UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO

MESTRADO EM MATEMÁTICA APLICADA À ECONOMIA E GESTÃO

A TEORIA DOS PROCESSOS PONTUAIS NA ESTIMAÇÃO EM MODELOS AUTOREGRESSIVOS DIRIGIDOS POR ERROS COM VARIÂNCIA INFINITA

Paula Cristina Martins dos Reis

Orientador: Prof.ª Doutora Luísa Canto e Castro Loura

Júri:

Doutor Carlos Alberto da Silva Ribeiro
Doutor Carlos Alberto dos Santos Braumann
Doutor Nuno Paulo de Arrobas Crato
Doutora Luísa Canto e Castro Loura

Setembro/2001

UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO

MESTRADO EM MATEMÁTICA APLICADA À ECONOMIA E GESTÃO

A TEORIA DOS PROCESSOS PONTUAIS NA ESTIMAÇÃO EM MODELOS AUTOREGRESSIVOS DIRIGIDOS POR ERROS COM VARIÂNCIA INFINITA

Paula Cristina Martins dos Reis

Orientador: Prof.ª Doutora Luísa Canto e Castro Loura

Júri:

Doutor Carlos Alberto da Silva Ribeiro

Doutor Carlos Alberto dos Santos Braumann

Doutor Nuno Paulo de Arrobas Crato

Doutora Luísa Canto e Castro Loura

Setembro/2001

Ao Afonso, ao Miguel e a ti Pedro.

A Teoria dos Processos Pontuais na Estimação em Modelos Autoregressivos dirigidos por Erros com Variância Infinita

Paula Cristina Martins dos Reis

Índice

Errata

Glossário de Termos e Abreviaturas

$\xrightarrow{d}$	converge em distribuição.
$\xrightarrow{p}$	converge em probabilidade.
$\Rightarrow$	converge fracamente.
$\xrightarrow{v}$	converge vagamente.
$\xrightarrow{q.c.}$	converge quase certamente.
$\stackrel{d}{=}$	igual em distribuição.
$i.i.d.$	independentes e identicamente distribuídas.
$i.o.$	para muitos e infinitos.
$q.c.$	quase certamente.
$F^{\leftarrow}$	inversa generalizada de F.
$L(\cdot)$	função de variação lenta em ∞.
$o_p(1)$	convergente em probabilidade para zero.
$O_{q.c.}(1)$	limitado quase certamente.
$\mathcal{B}(\mathbb{R})$	σ-álgebra dos boreleanos de $\mathbb{R}$.
$C_K^+(E)$	espaço das funções reais contínuas positivas definidas em E, com suporte compacto.
$D(\alpha)$	domínio de atracção de uma variável estável de índice α.
$M_p(E)$	espaço das medidas pontuais definidas em E.
$\mathcal{M}_p(E)$	σ-álgebra dos boreleanos de $M_p(E)$.
$M_+(E)$	espaço das medidas radon não negativas.
$\mathcal{M}_+(E)$	σ-álgebra dos boreleanos de $M_+(E)$.
$\psi_N(\cdot)$	funcional de Laplace do processo pontual N.
$RV\rho$	variação regular em ∞, com expoente de variação ρ.
AR	processo autoregressivo.
ARMA	processo autoregressivo e de médias móveis.
LAD	mínimos desvios absolutos (least absolute deviations).
LS	mínimos quadrados (least squares).
MA	processo de médias móveis.
PRM(μ)	Processo Pontual de Poisson de intensidade μ.

Resumo

A TEORIA DOS PROCESSOS PONTUAIS NA ESTIMAÇÃO EM MODELOS AUTOREGRESSIVOS DIRIGIDOS POR RUÍDOS COM VARIÂNCIA INFINITA

Paula Cristina Martins dos Reis

Mestrado em: Matemática Aplicada à Economia e Gestão

Orientador: Professora Luísa Canto e Castro Loura.

Provas concluídas em:

RESUMO

O estudo de séries temporais de natureza económica, física, social ou outra, nomeadamente no que diz respeito ao comportamento limite de somas e funções de covariância amostrais, pode ser feito com recurso às propriedades de convergência fraca de processos pontuais associados a modelos de médias móveis, gerados por erros com distribuição de variação regular de índice $-\alpha$ nas caudas, onde $\alpha > 0$. No âmbito deste trabalho são analisadas essas propriedades e utilizadas posteriormente no estudo do comportamento limite de um M-estimador dos parâmetros de um processo autoregressivo, cujas inovações apresentam variância infinita ou distribuição com caudas elevadas, isto é, a distribuição comum das inovações pertence agora ao domínio de atracção de uma lei estável não normal. Sob determinadas condições de momentos, a que obedece a distribuição dessas inovações próximo da origem, é estabelecida a convergência em distribuição do M-estimador em causa, o estimador LAD, para o mínimo de um processo estocástico definido no espaço das funções contínuas de $\mathbb{R}^p$ em $\mathbb{R}$.

Palavras chave: Processo pontual, processo de médias móveis, convergência vaga, convergência fraca, caudas com variação regular, processo autoregressivo, lei estável, domínio de atracção, M-estimação, estimador LAD (mínimos desvios absolutos).

Abstract

THE THEORY OF POINT PROCESSES ON AUTOREGRESSION ESTIMATION WITH INFINITE VARIANCE

Paula Cristina Martins dos Reis
Master of Science: Matemática Aplicada à Economia e Gestão
Superviser: Professora Luísa Canto e Castro Loura

ABSTRACT

The study of economic, physic, social or other nature time series, specially in the limit behaviour of sums and sample covariance functions, can be based on the proprieties of weak convergence of point processes, connected with moving averages models. Those models are driven by noise sequences with regulary varying tails of index $-\alpha$, where $\alpha > 0$. Within this work are analysed the refereed proprieties of weak convergence and used in the study of limit behaviour of an M-estimator for autoregressive process parameters. The innovations driving the AR process have heavy tails, i.e. the common distribution of innovations belongs to the domain of attraction of a non-normal stable law. Under specific moments conditions which involves the behaviour of distribution of this innovations near origin, is establish the convergence in distribution of the M-estimator (LAD estimator), for the minimum of a stochastic process defined in the space of continuous functions mapping $\mathbb{R}^p$ to $\mathbb{R}$.

Key words: Point processes, moving averages, vague convergence, weak convergence, regulary varying tails, autoregressive process, stable law, domain of attraction, M-estimation, LAD estimator (least absolute deviation).

Prefácio

Esta dissertação realizada no âmbito do curso de mestrado de Matemática Aplicada à Economia e Gestão, do Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa, tem por objectivo, estabelecer a ponte entre os processos autoregressivos e de médias móveis (ARMA) e a teoria geral dos *processos pontuais*. A escolha de tal objectivo, prende-se por um lado, com o reconhecimento claro da utilização frequente dos processos ARMA na modelização de séries temporais e por outro lado, na necessidade de se apresentar um suporte teórico alternativo, baseado concretamente na teoria dos processos pontuais e que contorne certo tipo de problemas, nomeadamente o problema da estimação, quando em particular os erros geradores dos processos têm distribuição com caudas elevadas, ou seja, apresentam variância infinita.

No primeiro capítulo, é exposto um conjunto de resultados que dizem respeito à teoria geral dos processos pontuais, com especial destaque para os *processos pontuais de poisson* (PRM), a *convergência vaga* e a *convergência fraca* de processos pontuais. Porque iremos estar a trabalhar em espaços métricos munidos das propriedades de completude e separabilidade, as noções de convergência fraca e em distribuição são equivalentes, sendo indiferente desse modo, estipular resultados em termos de uma ou outra convergência. Esta primeira parte do trabalho é baseada na sua maioria nos livros *Extreme Values, Regular Variation and Point Processes (*Resnick, 1987) e *Convergence of Probabilitiy Measures* (Billingsley, 1968).

No segundo capítulo, são construídos e interpretados processos pontuais associados a um modelo de médias móveis $\{X_k, k \geq 1\}$, cuja distribuição dos erroe apresenta as caudas com variação regular de índice $-\alpha$, onde $\alpha > 0$. Sob determinadas condições a que obedecem essas caudas, são estabelecidas e analisadas a convergência fraca desses processos pontuais, pondo-se ênfase à relação profunda entre convergência vaga e variação regular. Quando em particular $\alpha \in (0, 2]$, a distribuição dos erros pertence agora ao domínio de atracção de uma *lei estável* (para $\alpha = 2$, estamos perante a distribuição normal). O grosso destes resultados encontra-se no artigo *Limit Theory for Moving Averages of Random Variables with Regularly Tail Probabilities* de Davis e Resnick (1985).

A hipótese de normalidade dos erros geradores de um processo linear, frequentemente violada por determinados fenómenos que observamos e o reconhecimento de que muitos desses fenómenos sugerem a existência de um mecanismo aditivo subjacente, fez com que se pensasse em modelizar séries temporais, com recurso aos processos ARMA dirigidos por erros com variância infinita ou com distribuição de caudas pesadas, ou seja a distribuição das inovações pertence ao domínio de atracção de uma lei estável de índice $\alpha \in (0, 2)$. No terceiro capítulo é estudado precisamente o problema da estimação dos parâmetros de um processo autoregressivo, quando os erros verificam as condições anteriormente descritas. Nesse sentido, é proposto um M-estimador dos parâmetros do processo, o estimador LAD e analisado o seu desempenho através do seu comportamento limite. Sob determinadas condições de momentos a verificar pela distribuição dos erros próximo da origem, iremos ver que o estimador LAD converge em distribuição para o mínimo de um processo estocástico definido no espaço das funções contínuas de $\mathbb{R}^p$ em $\mathbb{R}$. O estudo deste estimador foi baseado fundamentalmente nos artigos *M-estimation For Autoregressions with Infinite Variance* de Davis e outros (1992) e *Inference for Linear Processes with Stable Noise* de Calder e Davis (1998).

Por último na secção das conclusões e dos desenvolvimentos futuros, expômos algumas reflexões sobre os diversos assuntos abordados ao longo deste trabalho, sem deixar de referir as motivações criadas para que se possa continuar a aprofundar todas as questões que envolvem esses mesmos assuntos.

Agradecimentos

O meu agradecimento, muito especial, à minha orientadora Prof.ª Luísa Canto e Castro Loura, pelos ensinamentos que me transmitiu ao longo deste trabalho. Reconheço que o seu entusiasmo e capacidade de realização reforçou o meu gosto pela investigação. Agradeço-lhe todo o seu incentivo, sugestões e críticas construtivas que me permitiram realizar esta dissertação com maior motivação e segurança.

O meu reconhecimento, igualmente, aos meus colegas Prof. Carlos Luz e Prof. Miguel Moreira, pelos seus contributos e pela disponibilidade que apresentaram.

Ao Departamento de Matemática e à Escola Superior de Tecnologia, agradeço a dispensa de serviço docente, o estímulo e o apoio constantes.

Finalmente, reconheço a compreensão familiar que tive e o incentivo que recebi para a realização deste trabalho. Agradeço particularmente aos meus pais por todo o apoio e carinho incessantes.

Capítulo 1

Resultados Gerais sobre Processos Pontuais

1.1 Conceitos Fundamentais

Suponhamos que um certo fenómeno aleatório tem ocorrências no intervalo real $(0, t]$ e que se pretende registar esse número de ocorrências. Se $N(t)$ representar esse número, o problema centra-se então na identificação da v.a. $N(t)$ ou $N((0, t])$. Ora $(0, t]$ é um intervalo de $\overline{\mathbb{R}}$ e é portanto natural que se queira estender o fenómeno em estudo a qualquer conjunto pertencente a uma σ-álgebra conveniente, gerada pelos intervalos da forma$(0, t]$. Assim poderíamos caracterizar $N(A)$, com A pertencente à referida σ-álgebra. Num contexto mais geral vamos considerar um espaço topológico E, com a classe $\mathcal{E}$ dos boreleanos de E, ou seja a σ-álgebra gerada pelos subconjuntos abertos de E. Mais concretamente, tomemos E como sendo um espaço topológico de Hausdorff[1], localmente compacto e com bases contáveis, i.e., um espaço de Hausdorff onde todo o elemento $x \in E$ tem uma vizinhança compacta, $\mathcal{V}_\delta(x)$, e todo o conjunto aberto G em E escreve-se na forma $G = \cup_{i \in I} G_i$, com $I \subset \mathbb{N}$ contável (finito ou numerável) e G_i subconjuntos abertos de E. Na nossa exposição iremos supor ainda que o espaço E é um subconjunto de um espaço euclidiano compacto de dimensão finita. Na classe $\mathcal{E}$ considerem-se as medidas de Dirac associadas a cada ponto $x \in E$,

$$\varepsilon_x(A) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases} \quad \text{para } A \in \mathcal{E} .$$

Facilmente se mostra que dada uma colecção contável de pontos de E,

$$\{x_1, x_2, ..., x_i, ...\}_{i \in I} ,$$

[1] Veja-se em apêndice.

não necessariamente distintos, a aplicação

$$\begin{array}{rccl} m: & \mathcal{E} & \longrightarrow & \overline{\mathbb{R}}_0^+ \\ & A & \longrightarrow & \sum_{i\in I} \varepsilon_{x_i}(A), \end{array}$$

é uma medida em $(E, \mathcal{E})$, designada por **medida pontual**. Para cada $A \in \mathcal{E}$, $m(A)$ representa então o número de pontos x_i que caem em A. Todas estas medidas pontuais podem ser reunidas num espaço simbolizado por $M_p(E)$. A classe dos boreleanos deste espaço representa-se por $\mathcal{M}_p(E)$ e define-se como a mais pequena σ-álgebra gerada por conjuntos de medidas definidos, para cada $B \in \mathcal{B}([0, \infty])$, da seguinte forma

$$\{m \in M_p(E) : m(A) \in B, \forall A \in \mathcal{E}\}.$$

As medidas pontuais com que iremos trabalhar são de *Radon*, i.e., $m(K) < \infty$, para todo o compacto $K \subset \mathcal{E}$.

Consideremos agora uma colecção contável de elementos aleatórias,

$$\{X_1, X_2 ..., X_i, ...\}_{i\in I},$$

que assumem valores em E, ou seja, para cada $i \in I$

$$\begin{array}{rccl} X_i: & \Omega & \longrightarrow & E \\ & w & \longrightarrow & X_i(w), \end{array}$$

é uma aplicação mensurável. Ora, para cada $w \in \Omega$, temos que o conjunto

$$\{X_1(w), X_2(w), ..., X_i(w), ...\}_{i\in I}$$

é uma colecção de pontos de E e por conseguinte podemos construir a medida pontual

$$\begin{array}{rccl} m_w: & \mathcal{E} & \longrightarrow & \overline{\mathbb{R}}_0^+ \\ & A & \longrightarrow & m_w(A) = \sum_{i\in I} \varepsilon_{X_i(w)}(A). \end{array}$$

Naturalmente que podemo-nos interrogar se, dada uma colecção de variáveis aleatórias, é então possível construir medidas pontuais que sejam igualmente aleatórias. A resposta é afirmativa e centra-se precisamente na teoria a que este trabalho se dedica: a teoria dos processos pontuais. Os próximos resultados ajudam-nos a esclarecer melhor estes conceitos.

Definição 1.1 *Um **processo pontual** em E é uma aplicação mensurável N, do espaço probabilístico $(\Omega, \mathcal{A}, P)$ para o espaço $(M_p(E), \mathcal{M}_p(E))$, ou seja, é um elemento aleatório de $M_p(E)$. A lei probabilística P_N do processo N, é a medida $P \circ N^{-1} = P[N \in \cdot]$ em $\mathcal{M}_p(E)$.*

Repare-se que se tomarmos $w \in \Omega$, $N(w,\cdot)$ é uma medida pontual em $M_p(E)$ e $N(w,A)$ representa o número de pontos em A para a realização w. Esquematicamente temos:

$$\begin{array}{rccc} N: & (\Omega,\mathcal{A},P) & \rightarrow & (M_p(E),\mathcal{M}_p(E)) \\ & w & \rightarrow & N(w,\cdot)=m_w \end{array}$$

Proposição 1.1 *N é um processo pontual sse a aplicação do espaço $(\Omega,\mathcal{A})$ para $([0,\infty],\mathcal{B}[0,\infty])$, que transforma $w \in \Omega$ em $N(w,A)$, com $A \in \mathcal{E}$ é uma aplicação mensurável.*

Dem. Esta demonstração segue de perto a apresentada em [24], pág.124.

Se N é um processo pontual, resulta por definição que a aplicação de $(\Omega,\mathcal{A})$ em $(M_p(E),\mathcal{M}_p(E))$ e que transforma w em $N(w,\cdot)$, é mensurável. Por outro lado a aplicação de $(M_p(E),\mathcal{M}_p(E))$ em $([0,\infty],\mathcal{B}[0,\infty])$ e que transforma m em $m(A)$, é igualmente mensurável por definição de $\mathcal{M}_p(E)$. Consequentemente a aplicação composta das duas anteriores é mensurável.

Reciprocamente, suponhamos que a aplicação de $(\Omega,\mathcal{A})$ para $([0,\infty],\mathcal{B}[0,\infty])$ é mensurável, ou seja, $\{w : N(w,A) \in B\} \in \mathcal{A}$, com $B \in \mathcal{B}[0,\infty]$ e $A \in \mathcal{E}$. Defina-se

$$\mathfrak{I} = \left\{F \in \mathcal{M}_p(E) : N^{-1}(F) \in \mathcal{A}\right\}.$$

É fácil de se verificar que $\mathfrak{I}$ é uma σ-álgebra. Por outro lado $\mathfrak{I}$ contém todos os conjuntos da forma $\{m : m(A) \in B\}$ uma vez que por hipótese tem-se

$$N^{-1}(\{m : m(A) \in B\}) = \{w : N(w,A) \in B\} \in \mathcal{A}.$$

Logo $\mathfrak{I} \supset \sigma\{\{m : m(A) \in B\}, A \in \mathcal{E}, B \in \mathcal{B}[0,\infty]\} = \mathcal{M}_p(E)$, o que prova que N é um processo estocástico pontual. ∎

Observação 1.1 Repare-se que N é processo pontual sse $N(A)$ é uma v.a. discreta com valores em $\overline{\mathbb{R}}_0^+$, para todo $A \in \mathcal{E}$.

Definição 1.2 *A* ***intensidade*** *de um processo N é a medida μ definida em $(E,\mathcal{E})$ por*

$$\begin{aligned} \mu(A) &= E[N(A)] = \sum_{n\in\mathbb{N}} nP[N(A)=n] \\ &= \int_\Omega N(w,A)\,P(dw) \\ &= \int_{M_p(E)} m(A)\,P_N(dm), \forall A \in \mathcal{E}. \end{aligned}$$

Sejam agora uma aplicação mensurável não negativa f definida de $(E, \mathcal{E})$ para o espaço $([0,\infty], \mathcal{B}[0,\infty])$ e $\{f_n\}_{n\in\mathbb{N}}$ uma sucessão de funções simples tal que $0 \leq f_n \uparrow f$ (sabe-se que existe). Ora, para cada $n \in \mathbb{N}$, f_n tem uma representação da forma:

$$f_n = \sum_{i=1}^{j_n} c_i^{(n)} I_{A_i^{(n)}},$$

com $A_i^{(n)} \in \mathcal{E}$ e com $\left\{A_i^{(n)}, i \leq j_n\right\}$ disjuntos. Por outro lado para cada $w \in \Omega$, $N_w = N(w,\cdot)$ é uma medida, pelo que se pode definir

$$N(w,f) = \int_E f dN_w \leq \infty.$$

Note-se agora que, para cada f, a aplicação

$$\begin{array}{rccc} N_f = N(f): & \Omega & \longrightarrow & \overline{\mathbb{R}}_0^+ \\ & w & \longrightarrow & N_f(w) = N(w,f) \end{array}$$

é uma v.a., pois $N_f = \lim_{n\to\infty} \uparrow N_{f_n}$(pelo teorema da convergência monótona) e N_{f_n} é uma combinação linear de v.a.'s. Tem-se por conseguinte,

$$N(f) = \lim_{n\to\infty} N(f_n) = \lim_{n\to\infty} \sum_{i=1}^{j_n} c_i^{(n)} N\left(A_i^{(n)}\right).$$

O valor esperado de $N(f)$ vem dado por

$$E[N(f)] = \mu(f) = \int_E f d\mu,$$

uma vez que,

$$\begin{aligned} E[N(f)] &= \lim_{n\to\infty} \uparrow E[N(f_n)] = \lim_{n\to\infty} \uparrow E\left[\sum_{i=1}^{j_n} c_i^{(n)} N\left(A_i^{(n)}\right)\right] \\ &= \lim_{n\to\infty} \uparrow \sum_{i=1}^{j_n} c_i^{(n)} \mu\left(A_i^{(n)}\right) = \lim_{n\to\infty} \uparrow \int_E f_n(x)\,\mu(dx) \\ &= \int_E \left(\lim_{n\to\infty} \uparrow f_n(x)\right) \mu(dx) = \int_E f(x)\,\mu(dx). \end{aligned}$$

A teoria dos processos pontuais recorre frequentemente ao *Funcional de Laplace,* o qual é fundamental para que se possa estabelecer uma vasta gama de resultados que dizem respeito entre outros, ao estudo da convergência fraca e convergência vaga de processos estocásticos pontuais (abordados atempadamente neste trabalho). Com vista a definir este funcional, considere-se Q uma medida de probabilidade em $(M_p(E), \mathcal{M}_p(E))$. *A transformada de Laplace de Q* define-se para cada $f : (E, \mathcal{E}) \to ([0,\infty], \mathcal{B}[0,\infty])$, com $f \geq 0$, por

$$\psi(f) = \int_{M_p(E)} \left(\exp\left\{ -\int_E f(x)\, m(dx) \right\} \right) Q(dm).$$

Temos então o seguinte resultado:

Definição 1.3 *Seja $N : (\Omega, \mathcal{A}) \to (M_p(E), \mathcal{M}_p(E))$ um processo pontual. O **Funcional de Laplace** de N é a transformada de Laplace da lei probabilística de N, ou seja:*

$$\begin{aligned} \Psi_N(f) &= E\{\exp[-N(f)]\} = \int_\Omega \exp\{-N(w,f)\}\, P(dw) \\ &= \int_{M_p(E)} \left\{ \exp\left[-\int_E f(x)\, m(dx) \right] P_N(dm) \right\}. \end{aligned}$$

Proposição 1.2 *O Funcional de Laplace ψ_N do processo N determina univocamente a lei de N.*

Dem. A demonstração pode ser consultada em [24], pág.129. ■

Exemplo 1.1 Considere-se a função simples $f : E \to [0,\infty]$ definida por

$$f(x) = \sum_{i=1}^{j} \lambda_i I_{A_i}(x), \text{ com } \{A_j, j \geq 1\} \subset \mathcal{E} \text{ e } \lambda_i \geq 0, i = 1, ...j.$$

Então

$$N(w,f) = \int f(x)\, N(w,dx) = \sum_{i=1}^{j} \lambda_i N(w, A_i)$$

e portanto o Funcional de N vem:

$$\Psi_N(f) = E\left\{ \exp\left[-\sum_{i=1}^{k} \lambda_i N(A_i) \right] \right\}.$$

Exemplo 1.2 Usando Ψ_N podem ser determinados alguns momentos de N. Por exemplo para $\forall f \geq 0$ mensurável, pode-se provar que

$$\mu(f) = E[N(f)] = \lim_{t \downarrow 0} \frac{1 - \Psi(tf)}{t}.$$

Para esse efeito basta observar que

$$N(f) = \lim_{t \downarrow 0} \uparrow \frac{1 - e^{-tN(f)}}{t};$$

tomando os valores esperados e usando o teorema da convergência monótona obtém-se o resultado.

1.2 Processos Pontuais de Poisson

Um classe de processos pontuais, de extrema importância para o trabalho que aqui é exposto, é a dos chamados processos de Poisson. Como iremos ver, nos próximos capítulos, eles constituem o suporte de toda a teoria limite desenvolvida para os processos de médias móveis e autoregressivos, quando estes últimos são dirigidos por ruídos com variância infinita.

Definição 1.4 *Seja μ uma medida de Radon em $\mathcal{E}$. Um processo pontual N é um **Processo de Poisson** ou um **PRM** (Poisson Random Measure), de intensidade μ e escreve-se **PRM(μ)**, se forem verificadas as condições:*

1. *$\forall A \in \mathcal{E}, \forall k \in \mathbb{N}$*

$$P[N(A)=k] = \begin{cases} \exp\{-\mu(A)\} \frac{[\mu(A)]^k}{k!} & se\ \mu(A) < \infty \\ 0 & se\ \mu(A) = \infty. \end{cases}$$

2. *Sendo $A_1, A_2, ..., A_j$, conjuntos mutuamente disjuntos em $\mathcal{E}$, para todo o $j \geq 1$, então as variáveis aleatórias $N(A_1), N(A_2), ..., N(A_j)$ são independentes.*

Observação 1.2 1. Repare-se que resulta de *1.* *que se* $\mu(A) = \infty$, então $N(A) = \infty$ quase certamente.

2. Para cada $A \in \mathcal{E}$, $N(A)$ é uma v.a. de Poisson de taxa $\mu(A)$.

3. Um PRM(μ) com $\mu(dx) = dx$ (medida de Lebesgue), diz-se um **PRM Homogêneo.**

Exemplo 1.3 Considere-se $\{E_i, i \geq 1\}$ uma sucessão de variáveis aleatórias i.i.d. de exponenciais unitárias e seja $\Gamma_n = \sum_{i=1}^n E_i$. Mostremos que o processo pontual $N = \sum_{n\geq 1} \varepsilon_{\Gamma_n}$ é um PRM homogéneo em $\mathbb{R}_0^+$. Dada a sucessão crescente de v.a.'s $\{\Gamma_n\}_{n\in\mathbb{N}}$, interessa-nos identificar a distribuição do n^{o}de vezes que os pontos Γ_n, caem num certo boreleano $B \in \mathcal{B}(\mathbb{R}_0^+)$. Ora, por se tratar dos boreleanos de $\mathbb{R}_0^+$ basta fazer essa identificação para B da forma $[0,t)$, com $t > 0$. Definindo então a medida,

$$\varepsilon_{\Gamma_n}([0,t)) = \begin{cases} 1 & \text{se } \Gamma_n < t \\ 0 & \text{se } \Gamma_n \geq t \end{cases}$$

vem,

$$\begin{aligned} P[N([0,t)) = k] &= P[\Gamma_{k+1} \geq t, \Gamma_k < t] \\ &= P[\Gamma_k < t] - P[\Gamma_{k+1} < t] \\ &= 1 - \sum_{j=0}^{k-1} e^{-t}\frac{t^j}{j!} - 1 + \sum_{j=0}^{k} e^{-t}\frac{t^j}{j!} \\ &= e^{-t}\frac{t^k}{k!}, \end{aligned}$$

pois $\Gamma_n = \sum_{i=1}^{n} E_i \frown Gama\,(n,1)$ e $F_X(x) = 1 - \sum_{j=0}^{n-1} e^{-\lambda x}\frac{(\lambda x)^j}{j!}$é a função de distribuição de uma v.a. gama de parâmetros (n,λ). A intensidade de N vem dada por

$$\begin{aligned} \mu([0,t)) &= t \Longleftrightarrow \int_0^t \mu(dx) = t \\ &\Longleftrightarrow \mu(dx) = dx, \end{aligned}$$

logo μ é a medida de Lebesgue nos boreleanos de $\mathbb{R}_0^+$. Finalmente falta analisar a independência das v.a.'s $N(A_1), ..., N(A_j)$, com $A_1, ..., A_j$, boreleanos de $\mathbb{R}_0^+$, mutuamente disjuntos. Prova-se, (Matthes, 1978), que essa independência é equivalente à indepedência dos incrementos $N[t_0, t_1), ..., N[t_{j-1}, t_j)$, com $t_0 < t_1 < ... < t_{j-1} < t_j$, sendo esta última por sua vez um resultado conhecido.

Exemplo 1.4 Considerem-se as variáveis aleatórias do exemplo anterior. Mostremos que o processo $N = \sum_{n\geq 1} \varepsilon_{\Gamma_n^{-1/\alpha}}$ é um PRM(μ) em $\overline{\mathbb{R}}^+$, com medida de intensidade dada por $\mu(dx) = \alpha x^{-\alpha-1}dx$, onde $\alpha > 0$.

Sendo $\{\Gamma_n, n \geq 1\}$ uma sucessão crescente por definição, a sucessão $\left\{\Gamma_n^{-1/\alpha}\right\}$ será decrescente para $\alpha > 0$. A identificação da distribuição do n^{o}de vezes que os pontos $\Gamma_n^{-1/\alpha}$ caem num certo boreleano de $\overline{\mathbb{R}}^+$, pode ser feita agora à custa de um boreleano da forma $B = (t,\infty]$, com $t > 0$. Tendo em conta que

$$\varepsilon_{\Gamma_n^{-1/\alpha}}((t,\infty]) = \begin{cases} 1 & \text{se } \Gamma_n^{-1/\alpha} > t \\ 0 & \text{se } \Gamma_n^{-1/\alpha} \leq t \end{cases}$$

vem,

$$\begin{aligned} P[N((t,\infty]) = k] &= P\left[\Gamma_k^{-1/\alpha} > t, \Gamma_{k+1}^{-1/\alpha} \leq t\right] \\ &= P\left[\Gamma_k^{-1/\alpha} > t\right] - P\left[\Gamma_{k+1}^{-1/\alpha} > t\right] \\ &= P\left[\Gamma_k \leq t^{-\alpha}\right] - P\left[\Gamma_{k+1} \leq t^{-\alpha}\right] \\ &= 1 - \sum_{j=0}^{k-1} e^{-t^{-\alpha}} \frac{(t^{-\alpha})^j}{j!} - 1 + \sum_{j=0}^{k} e^{-t^{-\alpha}} \frac{(t^{-\alpha})^j}{j!} \\ &= e^{-t^{-\alpha}} \frac{(t^{-\alpha})^k}{k!}. \end{aligned}$$

Então a intensidade μ vem então dada por

$$\begin{aligned} \mu([t,\infty]) &= \int_t^\infty \mu(dx) = t^{-\alpha}, \text{ donde,} \\ \mu(dx) &= \alpha x^{-\alpha-1}dx. \end{aligned}$$

Exemplo 1.5 Sejam $\{\delta_i, i \geq 1\}$ variáveis aleatórias i.i.d., com $P[\delta_i = 1] = p$ e $P[\delta_i = -1] = 1 - p, 0 \leq p \leq 1$. Considere-se ainda $\Gamma_n = \sum_{i=1}^{n} E_i$, onde $\{E_i, i \geq 1\}$ é a sequência de v.a.'s i.i.d. dos exemplos anteriores.. Admita-se que $\{\delta_i, i \geq 1\}$ e $\{E_i, i \geq 1\}$ são independentes. Verifiquemos que em $\overline{\mathbb{R}} \setminus \{0\}$, o processo

$$N^* = \sum_{n=1}^{\infty} \varepsilon_{\delta_n \Gamma_n^{-1/\alpha}}$$

é um PRM(μ), com medida de intensidade dada por

$$\mu(dx) = \alpha \left[p x^{-\alpha-1} I_{(0,\infty)}(x) + (1-p)(-x)^{-\alpha-1} I_{(-\infty,0)}(x) \right] dx.$$

A distribuição do nºde vezes que os pontos $\delta_n \Gamma_n^{-1/\alpha}$ estão num certo boreleano B de $\overline{\mathbb{R}} \setminus \{0\}$, bem como a intensidade do processo, podem ser determinadas tomando um boreleano da forma $B = (t, \infty] \cup [-\infty, -t)$, com $t > 0$. Temos então

$$\begin{aligned}
P[N^*((t,\infty] \cup [-\infty,-t)) = k] &= P\left[\left(\delta_k \Gamma_k^{-1/\alpha} > t \wedge \delta_{k+1} \Gamma_{k+1}^{-1/\alpha} \leq t\right)\right. \\
&\qquad \left. \vee \left(-\delta_k \Gamma_k^{-1/\alpha} < -t \wedge -\delta_{k+1} \Gamma_{k+1}^{-1/\alpha} \geq -t\right)\right] \\
&= P\left[\delta_k \Gamma_k^{-1/\alpha} > t \wedge \delta_{k+1} \Gamma_{k+1}^{-1/\alpha} \leq t\right] \\
&= P\left[\delta_k \Gamma_k^{-1/\alpha} > t\right] - P\left[\delta_{k+1} \Gamma_{k+1}^{-1/\alpha} > t\right] \\
&= P\left[\Gamma_k \leq t^{-\alpha}\right] - P\left[\Gamma_{k+1} \leq t^{-\alpha}\right] \\
&= e^{-t^{-\alpha}} \frac{(t^{-\alpha})^k}{k!}
\end{aligned}$$

Analisemos agora a probabilidade de caírem k pontos da forma $\delta_n \Gamma_n^{-1/\alpha}$ no intervalo $B_1 = (t, \infty]$. Considerando $N = \sum_{n \geq 1} \varepsilon_{\Gamma_n^{-1/\alpha}}$, vem que

$$\begin{aligned}
P[N^*(B_1) = k] &= \sum_{j \geq k} P[N^*((t,\infty]) = k \mid N((t,\infty]) = j] \\
&= \sum_{j \geq k} \binom{j}{k} p^k (1-p)^{j-k} e^{-t^{-\alpha}} \frac{(t^{-\alpha})^j}{j!} \\
&= \frac{(t^{-\alpha})^k}{k!} p^k e^{-t^{-\alpha}} \sum_{j \geq 0} \frac{(1-p)^j (t^{-\alpha})^j}{(j)!} \\
&= \frac{(t^{-\alpha})^k}{k!} p^k e^{-t^{-\alpha}} e^{(1-p)t^{-\alpha}} \\
&= \frac{(t^{-\alpha} p)^k}{k!} e^{-pt^{-\alpha}},
\end{aligned}$$

logo,

$$\mu\left((t,\infty]\right)=\int_t^{\infty}\mu\left(dx\right)=t^{-\alpha}p\Longleftrightarrow\mu\left(dx\right)=\alpha p x^{-\alpha-1}dx,$$

para $x>0$. Como $\mu\left((t,\infty]\cup[-\infty,-t)\right)=\mu\left((t,\infty]\right)+\mu\left([-\infty,-t)\right)$, vem

$$\begin{aligned}\mu\left([-\infty,-t)\right) &= \int_{-\infty}^{t}\mu\left(dx\right)=t^{-\alpha}-t^{-\alpha}p\\ &= (1-p)\,t^{-\alpha}\\ \Longleftrightarrow&\ \mu\left(dx\right)=\alpha\left(1-p\right)\left(-x\right)^{-\alpha-1}dx,\end{aligned}$$

para $x<0$, donde,

$$\begin{aligned}\mu\left(dx\right) &= \alpha\left(px^{-\alpha-1}I_{(0,\infty]}\left(x\right)+\left(1-p\right)\left(-x\right)^{-\alpha-1}I_{[-\infty,0)}\left(x\right)\right)dx\\ &= \alpha\left(px^{-\alpha-1}I_{(0,\infty)}\left(x\right)+\left(1-p\right)\left(-x\right)^{-\alpha-1}I_{(-\infty,0)}\left(x\right)\right)dx.\end{aligned}$$

Proposição 1.3 *O Funcional de Laplace de um PRM(μ) é dado por*

$$\Psi_N\left(f\right)=\exp\left\{-\int\left(1-e^{-f(x)}\mu\left(dx\right)\right)\right\},\forall f\geq 0,\ \textit{mensurável.}\tag{1.1}$$

Reciprocamente, todo o processo pontual com o Funcional de Laplace definido em (1.1) é um PRM(μ).

Dem. Esta demonstração segue de perto a apresentada em [24], pág.130.

Comecemos por demonstrar que todo o processo pontual que satisfaça 1. e 2. da definição anterior tem o funcional de Laplace da forma (1.1). Seja então N um PRM(μ). Se $c>0$, dado $A\in\mathcal{E}$ e $f\left(x\right)=cI_A\left(x\right)$ então $N\left(f\right)=cN\left(A\right)$, com $N\left(A\right)$ v.a. Poisson, vindo a partir de 1. da definição

$$\begin{aligned}\Psi_N\left(f\right) &= E\left\{\exp\left[-N\left(f\right)\right]\right\}=E\left\{\exp\left[-cN\left(A\right)\right]\right\}\\ &= \exp\left[\left(e^{-c}-1\right)\mu\left(A\right)\right]\\ &= \exp\left[-\int_E\left(1-e^{-f(x)}\right)\mu\left(dx\right)\right],\end{aligned}$$

que é a expressão definida em (1.1). Sejam agora $c_i\geq 0$ e $A_1,\ldots,A_j$ conjuntos disjuntos de $\mathcal{E}$, com $i\leq j$ e $j\geq 1$. Então $N\left(A_1\right),\ldots,N\left(A_j\right)$ são independentes e se $f(x) = \sum_{i=1}^{j} c_i I_{A_i}(x)$, tem-se

$$\begin{aligned}
\Psi_N(f) &= E\{\exp[-N(f)]\} = E\left\{\exp\left[-\sum_{i=1}^{j} c_i N(A_i)\right]\right\} \\
&= \prod_{i=1}^{j} \{\exp[-c_i N(A_i)]\} \\
&= \prod_{i=1}^{j} \left\{\exp\left[-\int_E \left(1 - e^{-c_i I_{A_i}(x)}\right)\mu(dx)\right]\right\} \\
&= \exp\left[-\int_E \left(\sum_{i=1}^{j} \left(1 - e^{-c_i I_{A_i}(x)}\right)\right)\mu(dx)\right] \\
&= \exp\left[-\int_E \left(1 - e^{-\sum_{i=1}^{j} c_i I_{A_i}(x)}\right)\mu(dx)\right] \\
&= \exp\left[-\int_E \left(1 - e^{-f(x)}\right)\mu(dx)\right],
\end{aligned}$$

que é a expressão definida em (1.1). Tomemos agora uma função qualquer f mensurável tal que $f \geq 0$. Sabemos que existe uma função simples f_n da forma

$$f_n = \sum_{i=1}^{j_n} c_i^{(n)} I_{A_i^{(n)}},$$

com $A_i^{(n)} \in \mathcal{E}$, $\left\{A_i^{(n)}, 1 \leq i \leq j_n\right\}$ disjuntos tal que $0 \leq f_n \uparrow f$ e $c_i^{(n)} \geq 0$. O teorema da convergência monótona garante que $N(f_n) \uparrow N(f)$ para todo $w \in \Omega$. Uma vez que $e^{-N(g)} \leq 1, \forall g \geq 0$ mensurável, então tem-se quando $n \to \infty$,

$$\Psi_N(f_n) = E\{\exp[-N(f_n)]\} \to E\{\exp[-N(f)]\} = \Psi_N(f).$$

Mas por outro lado f_n verifca (1.1) e portanto

$$\Psi_N(f_n) = \exp\left[-\int \left(1 - e^{-f_n}\right) d\mu\right].$$

Se $f_n \uparrow f$ então ter-se-á também $\left(1 - e^{-f_n}\right) \uparrow \left(1 - e^{-f}\right)$ e aplicando o teorema da convergência monótona vem que

$$\int_E \left(1 - e^{-f_n}\right) d\mu \uparrow \int_E \left(1 - e^{-f}\right) d\mu,$$

provando-se assim (1.1) para toda a função $f \geq 0$ mensurável.

