

Carla Manuela Ribeiro Henriques

**CONVERGÊNCIA FRACA DO
PROCESSO EMPÍRICO**

Dissertação na área de Probabilidades e Estatística
do Mestrado em Matemática Aplicada
da Faculdade de Ciências do Porto

Janeiro de 1996

Orientação
Paulo Oliveira

Agradeço ao Prof. Dr. Paulo Oliveira pelo conhecimento que me proporcionou, pela disponibilidade e pelo apoio e atenção dispensados ao longo deste trabalho. Agradeço à minha família e amigos a compreensão e o incentivo que me transmitiram.

ÍNDICE

	página
Principais notações	1
INTRODUÇÃO	2
CAPÍTULO I - Definições Básicas e Noções Preliminares	4
1. Convergência fraca em espaços métricos	4
2. Convergência fraca em espaços de Hilbert separáveis	7
3. Processo empírico	12
CAPÍTULO II - Convergência Fraca do Processo Empírico em $D[0,1]$	16
1. Convergência do processo empírico em $D[0,1]$ para variáveis φ -misturadoras	16
2. Convergência do processo empírico em $D[0,1]$ para variáveis associadas	23
CAPÍTULO III - Convergência Fraca do Processo Empírico em $L^2[0,1]$	31
1. Espaços auto-reprodutores	31
2. Considerações gerais	37
3. Compacidade relativa do processo empírico	41
4. Convergência do processo empírico	48
5. Apresentação de um exemplo concreto	55
6. Algumas aplicações	64
Referências	68

Principais notações

\mathbb{R}	Conjunto dos números reais
\mathbb{N}	Conjunto dos números naturais
M^\perp	Complemento ortogonal do subconjunto M
$\langle \cdot, \cdot \rangle_2$	Produto interno em $L^2[0,1]$
$\ \cdot \ _2$	Norma definida à custa do produto interno $\langle \cdot, \cdot \rangle_2$
λ	Medida de Lebesgue

INTRODUÇÃO

A convergência fraca de medidas de probabilidade em espaços de funções, tem sido objecto de estudo de muitos autores, como por exemplo Billingsley [3], Yu [23], Oliveira ([12] a [15]), Suquet ([12] a [15]), Parthasarathy [16], entre outros. O espaço $C[0,1]$ das funções contínuas e o espaço $D[0,1]$ das funções contínuas à direita e com limites à esquerda munido da topologia de Skorohod, são sem dúvida os mais utilizados para este estudo. No entanto, tanto um como outro apresentam algumas desvantagens: o espaço $C[0,1]$ não permite o estudo da convergência de funções aleatórias com descontinuidades; o espaço $D[0,1]$ não sendo um grupo topológico (e por isso não é também espaço vectorial), não permite a adição pontual de funções. Para além disso, uma condição importante para estabelecer a convergência fraca de uma sucessão de medidas de probabilidade, é a compacidade relativa desta sucessão, condição esta que é, muitas vezes, difícil de verificar nos espaços mencionados.

Neste trabalho pretende-se estudar a convergência fraca de um processo estocástico muito particular - o processo empírico. Para isso vamos considerá-lo como uma sucessão de funções aleatórias num outro espaço, o espaço $L^2[0,1]$ das funções de quadrado integráveis à Lebesgue. Apesar disso, apresentamos no capítulo II deste volume, o estudo desenvolvido por dois autores Billingsley [3] e Yu [23], relativo à convergência do processo empírico no espaço $D[0,1]$. Esta apresentação servirá para dar uma ideia de como se trata este problema no espaço $D[0,1]$, e também para podermos estabelecer alguns aspectos comparativos deste estudo com o estudo realizado no espaço $L^2[0,1]$. Veremos, por exemplo, um processo empírico particular, cuja convergência fraca poderá ser estabelecida usando os resultados em $L^2[0,1]$, mas para o qual os resultados em $D[0,1]$ não poderão ser utilizados. Veremos também que o espaço $L^2[0,1]$ se adapta com facilidade ao estudo de algumas funcionais estatísticas importantes. Por exemplo, o teste estatístico de Cramer-Von Mises é, de facto, a norma em $L^2[0,1]$ do processo empírico, e portanto os resultados assintóticos seguem-se da convergência em $L^2[0,1]$ deste processo. Algumas das funcionais de

Von Mises constituem outro exemplo de aplicação para o qual é suficiente a convergência do processo empírico em $L^2[0,1]$. Esta análise é feita no capítulo III, onde também se apresenta o estudo da convergência do processo empírico em $L^2[0,1]$. Este capítulo é baseado no artigo [13] de Oliveira e Suquet. No capítulo I encontram-se as noções e resultados necessários à compreensão dos capítulos seguintes.

CAPÍTULO I

Definições Básicas e Noções Preliminares

1. Convergência fraca em espaços métricos

Com o objectivo de inserir o leitor no campo de trabalho, apresentamos nesta secção um resumo dos resultados básicos, relativos à convergência fraca de sucessões de medidas de probabilidade em espaços métricos. Pretendemos também, estabelecer condições suficientes para a convergência fraca num espaço métrico particular - o espaço $D[0,1]$ com a topologia de Skorohod. Estes assuntos são tratados com detalhe nos capítulos 1 e 3 de (Billingsley [3]). Aqui limitamo-nos a enunciar alguns resultados e definições, com o único objectivo de facilitar a compreensão do capítulo II.

No que se segue, S representa um espaço métrico qualquer e \mathbf{S} a tribo de Borel que lhe está associada.

Definição 1.1: Sejam $P_n, n \in \mathbb{N}$, e P medidas de probabilidade em (S, \mathbf{S}) . Dizemos que $\{P_n\}$ converge fracamente para P e escrevemos $P_n \xrightarrow{w} P$, se $\int f dP_n$ converge para $\int f dP$, para toda a função f real contínua e limitada definida em S .

Consideramos agora uma sucessão $\{X_n\}$ de elementos aleatórios em (S, \mathbf{S}) e $\{P_n\}$ a sucessão das respectivas distribuições.

Definição 1.2: Dizemos que $\{X_n\}$ converge em distribuição para o elemento aleatório X e escrevemos $X_n \xrightarrow{D} X$, se $\{P_n\}$ converge fracamente para a distribuição de X .

Tendo em conta a definição anterior, todos os resultados que se seguem, relativos à convergência fraca, podem ser rescritos em termos de convergência em distribuição. Por isso, a expressão “convergência fraca” será muitas vezes utilizada tanto para medidas como para elementos aleatórios.

Seja agora h uma função mensurável de S num outro espaço métrico S' . Representamos por Ph^{-1} a medida imagem de P por h . É fácil verificar que se h for uma função contínua e se $\{P_n\}$ for uma sucessão de medidas de probabilidade fracamente convergente para P em (S, \mathcal{S}) , então $P_n h^{-1} \xrightarrow{w} Ph^{-1}$ em (S', \mathcal{S}') . O seguinte teorema permite-nos enfraquecer a condição de continuidade de h .

Teorema 1.1 [3]: Seja h uma função mensurável de S em S' e $\{P_n\}$ uma sucessão de medidas convergindo fracamente para P em (S, \mathcal{S}) . Se o conjunto de descontinuidades de h tiver medida P nula, então a sucessão $\{P_n h^{-1}\}$ converge fracamente para Ph^{-1} em (S', \mathcal{S}') .

Introduzimos de seguida dois conceitos básicos para o estudo da convergência fraca em espaços métricos.

Definição 1.3: Um conjunto Π de medidas de probabilidade em (S, \mathcal{S}) é relativamente compacto se, de toda a sucessão de elementos de Π , se pode extrair uma subsucessão fracamente convergente para alguma medida P , não necessariamente um elemento de Π .

Definição 1.4: Um conjunto Π de medidas de probabilidade em (S, \mathcal{S}) é fino se, para todo o ε positivo existe um conjunto compacto K tal que $P(K) > 1 - \varepsilon$, para todo o P de Π .

O seguinte teorema devido a Prokhorov relaciona os dois conceitos anteriores.

Teorema 1.2 [18]: Toda o conjunto fino de medidas de probabilidade em (S, \mathcal{S}) é relativamente compacto. Se S é separável e completo, então um conjunto de medidas de probabilidade em (S, \mathcal{S}) é fino se e só se for relativamente compacto.

Notemos que no presente trabalho consideramos apenas espaços métricos completos e separáveis, portanto sempre que nos referimos a um dos conceitos o outro deve ser entendido como equivalente.

Particularizemos agora o nosso estudo, considerando o espaço métrico $D[0,1]$. Definimos para cada $k \in \mathbb{N}$ e $t_1, \dots, t_k \in [0,1]$ a projecção Π_{t_1, \dots, t_k} de $D[0,1]$ em \mathbb{R}^k da forma habitual:

$$\Pi_{t_1, \dots, t_k}(x) = (x(t_1), \dots, x(t_k)).$$

Chamamos distribuições finitas de P às medidas imagem de P por Π_{t_1, \dots, t_k} e representamo-las por $P \Pi_{t_1, \dots, t_k}^{-1}$. Por Billingsley [3], se T_0 é um conjunto denso em $D[0,1]$ que contém $\{1\}$, então P é completamente determinada pelas suas distribuições finitas $P \Pi_{t_1, \dots, t_k}^{-1}$ com $t_1, \dots, t_k \in T_0$. Isto significa que não existe em $D[0,1]$ uma outra medida de probabilidade com aquelas distribuições finitas.

O teorema que se segue estabelece condições suficientes para a convergência fraca de uma sucessão de medidas de probabilidade em $D[0,1]$. Representamos por T_P o conjunto dos pontos $t \in [0,1]$ para os quais Π_t é contínua, excepto num conjunto de medida P nula. Notemos que os pontos 0 e 1 estão sempre em T_P , pois Π_0 e Π_1 são contínuas em $D[0,1]$

Teorema 1.3 [3]: Se $\{P_n\}$ é uma sucessão fina de medidas de probabilidade em $D[0,1]$ e se $P_n \Pi_{t_1, \dots, t_k}^{-1}$ converge fracamente para $P \Pi_{t_1, \dots, t_k}^{-1}$ para cada $k \in \mathbb{N}$ e $t_1, \dots, t_k \in T_P$, então $\{P_n\}$ converge fracamente para P .

Por este teorema, para demonstrar a convergência fraca de uma sucessão $\{P_n\}$, é suficiente mostrar que $\{P_n\}$ é fina e que as suas distribuições finitas convergem fracamente. O seguinte teorema estabelece condições suficientes para a fineza de uma sucessão $\{P_n\}$. Trata-se de uma versão do teorema 15.5 de Billingsley [3].

Teorema 1.4 [3]: Seja $\{P_n\}$ uma sucessão de medidas de probabilidade em $D[0,1]$. Suponhamos que:

$$(i) \forall \eta > 0 \exists a \in \mathbb{R} : P_n \{x : |x(0)| > a\} \leq \eta, \text{ para } n \geq 1;$$

$$(ii) \forall \varepsilon, \eta > 0 \exists \delta \in (0,1), \exists n_0 \in \mathbb{N} : P_n \{x : \sup_{t \leq s \leq t+\delta} |x(s)-x(t)| \geq \varepsilon\} \leq \eta \delta, \text{ para } n \geq n_0 \text{ e}$$

para todo $t \in [0,1]$. Então, a sucessão $\{P_n\}$ é fina.

Salientemos mais uma vez, que estes resultados podem ser traduzidos em termos de convergência em distribuição de uma sucessão de elementos aleatórios $\{X_n\}$, tendo em conta que esta sucessão se diz fina se a sucessão das respectivas distribuições o for. No teorema 1.3 a convergência fraca das distribuições finitas é equivalente à convergência em distribuição da sucessão da vectores aleatórios $(X_n(t_1), \dots, X_n(t_k))$.

2. Convergência fraca em espaços de Hilbert separáveis

Para o estudo a apresentar no capítulo III, referente à convergência do processo empírico em $L^2[0,1]$, necessitamos de estabelecer condições suficientes para a convergência fraca num espaço de Hilbert separável. Nesta secção expomos um resumo sobre este assunto, para o qual nos baseamos em Parthasarathy [16].

Representamos por X um espaço de Hilbert separável e por $M(X)$ o espaço das medidas de probabilidade sobre a tribo de Borel \mathbf{B}_X de X .

Definição 2.1: Para cada P de $M(X)$, a função característica de P , $\varphi_P(x)$ com $x \in X$, é definida por:

$$\varphi_P(x) = \int e^{i\langle x, y \rangle} P(dy) \quad , \quad x \in X \quad ,$$

onde $\langle \cdot, \cdot \rangle$ representa o produto interno em X .

As funções características, caracterizam completamente as medidas a que estão associadas, isto é, não pode haver duas medidas diferentes em $M(X)$ com iguais funções características (pag. 152 de [16]).

O seguinte teorema fornece condições suficientes para a convergência fraca de uma sucessão de medidas $\{P_n\}$ em $M(X)$.

Teorema 2.1 [16]: Se a sucessão $\{P_n\}$ de medidas em $M(X)$ é relativamente compacta e se $\varphi_{P_n}(x) \xrightarrow{n} \varphi(x)$ para todo o $x \in X$, então existe uma medida P em $M(X)$ tal que $\varphi_P(x) = \varphi(x)$ para todo o $x \in X$ e $P_n \xrightarrow{w} P$.

Demonstração:

Por hipótese, a sucessão é relativamente compacta, logo toda a subsucessão de $\{P_n\}$ contém outra subsucessão fracamente convergente para uma medida em $M(X)$. Para demonstrar que $\{P_n\}$ é fracamente convergente, temos de provar que todas as subsucessões de $\{P_n\}$ convergem fracamente para o mesmo limite. Vamos fazê-lo por redução ao absurdo. Suponhamos que $\{P_n\}$ não é fracamente convergente. Então, existem duas subsucessões $\{P_{n'}\}$ e $\{P_{n''}\}$ que convergem fracamente para duas medidas diferentes, P_1 e P_2 em $M(X)$. Por definição de convergência fraca vem:

$$\int e^{i\langle x,y \rangle} P_{n'}(dy) \xrightarrow{n} \int e^{i\langle x,y \rangle} P_1(dy) , \quad \forall x \in X$$

e

$$\int e^{i\langle x,y \rangle} P_{n''}(dy) \xrightarrow{n} \int e^{i\langle x,y \rangle} P_2(dy) , \quad \forall x \in X ,$$

isto é:

$$\varphi_{P_{n'}}(x) \xrightarrow{n} \varphi_{P_1}(x) , \quad \forall x \in X$$

e

$$\varphi_{P_{n''}}(x) \xrightarrow{n} \varphi_{P_2}(x) , \quad \forall x \in X .$$

Por hipótese $\varphi_{P_n}(x) \xrightarrow{n} \varphi(x)$ para todo o x de X . Então $\varphi_{P_1}(x) = \varphi_{P_2}(x) = \varphi(x)$ para todo o x de X . Como as funções características caracterizam as medidas, vem $P_1 = P_2$ em $M(X)$ o que é uma contradição. Fica assim provado o teorema. ■

Definimos função característica de um elemento aleatório Z de (X, \mathbf{B}_X) , como a

função característica da sua distribuição P_Z em $M(X)$. Assim:

$$\varphi_Z(x) = \varphi_{P_Z}(x) = E\{ e^{i\langle Z, x \rangle} \} , x \in X ,$$

donde concluímos que os produtos internos $\langle Z, x \rangle$, $x \in X$, caracterizam completamente a distribuição P_Z de Z . Notemos ainda, que se $\{Z_n\}$ for uma sucessão de elementos aleatórios em (X, \mathbf{B}_X) , a convergência pontual das funções características para a função característica de um elemento aleatório Z , é equivalente à convergência em distribuição dos produtos internos $\langle Z_n, x \rangle$ para $\langle Z, x \rangle$, $\forall x \in X$. Sendo assim, pelo teorema 2.1, para estabelecer a convergência em distribuição da sucessão $\{Z_n\}$, basta provar a convergência em distribuição das variáveis aleatórias $\langle Z_n, x \rangle$ com $x \in X$ e a compacidade relativa de $\{Z_n\}$.

Para o estudo da compacidade relativa iremos utilizar a condição de Prokhorov (teorema 1.13 de [18]), convenientemente modificada para evitar contra-exemplos triviais.

Teorema 2.2 [20]: Seja $(e_i)_{i \in \mathbb{N}}$ uma base ortonormada de X e Π um subconjunto de $M(X)$. Para todo o x de X definimos:

$$r_N^2(x) = \sum_{i=N}^{\infty} \langle x, e_i \rangle^2.$$

Suponhamos que :

$$(i) \sup_{P \in \Pi} \int \|x\|^2 dP(x) < +\infty ;$$

$$(ii) \lim_{N \rightarrow \infty} \sup_{P \in \Pi} \int r_N^2(x) dP(x) = 0.$$

Então Π é um subconjunto relativamente compacto de $M(X)$.

Demonstração:

Seja $\phi(N) = \sup_{P \in \Pi} \int r_N^2(x) dP(x)$. Por (i) $\phi(0) < +\infty$ e por (ii) $\lim_{N \rightarrow \infty} \phi(N) = 0$, o que

impede que $\phi(N) = +\infty$ para os primeiros valores de N .

Seja $\varepsilon > 0$, escolhamos uma sucessão de inteiros $(N_k)_{k \in \mathbb{N}}$ tal que $N_1 = 0$ e uma sucessão de reais estritamente positivos $(\Delta_k)_{k \in \mathbb{N}}$, tais que: $\sum_{k=1}^{\infty} \Delta_k \phi(N_k) < \varepsilon$.

Seja $K = \bigcap_{k=1}^{\infty} E_k$ com $E_k = \{x \in X : r_{N_k}^2(x) \leq \Delta_k^{-1}\}$. Vamos ver que K assim definido é um conjunto compacto. Notemos, para já, que K sendo intersecção de fechados é um conjunto fechado. Além disso, atendendo a que $N_1=0$ e representando por $\overline{B}(0, \Delta_1^{-1/2})$ a bola fechada de centro em zero e raio $\Delta_1^{-1/2}$, temos: $K \subset E_1 = \overline{B}(0, \Delta_1^{-1/2})$. Portanto K é um conjunto limitado.

