$, então $g(X_n) \Rightarrow g(X)$. Além disso, se g é limitada, então $Eg(X_n) \to Eg(X)$*

Sejam Ω um conjunto e $\mathcal{F}$ uma σ-álgebra de subconjuntos de Ω. Uma função μ, definida em $\mathcal{F}$ e com valores em $\mathbb{R}$, é uma *medida* se satisfaz os axiomas:

(i) $\mu(A) \geq \mu(\emptyset) = 0$ para todo $A \in \mathcal{F}$, e

(ii) Se $A_i \in \mathcal{F}$ é uma sequência contável de conjuntos disjuntos, então

$$\mu(\cup_i(A_i)) = \sum_i \mu(A_i)$$

Se $\mu(\Omega) = 1$, chamamos μ uma **medida de probabilidade**. Neste trabalho, medidas de probabilidade serão habitualmente denotadas por P. Ao par $(\Omega, \mathcal{F})$ chamamos *espaço mensurável*, enquanto a trinca $(\Omega, \mathcal{F}, P)$ denominamos *espaço de probabilidade*. Os subconjuntos que estão em $\mathcal{F}$ são denominados *eventos* e é somente a eles que se atribui probabilidade.

Teorema A.7. *Suponha $X, Y \geq 0$ ou $E|X|, E|Y| < \infty$.*

(a) $E(X + Y) = EX + EY$

(b) $E(aX + b) = aEX + b$

(c) Se $X \geq Y$, então $EX \geq EY$.

O teorema a seguir é fundamental no cálculo do valor esperado.

Teorema A.8. *(Mudança de Variável) Seja X um elemento aleatório de $(S, \mathcal{S})$, com distribuição μ, isto é, $\mu(A) = P(X \in A)$. Se f é uma função mensurável de $(S, \mathcal{S})$ para $(\mathbf{R}, \mathcal{R})$ tal que $f \geq 0$ ou $E|f(X)| < \infty$, então*

$$Ef(X) = \int_S f(y)\mu(dy) \tag{A.1}$$

Uma consequência do Teorema (A.8) é que podemos computar o valor esperado de funções de variáveis aleatórias calculando integrais sobre a reta real.

Definição A.1. Se k é um número inteiro positivo, então EX^k é chamado de **k-ésimo momento** de X. Se $EX^2 < \infty$, então a **variância** de X é definida por

$$\begin{aligned} var(X) &= E(X-\mu)^2 \\ &= EX^2 - 2\mu EX + \mu^2 \\ &= EX^2 - \mu^2. \end{aligned}$$

Da equação acima, é imediato que $var(X) \leq EX^2$. Como $E(aX+b) = aEX + b$ de (b) do Teorema (A.7), segue da definição que

$$\begin{aligned} var(aX+b) &= E(aX+b-E(aX+b))^2 \\ &= a^2E(X-EX)^2 \\ &= a^2 var(X). \end{aligned}$$

A.1.1 Independência

Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade. Os eventos $A_1, \ldots, A_n$ são **independentes** se para qualquer $I \subset \{1, \ldots, n\}$ temos

$$P(\cap_{i\in I} A_i) = \prod_{i\in I} P(A_i).$$

As variáveis aleatórias $X_1, \ldots, X_n$ são independentes se para todo $B_i \in \mathcal{R}$ com $i = 1, \ldots, n$ temos

$$P(\cap_{i=1}^n (X_i \in B_i)) = \prod_{i=1}^n P(X_i \in B_i).$$

Uma sequência infinita de objetos (eventos ou variáveis aleatórias) são independentes se qualquer subcoleção finita for independente. Dizemos que os eventos $A_1, \ldots, A_n$ são **independentes aos pares** se $P(A_i \cap A_j) = P(A_i)P(A_j)$ para todo $i \neq j$. Analogamente, as variáveis aleatórias $X_1, \ldots, X_n$ são ditas **independentes aos pares** se para todo $B_i \in \mathcal{R}$ com $i = 1, \ldots, n$ os eventos $A_1 = (X_1 \in B_1), \ldots, A_n = (X_n \in B_n)$ são independentes aos pares.

A.1.2 Teorema de Continuidade de Lévy

Teorema A.9. *(Teorema de Continuidade de Lévy) Sejam μ_n, $1 \leq n \leq \infty$, medidas de probabilidade com funções características ϕ_n. (i) Se $\mu_n \Rightarrow \mu_\infty$, então $\phi_n(t) \to \phi_\infty(t)$*

para todo t. (ii) Se $\phi_n(t)$ converge pontualmente para um limite $\phi(t)$ que é contínuo em 0*, então a sequência de distribuição μ_n é tight e converge fracamente para uma medida μ com função característica ϕ.*

A.1.3 Lei dos Grandes Números

Quando as variáveis aleatórias $X_1, X_2, \ldots$ são independentes e possuem a mesma distribuição dizemos que elas são *independentes e identicamente distribuídas*, abreviaremos por (*i.i.d.*).

Definição A.2. X_n converge para X **quase certamente** (*q.c.*) se $P(X_n \to X$ quando $n \to \infty) = 1$, isto é, o evento $A_0 = \{\omega : X_n(\omega) \to X(\omega)\}$ é de probabilidade 1.

Teorema A.10. *(Lei Forte dos Grandes Números) Se $X_1, X_2, \ldots$ são variáveis aleatórias independentes aos pares e identicamente distribuídas com $E|X_i| < \infty$, $EX_i = \mu$ e $S_n = X_1 + \cdots + X_n$, então*

$$\frac{S_n}{n} \to \mu \text{ } q.c. \quad quando \ n \to \infty.$$

Referências Bibliográficas

[1] Barreto-Souza, W.; Fontes, L. R. G. *Long range trap models on $\mathbb{Z}$ and quasistable processes.* Journal of Theoretical Probability, v. 28, n. 4, p. 1500-1519, 2015.

[2] Billingsley, P., *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.

[3] Borodin, A.N The asymptotic behavior of local times of recurrent random walks with infinite variance. *Theory of Probability and its Applications*,1985.

[4] Breiman, L. *Probability.* Addison-Wesley, Reading, MA, 1968.

[5] Durrett, R. *Probability: Theory and Examples*, Edition 4.1, Cambridge University, 2010.

[6] Fontes, L.R.G., Isopi, M., Newman, C.M.: *Random walks with strongly inhomogeneous rates and singular diffusions: convergence, localization and aging in one dimension.* Annals of Probability. 30, 579-604, 2002.

[7] Karatzas, I. & Shreve, S.E., *Brownian Motion and Stochastic Calculus*, Springer, 1991.

[8] Whitt, W.: *Stochastic-Process Limits.* Springer, New York, 2002.