Reciprocamente, suponhamos que N é um processo pontual que verifica (1.1). Tomando $f = \lambda I_A$ com $A \in \mathcal{E}$, tem-se

$$E\{\exp[-N(f)]\} = E\{\exp[-\lambda N(A)]\} = \exp\left[-\left(1-e^{-\lambda}\right)\mu(A)\right],$$

que representa o Funcional de Laplace de uma variável aleatória de Poisson com parâmetro $\mu(A)$. Assim a condição 1. da definição de PRM é verificada. Por outro lado para $A_1, ..., A_j$ disjuntos em $\mathcal{E}$ e considerando $f = \sum_{i=1}^{j} \lambda_i I_{A_i}, \lambda_i \geq 0$, obtemos

$$\begin{aligned}
E\{\exp[-N(f)]\} &= E\left\{\exp\left[-\sum_{i=1}^{j}\lambda_i N(A_i)\right]\right\} \\
&= \exp\left[-\int_E \left(1 - e^{\sum_{i=1}^{j}\lambda_i I_{A_i}}\right) d\mu\right] \\
&= \exp\left[-\int_E \left(\sum_{i=1}^{j}\left(1 - e^{-\lambda_i I_{A_i}}\right)\right) d\mu\right] \\
&= \prod_{i=1}^{j} \exp\left[-\left(1-e^{-\lambda_i}\right)\mu(A_i)\right] \\
&= \prod_{i=1}^{j} E\{\exp[-\lambda N(A_i)]\},
\end{aligned}$$

donde $N(A_1), ..., N(A_j)$ são independentes e por conseguinte a condicão 2 da definição é verificada. ■

Vamos ver agora como se pode construir um PRM(μ). Para esse efeito consideramos duas situações. Uma em que μ é finita e outra em que é infinita.

1. $\mu(E) < \infty$, i.e., μ é finita.

   Seja $\mu = c\upsilon$, onde υ é uma medida de probabilidade e $c > 0$. Considere-se um espaço de probabilidade onde as variáveis $M, X_1, X_2, ...$, são independentes, com $M \frown P(c)$ e $\{X_i\}_{i\geq 1}$ são elementos aleatórios i.i.d. em E com distribuição υ, ou seja, $P[X_i \in A] = \upsilon(A), \forall A \in \mathcal{E}$. Defina-se N em Ω por

   $$N = \begin{cases} \sum_{i=1}^{M} \varepsilon_{X_i} & \text{se } M > 0 \\ 0 & \text{se } M = 0 \end{cases}$$

   Verifiquemos que N é um Processo Pontual de Poisson, ou seja que verifica as condições 1. e 2. da definição.

a) Neste caso temos

$$\begin{aligned}
P[N(A)=k] &= P\left[\sum_{i=1}^{M}\varepsilon_{X_i}(A)=k\right] \\
&= \sum_{n\geq k} P\left[\sum_{i=1}^{n}\varepsilon_{X_i}(A)=k\right] P[M=n] \\
&= \sum_{n\geq k}\binom{n}{k}(\nu(A))^k(1-\nu(A))^{n-k}e^{-c}\frac{c^n}{n!} \\
&= \sum_{n\geq k}\frac{n!}{(n-k)!k!}(\nu(A))^k(1-\nu(A))^{n-k}e^{-c}\frac{c^n}{n!} \\
&= \frac{(c\nu(A))^k}{k!}e^{-c}\sum_{n\geq 1}\frac{[c(1-\nu(A))]^n}{n!} \\
&= \frac{(\mu(A))^k}{k!}e^{-c}e^{c(1-\nu(A))} = e^{-\mu(A)}\frac{(\mu(A))^k}{k!}.
\end{aligned}$$

Note-se que $N^* = \sum_{i=1}^{n}\varepsilon_{X_i}(A)$ representa o número de v.a X_i entre as n que estão em A e por tal N^* é uma binomial de parâmetros $(n, P[X_i \in A])$, ou seja, $N^* \frown B(n, \nu(A))$.

b) A independência de $N(A_1), N(A_2), ..., N(A_j)$, com $A_1, A_2, ..., A_j$, disjuntos em $\mathcal{E}$, prova-se facilmente, tendo em conta que o vector aleatório $(\sum_{i=1}^{n}\varepsilon_{X_i}(A_1), ..., \sum_{i=1}^{n}\varepsilon_{X_i}(A_j))$ tem distribuição multinomial. Omite-se aqui a sua resolução, a qual pode ser consultada em [24], pág.133.

2. $\mu(F) = \infty$

Vamos decompor μ na forma $\mu = \sum_{j=1}^{\infty}\mu_j$, procedendo do seguinte modo: tome-se uma partição de E, $\{A_j, j \geq 1\}$, constituída por conjuntos relativamente compactos e defina-se $\mu_j(B) = \mu(B \cap A_j), \forall B \in \mathcal{E}$. μ_j é finita uma vez que $\mu_j(E) = \mu(E \cap A_j) = \mu(A_j) < \infty$ (pois μ é Radon e A_j é relativamente compacto). Podemos assim construir processos PRM(μ_j), os quais serão designados por N_j. Supondo que estes processos são independentes uns dos outros, então o Processo Pontual de Poisson de intensidade μ a construir é da forma $N = \sum_{j=1}^{\infty} N_j$. Averiguemos que N é de facto um PRM(μ), verificando que o seu Funcional de Laplace se define como em

(1.1). Seja então $f \geq 0$, mensurável. Então

$$\begin{aligned}
\psi_N(f) &= E\{\exp[-N(f)]\} = E\left\{\exp\left[-\sum_{j=1}^{\infty} N_j(f)\right]\right\} \\
&= \lim_{r\to\infty} E\left\{\exp\left(-\sum_{j=1}^{r} N_j(f)\right)\right\} \\
&= \lim_{r\to\infty}\prod_{j=1}^{r} E\{\exp(-N_j(f))\}, \\
&= \lim_{r\to\infty}\prod_{j=1}^{r}\left\{\exp\left[-\int_E \left(1-e^{-f(x)}\right)\mu_j(dx)\right]\right\} \\
&= \lim_{r\to\infty}\left\{\exp\left[-\int_E \left(1-e^{-f(x)}\right)\left(\sum_{j=1}^{r}\mu_j\right)(dx)\right]\right\} \\
&= \exp\left[-\int_E \left(1-e^{-f(x)}\right)\left(\sum_{j=1}^{\infty}\mu_j\right)(dx)\right] \\
&= \exp\left[-\int_E \left(1-e^{-f(x)}\right)\mu(dx)\right] \\
&= \exp\left[-\int_E \left(1-e^{-f}\right)d\mu\right],
\end{aligned}$$

em que a quarta igualdade surge pela independência de $\{N_j, j \geq 1\}$ e a quinta pelo facto de N_j ser um PRM(μ_j). A última expressão é precisamente o funcional de Laplace definido como em (1.1).

Vimos como são caracterizados os Processos de Poisson e como inclusive, podem ser construídos.Veremos no próximo resultado, que a transformação de um PRM, por meio de uma aplicação mensurável, ainda é um PRM, isto é, o processo que se obtém aplicando uma transformação mensurável aos pontos que definem um PRM é ainda um PRM.

Proposição 1.4 *Sejam E_1 e E_2 dois espaços localmente compactos com bases contáveis, e $\mathcal{E}_1$ e $\mathcal{E}_2$ as respectivas σ-álgebras. Seja $T : (E_1, \mathcal{E}_1) \longrightarrow (E_2, \mathcal{E}_2)$ uma aplicação mensurável. Se N é um PRM(μ) em E_1, então*

$$\widetilde{N} := N \circ T^{-1}$$

é um PRM em E_2 com intensidade $\widetilde{\mu}$ definida por

$$\widetilde{\mu} = \mu \circ T^{-1}.$$

Se tivermos uma representação da forma

$$N = \sum_{i \geq 1} \varepsilon_{X_i}$$

então,

$$\widetilde{N} = N \circ T^{-1} = \sum_{i \geq 1} \varepsilon_{T(X_i)}.$$

Dem. A demonstração segue de perto a apresentada em [23], pág.134. Seja f_2 uma aplicação mensurável de E_2 em $[0, \infty)$. Então pelo teorema de transformação de integrais, vem

$$\begin{aligned}
\psi_{\widetilde{N}}(f_2) &= E\left\{\exp\left[-\widetilde{N}(f_2)\right]\right\} \\
&= E\left\{\exp\left[-\int_{E_2} f_2(x_2)\, N \circ T^{-1}(w, dx_2)\right]\right\} \\
&= E\left\{\exp\left[-\int_{E_1} f_2(T(x_1))\, N(w, dx_1)\right]\right\}.
\end{aligned}$$

Como $f_2 \circ T$ é não negativa e mensurável em E_1 e N é um $PRM(\mu)$ em E_1, então por (1.1) obtém-se

$$\begin{aligned}
\psi_{\widetilde{N}}(f_2) &= \exp\left[-\int_{E_1} \left(1 - e^{-f_2 \circ T}\right) d\mu\right] \\
&= \exp\left[-\int_{E_2} \left(1 - e^{-f_2(x)}\right) \mu \circ T^{-1}(dx)\right],
\end{aligned}$$

que é o funcional de Laplace do $PRM(\mu \circ T^{-1})$ em E_2. ∎

Proposição 1.5 *Sejam E_1 e E_2 dois espaços nas condições anteriores. Suponhamos que*

$$N_1 = \sum_{i \geq 1} \varepsilon_{X_i},$$

é um PRM(μ) em $(E_1, \mathcal{E}_1)$ e seja $K : E_1 \times \mathcal{E}_2 \longrightarrow [0,1]$ uma função de transição de $E_1 \longrightarrow E_2$, i.e., $K(\cdot, F_2)$ é $\mathcal{E}_1$-mensurável para todo $F_2 \in \mathcal{E}_2$ e $K_x(\cdot)$ é uma medida de probabilidade em $\mathcal{E}_2$, para todo $x \in E_1$. Seja $\{J_i, i \geq 1\}$ uma colecção de elementos aleatórios com valores em E_2, condicionalmente independentes, dado $\{X_n\}$, i.e.,

$$P[J_i \leq y \mid \{X_n\}, \{J_\alpha, \alpha \neq i\}] = P[J_i \leq y \mid X_i = x] = K_x(y)$$

ou,

$$
\begin{aligned}
P\left[J_i \in F_2 \mid \{X_n\}, \{J_\alpha, \alpha \neq i\}\right] &= P\left[J_i \in F_2 \mid X_i\right] \qquad (1.2) \\
&= K_{X_i}(F_2) = K(X_i, F_2), \forall_i, \forall F_2 \in \mathcal{E}_2.
\end{aligned}
$$

Então o processo bidimensional

$$
N^* = \sum_{i \geq 1} \varepsilon_{(X_i, J_i)}
$$

é um PRM(μ) em $E_1 \times E_2$ de intensidade μ^* onde

$$
\mu^*(dx, dy) = \mu(dx) K_x(dy).
$$

Caso especial: Se $\{J_i, i \geq 1\}$ forem i.i.d. e independentes de $\{X_i, i \geq 1\}$, com distribuição comum F, então $\sum_{i \geq 1} \varepsilon_{(X_i, J_i)}$ é um PRM(μ^*) com $\mu^* = \mu(dx) F(dy)$.

A demonstração desta proposição recorre a dois lemas, os quais passamos a apresentar. As respectivas demonstrações omitem-se, mas podem ser analisadas em [24], pág.135-36.

Lema 1.1 *Seja $f : E_1 \times E_2 \longrightarrow [0, \infty)$ uma função limitada. Então*

$$
E\left[f(X_i, J_i) \mid \{X_n\}\right] = \int_{E_2} f(X_i, y) K_{X_i}(dy), q.c. \qquad (1.3)
$$

Lema 1.2 *Se $g : E_1 \times E_2 \longrightarrow [0, 1]$ é $\mathcal{E}_1 \times \mathcal{E}_2$-mensurável, então*

$$
E\left(\prod_{i=1}^{\infty} g(X_i, J_i) \mid \{X_n\}\right) = \prod_{i=1}^{\infty} E\left(g(X_i, J_i) \mid \{X_n\}\right). \qquad (1.4)
$$

Passamos agora a expôr a demonstração da proposição anterior, tal como é apresentada em [24], pág.136.

Dem. Seja $f \geq 0$, $\mathcal{E}_1 \times \mathcal{E}_2$-mensurável. Então

$$
\begin{aligned}
\Psi_{N^*}(f) &= E\{\exp[-N^*(f)]\} = E\left\{\exp\left[-\sum_{i=1}^{\infty} f(X_i, J_i)\right]\right\} \\
&= E\left[\prod_{i=1}^{\infty} e^{-f(X_i, J_i)}\right] = E\left[E\left(\prod_{i=1}^{\infty} e^{-f(X_i, J_i)} \mid \{X_n\}\right)\right] \\
&= E\left[\prod_{i=1}^{\infty} E\left(e^{-f(X_i, J_i)} \mid \{X_n\}\right)\right], \text{ por (1.4)} \\
&= E\left[\prod_{i=1}^{\infty} \left(\int_{E_2} e^{-f(X_i, y)} K_{X_i}(dy)\right)\right], \text{ por (1.3)}.
\end{aligned}
$$

Considere-se $\theta(X_i) = \int_{E_2} e^{-f(X_i,y)} K_{X_i}(dy)$ tal que $0 \leq \theta \leq 1$. Temos então

$$\begin{aligned}
\psi_{N^*}(f) &= E\left\{\prod_{i=1}^{\infty} \theta(X_i)\right\} = E\left\{\exp\left[-\sum_{i=1}^{\infty}(-\log\theta(X_i))\right]\right\} \\
&= E\left\{\exp\left[-\int_{E_1}(-\log\theta(x)\,N_1 dx)\right]\right\} \\
&= \Psi_{N_1}(-\log\theta)
\end{aligned}$$

mas porque N_1é um PRM(μ) obtém-se a partir de (1.1)

$$\begin{aligned}
\Psi_{N^*}(f) &= \exp\left\{-\int_{E_1}\left(1 - e^{-(-\log\theta(x))}\right)\mu(dx)\right\} \\
&= \exp\left\{-\int_{E_1}(1-\theta(x))\,\mu(dx)\right\} \\
&= \exp\left\{-\int_{E_1}\left(1 - \int_{E_2} e^{-f(x,y)} K_x(dy)\right)\mu(dx)\right\} \\
&= \exp\left\{-\int_{E_1\times E_2}(1 - e^{-f(x,y)}\mu(dx)\,K_x(dy))\right\} \\
&= \exp\left\{-\int_{E_1\times E_2}(1 - e^{-f(x,y)}\mu^*(dx,dy))\right\},
\end{aligned}$$

que é o Funcional de Laplace do PRM(μ^*), onde $\mu^*(dx,dy) = \mu(dx)\,K_x(dy)$, provando-se assim o resultado. ■

1.3 Convergência Vaga

O estudo de certas propriedades inerentes aos processos estocásticos, tais como, o comportamento limite de extremos, de somas e covariâncias amostrais, e o próprio problema da estimação, recorre frequentemente à teoria da *convergência fraca* dos processos pontuais. Para se trabalhar com este tipo de convergência, tem toda a utilidade estabecer-se em $M_p(E)$, a noção de *convergência vaga* e a topologia associada, a *topologia vaga*. Como iremos ver, esta topologia confere a $M_p(E)$, o estatuto de espaço métrico, completo e separável.

Na nossa exposição apresentaremos inicialmente, alguns resultados para a topologia vaga referentes ao espaço $M_+(E)$, o espaço de todas as medidas de Radon não negativas, definidas em $(E,\mathcal{E})$. Veremos que toda a teoria da topologia vaga pode ser transferida de $M_+(E)$ para $M_p(E)$.

Comecemos por considerar novamente E, como um espaço topológico de Hausdorff, localmente compacto e com bases contáveis. Seja ρ uma métrica que torna E num espaço métrico, completo e separável. Seja ainda $C_K(E)$ o espaço das funções reais contínuas definidas em E com suporte compacto, isto é., $f \in C_K(E)$

sse f é continua e $\exists K \subseteq E$, compacto, tal que $f(x) = 0$ se $x \in K^c$. Tome-se $C_K^+(E)$, o subconjunto de $C_K(E)$, constituído por todas as funções contínuas não negativas de suporte compacto. Defina-se $\mathcal{M}_+(E)$ como a mais pequena σ-álgebra de subconjuntos de $M_+(E)$, que torna mensuráveis as aplicações da forma $H_f : M_+(E) \longrightarrow \mathbb{R}^+$ e que transforma qualquer medida m em $m(f) = \int_E f dm$, para funções $f \in C_K^+(E)$. Repare-se que à semelhança de $M_p(E)$, um elemento aleatório em $M_+(E)$ é uma aplicação mensurável do espaço probabilístico $(\Omega, \mathcal{A}, P)$ para o espaço $(M_+(E), \mathcal{M}_+(E))$.

Definição 1.5 *Dados μ_n e μ pertencentes a $M_+(E)$, diz-se que μ_n* ***converge vagamente para*** *μ, e escreve-se $\mu_n \xrightarrow{v} \mu$, se $\mu_n(f) \to \mu(f)$ em $\mathbb{R}$, para toda a função $f \in C_K^+(E)$ e onde $\mu_n(f) = \int_E f d\mu_n$.*

A **topologia vaga**, tal como é referido em [24], pág.140, pode ser construída em $M_+(E)$, definindo-se uma colecção de conjuntos abertos. Esses conjuntos são obtidos da seguinte forma: tome-se a sub-base em $M_+(E)$, constituída pelos conjuntos do tipo

$$\{\mu \in M_+(E) : s < \mu(f) < t\},$$

para alguma $f \in C_K^+(E)$ e com $s < t$. Considerem-se todas as intersecções finitas dos elementos da sub-base para se formar uma base. Então os conjuntos abertos são obtidos tomando as uniões dos elementos dessa base. Uma vizinhança para $\mu \in M_+(E)$ será então um conjunto da forma

$$\left\{\nu \in M_+(E) : \exists I \text{ finito}, f_i \in C_K^+(E),\ i \in I,\ |\nu(f_i) - \mu(f_i)| < \delta, i \in I\right\}.$$

Definida a topologia vaga em $M_+(E)$, pode formar-se naturalmente a σ-álgebra gerada pelos conjuntos abertos, i.e., a σ-álgebra dos boreleanos para a topologia vaga: $\mathcal{B}_v(M_+(E))$. Ora, Jagers (1974), provou que esta σ-álgebra coincide precisamente com $\mathcal{M}_+(E)$.

Interessa-nos agora estabelecer o conceito de convergência vaga considerando-se sucessões da forma $\{\mu_n(A), A \in \mathcal{E}\}$. Essa identificação é traduzida pelos resultados seguintes:

Proposição 1.6 *Sejam $\mu, \mu_1, \mu_2, ..., \mu_n, ...$, elementos em $M_+(E)$. Então são equivalentes as condições:*

a) $\mu_n \xrightarrow{v} \mu$.

b) $\mu_n(A) \to \mu(A)$, *para todo o conjunto A relativamente compacto para o qual se tenha $\mu(\delta A) = 0$ (a fronteira de A tem medida nula).*

c) Para todo o compacto K, tem-se

$$lim_{n\to\infty} (\sup \mu_n (K)) \leq \mu (K)$$

e para todo o aberto, relativamente compacto G

$$lim_{n\to\infty} (\inf \mu_n (G)) \geq \mu (G) .$$

Dem. Veja-se em [24], pág.142. ∎

A próxima proposição assegura-nos que todo o espólio topológico e métrico de $M_+ (E)$ pode ser transferido para $M_p (E)$.

Proposição 1.7 *O espaço das medidas pontuais $M_p (E)$ é fechado em $M_+ (E)$ para a convergência vaga.*

Dem. Veja-se em [24], pág.145. ∎

Temos agora a garantia de que dada uma sucessão de medidas m_n em $M_p (E)$, que convirja vagamente para uma medida m, essa medida também está em $M_p (E)$. O conceito de convergência vaga pode então ser enunciado em termos do espaço das medidas pontuais:

$$m_n \xrightarrow{v} m \text{ em } M_p (E) \text{ se } m_n (f) = \int f dm_n \to m (f) = \int f dm,$$

para toda a função $f \in C_K^+ (E)$. Uma outra consequência a reter da proposição anterior, é que a σ-álgebra dos boreleanos gerada pelos abertos para a topologia vaga em $M_p (E)$, representada por $\mathcal{B}_v (M_p (E))$, coincide com $\mathcal{M}_p (E)$, ou seja,

$$\mathcal{B}_v (M_p (E)) = \mathcal{M}_p (E) .$$

Deste modo, pode descrever-se de forma semelhante, uma estrutura de medida pontual definida à custa de uma ou de outra σ-álgebra.

Proposição 1.8 *A topologia vaga em $M_p (E)$ e em $M_+ (E)$ é metrizável, obtendo-se um espaço métrico, completo e separável.*

Uma métrica que concede a $M_p (E)$ e a $M_+ (E)$ as características mencionadas é a designada ***métrica vaga***

$$d (\mu, \mu') = \sum_{i=1}^{\infty} \frac{1}{2^i} [1 - \exp \{- |\mu (h_i) - \mu' (h_i)|\}] ,$$

com $\mu, \mu' \in M_p (E)$ (ou $M_+ (E)$) e $\mathcal{H} = \{h_i\}$ uma colecção contável densa de funções em $C_K^+ (E)$.

Dem. A demonstração pode ser examinada em [24], pág.147. ∎

Observação 1.3 Dada $\{\mu_n\}$ uma sequência de medidas em $M_p(E)$ ou em $M_+(E)$, sabemos que $\mu_n \xrightarrow{v} \mu$ sse $\mu_n(h_l) \rightarrow \mu(h_l)$ para toda a função $h_l \in \mathcal{H}$. Por outro lado resulta por definição da métrica vaga $d(\mu_n, \mu) \rightarrow 0$ sse $|\mu_n(h_l) - \mu(h_l)| \rightarrow 0$, ou seja podemos escrever que

$$\mu_n \xrightarrow{v} \mu \quad sse \quad d(\mu_n, \mu) \rightarrow 0.$$

A continuidade é um marco necessário ao estudo da convergência fraca de elementos aleatórios e que culmina com o famoso *teorema da aplicação contínua (Billingsley, 1968)*.O próximo resultado irá ajudar-nos a esclarecer esse teorema, quando os elementos em causa são processos pontuais.

Proposição 1.9 *Sejam E e E', dois espaços localmente compactos com bases contáveis. Suponhamos que a aplicação $T : E \rightarrow E'$ é contínua e satisfaz para todo o subconjunto compacto $K' \subset E'$*

$$T^{-1}(K') \text{ é compacto em } E. \tag{1.5}$$

Então a aplicação

$$\begin{array}{rccc} T^* : & M_+(E) & \rightarrow & M_+(E') \\ & \mu & \rightarrow & T^*(\mu) = \mu \circ T^{-1} \end{array}$$

é contínua. A restrição de T^* a $M_p(E)$ é dada por

$$\begin{array}{rccc} T^*_{|M_p(E)} : & M_p(E) & \rightarrow & M_p(E') \\ & \sum \varepsilon_{x_i} & \rightarrow & T^*(\sum \varepsilon_{x_i}) = \sum \varepsilon_{T(x_i)} \end{array}$$

e é contínua se a aplicação T^* sobre os pontos que caracterizam a medida pontual fôr contínua e verificar (1.5).

Dem. Sejam $\{\mu_n\}_{n\in\mathbb{N}}$ e μ elementos em $M_+(E)$ tal que $\mu_n \xrightarrow{v} \mu$. Dada a função $f \in C_K^+(E')$, tem-se

$$T^*(\mu_n(f)) = \mu_n \circ T^{-1}(f) = \mu_n(f \circ T).$$

Por (1.5) , a aplicação $f \circ T$ pertence a $C_K^+(E)$, pois o suporte de $f \circ T$ é dado por T^{-1}(suporte de f), sendo por sua vez compacto em E. Dado que $\mu_n \xrightarrow{v} \mu$, então

$$T^*(\mu_n(f)) = \mu_n(f \circ T) \rightarrow \mu(f \circ T) = \mu \circ T^{-1}(f),$$

e portanto,

$$T^*(\mu_n) \xrightarrow{v} T^*(\mu). \qquad \blacksquare$$

1.4 Convergência Fraca de Processos Pontuais

Vimos na secção anterior que o espaço $M_p(E)$ goza das propriedades de existência de métrica, de completude e de separabilidade. Ora, essas propriedades permitem estabelecer a equivalência entre a convergência fraca e a convergência em distribuição. Essa equivalência é aliás, válida para qualquer espaço que usufrua dessas mesmas características. Com vista a desenvolver o estudo da convergência fraca para medidas pontuais aleatórias, apresentamos inicialmente uma revisão deste tema.

Seja $\mathcal{X}$ um espaço métrico, completo e separável, munido de uma métrica ρ. Sejam $\mathcal{B}(\mathcal{X})$ a σ-álgebra dos borelcanos gerada pelos abertos de $\mathcal{X}$ e $(\Omega, \mathcal{A}, P)$ um espaço de probabilidade. Um elemento aleatório X em $\mathcal{X}$, é, por definição, uma aplicação mensurável do espaço $(\Omega, \mathcal{A})$ para o espaço $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. Alguns exemplos de elementos aleatórios são as variáveis aleatórias ($\mathcal{X} = \mathbb{R}$); os vectores aleatórios em $\mathcal{X} = \mathbb{R}^k$; as funções aleatórias no espaço $\mathcal{X} = C([0,\infty))$-espaço das funções contínuas em $[0,\infty)$; os processos pontuais no espaço $M_p(E)$, entre outros.Considere-se em $\mathcal{X}$, a sequência de elementos aleatórios $\{X_n\}_{n\in\mathbb{N}_0}$. Defina-se em $\mathcal{B}(\mathcal{X})$, a sucessão de medidas de probabilidade ou distribuições de X_n, por

$$P_n(B) = P \circ X_n^{-1}(B) = P[X_n \in B] = P[\{w \in \Omega : X_n(w) \in B\}], \qquad B \in \mathcal{B}(\mathcal{X}).$$

Diz-se que X_n **converge fracamente** para o elemento aleatório X_0, e escreve-se $X_n \Rightarrow X_0$ ou $P_n \Rightarrow P_0$ ou ainda $X_n \xrightarrow{d} X_0$, sse

$$E[f(X_n)] = \int_{\mathcal{X}} f(x)\, P_n(dx) \to \int_{\mathcal{X}} f(x)\, P_0(dx) = E[f(X_0)],$$

para toda a função $f \in C(\mathcal{X})$, a classe das funções reais, limitadas e contínuas, com valores em $\mathcal{X}$.

Uma outra forma de definir a convergência fraca é com base no conceito de *sucessão relativamente compacta.* Diz-se que a sucessão $\{X_n\}_{n\in\mathbb{N}_0}$ é **relativamente compacta** sse a sucessão das suas distribuições $\{P_n\}$, é relativamente compacta, i.e., sse toda a subsequência $\{X_{n'}\} \subset \{X_n\}$, contém uma subsequência $\{X_{n''}\} \subset \{X_{n'}\}$ que converge fracamente para algum elemento aleatório $X \in \mathcal{X}$, podendo X ser diferente de subsequência para subsequência. Ora, para que se tenha $X_n \Rightarrow X_0$ prova-se que é necessário e suficiente, que a sequência $\{X_n\}$ seja relativamente compacta e que todos os limites de subsequências de distribuições de X_n sejam iguais à distribuição de X_0. Como se vê pela definição, a relatividade compacta não é simples de ser verificada. Mas, o teorema de Prohorov (1956)[2], garante que sendo $\mathcal{X}$ um espaço métrico, completo e separável, pode estabelecer-se uma equivalência entre relatividade compacta e *Tightness*[3]. A sucessão $\{X_n\}_{n\in\mathbb{N}_0}$

[2] Veja-se em apêndice.

[3] Não se encontrou na literatura consultada um termo português correspondente ao conceito de *tight* (apertado), optando -se por se utilizar o termo em inglês.

diz-se **tight** sse a sucessão de distribuições $\{P_n\}$ é tight, i.e., sse para todo $\varepsilon > 0$, existe um compacto $K = K_\varepsilon$ em $\mathcal{B}(\mathcal{X})$, tal que $P(X_n \in K) > 1 - \varepsilon, \forall n \in \mathbb{N}_0$. A vantagem em utilizar-se o conceito de sucessão tight em substituição de sucessão relativamente compacta, está no facto de em espaços com as propriedades como as de $\mathcal{X}$ (existência de métrica, separabilidade e completude), ser normalmente mais fácil trabalhar-se com a caracterização de conjuntos compactos. O próprio conceito de sucessão tight, tem as suas ramificações e variantes, quando aplicado a um particular espaço métrico. É o caso do exemplo seguinte.

Exemplo 1.6 Seja $C = C([0,1])$,o espaço métrico,completo e separável, das funções contínuas definidas em $[0,1]$, munido da métrica usual definida neste espaço, a métrica do supremo:

$$\rho(x,y) = \sup_{0 \le t \le 1} |x(t) - y(t)|,$$

com $x, y \in C$. Seja ainda $\mathcal{B}(C)$, a σ-álgebra dos boreleanos de C, gerada pelos conjuntos abertos de C. Para cada $x \in C$, defina-se a função módulo de continuidade

$$\kappa_\delta(x) = \sup_{0 \le s,t \le 1, |s-t| < \delta} |x(t) - x(s)|, \text{ com } 0 < \delta < 1.$$

Considerem-se X_n e X elementos aleatórios de C, definidos num espaço de probabilidade $(\Omega, \mathcal{A}, P)$, ou seja, $X_n : \Omega \to C([0,1])$ e $X : \Omega \to C([0,1])$ tal que $X_n^{-1}(\mathcal{B}(C)) \subset \mathcal{A}$ e $X^{-1}(\mathcal{B}(C)) \subset \mathcal{A}$. Isto significa que para cada $w \in \Omega$, $X_n(\cdot, w)$ e $X(\cdot, w)$ são funções contínuas, e para cada $t \in [0,1]$, $X_n(t)$ e $X(t)$ são variáveis aleatórias. Nestas condições, diz-se que $X_n \xrightarrow{d} X$ ou que $X_n \Rightarrow X$ sse $\{X_n\}_{n \in \mathbb{N}_0}$ é tight e as distribuições dimensionais finitas de $X_n(t)$ convergirem (em distribuição ou fracamente) para as de $X(t)$, i.e., $\forall k \in \mathbb{N}, 0 \le t_1 \le ... \le t_k \le 1$

$$(X_n(t_1), ..., X_n(t_k)) \xrightarrow{d} (X(t_1), ..., X(t_k)) \text{ em } \left(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)\right).$$

Para provar a Tightness, utiliza-se uma outra caracterização de conjuntos compactos, para o caso especial do espaço C, a qual resulta de um teorema clássico da análise, o teorema de Arzelà-Ascol. O teorema referido diz que um subconjunto $A \in C([0,1])$, tem aderência $(\overline{A})$ compacta sse

$$\sup_{x \in A} |x(0)| < \infty \quad \text{e} \quad \lim_{\delta \to 0} \sup_{x \in A} \sup_{|s-t| < \delta} |x(t) - x(s)| = 0$$

Ora, com base neste resultado, pode provar-se[4] que $\{X_n\}_{n \in \mathbb{N}_0}$ é tight sse:

a) $\forall \varepsilon > 0, \exists a \ge 0$ tal que $P(|X_n(0)| > a) \le \varepsilon$,

[4] Veja-se por exemplo [2], pág.55.

b) $\forall \eta > 0, \forall \varepsilon > 0, \exists \delta = \delta(\eta, \varepsilon)$ com $0 < \delta < 1$ e $\exists n_0 = n_0(\eta, \varepsilon) \in \mathbb{N}$, tal que, para $n > n_0$ tem-se

$$P\left[\sup \kappa_{X_n}(\delta) \geq \eta\right] = P\left[\sup_{0 \leq s,t \leq 1, |s-t| < \delta} |X_n(t) - X_n(s)| \geq \eta\right] < \varepsilon.$$

Repare-se que este exemplo pode ser extendido a outros espaços tais como $C([a,b])$, ou $C(\mathbb{R}^p)$, com $p \geq 1$, pois é possível estabelecer-se um isomorfismo entre$(C([0,1]), \rho)$ e os referidos espaços, que preserva todas as propriedades inerentes à existência de métrica, completude e separabilidade.

Um dos teoremas fundamentais e de grande utilidade na teoria da convergência fraca, é o teorema da aplicação contínua, que consiste no seguinte:

Teorema 1.1 *(Teorema da aplicação contínua, Billingsley,1968). Sejam* $(\mathcal{X}_1, \rho_1)$ *e* $(\mathcal{X}_2, \rho_2)$ *dois espaços métricos e suponha-se que* $\{X_n\}_{n \in \mathbb{N}_0}$ *é uma sequência de elementos aleatórios de* $\mathcal{X}_1$ *tal que* $X_n \Rightarrow X_0$*, com* $X_0 \in \mathcal{X}_1$*. Se* $h : \mathcal{X}_1 \to \mathcal{X}_2$ *verificar*

$$P[X_0 \in D_h] = P[X_0 \in \{x_1 \in \mathcal{X}_1 : h \text{ é descontínua em } x_1\}] = 0$$

então

$$h(X_n) \Rightarrow h(X_0)$$

em $\mathcal{X}_2$.

Dem. Omite-se. Para um estudo mais aprofundado veja-se a demonstração em [2], págs 30-31. ∎

Voltemos agora ao espaço $M_p(E)$. Já sabemos que todo o processo pontual, é identificado pelo respectivo Funcional de Laplace e por conseguinte, é natural que se possa estabelecer a convergência fraca do processo, em termos do seu funcional. Mas, por outro lado, como foi visto anteriormente, a convergência fraca está fortemente ligada ao conceito de sucessão tight. Ora, é nesta relação estreita entre funcional de Laplace e sucessão tight, que se rege a convergência fraca de medidas aleatórias pontuais. Neste sentido, apresentamos o seguinte resultado:

Proposição 1.10 *Sejam* $\xi, \xi_0, \ldots, \xi_n, \ldots$*, elementos aleatórios de* $M_p(E)$*. Então*

$$\xi_n \Rightarrow \xi \quad sse \quad \Psi_{\xi_n}(f) \to \Psi_\xi(f),$$

para toda a função $f \in C_K^+(E)$.

Dem. Apresentamos aqui, apenas um breve esboço da demonstração, a qual pode ser examinada detalhadamente em [24], pág. 153. Suponha-se então que $\xi_n \Rightarrow \xi$ em $M_p(E)$. Para $f \in C_K^+(E)$ defina-se a aplicação contínua

$$\begin{array}{rccc} T_f: & M_p(E) & \rightarrow & \mathbb{R}_0^+ \\ & \mu & \rightarrow & T_f(\mu) = \mu(f) = \int_E f d\mu \end{array}$$

Ora, o teorema da aplicação contínua garante que se $\xi_n \Rightarrow \xi$ então

$$T_f(\xi_n) = \xi_n(f) \Rightarrow T_f(\xi) = \xi(f),$$

e pelo teorema da convergência dominada, vem,

$$\psi_{\xi_n}(f) = E\{\exp[-\xi_n(f)]\} \rightarrow E\{\exp[-\xi(f)]\} = \psi_\xi(f).$$

A recíproca prova-se com base no seguinte resultado: $\{\xi_n\}$ é tight em $M_p(E)$ sse $\{\xi_n(f)\}$ é tight em $\mathbb{R}$, para $f \in C_K^+(E)$. ■

Apresentamos por último, dois resultados essenciais na teoria dos processos pontuais, que se referem à relação de equivalência entre convergências fraca e vaga, permitindo desse modo, estabelecer a ponte entre processos pontuais e variação regular, assunto a que nos dedicaremos nos próximos capítulos.

Proposição 1.11 *Para cada natural n, seja $\{X_{n,k}, k \geq 1\}$, uma sucessão de elementos aleatórios i.i.d, que tomam valores em $(E, \mathcal{E})$. Defina-se $N_n := \sum_{k=1}^{n} \varepsilon_{X_{n,k}}$ e admita-se que $N \frown PRM(\mu)$, com μ medida de radon em $(E, \mathcal{E})$. Então*

$$N_n \Rightarrow N$$

em $M_p(E)$, sse

$$nP[X_{n,1} \in \cdot] \xrightarrow{v} \mu,$$

em E.

Dem. Seja $f \in C_K^+(E)$. Tem-se

$$\begin{aligned} \Psi_{N_n}(f) &= E\{\exp[-N_n(f)]\} = E\left\{\exp\left[-\sum_{k=1}^{n} f(X_{n,k})\right]\right\} \\ &= (E\{\exp[-f(X_{n,1})]\})^n \\ &= \left(1 - \frac{\int_E \left(1 - e^{-f(x)}\right) nP[X_{n,1} \in dx]}{n}\right)^n \\ &\rightarrow \exp\left\{-\int_E \left(1 - e^{-f(x)}\right) \mu(dx)\right\} = \Psi_N(f), \end{aligned}$$

sse $nP[X_{n,1} \in \cdot] \xrightarrow{v} \mu$. ■

Proposição 1.12 *Nas condições da proposição anterior, defina-se*

$$\xi_n := \sum_{k=1}^{\infty} \varepsilon_{\left(\frac{k}{n}, X_{n,k}\right)}$$

e admita-se que ξ é um PRM em $[0,\infty) \times E$ com intensidade $dt \times d\mu$. Então $\xi_n \Rightarrow \xi$ em $M_p([0,\infty) \times E)$ sse

$$nP\left[X_{n,1} \in \cdot\right] \xrightarrow{v} \mu \text{ em } E. \tag{1.6}$$

Dem. Esta demonstração segue de perto a apresentada em [24], pág.155.

Mostremos em primeiro lugar que se $nP\left[X_{n,1} \in \cdot\right] \xrightarrow{v} \mu$ então $\xi_n \Rightarrow \xi$, provando para esse efeito que o funcional de Laplace de ξ_n converge para o de ξ. Ora, para $f \in C_K^+([0,\infty) \times E)$, o funcional de Laplace de ξ_n é da forma

$$\begin{aligned} \Psi_{\xi_n}(f) &= E\left\{\exp\left[-\xi_n(f)\right]\right\} = E\left\{\exp\left[-\sum_k f\left(\frac{k}{n}, X_{n,k}\right)\right]\right\} \\ &= \prod_k \left(1 - \int_E \left(1 - e^{-f\left(\frac{k}{n}, x\right)}\right) P\left[X_{n,1} \in dx\right]\right). \end{aligned}$$

Provemos que $\Psi_{\xi_n}(f) \to \Psi_\xi(f)$ sse

$$-\log \Psi_{\xi_n}(f) = -\sum_k \log\left(1 - \int_E \left(1 - e^{-f\left(\frac{k}{n}, x\right)}\right) P\left[X_{n,1} \in dx\right]\right) \to -\log \Psi_\xi(f).$$

Defina-se λ_n por

$$\lambda_n(ds, dx) = \sum_k \varepsilon_{\frac{k}{n}}(ds)\, P\left[X_{n,1} \in dx\right].$$

De (1.6) segue-se

$$\lambda_n(ds, dx) \xrightarrow{v} ds\mu(dx).$$

Então,

$$\begin{aligned} &\sum_k \int_E \left(1 - e^{-f\left(\frac{k}{n}, x\right)}\right) P\left[X_{n,1} \in dx\right] \\ &= \int\int_{[0,\infty)\times E} \left(1 - e^{-f\left(\frac{k}{n}, x\right)}\right) d\lambda_n \\ &\underset{n\to\infty}{\longrightarrow} \int\int_{[0,\infty)\times E} \left(1 - e^{-f(s,x)}\right) ds\mu(dx) \\ &= -\log \Psi_\xi(f). \end{aligned} \tag{1.7}$$

Além disso, se o conjunto compacto K, for o suporte de f em $[0,\infty) \times E$, então existe um subconjunto compacto $A \subset E$ tal que por (1.6), vem

$$\sup_{k\geq 1} \int_E \left(1 - e^{-f\left(\frac{k}{n},x\right)}\right) P\left[X_{n,1} \in dx\right] \leq P\left[X_{n,1} \in A\right] \underset{n\to\infty}{\longrightarrow} 0. \tag{1.8}$$

Tendo em conta a expansão,

$$\log(1+z) = z\left(1+\varepsilon(z)\right), \qquad \text{com} \qquad |\varepsilon(z)| \leq |z| \qquad \text{se} \qquad |z| \leq 1/2,$$

vem por (1.7) e (1.8) e para n suficientemente grande,

$$\begin{aligned}
&\left| -\log \Psi_{\xi_n}(f) - \sum_k \int_E \left(1 - e^{-f\left(\frac{k}{n},x\right)}\right) P\left[X_{n,1} \in dx\right] \right| \\
\leq\ & \sum_k \left(\int_E \left(1 - e^{-f\left(\frac{k}{n},x\right)}\right) P\left[X_{n,1} \in dx\right] \right)^2 \\
\leq\ & \left(\sup_{k\geq 1} \int_E \left(1 - e^{-f\left(\frac{k}{n},x\right)}\right) P\left[X_{n,1} \in dx\right] \right) \\
& \times \sum_k \int_E \left(1 - e^{-f\left(\frac{k}{n},x\right)}\right) P\left[X_{n,1} \in dx\right] \underset{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

e por conseguinte $\psi_{\xi_n}(f) \to \psi_\xi(f)$.