Fixemos $m \in \mathbb{N}$ arbitrariamente. Seja M o subespaço de dimensão finita gerado pelos primeiros N_m vectores da base ortonormada $(e_i)_{i \in \mathbb{N}}$. Consideremos a aplicação, projecção ortogonal $P: X \rightarrow M$, que sabemos ser um operador linear limitado, portanto, transforma conjuntos limitados em conjuntos limitados. Representemos por K_1 a projecção ortogonal de K sobre M . Pelo que acabamos de dizer, K_1 é limitado num subespaço de dimensão finita M , então o seu fecho é compacto em M , logo em X . Isto é, K_1 é relativamente compacto em X . É sabido que um conjunto num espaço de Hilbert é relativamente compacto se e só se for totalmente limitado, isto é, se e só se para todo $\varepsilon > 0$ existir um número finito de pontos $I = \{y_1, \dots, y_k\}$ tal que as bolas abertas $B(y_i, \varepsilon)$ $i=1, \dots, k$, constituam uma cobertura finita desse conjunto. Então para K_1 existe um número finito de pontos $\{y_1, \dots, y_{k_m}\}$ tal que:

$$K_1 \subset \bigcup_{i=1}^{k_m} B(y_i, \Delta_m^{-1/2}).$$

Como $(e_i)_{i \in \mathbb{N}}$ é uma base ortonormada de X , temos para todo o $x \in K$,

$$x = \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i = x_1 + x_2 \quad \text{com} \quad x_1 = \sum_{i=1}^{N_m} \langle x, e_i \rangle e_i \in K_1$$

$$\text{e} \quad x_2 = \sum_{i=N_m+1}^{\infty} \langle x, e_i \rangle e_i.$$

Seja $K_2 = \{x_2 = \sum_{i=N_m+1}^{\infty} \langle x, e_i \rangle e_i : x \in K\}$, vamos verificar que K_2 está contido numa bola de centro zero e raio $\Delta_m^{-1/2}$. De facto, se x_2 for um elemento qualquer de K_2 , temos:

$$\|x_2\|^2 = \sum_{i=N_m+1}^{\infty} \langle x, e_i \rangle^2 \quad \text{com } x \in K.$$

Então, por definição de K , vem:

$$\|x_2\|^2 = \sum_{i=N_m+1}^{\infty} \langle x, e_i \rangle^2 \leq r_{N_m}^2(x) \leq \Delta_m^{-1},$$

pois $x \in E_m$, donde se segue que $x_2 \in \overline{B}(0, \Delta_m^{-1/2})$.

Considerando agora x um elemento arbitrário de K , temos:

$$x = x_1 + x_2, \quad \text{com } x_1 \in B(y_i, \Delta_m^{-1/2}) \text{ para algum } i \in \{1, \dots, k_m\}$$

$$\text{e } x_2 \in \overline{B}(0, \Delta_m^{-1/2}),$$

o que implica que $\|x - y_i\| \leq \|x_1 - y_i\| + \|x_2\| < 2\Delta_m^{-1/2}$. Isto é, $x \in B(y_i, 2\Delta_m^{-1/2})$ para algum $i \in \{1, \dots, k_m\}$. Então $K \subset \bigcup_{i=1}^{k_m} B(y_i, 2\Delta_m^{-1/2})$. Da escolha arbitrária de $m \in \mathbb{N}$, podemos concluir que K é totalmente limitado em X , ou seja, K é relativamente compacto em X . Como K é fechado, então fica provado que é compacto.

Seja agora $P \in \Pi$, utilizando a desigualdade de Tchebycheff vem:

$$P(X \setminus K) \leq \sum_{i=1}^{\infty} P(X \setminus E_i) = \sum_{i=1}^{\infty} P\{x \in X : r_{N_i}^2(x) > \Delta_i^{-1}\} \leq$$

$$\leq \sum_{i=1}^{\infty} \Delta_i \int r_{N_i}^2(x) dP(x) \leq \sum_{i=1}^{\infty} \Delta_i \sup_{P \in \Pi} \int r_{N_i}^2(x) dP(x)$$

$$= \sum_{i=1}^{\infty} \Delta_i \phi(N_i) \leq \varepsilon.$$

Podemos então concluir que, para todo o $\varepsilon > 0$ existe um conjunto compacto K , tal que $P(K) > 1 - \varepsilon$ para todo $P \in \Pi$, isto é, Π é um subconjunto fino de probabilidades em (X, \mathbf{B}_X) . Como X é completo, por ser de Hilbert, o conceito de fineza é equivalente ao de compacidade relativa, donde concluimos que Π é relativamente compacto. ■

3. Processo empírico

Vamos iniciar esta secção com a definição de função de distribuição empírica e de processo empírico. Para isso consideramos uma sucessão $\{X_n\}$ de variáveis aleatórias igualmente distribuídas, com função de distribuição F concentrada em $[0,1]$.

Definição 3.1: Chama-se função de distribuição empírica de X_1, \dots, X_n à função definida por:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x), \quad x \in [0,1]$$

onde $I(X_i \leq x)$ representa a indicatriz do conjunto $\{\omega \in \Omega: X_i(\omega) \leq x\}$.

Definição 3.2: O processo empírico associado à sucessão $\{X_n\}$ será representado por $\{Z_n\}$ e é definido da seguinte forma:

$$Z_n(x) = \sqrt{n} (F_n(x) - F(x)), \quad x \in [0,1].$$

Neste trabalho consideramos sempre que F é uma função de distribuição contínua. Vamos ver que no estudo da convergência do processo empírico, o caso mais importante é o do processo empírico uniforme, ou seja, o processo empírico associado a uma sucessão $\{X_n\}$ de variáveis aleatórias uniformemente distribuídas no intervalo $[0,1]$. Para tal, definimos a função Quantil $Q(y)$, $0 \leq y \leq 1$, por:

$$Q(y) = \inf\{x: F(x) \geq y\}, \quad 0 < y \leq 1$$

Notemos que, como F é uma função contínua, $F(Q(y))=y$ para todo o $y \in [0,1]$. Da continuidade de F , resulta ainda, que as variáveis aleatórias $U_n=F(X_n)$, $n \in \mathbb{N}$, são uniformemente distribuídas no intervalo $[0,1]$.

Seja então $F_n^*(y)$, $0 \leq y \leq 1$, a função de distribuição empírica de U_1, \dots, U_n e $\{Z_n^*\}$ o processo empírico associado a $\{U_n\}$. Para cada $y \in [0,1]$, temos por definição:

$$F_n^*(y) = \frac{1}{n} \sum_{i=1}^n I(U_i \leq y).$$

Como $I(U_i \leq y) = I(X_i \leq Q(y))$ com probabilidade 1, vem para cada $n \in \mathbb{N}$:

$$F_n^*(y) = F_n(Q(y)) \text{ com probabilidade 1,}$$

donde concluímos que $F_n^*(y)$ e $F_n(Q(y))$ têm a mesma distribuição: $F_n^*(y) \stackrel{D}{=} F_n(Q(y))$, para $y \in [0,1]$.

Relativamente aos processos empíricos, podemos da mesma forma concluir que, para cada $n \in \mathbb{N}$ e para cada $y \in [0,1]$, as variáveis aleatórias $Z_n^*(y)$ e $Z_n(Q(y))$ têm a mesma distribuição por serem iguais com probabilidade 1. Então, é fácil verificar que, quando consideradas como funções aleatórias em $D[0,1]$ ou em $L^2[0,1]$, $Z_n^*(\cdot)$ e $Z_n(Q(\cdot))$, vão ter a mesma distribuição. O que significa que basta estudar a convergência de $\{Z_n^*\}$, pois a partir deste é possível deduzir resultados análogos para $\{Z_n\}$.

Ao longo deste trabalho, preocupamo-nos apenas em deduzir resultados que estabeleçam a convergência fraca do processo empírico uniforme. Citamos de seguida um resultado clássico que estabelece a convergência deste processo, supondo a independência das variáveis aleatórias X_n . Este resultado é uma versão do teorema 16.4 de Billingsley [3] adaptada ao caso uniforme.

Teorema 3.1 [3]: Suponhamos que as variáveis aleatórias X_n , $n \in \mathbb{N}$, são independentes e uniformemente distribuídas em $[0,1]$. Seja $\{Z_n\}$ o processo empírico associado à sucessão $\{X_n\}$, então $Z_n \xrightarrow{D} Z$, onde Z é um elemento aleatório Gaussiano de $D[0,1]$ tal que:

$$E\{Z(t)\}=0;$$

$$\text{e } E\{Z(t)Z(s)\}=s(1-t) \text{ para } s \leq t.$$

Neste trabalho, pretendemos estabelecer a convergência do processo empírico uniforme, relaxando a condição de independência das variáveis aleatórias $\{X_n\}$. Teremos três formas de medir o grau de dependência entre as variáveis aleatórias $\{X_n\}$, que são apresentadas seguidamente.

Definição 3.3: Dizemos que uma sucessão $\{X_n\}$ de variáveis aleatórias estritamente estacionárias é φ -misturadora, se para cada $n \in \mathbb{N}$ e para cada $k \in \mathbb{N}$, se tem:

$$\sup\{ |P(B|A) - P(A)| : A \in \mathbf{F}_1^k, B \in \mathbf{F}_{k+n}^\infty \} = \varphi_n,$$

$$\text{com } \lim_{n \rightarrow \infty} \varphi_n = 0,$$

onde $P(B|A)$ é a probabilidade de B condicionada a A , \mathbf{F}_1^k a σ -álgebra gerada pelas variáveis aleatórias X_i , $i=1, \dots, k$ e \mathbf{F}_{k+n}^∞ gerada por X_i , $i=k+n, k+n+1, \dots$.

Definição 3.4: Dizemos que uma sucessão $\{X_n\}$ de variáveis aleatórias estritamente estacionárias é α -misturadora, se para cada $n \in \mathbb{N}$ e para cada $k \in \mathbb{N}$, se tem:

$$\sup\{ |P(A \cap B) - P(A)P(B)| : A \in \mathbf{F}_1^k, B \in \mathbf{F}_{k+n}^\infty \} = \alpha_n,$$

$$\text{com } \lim_{n \rightarrow \infty} \alpha_n = 0,$$

onde \mathbf{F}_1^k e \mathbf{F}_{k+n}^∞ são definidas como anteriormente.

Definição 3.5: Dizemos que uma sucessão $\{X_n\}$ de variáveis aleatórias é associada, se:

$$\text{Cov}(f(X_{i_1}, \dots, X_{i_m}), g(X_{i_1}, \dots, X_{i_m})) \geq 0,$$

para todo o subconjunto finito de índices $\{i_1, \dots, i_m\} \subset \mathbb{N}$ e para todo o par de funções f e g não decrescentes em cada variável, para as quais aquela covariância exista.

Notemos que na última definição é indiferente escolher f e g ambas não decrescentes ou ambas não crescentes em cada variável. É de notar, também, que se a sucessão $\{X_n\}$ de variáveis aleatórias reais é associada, então para cada par de índices $i, j \in \mathbb{N}$ e para quaisquer $a_i, a_j \in \mathbb{R}$, escolhendo $f(x, y) = I_{[a_i, +\infty)}(x)$ e $g(x, y) = I_{[a_j, +\infty)}(y)$, vem:

$$P(X_i > a_i, X_j > a_j) - P(X_i > a_i)P(X_j > a_j) \geq 0,$$

ou equivalentemente:

$$P(X_i \leq a_i, X_j \leq a_j) - P(X_i \leq a_i)P(X_j \leq a_j) \geq 0.$$

CAPÍTULO II

Convergência Fraca do Processo Empírico em $D[0,1]$

Neste capítulo vamos fazer o estudo da convergência do processo empírico uniforme $\{Z_n\}$, encarando-o como uma sucessão de funções aleatórias no espaço de Skorohod $D[0,1]$. Para isso, vamos considerar dois graus de dependência entre as variáveis aleatórias X_n , $n \in \mathbb{N}$, às quais está associado o processo empírico uniforme $\{Z_n\}$: φ -misturação e associação (ver secção 3 do capítulo I). Dividimos este capítulo em duas partes: a primeira trata o problema para variáveis aleatórias φ -misturadoras, para a qual nos baseamos em Billingsley [3]; o estudo apresentado na segunda parte é devido a Yu [23] e trata o problema para variáveis aleatórias associadas.

Não se pretende explorar este assunto exhaustivamente, mas apenas, dar uma ideia das técnicas utilizadas neste espaço para estabelecer a convergência fraca do processo empírico uniforme. Teremos oportunidade de verificar que estas técnicas são bastante mais complicadas do que as utilizadas no espaço $L^2[0,1]$.

1. Convergência do processo empírico em $D[0,1]$ para variáveis φ -misturadoras

Nesta secção consideramos sempre, salvo menção em contrário, uma sucessão $\{X_n\}$ φ -misturadora de variáveis aleatórias estritamente estacionárias. Pretendemos obter condições suficientes para a convergência do processo empírico uniforme $\{Z_n\}$. Começamos por enunciar alguns resultados que serão necessários para estabelecer a convergência deste processo no espaço $D[0,1]$.

O teorema que se segue é devido a Billingsley.

Teorema 1.1 (teorema 20.1 de [3]): Suponhamos que a sucessão $\{X_n\}$ é estritamente estacionária e φ -misturadora, cujos coeficientes de mistura φ_n satisfazem a: $\sum_{i=1}^{\infty} \varphi_n^{1/2} < \infty$. Se X_0 tiver esperança nula e variância finita, então a série em:

$$\sigma^2 = E\{X_0^2\} + 2 \sum_{i=1}^{\infty} E\{X_0 X_i\}$$

é absolutamente convergente. Se ainda $\sigma^2 > 0$, então a sucessão de elementos aleatórios de $D[0,1]$ definida por:

$$Y_n(t) = \frac{1}{\sigma\sqrt{n}} S_{[nt]} \quad , \quad 0 \leq t \leq 1 \quad ,$$

onde $[nt]$ representa o maior inteiro não superior a nt e $S_{[nt]} = X_1 + \dots + X_{[nt]}$, converge em distribuição para o movimento Browniano W .

Deste resultado decorre um teorema limite central, invocando a continuidade da projecção Π_1 (ver secção 1 do capítulo I). Isto é, nas condições do teorema anterior a variável aleatória $\frac{1}{\sqrt{n}}(X_1 + \dots + X_n)$ converge em distribuição para uma variável aleatória Gaussiana centrada com variância dada por:

$$\sigma^2 = E\{X_0^2\} + 2 \sum_{i=1}^{\infty} E\{X_0 X_i\}.$$

É de salientar que Ibragimov tinha demonstrado este teorema limite central em [7]. Mais, Davydov demonstrou em [5] o mesmo resultado que Billingsley independentemente dele.

O teorema de Cramér-Wold (teorema 7.7 de [3]) permite reduzir a convergência em distribuição de vectores aleatórios à convergência em distribuição das variáveis aleatórias obtidas por combinação linear das coordenadas dos vectores. Aplicando

este resultado prova-se que, se $\{X_n^{(1)}\}, \dots, \{X_n^{(r)}\}$ forem r sucessões de variáveis aleatórias que verificam as condições do teorema 1.1, então a sucessão de vectores aleatórios $\frac{1}{\sqrt{n}} \sum_{k=1}^{\infty} (X_k^{(1)}, \dots, X_k^{(r)})$ converge em distribuição para um vector aleatório

Gaussiano centrado com covariância dada por:

$$\sigma_{ij} = E\{X_0^{(i)} X_0^{(j)}\} + \sum_{k=1}^{\infty} E\{X_0^{(i)} X_k^{(j)}\} + \sum_{k=1}^{\infty} E\{X_k^{(i)} X_0^{(j)}\},$$

onde as séries são absolutamente convergentes.

Enunciamos de seguida um lema que estabelece uma majoração que utilizaremos para demonstrar a convergência fraca do processo empírico uniforme.

Lema 1.1 (Billingsley [3], lema 1, pag.195): Suponhamos que $\{X_n\}$ é uma sucessão φ -misturadora de variáveis aleatórias estritamente estacionárias, tal que

$|X_0| \leq 1$ com probabilidade 1, $E\{X_0\} = 0$ e $\sum_{k=0}^{\infty} k^2 \varphi_k^{1/2} < \infty$. Então:

$$E\{S_n^4\} \leq K_1 [n^2 E\{X_0^2\} + n E\{X_0^2\}] \left[\sum_{k=0}^{\infty} (k+1)^2 \varphi_k^{1/2} \right]^2, \quad (1.1)$$

onde K_1 é uma constante positiva.

Para dar-mos uma ideia de como este lema se demonstra, necessitamos da desigualdade estabelecida no lema seguinte.

Lema 1.2 (Billingsley [3], lema 1, pag. 170): Seja $\{X_n\}$ uma sucessão φ -misturadora de variáveis aleatórias estritamente estacionárias. Suponhamos que Y_1 é uma variável aleatória mensurável relativamente a F_1^k e Y_2 mensurável relativamente a F_{k+n}^{∞} ($n \geq 0$). Se $E\{|Y_1|^r\} < \infty$ e $E\{|Y_2|^s\} < \infty$ onde $r, s > 1$ e $\frac{1}{r} + \frac{1}{s} = 1$, então:

$$|E\{Y_1 Y_2\} - E\{Y_1\}E\{Y_2\}| \leq 2 \varphi_n^{1/r} E^{1/r}\{|Y_1|^r\} E^{1/s}\{|Y_2|^s\}.$$

Tendo em conta a estacionaridade da sucessão $\{X_n\}$ é fácil verificar que:

$$E\{S_n^4\} \leq 4! n \sum_{\substack{i,j,k \geq 0 \\ i+j+k \leq n}} |E\{X_0 X_i X_{i+j} X_{i+j+k}\}|. \quad (1.2)$$

As três desigualdades que se seguem decorrem da desigualdade estabelecida no lema 1.2.

$$|E\{X_0 (X_i X_{i+j} X_{i+j+k})\}| \leq 2\varphi_i^{1/2} E\{X_0^2\}; \quad (1.3)$$

$$|E\{(X_0 X_i X_{i+j} X_{i+j+k})\}| \leq 2\varphi_k^{1/2} E\{X_0^2\}; \quad (1.4)$$

$$|E\{(X_0 X_i) (X_{i+j} X_{i+j+k})\}| \leq 4\varphi_i^{1/2} \varphi_k^{1/2} E^2\{X_0^2\} + 2\varphi_j^{1/2} E\{X_0^2\}. \quad (1.5)$$

Aplicando (1.3), (1.4) e (1.5) de forma adequada a (1.2) obtêm-se:

$$E\{S_n^4\} \leq K n \left(E^2\{X_0^2\} \sum_{i,k \leq j} \varphi_i^{1/2} \varphi_k^{1/2} + 3 E\{X_0^2\} \sum_{j,k \leq i} \varphi_i^{1/2} \right),$$

onde K é uma constante positiva e os índices dos somatórios obedecem a: $i, j, k \geq 0$ e $i+j+k \leq n$. A desigualdade (1.1) segue-se imediatamente da desigualdade anterior e do facto da série envolvida ser convergente.