Reciprocamente, suponhamos que $\psi_{\xi_n}(f) \to \psi_\xi(f)$, qualquer que seja a função $f \in C_K^+([0,\infty) \times E)$. Seja $f(s,x) = I_{[0,1]}(s)\, g(x)$, onde $g \in C_K^+(E)$. Repare-se que esta função f não está definida em $C_K^+([0,\infty) \times E)$. Mas tomando as desigualdades

$$h_l^-(s)\, g(x) \leq f(s,x) \leq h_l^+(s)\, g(x),$$

onde $0 \leq h_l^-(s) \uparrow I_{[0,1]}(s)$ e $h_l^+(s) \downarrow I_{[0,1]}(s)$ quando $l \to \infty$ e para $h_l^\pm(s)$ pertencente a $C_K^+([0,\infty) \times E)$, pode provar-se, seguindo o mesmo raciocínio anterior, que

$$\begin{aligned}
\psi_{\xi_n}(f) &= E\left\{\exp\left[-\xi_n(f)\right]\right\} = E\left\{\exp\left[-\sum_{k=1}^{n} g\left(X_{n,k}\right)\right]\right\} \\
&\to \exp\left[-\int_E \left(1 - e^{-g(x)} d\mu\right)\right],
\end{aligned}$$

ou seja, $\sum_{k=1}^{n} \varepsilon_{X_{n,k}} \Rightarrow PRM(\mu)$. Logo conclui-se que

$$nP\left[X_{n,1} \in \cdot\right] \xrightarrow{v} \mu,$$

pela proposição 1.11. ■

Capítulo 2

Processos Pontuais Limite associados a Modelos de Médias Móveis dirigidos por Erros com Distribuição de Variação Regular nas Caudas

Os processos autoregressivos e de médias móveis (ARMA), são utilizados frequentemente para modelarem séries temporais e tornam-se particularmente interessantes quando são dirigidos por erros que apresentem nas caudas distribuição de *variação regular*. Esse interesse advém de se poderem extrapolar propriedades que se prendem com o comportamento limite desses processos, tais como o estudo de extremos e de somas e funções de covariâncias amostrais. Neste capítulo restringiremos o nosso estudo à análise da teoria limite de processos pontuais, quando associados a modelos de médias móveis estacionários, gerados por erros com distribuição de variação regular de índice $-\alpha$, com $\alpha > 0$. Para já apresentamos uma breve revisão de alguns resultados das funções de variação regular de variávies reais, que irão ser utilizados ao longo deste trabalho.

2.1 Funções de Variação Regular. Conceitos Básicos

Definição 2.1 *Seja $U : \overline{\mathbb{R}}^+ \longrightarrow \overline{\mathbb{R}}^+$ uma função mensurável. Diz-se que U é de* ***variação regular*** *em ∞ com índice $\rho \in \mathbb{R}$ e escreve-se $U \in RV_\rho$, sse*

$$\lim_{t \to \infty} \frac{U(tx)}{U(t)} = x^\rho,$$

*para todo $x > 0$. O expoente ρ também é designado por **expoente de variação regular**.*

Vemos pela definição que as funções de variação regular comportam-se assintoticamente como as funções polinomiais ($\rho > 0$) Quando $\rho = 0$, a função U diz-se de **variação lenta** em ∞. As funções de variação lenta são normalmente representadas na literatura por $L(x)$ e verificam portanto

$$\lim_{t\to\infty} \frac{L(tx)}{L(t)} = 1.$$

Se $U(x) \in RV_\rho$, então $U(x)/x^\rho \in RV_0$, pois

$$\lim_{t\to\infty} \frac{U(tx)/(tx)^\rho}{U(t)/t^\rho} = \lim_{t\to\infty} \frac{x^\rho}{x^\rho} = 1.$$

Fazendo $L(x) = U(x)/x^\rho$, podemos escrever $U(x) = x^\rho L(x)$, o que significa que qualquer função de variação regular de expoente ρ pode ser identificada pela função de variação lenta que lhe está associada. Doravante sempre que escrevermos $f(x) \sim g(x)$, significa que $\lim_{x\to\infty} f(x)/g(x) = 1$. Note-se que a definição anterior pode ser extendida à origem: U é de variação regular em 0 sse $U(x^{-1})$ é de variação regular em ∞. Alguns exemplos de funções de variação regular em ∞ são a própria função x^ρ e a função $(1+x^2)^p$ com expoente de variação $2p$. Todas as potências de $|\log x|$ são de variação lenta em 0 e em ∞. Na Teoria das Probabilidades também encontramos alguns exemplos de funções de distribuição em que as caudas apresentam variação regular em ∞. É o caso das funções de distribuição que verificam a propriedade

$$1 - F(x) \sim Cx^{-\alpha}, x \to \infty, \text{ com } C > 0 \text{ e } \alpha > 0.$$

Observe-se que, para $0 < \alpha \leq 2$, estas funções de distribuição pertencem à classe das *leis estáveis*.

Proposição 2.1 (de Haan,1970; Feller,1971) *a) $U : \overline{\mathbb{R}}^+ \longrightarrow \overline{\mathbb{R}}^+$ é de variação regular se existe uma função h, tal que para todo $x > 0$*

$$\lim_{x\to\infty} \frac{U(tx)}{U(t)} = h(x).$$

Neste caso $h(x) = x^\rho$ para algum $\rho \in \mathbb{R}$ e $U \in RV_\rho$.

b) Seja $U : \overline{\mathbb{R}}^+ \longrightarrow \overline{\mathbb{R}}^+$ uma função monótona e mensurável. Suponha-se que existem duas sucessões de termos positivos, $\{\lambda_n\}$ e $\{a_n\}$, tais que

$$\frac{\lambda_{n+1}}{\lambda_n} \to 1 \quad e \quad a_n \to \infty. \tag{2.1}$$

Se para todo $x > 0$ existir uma função $H(x)$ finita e positiva, tal que

$$\lim_{n\to\infty} \lambda_n U(a_n x) = H(x), \tag{2.2}$$

então U é de variação regular de expoente ρ, para algum $\rho \in \mathbb{R}$, tendo-se além disso $H(x) = Cx^\rho, C > 0$.

Dem. Esta demonstração segue de perto a apresentada em [24], pág. 14.

a) A função h é mensurável, pois é o limite de funções mensuráveis. Por outro lado para $x > 0$ e $y > 0$, tem-se

$$\frac{U(txy)}{U(t)} = \frac{U(txy)}{U(tx)} \times \frac{U(tx)}{U(txy)}$$

e fazendo $t \to \infty$, vem

$$h(xy) = h(y)\, h(x),$$

logo $h(x) = x^\rho$.

b) Assuma-se que U é não decrescente. Como $a_n \to \infty$, para cada t existe uma função $n(t)$, definida por

$$n(t) = \inf\{m : a_{m+1} > t\}$$

e tal que $a_{n(t)} \leq t < a_{n(t)+1}$. Pela monotonia da função U e para $x > 0$ tem-se

$$\left(\frac{\lambda_{n(t)+1}}{\lambda_{n(t)}}\right)\left(\frac{\lambda_{n(t)}U\left(a_{n(t)}x\right)}{\lambda_{n(t)+1}U\left(a_{n(t)+1}\right)}\right) \leq \frac{U(tx)}{U(t)} \leq \left(\frac{\lambda_{n(t)}}{\lambda_{n(t)+1}}\right)\left(\frac{\lambda_{n(t)}U\left(a_{n(t)+1}x\right)}{\lambda_{n(t)+1}U\left(a_{n(t)}\right)}\right).$$

Tomando $t \to \infty$ e usando (2.1) e (2.2) resulta que $\lim\limits_{t\to\infty} \frac{U(tx)}{U(t)} = 1.H(x)/H(1)$. O resto da demonstração resulta de a). ■

Exemplo 2.1 Considere-se $\{X_n, n \geq 1\}$ uma sequência i.i.d. de v.a.'s com distribuição comum F e seja

$$M_n = \max_{1\leq i\leq n} X_i.$$

A função de distribuição de M_n é dada por

$$P[M_n \leq x] = P\left\{\bigcap_{i=1}^{n} [X_i \leq x]\right\} = \prod_{i=1}^{n} P[X_i \leq x] = F^n(x).$$

Suponhamos que existe uma sequência de constantes positivas $\{a_n\}$ tais que

$$P\left[a_n^{-1}M_n \leq x\right] = F^n(a_n x) \longrightarrow \Phi_\alpha(x) = \exp\left\{-x^{-\alpha}\right\}, \quad \text{com } x \geq 0 \text{ e } \alpha > 0. \tag{2.3}$$

Se tomarmos logarítmos em (2.3), para $x > 0$, obtemos

$$\lim_{n\to\infty} n\,(-\log F(a_n x)) = x^{-\alpha}.$$

Utilizando a relação $-\log(1-z) \sim z$, quando $z \to 0$, no limite anterior ficamos com

$$\begin{aligned}\lim_{n\to\infty} n\,(-\log F(a_n x)) &= \lim_{n\to\infty} n\,(-\log(1-(1-F(a_n x)))) \\ &= \lim_{n\to\infty} n\,(1-F(a_n x)) = x^{-\alpha}. \end{aligned} \tag{2.4}$$

Usando agora a alínea b) da proposição anterior e o limite em (2.4), resulta que

$$1 - F(x) = x^{-\alpha}L(x), \quad x \to \infty, \quad \text{para algum } \alpha > 0.$$

Pode-se provar (veja-se por exemplo em [13] ou em [24]), que uma sucessão $\{a_n\}$ que satisfaz (2.4), é determinada pela relação assintótica

$$a_n \sim \left(\frac{1}{(1-F)}\right)^{\leftarrow}(n) = \inf\left\{x : \frac{1}{1-F(x)} \geq n\right\} = \inf\left\{x : 1 - F(x) \leq \frac{1}{n}\right\}.$$

Por outro lado em (2.4), se fizermos $x = 1$, vem que $\lim_{n\to\infty} n\,(1 - F(a_n)) = 1$, donde $1 - F(a_n) \sim \frac{1}{n}$. Mas uma vez que $1 - F(a_n) \sim (a_n)^{-\alpha} L(a_n)$, tem-se que

$$(a_n)^{-\alpha} L(a_n) \sim \frac{1}{n}, \text{ donde } a_n \sim n^{1/\alpha} L(n).$$

O teorema seguinte, conhecido por teorema de Karamata, é muito utilizado no estudo de limites que envolvam integrais de funções de variação regular. Nesse teorema assume-se que as funções são localmente integráveis. A respectiva demonstração pode ser analisada por exemplo em [24], págs. 18 e 19.

Teorema 2.1 (Karamata) *(i) Se $\rho \geq -1$ e $U \in RV_\rho$ então*

$$\int_0^x U(t)\,dt \in RV_{\rho+1},$$

verificando-se além disso

$$\lim_{x\to\infty} \frac{xU(x)}{\int_0^x U(t)\,dt} = \rho + 1.$$

Se $\rho < -1$ (ou se $\rho = -1$ e existir $\int_x^\infty U(t)\,dt$), então $U \in RV_\rho$ implica $\int_x^\infty U(t)\,dt \in RV_{\rho+1}$ e

$$\lim_{x\to\infty} \frac{xU(x)}{\int_x^\infty U(t)\,dt} = -\rho - 1.$$

(ii) Se U satisfaz

$$\lim_{x\to\infty} \frac{xU(x)}{\int_0^x U(t)\,dt} = \lambda \in (0,\infty),$$

então $U \in RV_{\lambda-1}$. Se $\int_x^\infty U(t)\,dt < \infty$ e

$$\lim_{x\to\infty} \frac{xU(x)}{\int_x^\infty U(t)\,dt} = \lambda \in (0,\infty),$$

então $U \in RV_{\lambda-1}$.

Corolário 2.1 *(Reresentação de Karamata). L é uma função de variação lenta ssc puder ser representada na forma*

$$L(x) = c(x)\exp\left\{\int_1^x t^{-1}\varepsilon(t)\,dt\right\},$$

para todo $x > 0$ e onde $c : \overline{\mathbb{R}}^+ \longrightarrow \overline{\mathbb{R}}^+$, $\varepsilon : \overline{\mathbb{R}}^+ \longrightarrow \overline{\mathbb{R}}^+$ tal que

$$\lim_{x\to\infty} c(x) = c \in (0,\infty) \quad e \quad \lim_{t\to\infty} \varepsilon(t) = 0.$$

Uma das propriedades fundamentais de que as funções de variação regular gozam, são as chamadas desigualdades de Potter ou *Potter Bounds.* Para um estudo mais aprofundado destas desigualdades aconselha-se a consulta do Teorema de Potter em [13], pág 25. Destacamos aqui apenas uma variante dessas propriedades, apresentada em [16], pág.106, por sinal a de maior utilidade para os resultados que iremos desenvolver no próximo capítulo.

Proposição 2.2 *Seja $U : \overline{\mathbb{R}}^+ \longrightarrow \overline{\mathbb{R}}^+$ uma função de variação regular, com expoente $\rho \in \mathbb{R}$. Então para todo $\varepsilon > 0$, existe um $t_0 = t_0(\varepsilon)$ tal que, para todo $t \geq t_0$ e $tx \geq t_0$, tem-se*

$$(1-\varepsilon)\,x^\rho \exp(-\varepsilon\,|\log x|) \leq \frac{U(tx)}{U(t)} \leq (1+\varepsilon)\,x^\rho \exp(\varepsilon\,|\log x|).$$

Dem. Utilizando a representação de Karamata, defina-se para $x \geq 1$

$$\frac{U(tx)}{U(t)} = \frac{c(tx)}{c(t)} \exp\left\{\int_1^x s^{-1}\rho(ts)\,ds\right\}.$$

Escolhendo um t_0 tal que, para $t \geq t_0$ e $tx \geq t_0$ implica que $\rho - \varepsilon < \rho(ts) < \rho + \varepsilon$, donde

$$\frac{c(tx)}{c(t)} \exp\left\{\int_1^x s^{-1}(\rho-\varepsilon)\,ds\right\} \leq \frac{U(tx)}{U(t)} \leq \frac{c(tx)}{c(t)} \exp\left\{\int_1^x s^{-1}(\rho+\varepsilon)\,ds\right\}$$

e portanto

$$\frac{c(tx)}{c(t)} \exp\{-(\rho-\varepsilon)\log x\} \leq \frac{U(tx)}{U(t)} \leq \frac{c(tx)}{c(t)} \exp\{(\rho+\varepsilon)\log x\}.$$

Como

$$\begin{aligned} \frac{c(tx)}{c(t)} \exp\{(\rho+\varepsilon)\log x\} &\leq \left(\frac{c(tx)}{c(t)} + \varepsilon\right) x^{\rho} \exp\{\varepsilon \log x\} \\ &\sim (1+\varepsilon)\, x^{\rho} \exp\{\varepsilon \log x\} \end{aligned}$$

e também,

$$\begin{aligned} \frac{c(tx)}{c(t)} \exp\{-(\rho-\varepsilon)\log x\} &\geq \left(\frac{c(tx)}{c(t)} - \varepsilon\right) x^{\rho} \exp\{-\varepsilon \log x\} \\ &\sim (1-\varepsilon)\, x^{\rho} \exp\{-\varepsilon \log x\}, \end{aligned}$$

vem que

$$(1-\varepsilon)\, x^{\rho} \exp\{-\varepsilon \log x\} \leq \frac{U(tx)}{U(t)} \leq (1+\varepsilon)\, x^{\rho} \exp\{\varepsilon \log x\}.$$

Para $0 < x < 1$, a resolução é idêntica, tendo-se apenas neste caso a diferença de que $|\log x| = -\log x$. ∎

2.2 Teoria Limite para Processos Pontuais associados a Modelos de Médias Móveis, dirigidos por Ruídos com Distribuição de Variação regular nas Caudas

Consideremos uma sequência i.i.d. de v.a.'s, $\{Z_k\}_{k\in\mathbb{Z}}$, com distribuição comum F, cujas caudas verificam

$$(1 - F(x) + F(-x)) = P[|Z_k| > x] = x^{-\alpha} L(x), x \to \infty, \tag{2.5}$$

e

$$\frac{P[Z_k > x]}{P[|Z_k| > x]} \underset{x\to\infty}{\longrightarrow} p \quad \text{e} \quad \frac{P[Z_k \le -x]}{P[|Z_k| > x]} \underset{x\to\infty}{\longrightarrow} q, \tag{2.6}$$

onde $L(x)$ é uma função de variação lenta em ∞, $\alpha > 0$ e $0 \le p \le 1$, com $p + q = 1$. Seja $\{X_k, k \ge 1\}$ o processo linear de médias móveis, estritamente estacionário, gerado pela sequência $\{Z_k\}$ e definido por

$$X_k := \sum_{j=0}^{\infty} c_j Z_{k-j}, \quad k \ge 1, \tag{2.7}$$

onde $\{c_j\}$ é uma sucessão de constantes que satisfaz

$$\sum_{j=0}^{\infty} |c_j|^{\delta} < \infty, \quad \text{para algum } \delta, \text{ tal que } 0 < \delta < \min(\alpha, 1). \tag{2.8}$$

Seja ainda $\{a_n\}_{n\in\mathbb{N}}$, uma sucessão de constantes positivas, definida por

$$a_n = \inf\left\{x : P[|Z_1| > x] \le \frac{1}{n}\right\} \tag{2.9}$$

e que verifica

$$nP[|Z_1| > a_n x] \underset{n\to\infty}{\longrightarrow} Cx^{-\alpha}, \quad \text{para } x > 0, C > 0. \tag{2.10}$$

Recorde-se que, tal como foi visto no exemplo 2.1, $a_n \sim n^{1/\alpha} L(n)$.

Uma consequência imediata de (2.8), é a convergência absoluta com probabilidade um, do processo linear $\{X_k\}$, pois uma vez que que $0 < \delta < \min(\alpha, 1)$, tem-se pela desigualdade triângular, $|x+y|^{\delta} \le |x|^{\delta} + |y|^{\delta}$, que

$$E\left(|X_k|^{\delta}\right) \le \sum_{j=0}^{\infty} |c_j|^{\delta} E\left(|Z_{k-j}|^{\delta}\right) = E\left(|Z_1|^{\delta}\right) \sum_{j=0}^{\infty} |c_j|^{\delta} < \infty,$$

o que mostra que $X_k := \sum_{j=0}^{\infty} c_j Z_{k-j}$ é finito quase certamente. Também se prova que a sequência dos erros $\{Z_k\}$ verifica

$$\begin{cases} E\left(|Z_1|^{\delta}\right) = \infty & se\ \delta \ge \alpha \\ E\left(|Z_1|^{\delta}\right) < \infty & se\ \delta < \alpha, \end{cases}$$

e por conseguinte $var(Z_k) = \infty$ se $0 < \alpha < 2$ e $E(|Z_l|) < \infty$ se $\alpha > 1$.

Nas condições anteriormente descritas, Cline (1983) provou[1] o seguinte resultado

$$\lim_{x\to\infty}\frac{P[|X_k|>x]}{P[|Z_1|>x]}=\lim_{x\to\infty}\frac{P\left[\sum_{j=0}^{\infty}|c_j||Z_{k-j}|>x\right]}{P[|Z_1|>x]}=\sum_{j=0}^{\infty}|c_j|^{\alpha}<\infty, \qquad (2.11)$$

o que revela que as caudas da distribuição de X_k, comportam-se do mesmo modo como as de Z_k. As relações assintóticas de variação regular implícitas em (2.5) e (2.6) e que envolvem as caudas, direita e esquerda, da distribuição de Z_k, levam naturalmente a considerarem-se espaços métricos do tipo $[-\infty,\infty]^s\setminus\{0\}$, para algum $s\geq 1$. Ora, nestes espaços se tomarmos colecções de variáveis aleatórias, definidas à custa da sequência $\{Z_k\}$, poderemos construir processos pontuais e estudar as suas propriedades limite, no que diz respeito à convergência fraca dos mesmos. Relembrando que os espaços métricos, onde esses processos pontuais estão definidos, os espaços das medidas pontuais aleatórias, gozam das propriedades de completude e separabilidade, ao implementarmos resultados em termos de convergência fraca, estamos a estabelecer de forma equivalente, propriedades inerentes à convergência em distribuição desses mesmos processos. Para começar, vamos considerar para cada n,a sequência de elementos $\left\{\left(\frac{k}{n},a_n^{-1}Z_k\right),k\geq 1\right\}$, definida no espaço $E=[0,\infty)\times\overline{\mathbb{R}}\setminus\{0\}$, e com ela tomar o processo pontual bidimensional $\sum_{k=1}^{\infty}\varepsilon_{\left(k/n,a_n^{-1}Z_k\right)}$. A convergência fraca deste processo, foi analisada por vários autores, tais como, Mori e Odaira (1976), Weissman(1975) e Resnick(1984), os quais provaram que no espaço $M_p\left([0,\infty)\times\overline{\mathbb{R}}\setminus\{0\}\right)$ tem-se

$$\sum_{k=1}^{\infty}\varepsilon_{\left(\frac{k}{n},\frac{Z_k}{a_n}\right)}\Rightarrow\sum_{k=1}^{\infty}\varepsilon_{(t_k,j_k)}, \qquad (2.12)$$

onde o processo $\sum_{k=1}^{\infty}\varepsilon_{(t_k,j_k)}$, é um $PRM(\mu)$ de medida μ, definida no conjunto $[0,\infty)\times\overline{\mathbb{R}}\setminus\{0\}$ por

$$\begin{aligned}\mu(dt,dx) &= \lambda(dt)\times v(dx)\\ &= dt\times v(dx)\end{aligned} \qquad (2.13)$$

e com $v(dx)=\alpha\left(px^{-\alpha-1}1_{(0,\infty)}(x)+(1-p)(-x)^{-\alpha-1}1_{(-\infty,0)}(x)\right)dx$. Tendo em conta os resultados desenvolvidos no exemplo 1.5, o processo limite em (2.12) também pode ser escrito como

$$\sum_{k=1}^{\infty}\varepsilon_{\left(t_k,\delta_k\Gamma_k^{-1/\alpha}\right)},$$

onde $\{\delta_k,k\geq 1\}$ é uma sequência de v.a.'s i.i.d que verifica

$$P[\delta_k=1]=p \text{ e } P[\delta_k=-1]=1-p, \qquad (2.14)$$

[1] Veja-se em [6], p. 83-24.

e $\{\Gamma_k, k \geq 1\}$ a sequência de v.a.'s que satisfaz

$$\Gamma_k = E_1 + ... + E_k, \tag{2.15}$$

com $\{E_k\}_{k\geq 1}$ sucessão i.i.d. de exponenciais unitárias. Observe-se que, pela proposição 1.12, a convergência em (2.12) tem lugar em $M_p\left([0,\infty) \times \overline{\mathbb{R}}\backslash\{0\}\right)$ sse $nP\left[a_n^{-1}Z_1 \in \cdot\right] \xrightarrow{v} \nu(\cdot)$ em $\overline{\mathbb{R}}\backslash\{0\}$. É precisamente este equilíbrio entre convergência fraca e convergência vaga, traduzido pelo resultado expresso em (2.12), que servirá de base para a determinação de limites de outros processos pontuais, e conduzirá simultaneamente a um resultado semelhante, envolvendo os elementos X's. Com esse propósito passamos agora a considerar o espaço topológico de Hausdorff, localmente compacto e com bases contáveis, $E = [0,\infty) \times \overline{\mathbb{R}}^m\backslash\{0\}$, em que m é um inteiro fixo, positivo e maior do que um. Consideremos ainda $\mathcal{E}$, a σ-álgebra dos borelcanos de E, tal que os conjuntos compactos de $\overline{\mathbb{R}}^m\backslash\{0\}$, são os compactos de $\overline{\mathbb{R}}^m$ "limitados a partir da origem" e que não contenham a origem (i.e. são conjuntos relativamente compactos e que não contêm a origem). Uma classe conveniente de conjuntos que geram $\mathcal{E}$, é a classe $\mathcal{S}$ constituída por todos os conjuntos $(m+1)$-dimensionais da forma

$$B = [b_0, c_0) \times (b_1, c_1] \times ... \times (b_m, c_m], \tag{2.16}$$

onde $[b_0, c_0)$ é um semi-aberto de $\mathbb{R}_0^+$ com $b_0 \geq 0$, o rectângulo

$$(b_1, c_1] \times ... \times (b_m, c_m]$$

é limitado a partir da origem e não contém a origem $0 = (0, ..., 0)$ e $b_i < c_i$, com $b_i \neq 0$ e $c_i \in \overline{\mathbb{R}}\backslash\{0\}$ para $i = 1, ..., m$. Estes conjuntos B, atribuem a $\mathcal{S}$ a propriedade de semianel-DC[2], condição necessária , como iremos ver, para que possamos construir uma medida pontual conveniente ao estudo de convergências envolvendo o processo de médias moveis $\{X_k\}$. Para além disso, estes conjuntos B também verificam as condições

$$(b_1, c_1] \times ... \times (b_m, c_m] \cap \{y\mathbf{e}_i : y \in \overline{\mathbb{R}}\} = \emptyset, \quad \text{para } i = 1, ..., m \tag{2.17}$$

ou

$$\exists i' \text{ tal que } (b_1, c_1] \times ... \times (b_m, c_m] \cap \{y\mathbf{e}_i : y \in \overline{\mathbb{R}}\} = \begin{cases} (b_{i'}, c_{i'}] & se\ i = i' \\ \emptyset & se\ i \neq i', \end{cases} \tag{2.18}$$

em que $\mathbf{e}_i = (0, ..., 0, 1, 0, ..., 0)$ é o vector da base canónica de $\mathbb{R}^m$, tendo a i-ésima componente igual a 1 e as restantes iguais a zero. Repare-se que por (2.17) e (2.18), o rectângulo m-dimensional $(b_1, c_1] \times ... \times (b_m, c_m]$ ou não intersecta

[2] Veja-se em apêndice.

nenhum eixo, ou intersecta exatamente o eixo i', não intersectando qualquer um dos outros. Repare-se ainda que na condição (2.18), tem-se $b_i < 0 < c_i$ para $i \neq i'$ e $0 \notin (b_{i'}, c_{i'}]$. Convencionaremos que $y\mathbf{e}_i = (0, ..., 0, \pm\infty, 0, ..., 0)$, sempre que $y = \pm\infty$. Com base nos vectores $\mathbf{Z}^{(k)} = (Z_k, Z_{k-1}, ..., Z_{k-m+1})$ e $\mathbf{e}_i$, constituam-se as colecções de pontos $\left\{\frac{k}{n}, a_n^{-1}\mathbf{Z}^{(k)}, k \geq 1\right\}$ e $\{t_k, j_k\mathbf{e}_i\}$, definidas no espaço $E = [0, \infty) \times \overline{\mathbb{R}}^m \backslash \{0\}$. A partir destas colecções de pontos podemos definir em $M_p\left([0, \infty) \times \overline{\mathbb{R}}^m \backslash \{0\}\right)$ os seguintes processos pontuais

$$I_n := \sum_{k=1}^{\infty} \varepsilon_{\left(\frac{k}{n}, a_n^{-1}\mathbf{Z}^{(k)}\right)} \quad \text{e} \quad I := \sum_{k=1}^{\infty}\sum_{i=1}^{m} \varepsilon_{(t_k, j_k\mathbf{e}_i)}. \tag{2.19}$$

Para melhor compreendermos a estrutura do processo I, imaginemos que a componente t_k é desprezada temporariamente. Tomemos os j_k's e deixemo-los cair no eixo $\mathbf{e}_1$. Pegando novamente nos mesmos pontos j_k's, façamo-los cair agora pelo eixo $\mathbf{e}_2$. Este procedimento é repetido deterministicamente, para os restantes eixos. Com vista a provar que o processo I_n converge em distribuição para o processo I, apresentamos algumas propriedades destes processos, reunidas na seguinte proposição:

Proposição 2.3 *Dados os processos pontuais definidos em (2.19), são válidas as propriedades:*

a) Para todo o $B \in S$,

$$P[I(\delta B) = 0] = 1.$$

b) Para todo o B que satisfaça (2.17),

$$P[I(B) = 0] = 1 \text{ e } E[I_n(B)] \underset{n}{\rightarrow} 0.$$

c) Para qualquer B que satisfaça (2.18),

$$P[I(B) = 0] = \exp\{-\mu([b_0, c_0) \times (b_{i'}, c_{i'}])\},$$

e além disso

$$E[I_n(B)] \underset{n}{\rightarrow} \mu([b_0, c_0) \times (b_{i'}, c_{i'}]),$$

onde $\mu(dt, dx) = dt \times \nu(dx)$ é a intensidade do PRM definido em (2.12).

Dem. a) Tome-se um boreleano qualquer B, que verifique (2.17). Ora, é evidente que não há pontos da forma $(t_k, j_k\mathbf{e}_i)$ que caiam na fronteira de B, já que os pontos j_k's estão localizados sobre os eixos e B não intersecta os eixos. Por conseguinte $I(\delta B) = 0$, tendo-se $P[I(\delta B) = 0] = 1$. Tomemos agora $B \in S$, tal que (2.18) é verificada. Estamos interessados em identificar o n^{o}de pontos $(t_k, j_k \mathrm{e}_i)$ que estão na fronteira de B. Neste caso como há intersecção com o eixo i' temos

$$
\begin{aligned}
I(\delta B) &= \sum_{k=1}^{\infty} \sum_{i=1}^{m} \varepsilon_{(t_k, j_k \mathrm{e}_i)} (\delta B) \\
&= \sum_{k=1}^{\infty} \left(\varepsilon_{(t_k, j_k \mathrm{e}_1)} (\delta B) + .. + \varepsilon_{(t_k, j_k \mathrm{e}_{i'})} (\delta B) + .. + \varepsilon_{(t_k, j_k \mathrm{e}_m)} (\delta B) \right) \\
&= \sum_{k=1}^{\infty} \varepsilon_{(t_k, j_k)} (\{b_0, c_0\} \times \{b_{i'}, c_{i'}\}) \\
&\leq \sum_{k=1}^{\infty} \varepsilon_{j_k} (\{b_{i'}, c_{i'}\}) = 0, \qquad q.c.
\end{aligned}
$$

porque $\sum_{k=1}^{\infty} \varepsilon_{j_k}$ não tem átomos, já que é um $PRM(\nu)$, com ν definida em (2.12).

b) Seja $I(B) = \sum_{k=1}^{\infty} \sum_{i=1}^{m} \varepsilon_{(t_k, j_k \mathrm{e}_i)} (B)$ tal que B verifica (2.17). É evidente que $I(B) = 0$, pois B não intersecta os eixos e por tal, os pontos $(t_k, j_k \mathrm{e}_i)$ nunca estarão em B, donde $I(B) = 0$ com probabilidade um. Calculemos agora o valor esperado de $I_n(B)$.

$$
\begin{aligned}
E[I_n(B)] &= E\left[\sum_{k=1}^{\infty} \varepsilon_{\left(\frac{k}{n}, a_n^{-1} \mathrm{Z}^{(k)}\right)} (B) \right] \\
&= \sum_{k} \left(1.P\left[\left(\frac{k}{n}, \frac{\mathrm{Z}^{(k)}}{a_n} \right) \in B \right] + 0.P\left[\left(\frac{k}{n}, \frac{\mathrm{Z}^{(k)}}{a_n} \right) \notin B \right] \right) \\
&= \sum_{k : \frac{k}{n} \in [b_0, c_0)} P\left(\frac{\mathrm{Z}^{(k)}}{a_n} \in (b_1, c_1] \times \ldots \times (b_m, c_m] \right) \\
&= \sum_{k : \frac{k}{n} \in [b_0, c_0)} P\left(\frac{\mathrm{Z}_k}{a_n} \in (b_1, c_1], \ldots, \frac{\mathrm{Z}_{k-m+1}}{a_n} \in (b_m, c_m] \right),
\end{aligned}
$$

e porque os Z_k's são independentes., o valor esperado vem

$$
\begin{aligned}
E[I_n(B)] &= \sum_{k : \frac{k}{n} \in [b_0, c_0)} \prod_{i=1}^{m} P\left(\frac{Z_{k-i+1}}{a_n} \in (b_i, c_i] \right) \\
&= \sum_{k : \frac{k}{n} \in [b_0, c_0)} \prod_{i=1}^{m} P\left(\frac{Z_1}{a_n} \in (b_i, c_i] \right), \qquad (2.20)
\end{aligned}
$$

em que a última igualdade surge pelo facto dos Z_k's serem identicamente distribuídos. Temos agora duas situações a analisar, uma em que $0 < b_i < c_i$ e outra em que $b_i < c_i < 0$. Se $0 < b_i < c_i$, (2.20) vem

$$\begin{aligned} E\left[I_n(B)\right] &= \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(\frac{Z_1}{a_n} \leq c_i \wedge \frac{Z_1}{a_n} > b_i\right) \\ &= \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} \left[P\left(\frac{Z_1}{a_n} > b_i\right) - P\left(\frac{Z_1}{a_n} > c_i\right)\right] \end{aligned}$$

donde,

$$\begin{aligned} E\left[I_n(B)\right] &\leq \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(\frac{Z_1}{a_n} > b_i\right) = \sum_{k:\frac{k}{n}\in(b_0,c_0]} \prod_{i=1}^{m} P\left(\frac{Z_1}{a_n} > |b_i|\right) \\ &= \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(\frac{Z_1}{a_n} > \min\left(|b_i|,|c_i|\right)\right) \\ &\leq \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(\left|\frac{Z_1}{a_n}\right| > \min\left(|b_i|,|c_i|\right)\right) \qquad (2.21) \end{aligned}$$

Se $b_i < c_i < 0$, (2.20) vem

$$\begin{aligned} E\left[I_n(B)\right] &= \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(\frac{Z_1}{a_n} \leq c_i \wedge \frac{Z_1}{a_n} > b_i\right) \\ &= \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(-\frac{Z_1}{a_n} > -c_i \wedge -\frac{Z_1}{a_n} \leq -b_i\right) \\ &= \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} \left[P\left(-\frac{Z_1}{a_n} > -c_i\right) - P\left(-\frac{Z_1}{a_n} > -b_i\right)\right] \\ &\leq \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(-\frac{Z_1}{a_n} > -c_i\right) = \sum_{k:\frac{k}{n}\in(b_0,c_0]} \prod_{i=1}^{m} P\left(-\frac{Z_1}{a_n} > |c_i|\right) \\ &= \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(-\frac{Z_1}{a_n} > \min\left(|b_i|,|c_i|\right)\right) \\ &\leq \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(\left|\frac{Z_1}{a_n}\right| > \min\left(|b_i|,|c_i|\right)\right) \qquad (2.22) \end{aligned}$$

De (2.21) e (2.22), e pela definição de a_n, obtemos finalmente

$$\begin{aligned} & \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(\left|\frac{Z_1}{a_n}\right| > \underbrace{\min\left(|b_i|,|c_i|\right)}_{d_i}\right) \\ \sim & \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} \frac{d_i^{-\alpha}}{n} \longrightarrow 0, \end{aligned}$$

quando $n \longrightarrow \infty$ e para qualquer $m \geq 2$. Ora, como $0 \leq E\left[I_n(B)\right] \leq 0$, vem que $E\left[I_n(B)\right] \longrightarrow 0$.

c) Tomemos agora $B \in \mathcal{S}$, tal que (2.18) é verificada. Determinemos a probabilidade da v.a. $I(B)$ ser nula.Tem-se por (2.12) e 2.20

$$\begin{aligned} P\left[I(B)=0\right] &= P\left[\sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{(t_k, j_k e_i)}(B)=0\right] \\ &= P\left[\sum_{k=1}^{\infty}\varepsilon_{(t_k, j_k)}\left([b_0,c_0)\times(b_{i'},c_{i'}]\right)=0\right] \\ &= \exp\left\{-\mu\left([b_0,c_0)\times(b_{i'},c_{i'}]\right)\right\}. \end{aligned}$$

Repare-se que μ, avaliada no boreleano bidimensional $[b_0,c_0)\times(b_{i'},c_{i'}]$, é dada por

$$\begin{aligned} \mu\left([b_0,c_0)\times(b_{i'},c_{i'}]\right) &= \int_{b_0}^{c_0}\int_{b_{i'}}^{c_{i'}} dt\nu(dx) \\ &= \int_{b_0}^{c_0}\int_{b_{i'}}^{c_{i'}} \left(\alpha p x^{-\alpha-1} I_{(0,\infty)}(x)\right. \\ &\quad \left. + (1-p)(-x)^{-\alpha-1} I_{(-\infty,0)}(x)\right) dxdt \\ &= \begin{cases} \int_{b_0}^{c_0}\int_{b_{i'}}^{c_{i'}} \alpha p x^{-\alpha-1} dxdt, & \text{se } 0 < b_{i'} < c_{i'} \\ \int_{b_0}^{c_0}\int_{b_{i'}}^{c_{i'}} \alpha q (-x)^{-\alpha-1} dxdt, & \text{se } b_{i'} < c_{i'} < 0 \end{cases} \\ &= \begin{cases} -p(c_0-b_0)\left(c_{i'}^{-\alpha}-b_{i'}^{-\alpha}\right), & \text{se } 0 < b_{i'} < c_{i'} \\ q(c_0-b_0)\left((-c_{i'})^{-\alpha}-(-b_{i'})^{-\alpha}\right), & \text{se } b_{i'} < c_{i'} < 0. \end{cases} \end{aligned}$$

Determinemos agora, o valor esperado de $I_n(B)$. Temos

$$\begin{aligned} E[I_n(B)] &= \sum_{k:\frac{k}{n}\in[b_0,c_0)} \prod_{i=1}^{m} P\left(\frac{Z_1}{a_n}\in(b_i,c_i]\right) \\ &= \sum_{k:\frac{k}{n}\in[b_0,c_0)} \underbrace{\prod_{i\neq i'} P\left(\frac{Z_1}{a_n}\in(b_i,c_i]\right)}_{(1)} P\left(\frac{Z_1}{a_n}\in(b_{i'},c_{i'}]\right). \end{aligned}$$

Averiguemos a probabilidade em (1). Neste caso como $b_i < 0 < c_i$, então

$$P\left(\frac{Z_1}{a_n}\in(b_i,c_i]\right) = 1 - P(Z_1 \leq a_n b_i \vee Z_1 > a_n c_i) \longrightarrow 1,$$

pois tomando $d_i = \min(|b_i|, c_i)$, tem-se que

$$\begin{aligned} P(Z_1 \leq a_n b_i \vee Z_1 > a_n c_i) &\leq P(Z_1 \leq -a_n d_i \vee Z_1 > a_n d_i) \\ &= P(|Z_1| > a_n d_i) \underset{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

por definição de a_n. Logo, $\prod_{i\neq i'} P\left(\frac{Z_1}{a_n}\in(b_i,c_i]\right) \to 1$. Tendo em conta que

$$\sum_{k:\frac{k}{n}\in[b_0,c_0)} 1 \sim n(c_0 - b_0),$$

o valor esperado vem então

$$E[I_n(B)] \sim (c_0 - b_0)\, nP\left(\frac{Z_1}{a_n}\in(b_{i'},c_{i'}]\right), \tag{2.23}$$

e novamente, temos que desdobrar o estudo em dois casos. Se $0 < b_{i'} < c_{i'}$, obtemos

$$\begin{aligned} nP\left(\frac{Z_1}{a_n}\in(b_{i'},c_{i'}]\right) &= nP\left(\frac{Z_1}{a_n} > b_{i'} \wedge \frac{Z_1}{a_n} \leq c_{i'}\right) \\ &= n\left[P\left(\frac{Z_1}{a_n} > b_{i'}\right) - P\left(\frac{Z_1}{a_n} > c_{i'}\right)\right] \\ &= nP(|Z_1| > a_n b_{i'}) \frac{P(Z_1 > a_n b_{i'})}{P(|Z_1| > a_n b_{i'})} \\ &\quad - nP(|Z_1| > a_n c_{i'}) \frac{P(Z_1 > a_n c_{i'})}{P(|Z_1| > a_n c_{i'})} \\ &\longrightarrow b_{i'}^{-\alpha} p - c_{i'}^{-\alpha} p = -p\left(c_{i'}^{-\alpha} - b_{i'}^{-\alpha}\right) = \nu((b_{i'},c_{i'}]), \end{aligned}$$

por (2.6), (2.10) e tomando $n \to \infty$. Se $b_{i'} < c_{i'} < 0$, obtemos

$$\begin{aligned} nP\left(\frac{Z_1}{a_n} \in (b_{i'}, c_{i'}]\right) &= n\left[P\left(\frac{Z_1}{a_n} \leq c_{i'}\right) - P\left(\frac{Z_1}{a_n} \leq b_{i'}\right)\right] \\ &= nP\left(|Z_1| > -a_n c_{i'}\right) \frac{P\left(Z_1 \leq a_n c_{i'}\right)}{P\left(|Z_1| > -a_n c_{i'}\right)} \\ &\quad -nP\left(|Z_1| > -a_n b_{i'}\right) \frac{P\left(Z_1 \leq a_n b_{i'}\right)}{P\left(|Z_1| > -a_n b_{i'}\right)} \\ &\to q\left((-c_{i'})^{-\alpha} - (-b_{i'})^{-\alpha}\right) = q\nu\left((b_{i'}, c_{i'}]\right) \end{aligned}$$

por (2.6), (2.10) e tomando $n \to \infty$. Retomando o valor esperado em (2.23), vem finalmente

$$\begin{aligned} E\left[I_n(B)\right] &\sim (c_0 - b_0)\, nP\left(\frac{Z_1}{a_n} \in (b_{i'}, c_{i'}]\right) \\ &\to \mu\left([b_0, c_0) \times (b_{i'}, c_{i'}]\right) \end{aligned}$$ ■

Em $M_p\left([0,\infty) \times \overline{\mathbb{R}}^m \setminus \{0\}\right)$ consideremos o novo processo pontual,

$$I_n^* := \sum_{k=1}^{\infty}\sum_{i=1}^{m} \varepsilon_{\left(\frac{k}{n}, a_n^{-1} Z_k \mathbf{e}_i\right)}.$$

Observe-se que para cada $B \in \mathcal{S}$, $I_n^*(B) \leq I_n(B)$. A próxima proposição assegura-nos que $I_n^*(B) - I_n(B)$ é um $o_p(1)$ e é a partir desta "cumplicidade" entre estes dois processos, que chegaremos à convergência fraca do processo I_n.

Proposição 2.4 *Sejam I_n e I_n^*, os processos pontuais definidos anteriormente. Então para cada $B \in \mathcal{S}$, verifica-se que*

$$I_n(B) - I_n^*(B) \xrightarrow{P} 0 \tag{2.24}$$

Dem. Esta demonstração tem por base o seguinte resultado da teoria das probabilidades: dadas as v.a.'s $X_1, X_2, \ldots$, não negativas e pertencentes a um mesmo espaço de probabilidade $(\Omega, \mathcal{A}, P)$, se $E[X_n] \to 0$ então X_n é um $o_p(1)$. Por simplicidade de cálculos admita-se que $B = [0,1) \times (b_1, c_1] \times \ldots \times (b_m, c_m]$, (para os restantes casos a demonstração é semelhante).

Suponhamos primeiro que B satisfaz (2.17). Pela proposição 2.3 vimos que $E\left[I_n(B)\right] \underset{n\to\infty}{\longrightarrow} 0$. Por outro lado pela definição de I_n^*, tem-se

$$\begin{aligned} I_n^*(B) &= \sum_{k=1}^{\infty}\sum_{i=1}^{m} \varepsilon_{\left(\frac{k}{n}, a_n^{-1} Z_k \mathbf{e}_i\right)}(B) \\ &= \sum_{k=1}^{\infty}\left(\varepsilon_{\left(\frac{k}{n}, a_n^{-1} Z_k \mathbf{e}_1\right)}(B) + \ldots + \varepsilon_{\left(\frac{k}{n}, a_n^{-1} Z_k \mathbf{e}_m\right)}(B)\right) \\ &= 0, \text{ pois } B \text{ não intersecta os eixos.} \end{aligned}$$

Como $0 \leq I_n(B) - I_n^*(B)$, então vem que $0 \leq E[I_n(B) - I_n^*(B)] \to 0$, quando $n \to \infty$, donde $I_n(B) - I_n^*(B) \xrightarrow{P} 0$.

Suponhamos agora que B satisfaz (2.18). Vimos que o valor esperado de $I_n(B)$ tem por limite em n, a medida $\mu((0,1) \times (b_{i'}, c_{i'}])$. Quanto ao valor esperado de $I_n^*(B)$, tem-se

$$\begin{aligned}
E[I_n^*(B)] &= E\left[\sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{\left(\frac{k}{n}, a_n^{-1}Z_k\mathbf{e}_i\right)}(B)\right] \\
&= \sum_{k=1}^{\infty} P\left[\left(\frac{k}{n}, a_n^{-1}Z_k\right) \in [0,1) \times (b_{i'}, c_{i'}]\right] \\
&= \sum_{k:\frac{k}{n}\in[0,1)}^{\infty} P\left(a_n^{-1}Z_1 \in (b_{i'}, c_{i'}]\right) \\
&\to \mu([0,1) \times (b_{i'}, c_{i'}]),
\end{aligned}$$

atendendo à demonstração de $E[I_n(B)]$, da alínea c) da proposição 2.3. Note-se agora que a v.a. $I_n(B)$ pode ser desdobrada na seguinte soma

$$I_n(B) = \underbrace{\sum_{k=1}^{i'-1}\varepsilon_{\left(\frac{k}{n}, a_n^{-1}\mathbf{Z}^{(k)}\right)}(B)}_{I_{1,n}(B)} + \underbrace{\sum_{k=i'}^{\infty}\varepsilon_{\left(\frac{k}{n}, a_n^{-1}\mathbf{Z}^{(k)}\right)}(B)}_{I_{2,n}(B)}. \tag{2.25}$$

Determinando o valor esperado de $I_{1,n}(B)$, obtém-se

$$\begin{aligned}
E[I_{1,n}(B)] &= E\left[\sum_{k=1}^{i'-1}\varepsilon_{\left(\frac{k}{n}, a_n^{-1}\mathbf{Z}^{(k)}\right)}(B)\right] = \sum_{k:\frac{k}{n}\in[0,1)}^{i'-1}\prod_{i=1}^{m} P\left(\frac{Z_1}{a_n} \in (b_i, c_i]\right) \\
&= \sum_{k:\frac{k}{n}\in[0,1)}^{i'-1} \underbrace{\prod_{i\neq i'}^{m} P\left(\frac{Z_1}{a_n} \in (b_i, c_i]\right)}_{\downarrow \, 1} P\left(\frac{Z_1}{a_n} \in (b_{i'}, c_{i'}]\right) \\
&\to \sum_{k:\frac{k}{n}\in[0,1)}^{i'-1} P\left(\frac{Z_1}{a_n} \in (b_{i'}, c_{i'}]\right) \\
&\to (i'-1)\, P\left(\frac{Z_1}{a_n} \in (b_{i'}, c_{i'}]\right) \to 0,
\end{aligned}$$

recorrendo novamente à demonstração da proposição 2.3 c). Tomando, agora os valores esperados em (2.25), concluímos que

$$0 \leq E[I_n(B)] = E[I_{1,n}(B)] + E[I_{2,n}(B)],$$

donde,

$$E\left[I_{2,n}(B)\right] \underset{n\to\infty}{\longrightarrow} \mu\left([0,1)\times(b_{i'},c_{i'}]\right). \tag{2.26}$$

Por outro lado, fazendo $k=j+i'-1$, a v.a. $I_{2,n}(B)$ vem escrita na forma

$$\begin{aligned} I_{2,n}(B) &= \sum_{j=1}^{\infty}\varepsilon_{\left(\frac{j+i'-1}{n},a_n^{-1}Z^{(j+i'-1)}\right)}(B) \\ &\stackrel{d}{=} \sum_{j=1}^{\infty}\varepsilon_{\left(\frac{j+i'-1}{n},a_n^{-1}Z_j\right)}\left([0,1)\times(b_{i'},c_{i'}]\right) \\ &\stackrel{d}{=} \sum_{j=i'}^{\infty}\varepsilon_{\left(\frac{j}{n},a_n^{-1}Z_j\right)}\left([0,1)\times(b_{i'},c_{i'}]\right) \\ &\le \sum_{j=1}^{\infty}\varepsilon_{\left(\frac{j}{n},a_n^{-1}Z_j\right)}\left([0,1)\times(b_{i'},c_{i'}]\right) = I_n^*(B), \end{aligned} \tag{2.27}$$

já que

$$\begin{aligned} I_n^*(B) &= \sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{\left(\frac{k}{n},a_n^{-1}Z_k\mathbf{e}_i\right)}(B) \\ &= \sum_{k=1}^{\infty}\left(\varepsilon_{\left(\frac{k}{n},a_n^{-1}Z_k\mathbf{e}_1\right)}(B)+\dots+\varepsilon_{\left(\frac{k}{n},a_n^{-1}Z_k\mathbf{e}_m\right)}(B)\right) \\ &= \sum_{k=1}^{\infty}\varepsilon_{\left(\frac{k}{n},a_n^{-1}Z_k\right)}\left([0,1)\times(b_{i'},c_{i'}]\right). \end{aligned}$$

Ora, como $0\le I_n(B)-I_n^*(B)$, então

$$\begin{aligned} E\left[I_n(B)-I_n^*(B)\right] &= E\left[I_{1,n}(B)+I_{2,n}(B)-I_n^*(B)\right] \\ &= E\left[I_{2,n}(B)-I_n^*(B)\right]+E\left[I_{1,n}(B)\right] \underset{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

Logo, $I_n(B)-I_n^*(B)\to 0$ em probabilidade. ∎

Provado que $I_n(B)-I_n^*(B)=o_p(1)$, o próximo passo é, como dissémos, estipular a convergência fraca do processo I_n para o processo I. Se mostrarmos que para todo o rectângulo $B\in\mathcal{S}$, $I_n^*(B)$ converge fracamente para $I(B)$, então tem-se naturalmente $I_n(B)\Rightarrow I(B)$. No suporte deste resultado, estão três elementos fundamentais: a convergência dos processos em (2.12), o teorema 1.1 da aplicação contínua e o teorema 4.2 apresentado em [19], pág. 32 e que pode ser consultado em apêndice. Neste último teorema, define-se a convergência em distribuição, quando a sucessão $\{X_n\}$ é uma sucessão de medidas aleatórias. Pelo que vimos no capítulo anterior, sabemos que $X_n\xrightarrow{d}X$, com X medida aleatória, sse

$E[f(X_n)] \to E[f(X)]$, para qualquer função f contínua e limitada. No entanto é possível caracterizar a convergência em distribuição em função dos vectores aleatórios k-dimensionais $(X_n(A_1), ..., X_n(A_k))$ e $(X(A_1), ..., X(A_k))$, em que $A_1, ..., A_k$ são conjuntos pertencentes à σ-álgebra dos boreleanos. É precisamente esta "nova" descrição de convergência em distribuição de medidas aleatórias, que vamos utilizar para demonstrar a convergência do processo I_n para o processo I, apresentada na próxima proposição.

Teorema 2.2 *Seja $\{Z_k\}$ a sequência de variáveis i.i.d. que verificam (2.5), (2.6) e (2.10). Então no espaço $M_p\left([0,\infty) \times \overline{\mathbb{R}}^m \setminus \{0\}\right)$, com $m > 1$,*

$$\sum_{k=1}^{\infty} \varepsilon_{\left(\frac{k}{n}, a_n^{-1} \mathbf{Z}^{(k)}\right)} \Rightarrow \sum_{k=1}^{\infty} \sum_{i=1}^{m} \varepsilon_{(t_k, j_k \mathbf{e}_i)}, \tag{2.28}$$

onde os pontos da sequência $\{(t_k, j_k)\}_{k \geq 1}$, formam o processo pontual de Poisson definido em (2.12).

Dem. Consideremos a aplicação contínua

$$\begin{array}{rccc} T: & [0,\infty) \times \overline{\mathbb{R}} \setminus \{0\} & \longrightarrow & [0,\infty) \times \overline{\mathbb{R}}^m \setminus \{\overline{0}\} \\ & (u_k, v_k) & \longrightarrow & (u_k, v_k \mathbf{e}_i) \end{array}$$

Se provarmos que para todo o compacto $K \subset [0,\infty) \times \overline{\mathbb{R}}^m \setminus \{0\}$, a imagem inversa de K, $T^{-1}(K)$, é um compacto de $\mathbb{R}_0^+ \times \overline{\mathbb{R}} \setminus \{0\}$, então pela proposição 1.9 sabemos que a aplicação induzida

$$\begin{array}{rccc} \widetilde{T}: & M_p\left([0,\infty) \times \overline{\mathbb{R}} \setminus \{0\}\right) & \longrightarrow & M_p\left([0,\infty) \times \{\overline{\mathbb{R}}^m \setminus \{\overline{0}\}\}\right) \\ & \sum_{k=1}^{\infty} \varepsilon_{(u_k, v_k)} & \longrightarrow & \sum_{k=1}^{\infty} \varepsilon_{T(u_k, v_k)} = \sum_{k=1}^{\infty} \varepsilon_{(u_k, v_k \mathbf{e}_i)} \end{array}$$

é uma aplicação contínua. Seja então K um compacto qualquer, subconjunto de $[0,\infty) \times \overline{\mathbb{R}}^m \setminus \{0\}$. Para que $T^{-1}(K)$ seja compacto em $[0,\infty) \times \overline{\mathbb{R}} \setminus \{0\}$, é necessário e suficiente que $T^{-1}(K)$ seja fechado e limitado a partir da origem, não contendo a origem. Ora uma vez que T é contínua, resulta que $T^{-1}(K)$ é fechado. Por outro lado sendo $T^{-1}(K) = \{x : \exists y \in K, T(x) = y\}$, qualquer elemento x em $T^{-1}(K)$, verifica $\|x\| = \|(u_k, v_k)\| = \|(u_k, v_k \mathbf{e}_i)\| = \|T(x)\| > M$, para algum $M > 0$, já que K é limitado a partir da origem, ou seja provámos que $T^{-1}(K)$ também é limitado. Concluímos assim que a imagem inversa de um compacto em $[0,\infty) \times \overline{\mathbb{R}}^m \setminus \{0\}$, também é um compacto em $[0,\infty) \times \overline{\mathbb{R}} \setminus \{0\}$ e por conseguinte a aplicação $\widetilde{T}$ é contínua. Com base em $\widetilde{T}$, podemos construir as aplicações, igualmente contínuas,

$$M_p\left([0,\infty) \times \overline{\mathbb{R}} \setminus \{0\}\right) \xrightarrow{\widetilde{T}_1} \left[M_p\left([0,\infty) \times \overline{\mathbb{R}} \setminus \{0\}\right)\right]^m \xrightarrow{\widetilde{T}_2} M_p\left([0,\infty) \times \overline{\mathbb{R}}^m \setminus \{0\}\right)$$

definidas por

$$\sum_{k=1}^{\infty}\varepsilon_{(u_k,v_k)} \mapsto \left(\sum_{k=1}^{\infty}\varepsilon_{(u_k,v_k\mathbf{e}_1)},\ldots,\sum_{k=1}^{\infty}\varepsilon_{(u_k,v_k\mathbf{e}_m)}\right) \mapsto \sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{(u_k,v_k\mathbf{e}_i)},$$

resultando que a aplicação composta $h = \widetilde{T}_2 \circ \widetilde{T}_1$ é também contínua. Ora, sabemos pelo teorema 1.1, da aplicação contínua de Billingsley, que se

$$\sum_{k=1}^{\infty}\varepsilon_{\left(\frac{k}{n},\frac{Z_k}{a_n}\right)} \Rightarrow \sum_{k=1}^{\infty}\varepsilon_{(t_k,j_k)} \text{ então } h\left(\sum_{k=1}^{\infty}\varepsilon_{\left(\frac{k}{n},\frac{Z_k}{a_n}\right)}\right) \Rightarrow h\left(\sum_{k=1}^{\infty}\varepsilon_{(t_k,j_k)}\right),$$

isto é,

$$I_n^* = \sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{\left(\frac{k}{n},\frac{Z_k}{a_n}\mathbf{e}_i\right)} \Rightarrow \sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{(t_k,j_k\mathbf{e}_i)} = I.$$

Mas o nosso interesse está em estabelecer esta convergência para o processo I_n e isso pode ser conseguido, como já foi referido, recorrendo à proposição anterior e ao teorema 4.2 em [19]. Tomando conjuntos $B_1,...,B_j \in \mathcal{S}$, com $j \geq 1$, o referido teorema garante que $I_n^* \Rightarrow I$ ssc

$$\left(I_n^*(B_1),...,I_n^*(B_j)\right) \Rightarrow \left(I(B_1),...,I(B_j)\right),$$

tendo-se consequentemente $I_n^*(B_1) \Rightarrow I(B_1),...,I_n^*(B_j) \Rightarrow I(B_j)$. Uma vez que $I_n(B) - I_n^*(B) \xrightarrow{P} 0, \forall B$, então $I_n(B_1) \Rightarrow I(B_1),...,I_n(B_j) \Rightarrow I(B_j)$, ou seja

$$\left(I_n(B_1),...,I_n(B_j)\right) \Rightarrow \left(I(B_1),...,I(B_j)\right)$$

provando-se assim que $I_n \Rightarrow I$. ■

Com vista a definir a convergência fraca do processo pontual associado ao processo linear $\{X_k\}$, apresentamos os seguintes lemas:

Lema 2.1 *Se a sequência de constantes $\{c_j\}$ satisfaz (2.8), então para qualquer $\gamma > 0$*

$$\lim_{m\to\infty}\limsup_{n\to\infty} P\left[a_n^{-1}\max_{1\leq k\leq n}\left|\sum_{j=0}^{m}c_jZ_{k-j} - X_k\right| > \gamma\right] = 0 \tag{2.29}$$

Dem. Tendo em conta que $X_k = \sum_{j=0}^{\infty}c_jZ_{k-j}$, a probabilidade em (2.29) vem dada por

$$P\left[a_n^{-1}\max_{1\leq k\leq n}\left|\sum_{j=0}^{m}c_jZ_{k-j} - X_k\right| > \gamma\right] = P\left[a_n^{-1}\max_{1\leq k\leq n}\left|\sum_{j>m}c_jZ_{k-j}\right| > \gamma\right].$$

Como a sequência $\left\{\sum_{j>m} c_j Z_{k-j},\ k = 1, ..., n\right\}$ é estacionária, a probabilidade anterior é limitada por

$$\begin{aligned} & nP\left[a_n^{-1}\left|\sum_{j>m} c_j Z_{k-j}\right| > \gamma\right] \\ = & \frac{P\left[a_n^{-1}\left|\sum_{j>m} c_j Z_{k-j}\right| > \gamma\right]}{P\left[|Z_1| > a_n\gamma\right]} nP\left[|Z_1| > a_n\gamma\right] \\ \to & \gamma^{-\alpha}\sum_{j>m} |c_j|^{\alpha}, \end{aligned}$$

por (2.11), (2.10) e tomando $n \to \infty$. Fazendo agora $m \to \infty$, a expressão anterior tende para zero, pois $R_m = \sum_{j>m} |c_j|^{\alpha}$ é o resto de uma série convergente. ∎

Lema 2.2 *(Teorema 4.2, Billingsley (1968))*
Sejam X_{un}, X_u, Y_n e X, elementos aleatórios de um espaço métrico, completo e separável $\mathcal{X}$, com métrica ρ, tal que para cada n, $X_{un}, Y_n, u \geq 1$, estão definidos no mesmo domínio. Suponha-se que para cada u, quando $n \to \infty$

$$X_{un} \Rightarrow X_u$$

e quando $u \to \infty$

$$X_u \Rightarrow X.$$

Suponha-se além disso que $\forall \varepsilon > 0$

$$\lim_{u\to\infty} \limsup_{n\to\infty} P\left[\rho\left(X_{un}, Y_n\right) > \varepsilon\right] = 0.$$

Então, quando $n \to \infty$

$$Y_n \Rightarrow X.$$

Dem. A demonstração segue de perto a apresentada em [24], pág.231. Pode consultar outra demonstração alternativa em [2], pág.25.

Pretende-se mostrar que $Y_n \Rightarrow X$, ou seja, $\lim_{n\to\infty} E\left[f\left(Y_n\right)\right] = E\left[f\left(X\right)\right]$, para toda a função f contínua e limitada em $\mathcal{X}$. Mas segundo o teorema 2.1, relativo a propriedades da convergência fraca, apresentado em [2], pág.11, é suficiente considerar f limitada e uniformemente contínua. Tendo em conta que

$$\begin{aligned} \left|E\left[f\left(Y_n\right)\right] - E\left[f\left(X\right)\right]\right| \leq & \ E\left|f\left(Y_n\right) - f\left(X_{un}\right)\right| \\ & + \left|E\left[f\left(X_{un}\right)\right] - E\left[f\left(X_u\right)\right]\right| + \left|E\left[f\left(X_u\right)\right] - E\left[f\left(X\right)\right]\right|, \end{aligned}$$

vem que

$$
\begin{aligned}
& \limsup_{n\to\infty} \left|E\left[f\left(Y_n\right)\right]-E\left[f\left(X\right)\right]\right| \\
\leq\ & \lim_{u\to\infty} \limsup_{n\to\infty} E\left|f\left(Y_n\right)-f\left(X_{un}\right)\right| \\
\leq\ & \lim_{u\to\infty} \limsup_{n\to\infty} E\left(\left|f\left(Y_n\right)-f\left(X_{un}\right)\right| ; \rho\left(Y_n, X_{un}\right) \leq \varepsilon\right) \\
& + \lim_{u\to\infty} \limsup_{n\to\infty} 2 \sup_{x\in\mathcal{X}} \left|f\left(x\right)\right| P\left[\rho\left(Y_n, X_{un}\right) > \varepsilon\right] \\
\leq\ & \sup\left\{\left|f\left(x\right)-f\left(y\right)\right| : \rho\left(x,y\right) \leq \varepsilon\right\} + 0 \rightarrow 0,
\end{aligned}
$$

quando $\varepsilon \to 0$, uma vez que f é uniformemente contínua. ■

Estamos agora aptos a provar a convergência do processo pontual definido com base na sequência $\{X_k\}_{k\in\mathbb{N}}$.

Teorema 2.3 *Suponhamos que* $\{a_n\}$ *satisfaz (2.10),* $\{c_j\}$ *satisfaz (2.8) e* $\{Z_k\}$ *satisfaz (2.5) e (2.6). Sejam* $\{X_k\}$*, o processo linear definido em (2.7) e* $\{(t_k, j_k)\}$ *a sequência de pontos do* $PRM(\mu)$ *definido em (2.12). Então no espaço das medidas pontuais,* $M_p\left([0,\infty)\times\overline{\mathbb{R}}\setminus\{0\}\right)$*, quando* $n\to\infty$*, verifica-se*

$$\zeta_n := \sum_{k=1}^{\infty} \varepsilon_{\left(\frac{k}{n}, \frac{X_k}{a_n}\right)} \Rightarrow \sum_{k=1}^{\infty}\sum_{i=0}^{\infty} \varepsilon_{\left(t_k, j_k c_i\right)} =: \zeta \qquad (2.30)$$

Dem. Comece-se por considerar a aplicação

$$
\begin{array}{llll}
T: & [0,\infty)\times\overline{\mathbb{R}}^m\setminus\{0\} & \longrightarrow & [0,\infty)\times\overline{\mathbb{R}}\setminus\{0\} \\
& \left(u_k, z_k, z_{k-1}, \ldots, z_{k-m+1}\right) & \longrightarrow & \left(u_k, \sum_{i=0}^{m-1} c_i z_{k-i}\right).
\end{array}
$$

Facilmente se vê que T não está bem definida (tome-se por exemplo $z_1=\infty$ e $z_2=-\infty$). A alternativa é construir a partir de T, uma outra aplicação que consiga garantir que a soma dos z's não seja nula e que simultaneamente contorne situações do tipo $\infty-\infty$. Defina-se então uma nova aplicação

$$T^*: [0,\infty)\times\overline{\mathbb{R}}^m\setminus\{0\} \longrightarrow [0,\infty)\times\overline{\mathbb{R}}\setminus\{0\}$$

definida por

$$
\begin{aligned}
\left(u_k, z_k, z_{k-1}, \ldots, z_{k-m+1}\right) & \longmapsto T^*\left(u, \mathbf{z}\right) \\
& = \begin{cases}
\left(u_k, \sum_{i=0}^{m-1} c_i z_{k-i}\right) & \text{se } \max\limits_{0\leq i\leq m-1} \left|z_{k-i}\right| < \infty \text{ e} \\
& \text{se } \sum_{i=0}^{m-1} c_i z_{k-i} \neq 0 \\
\left(u_k, \infty\right) & \text{caso contrário,}
\end{cases}
\end{aligned}
$$

e considere-se a aplicação $\widetilde{T^*}$ induzida por T^* em $M_p\left([0,\infty)\times\overline{\mathbb{R}}^m\setminus\{0\}\right)$, dada por

$$\widetilde{T^*}:\ M_p\left([0,\infty)\times\overline{\mathbb{R}}^m\setminus\{\overline{0}\}\right)\longrightarrow M_p\left([0,\infty)\times\overline{\mathbb{R}}\setminus\{0\}\right)$$
$$\sum_{k=1}^{\infty}\varepsilon_{(u_k,z_k,z_{k-1},\dots,z_{k-m+1})}\longrightarrow\begin{cases}\sum\limits_{k=1}^{\infty}\varepsilon_{\left(u_k,\sum_{i=0}^{m-1}c_iz_{k-i}\right)}\\ \sum_{k=1}^{\infty}\varepsilon_{(u_k,\infty)}\end{cases},$$

em que o 1ºramo é verificado se $\max\limits_{0\le i\le m-1}|z_{k-i}|<\infty$ e se $\sum_{i=0}^{m-1}c_iz_{k-i}\neq 0$. Se mostrarmos que $\widetilde{T^*}$ é quase certamente contínua com respeito à distribuição de $\sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{(t_k,j_k\mathbf{e}_i)}$, então o teorema 1.1, da aplicação contínua e (2.28), certificam a convergência

$$\widetilde{T^*}\left(\sum_{k=1}^{\infty}\varepsilon_{\left(\frac{k}{n},a_n^{-1}\mathbf{Z}^{(k)}\right)}\right)\Rightarrow\widetilde{T^*}\left(\sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{(t_k,j_k\mathbf{e}_i)}\right),$$

isto é,

$$\sum_{k=1}^{\infty}\varepsilon_{\left(\frac{k}{n},a_n^{-1}\sum_{i=0}^{m-1}c_iZ_{k-i}\right)}\Rightarrow\underbrace{\sum_{k=1}^{\infty}\sum_{i=0}^{m-1}\varepsilon_{(t_k,c_ij_k)}}_{\xi_m},\tag{2.31}$$

em que

$$\widetilde{T^*}\left(\sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{(t_k,j_k\mathbf{e}_i)}\right)=\widetilde{T^*}\left(\sum_{k=1}^{\infty}\sum_{i=0}^{m-1}\varepsilon_{(t_k,j_k\mathbf{e}_{i+1})}\right)$$
$$=\sum_{k=1}^{\infty}\sum_{i=0}^{m-1}\varepsilon_{T(t_k,j_k\mathbf{e}_{i+1})}=\sum_{k=1}^{\infty}\sum_{i=0}^{m-1}\varepsilon_{(t_k,c_ij_k)}.$$

Repare-se em primeiro lugar que, uma vez que é nula a probabilidade

$$P\left[\sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{(t_k,j_k\mathbf{e}_i)}\left([0,\infty)\times\left\{\mathbf{z}\in\overline{\mathbb{R}}^m\setminus\{0\}:\max_{0\le i\le m}|z_{k-i}|=\infty\right\}\right)>0\right],$$

então $\widetilde{T^*}$ é quase certamente contínua, se fôr contínua para qualquer medida m pertencente a $M_p\left([0,\infty)\times\overline{\mathbb{R}}^m\setminus\{0\}\right)$ e que verifique

$$m\left([0,\infty)\times\left\{\mathbf{z}\in\overline{\mathbb{R}}^m\setminus\{0\}:\max_{0\le i\le m}|z_{k-i}|=\infty\right\}\right)=0.$$

Suponhamos então que $f \in C_K^+\left([0,\infty)\times\overline{\mathrm{IR}}\setminus\{0\}\right)$ tal que o seu suporte compacto K, está contido no conjunto

$$D_\delta = [0,b)\times\left\{x\in\overline{\mathrm{IR}}\setminus\{0\}:|x|>\delta\right\} = [0,b)\times D_\delta^1 \tag{2.32}$$

Admita-se ainda que $m_n \xrightarrow{v} m$. Queremos mostrar que $\widetilde{T^\bullet}(m_n) \xrightarrow{v} \widetilde{T^\bullet}(m)$, ou seja, que

$$\widetilde{T^\bullet} m_n(f) := m_n \circ (T^\bullet)^{-1}(f) = m_n(f\circ T^\bullet) \to m(f\circ T^\bullet) =: \widetilde{T^\bullet} m(f),$$

e que a imagem inversa por meio da aplicação $T^\bullet$ do suporte compacto de f ainda seja um compacto. Uma vez que o suporte K de f está contido em D_δ, então o suporte de $f\circ T^\bullet = (T^\bullet)^{-1}(K)$ vai estar contido no conjunto

$$\mathrm{IR}_0^+\times\left(\left\{\mathbf{z}\in\overline{\mathrm{IR}}^m\setminus\{0\}:\left|\sum_{i=0}^{m-1}c_iZ_{k-i}\right|>\delta\right\}\cup\left\{\mathbf{z}\in\overline{\mathrm{IR}}^m\setminus\{0\}:\max_{0\le i\le m}|z_{k-i}|=\infty\right\}\right).$$

Para provarmos que $(T^\bullet)^{-1}(K)$ é compacto procede-se de forma análoga ao que foi feito na demonstração do teorema 2.2. Por outro lado, é fácil de ver que $f\circ T^\bullet$ é contínua em $[0,\infty)\times\mathrm{IR}^m\setminus\{0\}$. Escolha-se então um compacto K' que contenha o suporte de $f\circ T^\bullet$ e tal que $m_n(K') = m(K')$, para n suficientemente grande. Defina-se $P_n = \frac{m_n(\cdot)}{m(K')}$, para $n\geq 1$. Então, quando $n\to\infty$, tem-se

$$\widetilde{T^\bullet} m_n(f) = m(K')\int f\circ T^\bullet dP_n \to m(K')\int f\circ T^\bullet dP = \widetilde{T^\bullet} m(f),$$

pelo teorema 1.1 e uma vez que $f\circ T^\bullet$ é quase certamente contínua com respeito à medida de probabilidade P.

Note-se agora que quando $m\to\infty$, tem-se, quase certamente

$$\xi_m := \sum_{k=1}^{\infty}\sum_{i=0}^{m-1}\varepsilon_{(t_k,c_ij_k)} \to \sum_{K=1}^{\infty}\sum_{i=0}^{\infty}\varepsilon_{(t_k,c_ij_k)} =: \xi, \tag{2.33}$$

pontualmente na métrica vaga, ou seja, para quase todo o $w\in\Omega$,

$$\xi_m(w,\cdot) = \sum_{k=1}^{\infty}\sum_{i=0}^{m-1}\varepsilon_{(t_k,c_ij_k)}(w,\cdot) \xrightarrow{v} \sum_{K=1}^{\infty}\sum_{i=0}^{\infty}\varepsilon_{(t_k,c_ij_k)}(w,\cdot) = \xi(w,\cdot).$$

Então por (2.28), (2.33) e pelo lema 2.2, é suficiente mostrar que a convergência em (2.30) dá-se desde que $\forall\eta>0$

$$\lim_{n\to\infty}\limsup_{n\to\infty} P\left[\rho\left(\sum_{k=1}^{\infty}\varepsilon_{\left(\frac{k}{n},a_n^{-1}\sum_{i=0}^{m-1}c_iZ_{k-i}\right)},\sum_{k=1}^{\infty}\varepsilon_{\left(\frac{k}{n},\frac{X_k}{a_n}\right)}\right)>\eta\right]=0,$$

em que ρ é a métrica induzida pela topologia vaga em $M_p\left([0,\infty)\times\overline{\mathbb{R}}\setminus\{0\}\right)$. Para se ter o limite anterior, Davis e Resnick, no seu artigo em [7], afirmam com base num resultado apresentado em Kallenberg (1976, pag.95)[3], que é suficiente provar que dado qualquer $f\in C_K^+([0,\infty)\times\overline{\mathbb{R}}\setminus\{0\})$

$$\lim_{m\to\infty}\limsup_{n\to\infty}P\left[\left|\sum_{k=1}^{\infty}f\left(\frac{k}{n},a_n^{-1}\sum_{i=0}^{m-1}c_iZ_{k-i}\right)-\sum_{k=1}^{\infty}f\left(\frac{k}{n},a_n^{-1}X_k\right)\right|>\eta\right]=0. \tag{2.34}$$

Suponhamos então que o suporte de f está contido no conjunto D_δ definido em (2.32) e seja

$$w(\gamma)=\sup\{|f(t,x)-f(t,y)|:x,y\in(0,\infty)\,,\ ou\ \ x,y\in(-\infty,0)\ \ e\ |x-y|\leq\gamma\},$$

com $t\geq 0$. Repare-se que, como f é contínua num suporte compacto, então f é uniformemente contínua e portanto se $\gamma\to 0$, tem-se que $|f(t,x)-f(t,y)|\to 0$, donde $w(\gamma)\to 0$. Tome-se o conjunto

$$A_n=\{(W_{nk},Y_{nk}):\max_{1\leq k\leq n}(|W_{nk}-Y_{nk}|)\leq\gamma\}$$

com $W_{nk}=a_n^{-1}\sum_{i=0}^{m-1}c_iZ_{k-i}$ e $Y_{nk}=a_n^{-1}X_k$. Em A_n suponha-se que $\gamma<\frac{\delta}{2}$. Então se $|W_{nk}|\leq\frac{\delta}{2}$, como $\max\limits_{1\leq k\leq n}(|W_{nk}-Y_{nk}|)\leq\gamma$, vem

$$\begin{aligned}|W_{nk}-Y_{nk}| &\leq \gamma\Leftrightarrow W_{nk}-\gamma\leq Y_{nk}\leq W_{nk}+\gamma\Leftrightarrow-\frac{\delta}{2}-\frac{\delta}{2}\leq Y_{nk}\leq\frac{\delta}{2}+\frac{\delta}{2}\\ &\Leftrightarrow |Y_{nk}|\leq\delta,\end{aligned}$$

logo $W_{nk},Y_{nk}\notin D_\delta^1$ e portanto $\left(\frac{k}{n},W_{nk}\right),\left(\frac{k}{n},Y_{nk}\right)\notin D_\delta$, o que significa que $\left(\frac{k}{n},W_{nk}\right),\left(\frac{k}{n},Y_{nk}\right)$ não estão no suporte de f. Tem-se neste caso a igualdade $f(\frac{k}{n},W_{nk})=f\left(\frac{k}{n},Y_{nk}\right)=0$, donde o limite em (2.34) é nulo. Suponha-se agora que $|W_{nk}|>\frac{\delta}{2}$ (e portanto maior que γ). Temos então

$$\begin{aligned}|W_{nk}| &> \frac{\delta}{2}\wedge|W_{nk}-Y_{nk}|\leq\gamma\Leftrightarrow\left(W_{nk}>\frac{\delta}{2}\vee W_{nk}<-\frac{\delta}{2}\right)\wedge(|W_{nk}-Y_{nk}|\leq\gamma)\\ &\Leftrightarrow \underbrace{\left(W_{nk}>\frac{\delta}{2}\wedge(|W_{nk}-Y_{nk}|\leq\gamma)\right)}_{(1)}\vee\underbrace{\left(W_{nk}<-\frac{\delta}{2}\wedge(|W_{nk}-Y_{nk}|\leq\gamma)\right)}_{(2)}\end{aligned}$$

Resolvendo (1), obtemos

$$\left(W_{nk}>\frac{\delta}{2}\right)\wedge(Y_{nk}\leq W_{nk}+\gamma)\wedge(Y_{nk}\geq W_{nk}-\gamma)\,,$$

[3] Na impossibilidade da consulta da edição de 1976, por não se encontrar disponível no mercado bibliotecário o livro em questão, não é apresentado o resultado em causa. Curiosamente, na edição de 1983, não há qualquer referência a esse mesmo resultado.

logo,

$$Y_{nk} \geq \frac{\delta}{2} - \gamma > \gamma - \gamma = 0,$$

vindo então que $W_{nk}, Y_{nk} \in (0, \infty)$, ou seja, são ambos positivos. Resolvendo (2) temos

$$\left(W_{nk} < -\frac{\delta}{2}\right) \wedge (Y_{nk} \leq W_{nk} + \gamma) \wedge (Y_{nk} \geq W_{nk} - \gamma),$$

logo,

$$Y_{nk} \leq -\frac{\delta}{2} + \gamma < 0,$$

tendo-se consequentemente que $W_{nk}, Y_{nk} \in (-\infty, 0)$, ou seja são ambos negativos. Em qualquer dos casos, chegamos à conclusão que para $|W_{nk}| > \frac{\delta}{2}$ e $|W_{nk} - Y_{nk}| \leq \gamma$, tem-se que $W_{nk}, Y_{nk} \in (0, \infty)$ ou $W_{nk}, Y_{nk} \in (-\infty, 0)$, donde $\left|f\left(\frac{k}{n}, W_{nk}\right) - f\left(\frac{k}{n}, Y_{nk}\right)\right| \leq w(\gamma)$. Retomando agora o limite em (2.34), vem

$$\begin{aligned}
&= \lim_{m\to\infty} \limsup_{n\to\infty} P\left\{\left[\left|\sum_{k=1}^{\infty} f\left(\frac{k}{n}, W_{nk}\right) - \sum_{k=1}^{\infty} f\left(\frac{k}{n}, Y_{nk}\right)\right| > \eta\right] \cap A_n\right\} + \\
&\quad + \underbrace{\lim_{m\to\infty} \limsup_{n\to\infty} P\left\{\left[\left|\sum_{k=1}^{\infty} f\left(\frac{k}{n}, W_{nk}\right) - \sum_{k=1}^{\infty} f\left(\frac{k}{n}, Y_{nk}\right)\right| > \eta\right] \cap A_n^c\right\}}_{(3)}.
\end{aligned}$$

Note-se que o limite em (3) é nulo pois

$$\begin{aligned}
& P\left\{\left[\left|\sum_{k=1}^{\infty} f\left(\frac{k}{n}, W_{nk}\right) - \sum_{k=1}^{\infty} f\left(\frac{k}{n}, Y_{nk}\right)\right| > \eta\right] \cap (W_{nk}, Y_{nk}) \in A_n^c\right\} \\
\leq\ & P\left[(W_{nk}, Y_{nk}) \in A_n^c\right] \to 0,
\end{aligned}$$

por(2.29). Ficamos então apenas com o limite

$$\begin{aligned}
& \lim_{m\to\infty} \limsup_{n\to\infty} P\left\{\left[\left|\sum_{k=1}^{\infty} f\left(\frac{k}{n}, W_{nk}\right) - \sum_{k=1}^{\infty} f\left(\frac{k}{n}, Y_{nk}\right)\right| > \eta\right] \cap A_n\right\} \\
=\ & \lim_{m\to\infty} \limsup_{n\to\infty} P\left\{\left[\left|\sum_{k=1}^{\infty} f\left(\frac{k}{n}, W_{nk}\right) - f\left(\frac{k}{n}, Y_{nk}\right)\right| > \eta\right]] \mid A_n\right\} P(A_n) \\
\leq\ & \lim_{m\to\infty} \limsup_{n\to\infty} P\left\{\left[\left|\sum_{k:|W_{nk}|>\frac{\delta}{2}, \frac{k}{n}\in[0,b]}^{\infty} f\left(\frac{k}{n}, W_{nk}\right) - f\left(\frac{k}{n}, Y_{nk}\right)\right| > \eta\right] \mid A_n\right\} \\
\leq\ & \lim_{m\to\infty} \limsup_{n\to\infty} P\left\{\left[\sum_{k:|W_{nk}|>\frac{\delta}{2}, \frac{k}{n}\in[0,b]}^{\infty} \left|f\left(\frac{k}{n}, W_{nk}\right) - f\left(\frac{k}{n}, Y_{nk}\right)\right| > \eta\right] \mid A_n\right\}.
\end{aligned}$$