Os resultados anteriores vão-nos permitir estabelecer condições suficientes para a convergência fraca do processo empírico uniforme $\{Z_n\}$. O teorema que se segue é uma versão do teorema 22.1 de Billingsley [3] adaptada ao caso uniforme.

Teorema 1.2: Seja $\{X_n\}$ uma sucessão φ -misturadora e estritamente estacionária de variáveis aleatórias uniformemente distribuídas em $[0,1]$, cujos coeficientes de mistura φ_n satisfazem a: $\sum_{n=1}^{\infty} n^2 \varphi_n^{1/2} < \infty$. Então, o processo empírico uniforme $\{Z_n\}$ converge em distribuição para um elemento aleatório Z em $D[0,1]$, Gaussiano centrado com função de covariância dada por:

$$\Gamma(s,t) = E\{g_s(X_0) g_t(X_0)\} + \sum_{k=1}^{\infty} E\{g_s(X_0) g_t(X_k)\} + \sum_{k=1}^{\infty} E\{g_s(X_k) g_t(X_0)\}, \quad (1.6)$$

onde $g_t(x) = I_{[0,t]}(x) - x$ e as séries são absolutamente convergentes.

Demonstração:

Já sabemos que a convergência fraca de $\{Z_n\}$ para Z em $D[0,1]$, decorre da convergência em distribuição de $(Z_n(t_1), \dots, Z_n(t_k))$ para $(Z(t_1), \dots, Z(t_k))$ para cada subconjunto finito $\{t_1, \dots, t_k\} \subset [0,1]$ e da fineza da sucessão $\{Z_n\}$. Resolvemos primeiro o problema da convergência em distribuição dos vectores $(Z_n(t_1), \dots, Z_n(t_k))$. Para isso, rescrevemos, para $s \in [0,1]$, $Z_n(s)$ da seguinte forma:

$$Z_n(s) = \frac{1}{\sqrt{n}} \sum_{i=1}^n g_s(X_i).$$

As hipóteses deste teorema garantem que as sucessões $\{g_{t_1}(X_n)\}, \dots, \{g_{t_k}(X_n)\}$ verificam as condições do teorema 1.1. Logo, atendendo à versão multidimensional deste teorema, a sucessão de vectores aleatórios

$$(Z_n(t_1), \dots, Z_n(t_k)) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (g_{t_1}(X_i), \dots, g_{t_k}(X_i))$$

converge em distribuição para um vector aleatório Gaussiano centrado com covariância dada por:

$$\sigma_{ij} = E\{g_{t_i}(X_0) g_{t_j}(X_0)\} + \sum_{k=1}^{\infty} E\{g_{t_i}(X_0) g_{t_j}(X_k)\} + \sum_{k=1}^{\infty} E\{g_{t_i}(X_k) g_{t_j}(X_0)\},$$

onde as séries são absolutamente convergentes. Notemos que $\sigma_{ij} = \Gamma(t_i, t_j)$ onde Γ é dada por (1.6).

Resta-nos demonstrar que $\{Z_n\}$ é uma sucessão fina em $D[0,1]$. Para isso vamos utilizar o teorema 1.4 do capítulo I. Isto é, precisamos de mostrar que as duas condições seguintes se verificam:

$$(C1): \forall \eta > 0 \exists a \in \mathbb{R} : P\{|Z_n(0)| > a\} \leq \eta, \text{ para } n \geq 1;$$

$$(C2): \forall \varepsilon, \eta > 0 \exists \delta \in (0,1), \exists n_0 \in \mathbb{N} : P\{\sup_{s \leq t \leq s+\delta} |Z_n(t) - Z_n(s)| \geq \varepsilon\} \leq \eta \delta, \text{ para } n \geq n_0 \text{ e}$$

para todo $t \in [0,1]$.

A condição (C1) é facilmente verificada pois para cada $n \in \mathbb{N}$, $Z_n(0) = 0$ com probabilidade 1.

Para verificar a condição (C2), fixemos $\varepsilon, \eta > 0$ arbitrariamente. Como X_0 é uniformemente distribuída em $[0, 1]$, temos:

$$E\{|g_t(X_0) - g_s(X_0)|^2\} \leq |t-s|. \quad (1.7)$$

Aplicando o lema 1.1 à sucessão $\{g_t(X_n) - g_s(X_n)\}$, segue-se que:

$$E\left\{\left|\sum_{i=1}^n (g_t(X_i) - g_s(X_i))\right|^4\right\} \leq K_1 \left[n^2 E^2\{|g_t(X_0) - g_s(X_0)|^2\} + n E\{|g_t(X_0) - g_s(X_0)|^2\}\right] \times \\ \times \left[\sum_{k=1}^{\infty} (k+1)^2 \varphi_k^{1/2}\right]^2,$$

onde K_1 é uma constante positiva. No que se segue K_1 denotará sempre uma constante positiva, embora possa assumir valores diferentes ao longo da demonstração. Por hipótese a série de termos não negativos $\sum_{n=1}^{\infty} n^2 \varphi_n^{1/2}$ é convergente, o que implica que a série envolvida na última desigualdade também converge. Então, por (1.7) vem:

$$E\left\{\left|\sum_{i=1}^n (g_t(X_i) - g_s(X_i))\right|^4\right\} \leq K_1(n^2(t-s)^2 + n|t-s|),$$

onde K_1 depende apenas dos coeficientes φ_n .

Se $\frac{\varepsilon}{n} \leq t-s$ (assumimos que $\varepsilon < 1$) então:

$$E\{|Z_n(t) - Z_n(s)|^4\} = \frac{1}{n^2} E\left\{\left|\sum_{i=1}^n (g_t(X_i) - g_s(X_i))\right|^4\right\} \leq \frac{2K_1}{\varepsilon}(t-s)^2. \quad (1.8)$$

Seja agora $p \in [0, 1]$ tal que $\frac{\varepsilon}{n} \leq p$. Consideramos as variáveis aleatórias $Z_n(s+ip) - Z_n(s+(i-1)p)$ com $i=1, \dots, m$, sendo m um inteiro positivo. Aplicamos o

teorema 12.2 de Billingsley [3] sendo as condições deste teorema verificadas por (1.8) tomando $\gamma = 4$, $\alpha = 2$ e $u_i = p(2 K_1)^{1/2} \varepsilon^{-1/2}$, $i=1, \dots, m$. Então por este teorema vem, para todo o $\lambda > 0$:

$$P\left\{ \max_{i \leq m} |Z_n(s+ip) - Z_n(s)| \geq \lambda \right\} \leq \frac{K_1}{\varepsilon \lambda^4} m^2 p^2. \quad (1.9)$$

Em seguida vamos verificar que:

$$|Z_n(t) - Z_n(s)| \leq |Z_n(s+p) - Z_n(s)| + p\sqrt{n}, \quad s \leq t \leq s+p. \quad (1.10)$$

Para isso, tomamos $s = 0$ para simplificar as notações. Seja $U_n(t)$ o número de X_i 's entre X_1, \dots, X_n que verificam $X_i \leq t$, isto é: $U_n(t) = \sum_{i=1}^n I(X_i \leq t)$. Então (1.10) para $s=0$ é equivalente a:

$$|U_n(t) - nt| \leq |U_n(p) - np| + np, \quad 0 \leq t \leq p, \quad (1.11)$$

já que $\sqrt{n} Z_n(t) = U_n(t) - nt$, para cada $t \in [0, 1]$ e tanto $Z_n(0)$ como $U_n(0)$ são nulos com probabilidade 1. Para provar (1.10) basta então verificar (1.11). Ora, por definição de $U_n(t)$ vem:

$$U_n(t) - nt \leq U_n(p) - np = U_n(p) - np + n(p-t) \leq |U_n(p) - np| + np$$

e

$$U_n(t) - nt \geq -nt \geq -np \geq -np - |U_n(p) - np|,$$

donde sai (1.11).

Por (1.10) e para $t \in [s+(i-1)p, s+ip]$ para algum $i=1, \dots, m$, vem:

$$\begin{aligned} |Z_n(t) - Z_n(s)| &\leq |Z_n(t) - Z_n(s+(i-1)p)| + |Z_n(s+(i-1)p) - Z_n(s)| \\ &\leq |Z_n(s+ip) - Z_n(s+(i-1)p)| + p\sqrt{n} + |Z_n(s+(i-1)p) - Z_n(s)| \\ &\leq |Z_n(s+ip) - Z_n(s)| + 2|Z_n(s+(i-1)p) - Z_n(s)| + p\sqrt{n}, \end{aligned}$$

donde se segue que:

$$\sup_{s \leq t \leq s+mp} |Z_n(t) - Z_n(s)| \leq 3 \max_{i \leq m} |Z_n(s+ip) - Z_n(s)| + p\sqrt{n}. \quad (1.12)$$

Se $\frac{\varepsilon}{n} \leq p < \frac{\varepsilon}{\sqrt{n}}$, então (1.9) verifica-se e segue-se da desigualdade anterior que:

$$P\left\{ \sup_{s \leq t \leq s+mp} |Z_n(t) - Z_n(s)| \geq 4\varepsilon \right\} \leq P\left\{ \max_{i \leq m} |Z_n(s+ip) - Z_n(s)| \geq \varepsilon \right\} \leq \frac{K_1 m^2}{\varepsilon^5} p^2.$$

Escolhemos δ de forma a que $\frac{K_1 \delta}{\varepsilon^5} < \eta$. Para n suficientemente grande, existe um inteiro m tal que $(\delta/\varepsilon)\sqrt{n} < m \leq (\delta/\varepsilon)n$. Isto é equivalente a dizer que existem p e m tais que $\varepsilon/n \leq p < \varepsilon/\sqrt{n}$ e $mp = \delta$. Então da desigualdade anterior vem:

$$P\left\{ \sup_{s \leq t \leq s+\delta} |Z_n(t) - Z_n(s)| \geq 4\varepsilon \right\} \leq \eta\delta,$$

que é a menos de uma constante a condição (C2). Concluimos então que $\{Z_n\}$ é uma sucessão fina terminando assim a demonstração. ■

Para terminar esta secção, notemos que a função de covariância de Z dada por (1.6), pode ser rescrita da seguinte maneira:

$$\Gamma(s,t) = \min\{s,t\} - s t + \sum_{k=1}^{\infty} \text{Cov}(I(X_0 \leq s), I(X_k \leq t)) + \sum_{k=1}^{\infty} \text{Cov}(I(X_0 \leq t), I(X_k \leq s)).$$

2. Convergência do processo empírico em $D[0,1]$ para variáveis associadas

Nesta secção consideramos $\{X_n\}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas e uniformemente distribuídas em $[0,1]$. Para

demonstrar o teorema que estabelece a convergência fraca do processo $\{Z_n\}$ associado à sucessão $\{X_n\}$, necessitaremos dos resultados que se seguem.

Lema 2.1 [3]: Se $\{X_n\}$ é uma sucessão estritamente estacionária de variáveis aleatórias associadas e uniformemente distribuídas em $[0,1]$, e se existe uma constante v tal que:

$$\sum_{n=1}^{\infty} n^{13/2+v} \text{Cov}(X_0, X_n) < \infty,$$

então, para todo o $n \geq 1$, tem-se:

$$E \left\{ \sum_{i=1}^n (I(s < X_i \leq t) - (t-s)) \right\}^4 \leq K_1 n^2 (n^{-1/2-v_1} + (t-s)^{6/5}),$$

onde $v_1 = \min\{v/3, 1/5\}$ e K_1 é uma constante positiva.

Para demonstrar este lema, Yu recorre a uma série de outros resultados, cujas demonstrações envolvem manipulações bastante técnicas, e onde se obtêm majorações de esperanças e desigualdades sobre probabilidades de conjuntos, à custa essencialmente da estacionaridade e do facto de que, por associação, para todo o $i, j \in \mathbb{N}$ e $a_i, a_j \in \mathbb{R}$, $P(X_i > a_i, X_j > a_j) - P(X_i > a_i) P(X_j > a_j) \geq 0$.

A desigualdade estabelecida no lema seguinte, foi demonstrada por Yu em [23] e mais tarde, com argumentos mais simples, por Oliveira e Suquet em [14]. A demonstração que expomos aqui é devida a estes dois últimos autores.

Lema 2.2 [14]: Se U e V são variáveis aleatórias associadas e uniformemente distribuídas em $[0,1]$, então:

$$\text{Cov}(I(U \leq s), I(V \leq t)) \leq (3/2)^{1/3} \text{Cov}^{1/3}(U, V), \quad (s, t) \in [0,1]^2.$$

Demonstração:

Seja $g(s,t)=\text{Cov}(I(U\leq s),I(V\leq t))$ com $(s,t) \in [0,1]^2$. Esta função é não negativa por associação. Pela igualdade de Hoeffding [10] temos:

$$\text{Cov}(U,V) = \int_{[0,1]^2} g(s,t) ds dt. \quad (2.1)$$

Tendo em conta que as margens do vector aleatório (U,V) são uniformemente distribuídas em $[0,1]$, é fácil verificar que, para s,s',t e t' em $[0,1]$, vem:

$$|g(s,t) - g(s',t')| \leq |s-s'| + |t-t'|.$$

A função g é contínua num compacto, logo atinge aí um máximo m . Seja (s_0,t_0) o ponto onde g atinge o seu máximo. Utilizando a desigualdade anterior, vem para todo $(s,t) \in [0,1]^2$:

$$g(s,t) \geq m - |s-s_0| - |t-t_0|.$$

Usando esta desigualdade no quadrado

$$S = \{(s,t) \in [0,1]^2: |s-s_0| + |t-t_0| \leq m\},$$

encontramos um minorante para o integral em (2.1). De facto,

$$\text{Cov}(U,V) \geq \int_S g(s,t) ds dt \geq \int_S m - |s-s_0| - |t-t_0| ds dt = 1/3 (m\sqrt{2})^2 m,$$

pois o último integral é o volume de uma pirâmide de base S e vértice (s_0,t_0,m) . Concluimos o que pretendíamos, isto é:

$$\text{Cov}(I(U\leq s),I(V\leq t)) = g(s,t) \leq m \leq (3/2)^{1/3} \text{Cov}^{1/3}(U,V). \quad \blacksquare$$

Sabemos que a convergência fraca de $\{Z_n\}$ em $D[0,1]$, decorre da fineza da sucessão $\{Z_n\}$ e da convergência em distribuição dos vectores aleatórios $(Z_n(t_1), \dots, Z_n(t_k))$ para cada $k \in \mathbb{N}$ e $t_1, \dots, t_k \in [0,1]$. Para estabelecer a última condição, Yu recorre a um teorema devido a Burton [4], o qual faz uso da definição que se segue.

Definição 2.1: Dizemos que uma sucessão $\{Y_n\}$ de vectores aleatórios em \mathbb{R}^d é fracamente associada, se para todo o subconjunto finito de índices $\{i_1, \dots, i_m\} \subset \mathbb{N}$ e para todo o k tal que $1 \leq k < m$, se tem :

$$\text{Cov}(f(Y_{i_1}, \dots, Y_{i_k}), g(Y_{i_{k+1}}, \dots, Y_{i_m})) \geq 0,$$

para quaisquer funções $f: \mathbb{R}^{kd} \rightarrow \mathbb{R}$ e $g: \mathbb{R}^{(m-k)d} \rightarrow \mathbb{R}$ não decrescentes em cada variável para as quais aquela covariância exista.

O teorema que se segue é uma versão simplificada do teorema devido a Burton a que nos referimos anteriormente, e trata-se de um teorema limite central para vectores aleatórios.

Teorema 2.1 [4]: Se $\{Y_n\}$ é uma sucessão de vectores aleatórios em \mathbb{R}^d , estritamente estacionária e fracamente associada, cujos vectores aleatórios são centrados e tal que:

$$E\|Y_0\|^2 < \infty$$

e

$$\sigma^2 = E\|Y_0\|^2 + 2 \sum_{i=1}^{\infty} \sum_{j=1}^d E\{Y_0^{(j)} Y_i^{(j)}\} < +\infty,$$

então $n^{-1/2}(Y_1 + \dots + Y_n)$ converge em distribuição para um vector aleatório Gaussiano centrado com matriz de covariâncias $\Gamma = [\sigma_{ij}]$, onde:

$$\sigma_{ij} = E\{Y_0^{(i)} Y_0^{(j)}\} + \sum_{k=1}^{\infty} \left(E\{Y_0^{(i)} Y_k^{(j)}\} + E\{Y_0^{(j)} Y_k^{(i)}\} \right).$$

Finalmente, apresentamos o teorema de Yu que estabelece a convergência do processo empírico uniforme $\{Z_n\}$.

Teorema 2.2 [23]: Se $\{X_n\}$ é uma sucessão estritamente estacionária e associada de variáveis aleatórias uniformemente distribuídas em $[0,1]$, e se existe uma constante v positiva tal que

$$\sum_{n=1}^{\infty} n^{13/2+v} \text{Cov}(X_0, X_n) < \infty,$$

então o processo empírico $\{Z_n\}$ associado à sucessão $\{X_n\}$, converge em distribuição para um elemento aleatório Z em $D[0,1]$, Gaussiano centrado com covariância definida por:

$$\Gamma(s,t) = \min\{s,t\} - s t + \sum_{k=1}^{\infty} \text{Cov}(I(X_0 \leq s), I(X_k \leq t)) + \sum_{k=1}^{\infty} \text{Cov}(I(X_0 \leq t), I(X_k \leq s)) \quad (2.2)$$

onde as séries são absolutamente convergentes.