Ora, verificando-se A_n, vimos que $w(\gamma) \geq \left|f\left(\frac{k}{n}, W_{nk}\right) - f\left(\frac{k}{n}, Y_{nk}\right)\right| > \eta$, donde

$$\sum_{k:|W_{nk}|>\frac{\delta}{2},\frac{k}{n}\in[0,b]}^{\infty} w(\gamma) \geq \sum_{k:|W_{nk}|>\frac{\delta}{2},\frac{k}{n}\in[0,b]}^{\infty} \left|f\left(\frac{k}{n}, W_{nk}\right) - f\left(\frac{k}{n}, Y_{nk}\right)\right| > \eta.$$

Logo,

$$\begin{aligned}
&\lim_{m\to\infty}\limsup_{n\to\infty} P\left\{\left[\sum_{k:|W_{nk}|>\frac{\delta}{2},\frac{k}{n}\in[0,b]}^{\infty} \left|f\left(\frac{k}{n}, W_{nk}\right) - f\left(\frac{k}{n}, Y_{nk}\right)\right| > \eta\right] \mid A_n\right\} \\
\leq\ &\lim_{m\to\infty}\limsup_{n\to\infty} P\left[\sum_{k:|W_{nk}|>\frac{\delta}{2},\frac{k}{n}\in[0,b]}^{\infty} w(\gamma) > \eta\right] \\
=\ &\lim_{m\to\infty}\limsup_{n\to\infty} P[w(\gamma)\sum_{k=1}^{\infty}\varepsilon_{(\frac{k}{n}, W_{nk})}\left([0,b]\times\left\{z:|z|>\frac{\delta}{2}\right\}\right) > \eta] \\
=\ &\lim_{m\to\infty} P[w(\gamma)\sum_{k=1}^{\infty}\sum_{i=0}^{m-1}\varepsilon_{(t_k, c_i j_k)}\left([0,b]\times\left\{z:|z|>\frac{\delta}{2}\right\}\right) > \eta],
\end{aligned}$$

pelo teorema 2.2 e por (2.31). Atendendo a (2.33), o limite anterior vem igual a

$$P[w(\gamma)\sum_{k=1}^{\infty}\sum_{i=0}^{\infty}\varepsilon_{(t_k, c_i j_k)}\left([0,b]\times\left\{z:|z|>\frac{\delta}{2}\right\}\right) > \eta].$$

Uma vez que $\sum_{k=1}^{\infty}\sum_{i=0}^{\infty}\varepsilon_{(t_k,c_i j_k)}\left([0,b]\times\left\{z:|z|>\frac{\delta}{2}\right\}\right) < \infty$ quase certamente, fazendo $\gamma \to 0$,prova-se (2.34) ■

Uma consequência do teorema anterior é que para toda a função f contínua em $[0,\infty)\times \mathbb{R}\setminus\{0\}$ de suporte compacto,

$$\varsigma_n^{\cdot}(f) \xrightarrow{d} \varsigma(f),$$

ou seja,

$$\sum_{k=1}^{\infty} f\left(\frac{k}{n}, \frac{X_k}{a_n}\right) \xrightarrow{d} \sum_{k=1}^{\infty}\sum_{i=0}^{\infty} f(t_k, j_k c_i). \tag{2.35}$$

Os resultados que vimos até aqui envolvem apenas processos pontuais construídos com base numa sequência do tipo $\{(b_k, Y_k)\}$, com $\{b_k\}$, constantes não negativas em $[0,\infty)$ e $\{Y_k\}$ uma sucessão de v.a.'s aleatórias. É nosso interesse porém, analisar a convergência de processos em que intervenham sequências bidimensionais da forma $\{(Z_k, Y_k)\}$, onde cada uma das componentes é agora representada por v.a.'s. Com vista a definir esses processos consideremos adicionalmente a sequência de variáveis aleatórias i.i.d. $\{Z_k\}$ e a sequência $\{\Gamma_k\}$, definidas em (2.14) e (2.15), respectivamente. Tomando estes novos elementos aleatórios definam-se os processos

$$\zeta_n := \sum_{k=1}^{n} \varepsilon_{\left(Z_k, a_n^{-1} X_k\right)} \quad \text{e} \quad \zeta := \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \varepsilon_{\left(Z_k, \delta_k c_i \Gamma_k^{-1/\alpha}\right)}.$$

Neste momento o principal resultado que pretendemos analisar centra-se na convergência do processo ζ_n para o processo limite ζ. A linha de pensamento que nos vai guiar para atingir esse objectivo, é aquela que orientou o rol de proposições e teoremas que temos vindo a analisar neste capítulo. Todos os resultados que iremos obter em função destes dois processos, não serão mais do que uma extensão daqueles que até aqui foram estudados, razão pela qual optaremos oportunamente e conscientemente, por suprimir questões que se prendam com detalhes técnicos, evitando-se desse modo que se caia numa mera repetição de cálculos. Uma vez que vamos trabalhar com os pontos $\{(Z_k, \cdot)\}$, é conveniente extender o domínio das nossas variáveis a um espaço do tipo $\mathbb{R} \times (\cdot)$. Além disso vamos admitir que o processo linear $\{X_k\}$ passa a estar definido na forma

$$X_k := \sum_{j=1}^{\infty} c_j Z_{k-j}, \quad k \geq 1, \tag{2.36}$$

onde agora a somatório tem início em $j = 1$, em substituição de $j = 0$. Na prática isto equivale a retirar o primeiro termo do processo, mas este procedimento é necessário para garantir a independência de X_k e Z_k, já que sem ela não é possível estipularmos os resultados a que nos propomos. É claro que nas séries de números reais definidas em (2.8) e (2.11) passa-se igualmente a iniciar j em 1.

Vimos no início do capítulo que o processo $\sum_{k\geq 1} \varepsilon_{\left(\frac{k}{n}, \frac{Z_k}{a_n}\right)}$ converge fracamente para o processo $\sum_{k\geq 1} \varepsilon_{(t_k, j_k)}$, onde o último é um PRM em $[0,\infty) \times \overline{\mathbb{R}} \backslash \{0\}$ com medida de intensidade $\mu(dt, dx) = dt \times \upsilon(dx)$, com υ definida como em (2.13). Vimos também que essa convergência é equivalente a ter-se $nP\left[a_n^{-1} Z_1\right] \xrightarrow{v} \upsilon$. O nosso objectivo agora, é estabelecer um resultado semelhante, quando os processos em causa são

$$N_n := \sum_{k=1}^{n} \varepsilon_{\left(Z_k, a_n^{-1} Z_{k-1}\right)} \quad \text{e} \quad N := \sum_{k=1}^{\infty} \varepsilon_{\left(Z_k, \delta_k \Gamma_k^{-1/\alpha}\right)}.$$

Recorde-se que no exemplo 1.5, provámos que o processo $\sum_{k\geq 1} \varepsilon_{\delta_k \Gamma_k^{-1/\alpha}}$ é um PRM em $\overline{\mathbb{R}} \backslash \{0\}$, precisamente com medida de intensidade $\upsilon(dx)$. Por analogia à proposição 1.5, concluímos que o processo $\sum_{k\geq 1} \varepsilon_{\left(Z_k, \delta_k \Gamma_k^{-1/\alpha}\right)}$ é um $PRM(\mu)$ com μ dada por

$$\begin{aligned}
\mu(dz, dx) &= F(dz) \times \upsilon(dx) \\
&= F(dz) \times \alpha\left(px^{-\alpha-1} I_{(0,\infty)}(x) + (1-p)(-x)^{-\alpha-1} I_{(-\infty,0)}(x)\right) dx,
\end{aligned}$$

em que F é a lei de distribuição comum das variáveis Z_k's. Em que condições e quando é que se tem $N_n \Rightarrow N$ são questões que se colocam e que podem ser esclarecidas, bem como fundamentadas na proposição que se segue.

Proposição 2.5 *Sejam E_1 e E_2, espaços localmente compactos e com bases contáveis. Considere-se as sequências de elementos aleatórios i.i.d $\{Z_k\}_{k\in\mathbf{Z}}$ definida no espaço E_1 e $\{X_{n,k}, k \geq 1\}$, com $n \in \mathbb{N}$ e $\{W_k, k \geq 1\}$ definidas em E_2, tal que $\{X_{n,k}\}$, $\{Z_k\}$ e $\{W_k\}$ são independentes. Definam-se os processos*

$$N_{n,1} := \sum_{k=1}^{n} \varepsilon_{X_{n,k}} \quad e \quad N_1 := \sum_{k=1}^{\infty} \varepsilon_{W_k},$$

em que N_1 é um PRM (ν), com ν definida em E_2. Nestas condições se o processo $N_{n,1}$ converge fracamente para N_1, então

$$N_n := \sum_{k=1}^{n} \varepsilon_{\left(Z_k, X_{n,k}\right)} \Rightarrow \sum_{k=1}^{\infty} \varepsilon_{(Z_k, W_k)} =: N.$$

Dem. Recorde-se que $N_n \Rightarrow N$ sse $\Psi_{N_n}(f) \to \Psi_N(f)$, $\forall f \in C_k^+(E)$ ou ainda sse $-\log \Psi_{N_n}(f) \to -\log \Psi_N(f)$. Dado $f \in C_k^+(E_1 \times E_2)$ tem-se

$$\begin{aligned}
\Psi_{N_n}(f) &= E\left[\exp\left(-N_n(f)\right)\right] = E\left[\exp\left\{-\sum_{k=1}^{n} f\left(Z_k, X_{n,k}\right)\right\}\right] \\
&= \prod_{k=1}^{n} E\left(e^{-f\left(Z_k, X_{n,k}\right)}\right) \\
&= \prod_{k=1}^{n}\left[1 - \int_{E_1}\int_{\mathbf{R}}\left(1 - e^{-f(z,x)}\right) P\left(Z_k \in dz\right) P\left(X_{n,k} \in dx\right)\right].
\end{aligned}$$

Então

$$\begin{aligned}
-\log \Psi_{N_n}(f) &= -\sum_{k=1}^{n} \log\left[1 - \int_{E_1}\int_{\mathbf{R}}\left(1 - e^{-f(z,x)}\right) P\left(Z_k \in dz\right) P\left(X_{n,k} \in dx\right)\right] \\
&= -n \log\left[1 - \int_{E_1}\int_{\mathbf{R}}\left(1 - e^{-f(z,x)}\right) P\left(Z_1 \in dz\right) P\left(X_{n,1} \in dx\right)\right],
\end{aligned}$$

pois tanto a sequência $\{X_{n,k}\}$ como a sequência $\{Z_k\}$, são constituídas por v.a.'s i.i.d.. Designando

$$A_n := -\int_{E_1}\int_{\mathbf{R}}\left(1 - e^{-f(z,x)}\right) P(Z_1 \in dz) P(X_{n,1} \in dx),$$

se provarmos que $|-\log \Psi_{N_n}(f) + nA_n| \to 0$ e que $nA_n \to \log \Psi_N(f)$, então temos o resultado pretendido, ou seja, $-\log \Psi_{N_n}(f) \to -\log \Psi_N(f)$. Tendo em conta a expansão

$$\log(1+y) = y(1+\varepsilon(y)), \text{ onde } |\varepsilon(y)| \leq |y| \text{ se } |y| \leq \frac{1}{2},$$

obtem-se

$$\begin{aligned} |-\log \Psi_{N_n}(f) + nA_n| &= |-n\log(1+A_n) + nA_n| = n\,|\log(1+A_n) - A_n| \\ &= n\,|A_n\,(1+\varepsilon\,(A_n)) - A_n| = |nA_n|\,|\varepsilon(A_n)|\,. \end{aligned}$$

Por outro lado a proposição 1.11, garante que $N_{n,1} \Rightarrow N_1$ sse $nP\left[X_{n,1} \in \cdot\right] \xrightarrow{v} \nu\,(\cdot)$, donde

$$\begin{aligned} |nA_n| &\leq \int_{E_1}\int_{E_2} P(Z_1 \in dz)nP(X_{n,1} \in dx) \\ &\to \int_{E_1}\int_{E_2} P(Z_1 \in dz)\nu\,(dx) = \int_{E_2} \nu\,(dx) \\ &= a \in \mathbb{R}. \end{aligned}$$

Ora, se $|nA_n| \to a$, então $|A| \to 0$ e por conseguinte $|\varepsilon\,(A_n)| \to 0$, o que justifica que $|-\log \Psi_{N_n}(f) + nAn| \to 0$. Vejamos agora que $nA_n \to \log \Psi_N(f)$. Temos então

$$\begin{aligned} nA &= -n\int_{E_1}\int_{E_2} \left(1 - e^{-f((z,x))}\right) P(Z_1 \in dz)P(X_{n,1} \in dx) \\ &= -\int_{E_1}\int_{E_2} \left(1 - e^{-f(z,x)}\right) P(Z_1 \in dz)nP(X_{n,1} \in dx) \\ &\to \int_{E_1}\int_{E_2} \left(1 - e^{-f(z,x)}\right) P[Z_1 \in dz]\nu(dx), \end{aligned}$$

quando $n \to \infty$. Por sua vez,

$$\begin{aligned} \Psi_N(f) &= E(\exp\{-N(f)\}) = E\left[\exp\left\{-\sum_{k=1}^{\infty} f\,(Z_k, W_k)\right\}\right] \\ &= E\left[\prod_{k=1}^{\infty} e^{-f(Z_k,W_k)}\right] = E\left[E\left(\prod_{k=1}^{\infty} e^{-f(Z_k,W_k)}/\{W_k\}\right)\right] \\ &= E\left[\prod_{k=1}^{\infty}\int_{E_1} e^{-f(z,W_k)}K\,(dz, W_k)\right], \end{aligned} \tag{2.37}$$

onde $K\,(dz, W_k) = P\left[Z_k \in dz \mid \{W_k\}, \{Z_i, i \neq k\}\right] = P\left[Z_k \in dz\right] = P\left[Z_1 \in dz\right]$, pela independência e estacionaridade de $\{Z_k\}$. Defina-se

$$\theta(W_k) = \int_{E_2} e^{-f((z,w))}K\,(dz, W_k)\,, \qquad 0 \leq \theta \leq 1.$$

Então (2.37) vem dado por

$$\begin{aligned}E\left[\prod_{k=1}^{\infty}\theta(W_k)\right] &= E\left[\exp\left\{-\sum_{k=1}^{\infty}(-\log\theta(W_k))\right\}\right]\\ &= E\left[\exp\left\{-\int_{E_1}(-\log\theta(w)N_1(dw)\right\}\right] = \Psi_{N_1}(-\log\theta).\end{aligned}$$

Mas porque N_1 é um $PRM(\mu)$ então na expressão anterior obtemos

$$\begin{aligned}\Psi_{N_1}(-\log\theta) &= \exp\left\{-\int_{E_1}\left(1-e^{-(-\log\theta(w))}\right)\nu(dw)\right\}\\ &= \exp\left\{-\int_{E_1}(1-\theta(w)\nu(dw)\right\}\\ &= \exp\left\{-\int_{E_1}\left(1-\int_{E_2}e^{-f(z,w)}\right)K(dz,w)\nu(dw)\right\}\\ &= \exp\left\{-\int_{E_1}\left(1-\int_{E_2}e^{-f(z,w)}\right)K(dz,w)\nu(dw)\right\}\\ &= \exp\left\{-\int_{E_1}\int_{E_2}\left(1-e^{-f(z,w)}\right)P(Z_1\in dz)\nu(dw)\right\},\end{aligned}$$

pelo teorema de Fubini. Logo

$$\log\Psi_N(f) = -\int_{E_1}\int_{E_2}\left(1-e^{-f(z,w)}\right)P(Z_1\in dz)\nu(dw),$$

provando-se assim que $nA_n \longrightarrow \log\Psi_N(f)$, ou seja $\log\Psi_{N_n}(f)\longrightarrow\log\Psi_N(f)$, donde $N_n\Rightarrow N$ ■

Um corolário a reter desta proposição é que

$$\sum_{k=1}^{n}\varepsilon_{\left(Z_k,a_n^{-1}Z_{k-1}\right)} \Rightarrow \sum_{k=1}^{\infty}\varepsilon_{\left(Z_k,\delta_k\Gamma_k^{-1/\alpha}\right)}. \tag{2.38}$$

Note-se que no processo $\sum_{k=1}^{\infty}\varepsilon_{\left(Z_k,a_n^{-1}Z_{k-1}\right)}$, é indiferente tomar a v.a. Z_{k-1} ou $Z_{k-j}, j\neq 1$, pois os Z_k's têm idêntica distribuição para todo $k\in\mathbb{Z}$. Este resultado representa a peça fundamental do "puzzle" constituído por um vasto conjunto de propriedades a que obedecem os novos processos pontuais com que iremos trabalhar. Nesse sentido vamos considerar o espaço topológico de Hausdorff, $E=\mathbb{R}\times\overline{\mathbb{R}}^m\setminus\{0\}$, localmente compacto e com bases contáveis, tal que a σ-álgebra produto $\mathcal{E}$ é gerada por rectângulos $(m+1)$ dimensionais da forma

$$B=[b_0,c_0)\times(b_1,c_1]\times...\times(b_m,c_m],$$

onde $[b_0,c_0)$ é um semi-aberto de $\mathbb{R}$, o rectângulo $(b_1,c_1]\times...\times(b_m,c_m]$ é limitado a partir da origem e não contém a origem $\overline{0}=(0,...,0)$ e $b_i<c_i$, com

$b_i \neq 0$ e $c_i \in \overline{\mathbb{R}} \backslash \{0\}$ para $i = 1, ..., m$. À semelhança de (2.16), o conjuntos B verificam as condições (2.17) e (2.18), constituindo também um semianel DC, que representaremos igualmente por $\mathcal{S}$.

No espaço $M_p\left(\mathbb{R} \times \overline{\mathbb{R}}^m \backslash \{0\}\right)$ definam-se os processos

$$I_n := \sum_{k=1}^{n} \varepsilon_{\left(Z_k, \frac{\mathbf{Z}^{(k-1)}}{a_n}\right)} \quad \text{e} \quad I := \sum_{k=1}^{\infty} \sum_{i=1}^{m} \varepsilon_{\left(Z_k, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_i\right)}, \tag{2.39}$$

sendo agora $\mathbf{Z}^{(k-1)} = (Z_{k-1}, Z_{k-2}, ..., Z_{k-m}) \in \overline{\mathbb{R}}^m \backslash \{0\}$. Tal como se fez na proposição 2.3, vamos ver que I_n e I , não possuem pontos na fronteira do conjunto B e que os seus valores esperados são finitos.

Proposição 2.6 *Considerem-se os processos pontuais em (2.39), definidos no espaço* $M_p\left(\mathbb{R} \times \overline{\mathbb{R}}^m \backslash \{0\}\right)$. *Então*

a) $P[I(\delta B) = 0] = 1$ *para todo o* $B \in \mathcal{S}$.

b) $P[I(B) = 0] = 1$ *e* $E[I_n(B)] \underset{n}{\rightarrow} 0$, *para todo o* B *que satisfaça (2.17).*

c) $P[I(B) = 0] = \exp\{-\mu([b_0, c_0) \times (b_{i'}, c_{i'}])\}$ *e*

$$E[I_n(B)] \underset{n}{\rightarrow} \mu([b_0, c_0) \times (b_{i'}, c_{i'}]),$$

para todo o B *que satisfaça (2.18), onde* $\mu(dz, dx) = F(dz) \times \nu(dx)$ *é a intensidade do PRM definido em (2.38).*

Dem. a) A demonstração segue minuciosamente de perto a da proposição 2.3, razão pela qual se omitem os cálculos.

b) Dado o boreleano B que satisfaça (2.17), é óbvio que $P[I(B) = 0]$, uma vez que os pontos da forma $\delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_i$ estão sobre o eixo $\mathbf{e}_i$ e B não intersecta qualquer eixo. Quanto ao valor esperado de $I_n(B)$ temos

$$\begin{aligned}
E[I_n(B)] &= E\left[\sum_{k=1}^{n} \varepsilon_{\left(Z_k, \frac{\mathbf{Z}^{(k-1)}}{a_n}\right)}(B)\right] \\
&= \sum_{k=1}^{n} P\left[\left(Z_k, \frac{\mathbf{Z}^{(k-1)}}{a_n}\right) \in [b_0, c_0) \times (b_1, c_1] \times ... \times (b_m, c_m]\right] \\
&= \sum_{k=1}^{n} P[Z_1 \in [b_0, c_0)] \prod_{i=1}^{m} P\left(\frac{Z_1}{a_n} \in (b_i, c_i]\right) \\
&\leq \sum_{k=1}^{n} P[Z_1 \in [b_0, c_0)] \prod_{i=1}^{m} P\left(\left|\frac{Z_1}{a_n}\right| > \min(|b_i|, |c_i|)\right) \\
&= P[Z_1 \in [b_0, c_0)]\, nP\left(\left|\frac{Z_1}{a_n}\right| > d_i\right) \prod_{i=2}^{m} P\left(\left|\frac{Z_1}{a_n}\right| > d_i\right),
\end{aligned}$$

onde $d_i = \min(|b_i|, |c_i|)$. Como $\prod_{i=2}^{m} P\left(\left|\frac{Z_1}{a_n}\right| > d_i\right) \underset{n\to\infty}{\longrightarrow} 0$ e por outro lado $nP\left(\left|\frac{Z_1}{a_n}\right| > d_i\right) \to d_i^{-\alpha}$, então $E[I_n(B)] \to 0$.

c) Seja $B \in \mathcal{S}$, tal que (2.18) é satisfeita. Como o processo I é um PRM, resulta por definição que

$$\begin{aligned} P\left[\sum_{k=1}^{\infty}\sum_{i=1}^{m}\varepsilon_{\left(Z_k,\delta_k\Gamma_k^{-1/\alpha}\mathbf{e}_i\right)}(B) = 0\right] &= P\left[\sum_{k=1}^{\infty}\varepsilon_{\left(Z_k,\delta_k\Gamma_k^{-1/\alpha}\right)}([b_0,c_0)\times(b_{i'},c_{i'}]) = 0\right] \\ &= \exp\{-\mu([b_0,c_0)\times(b_{i'},c_{i'}])\}. \end{aligned}$$

Por outro lado,

$$\begin{aligned} E[I_n(B)] &= \sum_{k=1}^{n} P[Z_1\in[b_0,c_0)]\prod_{i=1}^{m} P\left(\frac{Z_1}{a_n}\in(b_i,c_i]\right) \\ &= \sum_{k=1}^{n} P[Z_1\in[b_0,c_0)]\,P\left(\frac{Z_1}{a_n}\in(b_{i'},c_{i'}]\right)\prod_{i\neq i'}^{m} P\left(\frac{Z_1}{a_n}\in(b_i,c_i]\right) \\ &= P[Z_1\in[b_0,c_0)]\,nP\left(\left|\frac{Z_1}{a_n}\right|\in(b_{i'},c_{i'}]\right)\underbrace{\prod_{i\neq i'}^{m} P\left(\frac{Z_1}{a_n}\in(b_i,c_i]\right)}_{\hookrightarrow 1} \\ &\to [F(c_0)-F(b_0)]\,\nu((b_{i'},c_{i'}]) = \mu([b_0,c_0)\times(b_{i'},c_{i'}]), \end{aligned}$$

tomando $n\to\infty$. Repare-se que a medida do boreleano $[b_0,c_0)\times(b_{i'},c_{i'}]$, vem dada por

$$\int_{b_0}^{c_0}\int_{b_{i'}}^{c_{i'}} F(dz)\,\nu(dx) = \begin{cases} -p\,(F(c_0)-F(b_0))\left(c_{i'}^{-\alpha}-b_{i'}^{-\alpha}\right), & \text{se } 0<b_{i'}<c_{i'} \\ q\,(F(c_0)-F(b_0))\left((-c_{i'})^{-\alpha}-(-b_{i'})^{-\alpha}\right), & \text{se } b_{i'}<c_{i'}<0. \end{cases}$$

(■)

Considere-se agora o processo $I_n^* := \sum_{k=1}^{n}\sum_{i=1}^{m}\varepsilon_{\left(Z_k,a_n^{-1}Z_{k-1}\mathbf{e}_i\right)}$. Tomando $B\in\mathcal{S}$, vamos provar que a v.a. $I_n(B)-I_n^*(B)$ é um $o_p(1)$ e que $I_n^*(B)$ converge fracamente para I, tendo-se consequentemente que $I_n \Rightarrow I$.

Proposição 2.7 *Em $M_p\left(\mathbb{R}\times\overline{\mathbb{R}}^m\setminus\{0\}\right)$ considerem-se os processos I_n e I_n^*. Então para cada $B\in\mathcal{S}$, verifica-se que*

$$I_n(B)-I_n^*(B)\to 0,$$

em probabilidade.

Dem. Se B verifica (2.17), tem-se

$$E\left[I_n^*(B)\right] = E\left[\sum_{k=1}^{n}\sum_{i=1}^{m}\varepsilon_{\left(Z_k, a_n^{-1}Z_{k-1}\mathbf{e}_i\right)}(B)\right] = 0,$$

pois B não intersecta os eixos. Neste caso temos $0 \leq I_n(B) - I_n^*(B)$ e $0 \leq E\left[I_n(B) - I_n^*(B)\right] \longrightarrow 0$, quando $n \to \infty$, donde $I_n(B) - I_n^*(B) \xrightarrow{P} 0$. Se B verifica (2.18) então

$$\begin{aligned}
E\left[I_n^*(B)\right] &= E\left[\sum_{k=1}^{n}\sum_{i=1}^{m}\varepsilon_{\left(Z_k, a_n^{-1}Z_{k-1}\mathbf{e}_i\right)}(B)\right] \\
&= \sum_{k=1}^{n} P\left[\left(Z_k, a_n^{-1}Z_{i'-1}\right) \in [b_0, c_0) \times (b_{i'}, c_{i'}]\right] \\
&= \sum_{k=1}^{n} P\left[Z_1 \in [b_0, c_0)\right] P\left(\frac{Z_1}{a_n} \in (b_{i'}, c_{i'}]\right) \\
&= P\left[Z_1 \in [b_0, c_0)\right] nP\left(\frac{Z_1}{a_n} \in (b_{i'}, c_{i'}]\right) \\
&\longrightarrow \mu\left([b_0, c_0) \times (b_{i'}, c_{i'}]\right).
\end{aligned}$$

Por outro lado a v.a. $I_n(B)$ é igual à seguinte soma

$$I_n(B) = \underbrace{\sum_{k=1}^{i'-1}\varepsilon_{\left(Z_k, a_n^{-1}\mathbf{Z}^{(k-1)}\right)}(B)}_{I_{1,n}(B)} + \underbrace{\sum_{k=i'}^{n}\varepsilon_{\left(Z_k, a_n^{-1}\mathbf{Z}^{(k-1)}\right)}(B)}_{I_{2,n}(B)}.$$

Calculando o valor esperado de $I_{1,n}(B)$, obtemos

$$\begin{aligned}
E\left[I_{1,n}(B)\right] &= E\left[\sum_{k=1}^{i'-1}\varepsilon_{\left(Z_k, a_n^{-1}\mathbf{Z}^{(k-1)}\right)}(B)\right] \\
&= \sum_{k=1}^{i'-1} P\left[Z_1 \in [b_0, c_0)\right] \prod_{i=1}^{m} P\left(\frac{Z_1}{a_n} \in (b_i, c_i]\right) \\
&= \sum_{k=1}^{i'-1} P\left[Z_1 \in [b_0, c_0)\right] P\left(\frac{Z_1}{a_n} \in (b_{i'}, c_{i'}]\right) \prod_{i \neq i'}^{m} P\left(\frac{Z_1}{a_n} \in (b_i, c_i]\right) \\
&= (i'-1)\, P\left[Z_1 \in [b_0, c_0)\right] P\left(\frac{Z_1}{a_n} \in (b_{i'}, c_{i'}]\right) \underbrace{\prod_{i \neq i'}^{m} P\left(\frac{Z_1}{a_n} \in (b_i, c_i]\right)}_{\hookrightarrow 1} \\
&\longrightarrow 0,
\end{aligned}$$

por definição de a_n. Como $E[I_n(B)] \to \mu([b_0,c_0) \times (b_{i'},c_{i'}])$, então obtemos $E[I_{2,n}(B)] \underset{n\to\infty}{\longrightarrow} \mu([b_0,c_0) \times (b_{i'},c_{i'}])$. Além disso tem-se que

$$\begin{aligned} I_{2,n}(B) &= \sum_{k=i'}^{n} \varepsilon_{\left(Z_k, a_n^{-1}\mathbf{Z}^{(k-1)}\right)}(B) \stackrel{d}{=} \sum_{j=1}^{n-i'+1} \varepsilon_{\left(Z_j, a_n^{-1}\mathbf{Z}^{(j-1)}\right)}(B) \\ &= \sum_{j=1}^{n-i'+1} \varepsilon_{\left(Z_j, a_n^{-1}Z_{j-1}\right)}([b_0,c_0) \times (b_{i'},c_{i'}]) \\ &\leq \sum_{j=1}^{n} \varepsilon_{\left(Z_j, a_n^{-1}Z_{j-1}\right)}([b_0,c_0) \times (b_{i'},c_{i'}]) = I_n^*(B). \end{aligned}$$

Logo $0 \leq E[I_n(B) - I_n^*(B)] = E[I_{2,n}(B) - I_n^*(B)] + E[I_{1,n}(B)] \underset{n\to\infty}{\longrightarrow} 0$, resultando que $I_n^*(B) - I_n(B) \to 0$, em probabilidade. ∎

Teorema 2.4 *Seja $\{Z_k\}$ a sequência de variáveis i.i.d. que verificam (2.5), (2.6) e (2.10). Considerem-se ainda as sequências $\{\delta_k\}_{k\geq 1}$ e $\{\Gamma_k\}_{k\geq 1}$, que satisfazem respectivamente (2.14) e (2.15). Então no espaço $M_p\left(\mathbb{R} \times \overline{\mathbb{R}}^m \backslash \{0\}\right)$, $m > 1$,*

$$\sum_{k=1}^{n} \varepsilon_{\left(Z_k, a_n^{-1}\mathbf{Z}^{(k-1)}\right)} \Rightarrow \sum_{k=1}^{\infty}\sum_{i=1}^{m} \varepsilon_{\left(Z_k, \delta_k \Gamma_k^{-1/\alpha}\mathbf{e}_i\right)}. \tag{2.40}$$

Dem. A demonstração é análoga à do teorema 2.2, por isso apresentamos aqui apenas as principais linhas de raciocínio que conduzem ao resultado (2.40). Vimos que o processo $\sum_{k=1}^{n} \varepsilon_{\left(Z_k, a_n^{-1}Z_{k-1}\right)}$ converge fracamente para o processo $\sum_{k=1}^{\infty} \varepsilon_{\left(Z_k, \delta_k \Gamma_k^{-1/\alpha}\right)}$. Por outro lado a composição das aplicações contínuas

$$\sum_{k=1} \varepsilon_{(u_k,v_k)} \mapsto \left(\sum_{k=1} \varepsilon_{(u_k,v_k\mathbf{e}_1)}, \ldots, \sum_{k=1} \varepsilon_{(u_k,v_k\mathbf{e}_m)}\right) \mapsto \sum_{k=1}\sum_{i=1}^{m} \varepsilon_{(u_k,v_k\mathbf{e}_i)}$$

é uma aplicação contínua de $M_p\left(\mathbb{R} \times \overline{\mathbb{R}} \backslash \{0\}\right)$ para $M_p\left(\mathbb{R} \times \overline{\mathbb{R}}^m \backslash \{0\}\right)$, que designaremos por h. Então pelo teorema 1.1, da aplicação contínua tem-se

$$h\left(\sum_{k=1}^{n} \varepsilon_{\left(Z_k, a_n^{-1}Z_{k-1}\right)}\right) \Rightarrow h\left(\sum_{k=1}^{\infty} \varepsilon_{\left(Z_k, \delta_k \Gamma_k^{-1/\alpha}\right)}\right),$$

ou seja

$$I_n^* = \sum_{k=1}^{n}\sum_{i=1}^{m} \varepsilon_{\left(Z_k, a_n^{-1}Z_{k-1}\mathbf{e}_i\right)} \Rightarrow \sum_{k=1}^{\infty}\sum_{i=1}^{m} \varepsilon_{\left(Z_k, \delta_k \Gamma_k^{-1/\alpha}\mathbf{e}_i\right)} = I.$$

Aplicando o teorema 4.2 em [19], pág.32, prova-se que $I_n^* \Rightarrow I^*$ sse

$$\left(I_n^*(B_1), ..., I_n^*(B_j)\right) \Rightarrow \left(I^*(B_1), ..., I^*(B_j)\right),$$

para quaisquer $B_1, ..., B_j \in \mathcal{S}$, com $j \geq 1$. Ora, como $I_n(B) - I_n^*(B)$ converge em probabilidade para zero para qualquer $B \in \mathcal{S}$, vem que

$$\left(I_n(B_1), ..., I_n(B_j)\right) \Rightarrow \left(I^*(B_1), ..., I^*(B_j)\right),$$

donde pelo mesmo teorema resulta que $I_n \Rightarrow I$. ■

O próximo e último teorema que apresentamos neste capítulo consolida a utilização da teoria dos processos pontuais no estudo da teoria limite para modelos de Médias Móveis, quando a distribuição dos erros pertence ao domínio de atracção das estáveis de índice $\alpha \in (0,2)$. Aliás este teorema é o motor dos princípios teóricos que envolvem o estimador LAD, quando os processos em causa são os autoregressivos. Sobre este tema falaremos no capítulo que se segue.

Teorema 2.5 *Suponhamos que $\{a_n\}$ satisfaz (2.10), $\{c_j\}$ satisfaz (2.8) e $\{Z_k\}$ satisfaz (2.5) e (2.6). Sejam $\{X_k\}$, o processo linear definido em (2.36), $\{\delta_k\}$ a sequência de v.a.'s i.i.d. que verifica (2.14) e $\{\Gamma_k\}$ a sequência que satisfaz (2.15). Suponha-se ainda que $\{Z_k\}$, $\{\delta_k\}$ e $\{E_i\}$ são independentes. Então em $M_p(\mathbb{R} \times \overline{\mathbb{R}} \setminus \{0\})$, quando $n \to \infty$, tem-se*

$$\sum_{k=1}^{n} \varepsilon_{\left(Z_k, a_n^{-1} X_k\right)} \Rightarrow \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} \varepsilon_{\left(Z_k, \delta_k c_i \Gamma_k^{-1/\alpha}\right)}. \tag{2.41}$$

Dem. Vimos no teorema anterior que

$$\sum_{k=1}^{n} \varepsilon_{\left(Z_k, a_n^{-1} \mathbf{Z}^{(k-1)}\right)} \Rightarrow \sum_{k=1}^{\infty} \sum_{i=1}^{m} \varepsilon_{\left(Z_k, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_i\right)}.$$

Por outro lado a aplicação $\widetilde{T^*}$ de $M_p(\mathbb{R} \times \overline{\mathbb{R}}^m \setminus \{0\})$ para $M_p(\mathbb{R} \times \overline{\mathbb{R}} \setminus \{0\})$, definida por

$$\sum_{k=1} \varepsilon_{(u_k, z_{k-1}, z_{k-2}, ..., z_{k-m})} \longmapsto \begin{cases} \sum_{k=1} \varepsilon_{\left(u_k, \sum_{i=1}^{m} c_i z_{k-i}\right)} \\ \sum_{k=1} \varepsilon_{(u_k, \infty)} \end{cases},$$

é quase certamente contínua com respeito à distribuição de $\sum_{k=1}^{\infty} \sum_{i=1}^{m} \varepsilon_{\left(Z_k, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_i\right)}$. Então pelo teorema 1.1, vem

$$\widetilde{T^*}\left(\sum_{k=1}^{n} \varepsilon_{\left(Z_k, a_n^{-1} \mathbf{Z}^{(k-1)}\right)}\right) \Rightarrow \widetilde{T^*}\left(\sum_{k=1}^{\infty} \sum_{i=1}^{m} \varepsilon_{\left(Z_k, \delta_k \Gamma_k^{-1/\alpha} \mathbf{e}_i\right)}\right),$$

isto é,

$$\sum_{k=1}^{n} \varepsilon_{\left(Z_k, a_n^{-1} \sum_{i=1}^{m} c_i z_{k-i}\right)} \Rightarrow \sum_{k=1}^{\infty}\sum_{i=1}^{m} \varepsilon_{\left(Z_k, \delta_k c_i \Gamma_k^{-1/\alpha}\right)}.$$

Tomando $m \to \infty$, tem-se quase certamente que

$$\sum_{k=1}^{\infty}\sum_{i=1}^{m} \varepsilon_{\left(Z_k, \delta_k c_i \Gamma_k^{-1/\alpha}\right)} \to \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \varepsilon_{\left(Z_k, \delta_k c_i \Gamma_k^{-1/\alpha}\right)},$$

pontualmente para a métrica vaga. Então aplicando o lema 2.2, a convergência em (2.41) ocorre desde que se verifique a condição

$$\lim_{m\to\infty} \limsup_{n\to\infty} P\left[\left|\sum_{k=1}^{\infty} f\left(Z_k, a_n^{-1}\sum_{i=1}^{m} c_i Z_{k-i}\right) - \sum_{k=1}^{\infty} f\left(Z_k, a_n^{-1} X_k\right)\right| > \eta\right] = 0, \tag{2.42}$$

para toda a função $f \in C_K^+([0,\infty) \times \overline{\mathbb{R}} \setminus \{0\})$. Admitindo que o suporte de f está contido no conjunto

$$D = [a,b] \times \left\{x \in \overline{\mathbb{R}} \setminus \{0\} : |x| > \delta, \text{ com } \delta > 0\right\}, a, b \in \mathbb{R}$$

e definindo-se

$$w(\gamma) = \sup\{|f(t,x) - f(t,y)| : x, y \in (0,\infty) \quad ou \quad x, y \in (-\infty, 0) \quad e \ |x-y| \leq \gamma\},$$

com $t \in \mathbb{R}$, prova-se após alguns cálculos, à semelhança dos que foram efectuados na demonstração do teorema 2.3, que dado o conjunto

$$A_n = \{(W_{nk}, Y_{nk}) : \max_{1\leq k\leq n} (|W_{nk} - Y_{nk}|) \leq \gamma\},$$

com $W_{nk} = a_n^{-1} \sum_{i=1}^{m} c_i Z_{k-i}$ e $Y_k = a_n^{-1} X_k$, o limite em (2.42) é nulo, provando-se desse modo (2.41). ∎

Resulta como consequência deste teorema que

$$\sum_{k=1}^{n} f\left(Z_k, a_n^{-1} X_k\right) \xrightarrow{d} \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} f\left(Z_k, \delta_k c_i \Gamma_k^{-1/\alpha}\right), \tag{2.43}$$

para qualquer função f contínua em $\mathbb{R} \times \overline{\mathbb{R}} \setminus \{0\}$ de suporte compacto.

Capítulo 3

M-Estimação para Processos Autoregressivos com Variância Infinita. Estimador LAD

No capítulo anterior, analisámos algumas propriedades relativas à convergência fraca de processos pontuais, quando associados a processos lineares de médias móveis, os quais por sua vez apresentavam a particularidade de serem gerados por ruídos com distribuição de variação regular de índice $-\alpha$ ($\alpha > 0$) nas caudas e estas obedecerem por sua vez às condições (2.5) e (2.6). Também vimos que quando esse expoente de variação, verifica $0 < \alpha < 2$, os erros têm variância infinita. Uma classe importante de distribuições que satisfazem precisamente (2.5) e (2.6), para $\alpha \in (0, 2)$, é classe *das distribuições estáveis ou* das *leis estáveis.* Embora possa parecer pouco provável, a existência de distribuições empíricas com variância infinita, muitos estudos mostraram, que certos fenómenos, quer sejam de natureza económica, social ou outra, sugeriam o uso de modelos ARMA com variância infinita. Entre esses estudos, salientamos as flutuações de preços de mercado (Fama, 1965), que sendo variáveis limitadas, tomam com muita frequência valores nas caudas da distribuição, eliminando assim peremptoriamente a hipótese de normalidade. Também os sinais de telefone, estudados por Stuck e Kleiner (1974), mostraram claramente a necessidade de se recorrer a modelos com variância infinita. Quando se entra no domínio da estimação de um processo linear, usualmente recorre-se às condições de segunda ordem do processo, descritas a partir da função de covariândia ou da função de densidade espectral, para se obterem os estimadores dos parâmatros. Porém, quando um processo é dirigido por inovações com variância infinita, esses estimadores perdem algumas das suas propriedades, nomeadamente a consistência, deixando por isso de serem óptimos. A *M-estimação* aparece assim como uma alternativa, que inclui entre outros, métodos conhecidos de estimação, tais como o método dos mínimos quadrados (LS-least-squares) e o método dos mínimos desvios absolutos (LAD-*least-absolute deviations*). Neste capítulo vamos restringir o nosso estudo ao comportamento limite do estimador-LAD dos parâmetros de um processo linear autoregressivo, quando dirigido por erros com variância infinita, isto é, vamos admitir que a distribuição comum desses erros fazem parte do domínio de atracção das leis estáveis de índice $\alpha \in (0,2)$. Para já apresentamos alguns conceitos básicos relacionados com as variáveis estáveis que podem ser analisados com mais rigor no sexto capítulo em [13].

3.1 Variáveis Estáveis. Conceitos básicos

Uma **variável aleatória** Z diz-se **estável**, ou tem **distribuição estável** sse para todo o inteiro positivo n, existam constantes $a_n > 0$ e $b_n \in \mathbb{R}$, tais que

$$Z_1 + ... + Z_n \stackrel{d}{=} a_n Z + b_n, \tag{3.1}$$

quaisquer que sejam as variáveis aleatórias i.i.d $Z_1, ..., Z_n$ com a mesma distribuição F de Z. Quando $b_n = 0$, a variável aleatória Z diz-se **estritamente estável** ou **simétrica estável**. As constantes normalizadoras a_n, tomam a forma assintótica $a_n \sim n^{1/\alpha}$, quando $0 < \alpha \leq 2$, sendo a constante α designada por **expoente** ou **índice característico** da v.a. Z. Além da propriedade aditiva expressa em (3.1), de que gozam estas variáveis, uma outra propriedade interessante é a possibilidade de se escolher uma constante centralizadora b, de forma a que a v.a. $Z - b$, seja estritamente estável, isto quando $\alpha \neq 1$. Outras propriedades importantes que satisfazem estas variáveis, são apresentadas de seguida.

3.1.1 Algumas Propriedades das Variáveis Estáveis

1. A função característica, $\phi(u) = E[\exp(iuZ)]$, de uma variável estável de expoente característico $\alpha \in (0,2]$, é dada por

$$\phi(u) = \begin{cases} \exp\{iu\beta - d|u|^{\alpha}(1 - i\theta sgn(u)\tan(\pi/2))\} & se \quad \alpha \neq 1 \\ \exp\{iu\beta - d|u|(1 + i\theta(2/\pi) sgn(u)\ln|u|)\} & se \quad \alpha = 1 \end{cases},$$

   onde $sgn(u) = \frac{u}{|u|}$, se $u \neq 0$ e zero caso contrário. Os reais $\beta \in \mathbb{R}$, $d^{1/\alpha} \in [0, \infty)$ e $\theta \in [-1, 1]$, são designados respectivamente por parâmetros de localização, de escala e de simetria. Quando $\beta = 0$ e $d = 1$, a v.a. **estável** Z, diz-se **standard**.

2. Se $\alpha = 2$ então $Z \sim N(\beta, 2d)$.

3. Todas as variávies estáveis de índice $\alpha \in (0,2)$ apresentam variância infinita.

4. Se $\theta = 0$, então a distribuição de Z é simétrica em relação a β. As distribuições estáveis simétricas ($\beta = 0$), têm a função caracterítica do tipo $\phi(u) = e^{-d|u|^{\alpha}}$, para $\alpha \neq 1$.

5. Se $\alpha = 1$ e $\theta = 0$, Z tem distribuição de Cauchy, com função densidade de probabilidade dada por

$$f(x) = d/\pi \left[d^2 + (x - \beta)^2\right]^{-1}, x \in \mathbb{R}.$$

6. Se $\alpha = 1/2$ e $\beta = 1$, Z tem distribuição de Lévy.

7. Toda a distribuição estável de expoente característico $\alpha \in (0, 2)$, tem momentos absolutos finitos de ordem $\delta \in (0, \alpha)$.

8. A função característica de uma variável estável, é absolutamente integrável, o que implica que toda a lei estável é absolutamente contínua.

9. Se Z é estritamente estável com expoente característico α, então

$$s^{1/\alpha} Z_1 + t^{1/\alpha} Z_2 \stackrel{d}{=} (s + t)^{1/\alpha} Z, \forall s > 0, \forall t > 0,$$

com Z_1, Z_2 e Z, v.a.'s i.i.d. Esta propriedade implica no caso geral, que todas as combinações lineares de estáveis da forma $\lambda_1 Z_1 + \lambda_2 Z_2$, ainda seja uma estável.

Gnedenko (1968), mostrou que as distribuições estáveis são as únicas distribuições limite de somas de v.a.'s i.i.d., quando convenientemente normadas. Ora, esse resultado fez com que surgisse naturalmente uma nova descrição das leis estáveis. Assim, dadas as v.a.'s i.i.d. $Z_1, ..., Z_n, Z$, se para cada $n \in \mathbb{N}$, existirem constantes $a_n > 0$ e $b_n \in \mathbb{R}$, que verificam

$$\frac{\sum_{k=1}^{n} Z_k - b_n}{a_n} \xrightarrow{d} Y,$$

com Y v.a. não degenerada, então Y é estável ou tem distribuição estável. Neste caso, $Z_k \stackrel{d}{=} Z$ diz-se que pertence ao **domínio de atracção** da v.a. estável Y e simbolicamente escreve-se $Z \in D(Y)$ ou ainda $Z \in D(\alpha)$. As constantes normalizadoras são agora definidas por

$$a_n = n^{1/\alpha} L(n),$$

onde $L(\cdot)$ é uma função de variação lenta . Se $L(n) = C \in \mathbb{R}^+$, então a v.a. Z, diz-se que pertence ao domínio de atracção standard da v.a. Y. As caudas da função de distribuição $F_Z(x)$ de uma variável estável Z, com $0 < \alpha < 2$, também conhecidas por **caudas elevadas**, satisfazem para todo $x > 0$

$$1 - F_Z(x) + F_Z(-x) = P(|Z| > x) = x^{-\alpha} L(x), x \to \infty$$

e

$$\frac{F_Z(-x)}{1-F_Z(x)} = \frac{P[Z \leq -x]}{P[Z > x]} \to \frac{c_1}{c_2}, \; x \to \infty,$$

com $c_1, c_2 \geq 0$ e $c_1 + c_2 > 0$. Se tomarmos $c_1 = q$ e $c_2 = p$, tal que $p + q = 1$ e $0 \leq p \leq 1$, vemos que as condições (2.5) e (2.6), verificadas pela sequência $\{Z_k\}$, geradora do processo linear médias móveis definido em (2.7), são necessárias e suficientes para que $Z_k \in D(\alpha)$, sempre que $\alpha \in (0,2)$.

3.2 M-Estimação para Autoregressões

Seja $\{X_t\}$ é um processo autoregressivo causal de ordem p, AR(p), da forma

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + Z_t, \tag{3.2}$$

onde $\{Z_t\}_{t\in\mathbf{Z}}$ é uma sequência de variáveis aleatórias i.i.d. tais que $Z_t \in D(\alpha)$. Dadas as observações $X_1, ..., X_n$, geradas a partir do modelo em (3.2), uma M-estimativa $\widehat{\phi}_M$, do vector dos verdadeiros parâmetros $\phi = (\phi_1, ..., \phi_p)'$, é o vector que minimiza com respeito a β, a função objectivo

$$\sum_{t=1}^{n} \rho[Z_t(\beta)],$$

onde $\rho(\cdot)$ é uma função perda e $\{Z_t(\beta)\}$ são estimativas dos erros $\{Z_t\}$. Estas estimativas podem ser calculadas a partir de um qualquer vector particular de parâmetros $\beta = (\beta_1, ..., \beta_p)'$, da seguinte forma

$$\begin{aligned}
Z_1(\beta) &= X_1 \\
Z_2(\beta) &= X_2 - \beta_1 X_1 \\
Z_2(\beta) &= X_3 - \beta_2 X_2 - \beta_1 X_1 \\
&\vdots \\
Z_n(\beta) &= X_n - \beta_1 X_{n-1} - \beta_2 X_{n-2} - ... - \beta_p X_{n-p}.
\end{aligned}$$

Os casos em que $\rho(x) = x^2$ e $\rho(x) = |x|$, correspondem respectivamente aos métodos de estimação dos mínimos quadrados e aos mínimos desvios absolutos. Para se analisar o comportamento assintótico do vector de parâmetros $\widehat{\phi}_M$, o procedimento usual, passa pela determinação das derivadas parciais da função objectivo

$$\sum_{t=1}^{n} \rho(X_t - \beta_1 X_{t-1} - ... - \beta_p X_{t-p}),$$

obrigando a que as estimativas $\widehat{\phi}_1, ..., \widehat{\phi}_p$, verifiquem

$$\sum_{t=1}^{n} X_{t-j}\psi\left(X_t - \widehat{\phi}_1 X_{t-1} - ... - \widehat{\phi}_p X_{t-p}\right) = 0, \quad \text{para } j = 1, ..., p, \tag{3.3}$$

sendo $\psi(\cdot)$ a derivada da função $\rho(\cdot)$. Dois problemas se levantam à priori com este tipo de procedimento. Em primeiro lugar se $\rho(x) = |x|$, e nesse caso estamos na situação do estimador LAD, ρ não é diferenciável na origem sendo necessário por isso impôr condições de momentos que envolvam o comportamento da distribuição de Z_t, perto da origem. Em segundo, se no caso da variância finita, as equações em (3.3), levam a que $\sqrt{n}\left(\widehat{\phi}_M - \phi\right) \xrightarrow{d} \mathbf{X}$, com $\mathbf{X}$ um vector normal, no caso em que a variância é infinita não é possível, de uma forma geral, estabelecer-se um resultado similar por este procedimento, nem tão pouco, obter-se uma representação concreta em termos da distribuição de $\mathbf{X}$. Iremos ver para o caso particular do estimador LAD, como contornar este problema encontrando uma forma alternativa.

Embora esteja fora do âmbito deste trabalho o estudo das propriedades assintóticas do estimador LS, para processos lineares autoregressivos dirigidos por inovações com variância infinita, não queremos deixar de registar alguns resultados, em termos de consistência e rácios de convergência, que têm vindo a ser desenvolvidos para este estimador e naturalmente para o estimador LAD. Assim, Hannan e Kanter (1977) mostraram que para $\delta > \alpha$,

$$n^{1/\delta}\left(\widehat{\phi}_{LS} - \phi\right) \to 0 \quad \text{q.c.}$$

Davis e Resnick (1985-86) provaram que existe uma função de variação lenta em ∞ ,$L_0(n)$, tal que

$$n^{1/\alpha} L_0(n)\left(\widehat{\phi}_{LS} - \phi\right) \xrightarrow{d} \mathbf{Y},$$

em que $\mathbf{Y}$ é o quociente entre duas variáveis aleatórias estáveis. Gross e Steiger (1979) analisaram a consistência do estimador LAD, sob a hipótese do valor de $E(|Z_t|)$ ser convergente e Z_t ter uma única mediana na origem, chegando à conclusão de que nestas condições o estimador é fortemente consistente. An e Chen (1982) provaram que para $\delta > \alpha$

$$n^{1/\delta}\left(\widehat{\phi}_{LAD} - \phi\right) \xrightarrow{P} 0,$$

e determinaram um rácio de convergência para $\widehat{\phi}_{LAD}$, quando Z_t tem uma única mediana na origem e $1 < \alpha < 2$, ou quando Z_t tem distribuição de Cauchy centrada na origem. Na secção seguinte vamos mostrar que, sob determinadas condições

$$a_n\left(\widehat{\phi}_{LAD} - \phi\right) \xrightarrow{d} \xi,$$

onde ξ é o mínimo de um certo processo estocástico e a velocidade de convergência é de ordem $n^{1/\alpha}L(n)$.

Vários estudos de simulação para processos AR(p), foram desenvolvidos com o objectivo de se compararem estes dois estimadores. Assim o fizeram Gross e Steiger (1979) e Knight (1987), chegando à conclusão que para $\alpha \in [1,2)$, as estimativas LAD obtidas eram de uma forma geral melhores, isto é, mais próximas dos verdadeiros parâmetros, do que as estimativas LS. Conclusões semelhantes foram obtidas por Bloomfield e Steiger (1983), para $\alpha < 1$. Estes resultados podem ser parcialmente justificados, se pensarmos que perante erros com caudas elevadas, o estimador dos mínimos quadrados ao minimizar a soma dos quadrados das estimativas dos erros, pode centrar-se em alguns erros demasiado grandes e desprezar as restantes observações de erros, fazendo deste modo com que os parâmetros sejam subestimados. Estudos de simulação mais recentes, têm revelado que embora ambos os métodos de estimação, LS e LAD, sejam computacionalmente eficientes e fáceis de implementar, como referem Calder e Davis em [5], o primeiro revela-se inadequado face à estimação de parâmetros de processos ARMA, quando as inovações têm caudas elevadas. Para uma leitura mais cuidada e aprofundada destes assuntos, veja-se por exemplo [9],[10] e [5].

3.3 Estimador LAD

Seja $\{X_t\}$ o processo causal AR(p) definido em (3.2), dirigido pela sequência de erros $\{Z_t\}_{t\in\mathbf{Z}}$, onde $Z_t \in D(\alpha)$, com $\alpha \in (0,2)$, ou seja Z_t satisfaz as condições (2.5) e (2.6) para $0 < \alpha < 2$. Considere-se o polinómio autoregressivo $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p$ e que verifica $\phi(z) \neq 0$, para todo o complexo z, com $|z| \leq 1$. Como o processo é causal, então pode ser representado por um processo linear da forma

$$X_t = \sum_{i=0}^{\infty} \psi_i Z_{t-i}, \tag{3.4}$$

onde $\{\psi_i, i = 0, 1, \ldots\}$ são os coeficientes de z^i na expansão da série de potências de $\phi^{-1}(z)$. Neste caso, tal como vimos no capítulo anterior, o processo $\{X_t\}$ verifica

$$\lim_{x\to\infty} \frac{P[|X_t| > x]}{P[|Z_1| > x]} = \lim_{x\to\infty} \frac{P\left[\sum_{i=0}^{\infty} |\psi_i| \, |Z_{t-i}| > x\right]}{P[|Z_1| > x]} = \sum_{i=0}^{\infty} |\psi_i|^{\alpha} < \infty, \tag{3.5}$$

o que mostra que a distribuição de X_t tem um comportamento assintótico análogo ao da distribuição de Z_t. Além disso $\{X_t\}$ converge absolutamente com probabilidade um. Considere-se ainda a sucessão de constantes positivas $\{a_n\}$, definida por

$$a_n = \inf\left\{x : P[|Z_1| > x] \leq \frac{1}{n}\right\}$$

e que verifica (2.10).

Pelo que vimos na secção anterior uma LAD-estimativa $\widehat{\phi}_{LAD} = \left(\widehat{\phi}_1, ..., \widehat{\phi}_p\right)'$ do vector de parâmetros $\phi = (\phi_1, ..., \phi_p)'$, minimiza a função objectivo

$$\sum_{t=1}^{n} |X_t - \beta_1 X_{t-1} - ... - \beta_p X_{t-p}|, \tag{3.6}$$

ou ainda a função objectivo modificada

$$\sum_{t=1}^{n} \left[|X_t - \beta_1 X_{t-1} - ... - \beta_p X_{t-p}| - |Z_t|\right].$$

Se somarmos e subtraírmos $Z_t = X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p}$ no somatório anterior, este pode ser escrito ainda na forma

$$\sum_{t=1}^{n} \left[|Z_t - a_n(\beta_1 - \phi_1) a_n^{-1} X_{t-1} - ... - a_n (\beta_p - \phi_p) a_n^{-1} X_{t-p}| - |Z_t|\right].$$

Fazendo agora $u_j = a_n (\beta_j - \phi_j)$, para $j = 1, ..., p$, obtemos para $\mathbf{u} = (u_1, ..., u_p)'$ pertencente a $\mathbb{R}^p$

$$W_n(\mathbf{u}) := \sum_{t=1}^{n} \left[|Z_t - u_1 a_n^{-1} X_{t-1} - ... - u_p a_n^{-1} X_{t-p}| - |Z_t|\right]. \tag{3.7}$$

Defina-se

$$\begin{aligned} Y_{nt}(\mathbf{u}) &= a_n^{-1}(u_1 X_{t-1} + ... + u_p X_{t-p}) \\ &= a_n^{-1} Y_t(\mathbf{u}) \end{aligned} \tag{3.8}$$

então o processo em (3.7), pode ser reescrito na forma

$$\begin{aligned} W_n(\mathbf{u}) &= \sum_{t=1}^{n} [|Z_t - Y_{nt}(\mathbf{u})| - |Z_t|] \\ &= \sum_{t=1}^{n} \left[|Z_t - a_n^{-1} Y_t(\mathbf{u})| - |Z_t|\right]. \end{aligned} \tag{3.9}$$

Vemos assim que minimizar (3.6) em relação a β é equivalente a minimizar (3.7) ou (3.9) relativamente a $\mathbf{u}$. Note-se que para cada vector $\mathbf{u} \in \mathbb{R}^p$, a variável $Y_t(\mathbf{u})$ pode ser escrita com base nas variáveis Z_t's por,

$$\begin{aligned} Y_t(\mathbf{u}) &= u_1 X_{t-1} + ... + u_p X_{t-p} \\ &= u_1 \sum_{i=0}^{\infty} \psi_i Z_{t-1-i} + ... + u_p \sum_{i=0}^{\infty} \psi_i Z_{t-p-i} \\ &= \sum_{i=1}^{\infty} c_i(\mathbf{u}) Z_{t-i}, \end{aligned} \tag{3.10}$$

onde a sucessão de constantes $\{c_i(\mathrm{u})\}$ definida por

$$c_i(\mathrm{u}) = \psi_{i-1}u_1 + \ldots + \psi_{i-p}u_p,$$

satisfaz

$$\sum_{i=1}^{\infty} |c_i(\mathrm{u})|^{\delta} < \infty, \qquad \text{para algum } \delta, \text{ tal que } 0 < \delta < \min(\alpha, 1).$$

pPor outro lado o processo linear $Y_t(\cdot)$ sendo dirigido pelas inovações Z_t's, a sua distribuição nas caudas comporta-se também de forma semelhante à de Z_t, tendo por isso caudas elevadas. Um resultado análogo a (3.5), pode então ser deduzido para $Y_t(\cdot)$, tendo-se assim

$$\lim_{x\to\infty} \frac{P[|Y_t| > x]}{P[|Z_1| > x]} = \lim_{x\to\infty} \frac{P[\sum_{i=1}^{\infty} |c_i(\mathrm{u})| |Z_{t-i}| > x]}{P[|Z_1| > x]} = \sum_{i=1}^{\infty} |c_i(\mathrm{u})|^{\alpha} < \infty \qquad (3.11)$$

Por simplificação de escrita e não havendo risco de confusão, optaremos, por vezes, por escrever apenas c_i em vez de $c_i(\mathrm{u})$. Repare-se que W_n é um elemento aleatório ou função aleatória em $(C(\mathbb{R}^p), \rho)$, o espaço das funções contínuas de $\mathbb{R}^p$ em $\mathbb{R}$, munido da métrica do supremo. Isto significa que dado o espaço probabilístico $(\Omega, \mathcal{A}, P)$, W_n é uma aplicação mensurável de Ω em $C(\mathbb{R}^p)$ e por tal, para cada $w \in \Omega$, $W_n(\cdot, w)$ é uma função contínua pertencente a $C(\mathbb{R}^p)$ e para cada $\mathrm{u} \in \mathbb{R}^p$ $W_n(\mathrm{u})$ é uma variável aleatória. Vamos ver que sob determinadas condições de momentos com respeito às inovações $\{Z_t\}$, o processo $W_n(\cdot)$ converge em distribuição em $C(\mathbb{R}^p)$ para o processo $W(\cdot)$ definido para cada $\mathrm{u} \in \mathbb{R}^p$ por

$$\begin{aligned} W(\mathrm{u}) &= \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left[\left|Z_k - (\psi_{i-1}u_1 + \ldots + \psi_{i-p}u_p)\,\delta_k\Gamma_k^{-1/\alpha}\right| - |Z_k|\right] \\ &= \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left[\left|Z_k - c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha}\right| - |Z_k|\right], \end{aligned} \qquad (3.12)$$

o qual por sua vez tem um único mínimo com probabilidade um.

Comecemos então por considerar $\{\delta_k\}$ a sequência de v.a.'s i.i.d. que satisfazem (2.14) e a soma $\Gamma_k = E_1 + \ldots + E_k$, com $\{E_i, i \geq 1\}$ sequência de exponenciais unitárias i.i.d., tal que $\{Z_k\}$, $\{\delta_k\}$ e $\{\Gamma_k\}$ são independentes. No teorema 2.5, estabelecemos que no espaço das medidas pontuais aleatórias definidas em $E = \mathbb{R} \times \overline{\mathbb{R}} \backslash \{0\}$,

$$\zeta_n := \sum_{k=1}^{n} \varepsilon_{(Z_k, a_n^{-1}X_k)} \xrightarrow{d} \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \varepsilon_{\left(Z_k, \delta_k c_i \Gamma_k^{-1/\alpha}\right)} =: \zeta,$$

onde aqui $\{X_k\}$ é o processo linear de médias móveis definido em (2.36) e a sequência de constantes positivas $\{c_i\}$, verifica (2.8). Vimos além disso, como consequência dessa convergência em distribuição, que para qualquer função f contínua em $\mathbb{R} \times \overline{\mathbb{R}} \backslash \{0\}$ de suporte compacto, $\zeta_n(f)$ converge em distribuição para $\zeta(f)$, isto é,

$$\sum_{k=1}^{n} f\left(Z_k, a_n^{-1} X_k\right) \xrightarrow{d} \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} f\left(Z_k, \delta_k c_i \Gamma_k^{-1/\alpha}\right). \tag{3.13}$$

Se tomássemos a v.a. $Y_t(\mathrm{u})$, definida em (3.10) e se fizéssemos

$$f(x,y) = |x-y| - |x|,$$

com $x \in \mathbb{R}$ e $y \in \overline{\mathbb{R}} \backslash \{0\}$, então obteríamos a partir de (3.13)

$$\underbrace{\sum_{t=1}^{n} \left[|Z_t - Y_{nt}(\mathrm{u})| - |Z_t|\right]}_{W_n(\mathrm{u})} \xrightarrow{d} \underbrace{\sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left[\left|Z_k - c_i(\mathrm{u})\, \delta_k \Gamma_k^{-1/\alpha}\right| - |Z_k|\right]}_{W(\mathrm{u})} \tag{3.14}$$

chegando deste modo rapidamente ao resultado que nos propusémos demonstrar. Contudo deparamo-nos com um pequeno problema: é que apesar da função módulo ser contínua, não tem suporte compacto. Mas por outro lado, sendo $f(x,y)$ contínua, o resultado (3.13), será válido para funções do tipo

$$g(x,y) = f(x,y)\, I(|x| < M)\, I(|y| > \delta),$$

onde $g(\cdot)$ tem agora suporte compacto, quaisquer que sejam os reais positivos M e δ. Ora, isto leva-nos a concluir que para provarmos a convergência em (3.14), teremos que recorrer a funções do tipo de $g(\cdot)$ e é precisamente essa a estratégia a seguir. Primeiro vamos ver quais as condições que garantem a existência de $W(\cdot)$, tendo estas naturalmente em conta, como já foi referido, o comportamento da distribuição das Z_t's próximo da origem.

Proposição 3.1 *Considere-se $\{Z_k\}$ a sequência de variáveis aleatórias reais i.i.d. que verificam $Z_k \in D(\alpha)$, apresentando mediana nula se $\alpha \geq 1$. Sejam $\{\delta_k\}$ e $\{\Gamma_k\}$ as sequências de v.a.'s que satisfazem respectivamente (2.14) e (2.15). Admita-se que $\{Z_k\}$, $\{\delta_k\}$ e $\{\Gamma_k\}$ são independentes. Defina-se*

$$W = \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left[\left|Z_k - c_i \delta_k \Gamma_k^{-1/\alpha}\right| - |Z_k|\right],$$

onde $\sum_{i=1}^{\infty} |c_i| < \infty$. Então

a) Se $\alpha < 1$, W é finita com probabilidade 1.

b) Se $\alpha > 1$, W é finita com probabilidade 1 sse $E\left(|Z_1|^{1-\alpha}\right) < \infty$.

c) Se $\alpha = 1$, W é finita com probabilidade 1 sse $-E\left(\ln|Z_1|\right) < \infty$.

Dem. Vamos provar em primeiro lugar que $\sum_{k=1}^{\infty}\left|\Gamma_k^{-1/\alpha} - k^{-1/\alpha}\right| < \infty$ *q.c.*, para qualquer $\alpha \in (0,2)$.

Seja $f(x) = x^{-1/\alpha}$. Expandindo esta função em torno do ponto k, vem

$$f(x) = k^{-1/\alpha} + (x-k)\left(\frac{-1}{\alpha}\right)k^{-1/\alpha-1} + O(x-k)^2,$$

e consequentemente

$$x^{-1/\alpha} - k^{-1/\alpha} = (x-k)\left(\frac{-1}{\alpha}\right)k^{-1/\alpha-1} + o_p(1). \tag{3.15}$$

Por outro lado, sendo $\Gamma_k = \sum_{i=1}^{k} E_i$, a lei dos logarítmos iterados[1] garante que

$$\limsup_{n\to\infty}\frac{\Gamma_k - k}{\sqrt{2k\log\log k}} = 1 \ q.c. \tag{3.16}$$

Resulta então de (3.15) e de (3.16) que

$$\begin{aligned}
\Gamma_k^{-1/\alpha} - k^{-1/\alpha} &= \frac{-1}{\alpha}\sqrt{2}\left(\log\log k\right)^{1/2} k^{1/2}k^{(-1/\alpha)-1}\left(1 + o_p(1)\right) \\
&= \frac{-1}{\alpha}\sqrt{2}k^{-1/2-1/\alpha}\left(\log\log k\right)^{1/2}\left(1 + o_p(1)\right) \\
&\Leftrightarrow \frac{\Gamma_k^{-1/\alpha} - k^{-1/\alpha}}{-1/\alpha\sqrt{2}k^{-1/2-1/\alpha}\left(\log\log k\right)^{1/2}} = 1 + o_p(1),
\end{aligned}$$

quase certamente. Consequentemente

$$\left|\Gamma_k^{-1/\alpha} - k^{-1/\alpha}\right| = O\left(k^{-1/2-1/\alpha}\left(\log\log k\right)^{1/2}\right) \quad \text{q.c.},$$

donde,

$$\sum_{k=1}^{\infty}\left|\Gamma_k^{-1/\alpha}(w) - k^{-1/\alpha}\right| \leq \sum_{k=1}^{\infty}\frac{\left(\log\log k\right)^{1/2}}{k^{\beta}}, \tag{3.17}$$

para quase todo o $w \in \Omega$ e para $\beta = 1/\alpha + 1/2 > 1$. Notando agora que

$$\sqrt{\log\log x} \leq \sqrt{\log x} \leq \log x,$$

[1] Veja-se em apêndice.

pois $\log x \leq x, \forall x \geq 1$ e tendo em conta que

$$\int_1^\infty \left(x^{-\beta} \log x\right) dx < \infty,$$

segue-se que, aplicando o critério do integral impróprio à série em (3.17),

$$\sum_{k=1}^{\infty} \left|\Gamma_k^{-1/\alpha} - k^{-1/\alpha}\right| < \infty \quad \text{q.c.},$$

para $0 < \alpha < 2$. Observe-se agora que

$$|a - b| - |a| \leq (|a - c| - |a|) + |b - c|, \forall a, b, c$$

e aplicando esta desigualdade a W temos

$$\begin{aligned} W &\leq \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left[\left|Z_k - c_i\delta_k k^{-1/\alpha}\right| - |Z_k|\right] + \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left|c_i\delta_k\Gamma_k^{-1/\alpha} - c_i\delta_k k^{-1/\alpha}\right| \\ &= \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left[\left|Z_k - c_i\delta_k k^{-1/\alpha}\right| - |Z_k|\right] + \sum_{k=1}^{\infty} \left|\Gamma_k^{-1/\alpha} - k^{-1/\alpha}\right| \sum_{i=1}^{\infty} |c_i| \\ &= W_1 + W_2. \end{aligned}$$

Como $W_2 = \sum_{k\geq 1} \left|\Gamma_k^{-1/\alpha} - k^{-1/\alpha}\right| \sum_{i\geq 1} |c_i| < \infty$, então a convergência de W só depende agora de W_1, sendo necessário considerar 3 situações. Primeito vamos admitir que $0 < \alpha < 1$. Ora,

$$\begin{aligned} W_1 &= \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left[\left|Z_k - c_i\delta_k k^{-1/\alpha}\right| - |Z_k|\right] \leq \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left|c_i\delta_k k^{-1/\alpha}\right| \\ &= \sum_{k=1}^{\infty} k^{-1/\alpha} \sum_{i=1}^{\infty} |c_i| < \infty, \end{aligned}$$

donde $W < \infty$, quase certamente. Seja agora $1 \leq \alpha < 2$. Aplicando a igualdade

$$|x - y| - |x| = y\left(I(x < 0) - I(x > 0)\right) + 2(y - x)\left(I(y > x > 0) - I(y < x < 0)\right),$$

válida para qualquer $x \neq 0$, vem que

$$\begin{aligned} W_1 &= \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left[c_i\delta_k k^{-1/\alpha} \left(I(Z_k > 0) - I(Z_k < 0)\right)\right] \\ &+ 2\sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left[\left(c_i\delta_k k^{-1/\alpha} - Z_k\right) \times \right. \\ &\times \left. \left[I\left(c_i\delta_k k^{-1/\alpha} > Z_k > 0\right) - I\left(c_i\delta_k k^{-1/\alpha} < Z_k < 0\right)\right]\right] \end{aligned} \tag{3.18}$$

Analisemos se o primeiro somatório é finito, começando por calcular o seguinte valor esperado

$$E\left[c_i\delta_k k^{-1/\alpha}\left(I\left(Z_k>0\right)-I\left(Z_k<0\right)\right)\right] = c_i k^{-1/\alpha}E\left[\delta_k\left(I\left(Z_k>0\right)-I\left(Z_k<0\right)\right)\right] = 0,$$

pois $\{\delta_k\}$ e $\{(I(Z_k>0)-I(Z_k<0))\}$ não são correlacionados e

$$E\left[I\left(Z_k>0\right)-I\left(Z_k<0\right)\right]=0,$$

já que por hipótese $\{Z_k\}$ tem mediana nula. Então,

$$\begin{aligned}
&\sum_{k=1}^{\infty}\sum_{i=1}^{\infty} var\left[c_i\delta_k k^{-1/\alpha}\left(I\left(Z_k>0\right)-I\left(Z_k<0\right)\right)\right]\\
=\;&\sum_{k=1}^{\infty}\sum_{i=1}^{\infty} E\left[c_i^2\delta_k^2 k^{-2/\alpha}\left(I\left(Z_k>0\right)-I\left(Z_k<0\right)\right)^2\right]\\
\leq\;&\sum_{k=1}^{\infty}\sum_{i=1}^{\infty} E\left(c_i^2\delta_k^2 k^{-2/\alpha}\right)=\sum_{k=1}^{\infty}k^{-2/\alpha}\sum_{i=1}^{\infty}c_i^2<\infty,
\end{aligned}$$

pois $\sum_{k\geq 1}k^{-2/\alpha}<\infty$ para $\alpha\in[1,2)$ e se $\sum_{i\geq 1}|c_i|<\infty$, tem-se a partir de certa ordem que $|c_i|<1$, o que implica que $c_i^2<|c_i|$, vindo portanto $\sum_{i\geq 1}c_i^2<\infty$. Logo

$$\sum_{k=1}^{\infty}\sum_{i=1}^{\infty} c_i\delta_k k^{-1/\alpha}\left(I\left(Z_k>0\right)-I\left(Z_k<0\right)\right)<\infty \quad q.c.$$

Estudemos agora a convergência quase certa do segundo somatório em (3.18).

$$\begin{aligned}
&\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left(c_i\delta_k k^{-1/\alpha}-Z_k\right)\left[I\left(c_i\delta_k k^{-1/\alpha}>Z_k>0\right)-I\left(c_i\delta_k k^{-1/\alpha}<Z_k<0\right)\right]\\
=\;&\underbrace{\sum_{i=1}^{\infty}\sum_{k=1}^{\infty}\left(c_i\delta_k k^{-1/\alpha}-Z_k\right)I\left(c_i\delta_k k^{-1/\alpha}>Z_k>0\right)}_{(1)}\\
&\underbrace{-\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left(c_i\delta_k k^{-1/\alpha}-Z_k\right)I\left(c_i\delta_k k^{-1/\alpha}<Z_k<0\right).}_{(2)}
\end{aligned}$$

Para garantir a convergência de (1), é suficiente provar a convergência de

$$\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left(|c_i|\,k^{-1/\alpha}-Z_k\right)I\left(|c_i|\,k^{-1/\alpha}>Z_k>0\right) \tag{3.19}$$

e neste caso como temos uma soma infinita de variáveis aleatórias não negativas, a convergência dá-se desde que o valor esperado dessa soma seja finito. Sendo $F(\cdot)$ a função de distribuição de Z_k, vem

$$
\begin{aligned}
& E\left[\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left(|c_i|\,k^{-1/\alpha}-Z_k\right)I\left(|c_i|\,k^{-1/\alpha}>Z_k>0\right)\right] \\
= & \sum_{k=1}^{\infty}\sum_{i=1}^{\infty}E\left[\left(|c_i|\,k^{-1/\alpha}-Z_k\right)I\left(|c_i|\,k^{-1/\alpha}>Z_k>0\right)\right] \\
= & \sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\int_0^{|c_i|k^{-1/\alpha}}\left(|c_i|\,k^{-1/\alpha}-z\right)F(dz) \\
\sim & \sum_{i=1}^{\infty}\int_1^{\infty}\int_0^{|c_i|k^{-1/\alpha}}\left(|c_i|\,k^{-1/\alpha}-z\right)F(dz)\,dk.
\end{aligned}
$$

Tomando $u=|c_i|\,k^{-1/\alpha}$ e tendo em conta que $dk=(-\alpha)\,|c_i|^{\alpha}\,u^{-\alpha-1}du$ e que $0<z<u<|c_i|$, a expressão anterior toma a forma

$$
\begin{aligned}
& \sum_{i=1}^{\infty}\int_{|c_i|}^{0}\int_{z}^{|c_i|}(u-z)\,(-\alpha)\,|c_i|^{\alpha}\,u^{-\alpha-1}du F(dz) \\
= & \alpha\sum_{i=1}^{\infty}|c_i|^{\alpha}\int_0^{|c_i|}\int_z^{|c_i|}(u-z)\,u^{-\alpha-1}du F(dz) \\
= & \alpha\sum_{i=1}^{\infty}|c_i|^{\alpha}\int_0^{|c_i|}\int_z^{|c_i|}\left(u^{-\alpha}-zu^{-\alpha-1}\right)du F(dz). \qquad (3.20)
\end{aligned}
$$

Para $1<\alpha<2$, obtém-se integrando em (3.20)

$$
\begin{aligned}
& \alpha\sum_{i=1}^{\infty}|c_i|^{\alpha}\int_0^{|c_i|}\left[\frac{u^{-\alpha+1}}{-\alpha+1}-z.\frac{u^{-\alpha}}{-\alpha}\right]_z^{|c_i|}F(dz) \\
= & \alpha\sum_{i=1}^{\infty}|c_i|^{\alpha}\int_0^{|c_i|}\left[\frac{|c_i|^{1-\alpha}}{1-\alpha}-z\frac{|c_i|^{-\alpha}}{-\alpha}+\left(\frac{1}{\alpha-1}-\frac{1}{\alpha}\right)z^{1-\alpha}\right]F(dz).
\end{aligned}
$$

Como $E(|Z_1|)<\infty$ então $\int_0^{|c_i|}zF(dz)<\infty$, o que significa que a convergência do integral anterior irá depender exclusivamente do integral $\int_0^{|c_i|}z^{1-\alpha}F(dz)$ ser finito. Este, por sua vez, prova-se[2] que é finito se

$$E\left(|Z_1|^{1-\alpha}\right) < \infty.$$

Consideremos agora $\alpha = 1$ em (3.20). Vem então que

$$\begin{aligned}
&\sum_{i=1}^{\infty} |c_i| \int_0^{|c_i|} \int_z^{|c_i|} \left(\frac{1}{u} - zu^{-2}\right) du F(dz) \\
&= \sum_{i=1}^{\infty} |c_i| \int_0^{|c_i|} \left(\ln|c_i| + z\frac{1}{|c_i|} - \ln|z| - 1\right) F(dz), \qquad (3.21)
\end{aligned}$$

o que mostra que a natureza do integral em (3.21) é equivalente à dos integrais $\int_0^{|c_i|} zF(dz)$ e $\int_0^{|c_i|} \log|z|\, F(dz)$. Quanto ao primeiro é fácil de ver que

$$\begin{aligned}
\int_0^{|c_i|} zF(dz) &= \left[z\left(1 - F(z)\right)\right]_0^{|c_i|} + \int_0^{|c_i|} \frac{z^2}{2}\left(1 - F(z)\right) dz \\
&< C + \int_0^{|c_i|} \frac{z^2}{2} dz < \infty.
\end{aligned}$$

Logo o integral em (3.21) é finito sse também o for $\int_0^{|c_i|} \ln|z|\, F(dz)$ e este por sua vez, também é finito[3] se

$$-E\left(\ln|Z_1|\right) < \infty.$$

Analisando agora de forma semelhante a convergência de $\int_{-|c_i|}^{0} |z|^{1-\alpha} F(dz)$ e $\int_{-|c_i|}^{0} \ln|z|\, F(dz)$, concluímos assim que a convergência absoluta de (1) ocorre com probabilidade um sse

$$\begin{cases} E\left(|Z_1|^{1-\alpha}\right) < \infty & \text{se } 1 < \alpha < 2 \\ -E\left(\ln|Z_1|\right) < \infty & \text{se } \alpha = 1 \end{cases}.$$

[2] Veja-se em apêndice.
[3] Veja-se em apêndice.

Falta analisar por último a existência de (2). Observe-se que

$$
\begin{aligned}
& -\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left(c_i\delta_k k^{-1/\alpha} - Z_k\right) I\left(c_i\delta_k k^{-1/\alpha} < Z_k < 0\right) \\
= & \sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left(-c_i\delta_k k^{-1/\alpha} + Z_k\right) I\left(-c_i\delta_k k^{-1/\alpha} > -Z_k > 0\right) \\
= & \sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left(\left|c_i\right| k^{-1/\alpha} - (-Z_k)\right) I\left(\left|c_i\right| k^{-1/\alpha} > -Z_k > 0\right).
\end{aligned}
$$

Tomando a nova variável $Z_k^* = -Z_k$, chegamos à expressão (1), pelo que se demonstra o resultado pretendido de forma análoga. ∎

Como corolário desta proposição, resulta que o processo $W(\cdot)$ existe e está bem definido quase certamente, como elemento aleatório do espaço $\left(C\left(\mathbb{R}^p\right), \rho\right)$. Na próxima proposição são estabelecidos alguns resultados assintóticos relativos a um processo linear de médias móveis estacionário, pondo-se em relevo as propriedades de variação regular nas caudas da respectiva função de distribuição.