Demonstração:

Tomemos um subconjunto finito $\{t_1, \dots, t_k\} \subset [0,1]$. Para provar a convergência em distribuição de $(Z_n(t_1), \dots, Z_n(t_k))$ para $(Z(t_1), \dots, Z(t_k))$ utilizamos o teorema 2.1, considerando $Y_n = (I(X_n \leq t_1) - t_1, I(X_n \leq t_2) - t_2, \dots, I(X_n \leq t_k) - t_k)$.

Das condições impostas à sucessão $\{X_n\}$ segue-se a estacionaridade e a associação fraca da sucessão de vectores $\{Y_n\}$. É também fácil verificar que $E\|Y_0\|^2 < \infty$ e que Y_0 é centrado. Para aplicar o teorema 2.1 precisamos ainda verificar que, para todo o $j \in \{1, \dots, d\}$, a série:

$$\sum_{i=1}^{\infty} E\{Y_0^{(j)} Y_i^{(j)}\} = \sum_{i=1}^{\infty} \text{Cov}(I(X_0 \leq t_j) - t_j, I(X_i \leq t_j) - t_j)$$

é convergente. Notemos que, esta é a série envolvida em (2.2), e que é de termos não negativos por associação da sucessão $\{X_n\}$. Para provar a convergência desta série recorreremos ao lema 2.2. Por este lema, para todo o $n \in \mathbb{N}$ temos:

$$\text{Cov}(I(X_0 \leq s), I(X_n \leq t)) \leq \left(\frac{3}{2}\right)^{1/3} \text{Cov}^{1/3}(X_0, X_n).$$

Então para estabelecer a convergência da série envolvida em (2.2) basta mostrar que a série $\sum_{n=1}^{\infty} \text{Cov}^{1/3}(X_0, X_n)$ é convergente, já que são ambas séries de termos não negativos por associação. Por hipótese:

$$\sum_{n=1}^{\infty} n^{13/2+v} \text{Cov}(X_0, X_n) < \infty,$$

o que implica que:

$$\sum_{n=1}^{\infty} n^2 \log^{2+v}(n+1) \text{Cov}(X_0, X_n) < \infty.$$

Pela desigualdade de Hölder com $p=3$ e $q=3/2$ vem:

$$\sum_{n=1}^{\infty} \text{Cov}^{1/3}(X_0, X_n) \leq \left(\sum_{n=1}^{\infty} n^2 \log^{2+v}(n+1) \text{Cov}(X_0, X_n) \right)^{1/3} \times \left(\sum_{n=1}^{\infty} n^{-1} \log^{-(2+v)/2}(n+1) \right)^{2/3}.$$

As duas séries envolvidas no segundo membro da desigualdade anterior são convergentes, donde decorre a convergência da série $\sum_{n=1}^{\infty} \text{Cov}^{1/3}(X_0, X_n)$.

Verificadas as condições do teorema (2.1) podemos concluir que,

$$n^{-1/2}(Y_1 + \dots + Y_n) = (Z_n(t_1), \dots, Z_n(t_k))$$

converge em distribuição para um vector aleatório Gaussiano centrado com covariância dada por:

$$\begin{aligned} \sigma_{ij} &= E\{Y_0^{(i)} Y_0^{(j)}\} + \sum_{k=1}^{\infty} (E\{Y_0^{(i)} Y_k^{(j)}\} + E\{Y_0^{(j)} Y_k^{(i)}\}) \\ &= \min\{t_i, t_j\} - t_i t_j + \sum_{k=1}^{\infty} (P(X_0 \leq t_i, X_k \leq t_j) - t_i t_j) + \sum_{k=1}^{\infty} (P(X_0 \leq t_j, X_k \leq t_i) - t_i t_j) \\ &= \min\{t_i, t_j\} - t_i t_j + \sum_{k=1}^{\infty} (\text{Cov}(I(X_0 \leq t_i), I(X_k \leq t_j)) + \text{Cov}(I(X_0 \leq t_j), I(X_k \leq t_i))), \end{aligned}$$

que é exactamente a covariância entre $Z(t_i)$ e $Z(t_j)$ dada por (2.2). Está então provado que $(Z_n(t_1), \dots, Z_n(t_k)) \xrightarrow{D} (Z(t_1), \dots, Z(t_k))$ para todo o subconjunto finito $\{t_1, \dots, t_k\} \subset [0, 1]$.

A técnica que vamos usar para demonstrar que a sucessão $\{Z_n\}$ é fina, é a mesma que seguimos na demonstração do teorema 1.2 deste capítulo. Relembramos que é suficiente verificar a seguinte condição:

$$(C2): \forall \varepsilon, \eta > 0 \quad \exists \delta \in (0, 1), \exists n_0 \in \mathbb{N} : P\left\{ \sup_{s \leq t \leq s+\delta} |Z_n(t) - Z_n(s)| \geq \varepsilon \right\} \leq \eta \delta, \text{ para } n \geq n_0 \text{ e}$$

para todo $t \in [0, 1]$.

Pelo lema 2.1 temos, para todo $n \geq 1$:

$$E \left[\sum_{i=1}^{\infty} [I(s < X_i \leq t) - (t-s)] \right]^4 \leq K_1 n^2 (n^{-1/2-\nu_1} + (t-s)^{6/5}). \quad (2.3)$$

No que se segue K_1 é sempre uma constante positiva, embora possa tomar valores diferentes ao longo desta demonstração. De (2.3) segue-se, para todo $n \geq 1$ e $0 \leq s < t \leq 1$, que:

$$E \left\{ |Z_n(t) - Z_n(s)|^4 \right\} = \frac{1}{n^2} E \left\{ \sum_{i=1}^n [I(s < U_i \leq t) - (t-s)] \right\}^4 \leq K_1 n^2 (n^{-1/2-\nu_1} + (t-s)^{6/5}).$$

Seja $\varepsilon \in (0, 1)$ e $r_n = \frac{\varepsilon}{\sqrt{n}}$. Se $r_n \leq t-s$ vem da desigualdade anterior:

$$E \left\{ |Z_n(t) - Z_n(s)|^4 \right\} \leq \frac{K_1}{\varepsilon^{1+2\nu_1}} (t-s)^{1+\nu_1}. \quad (2.4)$$

Consideramos as variáveis aleatórias: $Z_n(s+i r_n) - Z_n(s+(i-1)r_n)$, $i=1, \dots, m$, onde m é um inteiro positivo. Vamos aplicar o teorema 12.2 de Billingsley [3], tal como fizemos na demonstração do teorema 1.2 deste capítulo. Por (2.4) as condições do

teorema 12.2 de Billingsley são verificadas tomando $\gamma=4$, $\alpha=1+\nu_1$ e $u_i = \frac{C_1^{1/(1+\nu_1)}}{\varepsilon^{(1+2\nu_1)/(1+\nu_1)}} r_n$

, com $i=1, \dots, m$. Então por este teorema vem, para todo $\lambda > 0$:

$$P\left\{ \max_{i \leq m} |Z_n(s+i r_n) - Z_n(s)| \geq \lambda \right\} \leq \frac{K_1}{\lambda^4 \varepsilon^{1+2\nu_1}} (m r_n)^{1+\nu_1}. \quad (2.5)$$

Para qualquer valor de $\eta > 0$ escolhemos um $\delta > 0$ tal que: $\frac{2^{1+v_1} K_1 \delta^{v_1}}{\varepsilon^{5+2v_1}} < \eta$ e $m_n = [\delta/r_n]$. Então para n suficientemente grande, temos $m_n \geq 1$ e $r_n m_n \leq \delta < (m_n + 1)r_n \leq 2m_n r_n \leq 2\delta$, o que nos permite escrever:

$$\sup_{s \leq t \leq s + (m_n + 1)r_n} |Z_n(t) - Z_n(s)| \geq \sup_{s \leq t \leq s + \delta} |Z_n(t) - Z_n(s)|,$$

pois $\delta < (m_n + 1)r_n$. Utilizando agora a desigualdade (1.12) verificada na secção anterior vem:

$$3 \max_{i \leq m_n + 1} |Z_n(s + ir_n) - Z_n(s)| + r_n \sqrt{n} \geq \sup_{s \leq t \leq s + (m_n + 1)r_n} |Z_n(t) - Z_n(s)|.$$

Tendo em conta as duas últimas desigualdades e (2.5), vem:

$$\begin{aligned} P\left\{ \sup_{s \leq t \leq s + \delta} |Z_n(t) - Z_n(s)| \geq 4\varepsilon \right\} &\leq P\left\{ 3 \max_{i \leq m_n + 1} |Z_n(s + ir_n) - Z_n(s)| + \varepsilon \geq 4\varepsilon \right\} \\ &\leq \frac{K_1}{\varepsilon^{5+2v_1}} [(m_n + 1)r_n]^{1+v_1} \\ &\leq \frac{K_1}{\varepsilon^{5+2v_1}} (2\delta)^{1+v_1} \leq \eta\delta, \end{aligned}$$

o que verifica a condição (C_2) , estabelecendo assim a fineza de $\{Z_n\}$. ■

CAPÍTULO III

Convergência Fraca do Processo Empírico em $L^2[0,1]$

1. Espaços auto-reprodutores

Para o estudo da convergência do processo empírico, encarado como uma função aleatória no espaço $L^2[0,1]$, vamos utilizar a teoria dos espaços auto-reprodutores, introduzida em 1943 por Aronszajn em [1] e complementada pelo mesmo autor em 1950 em [2].

Seja X um espaço topológico e \mathbf{B}_X a tribo de Borel que lhe está associada.

Definição 1.1: Uma função real K definida em $X \times X$ diz-se um núcleo reprodutor se é simétrica e semi-definida positiva, isto é:

$$(i) \quad \forall x, y \in X \quad K(x, y) = K(y, x)$$

$$(ii) \quad \forall n \in \mathbb{N}, \forall a_1, \dots, a_n \in \mathbb{R} \quad \forall x_1, \dots, x_n \in X \quad \sum_{i, j=1}^n a_i a_j K(x_i, x_j) \geq 0$$

Definição 1.2: Um espaço H_K de aplicações f de X em \mathbb{R} , diz-se um espaço de Hilbert de núcleo reprodutor K ou espaço auto-reprodutor associado a K , se:

$$(i) \quad \forall x \in X, \text{ a função } K(x, \cdot) \in H_K;$$

$$(ii) \quad \forall x \in X, \forall f \in H_K, f(x) = \langle f, K(x, \cdot) \rangle_K, \quad (\text{propriedade de auto-reprodução})$$

onde $\langle \cdot, \cdot \rangle_K$ representa o produto interno em H_K .

Por Aronszajn [1] sabemos que a todo o núcleo reprodutor K podemos associar um espaço de Hilbert auto-reprodutor H_K . A construção deste espaço é feita definindo no espaço H_0 , das combinações lineares finitas de $K(x, \cdot)$, um produto interno pondo: $\langle K(x, \cdot), K(y, \cdot) \rangle_K = K(x, y)$. O espaço H_K é pois o complemento de H_0 , que pode ser

construído juntando as funções de X em \mathbb{R} que sejam limites simples de sucessões de Cauchy em H_0 .

Por outro lado, vamos ver que, se H é um espaço verificando a definição 2.2, a função K que lhe está associada é, de facto, um núcleo reprodutor e além disso única. Veremos também que o espaço H_0 , das combinações lineares finitas de $K(x, \cdot)$, é denso em H . Começamos por provar que K é simétrica. Para isso consideramos x e y dois elementos quaisquer de X . Da definição 2.2 sabemos que as funções $f_y = K(y, \cdot)$ e $f_x = K(x, \cdot)$ pertencem a H , logo da propriedade de auto-reprodução segue-se:

$$K(y, x) = f_y(x) = \langle f_y, K(x, \cdot) \rangle_K = \langle K(x, \cdot), f_y \rangle_K = \langle f_x, K(y, \cdot) \rangle_K = f_x(y) = K(x, y),$$

donde K é de facto uma função simétrica.

Para concluirmos que K é um núcleo reprodutor, falta apenas verificar que é semi-definida positiva. Sejam $n \in \mathbb{N}$, $a_1, \dots, a_n \in \mathbb{R}$ e $x_1, \dots, x_n \in X$, então temos:

$$\sum_{i,j=1}^n a_i a_j K(x_i, x_j) = \sum_{i,j=1}^n a_i a_j \langle K(x_i, \cdot), K(x_j, \cdot) \rangle_K = \left\langle \sum_{i=1}^n a_i K(x_i, \cdot), \sum_{j=1}^n a_j K(x_j, \cdot) \right\rangle_K \geq 0,$$

portanto K é semi-definida positiva, logo é um núcleo reprodutor.

Finalmente, vamos verificar que K é único. Se de facto não o for, isto é, se existirem dois núcleos reprodutores K e K^* aos quais está associado o mesmo espaço auto-reprodutor H , vem pela propriedade de auto-reprodução:

$$\forall f \in H, \forall x \in X \quad f(x) = \langle f, K(x, \cdot) \rangle_K = \langle f, K^*(x, \cdot) \rangle_K.$$

Fixemos $x \in X$ arbitrariamente, então para todo o f de H , vem:

$$\langle f, K(x, \cdot) - K^*(x, \cdot) \rangle_K = 0 \Rightarrow K(x, \cdot) - K^*(x, \cdot) = 0 \Rightarrow K(x, \cdot) = K^*(x, \cdot),$$

isto é, $K(x, y) = K^*(x, y)$ para todo o $y \in X$. Da escolha arbitrária de x segue-se que $K = K^*$. Está assim provado que K é único.

Falta ainda verificar que H_0 é denso em H . Para isso vamos precisar do lema que se segue.

Lema 1.1 [9]: Para qualquer subconjunto $M \neq \emptyset$ de um espaço de Hilbert H , o subespaço gerado por M é denso em H se e só se $M^\perp = \{0\}$.

Por este lema H_0 é denso em H , pois o único elemento ortogonal a todas as funções $K(x, \cdot)$ é a função identicamente nula. De facto: $f(x) = \langle f, K(x, \cdot) \rangle_K = 0$, $\forall x \in X$.

A propriedade de auto-reprodução permite a transmissão de certas propriedades de K a todos os elementos de H_K . Por exemplo, se K for um núcleo reprodutor limitado, todas as funções do espaço H_K que lhe está associado, são também limitadas. De facto, se $f \in H_K$,

$$\begin{aligned} \sup_{x \in X} |f(x)| &= \sup_{x \in X} |\langle f, K(x, \cdot) \rangle_K| \leq \sup_{x \in X} (\|f\|_K \|K(x, \cdot)\|_K) \\ &= \sup_{x \in X} (\|f\|_K K(x, x)^{1/2}) \\ &\leq \|f\|_K \sup_{x \in X} |K|^{1/2}. \end{aligned}$$

Portanto $\|f\|_\infty = \sup_{x \in X} |f(x)| \leq \|f\|_K \sup_{x \in X} |K|^{1/2}$, onde $\|\cdot\|_K$ é a norma definida à custa do produto interno $\langle \cdot, \cdot \rangle_K$. Daqui concluímos que f é limitada.

Para o nosso estudo, vamos considerar $X=[0,1]$ e o núcleo K que admite a representação integral que se indica:

$$K(s,t) = 1 - \max(s,t) = \int_{[0,1]} I_{[s,1]}(u) I_{[t,1]}(u) \lambda(du). \quad (1.1)$$

Para verificar que K é de facto um núcleo reprodutor, basta mostrar que é semi-definida positiva, pois a sua simetria é evidente. Tendo em conta que as funções

do tipo $I_{[s,1]}(\cdot)$ estão em $L^2[0,1]$ e, representando o produto interno neste espaço por $\langle \cdot, \cdot \rangle_2$, vem:

$$\forall n \in \mathbf{N}, \forall a_1, \dots, a_n \in \mathbf{R}, \forall x_1, \dots, x_n \in [0,1]$$

$$\begin{aligned} \sum_{i,j=1}^n a_i a_j K(x_i, x_j) &= \sum_{i,j=1}^n a_i a_j \int_{[0,1]} I_{[x_i,1]}(u) I_{[x_j,1]}(u) \lambda(du) = \\ &= \sum_{i,j=1}^n a_i a_j \langle I_{[x_i,1]}(\cdot), I_{[x_j,1]}(\cdot) \rangle_2 = \\ &= \left\langle \sum_{i=1}^n a_i I_{[x_i,1]}(\cdot), \sum_{j=1}^n a_j I_{[x_j,1]}(\cdot) \right\rangle_2 \geq 0, \end{aligned}$$

donde K é semi-definida positiva.

Sabemos que à função K definida em (1.1) podemos associar um espaço de Hilbert auto-reprodutor H_K . Pretendemos agora, encontrar uma forma de explicitar qualquer função de H_K . Antes disso, notemos que o espaço F gerado por $M = \{I_{[s,1]}(\cdot), s \in [0,1]\}$ é denso em $L^2[0,1]$. Pelo lema 1.1 só precisamos de verificar que o único elemento ortogonal a todas as funções do conjunto M é o zero de $L^2[0,1]$. De facto, para $s \in [0,1]$, $\langle f, I_{[s,1]}(\cdot) \rangle_2 = \int_{[0,1]} f(u) I_{[s,1]}(u) \lambda(du) = 0$. Atendendo a que $s \in [0,1]$ é arbitrário, segue-se que $f = 0$.

O teorema que se segue é uma versão de um resultado de [21] devido a Suquet.

Teorema 1.1 [21]:

(i) h é uma função de H_K se e só se existe uma função g de $L^2[0,1]$ tal que:

$$h(x) = \int_{[x,1]} g(u) \lambda(du) ;$$

(ii) esta representação é única e define uma isometria Ψ de H_K sobre $L^2[0,1]$.

Demonstração:

Definimos uma aplicação Ψ de H_0 em $L^2[0,1]$ por:

$$\Psi\left(\sum_{i=1}^n a_i K(x_i, \cdot)\right) = \sum_{i=1}^n a_i I_{[x_i, 1]}(\cdot).$$

Como a expressão depende da representação escolhida para o elemento de H_0 , começamos por verificar que Ψ é, de facto, independente desta representação. Seja $h \in H_0$ tal que:

$$h = \sum_{i=1}^n a_i K(x_i, \cdot) = \sum_{j=1}^m b_j K(y_j, \cdot)$$

e g_1, g_2 elementos de $L^2[0,1]$ tais que:

$$g_1 = \sum_{i=1}^n a_i I_{[x_i, 1]}(\cdot), \quad g_2 = \sum_{j=1}^m b_j I_{[y_j, 1]}(\cdot).$$

Pretendemos mostrar que $g_1 = g_2 = \Psi(h)$. Ora para todo o $x \in [0,1]$ temos:

$$\begin{aligned} \langle g_1, I_{[x, 1]}(\cdot) \rangle_2 &= \sum_{i=1}^n a_i \int_{[0,1]} I_{[x_i, 1]}(u) I_{[x, 1]}(u) \lambda(du) \\ &= \sum_{i=1}^n a_i K(x_i, x) = \sum_{j=1}^m b_j K(y_j, x) \\ &= \int_{[0,1]} \sum_{j=1}^m b_j I_{[y_j, 1]}(u) I_{[x, 1]}(u) \lambda(du) \\ &= \langle g_2, I_{[x, 1]}(\cdot) \rangle_2. \end{aligned}$$

Donde concluímos que:

$$\langle g_1 - g_2, g_1 - g_2 \rangle_2 = \sum_{i=1}^n a_i \langle g_1 - g_2, I_{[x_i, 1]}(\cdot) \rangle_2 - \sum_{j=1}^m b_j \langle g_1 - g_2, I_{[y_j, 1]}(\cdot) \rangle_2 = 0,$$

isto é, $g_1 = g_2$ em $L^2[0,1]$.

É evidente que Ψ é um operador linear de H_0 em $L^2[0,1]$, e o seu contradomínio $\Psi(H_0)$ é o subespaço F introduzido atrás, logo denso em $L^2[0,1]$. Temos ainda:

$$\begin{aligned} \|h\|_K^2 &= \langle h, h \rangle_K = \sum_{i,j=1}^n a_i a_j K(x_i, x_j) = \int_{[0,1]} \left(\sum_{i=1}^n a_i I_{[x_i,1]}(u) \right)^2 \lambda(du) \\ &= \int_{[0,1]} [\Psi(h)(u)]^2 \lambda(du) = \|\Psi(h)\|_2^2, \end{aligned}$$

o que prova que Ψ é uma isometria.

Por ser uma isometria, Ψ é contínua, logo trata-se de um operador linear limitado. Como tal, Ψ admite uma extensão única a $\overline{H_0} = H_K$, que denotamos também por Ψ [9]. Facilmente se verifica que esta extensão é uma isometria bijectiva de $L^2[0,1]$ em H_K .

Seja agora h um elemento arbitrário de H_K . Vamos ver que h admite uma representação do tipo assinalado em (i). Sabemos que para todo o $x \in [0,1]$, $h(x) = \langle h, K(x, \cdot) \rangle_K$. Como Ψ é uma isometria vem:

$$h(x) = \langle \Psi(h), I_{[x,1]}(\cdot) \rangle_2 = \int_{[0,1]} \Psi(h)(u) I_{[x,1]}(u) \lambda(du) = \int_{[x,1]} \Psi(h)(u) \lambda(du).$$

Se esta representação não for única, deve existir em $L^2[0,1]$ outra função g tal que, para todo o $x \in [0,1]$, se tenha:

$$h(x) = \langle \Psi(h), I_{[x,1]}(\cdot) \rangle_2 = \langle g, I_{[x,1]}(\cdot) \rangle_2.$$

Temos então, para todo o $x \in [0,1]$: $\langle \Psi(h) - g, I_{[x,1]}(\cdot) \rangle_2 = 0$. Já vimos que a única função ortogonal a todas as funções $I_{[x,1]}(\cdot)$ com $x \in [0,1]$, é a função nula de $L^2[0,1]$, logo $\Psi(h) = g$ em $L^2[0,1]$, donde concluímos que aquela representação é única.

Falta provar a condição necessária de (i). Para isso consideramos g um elemento qualquer de $L^2[0,1]$. Então pela sobrejectividade de Ψ , existe uma função h em H_K tal que $\Psi(h)=g$. Assim:

$$\int_{[0,1]} g(u) I_{[x,1]}(u) \lambda(du) = \int_{[x,1]} \Psi(h)(u) I_{[x,1]}(u) \lambda(du) = h(x),$$

o que prova a condição necessária da alínea (i) deste teorema. ■

O teorema que acabamos de demonstrar permite-nos descrever o conjunto H_K da seguinte maneira:

$$H_K = \left\{ h: h(x) = \int_{[x,1]} g(u) \lambda(du), g \in L^2[0,1] \right\}.$$

Nas secções que se seguem consideramos muitas vezes a inversa de Ψ :

$$\begin{aligned} \Psi^{-1}: L^2[0,1] &\longrightarrow H_K \\ g &\longrightarrow h(x) = \int_{[x,1]} g(u) \lambda(du). \end{aligned}$$

Para finalizar esta secção, notemos que por K ser limitada, toda a função h de H_K é também limitada.

2. Considerações gerais

Vimos no capítulo I, que a convergência fraca de uma sucessão $\{Z_n\}$ para um elemento aleatório Z em $L^2[0,1]$, decorre da compacidade relativa da sucessão e da convergência em distribuição das variáveis aleatórias $\langle Z_n, g \rangle_2$ para $\langle Z, g \rangle_2$ com $g \in L^2[0,1]$.

Neste capítulo interessa-nos estudar a convergência fraca do processo empírico uniforme, que denotamos por $\{Z_n\}$, associado a uma sucessão $\{X_n\}$ de variáveis aleatórias uniformemente distribuídas no intervalo $[0,1]$. Da definição de processo empírico (secção 3 do capítulo I) vem, para $t \in [0,1]$:

$$Z_n(t) = \sqrt{n} (F_n(t) - \lambda[0,t]) ,$$

onde $F_n(t) = \frac{1}{n} \sum_{i=1}^n I_{[X_i,1]}(t)$ é a função de distribuição empírica de X_1, \dots, X_n .

Pela teorema 2.2 do capítulo I, se $(e_i)_{i \in \mathbb{N}}$ for uma base ortonormada de $L^2[0,1]$ e se:

$$(i) \sup_{n \geq 1} E\{\|Z_n\|_2^2\} < +\infty ; \quad (2.1)$$

$$(ii) \lim_{N \rightarrow +\infty} \sup_{n \geq 1} E\left\{ \sum_{i=N}^{\infty} \left(\int_{[0,1]} e_i Z_n d\lambda \right)^2 \right\} = 0 , \quad (2.2)$$

então $\{Z_n\}$ é relativamente compacta.

A base ortonormada que consideramos é o sistema de Haar $(e_n)_{n \in \mathbb{N}}$ definida à custa de $\theta(x) = I_{[0,1/2)}(x) - I_{[1/2,1]}(x)$, da seguinte maneira:

$$e_0(x) = 1$$

$$e_n(x) = 2^{j/2} \theta(2^j x - k) \text{ onde } n=2^j+k \text{ e } 0 \leq k < 2^j .$$

Para estas funções $\int_{[0,1]} e_m e_n d\lambda$ é igual a 1 se $m=n$ e igual a 0 se $m \neq n$.

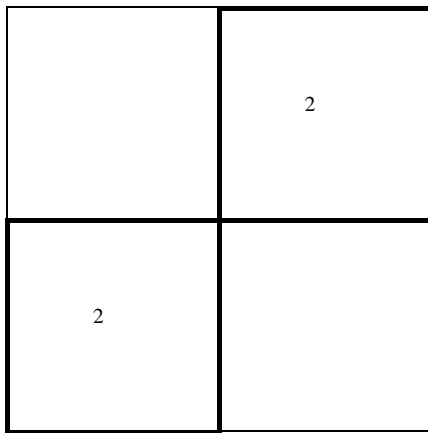
Definição 2.1: Uma base de Schauder num espaço normado X , é uma sucessão $(e_n)_{n \in \mathbb{N}}$ de elementos de X , tal que, para todo o $x \in X$, existe uma única sucessão de escalares $(\alpha_n)_{n \in \mathbb{N}}$ que verifica a seguinte condição: $\|x - \sum_{i=1}^n \alpha_i e_i\| \xrightarrow{n} 0$.

O sistema de Haar é uma base de Schauder e para toda a função contínua f definida em $[0,1]$, as séries de Haar de f são uniformemente convergentes, isto é, se $f \in C[0,1]$ então $\sum_{n=1}^{\infty} \langle f, e_n \rangle e_n$ é uniformemente convergente para f (capítulos 1 e 2 de [19]).

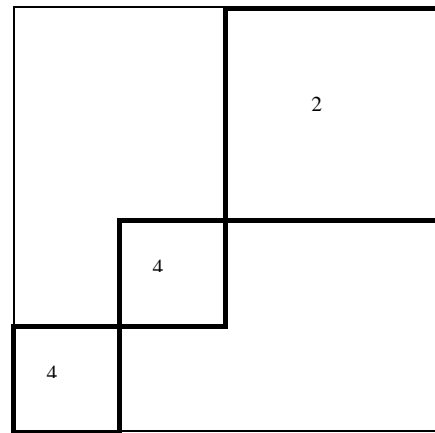
Para cada $m \in \mathbb{N}$, seja K_m o núcleo de Dirichlet definido por:

$$K_m(x,y) = \sum_{i=0}^m e_i(x) e_i(y), \quad (x,y) \in [0,1]^2.$$

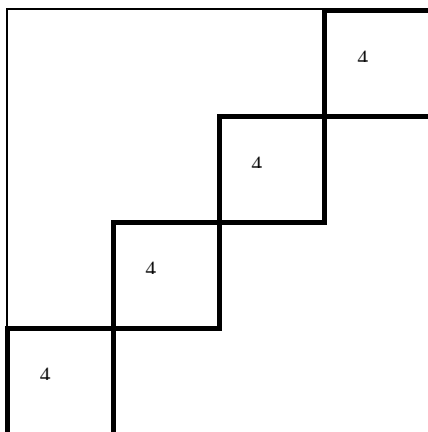
Designamos por C_m o suporte de K_m . As figuras que se seguem, mostram os suportes C_1, C_2, C_3, C_4, C_5 e C_6 e os respectivos valores de K_m nesses suportes:



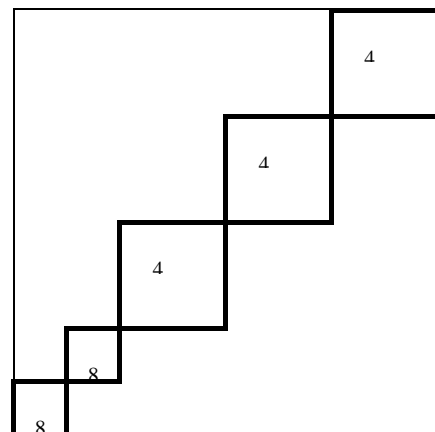
C_1



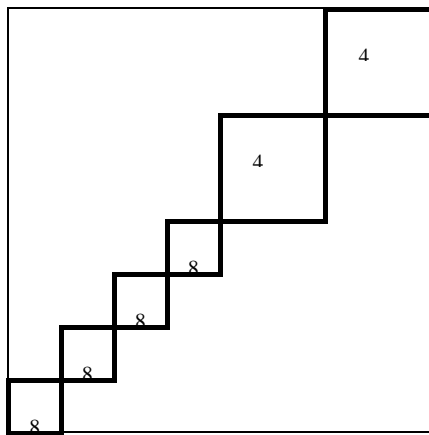
C_2



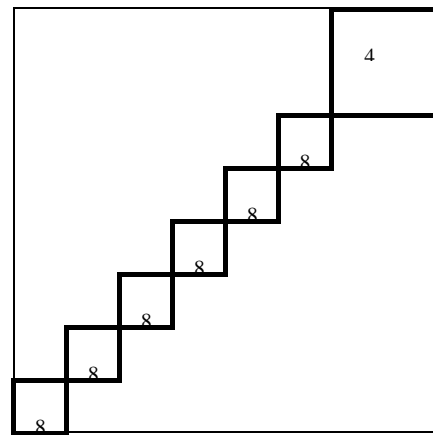
C_3



C_4



C_5



C_6

Como se pode verificar, para cada $m \in \mathbb{N}$, C_m é um conjunto de quadrados dispostos ao longo da diagonal de $[0,1]^2$. De C_m para C_{m+1} , um dos quadrados é dividido em quatro e rejeitam-se os dois quadrados mais pequenos que estão fora da diagonal. Além disso, K_m é constante em cada um dos quadrados de C_m e $\int_{C_m} K_m(x,y) d\lambda^2 = 1$ para cada $m \in \mathbb{N}$. Temos ainda, para $p > m$:

$$K_m(x,y) \leq K_p(x,y), (x,y) \in C_p. \quad (2.3)$$

Notemos que se f for uma função contínua em $[0,1]$, o integral

$$\int_{[0,1]} K_m(\cdot, y) f(y) \lambda(dy) = \sum_{i=0}^m \int_{[0,1]} e_i(y) f(y) \lambda(dy) e_i(\cdot) = \sum_{i=0}^m \langle e_i, f \rangle_2 e_i(\cdot)$$

converge uniformemente para f , por se tratar da série de Haar de f .

3. Compacidade relativa do processo empírico

No que se segue, $(e_n)_{n \in \mathbb{N}}$ representa a base de Haar definida na secção anterior, que é, como já dissemos, uma base ortonormada em $L^2[0,1]$.

Nesta secção pretendemos estabelecer condições suficientes para a compacidade relativa do processo empírico uniforme $\{Z_n\}$.

Teorema 3.1 [13]: Se a sucessão de funções $(L_n)_{n \in \mathbb{N}}$ definida por:

$$L_n(x,y) = \frac{1}{n} \sum_{j,k=1}^n [P(X_j \leq x, X_k \leq y) - P(X_j \leq x)P(X_k \leq y)],$$

converge uniformemente em $[0,1]^2$, então a sucessão $\{Z_n\}$ é relativamente compacta em $L^2[0,1]$.

Demonstração:

Para cada $n \in \mathbb{N}$, L_n é uma função contínua em $[0,1]^2$. De facto,

$$L_n(x,y) = \frac{1}{n} \sum_{j,k=1}^n [F_{(X_j, X_k)}(x,y) - x y],$$

onde $F_{(X_j, X_k)}$, a função de distribuição do vector (X_j, X_k) , é contínua, pois X_j e X_k têm funções de distribuição contínuas, já que são uniformemente distribuídas no intervalo $[0,1]$.

Sendo L o limite uniforme da sucessão $(L_n)_{n \geq 1}$ de funções contínuas, então L é também contínua. Além disso, L é definida no compacto $[0,1]^2$, logo é limitada, donde vem: $\|L\|_\infty < \infty$. Por outro lado,

$$\|L_n - L\|_\infty \xrightarrow{n} 0 \Rightarrow \|L_n\|_\infty \xrightarrow{n} \|L\|_\infty,$$

donde se conclui que a sucessão $\|L_n\|_\infty$ é limitada, isto é, $\sup_{n \in \mathbb{N}} \|L_n\|_\infty < +\infty$.

Para estabelecer a compacidade relativa de $\{Z_n\}$ precisamos de verificar as condições (2.1) e (2.2) assinaladas na secção anterior. Começamos pela condição (2.1). Por aplicação do teorema de Fubini, temos:

$$E\{\|Z_n\|_2^2\} = E\left\{\int_{[0,1]} Z_n(u)^2 \lambda(du)\right\} = \int_{[0,1]} E\{Z_n(u)^2\} \lambda(du).$$

Mas,

$$\begin{aligned} E\{Z_n(u)^2\} &= E\{[\sqrt{n}(F_n(u) - u)]^2\} = n \times \frac{1}{n^2} \sum_{i,j=1}^n P(X_i \leq u, X_j \leq u) - n u^2 \\ &= \frac{1}{n} \sum_{i,j=1}^n [P(X_i \leq u, X_j \leq u) - P(X_i \leq u)P(X_j \leq u)] \\ &= L_n(u, u). \end{aligned}$$

Donde vem:

$$\begin{aligned} E\{\|Z_n\|_2^2\} &= \int_{[0,1]} L_n(u, u) \lambda(du) \leq \\ &\leq \int_{[0,1]} \|L_n\|_\infty \lambda(du) \\ &\leq \sup_{n \geq 1} \|L_n\|_\infty < +\infty. \end{aligned}$$

Então $\sup_{n \geq 1} E\{\|Z_n\|_2^2\} < +\infty$, isto é, a condição (2.1) de compacidade relativa é verificada.