Proposição 3.2 *Seja $\{Y_t\}$ um processo linear de médias móveis estritamente estacionário, gerado pela sequência de v.a.'s i.i.d. $\{Z_t\}$ e que verificam (2.5) e (2.6), para $\alpha > 0$. Para cada $n \in \mathbb{N}$ considere-se a sequência $\{Y_{nt}\}$, com $Y_{nt} = a_n^{-1} Y_t$. Seja ainda $\{V_t\}$ uma sequência de v.a.'s i.i.d com média finita tal que para todo t, V_t e Y_t são independentes. Então para todo o $\delta > 0$ e $\eta > 0$*

a) $\limsup_{n\to\infty} P\left[\sum_{t=1}^{n} |V_t| \, |Y_{nt}|^\gamma I\left(|Y_{nt}| \leq \delta\right) > \eta\right] \leq \eta^{-1} C_1 E\left(|V_1|\right) \delta^{\gamma-\alpha}$, *para qualquer $\gamma > \alpha$.*

b) $\limsup_{n\to\infty} P\left[\sum_{t=1}^{n} |V_t| \, |Y_{nt}|^\gamma I\left(|Y_{nt}| > \delta\right) > \eta\right] \leq C_2 \delta^{-\alpha} P[|V_1| > 0]$, *para qualquer $\gamma > 0$.*

c) Se V_1 tem média nula e variância σ^2 e $1 \leq \alpha < 2$ então

$$
\begin{aligned}
var\left(\sum_{t=1}^{n} V_t Y_{nt} I\left(|Y_{nt}| \leq \delta\right)\right) &= n a_n^{-2} E\left[Y_1^2 I\left(|Y_1| \leq a_n \delta\right)\right] E\left(V_1^2\right) \\
&\sim C_3 E\left(V_1^2\right) \delta^{2-\alpha},
\end{aligned}
$$

quando $n \to \infty$, onde $C_1, C_2, C_3 \in \mathbb{R}$.

Dem. a) Seja $\delta > 0$. Para todo $\eta > 0$, vem pela desigualdade de Markov

$$
\begin{aligned}
& P\left[\sum_{t=1}^{n} |V_t| \, |Y_{nt}|^\gamma I\left(|Y_{nt}| \le \delta\right) > \eta\right] \\
\le\ & \eta^{-1} E\left[\sum_{t=1}^{n} |V_t| \left|a_n^{-1} Y_t\right|^\gamma I\left(|Y_t| \le a_n\delta\right)\right] \\
=\ & \eta^{-1} E\left(V_1\right) n a_n^{-\gamma} E\left[|Y_1|^\gamma I\left(|Y_1| \le a_n\delta\right)\right], \ V_t, Y_t \text{ ind.; estacion. } \{Y_t\} \\
=\ & \eta^{-1} E\left(V_1\right) n a_n^{-\gamma} E\left[|Y_1|^\gamma I\left(|Y_1|^\gamma \le (a_n\delta)^\gamma\right)\right] \\
=\ & \eta^{-1} E\left(V_1\right) n a_n^{-\gamma} \int_0^{(a_n\delta)^\gamma} y G(dy), \text{ onde G é a função de dist. de } |Y_1|^\gamma \\
=\ & \eta^{-1} E(V_1) n a_n^{-\gamma} \left(\left[-y\left(1 - G(y)\right)\right]_0^{(a_n\delta)^\gamma} + \int_0^{(a_n\delta)^\gamma} \left(1 - G(y)\right) dy \right) \\
=\ & \eta^{-1} E(V_1) \delta^\gamma \underbrace{n(1 - G(a_n\delta)^\gamma)}_{P[|Y_1|^\gamma > (a_n\delta)^\gamma]} \left(-1 + \frac{\int_0^{(a_n\delta)^\gamma} \left(1 - G(y)\right) dy}{(a_n\delta)^\gamma \left(1 - G(a_n\delta)^\gamma\right)} \right). \qquad (3.22)
\end{aligned}
$$

Repare-se agora que por (3.11) tem-se

$$\lim_{n\to\infty} nP\left[|Y_1| > a_n x\right] = \lim_{n\to\infty} nP\left[|Y_1|^\gamma > (a_n x)^\gamma\right] \sim \sum_{i=1}^{\infty} |c_i|^\alpha x^{-\alpha} = Cx^{-\alpha},$$

para qualquer $x > 0$, com

$$C = \sum_{i=1}^{\infty} |c_i|^\alpha < \infty.$$

Além disso como $P\left[|Y_1| > x\right] \in RV_{-\alpha}$, então $P\left[|Y_1|^\gamma > x\right] = P\left[|Y_1| > x^{1/\gamma}\right]$ é assintoticamente $x^{-\alpha/\gamma} L(x)$, ou seja $P\left[|Y_1|^\gamma > x\right] \in RV_{-\alpha/\gamma}$. Ora, como $\gamma > \alpha$, vem que $-\frac{\alpha}{\gamma} > -1$. Então tomando o limite em n e aplicando o teorema de Karamata (teorema 2.1) em (3.22) obtemos

$$C\eta^{-1} E(V_1) \delta^\gamma \delta^{-\alpha} \left(-1 + \frac{1}{-\alpha/\gamma + 1}\right) = \eta^{-1} C_1 E\left(|V_1|\right) \delta^{\gamma-\alpha},$$

com $C_1 \in \mathbb{R}$ e por conseguinte

$$\limsup_{n\to\infty} P\left[\sum_{t=1}^{n} |V_t| \, |Y_{nt}|^\gamma I\left(|Y_{nt}| \le \delta\right) > \eta\right] \le \eta^{-1} C_1 E\left(|V_1|\right) \delta^{\gamma-\alpha},$$

para qualquer para $\gamma > \alpha$.

b) Dados $\delta > 0$ e $\eta > 0$, tem-se agora

$$\begin{aligned} P\left[\sum_{t=1}^{n} |V_t| \, |Y_{nt}|^{\gamma} I\left(|Y_{nt}| > \delta\right) > \eta\right] &= P\left[\sum_{t=1}^{n} |V_t| \left|a_n^{-1} Y_t\right|^{\gamma} I\left(|Y_t| > a_n\delta\right)\right] \\ &\leq P\left[\bigcup_{t=1}^{n} \left(\{|Y_t| > a_n\delta)\} \cap \{|V_t| > 0\}\right)\right] \\ &\leq nP\left[|Y_1| > a_n\delta\right] P\left[|V_1| > 0\right]. \end{aligned}$$

Tomando $n \longrightarrow \infty$, obtemos

$$\begin{aligned} \limsup_{n\longrightarrow\infty} P\left[\sum_{t=1}^{n} |V_t| \, |Y_{nt}|^{\gamma} I\left(|Y_{nt}| > \delta\right) > \eta\right] &\leq P\left[|V_1| > 0\right] \sum_{i=1}^{\infty} |c_i|^{\alpha} \delta^{-\alpha} \\ &= C_2 P\left[|V_1| > 0\right] \delta^{-\alpha}, \end{aligned}$$

com $C_2 > 0$ e para todo $\gamma > 0$.

c) Suponha-se que $E(V_1) = 0$, $var(V_1) = \sigma^2$. Uma vez que V_t e Y_t são independentes, então para quaisquer $\delta > 0$ e $\eta > 0$, vem

$$E\left(V_t Y_{nt} I\left(|Y_{nt}| \leq \delta\right)\right) = a_n^{-1} E\left(V_t\right) E\left(Y_t I\left(|Y_t| \leq a_n\delta\right)\right) = 0,$$

logo a variância vem dada por

$$\begin{aligned} & var\left(\sum_{t=1}^{n} V_t Y_{nt} I\left(|Y_{nt}| \leq \delta\right)\right) \\ &= \sum_{t=1}^{n} var\left(V_t Y_{nt} I\left(|Y_{nt}| \leq \delta\right)\right) \\ &= \sum_{t=1}^{n} a_n^{-2} E\left(V_t^2\right) E\left(Y_t^2 I\left(|Y_t| \leq a_n\delta\right)\right) \\ &= n a_n^{-2} E\left(V_1^2\right) E\left(Y_1^2 I\left(|Y_1|^2 \leq (a_n\delta)^2\right)\right) \\ &= n a_n^{-2} E\left(V_1^2\right) \int_0^{(a_n\delta)^2} y G(dy), \text{ com G função de dist, de } |Y_1|^2 \\ &= n a_n^{-2} E\left(V_1^2\right) \left(\left[-y\left(1 - G(y)\right)\right]_0^{(a_n\delta)^2} + \int_0^{(a_n\delta)^2} \left(1 - G(y)\right) dy\right) \\ &= E\left(V_1^2\right) \delta^2 \underbrace{n\left(1 - G(a_n\delta)^2\right)}_{P\left[|Y_1|^2 > (a_n\delta)^2\right]} \left(-1 + \frac{\int_0^{(a_n\delta)^2} \left(1 - G(y)\right) dy}{(a_n\delta)^2 \left(1 - G(a_n\delta)^2\right)}\right). \end{aligned}$$

Notando que $P\left[|Y_1|^2 > x\right] \in RV_{-\alpha/2}$, então, para $1 \leq \alpha < 2$, tomando $n \longrightarrow \infty$ e aplicando novamente o teorema de Karamata, a expressão acima tem por limite em n

$$E\left(V_1^2\right)C\delta^{2-\alpha}\left(-1+\frac{1}{-\alpha/2+1}\right) = C_3E\left(V_1^2\right)\delta^{2-\alpha}, \text{ com } C_3 \in \mathbb{R}$$

obtendo-se assim o resultado. ∎

Finalmente temos criadas as condições para podermos estabelecer a convergência em distribuição do processo $W_n(\cdot)$ para o processo $W(\cdot)$. Relembrando o exemplo 1.6 do capítulo 1 e uma vez que estamos a trabalhar com elementos do espaço $C(\mathbb{R}^p)$, para provarmos este resultado teremos que mostrar que as distribuições conjuntas finitas de W_n convergem em distribuição para as de W e além disso que $\{W_n\}$ é tight, ou de forma equivalente, que a sequência$\{W_n\}$ seja relativamente compacta. Ou seja, a convergência em distribuição ocorre sse para todo o $k \in \mathbb{N}$

$$\left(W_n(\mathbf{u}_1),...,W_n(\mathbf{u}_k)\right) \xrightarrow{d} \left(W(\mathbf{u}_1),...,W(\mathbf{u}_k)\right), \forall \mathbf{u}_i \in \mathbb{R}^p, 1 \leq i \leq k \quad (3.23)$$

e se $\{W_n\}$ é tight. Recorde que $\{W_n\}$ é relativamente compacta sse para toda a subsequência $\{W_{n'}\} \subset \{W_n\}$, existe uma subsequência $\{W_{n''}\} \subset \{W_{n'}\}$, tal que $W_{n''}$ converge em distribuição para algum elemento aleatório de $C(\mathbb{R}^p)$, podendo esse elemento aleatório ser diferente de subsequência para subsequência. Esta propriedade de relatividade compacta do processo W_n, é necessária, pois a convergência em distribuição em (3.23), é estabelecida apenas em termos de vectores aleatórios de $\mathbb{R}^k$, não sendo por isso, suficiente para garantir a convergência em distribuição do processo W_n para o processo W, enquanto elementos aleatórios do espaço das funções contínuas de $\mathbb{R}^p$ em $\mathbb{R}$. É combinando precisamente estes dois ingredientes, convergência em distribuição das conjuntas finitas e relatividade compacta , que se garante que $W_n \xrightarrow{d} W$. Contudo, o conhecimento prévio da existência do processo limite W, no sentido em que ele está bem definido, é finito (quase certamente) e é único, prescinde naturalmente da demonstração da relatividade compacta ou tightness de W_n, uma vez que verificando-se (3.23), o processo W_n terá necessariamente de convergir em distribuição para W. Este é o procedimento que nos vai orientar para demonstrar o próximo teorema.

Teorema 3.1 *Seja $\{X_t\}$ um processo autoregressivo $AR(p)$ em que as inovações satisfazem $\{Z_t\} \in D(\alpha)$, apresentando mediana nula se $\alpha \geq 1$. Se*

(i) $\alpha < 1$ *ou*

(ii) $\alpha > 1$ *e* $E\left(|Z_1|^\beta\right) < \infty$, *para algum* $\beta < 1 - \alpha$ *ou*

(iii) $\alpha = 1$ *e* $-E(\ln|Z_1|) < \infty$,

então

$$W_n(\cdot) \xrightarrow{d} W(\cdot),$$

onde $W_n(\cdot)$ e $W(\cdot)$ são os processos definidos em (3.9) e (3.12) respectivamente.

Dem. Vamos analisar inicialmente o caso unidimensional, isto é, vamos ver primeiro que $W_n(\mathbf{u}) \xrightarrow{d} W(\mathbf{u})$ para $\mathbf{u} \in \mathbb{R}^p$. A situação multidimensional resulta por aplicação do teorema de *Cramer-Wold Device*[4].

(i) Admita-se então que $0 < \alpha < 1$. Note-se em primeiro lugar que para quaisquer reais positivos M e δ, W_n pode ser escrito na forma

$$\begin{aligned} W_n(\mathbf{u}) &= \underbrace{\sum_{t=1}^{n} [|Z_t - Y_{nt}(\mathbf{u})| - |Z_t|]\, I(|Z_t| \le M)\, I(|Y_{nt}(\mathbf{u})| > \delta)}_{W^1_{n,\delta,M}(\mathbf{u})} \\ &\quad \underbrace{+\sum_{t=1}^{n} [|Z_t - Y_{nt}(\mathbf{u})| - |Z_t|]\, (1 - I(|Z_t| \le M)\, I(|Y_{nt}(\mathbf{u})| > \delta)),}_{W^2_{n,\delta,M}(\mathbf{u})} \end{aligned}$$

i.e., $W_n(\mathbf{u}) - W^1_{n,\delta,M}(\mathbf{u}) = W^2_{n,\delta,M}(\mathbf{u})$. Por outro lado, por (3.13) sabemos que para todo $\alpha > 0$ e tomando $n \to \infty$

$$W^1_{n,\delta,M}(\mathbf{u}) \xrightarrow{d} \underbrace{\sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \left[\left|Z_k - c_i(\mathbf{u})\,\delta_k \Gamma_k^{-1/\alpha}\right| - |Z_k|\right] I(|Z_k| \le M)\, I\left(|c_i(\mathbf{u})|\, \Gamma_k^{-1/\alpha} > \delta\right)}_{W^1_{\delta,M}(\mathbf{u})}$$

Se provarmos que $W^1_{\delta,M}(\mathbf{u}) \xrightarrow{d} W(\mathbf{u})$, quando $M \to \infty$ e $\delta \to 0$, e que para todo o $\eta > 0$

$$\lim_{\delta \to 0} \lim_{M \to \infty} \limsup_{n \to \infty} P\left[\left|W_n(\mathbf{u}) - W^1_{n,\delta,M}(\mathbf{u})\right| > \eta\right] = 0, \tag{3.24}$$

então pelo Teorema 4.10 em Billingsley (1967), tal como refere Davis, R.A. e outros em [9], poderemos concluir que $W_n(\mathbf{u}) \xrightarrow{d} W(\mathbf{u})$. Convém realçar que apesar dos esforços efectuados na tentativa de se encontrar o referido teorema, tal não foi possível pelo total desconhecimento do livro (edição de 67) a que pertence o teorema em causa. Contudo é fácil de ver que se trata de uma variante do teorema 4.2. em Billingsley (1968), já apresentado neste trabalho, pelo lema 2.2, com a particularidade de que agora é necessário tomar-se $\delta \to 0$. Comecemos

[4] Veja-se em apêndice.

então por mostrar a convergência de $W^1_{\delta,M}(\mathrm{u})$ para $W(\mathrm{u})$. Neste caso tem-se pela desigualdade de Markov e para todo $\varepsilon > 0$,

$$P\left[\left|W^1_{\delta,M}(\mathrm{u}) - W(\mathrm{u})\right| > \varepsilon\right] \leq \varepsilon^{-1} E\left(\left|W^1_{\delta,M}(\mathrm{u}) - W(\mathrm{u})\right|\right). \qquad (3.25)$$

Como $\left|W^1_{\delta,M}(\mathrm{u}) - W(\mathrm{u})\right|$ é dado por

$$\left|\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left[\left|Z_k - c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha}\right| - |Z_k|\right]\left(1 - I(|Z_k| \leq M)\, I\left(|c_i(\mathrm{u})|\,\Gamma_k^{-1/\alpha} > \delta\right)\right)\right|,$$

e atendendo a que $||a| - |b|| \leq |a - b|, \forall a, b$, o valor esperado em (3.25) verifica

$$\begin{aligned} E\left(\left|W^1_{\delta,M}(\mathrm{u}) - W(\mathrm{u})\right|\right) &\leq E\left(\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}|c_i(\mathrm{u})|\,\Gamma_k^{-1/\alpha}\right) \\ &= \sum_{i=1}^{\infty}|c_i(\mathrm{u})|\, E\left(\sum_{k=1}^{\infty}\Gamma_k^{-1/\alpha}\right) < \infty, \end{aligned}$$

pois $\sum_{i=1}^{\infty}|c_i(\mathrm{u})| < \infty$ e o teorema de Kolmogorov[5] garante que a série $\sum_{k=1}^{\infty}\Gamma_k^{-1/\alpha}$ é um $O\left(k^{-1/\alpha}\right)$, já que se $(\Gamma_k - k)/k \to 1$ q.c., então também $\Gamma_k^{-1/\alpha}/k^{-1/\alpha} \to 1$ q.c. Como para $\alpha \in (0,1)$, $\sum_{k=1}^{\infty} k^{-1/\alpha} < \infty$ então tem-se consequentemente que $E\left(\sum_{k=1}^{\infty}\Gamma_k^{-1/\alpha}\right) < \infty$. Existindo o valor esperado e tomando primeiro $M \to \infty$ e depois $\delta \to 0$, vem que $E\left(\left|W^1_{\delta,M}(\mathrm{u}) - W(\mathrm{u})\right|\right) \to 0$, donde $P\left[\left|W^1_{\delta,M}(\mathrm{u}) - W(\mathrm{u})\right| > \varepsilon\right] \to 0$, ou seja, $W^1_{\delta,M}(\mathrm{u}) \xrightarrow{P} W(\mathrm{u})$, implicando deste modo que $W^1_{\delta,M}(\mathrm{u}) \xrightarrow{d} W(\mathrm{u})$. Note-se que a convergência do processo $W^1_{\delta,M}(\mathrm{u})$ para $W(\mathrm{u})$ é inclusive mais forte do que a convergência em probabilidade, tratando-se na verdade de uma convergência quase certa. Isto porque, como $I(|Z_k| \leq M) \xrightarrow{q.c.} 1$, quando $M \to \infty$ e como $I\left(|c_i(\mathrm{u})|\,\Gamma_k^{-1/\alpha} > \delta\right) \xrightarrow{q.c.} 1$ quando $\delta \to 0$, a convergência quase certa de $W^1_{\delta,M}(\mathrm{u})$ para $W(\mathrm{u})$ depende unicamente de se poder permutar o limite com os somatórios. Tal será válido se existir uma série de v.a.'s absolutamente convergente, que não dependenda de M e de δ e que majore a série dos módulos de $W^1_{\delta,M}(\mathrm{u})$, o que na verdade se verifica pois

$$\left|W^1_{\delta,M}(\mathrm{u})\right| \leq \sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left|\left|Z_k - c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha}\right| - |Z_k|\right| \leq \sum_{k=1}^{\infty}\sum_{i=1}^{\infty}|c_i(\mathrm{u})|\,\Gamma_k^{-1/\alpha} < \infty.$$

Estabelecida a convergência de $W^1_{\delta,M}(\mathrm{u})$ para $W(\mathrm{u})$, verifiquemos a validade de (3.24), ou seja que para todo o $\eta > 0$

$$\lim_{\delta\to 0}\lim_{M\to\infty}\limsup_{n\to\infty} P\left[\left|W_n(\mathrm{u}) - W^1_{n,\delta,M}(\mathrm{u})\right| > \eta\right] = 0. \qquad (3.26)$$

[5] Veja-se em apêndice.

Mostrar que este limite é nulo resume-se a provar que $W^2_{n,\delta,M}(\mathbf{u}) \xrightarrow{p} 0$, quando $n \to \infty, M \to \infty$ e depois $\delta \to 0$. Ora,

$$\begin{aligned}
\left|W^2_{n,\delta,M}(\mathbf{u})\right| &= \left|\sum_{t=1}^{n}\left[|Z_t - Y_{nt}(\mathbf{u})| - |Z_t|\right](1 - I(|Z_t| \le M) I(|Y_{nt}(\mathbf{u})| > \delta))\right| \\
&\le \sum_{t=1}^{n}|Y_{nt}(\mathbf{u})|\,|(1 - I(|Z_t| \le M) I(|Y_{nt}(\mathbf{u})| > \delta))| \\
&\le \sum_{t=1}^{n}|Y_{nt}(\mathbf{u})|\, I(|Y_{nt}(\mathbf{u})| \le \delta) + \\
&\qquad + \sum_{t=1}^{n}|Y_{nt}(\mathbf{u})|\, I(|Z_t| > M) I(|Y_{nt}(\mathbf{u})| > \delta),
\end{aligned}$$

em que na última desigualdade teve-se em conta que $1 - I_A I_B = 1 - I_{A\cap B} = I_{\overline{A\cap B}} = I_{\overline{A}\cup\overline{B}} = I_{\overline{B}} + I_{\overline{A}\cap B}$. Pela proposição 3.2 a), temos que

$$\limsup_{n\to\infty} P\left[\sum_{t=1}^{n}|Y_{nt}(\mathbf{u})|\, I(|Y_{nt}(\mathbf{u})| \le \delta) > \eta\right] \le \eta^{-1} C_1 \delta^{1-\alpha} \to 0 \tag{3.27}$$

quando $\delta \to 0$. Por outro lado, fazendo $V_t = I(|Z_t| > M)$ na mesma proposição alínea b), vem

$$\begin{aligned}
&\limsup_{n\to\infty} P\left[\sum_{t=1}^{n}|Y_{nt}(\mathbf{u})|\, I(|Z_t| > M) I(|Y_{nt}(\mathbf{u})| > \delta) > \eta\right] \\
\le\ & C_2 P[I(|Z_1| > M > 0) > 0]\,\delta^{-\alpha} \to 0,
\end{aligned} \tag{3.28}$$

quando $M \to \infty$. Combinando (3.27) e (3.28), tem-se que $\left|W^2_{n,\delta,M}(\mathbf{u})\right| \xrightarrow{p} 0$, sempre que $n \to \infty, M \to \infty$ e $\delta \to 0$ e uma vez que $W^1_{\delta,M}(\mathbf{u}) \xrightarrow{d} W(\mathbf{u})$, quando $M \to \infty$ e $\delta \to 0$, vem finalmente que $W_n(\mathbf{u}) \xrightarrow{d} W(\mathbf{u})$.

(ii) Consideremos agora o caso em que $1 < \alpha < 2$ e admita-se que $E\left(|Z_1|^\beta\right)$ é finito, para algum $\beta < 1 - \alpha < 0$. Como $E\left(|Z_1|^{1-\alpha}\right) < E\left(|Z_1|^\beta\right)$, tem-se que $E\left(|Z_1|^{1-\alpha}\right) < \infty$. Logo pela proposição 3.1 b), $W(\cdot)$ é finita e portanto está bem definida. Mostremos agora que $W_n(\mathbf{u}) \xrightarrow{d} W(\mathbf{u})$. Note-se em primeiro lugar que para $x \neq 0$ e para qualquer y, verifica-se

$$|x - y| - |x| = y\,(I(x < 0) - I(x > 0)) + 2\,(y - x)\,(I(y > x > 0) - I(y < x < 0)).$$

Aplicando esta igualdade a $W_n(\mathbf{u})$, obtemos

$$\begin{aligned} W_n(\mathbf{u}) &= \sum_{t=1}^{n}\left[|Z_t - Y_{nt}(\mathbf{u})| - |Z_t|\right] = \sum_{t=1}^{n} Y_{nt}(\mathbf{u})\left(I(Z_t<0) - I(Z_t>0)\right) \\ &\quad + 2\sum_{t=1}^{n}\left(Y_{nt}(\mathbf{u}) - Z_t\right)\left(I(Y_{nt}(\mathbf{u}) > Z_t > 0) - I(Y_{nt}(\mathbf{u}) < Z_t < 0)\right) \\ &= W_{1n}(\mathbf{u}) + W_{2n}(\mathbf{u}). \end{aligned}$$

Vejamos que $W_{1n}(\mathbf{u})$ converge em distribuição. Ora,

$$\begin{aligned} W_{1n}(\mathbf{u}) &= \sum_{t=1}^{n} Y_{nt}(\mathbf{u})\left(I(Z_t<0) - I(Z_t>0)\right) I(|Z_t| \le M)\, I(|Y_{nt}(\mathbf{u})| > \delta) \\ &\quad + \sum_{t=1}^{n} Y_{nt}(\mathbf{u})\left(I(Z_t<0) - I(Z_t>0)\right) \\ &\qquad \times \left(I(|Y_{nt}(\mathbf{u})| \le \delta) + I(|Z_t| > M)\, I(|Y_{nt}(\mathbf{u})| > \delta)\right) \\ &= W^1_{1n,\delta,M}(\mathbf{u}) + W^2_{1n,\delta,M}(\mathbf{u}). \end{aligned}$$

Em relação a $W^1_{1n,\delta,M}(\mathbf{u})$ temos

$$\begin{aligned} &\sum_{t=1}^{n} Y_{nt}(\mathbf{u})\left(I(Z_t<0) - I(Z_t>0)\right) I(|Z_t| \le M)\, I(|Y_{nt}(\mathbf{u})| > \delta) \\ &\xrightarrow[n\to\infty]{d} \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} c_i \delta_k \Gamma_k^{-1/\alpha}\left(I(Z_k<0) - I(Z_k>0)\right) I(|Z_k| \le M)\, I\left(|c_i|\,\Gamma_k^{-1/\alpha} > \delta\right) \\ &\xrightarrow[M\to\infty,\delta\to 0]{q.c.} \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} c_i(\mathbf{u})\, \delta_k \Gamma_k^{-1/\alpha}\left(I(Z_k<0) - I(Z_k>0)\right) = W_1(\mathbf{u}). \end{aligned}$$

A convergência em distribuição ocorre por (3.13) e a convergência quase certa resulta pelos mesmos argumentos apresentados anteriormente na situação em que $\alpha < 1$, tendo-se neste caso essa convergência garantida pelo facto de

$$\begin{aligned} &\left|\sum_{k=1}^{\infty}\sum_{i=1}^{\infty} c_i \delta_k \Gamma_k^{-1/\alpha}\left(I(Z_k<0) - I(Z_k>0)\right) I(|Z_k| \le M)\, I\left(|c_i|\,\Gamma_k^{-1/\alpha} > \delta\right)\right| \\ &\le \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} c_i \delta_k \Gamma_k^{-1/\alpha}\left(I(Z_k<0) - I(Z_k>0)\right) \end{aligned} \tag{3.29}$$

e esta última série ser finita pela demonstração da proposição 3.1 b). À semelhança do que foi feito para o caso em que $\alpha \in (0,1)$, se provarmos que o limite em (3.26) é nulo, mas agora considerando os elementos $W_{1n}(\mathbf{u})$ e $W^1_{1n,\delta,M}(\mathbf{u})$, resultará a convergência em distribuição do primeiro para $W_1(\mathbf{u})$. teremos que mostrar que

$$\lim_{\delta\to 0}\lim_{M\to\infty}\limsup_{n\to\infty} P\left[\left|W_{1n}(\mathbf{u}) - W^1_{1n,\delta,M}(\mathbf{u})\right| > \eta\right] = 0, \forall \eta > 0. \tag{3.30}$$

Tem-se então

$$\begin{aligned}
\left|W^2_{1n,\delta,M}(\mathbf{u})\right| &= \left|W_{1n}(\mathbf{u}) - W^1_{1n,\delta,M}(\mathbf{u})\right| \\
&\leq \left|\sum_{t=1}^{n} Y_{nt}(\mathbf{u})\left(I(Z_t<0) - I(Z_t>0)\right) I\left(|Y_{nt}(\mathbf{u})| \leq \delta\right)\right| \\
&\quad + \left|\sum_{t=1}^{n} Y_{nt}(\mathbf{u})\left(I(Z_t<0) - I(Z_t>0)\right) I\left(|Z_t| > M\right) I\left(|Y_{nt}(\mathbf{u})| > \delta\right)\right|.
\end{aligned}$$

Aplicando a desigualdade de Chebychev e a proposição 3.2 c), tomando

$$V_t = I(Z_t<0) - I(Z_t>0),$$

e atendendo a que $E\left(I(Z_t<0) - I(Z_t>0)\right) = 0$, pois por hipótese $\{Z_t\}$ têm mediana nula se $\alpha \geq 1$, temos

$$\begin{aligned}
&\limsup_{n\to\infty} P\left[\left|\sum_{t=1}^{n} Y_{nt}(\mathbf{u})\left(I(Z_t<0) - I(Z_t>0)\right) I\left(|Y_{nt}(\mathbf{u})| \leq \delta\right)\right| > \eta\right] \\
\leq\ & \limsup_{n\to\infty} \eta^{-2} var \sum_{t=1}^{n} Y_{nt}(\mathbf{u})\left(I(Z_t<0) - I(Z_t>0)\right) I\left(|Y_{nt}(\mathbf{u})| \leq \delta\right) \\
=\ & \eta^{-2} E\left(I(Z_1<0) - I(Z_1>0)\right)^2 n a_n^{-2} E\left(Y_1^2 I\left(|Y_1|^2 \leq (a_n\delta)^2\right)\right) \\
\sim\ & C_3 E\left(V_1^2\right)\delta^{2-\alpha} \to 0,
\end{aligned} \tag{3.31}$$

quando $n \to \infty$ e $\delta \to 0$. Por outro lado, pela mesma proposição alínea b) e tomando

$$V_t = \left(I(Z_t<0) - I(Z_t>0)\right) I\left(|Z_t| > M\right)$$

vem que,

$$\begin{aligned}
&\limsup_{n\to\infty} P\left[\left|\sum_{t=1}^{n} Y_{nt}(\mathbf{u})\left(I(Z_t<0) - I(Z_t>0)\right) I\left(|Z_t| > M\right) I\left(|Y_{nt}(\mathbf{u})| > \delta\right)\right| > \eta\right] \\
\leq\ & \limsup_{n\to\infty} P\left[\sum_{t=1}^{n} |Y_{nt}(\mathbf{u})| \left|\left(I(Z_t<0) - I(Z_t>0)\right) I\left(|Z_t| > M\right)\right| I\left(|Y_{nt}(\mathbf{u})| > \delta\right) > \eta\right] \\
\leq\ & C_2\delta^{-\alpha} P\left[\left|\left(I(Z_1<0) - I(Z_1>0)\right) I\left(|Z_1| > M\right)\right| > 0\right] \\
=\ & C_2\delta^{-\alpha} P\left[|Z_1| > M\right] \to 0,
\end{aligned} \tag{3.32}$$

quando $M \longrightarrow \infty$. Logo combinando (3.31) e (3.32) obtemos

$$\left|W_{1n}(\mathrm{u}) - W^1_{1n,\delta,M}(\mathrm{u})\right| = o_p(1),$$

quando $n \longrightarrow \infty$, $M \longrightarrow \infty$ e $\delta \longrightarrow 0$, provando-se deste modo o limite em (3.30) e resultando por tal que $W_{1n}(\mathrm{u})$ converge em distribuição para $W_1(\mathrm{u})$, ou seja

$$\sum_{t=1}^{n} Y_{nt}(\mathrm{u})\left(I\left(Z_t < 0\right) - I\left(Z_t > 0\right)\right) \xrightarrow{d} \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha}\left(I\left(Z_k < 0\right) - I\left(Z_k > 0\right)\right).$$

Averiguemos agora a convergência de $W_{2n}(\mathrm{u})$, desdobrando-o na seguinte soma

$$\begin{aligned}
W_{2n}(\mathrm{u}) &= 2\sum_{t=1}^{n}\left(Y_{nt}(\mathrm{u}) - Z_t\right)\left(I\left(Y_{nt}(\mathrm{u}) > Z_t > 0\right) - I\left(Y_{nt}(\mathrm{u}) < Z_t < 0\right)\right)\\
&= 2\sum_{t=1}^{n}\left(Y_{nt}(\mathrm{u}) - Z_t\right)\left(I\left(Y_{nt}(\mathrm{u}) > Z_t > 0\right) - I\left(Y_{nt}(\mathrm{u}) < Z_t < 0\right)\right)\\
&\qquad\qquad \times I\left(\left|Y_{nt}(\mathrm{u})\right| > \delta\right)\\
&\quad +2\sum_{t=1}^{n}\left(Y_{nt}(\mathrm{u}) - Z_t\right)\left(I\left(Y_{nt}(\mathrm{u}) > Z_t > 0\right) - I\left(Y_{nt}(\mathrm{u}) < Z_t < 0\right)\right)\\
&\qquad\qquad \times \left(1 - I\left(\left|Y_{nt}(\mathrm{u})\right| > \delta\right)\right)\\
&= W^1_{2n,\delta}(\mathrm{u}) + W^2_{2n,\delta}(\mathrm{u}).
\end{aligned}$$

Para analisar a convergência em distribuição de $W_{2n}(\mathrm{u})$ segue-se exactamente o mesmo raciocínio implementado para as situações anteriores. Temos então

$$\begin{aligned}
&\sum_{t=1}^{n}\left(Y_{nt}(\mathrm{u}) - Z_t\right)\left(I\left(Y_{nt}(\mathrm{u}) > Z_t > 0\right) - I\left(Y_{nt}(\mathrm{u}) < Z_t < 0\right)\right) I\left(\left|Y_{nt}(\mathrm{u})\right| > \delta\right)\\
&\underset{n\to\infty}{\xrightarrow{d}} \sum_{t=1}^{n}\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha} - Z_k\right) \times I\left(\left|c_i(\mathrm{u})\right|\Gamma_k^{-1/\alpha} > \delta\right)\\
&\qquad \times\left(I\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha} > Z_k > 0\right) - I\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha} < Z_k < 0\right)\right)\\
&\xrightarrow{q.c} \sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha} - Z_k\right) \times\\
&\times\left(I\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha} > Z_k > 0\right) - I\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha} < Z_k < 0\right)\right) = W_2(\mathrm{u}).
\end{aligned}$$

A convergência em distribuição resulta por aplicação de (3.13) quando tomamos $n \longrightarrow \infty$ e a convergência quase certa resulta quando $\delta \longrightarrow 0$, tendo-se neste caso essa convergência garantida por se verificar

$$\begin{aligned}
&\left|\sum_{t=1}^{n}\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha}-Z_k\right)I\left(\left|c_i(\mathrm{u})\right|\Gamma_k^{-1/\alpha}>\delta\right)\right.\\
&\left.\times\left(I\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha}>Z_k>0\right)-I\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha}<Z_k<0\right)\right)\right|\\
\le\ &\sum_{t=1}^{n}\left|\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha}-Z_k\right)\right|\times\\
&\times\left|I\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha}>Z_k>0\right)-I\left(c_i(\mathrm{u})\,\delta_k\Gamma_k^{-1/\alpha}<Z_k<0\right)\right|,
\end{aligned}$$

sendo este último somatório finito pela demonstração da proposição 3.1. Provemos agora que

$$\lim_{\delta\to0}\limsup_{n\to\infty}P\left[\left|W_{2n}(\mathrm{u})-W_{2n,\delta}^{1}(\mathrm{u})\right|>\eta\right]=0,\forall\eta>0, \tag{3.33}$$

isto é, que $\left|W_{2n,\delta}^{2}(\mathrm{u})\right|=o_p(1)$, quando $n\to\infty$ e $\delta\to0$. Note-se primeiro que

$$\begin{aligned}
W_{2n,\delta}^{2}(\mathrm{u}) &= \sum_{t=1}^{n}\left(Y_{nt}(\mathrm{u})-Z_t\right)I\left(Y_{nt}(\mathrm{u})>Z_t>0\right)I\left(\left|Y_{nt}(\mathrm{u})\right|\le\delta\right)\\
&\quad-\sum_{t=1}^{n}\left(Y_{nt}(\mathrm{u})-Z_t\right)I\left(Y_{nt}(\mathrm{u})<Z_t<0\right)I\left(\left|Y_{nt}(\mathrm{u})\right|\le\delta\right),
\end{aligned}$$

e por conseguinte para mostrarmos que $W_{2n,\delta}^{2}(\mathrm{u})$ é um $o_p(1)$, é suficiente verificar que qualquer um dos somatórios anteriores também são $o_p(1)$'s em δ e n, ou seja que sejam verificados os limites

$$\lim_{\delta\to0}\limsup_{n\to\infty}P\left[\sum_{t=1}^{n}\left(Y_{nt}(\mathrm{u})-Z_t\right)I\left(Y_{nt}(\mathrm{u})>Z_t>0\right)I\left(\left|Y_{nt}(\mathrm{u})\right|\le\delta\right)>\gamma\right]=0 \tag{3.34}$$

e

$$\lim_{\delta\to0}\limsup_{n\to\infty}P\left[\sum_{t=1}^{n}\left(Y_{nt}(\mathrm{u})-Z_t\right)I\left(Y_{nt}(\mathrm{u})<Z_t<0\right)I\left(\left|Y_{nt}(\mathrm{u})\right|\le\delta\right)<-\gamma\right]=0. \tag{3.35}$$

Para provar (3.34) observe-se que pela desigualdade de Markov e considerando $G(\cdot)$ a função de distribuição de $Y_{nt}(\mathrm{u})$ e $F(\cdot)$ a função de distribuição de Z_t, tem-se

$$
\begin{aligned}
& P\left[\sum_{t=1}^{n}\left(Y_{nt}(\mathbf{u})-Z_t\right) I\left(Y_{nt}(\mathbf{u})>Z_t>0\right) I\left(\left|Y_{nt}(\mathbf{u})\right| \leq \delta\right)>\gamma\right] \\
\leq\ & \gamma^{-1} E\left(\sum_{t=1}^{n}\left(Y_{nt}(\mathbf{u})-Z_t\right) I\left(Y_{nt}(\mathbf{u})>Z_t>0\right) I\left(\left|Y_{nt}(\mathbf{u})\right| \leq \delta\right)\right) \\
=\ & \gamma^{-1} n E\left[\left(Y_{nt}(\mathbf{u})-Z_t\right) I\left(\delta \geq Y_{nt}(\mathbf{u})>Z_t>0\right)\right] \\
=\ & \gamma^{-1} n \int_0^{\delta} \int_{a_n z}^{a_n \delta}\left(\frac{y}{a_n}-z\right) G(dy) F(dz) \\
\leq\ & \gamma^{-1} n a_n^{-1} \int_0^{\delta} \int_{a_n z}^{\infty} y G(dy) F(dz)=\gamma^{-1} n a_n^{-1} \int_0^{\delta} H\left(a_n z\right) F(dz), \qquad (3.36)
\end{aligned}
$$

em que $H(z)=\int_z^{\infty} yG(dy)<\infty$, pois $E(|Z_1|)<\infty$ para $\alpha>1$. Vamos ver precisamente que a expressão em (3.36) tende para zero em δ e n. Note-se que

$$
\begin{aligned}
H(z) &= \int_z^{\infty} yG(dy)=\left[-y(1-G(y))\right]_z^{\infty}+\int_z^{\infty}(1-G(y))\,dy \\
&= \lim_{y\to\infty}\left[-y(1-G(y))\right]+z(1-G(z))+\int_z^{\infty}(1-G(y))\,dy.
\end{aligned}
$$