Para verificar a condição (2.2) consideramos:

$$a_N = \sup_{n \geq 1} E\left\{\sum_{i=N}^{\infty} \left(\int_{[0,1]} e_i Z_n d\lambda\right)^2\right\} = \sup_{n \geq 1} E\left\{\sum_{i=N}^{\infty} \langle e_i, Z_n \rangle_2^2\right\}.$$

Para já, vamos mostrar que a sucessão de termos geral a_N é de Cauchy. Para $p > m$ vem:

$$\begin{aligned} |a_{m+1} - a_{p+1}| &\leq \sup_{n \geq 1} \left| E \left\{ \sum_{i=m+1}^{\infty} \langle e_i, Z_n \rangle_2^2 \right\} - E \left\{ \sum_{i=p+1}^{\infty} \langle e_i, Z_n \rangle_2^2 \right\} \right| \\ &= \sup_{n \geq 1} \left| E \left\{ \sum_{i=m+1}^p \left(\int_{[0,1]} e_i Z_n d\lambda \right)^2 \right\} \right|. \end{aligned}$$

Basta portanto mostrar que, para todo $\varepsilon > 0$, existe $m_0 \in \mathbb{N}$ tal que, se $p > m > m_0$ tem-se, para todo o $n \geq 1$:

$$E \left\{ \sum_{i=m+1}^p \left(\int_{[0,1]} e_i Z_n d\lambda \right)^2 \right\} < \varepsilon. \quad (3.1)$$

Aplicando o teorema de Fubini, vem:

$$\begin{aligned} E \left\{ \left(\int_{[0,1]} e_i Z_n d\lambda \right)^2 \right\} &= E \left\{ \int_{[0,1]^2} e_i(x) e_i(y) Z_n(x) Z_n(y) \lambda \otimes \lambda(dx, dy) \right\} \\ &= \int_{[0,1]^2} e_i(x) e_i(y) E \{ Z_n(x) Z_n(y) \} \lambda \otimes \lambda(dx, dy) \\ &= \int_{[0,1]^2} e_i(x) e_i(y) L_n(x, y) \lambda \otimes \lambda(dx, dy), \end{aligned}$$

donde,

$$\begin{aligned} E \left\{ \sum_{i=m+1}^p \left(\int_{[0,1]} e_i Z_n d\lambda \right)^2 \right\} &= \int_{[0,1]^2} \sum_{i=m+1}^p e_i(x) e_i(y) L_n(x, y) \lambda \otimes \lambda(dx, dy) \\ &= \int_{[0,1]^2} [K_p(x, y) - K_m(x, y)] L_n(x, y) \lambda \otimes \lambda(dx, dy). \end{aligned}$$

Dado $\varepsilon > 0$ arbitrário, a convergência uniforme de L_n para L , permite-nos escolher $n_0 \in \mathbb{N}$ tal que, para todo o $x, y \in [0, 1]$ se tenha:

$$n \geq n_0 \Rightarrow L(x, y) - \varepsilon \leq L_n(x, y) \leq L(x, y) + \varepsilon.$$

Usando as propriedades dos núcleos de Dirichlet K_m , mencionadas na secção anterior deste capítulo vem, para $n \geq n_0$,

$$\begin{aligned}
E \left\{ \sum_{i=m+1}^p \left(\int_{[0,1]} e_i Z_n d\lambda \right)^2 \right\} &= \int_{C_m} (K_p - K_m) L_n d\lambda^2 \\
&= \int_{C_m \setminus C_p} -K_m L_n d\lambda^2 + \int_{C_p} (K_p - K_m) L_n d\lambda^2 \\
&\leq \int_{C_m \setminus C_p} -K_m (L - \varepsilon) d\lambda^2 + \int_{C_p} (K_p - K_m) (L + \varepsilon) d\lambda^2 \\
&\leq \int_{C_m} -K_m L d\lambda^2 + \int_{C_p} K_p L d\lambda^2 + \varepsilon \int_{C_m} K_m d\lambda^2 + \varepsilon \int_{C_p} K_p d\lambda^2 \\
&= - \int_{[0,1]^2} K_m L d\lambda^2 + \int_{[0,1]^2} K_p L d\lambda^2 + 2\varepsilon . \tag{3.2}
\end{aligned}$$

Como L é uma função contínua, o integral

$$\int_{[0,1]} K_m(x,y) L(x,y) \lambda(dy) ,$$

converge uniformemente para $L(x,x)$ (ver secção 2 deste capítulo). Tanto K_m como L são funções limitadas. Podemos portanto, majorar aquele integral por uma constante, o que nos permite aplicar o teorema da convergência dominada e obter:

$$\int_{[0,1]^2} K_m(x,y) L(x,y) \lambda^2(dx,dy) \longrightarrow \int_{[0,1]} L(x,x) \lambda(dx) . \tag{3.3}$$

Então existe $m_0 \in \mathbb{N}$ tal que para $p > m \geq m_0$:

$$\left| \int_{[0,1]^2} K_m L d\lambda^2 - \int_{[0,1]} L d\lambda \right| < \varepsilon/2$$

e

$$\left| \int_{[0,1]^2} K_p L d\lambda^2 - \int_{[0,1]} L d\lambda \right| < \varepsilon/2 .$$

Donde se segue, somando e subtraindo $\int_{[0,1]} L \, d\lambda$ em (3.2) que, para $n \geq n_0$:

$$E \left\{ \sum_{i=m+1}^p \left(\int_{[0,1]} e_i Z_n \, d\lambda \right)^2 \right\} < 3\varepsilon .$$

Falta ainda controlar as esperanças correspondentes ao número finito de índices $n < n_0$. Fixemos um destes índices arbitrariamente. Então, uma vez que L_n é também uma função contínua e limitada, a convergência assinalada em (3.3) verifica-se com L_n em vez de L . Podemos então escolher $m_0(n) \in \mathbb{N}$ tal que, para $p > m \geq m_0(n)$, se verifique:

$$\begin{aligned} E \left\{ \sum_{i=m+1}^p \left(\int_{[0,1]} e_i Z_n \, d\lambda \right)^2 \right\} &= \\ &= \int_{[0,1]^2} K_p L_n \, d\lambda^2 - \int_{[0,1]^2} K_m L_n \lambda^2 \\ &= \left(\int_{[0,1]^2} K_p L_n \, d\lambda^2 - \int_{[0,1]} L_n \, d\lambda \right) - \left(\int_{[0,1]^2} K_m L_n \lambda^2 - \int_{[0,1]} L_n \, d\lambda \right) \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon . \end{aligned}$$

Então, tomando $m_0 = \max \{ m_0(1), m_0(2), \dots, m_0(n_0-1) \}$, vem para todo o $n < n_0$ e $p > m \geq m_0$:

$$E \left\{ \sum_{i=m+1}^p \left(\int_{[0,1]} e_i Z_n \, d\lambda \right)^2 \right\} < \varepsilon .$$

Com isto fica estabelecido que a sucessão de termo geral

$$a_N = \sup_{n \geq 1} E \left\{ \sum_{i=N}^{\infty} \left(\int_{[0,1]} e_i Z_n \, d\lambda \right)^2 \right\}$$

é de Cauchy, logo convergente, o que é o mesmo que dizer que

$$a_N(n) = E \left\{ \sum_{i=N}^{\infty} \left(\int_{[0,1]} e_i Z_n \, d\lambda \right)^2 \right\}$$

é uniformemente convergente relativamente a n . Como para cada $n \in \mathbb{N}$ fixo, se tem:

$$\lim_{N \rightarrow \infty} a_N(n) = 0,$$

pois $a_N(n)$ é o resto de ordem N de uma série convergente, então $a_N(n)$ converge uniformemente para zero, isto é:

$$\lim_{N \rightarrow \infty} \sup_{n \geq 1} E \left\{ \sum_{i=N}^{\infty} \left(\int_{[0,1]} e_i Z_n d\lambda \right)^2 \right\} = 0. \quad \blacksquare$$

A condição de compacidade relativa do teorema que acabamos de demonstrar, pode ser rescrita para sucessões estacionárias da forma que nos indica o seguinte teorema.

Teorema 3.2 [13]: Seja $\{X_n\}$ uma sucessão estritamente estacionária de variáveis aleatórias uniformemente distribuídas em $[0,1]$. Se,

$$\sum_{n=0}^{\infty} |P(X_n \leq x, X_0 \leq y) - xy|$$

é uniformemente convergente em $[0,1]^2$, então a sucessão $\{Z_n\}$ é relativamente compacta em $L^2[0,1]$.

Demonstração:

Pelo teorema 3.1 basta mostrar que L_n é uniformemente convergente.

$$\begin{aligned} \text{Seja } Q_{jk}(x,y) &= P(X_j \leq x, X_k \leq y) - P(X_j \leq x) P(X_k \leq y) \\ &= P(X_j \leq x, X_k \leq y) - xy. \end{aligned}$$

É fácil verificar que $L_n(x,y)$ se pode rescrever da seguinte maneira:

$$L_n(x,y) = \frac{1}{n} \sum_{j,k=1}^n Q_{jk}(x,y) = \frac{1}{n} \sum_{j=1}^n Q_{jj}(x,y) + \frac{1}{n} \sum_{1 \leq j < k \leq n} Q_{jk}(x,y) + \frac{1}{n} \sum_{1 \leq k < j \leq n} Q_{jk}(x,y).$$

Atendendo agora à estacionaridade de $\{X_n\}$, obtêm-se:

$$\begin{aligned} L_n(x,y) &= \frac{1}{n} \sum_{i=1}^n Q_{00}(x,y) + \frac{1}{n} \sum_{i=1}^{n-1} (n-i)Q_{0i}(x,y) + \frac{1}{n} \sum_{i=1}^{n-1} (n-i)Q_{i0}(x,y) \\ &= Q_{00}(x,y) + \sum_{i=1}^{n-1} \left(1 - \frac{i}{n}\right)(Q_{0i}(x,y) + Q_{i0}(x,y)). \end{aligned}$$

Como $Q_{0i}(x,y) = Q_{i0}(y,x)$, a convergência uniforme de $\sum_{i=0}^{\infty} |Q_{i0}(x,y)|$ em $[0,1]^2$

implica a convergência uniforme de $\sum_{i=0}^{\infty} |Q_{0i}(x,y)|$ e de $\sum_{i=0}^{\infty} |Q_{0i}(x,y) + Q_{i0}(x,y)|$.

Então,

$$\begin{aligned} \left| L_n - Q_{00} - \sum_{i=1}^{n-1} (Q_{i0} + Q_{0i}) \right| &= \left| \sum_{i=1}^{n-1} \frac{i}{n} (Q_{i0} + Q_{0i}) \right| \\ &\leq \sum_{i < \sqrt{n}} \frac{i}{n} |Q_{i0} + Q_{0i}| + \sum_{\sqrt{n} \leq i \leq n-1} \frac{i}{n} |Q_{i0} + Q_{0i}| \\ &\leq \frac{1}{\sqrt{n}} \sum_{i < \sqrt{n}} |Q_{i0} + Q_{0i}| + \sum_{i \geq \sqrt{n}} |Q_{i0} + Q_{0i}| \\ &\leq \frac{1}{\sqrt{n}} \sum_{i < \sqrt{n}} |Q_{i0}(x,y)| + \frac{1}{\sqrt{n}} \sum_{i < \sqrt{n}} |Q_{i0}(y,x)| + \\ &\quad + \sum_{i \geq \sqrt{n}} |Q_{i0}(x,y)| + \sum_{i \geq \sqrt{n}} |Q_{i0}(y,x)| \\ &\leq \frac{2}{\sqrt{n}} \sup_{(x,y)} \sum_{i=0}^{\infty} |Q_{i0}(x,y)| + 2 \sup_{(x,y)} \sum_{i \geq \sqrt{n}} |Q_{i0}(x,y)|. \end{aligned}$$

A função $\sum_{i=0}^{\infty} |Q_{i0}(x,y)|$ é contínua por ser limite uniforme de uma sucessão de funções contínuas, além disso está definida num compacto, logo é limitada. Então,

$$\frac{2}{\sqrt{n}} \sup_{(x,y)} \sum_{i=0}^{\infty} |Q_{i0}(x,y)| \xrightarrow{n} 0.$$

Por outro lado, da convergência uniforme da série $\sum_{i=0}^{\infty} |Q_{i0}(x,y)|$, segue-se imediatamente:

$$2 \sup_{(x,y) \text{ } i \geq \sqrt{n}} |Q_{i0}(x,y)| \xrightarrow{n} 0.$$

Portanto L_n é uniformemente convergente. ■

4. Convergência do processo empírico

Na secção anterior estudámos condições suficientes para a compacidade relativa do processo empírico uniforme. Resta-nos estabelecer condições suficientes para a convergência em distribuição das variáveis aleatórias $\langle Z_n, g \rangle_2$ com $g \in L^2[0,1]$.

Antes de mais, notemos que para $g \in L^2[0,1]$, temos por aplicação do teorema de Fubini:

$$\begin{aligned} \int_{[0,1]} g(t) \lambda[0,t] \lambda(dt) &= \int_{[0,1]} \int_{[0,1]} g(t) I_{[0,t]}(s) \lambda(ds) \lambda(dt) \\ &= \int_{[0,1]} \int_{[s,1]} g(t) \lambda(dt) \lambda(ds) \\ &= E \left\{ \int_{[X_i,1]} g(t) \lambda(dt) \right\}, \end{aligned}$$

pois a distribuição das variáveis aleatórias X_i coincide com a medida de Lebesgue λ em $[0,1]$. Temos então:

$$\int_{[0,1]} t g(t) \lambda(dt) = E \left\{ \int_{[X_i,1]} g(t) \lambda(dt) \right\}. \quad (4.1)$$

Dado $g \in L^2[0,1]$, se definirmos $h = \Psi^{-1}(g)$, onde Ψ é a isometria entre H_K e $L^2[0,1]$ introduzida na secção 1, temos:

$$\begin{aligned} \langle g, Z_n \rangle_2 &= \sqrt{n} \left[\left\langle g, \frac{1}{n} \sum_{i=1}^n I_{[X_i, 1]}(\cdot) \right\rangle_2 - \langle g, \lambda[0, \cdot] \rangle_2 \right] \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_{[X_i, 1]} g(t) \lambda(dt) - \sqrt{n} \int_{[0, 1]} g(t) \lambda[0, t] \lambda(dt) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n h(X_i) - \sqrt{n} \int_{[0, 1]} t g(t) \lambda(dt) . \end{aligned}$$

Por (4.1) vem:

$$\begin{aligned} \langle g, Z_n \rangle_2 &= \frac{1}{\sqrt{n}} \sum_{i=1}^n h(X_i) - \sqrt{n} E(h(X_i)) = \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n [h(X_i) - E(h(X_i))]. \end{aligned}$$

Assim o problema de provar a convergência em distribuição dos produtos internos $\langle Z_n, g \rangle_2$, reduz-se à demonstração de um teorema limite central para as variáveis aleatórias $h(X_j)$ com $h \in H_K$. Vimos na secção 1 que as funções de H_K são todas limitadas, então existe uma constante positiva C , que depende apenas de h , tal que $|h(X_j)| < C$ para todo $j \in \mathbb{N}$, isto é, as variáveis aleatórias $h(X_j)$ são uniformemente limitadas.

Pretendemos nesta secção, demonstrar dois teoremas que estabelecem a convergência fraca do processo empírico uniforme, considerando para isso dois tipos de dependência entre as variáveis aleatórias X_n : α -misturação e associação (ver secção 3 do capítulo I). Começamos por tratar o problema para variáveis aleatórias α -misturadas. Para isso vamos usar um teorema limite central de Doukhan, Massart, Rio [6], que aparece aqui numa versão reduzida, adaptada ao caso de variáveis aleatórias uniformemente limitadas.

Teorema 4.1: Seja $\{Y_n\}$ uma sucessão estritamente estacionária e α -misturadora de variáveis aleatórias centradas e uniformemente limitadas. Se $\sum_{n=1}^{\infty} \alpha_n < +\infty$ então,

$$\sigma^2 = E(Y_0^2) + 2 \sum_{n=1}^{\infty} E\{Y_0 Y_n\}$$

é convergente e, se $\sigma > 0$, então $n^{-1/2}(Y_1 + \dots + Y_n)$ converge em distribuição para uma variável aleatória Gaussiana centrada e com variância σ^2 .

Teorema 4.2 [13]: Suponhamos que as variáveis aleatórias X_n , $n \geq 0$ são estritamente estacionárias, α -misturadoras e uniformemente distribuídas no intervalo $[0,1]$. Se:

$$\sum_{n=1}^{\infty} \alpha_n < +\infty, \quad (4.2)$$

então o processo empírico uniforme associado a $\{X_n\}$ converge fracamente em $L^2[0,1]$ para um processo Gaussiano centrado e com função de covariância dada por:

$$\Gamma(s,t) = \min\{s,t\} - st + 2 \sum_{k=1}^{\infty} [P(X_0 \leq s, X_k \leq t) - st]. \quad (4.3)$$

Demonstração:

Como $h \in H_k$ é limitada, as variáveis aleatórias $Y_j = h(X_j) - E(h(X_j))$ são uniformemente limitadas. Além disso, segue-se das hipóteses do teorema que $\{Y_n\}$ é uma sucessão estritamente estacionária de variáveis aleatórias centradas. Como cada Y_j depende apenas de X_j , então a sucessão $\{Y_n\}$ é também α -misturadora. Podemos então aplicar o teorema 4.1 à sucessão $\{Y_n\}$, donde se segue que a sucessão

$$\frac{1}{\sqrt{n}} \left(\sum_{i=1}^n Y_i \right) = \frac{1}{\sqrt{n}} \sum_{i=1}^n [h(X_i) - E(h(X_i))]$$

converge para uma variável aleatória Gaussiana centrada cuja variância é dada por:

$$\sigma^2(h) = \text{Var}(h(X_0)) + 2 \sum_{k=1}^{\infty} \text{Cov}(h(X_0), h(X_k)),$$

onde a série envolvida é convergente.

Pelo teorema 3.2 a compacidade relativa segue-se da convergência uniforme da série:

$$\sum_{n=1}^{\infty} |Q_{n0}(x,y)|,$$

com Q_{n0} definido como anteriormente. Da definição de Q_{n0} e dos coeficientes de α -misturação segue-se que, para todo o $x,y \in [0,1]$:

$$\sum_{n=1}^{\infty} |Q_{n0}(x,y)| \leq \sum_{n=1}^{\infty} \alpha_n < \infty,$$

donde se conclui que a série $\sum_{n=0}^{\infty} |Q_{n0}|$ é uniformemente convergente em $[0,1]^2$, ficando assim estabelecida a compacidade relativa.

Para verificar que a expressão para a covariância do processo limite é dado por (4.3), tomamos $g = \Psi(h)$. Então utilizando o teorema de Fubini,

$$\begin{aligned} E\{h(X_0)^2\} &= \int_{[0,1]} h^2(u) \lambda(du) = \int_{[0,1]} \left(\int_{[u,1]} g(s) \lambda(ds) \right)^2 \lambda(du) \\ &= \int_{[0,1]} \int_{[0,1]} \int_{[0,1]} I_{[0,s]}(u) I_{[0,t]}(u) g(s) g(t) \lambda(du) \lambda(ds) \lambda(dt) \\ &= \int_{[0,1]^2} \min\{s,t\} g(s) g(t) \lambda^2(ds,dt) \end{aligned}$$

e por (4.1) vem:

$$\begin{aligned} E^2\{h(X_0)\} &= E^2 \left\{ \int_{[X_0,1]} g(t) \lambda(dt) \right\} \\ &= \left(\int_{[0,1]} t g(t) \lambda(dt) \right)^2 \\ &= \int_{[0,1]^2} s t g(s) g(t) \lambda^2(ds,dt), \end{aligned}$$

isto é,

$$\text{Var}(h(X_0)) = \int_{[0,1]^2} (\min\{s,t\} - st) g(s) g(t) \lambda^2(ds,dt).$$

Por outro lado, usando o mesmo tipo de raciocínio:

$$E\{h(X_0)\}E\{h(X_k)\} = \int_{[0,1]^2} s t g(s) g(t) \lambda^2(ds,dt)$$

e

$$E\{h(X_0)h(X_k)\} = \int_{[0,1]^2} [P(X_0 \leq s, X_k \leq t) - st] g(s) g(t) \lambda^2(ds,dt),$$

donde

$$\text{Cov}(h(X_0)h(X_k)) = \int_{[0,1]^2} [P(X_0 \leq s, X_k \leq t) - st] g(s) g(t) \lambda^2(ds,dt).$$

Finalmente vem:

$$\sigma^2(h) = \int_{[0,1]^2} \Gamma(s,t) g(s) g(t) \lambda^2(ds,dt) .$$

Seja agora Z um processo Gaussiano em $[0,1]$ com função de covariância Γ . $\langle Z, g \rangle_2$ é uma variável aleatória Gaussiana com variância:

$$E \left\{ \int_{[0,1]} Z(s) g(s) \lambda(ds) \right\}^2 = \int_{[0,1]^2} E\{Z(s)Z(t)\} g(s) g(t) \lambda^2(ds,dt) = \sigma^2(h).$$

Então a sucessão de variáveis aleatórias $\langle Z_n, g \rangle_2$ converge em distribuição para $\langle Z, g \rangle_2$, o que com a compacidade relativa de $\{Z_n\}$ implica que $\{Z_n\}$ converge fracamente para Z . ■

Tratemos agora o caso em que as variáveis aleatórias da sucessão $\{X_n\}$ são associados. Para provar a convergência dos produtos internos, vamos usar o seguinte resultado devido a Newman [11].

Teorema 4.3 [11]: Seja $\{X_n\}$ uma sucessão de variáveis aleatórias estritamente estacionárias e associadas e $Y_n = f(X_n)$, sendo f uma função absolutamente contínua.

Seja $\bar{Y}_n = \bar{f}(X_n)$ onde $\bar{f} = \int_{[0,1]} |f'(u)| du$. Se

$$\int_{[0,1]^2} f'(x) f'(y) \Gamma(x,y) \lambda^2(dx,dy) < +\infty, \quad (4.4)$$

onde Γ é definido por (4.3), então \bar{Y}_1 é de quadrado integrável e

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n (Y_j - E(Y_j))$$

converge fracamente para uma variável aleatória Gaussiana e centrada com variância

$$\sigma^2 = \int_{[0,1]^2} f'(x) f'(y) \Gamma(x,y) \lambda^2(dx,dy).$$

Teorema 4.4 [13]: Seja $\{X_n\}$ uma sucessão de variáveis aleatórias estritamente estacionárias e associadas com distribuição uniforme em $[0,1]$. Se a série em (4.3) converge uniformemente em $[0,1]^2$, então o processo empírico associado a $\{X_n\}$ converge fracamente em $L^2[0,1]$ para um processo Gaussiano e centrado com função covariância dada por (4.3).

Demonstração:

Pelo teorema 3.2 a compacidade relativa do processo empírico decorre da convergência uniforme da série

$$\sum_{k=1}^{+\infty} |P(X_k \leq x, X_0 \leq y) - xy|.$$

Como por hipótese a sucessão $\{X_n\}$ é associada, o termo geral desta série é não negativo, e portanto, trata-se da série envolvida em (4.3), que é uniformemente convergente por hipótese. Está assim provada a compacidade relativa do processo

empírico uniforme $\{Z_n\}$. Resta-nos provar a convergência dos produtos internos. Notemos que:

$$\begin{aligned} \langle Z_n, g \rangle_2 &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \left[\int_{[X_j, 1]} g(u) \lambda(du) - E \left\{ \int_{[X_j, 1]} g(u) \lambda(du) \right\} \right] \\ &= -\frac{1}{\sqrt{n}} \sum_{j=1}^n \left[\int_{[0, X_j]} g(u) \lambda(du) - E \left\{ \int_{[0, X_j]} g(u) \lambda(du) \right\} \right], \end{aligned}$$

pelo que iremos aplicar o teorema anterior com $Y_n = f(X_n) = \int_{[0, X_n]} g(u) \lambda(du)$.

A função f é absolutamente contínua em $[0, 1]$, por ser o integral indefinido de uma função g integrável à Lebesgue em $[0, 1]$. Para aplicar o teorema 4.3 falta verificar a condição (4.4):

$$\int_{[0, 1]^2} g(x) g(y) \Gamma(x, y) \lambda^2(dx, dy) < +\infty.$$

Ora, $\Gamma(x, y) = \min\{x, y\} - xy + 2 \sum_{k=1}^{\infty} Q_{0k}(x, y)$, onde $Q_{0k}(\cdot, \cdot)$ é a função contínua definida

anteriormente. Além disso a função $\theta(x, y) = \sum_{k=1}^{\infty} Q_{0k}(x, y)$ é contínua por ser o limite

uniforme de funções contínuas. Portanto $\Gamma(x, y)$ é contínua, logo é limitada no compacto $[0, 1]^2$, isto é, existe uma constante $c > 0$ tal que $\sup_{(x, y)} \Gamma(x, y) < c$. Tendo ainda

em consideração que $g \in L^2[0, 1]$, a condição (4.4) decorre com facilidade. De facto,

$$\int_{[0, 1]^2} g(x) g(y) \Gamma(x, y) \lambda^2(dx, dy) \leq c \int_{[0, 1]^2} g(x) g(y) \lambda^2(dx, dy) < \infty.$$

Verificadas as condições do teorema 4.3, podemos concluir que $\langle Z_n, g \rangle_2$ converge em distribuição para uma variável aleatória Gaussiana centrada com variância dada por:

$$\sigma^2 = \int_{[0, 1]^2} g(x) g(y) \Gamma(x, y) \lambda^2(dx, dy).$$

O resto da demonstração prossegue exactamente como no teorema 4.2. ■

Salientemos que os resultados demonstrados nesta secção e na anterior, relativos ao processo empírico uniforme, podem facilmente ser estendidos ao caso geral, desde que seja imposta a condição de que as variáveis aleatórias X_n , $n \in \mathbb{N}$, às quais está associado o processo empírico $\{Z_n\}$, tenham uma função de distribuição contínua e concentrada em $[0,1]$. A função de covariância do processo limite passaria a ser:

$$\Gamma(s,t) = \mu[0, \min(s,t)] - \mu[0,s] \mu[0,t] + 2 \sum_{k=1}^{\infty} [P(X_0 \leq s, X_k \leq t) - P(X_0 \leq s)P(X_k \leq s)]$$

onde μ é a distribuição das variáveis aleatórias X_i .

5 - Apresentação de um exemplo concreto

Vimos nas secções anteriores deste capítulo, condições suficientes para a compacidade relativa e para a convergência fraca do processo empírico, no caso em que as variáveis aleatórias $(X_n)_{n \geq 0}$ têm uma distribuição uniforme em $[0,1]$. Como já sabemos, este é o caso mais importante, pois esta é a distribuição das variáveis aleatórias $F(X_n)$, onde F é a função de distribuição contínua de X_n .

Vamos apresentar uma classe de sucessões, para as quais a condição de compacidade relativa do teorema 3.2 (que é também condição suficiente para a convergência do processo empírico quando $(X_n)_{n \in \mathbb{N}}$ é associada), é equivalente a uma condição mais fraca que a obtida por Yu e apresentada no teorema 2.2 do capítulo II. Daremos também um exemplo de uma sucessão pertencente a esta classe que verifica a condição do teorema 3.2, mas não a condição obtida por Yu.

No que se segue, consideramos sempre sucessões de variáveis aleatórias estritamente estacionárias e associadas.

A condição do teorema 3.2 aplicada à sucessão $(F(X_n))_{n \geq 0}$ exige a convergência uniforme em $[0,1]^2$ da série:

$$\sum_{n=0}^{\infty} \left| P(F(X_n) \leq x, F(X_0) \leq y) - P(F(X_n) \leq x) P(F(X_0) \leq y) \right|. \quad (5.1)$$

Antes de mais, vamos ver que esta condição implica que:

$$\sum_{n=1}^{\infty} \text{Cov}(F(X_0), F(X_n)) < \infty \quad (5.2)$$

Utilizando a igualdade de Hoeffding, temos:

$$\sum_{n=1}^{\infty} \text{Cov}(F(X_0), F(X_n)) = \sum_{n=1}^{\infty} \int_{[0,1]^2} P(F(X_0) \leq y, F(X_n) \leq x) - P(F(X_n) \leq x) P(F(X_0) \leq y) \, dx \, dy.$$

Podemos agora aplicar o teorema da convergência monótona, pois a função integranda é não negativa por associação. Assim:

$$\sum_{n=1}^{\infty} \text{Cov}(F(X_0), F(X_n)) = \int_{[0,1]^2} \sum_{n=1}^{\infty} |H_{n0}(x,y)| \, dx \, dy,$$

onde $H_{n0}(x,y) = P(F(X_n) \leq x, F(X_0) \leq y) - P(F(X_n) \leq x) P(F(X_0) \leq y) \geq 0$.

Mais uma vez, a convergência uniforme da série integranda e a continuidade das funções $H_{n0}(\cdot, \cdot)$ implicam que a dita série é uma função limitada em $[0,1]^2$, logo (5.2) é verificada.

Por outro lado, é evidente, atendendo à associação, que a condição obtida por Yu também implica a condição (5.2). Isto é:

$$\sum_{n=1}^{\infty} n^{13/2 + \nu} \text{Cov}(F(X_0), F(X_n)) < +\infty \Rightarrow \sum_{n=1}^{\infty} \text{Cov}(F(X_0), F(X_n)) < +\infty.$$

Seja agora $(X_n)_{n \geq 0}$ um processo estocástico Gaussiano, centrado, estacionário e associado. Por Pitt [17] a associação deste processo é equivalente a: $\text{Cov}(X_i, X_j) \geq 0$, $\forall i, j \geq 0$.

Suponhamos que $\text{Var}(X_n) = 1$, $n \geq 0$, e denotemos por ρ_n o coeficiente de correlação entre X_0 e X_n . Suponhamos ainda que $\rho_n \xrightarrow{n} 0$.

A função densidade de (X_0, X_n) pode escrever-se:

$$f_{0n}(x, y) = \frac{1}{2\pi\sqrt{1-\rho_n^2}} \exp\left(\frac{1}{2(1-\rho_n^2)}(x^2 + y^2 - 2\rho_n xy)\right),$$

e portanto temos:

$$\begin{aligned} \text{Cov}(F(X_0), F(X_n)) &= \\ &= \int_{\mathbb{R}^2} F(x)F(y) f_{0n}(x, y) dx dy - \int_{\mathbb{R}} F(x) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \times \int_{\mathbb{R}} F(y) \\ &\quad \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \\ &= \int_{\mathbb{R}^2} F(x)F(y) \left[f_{0n}(x, y) - \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right) \right] dx dy. \end{aligned}$$

$$\text{Façamos: } A_n(x, y) = f_{0n}(x, y) - \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right).$$

Utilizando a regra de Cauchy e tendo em conta que $\rho_n \xrightarrow{n} 0$, vem:

$$\lim_{n \rightarrow \infty} \frac{A_n(x, y)}{\rho_n} = \frac{xy}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right). \quad (5.3)$$

Pretendemos mostrar que as séries $\sum_{n=1}^{\infty} \text{Cov}(F(X_0), F(X_n))$ e $\sum_{n=1}^{\infty} \text{Cov}(X_0, X_n)$ são da mesma natureza. Para isso basta mostrar que o quociente entre os respectivos termos gerais converge para um limite finito e estritamente positivo, atendendo a que

se tratam de séries de termos não negativos, já que as variáveis aleatórias intervenientes são associadas. Para o estudo do quociente referido necessitamos de invocar o teorema da convergência dominada, o que nos obrigará a majorar uniformemente o quociente $\left| \frac{A_n(x,y)}{\rho_n} \right|$. Para estabelecer essa majoração, vamos considerar a função:

$$B_n(x,y) = 2\pi A_n(x,y) \exp\left(\frac{x^2 + y^2}{2}\right)$$

e procurar para esta um majorante uniforme e um minorante uniforme. Notemos que $B_n(x,y)$ pode ser rescrito da seguinte maneira:

$$B_n(x,y) = \frac{1}{\sqrt{1-\rho_n^2}} \exp\left(-\frac{(x^2 + y^2)\rho_n^2}{2(1-\rho_n^2)} + \frac{\rho_n xy}{1-\rho_n^2}\right) - 1. \quad (5.4)$$

Atendendo à convergência para zero de ρ_n , é fácil verificar que a sucessão de termo geral $\frac{1}{\rho_n^2 \sqrt{1-\rho_n^2}} - \frac{1}{\rho_n^2}$ é convergente. É por isso limitada, donde se conclui que existe uma constante $C_1 > 0$ tal que:

$$\frac{1}{\sqrt{1-\rho_n^2}} \leq C_1 \rho_n^2 + 1.$$

Por outro lado, se $\rho_n \xrightarrow{n} 0$, então existe uma ordem $n_0 \in \mathbb{N}$ a partir do qual se verifica sempre $1 - \rho_n^2 \geq 0$. Então, para $n \geq n_0$, temos:

$$\exp\left(-\frac{(x^2 + y^2)\rho_n^2}{2(1-\rho_n^2)}\right) \leq 1.$$

Estas considerações permitem-nos deduzir que existe uma constante $C_1 > 0$ tal que:

$$B_n(x,y) \leq (C_1 \rho_n^2 + 1) \exp\left(\frac{\rho_n xy}{1-\rho_n^2}\right) - 1, \quad n \geq n_0.$$

Agora, usando a desigualdade $e^u \leq 1 + |u| e^{|u|}$, vem para $n \geq n_0$:

$$B_n(x,y) \leq (C_1 \rho_n^2 + 1) \left[1 + \frac{\rho_n |xy|}{1-\rho_n^2} \exp\left(\frac{\rho_n}{1-\rho_n^2} |xy|\right) \right] - 1,$$

isto é:

$$B_n(x,y) \leq \rho_n \left(C_1 \rho_n + \frac{1+C_1 \rho_n^2}{1-\rho_n^2} |xy| \exp\left(\frac{\rho_n}{1-\rho_n^2} |xy|\right) \right).$$

Notemos que podemos escolher $n_0 \in \mathbb{N}$ de forma a que, para $n \geq n_0$, $0 \leq \rho_n \leq \frac{1}{3}$, o que

implica: $\frac{\rho_n}{1-\rho_n^2} \leq \frac{1}{2}$. Então:

$$\frac{B_n(x,y)}{\rho_n} \leq C_1 \rho_n + \frac{1+C_1 \rho_n^2}{1-\rho_n^2} |xy| \exp\left(\frac{|xy|}{2}\right), \quad n \geq n_0.$$

Finalmente, do facto das sucessões ρ_n e $\frac{1+C_1 \rho_n^2}{1-\rho_n^2}$ serem convergentes, logo

limitadas, é possível escolher constantes C_2 e C_3 positivas tais que:

$$\frac{A_n(x,y)}{\rho_n} \leq \left[C_2 + C_3 |xy| \exp\left(\frac{|xy|}{2}\right) \right] \exp\left(-\frac{x^2+y^2}{2}\right). \quad (5.5)$$

Para encontrar um minorante, consideramos mais uma vez, que existe uma ordem $n_0 \in \mathbb{N}$ a partir da qual $0 \leq \rho_n \leq \frac{1}{3}$. Então, para $n \geq n_0$, temos: $\frac{1}{\sqrt{1-\rho_n^2}} \geq 1$ e de

(5.4) vem:

$$B_n(x,y) \geq \exp\left(\rho_n \frac{xy}{1-\rho_n^2} - \rho_n \frac{(x^2+y^2)\rho_n}{2(1-\rho_n^2)}\right) - 1.$$

Usando agora a desigualdade $e^u \geq 1+u$, $u \in \mathbb{R}$, segue-se que:

$$B_n(x,y) \geq \rho_n \left(\frac{xy}{1-\rho_n^2} - \frac{(x^2+y^2)\rho_n}{2(1-\rho_n^2)} \right),$$

o que implica que:

$$\frac{A_n(x,y)}{\rho_n} \geq \left[\frac{xy}{1-\rho_n^2} - \frac{\rho_n}{2(1-\rho_n^2)}(x^2+y^2) \right] \exp\left(-\frac{x^2+y^2}{2}\right) \times \frac{1}{2\pi}.$$

Atendendo a que $xy \geq -\frac{x^2+y^2}{2}$ vem:

$$\frac{A_n(x,y)}{\rho_n} \geq -\frac{x^2+y^2}{2(1-\rho_n^2)}(1+\rho_n) \exp\left(-\frac{x^2+y^2}{2}\right) \times \frac{1}{2\pi}.$$

Mais uma vez, da convergência de $\frac{1+\rho_n}{2(1-\rho_n^2)}$ segue-se a existência de $C_4 > 0$ tal que:

$$\frac{A_n(x,y)}{\rho_n} \geq -C_4(x^2+y^2) \exp\left(-\frac{x^2+y^2}{2}\right). \quad (5.6)$$

De (5.5) e (5.6) obtém-se a seguinte majoração uniforme:

$$\left| \frac{A_n(x,y)}{\rho_n} \right| \leq C(1+x^2+y^2) \exp\left(-\frac{x^2+y^2}{4}\right). \quad (5.7)$$

Relembremos que o nosso objectivo é provar que as séries $\sum_{n=1}^{\infty} \text{Cov}(X_0, X_n)$ e

$\sum_{n=1}^{\infty} \text{Cov}(F(X_0), F(X_n))$ são da mesma natureza. A majoração obtida em (5.7)

permite-nos então usar o teorema da convergência dominada. Tendo em conta (5.3)

vem:

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{\text{Cov}(F(X_0), F(X_n))}{\text{Cov}(X_0, X_n)} &= \lim_{n \rightarrow +\infty} \int_{\mathbb{R}^2} F(x)F(y) \frac{A_n(x,y)}{\rho_n} dx dy \\ &= \int_{\mathbb{R}^2} F(x)F(y) \frac{xy}{2\pi} \exp\left(-\frac{x^2+y^2}{2}\right) dx dy \end{aligned}$$

$$= \left[\int_{\mathbb{R}} F(x) \frac{x}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \right]^2$$

$$= E^2\{X_n F(X_n)\} \neq 0,$$

o que mostra que as séries são da mesma natureza.