Como a distribuição da variável Y_{nt} comporta-se nas caudas de forma semelhante à distribuição de Z_t, então $y(1-G(y))=y^{-\alpha+1}L(y)$, pelo que, para $\alpha>1$, é nulo o seu limite em infinito. Por outro lado

$$
z(1-G(z))+\int_z^{\infty}(1-G(y))\,dy=z(1-G(z))\left(1+\frac{\int_z^{\infty}(1-G(y))\,dy}{z(1-G(z))}\right).
$$

Aplicando o teorema de Karamata alínea b) ao quociente acima, obtemos

$$
\frac{\int_z^{\infty}(1-G(y))\,dy}{z(1-G(z))} \sim \frac{1}{\alpha-1},
$$

o leva a que $H(z)$ possa ser escrito assintoticamente na forma

$$
H(z) \sim \frac{\alpha}{\alpha-1} z(1-G(z)).
$$

Como além disso $n(1 - G(a_n z)) \underset{n\to\infty}{\longrightarrow} Cz^{-\alpha}$, com $C > 0$, ter-se-á que

$$na_n^{-1} H(a_n) = na_n^{-1} \frac{\alpha}{\alpha - 1} a_n (1 - G(a_n)) \underset{n\to\infty}{\longrightarrow} C_1,$$

com $C_1 > 0$. Por outro lado a proposição 2.2 garante que sendo $H(z) \in RV_{1-\alpha}$ e para todo $\varepsilon > 0$, existe um $t_0 = t_0(\varepsilon)$, tal que para $t \geq t_0$ e $tz \geq t_0$ tem-se em particular

$$\frac{H(tz)}{H(t)} \leq (1+\varepsilon) z^{1-\alpha} \exp(-\varepsilon \log z) = (1+\varepsilon) z^{1-\alpha-\varepsilon}, \text{ para } 0 < z < 1.$$

Então a expressão (3.36) vem

$$\begin{aligned}
\gamma^{-1} na_n^{-1} \int_0^\delta H(a_n z) F(dz) &= \gamma^{-1} \int_0^\delta \frac{H(a_n z)}{H(a_n)} na_n^{-1} H(a_n) F(dz) \\
&\leq \gamma^{-1} \int_0^\delta (1+\varepsilon) z^{1-\alpha-\varepsilon} na_n^{-1} H(a_n) F(dz) \\
&\underset{n\to\infty}{\longrightarrow} C_1 \int_0^\delta (1+\varepsilon) z^{1-\alpha-\varepsilon} F(dz) \qquad (3.37)
\end{aligned}$$

Existindo um $\beta < 1 - \alpha$ tal que $E\left[|Z_1|^\beta\right] < \infty$ e tomando $\varepsilon = 1 - \alpha - \beta$ em (3.37) temos

$$\int_0^\delta (1+\varepsilon) z^{1-\alpha-\varepsilon} F(dz) \leq (2 - \alpha - \beta) \int_0^\delta z^\beta F(dz) \longrightarrow 0,$$

quando $\delta \longrightarrow 0$, provando-se assim que o limite em (3.34) é nulo. Para determinar o limite em (3.35) note-se que

$$\begin{aligned}
& P\left[\sum_{t=1}^n (Y_{nt}(\mathbf{u}) - Z_t) I(Y_{nt}(\mathbf{u}) < Z_t < 0) I(|Y_{nt}(\mathbf{u})| \leq \delta) < -\gamma\right] \\
= \;& P\left[\sum_{t=1}^n (Z_t - Y_{nt}(\mathbf{u})) I(-\delta \leq Y_{nt}(\mathbf{u}) < Z_t < 0) > \gamma\right] \\
= \;& P\left[\sum_{t=1}^n (Z_t - Y_{nt}(\mathbf{u})) I(\delta > -Y_{nt}(\mathbf{u}) > -Z_t > 0) > \gamma\right].
\end{aligned}$$

Tomando-se $Y_{nt}^* = -Y_{nt}$ e $Z_t^* = -Z_t$ ficamos com uma probabilidade definida como em (3.34), sendo por tal, minuciosamente semelhante a demonstração de que a referida probabilidade tende para zero. Por essa razão os cálculos são suprimidos. Ora, uma vez provado que o limite em (3.33) é nulo, vem que $W_{2n}(\mathbf{u}) \xrightarrow{d} W_2(\mathbf{u})$, ou seja,

$$
\begin{aligned}
&2\sum_{t=1}^{n}\left(Y_{nt}(\mathbf{u}) - Z_t\right)\left(I\left(Y_{nt}(\mathbf{u}) > Z_t > 0\right) - I\left(Y_{nt}(\mathbf{u}) < Z_t < 0\right)\right) \\
&\xrightarrow{d} 2\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left[\left(c_i(\mathbf{u})\,\delta_k\Gamma_k^{-1/\alpha} - Z_k\right) \times\right. \\
&\qquad\qquad \left.\times\left(I\left(c_i(\mathbf{u})\,\delta_k\Gamma_k^{-1/\alpha} > Z_k > 0\right) - I\left(c_i(\mathbf{u})\,\delta_k\Gamma_k^{-1/\alpha} < Z_k < 0\right)\right)\right],
\end{aligned}
$$

vindo finalmente,

$$
\begin{aligned}
W_n(\mathbf{u}) &= \sum_{t=1}^{n}\left[\left|Z_t - Y_{nt}(\mathbf{u})\right| - \left|Z_t\right|\right] \\
&\xrightarrow{d} \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} c_i(\mathbf{u})\,\delta_k\Gamma_k^{-1/\alpha}\left(I\left(Z_k < 0\right) - I\left(Z_k > 0\right)\right) \\
&\quad +2\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left[\left(c_i(\mathbf{u})\,\delta_k\Gamma_k^{-1/\alpha} - Z_k\right) \times\right. \\
&\qquad\qquad \left.\times\left(I\left(c_i(\mathbf{u})\,\delta_k\Gamma_k^{-1/\alpha} > Z_k > 0\right) - I\left(c_i(\mathbf{u})\,\delta_k\Gamma_k^{-1/\alpha} < Z_k < 0\right)\right)\right] \\
&= \sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\left[\left|Z_k - c_i(\mathbf{u})\,\delta_k\Gamma_k^{-1/\alpha}\right| - \left|Z_k\right|\right].
\end{aligned}
$$

(iii) O terceiro e último caso a analisar é quando $\alpha = 1$. Como por hipótese $-E(\ln|Z_1|) < \infty$, então pela proposição 3.1 $W(\mathbf{u})$ está bem definida. Todos os cálculos e argumentos utililizados na demonstração da convergência em distribuição de $W_n(\mathbf{u})$ no caso em que $1 < \alpha < 2$, são válidos para $\alpha = 1$. A excepção está no cálculo dos limites em (3.34) e (3.35), uma vez que agora $E(|Z_1|) = \infty$. A alternativa é calcular por definição o primeiro valor esperado em (3.36)., provando que ele tende para zero quando $n \longrightarrow \infty$ e $\delta \longrightarrow 0$. Temos então

$$\begin{aligned}
& E\left(\sum_{t=1}^{n}\left(Y_{nt}(\mathbf{u})-Z_t\right)I\left(Y_{nt}(\mathbf{u})>Z_t>0\right)I\left(\left|Y_{nt}(\mathbf{u})\right|\leq\delta\right)\right)\\
= \;& nE\left[\left(Y_{nt}(\mathbf{u})-Z_t\right)I\left(\delta\geq Y_{nt}(\mathbf{u})>Z_t>0\right)\right]\\
= \;& n\int_0^{\delta}\int_{a_nz}^{a_n\delta}\left(\frac{y}{a_n}-z\right)G(dy)\,F(dz)=na_n^{-1}\int_0^{\delta}\int_{a_nz}^{a_n\delta}\left(y-a_nz\right)G(dy)\,F(dz)\\
= \;& na_n^{-1}\int_0^{\delta}\left(\left[-\left(y-a_nz\right)\left(1-G(y)\right)\right]_{a_nz}^{a_n\delta}+\int_{a_nz}^{a_n\delta}\left(1-G(y)\right)dy\right)F(dz)\\
= \;& -n\int_0^{\delta}(\delta-z)\left(1-G\left(a_n\delta\right)\right)F(dz)+na_n^{-1}\int_0^{\delta}\int_{a_nz}^{a_n\delta}\left(1-G(y)\right)dyF(dz).
\end{aligned}$$

Em relação à primeira parcela note-se que

$$-n\int_0^{\delta}(\delta-z)\left(1-G\left(a_n\delta\right)\right)F(dz)\underset{n\to\infty}{\longrightarrow}-C\int_0^{\delta}\left(1-\frac{z}{\delta}\right)F(dz),\quad C\in\mathbb{R}.$$

Como $0<z<\delta$, então $0<1-\frac{z}{\delta}<1$, vindo

$$0<\int_0^{\delta}\left(1-\frac{z}{\delta}\right)F(dz)<\int_0^{\delta}F(dz)<\infty.$$

Tomando $\delta\longrightarrow 0$, tem-se $\int_0^{\delta}F(dz)\longrightarrow 0$ e portanto $\int_0^{\delta}\left(1-\frac{z}{\delta}\right)F(dz)\longrightarrow 0$, logo $-\int_0^{\delta}\left(1-\frac{z}{\delta}\right)F(dz)\longrightarrow 0$. Fazendo a mudança de variável $y=a_nt$, na segunda parcela, obtem-se

$$\begin{aligned}
na_n^{-1}\int_0^{\delta}\int_{a_nz}^{a_n\delta}\left(1-G(y)\right)dyF(dz) &= n\int_0^{\delta}\int_z^{\delta}\left(1-G\left(a_nt\right)\right)dtF(dz)\\
&\underset{n\to\infty}{\longrightarrow} C\int_0^{\delta}\int_z^{\delta}t^{-1}dtF(dz)\\
&= C\int_0^{\delta}\left(\log\delta-\log|z|\right)F(dz).
\end{aligned}$$

Uma vez que $0 < z < \delta$, então $\log z = \log|z| < \log\delta$, donde $\log\delta - \log|z| > 0$. Ora nesse caso é então válido

$$\int_0^\delta (\log\delta - \log|z|)\, F(dz) > 0.$$

Por outro lado se $|z| < \delta$, então $\frac{\delta}{|z|} > 1$, implicando que $\log\frac{\delta}{z} < \frac{\delta}{z}$. Logo

$$\int_0^\delta (\log\delta - \log|z|)\, F(dz) \int_0^\delta \log\frac{\delta}{|z|} F(dz) < \int_0^\delta \frac{\delta}{|z|} F(dz) = \delta \int_0^\delta \frac{1}{|z|} F(dz) < \infty,$$

uma vez que $z \neq 0$. Ou seja, concluímos que

$$0 < \int_0^\delta (\log\delta - \log|z|)\, F(dz) < \delta \int_0^\delta \frac{1}{|z|} F(dz) < \infty.$$

Tomando $\delta \longrightarrow 0$ vem que $\int_0^\delta (\log\delta - \log|z|)\, F(dz) \longrightarrow 0$, o que garante que o limite em (3.34) é nulo para $\alpha = 1$. (3.35) obtém-se de forma análoga. Logo ficou provado que $W_n(\mathbf{u}) \xrightarrow{d} W(\mathbf{u})$ para $\mathbf{u} \in \mathbb{R}^p$ e $0 < \alpha < 2$. ∎

Com o próximo lema, temos finalmente reunidas as condições, para mostrarmos a convergência em distribuição do estimador LAD para o mínimo em $\mathbf{u}$, do processo estocástico $W(\cdot)$.

Lema 3.1 *Sejam $\{V_n(\cdot)\}$ e $V(\cdot)$ processos estocásticos em $\mathbb{R}^p$ e suponha-se que $V_n(\cdot) \xrightarrow{d} V(\cdot)$ em $C(\mathbb{R}^p)$. Considerem-se ainda*

$$\xi_n = \min_{\mathbf{u}} V_n(\mathbf{u}) \quad e \quad \xi = \min_{\mathbf{u}} V(\mathbf{u}).$$

Se $V_n(\cdot)$ é convexa para todo o n e ξ é único com probabilidade 1, então $\xi_n \xrightarrow{d} \xi$ em $\mathbb{R}^p$.

Dem. O Teorema de Representação de Skorohod[6] garante que existe um espaço de probabilidade (Ω, A, P) e elementos $\{V_n^*(\cdot)\}$ e $V^*(\cdot)$ em $(C(\mathbb{R}^p), \rho)$, definidos nesse mesmo espaço probabilístico tais que $V_n^*(\cdot) \stackrel{d}{=} V_n(\cdot)$, para todo o n e $V^*(\cdot) \stackrel{d}{=} V(\cdot)$, tais que

$$\sup_{\mathbf{u}\in K} |V_n^*(\mathbf{u}) - V^*(\mathbf{u})| \rightarrow 0, \tag{3.38}$$

[6]Veja-se em apêndice.

quase certamente, para todo o compacto $K \subset \mathbb{R}^p$. Sejam ξ_n^* a variável aleatória que representa o mínimo em $\mathbf{u}$ de $V_n^*(\mathbf{u})$ e ξ^* a v.a. que minimiza $V^*(\mathbf{u})$. Se mostrarmos que $\xi_n^* \overset{q.c.}{\rightarrow} \xi^*$, então teremos seguramente que $\xi_n \overset{d}{\rightarrow} \xi$. Isto acontece porque a distribuição de probabilidade de ξ_n^* deduz-se unicamente a partir das distribuições dimensionais finitas de $V_n^*(\cdot)$, ou seja, a partir das distribuições conjuntas com respeito a $V_n^*(\mathbf{u}_1), V_n^*(\mathbf{u}_2), ..., V_n^*(\mathbf{u}_p)$. Como estas distribuições coincidem com as de $V_n(\mathbf{u}_1), V_n(\mathbf{u}_2), ..., V_n(\mathbf{u}_p)$, ter-se-á $\xi_n^* \overset{d}{\rightarrow} \xi$. Ora, tendo-se $\xi_n^* \overset{q.c.}{\rightarrow} \xi^*$ então também teremos $\xi_n^* \overset{d}{\rightarrow} \xi^*$, o que justifica a convergência em distribuição de ξ_n para ξ. Note-se que a continuidade de $V_n^*(\cdot)$ e $V^*(\cdot)$ em $\mathbb{R}^p$ é garantida pelo próprio teorema de Skorohod. Por outro lado, usando o mesmo teorema consegue-se provar que $V_n^*(\cdot)$ é q.c. convexa, já que sendo $V_n(\cdot)$ convexa por hipótese tem-se por definição:

$$V_n(\lambda\mathbf{u}_1 + (1-\lambda)\mathbf{u}_2) \leq \lambda V_n(\mathbf{u}_1) + (1-\lambda)V_n(\mathbf{u}_2), \forall \mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^p, 0 \leq \lambda \leq 1.$$

Então

$$P[V_n(\lambda\mathbf{u}_1 + (1-\lambda)\mathbf{u}_2) \leq \lambda V_n(\mathbf{u}_1) + (1-\lambda)V_n(\mathbf{u}_2)] = 1,$$

donde

$$P[V_n^*(\lambda\mathbf{u}_1 + (1-\lambda)\mathbf{u}_2) \leq \lambda V_n^*(\mathbf{u}_1) + (1-\lambda)V_n^*(\mathbf{u}_2)] = 1,$$

isto é, $V_n^*(\cdot)$ é q.c. convexa. Com vista a mostrarmos que $\xi_n^* \overset{q.c.}{\rightarrow} \xi^*$ considere-se, para $w \in \Omega$, o conjunto

$$B_\gamma(w) = \{\mathbf{u} \in \mathbb{R}^p : \|\mathbf{u} - \xi^*(w)\| = \gamma\}, \text{ para qualquer } \gamma > 0.$$

Ora, por um resultado da teoria das probabilidades[7] sabemos que $\xi_n^* \overset{q.c.}{\rightarrow} \xi^*$ sse

$$P[\|\xi_n^* - \xi^*\| > \gamma, i.o] = 0, \text{ para todo } \gamma > 0.$$

Por absurdo admita-se então que a probabilidade anterior é não nula e que (3.38) é verificado. Isto significa que o conjunto $C = \{w : \|\xi_n^*(w) - \xi^*(w)\| > \gamma, i.o\}$ é não vazio e tem probabilidade não nula. Por outro lado

$$V_n^*(\xi^*) \overset{q.c.}{\rightarrow} V^*(\xi^*), \tag{3.39}$$

já que a partir de (3.38) tem-se $\forall\mathbf{u}, V_n^*(\mathbf{u}) \overset{q.c.}{\rightarrow} V^*(\mathbf{u})$, bastando tomar $K = \{\mathbf{u}\}$. Além disso em B_γ verifica-se que

$$V_n^*(\mathbf{u}) \rightarrow V^*(\mathbf{u}), \text{ uniformemente.} \tag{3.40}$$

[7] Veja-se em apêndice.

Defina-se em Ω, o conjunto $A = \{w : V_n^*(\xi^*(w), w) \to V^*(\xi^*(w), w)\}$ para o qual se tem $P(A) = 1$. Note-se que $P(A \cap C) \neq 0$, pois

$$\underbrace{P(C)}_{\neq 0} = P(A \cap C) + \underbrace{P(\overline{A} \cap C)}_{<P(\overline{A})=0},$$

pelo que $A \cap C \neq \emptyset$. Tome-se $w \in A \cap C$ e $\mathbf{u} \in B_\gamma(w)$. Então tem-se por (3.39) e (3.40), para cada $\delta > 0$, tem-se a partir de certa ordem

$$V_n^*(\xi^*, w) < V^*(\xi^*, w) + \frac{\delta}{2}, \tag{3.41}$$

e

$$V_n^*(\mathbf{u}, w) > V^*(\mathbf{u}, w) - \frac{\delta}{2}. \tag{3.42}$$

Tomando $\delta = V^*(\mathbf{u}, w) - V^*(\xi^*, w)$, vem por (3.41)

$$\begin{aligned} V_n^*(\xi^*, w) &< V^*(\xi^*, w) + \frac{V_n^*(\mathbf{u}, w) - V^*(\xi^*, w)}{2} \\ &\Leftrightarrow 2V_n^*(\xi^*, w) < V^*(\xi^*, w) + V^*(\mathbf{u}, w), \end{aligned}$$

e por (3.42)

$$\begin{aligned} V_n^*(\mathbf{u}, w) &> V^*(\mathbf{u}, w) - \frac{V^*(\mathbf{u}, w) - V^*(\xi^*, w)}{2} \\ &\Leftrightarrow 2V_n^*(\mathbf{u}, w) > V^*(\mathbf{u}, w) + V^*(\xi^*, w), \end{aligned}$$

o que implica que se tenha

$$2V_n^*(\mathbf{u}, w) > 2V^*(\xi^*, w).$$

Mas por outro lado

$$V_n^*(\xi^*, w) \geq V_n^*(\xi_n^*, w),$$

por ξ_n^* ser mínimo de $V_n^*(\cdot)$, donde para infinitos n

$$V_n^*(\mathbf{u}, w) > V_n^*(\xi^*, w) \geq V_n^*(\xi_n^*, w). \tag{3.43}$$

Seja agora o conjunto

$$D = \{w : \textit{a condição (3.43) é válida}\},$$

vem então que $P(D) \geq P(A \cap C) > 0$. Escolhendo um vector $\mathbf{u}$ em B_γ de modo que $\boldsymbol{\xi}^*(w), \boldsymbol{\xi}_n^*(w)$ e $\mathbf{u}$ sejam colineares, as desigualdades em (3.43), indicam que $V_n^*(\cdot, w)$ não é convexa para $w \in D$, o que contraria o facto de $V_n^*(\cdot, w)$ ser quase certamente convexa. ■

Teorema 3.2 *Sob as condições do teorema 3.1, se $W(\cdot)$ tem um único mínimo q.c., então*

$$a_n(\widehat{\phi} - \phi) \xrightarrow{d} \xi,$$

onde $\xi = \min W(\cdot)$ e $\widehat{\phi}$ é o estimador LAD de ϕ.

Dem. A demonstração resulta do teorema 3.1 e do lema anterior. ∎

Conclusões e Desenvolvimentos Futuros

A estimação dos parâmetros dos processos ARMA requer uma atenção cuidada e especial quando tais processos são dirigidos por inovações com caudas elevadas. Perante erros com variância infinita, como já anteriormente referimos, os estimadores usuais, determinados com base nas condições de segunda ordem descritas pelas funções de densidade espectral ou de autocovariânca, perdem as suas propriedades assintóticas, obtendo desse modo uma má *perfomance*. Como estimar este tipo de modelos, é naturalmente uma questão pertinente, que tem encontrado resposta ao longo dos anos no desenvolvimento de diversos métodos de estimação. Neste trabalho quisémos apresentar algumas ideias sobre um desses procedimentos de estimação, concretamente a M-estimação, quando o processo em estudo era um modelo autoregressivo de ordem p. Na classe dos M-estimadores, optámos pelo estimador LAD, por ser este o menos conhecido, face por exemplo, ao veterano estimador dos mínimos quadrados. Além disso, o facto deste estimador ser obtido a partir da função módulo e desta não ser diferenciável na origem, despertou o nosso interesse em determinar analiticamente o seu comportamento limite. Como a avaliação do desempenho de um estimador é feita precisamente analisando esse comportamento limite, preocupámo-nos apenas em determinar e comprender um conjunto de resultados que sustentam tal comportamento. Em concreto vimos que, impostas à sequência dos erros certas condições de momentos, que envolvem o comportamento da respectiva função de distribuição junto da origem, o estimador LAD converge em distribuição, a uma taxa de velocidade de ordem $n^{1/\alpha}L(n)$, com $\alpha \in (0, 2)$, para um vector aleatório ξ que representa o mínimo de um processo estocástico, definido no espaço das funções contínuas de $\mathbb{R}^p$ em $\mathbb{R}$. É indiscutível a importância da teoria dos processos pontuais na obtenção deste resultado. Aliás os dois primeiros capítulos espelham exactamente essa importância. No primeiro, vimos como se constrói toda uma teoria que engloba entre outros, os conceitos de convergência vaga e fraca de processos pontuais. No segundo, vimos como esses conceitos se aplicam no caso concreto em que os processos pontuais são formados com base numa sequência de ruídos, geradores de um processo linear de médias móveis $\{X_k\}$ e que têm a função de distribuição comum com as caudas de variação regular de índice $-\alpha$, para $\alpha > 0$. Embora não faça parte do âmbito deste trabalho, é importante salientar mais uma vez, que com este tipo de processos pontuais, podem ser estudados entre outros, o comportamento limite de extremos e de somas e funções de covariâncias amostrais do processo $\{X_k\}$, tendo-se deste modo criada uma estrutura sólida de resultados que permita transportá-los para o plano da estimação e da previsão. Se juntarmos a tudo isto, o facto das caudas da distribuição desses erros verificarem a propriedade (2.6) e se restringirmos o expoente de variação α ao intervalo $(0, 2)$, acabamos por lidar com processos lineares gerados por erros com variância infinita, tema a que nos dedicámos no terceiro capítulo.

Embora já exista um número razoável de estudos de simulação que avaliem as Lad-estimativas dos parâmetros de processos ARMA com variância infinita, pretendemos a partir deste trabalho implementar mais experiências a esse nível e que possam embarcar eventualmente outros estimadores, como o estimador dos mínimos quadrados. Outra questão interessante que gostaríamos de desenvolver, prende-se com a possibilidade de se poder alargar o estudo da M-estimação a funções de perda do tipo $\rho(x) = |x|^\gamma$, com $\gamma > 0$, numa tentativa de se encontrar o γ óptimo, que torne as propriedades de consistência dos estimadores mais fortes. Alguns resultados já foram desenvolvidos por Davis e outros (1992) para processos autoregressivos, no entanto da bibliografia consultada nestas áreas, não encontrámos referências quer a estudos teóricos, quer a estudos de simulação, que tratem esta questão para os modelos ARMA em geral. Apesar de termos provado a existência da lei limite não degenerada do vector aleatório ξ, não existe até ao momento nenhum método simples que determine essa distribuição limite. Sabe-se que a distribuição em causa depende fortemente da distribuição das inovações e da função perda $\rho(x) = |x|$, mas ao que parece a sua expressão é extremamente complicada. Davis e Wu (1994) trabalharam neste assunto utilizando técnicas amostrais de aproximação à distribuição de $(\widehat{\phi}_M - \phi)$ na classe dos M-estimadores, sem contudo obterem grandes resultados. Fica lançada mais uma questão em aberto para um possível desenvolvimento futuro.

Apêndice I: Espaços Topológicos

1. **Espaços de Hausdorff localmente compactos**

   Dados um conjunto E e a respectiva σ-álgebra $\mathcal{E}$, gerada pelos subconjuntos abertos de E, diz-se que E é um *espaço topológico de Hausdorff*, sse for um espaço topológico que verifica a propriedade: se x_1, x_2 são elementos de E, então existem conjuntos disjuntos G_1, G_2 em E, tais que $x_1 \in G_1$ e $x_2 \in G_2$. Um espaço topológico de Hausdorff localmente compacto, verifica a propriedade adicional de todo o elemento $x \in E$, ter uma vizinhança compacta $\mathcal{V}_\delta(x)$. Recorde-se que um conjunto $A \subset E$ é compacto por definição, sse toda a cobertura de abertos de A, contém uma subcobertura finita. Recorde-se ainda que $B \subset E$, diz-se *limitado* ou *relativamente compacto* sse o seu conjunto aderente $\overline{B}$ é compacto em E.

2. **Espaços métricos, completos e separáveis**

   Um espaço métrico $\{E, \rho\}$, diz-se *separável* por definição se contém um subconjunto denso contável, i.e., se existe um conjunto contável $D \subset E$, tal que $\overline{D} = E$. O espaço E diz-se completo se toda a sequência de Cauchy em E é uma sequência convergente.

3. **DC-semianéis**

   Dado um espaço topológico de Hausdorff, E, localmente compacto e com bases contáveis e $\mathcal{E}$ a σ-álgebra gerada pelos subconjuntos abertos de E, considere-se o *anel* $\mathfrak{B}$ formado por todos os conjuntos relativamente compactos de $\mathcal{E}$. Uma subclasse $\mathcal{S} \subset \mathfrak{B}$, diz-se um *DC-semianel* (D de "dissecting" e C de "covering"), se fôr um semianel e verificar a propriedade: para todo $S \in \mathcal{S}$ e para todo o $\varepsilon > 0$, existe uma cobertura finita de S formada por subconjuntos de $\mathcal{S}$, de diâmetro menor que ε. Recorde-se que $\mathcal{S}$ é um semianel se fôr fechado para a intersecção finita de conjuntos e se toda a diferença própria entre dois conjuntos de $\mathcal{S}$, puder ser escrita como uma união finita de conjuntos disjuntos de $\mathcal{S}$. Prova-se que em $\mathbb{R}$, os DC-semianéis são famílias de intervalos.

Apêndice II: Alguns Resultados Importantes

1. Teorema de Prohorov (1956): *Seja* $\{\mathcal{X}, \rho\}$, *um espaço métrico, completo e separável e* $\{\mu_n\}$ *uma sequência de medidas de probabilidade em* $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, *com* $B(\mathcal{X})$ *a* σ*-álgebra de borel de* X. *Então* $\{\mu_n\}$ *é relativamente compacta sse é tight.*

   Dem. Veja-se em [2], pág.240. ∎

2. Teorema 4.2 (Kallenberg,1983): *Sejam* E *um espaço topológico de Hausdorff, localmente compacto e com bases contáveis,* $\mathcal{E}$ *a respectiva* σ*-álgebra de borel e* $\mathfrak{B}$ *a classe de todos os conjuntos relativamente compactos de* E. *Defina-se*

$$\mathfrak{B}_\xi = \{B \in \mathfrak{B} : \xi(\partial B) = 0 \quad q.c.\}.$$

   Dadas $\xi, \xi_1, \xi_2, ...$ *medidas aleatórias em* E *e* $\mathcal{S}$ *um DC-semianel contido em* $\mathfrak{B}_\xi$, *então são equivalentes a condições*

   (a) $\xi_n \xrightarrow{d} \xi$,

   (b) $\xi_n(f) \xrightarrow{d} \xi(f)$, *para toda a função* $f \in f \in C_K^+(E)$,

   (c) $(\xi_n(B_1), ..., \xi_n(B_k)) \xrightarrow{d} (\xi(B_1), ..., \xi(B_k))$, *para quaisquer conjuntos* $B_1, ..., B_k \in \mathcal{S}$, $k \geq 1$.

   Dem. A demonstração pode ser consultada em [19], pág.32. ∎

3. Lei dos logarítmos iterados de Hartman e Winter (1941)

   (a) *Dadas* $X_1, X_2, ...$ *variáveis aleatórias i.i.d., com média* μ *e variância finita* σ^2, *então*

$$\limsup_{n\to\infty} \frac{S_n - n\mu}{\sqrt{2\sigma^2 n \log\log n}} = 1 \quad \text{q.c.},$$

onde $S_n = X_1 + ... + X_n$.

Dem. A demonstração pode ser consultada por exemplo em [2]. ∎

4. Proposição: *Dadas* $X_1, X_2, ...$ *v.a.'s não correlacionadas, onde* $E(X_k) = 0$, *para* $k = 1, 2, ...,$ *e* $\sum_{k=1}^{\infty} E(X_k^2) < \infty$, *então* $\sum_{k=1}^{\infty} X_k$ *converge quase certamente.*

   Dem. A demonstração pode ser consultada por exemplo em [2]. ∎

5. Proposição: *Sejam* $X_1, X_2, ...$ *v.a.'s definidas num mesmo espaço de probabilidade* $(\Omega, \mathcal{A}, P)$. *Se* $\sum_{n\geq 1} E(|X_n|) < \infty$, então $\sum_{n\geq 1} X_n$ *converge em valor absoluto quase certamente e*

$$E\left(\sum_{n\geq 1} X_n\right) = \sum_{n\geq 1} E(X_n).$$

   Dem. A demonstração pode ser consultada por exemplo em [2]. ∎

6. Teorema de Kolmogorov (1933): Se $X_1, X_2, ...$ *são v.a.'s i.i.d., com média finita* μ, então

$$\frac{1}{n}\sum_{k=1}^{n} X_k \to \mu, \quad q.c.$$

   Dem. Veja-se por exemplo [2] ou em [14]. ∎

7. Teorema de Cramer-Wold (*Cramer-Wold Device*): *dados os vectores aleatórios* $X_n = (X_{n1}, ..., X_{nk})$ *e* $X = (X_1, ..., X_k)$, *então* $X_n \xrightarrow{d} X$ *sse* $\forall k \in \mathbb{N}$

$$\sum_{j=1}^{k} c_j X_{nj} \xrightarrow{d} \sum_{j=1}^{k} c_j X_j,$$

   para todo o vector de constantes $c = (c_1, ..., c_k) \in \mathbb{R}^k$.

   Dem. Veja-se em [2], pág.49. ∎

8. Teorema de Skorohod (1956): *Sejam* X_n e X *elementos aleatórios de um espaço métrico, completo e separável* $\{\mathcal{X}, \rho\}$, *podendo eventualmente estarem definidos em espaços de probabilidade diferentes. Suponha-se que* $X_n \xrightarrow{d} X$. *Então existe um espaço probabilístico* $(\Omega, \mathcal{A}, P)$ *e elementos aleatórios* X_n' *e* X' *em* $\{\mathcal{X}, \rho\}$, *todos definidos sobre o mesmo espaço* $(\Omega, \mathcal{A}, P)$, *tais que* $X' \stackrel{d}{=} X$ *e* $X_n' \stackrel{d}{=} X_n$, *para* $n = 1, 2, ...,$ *e* $\rho(X_n', X') \to 0$ *quase certamente.*

   Dem. Aconselha-se a consulta da demonstração em [25]. ∎

9. Proposição: *Dadas* X, $X_1, X_2, \ldots$, *variáveis aleatórias todas definidas sobre o mesmo espaço de probabilidade* $(\Omega, \mathcal{A}, P)$, *então* $X_n \to X$ *quase certamente sse* $P(|X_n - X| > \varepsilon \text{ i.o.}) = 0$, *para todo* $\varepsilon > 0$.

   Dem. A demonstração pode ser consultada em [2] ou em [14]. ■

10. Dada uma sequência $\{Z_k\}$ de v.a.'s i.i.d. que verificam (2.5) e (2.6). Então

    (a) $E\left[|Z_1|^{1-\alpha}\right] < \infty$ sse para $\alpha > 1$ $\int_0^a z^{1-\alpha} F(dz)$ é finito, com $a \in \mathbb{R}$.

    (b) $-E[\log|Z_1|] < \infty$ sse para $\alpha = 1$ $\int_0^a - \ln|z| F(dz)$é finito, com $a \in \mathbb{R}$.

    Dem. a) A demonstração é feita tendo em conta que $a > 0$. A situação em que $a < 0$ prova-se de forma semelhante. Se $\alpha > 1$, tem-se

$$E\left[|Z_1|^{1-\alpha}\right] = \int_{-\infty}^{\infty} |z|^{1-\alpha} F(dz) = \int_{-\infty}^{0} -z^{1-\alpha} F(dz) + \int_{0}^{\infty} z^{1-\alpha} F(dz),$$

em que $F(\cdot)$ é a função de distribuição comum das v.a's Z_k's. Notando que

$$\int_{0}^{\infty} z^{1-\alpha} F(dz) = \int_{0}^{a} z^{1-\alpha} F(dz) + \int_{a}^{\infty} z^{1-\alpha} F(dz),$$

tem se por um lado que

$$\int_{a}^{\infty} z^{1-\alpha} F(dz) = \left[-z^{1-\alpha}(1 - F(z))\right]_a^{\infty} + (1-\alpha)\int_a^{\infty} z^{-\alpha}(1 - F(z))\, dz.$$

Mas como

$$\lim_{t\to\infty} \left[-z^{1-\alpha}(1 - F(z))\right]_a^t = \lim_{t\to\infty} \left[-Cz^{-\alpha} L(z)\right]_a^t = C_1,$$

onde $C, C_1 \in \mathbb{R}$ e $L(z)$ é uma função de variação lenta, e além disso

$$\int_a^{\infty} z^{-\alpha}(1 - F(z))\, dz \leq \int_a^{\infty} z^{-\alpha} dz < \infty,$$

então a existência de $E\left[|Z_1|^{1-\alpha}\right]$ vai depender da existência de $\int_{0^+} z^{1-\alpha} F(dz)$.

b)Se $\alpha = 1$ tem-se que

$$E[-\ln|Z_1|] = \int_{-\infty}^{\infty} - \ln|z| F(dz) = \int_{-\infty}^{0} - \ln(-z) F(dz) + \int_{0}^{\infty} - \ln z F(dz).$$

Analisemos a convergência do integral $\int_0^\infty -\ln z F(dz)$, já que a convergência de $\int_{-\infty}^0 -\ln(-z) F(dz)$ estuda-se de forma análoga. Ora para $a > 0$ tem-se

$$\int_0^\infty -\ln z F(dz) = \int_0^a -\ln z F(dz) + \int_a^\infty -\ln z F(dz).$$

Por outro lado

$$\int_a^\infty -\ln z (1 - F(z)) F(dz) = [\ln z (1 - F(z))]_a^\infty - \int_a^\infty \frac{1}{z} (1 - F(z)) dz,$$

onde

$$[\ln z (1 - F(z))]_a^\infty = \lim_{t \to \infty} \left[\frac{\ln z}{z} z (1 - F(z)) \right]_a^t = C \in \mathbb{R}.$$

Como

$$\int_a^\infty \frac{1}{z} (1 - F(z)) F(dz) = \int_a^\infty \frac{1}{z^2} z (1 - F(z)) dz < \infty,$$

conclui-se que a existência de $E[-\ln |Z|]$ vai depender de $\int_{0^+} -\ln z F(dz) < \infty$. ■

Bibliografia

[1] An, H.Z. e Chen, Z.G. (1982), On Convergence of LAD Estimates in Autoregression with Infinite Variance, *Journal of Multivariate Analysis*, 12, p. 335-345.

[2] Billingsley, P. (1968), *Convergence of Probabilitiy Measures*, Wiley, New York.

[3] Bloomfield, P. e Stieger, W. (1983), *Least Absolute Deviations: Theory, Applications and Algorithms*, Boston, MA, Birkhäuser.

[4] Brocwell, P.J. e Davis, R.A. (1991), *Time Series: theory and Methods*, 2ª edição, Springer, New York.

[5] Calder, M. e Davis, R.A. (1998), Inference for Linear Processes with Stable Noise, In Adler, R.J., Feldman, R.E. e M.S. Taqqu, editors, *A Practical Guide to Heavy Tails*, p. 159-176, Birhäuser.

[6] Cline, D., Infinite Series of Random Variables with Regulary Varying Tails, *Tech. Reports*, Institute of Applied Mathematics and Statistics, University of British Columbia, p.83-24, 1983.

[7] Davis, R.A. e Resnick, S. (1985a), Limit Theory for Moving Averages of Random Variables with Regularly Tail Probabilities, *The Annals of Probability*, 13, p. 179-195.

[8] Davis, R.A. e Resnick, S. (1986), Limit Theory for the Sample Covariance and Correlation Functions of Moving Averages, *The Annals of Statistics*, 14, p. 533-558.

[9] Davis, R.A., Knight, K. e Liu, J. (1992), M-estimation For Autoregressions with Infinite Variance, *Stochastic Processes and their Applications*, 40, p. 145-180.

[10] Davis, R.A. (1996), Gauss-Newton and M-Estimation for ARMA Processes with Infinite Variance, *Stochastic Processes and their Applications*, 63, p. 75-95.

[11] Davis, R. e Wu, W. (1994), Bootstrapping M-estimates in Regression and Autoregression with Infinite Variance, *Preprint, Colorado State Univversity, Dept. of Statistics.*

[12] Fama, E. (1965), Behavior of Stock Market Prices, *J. Bus U. Chicago*, 38, p. 34-105.

[13] Feller, W. (1971), *An introdution to Probability Theory and its Applications*, vol.2, 2ª edição, Wiley, New York.

[14] Gnedenko, B.V. e Kolmogorov, A. N. (1954), *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley.

[15] Gross, S. e Steiger, W.L. (1979), Least Absolute Deviations Estimates in Autoregression with Infinite Variance, *Journal of Applied Probability*, 16, p. 104-116.

[16] Haan, L. de e Resnick, S. (1996), Second-Order Regular Variation and Rates of Convergence in Extreme-Value Theory, *The Annals of Probability*, 24, p. 97-124.

[17] Hannan, E.J. e Kanter, M. (1977), Autoregressive Processes with Infinite Variance, *Journal of Applied Probability*, 14, p. 411-415.

[18] Jagers, P. (1974), Aspects of Random Measures and Point Processes, *Advances in Applied Probability*, 3, P. Ney e S. Port. Marcel Dekker, New York.

[19] Kallenberg, O. (1983), *Random Measures*, Akademie Verlag, Berlin.

[20] Knight, K. (1987), Rate of Convergence of Centered Estimates Parameters for Infinite Variance Autoregressions, *Journal of Time Series Analysis*, 8, p. 51-60.

[21] Mathes, K., Kerstan J. e Mecke J. (1978), *Infinitely Divisible Point Processes*, Wiley Chichester.

[22] Mori, T. e Oodaira, H. (1976), A Funcional Law of the Iterared Logarithm for Sample Sequences, *Yokohama Mathe. Journal*, 24, p. 35-49.

[23] Resnick, S. (1986), Point Processes, Regular Variation and Weak Convergence, *Advances in Applied Probability*, 18, p. 66-138.

[24] Resnick, S. (1987), *Extreme Values, Regular Variation and Point Processes*, Springer, New York.

[25] Skorohod, A.V. (1956), Limit Theorems for Stochastics Processes, *Theor. Probability Applied*, 1, p. 261-290.

[26] Stuck, B.W. e Kleiner, B. (1974), A Statistical Analysis of Telephone Noise, *The Bell System Technical Journal*, 53, p. 1263-1320.

[27] Weissman, I. (1975a), Multivariate Extremal Processes Generated by Independent Non -Identically Distributed Random Variables, *Journal of Applied Probability*, 12, p. 477-487.

[28] Weissman, I. (1975b), On Weak Convergence of extremal Processes, *The Annals of Probability*, 4, p. 470-473.