De seguida vamos provar que a condição:

$$\sum_{n=1}^{\infty} \text{Cov}(X_0, X_n) < +\infty, \quad (5.10)$$

implica a convergência uniforme da série referida em (5.1). Temos:

$$H_{n0}(x,y) = P(F(X_n) \leq x, F(X_0) \leq y) - P(F(X_n) \leq x) P(F(X_0) \leq y)$$

$$= E\{I(F(X_n) \leq x) I(F(X_0) \leq y)\} - E\{I(F(X_n) \leq x)\} E\{I(F(X_0) \leq y)\}.$$

Considerando a função $Q(x) = \inf\{y: F(y) \geq x\}$ introduzida no capítulo I, tem-se:

$$I(F(X_i) \leq x) = I(X_i \leq Q(x)) \quad \text{com probabilidade 1.}$$

Então,

$$H_{n0}(x,y) = E\{I(X_0 \leq x') I(X_n \leq y')\} - E\{I(X_0 \leq x')\} E\{I(X_n \leq y')\},$$

com $x'=Q(x)$ e $y'=Q(y)$. Portanto, podemos escrever:

$$H_{n0}(x,y) = \int_{\mathbb{R}^2} I_{[-\infty, x']}(s) I_{[-\infty, y']}(t) f_{0n}(s,t) ds dt -$$

$$- \int_{\mathbb{R}} I_{[-\infty, x']}(s) \frac{1}{\sqrt{2\pi}} \exp(-s^2/2) ds \times \int_{\mathbb{R}} I_{[-\infty, y']}(t) \frac{1}{\sqrt{2\pi}} \exp(-t^2/2) dt$$

$$= \int_{\mathbb{R}^2} I_{[-\infty, x']}(s) I_{[-\infty, y']}(t) A_n(s,t) ds dt.$$

Utilizando a majoração (5.7) obtemos:

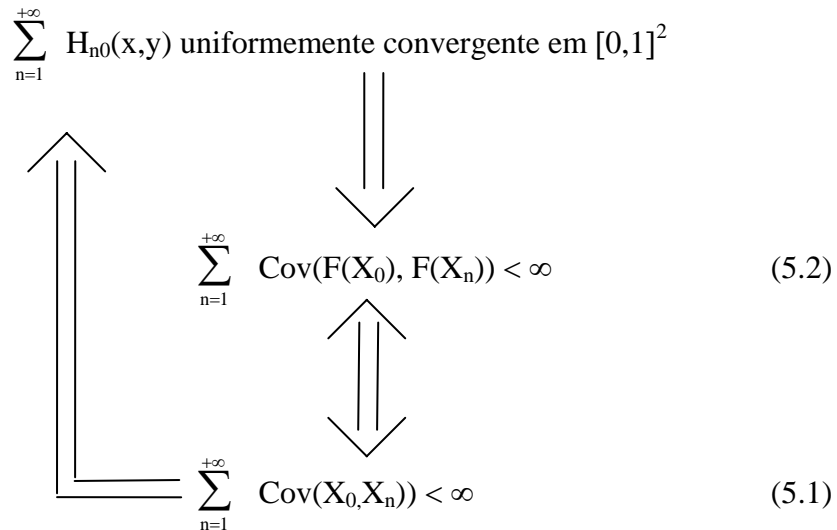
$$0 \leq H_{n0} \leq \rho_n \int_{\mathbb{R}^2} C(1+s^2+t^2) \exp\left(-\frac{s^2+t^2}{4}\right) ds dt.$$

Então existe uma constante C' tal que:

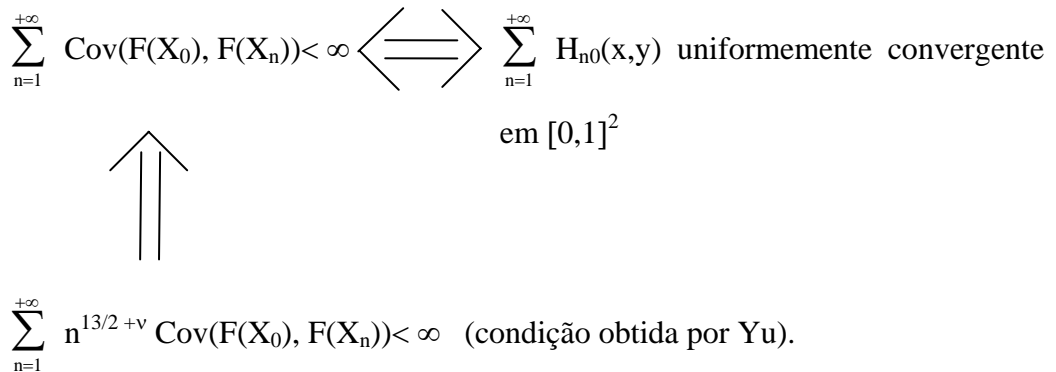
$$\sum_{n=1}^{\infty} H_{n0}(x,y) \leq C' \sum_{n=1}^{\infty} \rho_n, \quad \forall (x,y) \in [0,1]^2.$$

Como por (5.10) a série $\sum_{n=1}^{\infty} \rho_n$ é convergente, então a série $\sum_{n=1}^{\infty} H_{n0}(x,y)$ é uniformemente convergente como pretendíamos demonstrar.

Temos então o seguinte esquema de implicações:



Resumindo:



Obtivemos então uma classe de processos estocásticos de índice discreto, para os quais a condição do teorema 3.2 é equivalente a uma condição mais fraca que a obtida por Yu. Isto permite-nos encontrar dentro desta classe, exemplos de processos estocásticos que verificam a condição do teorema 3.2, mas não a condição obtida por Yu. É o caso, por exemplo, do processo estocástico Gaussiano e centrado $(X_n)_{n \geq 0}$ com função de covariância dada por: $\text{Cov}(X_i, X_j) = \frac{1}{1 + (j-i)^2}$, $i, j \in \mathbb{N}$.

Segundo Pitt [17], as variáveis aleatórias são associadas pois $\text{Cov}(X_i, X_j) \geq 0$ para todo $i, j \in \mathbb{N}$. Além disso, é um processo estacionário cuja sucessão $\rho_n = \text{Cov}(X_0, X_n)$ converge para zero. É portanto um elemento da classe de processos estocásticos em estudo. Provámos anteriormente que o limite de $\frac{\text{Cov}(F(X_0), F(X_n))}{\text{Cov}(X_0, X_n)}$ é uma constante real positiva, então claramente:

$$\text{Cov}(F(X_0), F(X_n)) = c \text{Cov}(X_0, X_n) = c (1+n^2)^{-1},$$

onde $c \in \mathbb{R}^+$.

A série $\sum_{n=1}^{\infty} \text{Cov}(F(X_0), F(X_n)) = c \sum_{n=1}^{\infty} \frac{1}{1+n^2}$ é convergente. Isto equivale a dizer, como vimos anteriormente, que a série $\sum_{n=1}^{\infty} H_{n0}(x, y)$ é uniformemente convergente em $[0, 1]^2$, logo a condição do teorema 3.2 é verificada.

Da implicação que se segue,

$$\sum_{n=1}^{\infty} n^{13/2+\nu} \text{Cov}(F(X_0), F(X_n)) < +\infty \Rightarrow \sum_{n=1}^{\infty} n^{1/3} \text{Cov}(F(X_0), F(X_n)) < +\infty,$$

resulta, que se a série do lado direito da implicação não for convergente, a condição obtida por Yu não é verificada. De facto, utilizando o critério da razão para séries de

termos não negativos, é fácil verificar que a série,

$$\sum_{n=1}^{\infty} n^{1/3} \text{Cov}(F(X_0), F(X_n)) = c^{1/3} \sum_{n=1}^{\infty} \frac{1}{(1+n^2)^{1/3}}$$

é divergente, logo a condição de Yu não é verificada por este processo estocástico.

Acabamos de apresentar um exemplo de uma sucessão de variáveis aleatórias $(X_n)_{n \geq 0}$ cuja convergência fraca do processo empírico que lhe está associado, pode ser estabelecida pelos teoremas apresentados neste capítulo, mas não pelos teoremas apresentados no capítulo II. Na secção seguinte apresentam-se algumas aplicações do estudo feito neste capítulo.

6. Algumas aplicações

Para apresentar uma das aplicações, necessitamos de recorrer a alguns resultados devidos a Suquet, relativos a espaços auto-reprodutores e medidas aleatórias [22]. Por isso, expomos de seguida, uma síntese desses resultados.

Seja X um espaço topológico, \mathbf{B}_X a tribo de Borel que lhe está associada e M o espaço das medidas com sinal, limitadas, definidas sobre o espaço mensurável (X, \mathbf{B}_X) . Seja H_K o espaço auto-reprodutor associado a um núcleo K . Por Suquet [22], a aplicação Φ definida por:

$$\Phi: M \longrightarrow H_K$$

$$\mu \longrightarrow \Phi(\mu) = \int K(\cdot, t) \mu(dt),$$

é injectiva desde que a seguinte implicação se verifique:

$$\int K(s, t) \mu \otimes \mu(ds, dt) = 0 \Rightarrow \mu \equiv 0. \quad (6.1)$$

Se considerarmos $X=[0,1]$, K o núcleo definido em (1.1) e M o espaço das medidas com sinal, limitadas e que dão massa zero ao ponto 1, então a implicação (6.1) verifica-se o que garante a injectividade de Φ .

A aplicação Φ permite-nos obter o produto escalar de H_K sob a forma de um integral, se pelo menos um dos elementos envolvidos no cálculo do produto escalar pertencer a $\Phi(M)$ [22]:

$$\forall f \in H_K \quad \forall \mu \in M \quad \langle f, \Phi(\mu) \rangle_K = \int_{[0,1]} f \, d\mu . \quad (6.2)$$

Então para todas as medidas $\mu \in M$ temos, por aplicação do teorema de Fubini:

$$\begin{aligned} \Phi(\mu)(s) &= \int_{[0,1]} K(s,t) \, \mu(dt) = \\ &= \int_{[0,1]} \int_{[0,1]} I_{[s,1]}(u) \, I_{[t,1]}(u) \, \lambda(du) \, \mu(dt) \\ &= \int_{[s,1]} \int_{[0,1]} I_{[0,u]}(t) \, \mu(dt) \, \lambda(du) \\ &= \int_{[s,1]} \mu[0,u] \, \lambda(du) \end{aligned} \quad (6.3)$$

Estamos agora em condições de apresentar a primeira aplicação do estudo feito nas secções anteriores.

Seja F uma função de distribuição concentrada em $[0,1]$. Definimos a funcional:

$$T(F) = \int g(x) \, F(dx),$$

para alguma função g para a qual aquele integral exista qualquer que seja a função de distribuição F . Seja F_n a função de distribuição empírica associada a F , então por (6.2)

vem para $g \in H_K$:

$$\sqrt{n} [T(F) - T(F_n)] = \sqrt{n} \int_{[0,1]} g(x)(F - F_n)(dx) = \sqrt{n} \langle g, \Phi(\mu_{F-F_n}) \rangle_K, \quad (6.4)$$

onde μ_{F-F_n} é a medida de Lebesgue-Stieltjes associada a $(F - F_n)$. Mas por (6.3), para todo o $s \in [0,1]$:

$$\Phi(\mu_{F-F_n})(s) = \int_{[s,1]} \mu_{F-F_n} [0, u] \lambda(du) = \int_{[s,1]} (F-F_n)(u) \lambda(du) = \Psi^{-1}(F-F_n)(s),$$

onde Ψ é a isometria entre H_K e $L^2[0,1]$ introduzida na secção 1 deste capítulo. Retomemos a igualdade (6.4):

$$\sqrt{n} [T(F) - T(F_n)] = \sqrt{n} \langle g, \Psi^{-1}(F-F_n) \rangle_K = \sqrt{n} \langle -g', (F-F_n) \rangle_2 = \langle g', Z_n \rangle_2,$$

onde Z_n é o processo empírico associado a F e $g(s) = \int_{[s,1]} -g'(t) \lambda(dt) = \Psi^{-1}(-g')$.

Então os teoremas (4.2) e (4.4) podem ser usados para estabelecer a convergência fraca de $\sqrt{n} [T(F) - T(F_n)]$ para $\langle g', Z \rangle_2$, onde Z é o limite em $L^2[0,1]$ do processo empírico. Estas funcionais são casos particulares das funcionais de Von Mises.

Outra aplicação do estudo feito neste capítulo é o teste estatístico de Cramer-Von Mises ω^2 . Este pretende testar a hipótese de F_0 ser a função da distribuição das variáveis aleatórias $X_n, n \in \mathbb{N}$, e é definido por:

$$\omega_n^2 = n \int_{[0,1]} (F_n(t) - t)^2 dt,$$

onde F_n é a função da distribuição empírica de $F_0(X_1), \dots, F_0(X_n)$. Da definição de ω_n^2 , é imediato que este é a norma em $L^2[0,1]$ do processo empírico uniforme. Portanto, verificadas as condições dos teoremas (4.2) ou (4.4), podemos concluir que ω_n^2 converge em distribuição para $\omega^2 = \int B^2(t) dt$, onde B é o limite em $L^2[0,1]$ do processo empírico uniforme.

Seguindo agora Khmaladze [8], suponhamos que temos uma família de funções $\{F(x,\theta), \theta \in \Theta\}$ e pretendemos fazer um teste de identificação da função de distribuição F dentro dessa família. O teste estatístico estudado em [8] é:

$$\hat{\omega}_n^2 = n \int (F_n(t) - G(t, \hat{\theta}_n))^2 G(dt, \hat{\theta}_n),$$

onde $G(t, \theta) = F(F^{-1}(t, \theta_0), \theta)$, $\hat{\theta}_n$ é um estimador do parâmetro θ e θ_0 o valor a ser testado. Então, impondo algumas condições técnicas, indicadas em [8], e considerando $\mu_n(t) = \sqrt{n} [F_n(t) - G(t, \hat{\theta}_n)]$, é possível verificar que a convergência do processo empírico, implica a convergência em distribuição em $L^2[0,1]$ de $\mu_n(t)$ para o processo $B(t) + g(t) \int l'(t) B(t) dt$, onde $g(t) = \frac{\partial G}{\partial \theta}(t, \theta_0)$ e $l(t) = l(F^{-1}(t, \theta_0), \theta_0)$ para alguma função $l(x, \theta)$ verificando as condições impostas em [8].

Referências

- [1] Aronszajn, N., La théorie des noyaux reproduisants et applications, Proceedings of the Cambridge Philosophical Society 39 (1943), 133-153.
- [2] Aronszajn, N., The theory of reproducing kernels, Transactions of the American Mathematical Society 68 (1950) 337-404.
- [3] Billingsley, P., Convergence of probability measures, Wiley (1968).
- [4] Burton, R. M., Dabrowski, A. R., Dehling, H., An invariance principle for weakly associated random variables, Stochastic Processes Appl. 23 (1986), 301-306.
- [5] Davydov, Y., Convergence of distributions generated by stationary stochastic processes, Th. Probab. Appl. 13 (1968), 691-696.
- [6] Doukhan, P., Massart, P., Rio, E., The functional central limit theorem for strongly mixing processes, Ann. Inst. Henri Poincaré, Probab. Stat., 30 (1994), 63-82.
- [7] Ibragimov, I. A., Some limit theorems for stationary processes, Th. Probab. Appl. 7 (1962), 349-382.
- [8] Khmaladze, E. V., The use of ω^2 tests for testing parametric hypothesis, Th. Probab. Appl. 24 (1979), 283-301.
- [9] Kreyszig, E., Introductory functional analysis with applications, Wiley (1978).
- [10] Lehman, E., Some concepts of dependence, Ann. Math. Statist. 37 (1966), 1137-1153.
- [11] Newman, C., Asymptotic independence and limit theorems for positively and negatively dependent random variables, Inequalities in Statistics and Probability, IMS Lect. Notes - Monograph Series 5 (1984), 127-140.
- [12] Oliveira, P. E., Suquet, C., Auto-reproducing spaces and invariance principles in $L^2[0,1]$, Publ. IRMA Lille 32 (1993), III.

- [13] Oliveira, P. E., Suquet, C., $L^2[0,1]$ weak convergence of the empirical process for dependent variables, Actes des XVèmes Rencontres Franco-Belges de Statisticiens (Ondelettes et Statistique), Lecture Notes in Statistics 103, Wavelets and Statistics, Ed. A. Antoniadis, G. Oppenheim, (1995).
- [14] Oliveira, P. E., Suquet, C., Empirical process under positive dependence in $L^p[0,1]$, Publ. IRMA Lille 37 (1995), IV, preprint.
- [15] Oliveira, P. E., Suquet, C., An invariance principle in $L^2[0,1]$ for non stationary ϕ -mixing sequences, Comment. Math. Univ. Carolinae 36 (1995), 293-302.
- [16] Parthasaraty, K. R., Probability measures on metric spaces, Academic Press (1967).
- [17] Pitt, L., D., Positively correlated normal variables are associated, Ann. Probab. 10 (1982), 496-499.
- [18] Prokhorov, Y. V., Convergence of random processes and limit theorems in probability theory, Theory Probab. Appl. 1 (1956), 157-214.
- [19] Semadeni, Z., Schauder bases in Banach spaces of continuous functions, Springer (1982).
- [20] Suquet, C., Relectures des critères de relative compacité d'une famille de probabilités sur un espace de Hilbert, Publ. IRMA Lille 28-III (1992), preprint.
- [21] Suquet, C., Distances euclidiennes sur les mesures signées et application à des théorèmes de Berry-Esséen, Publ. IRMA Lille 34-IV (1994), preprint.
- [22] Suquet, C., Espaces autoreproduisants et mesures aléatoires, Thèse de 3^o cycle, Lille (1986).
- [23] Yu, H., A Glivenko-Cantelli lemma and weak convergence for empirical processes of associated sequences, Probab. Theory Relat. Fields 95 (1993), 357-370.