Cecília Maria Vasconcelos Costa e Castro de Azevedo

Estimação do Operador de Covariância do Limite do Processo Empírico

Dissertação apresentada para obtenção do grau de Doutor em Ciências.

Universidade do Minho
Braga, Maio de 2002

Conteúdo

Resumo vii

Abstract ix

Introdução xi

1 Preliminares 1

1.1 Definições e Propriedades 3

1.2 O Lema de Hoeffding 5

1.2.1 Independência e Associação. A desigualdade de Newman. 8

1.2.2 A generalização de Yu. 13

1.3 Outras definições de associação 17

1.4 Alguns exemplos 21

1.5 Máximo das somas parciais. 27

1.6 A Lei Forte dos Grandes Números. 28

1.7 Convergência do Processo Empírico 35

1.7.1 Introdução 35

1.7.2 Convergência em $L^2[0,1]$ 44

2 A função de distribuição bivariada 53

2.1 Definições e Hipóteses Prévias. 55

2.2 Consistência do Estimador 57

2.2.1 Uma nota sobre velocidades de convergência. . . . 65
2.3 O Erro Quadrático Médio do Estimador 75
2.4 Convergência em distribuição de $\widehat{F}_{n,k}$ 90

3 O operador de covariância **105**
3.1 Definições e resultados preliminares 112
3.2 Consistência de $\widehat{\Gamma}_n$ 120
3.3 Os valores próprios de $\widehat{\Gamma}_n$ 128
3.4 Um método numérico 133

4 Ilustração dos resultados **137**
4.1 Preliminares 138
4.1.1 Optimização de $\{h_n\}_{n\in\mathbb{N}}$. 138
4.1.2 A escolha para $\{q_n\}_{n\in\mathbb{N}}$. 141
4.1.3 As funções núcleo. 142
4.2 Simulações 143
4.3 O maior valor próprio. 159
4.3.1 Conclusões 161

5 Apêndice **163**
5.1 Programas para as aproximações das funções 163
5.2 Cálculo do maior valor próprio 168

Bibliografia **171**

É autorizada a reprodução integral desta tese apenas para efeitos de investigação.

À Carolina, Marta e Cai

Agradecimentos

Ao terminar esta dissertação, agradeço a todos os que, das mais variadas formas, contribuíram para a sua realização. Quero, no entanto, enfatizar as pessoas e entidades que, pela sua relevância, não esquecerei.

Ao Professor Doutor Paulo Eduardo Oliveira quero agradecer ter aceite orientar esta tese, bem como todo o apoio científico concedido nos assuntos que estudei. Sabiamente me soube indicar onde estaria o caminho que pretendia na resolução de problemas que foram surgindo.

À Professora Doutora Emilia Athayde agradeço a disponibilidade que sempre teve para comigo, assim como a sua preocupação em me facultar todos os meios técnicos necessários ao desenvolvimento deste trabalho.

Aos meus colegas e amigos envio uma palavra de amizade por todas as formas de apoio que me proporcionaram.

Agradeço também ao Departamento de Matemática da Universidade do Minho todo o apoio financeiro bem como a concessão da dispensa de serviço, sem a qual a elaboração desta tese não seria, certamente, possível.

Ao CMAT – Centro de Matemática da Universidade do Minho agradeço todo o apoio financeiro prestado, bem como todos os meios científicos disponibilizados.

Cecília Azevedo

Braga, Maio de 2002.

Resumo

Com este trabalho pretendemos contribuir para a obtenção de majorações das probabilidades de cauda de estatísticas de decisão que dependem do processo empírico quando este é construído a partir de variáveis aleatórias com um determinado tipo de dependência positiva. A estatística de Cramèr-von Mises sendo, como se sabe, a norma do processo empírico relativamente à topologia de L^2, serviu-nos como referência principal neste trabalho. No entanto, a maior parte dos resultados obtidos podem ser aplicados a outras estatísticas. Sob algumas condições de regularidade da sucessão associada e estritamente estacionária $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias reais com função de distribuição F, a convergência em distribuição do processo empírico em $L^2[0,1]$ provada por Oliveira e Suquet em 1998, abre caminho ao estudo das probabilidades de cauda atrás referidas, via processo limite deste processo empírico. Tal processo limite é um processo Gaussiano centrado com função de covariância Γ dada por

$$\begin{aligned}\Gamma(s,t) &= F(\min\{s,t\}) - F(s)F(t) + \\ &+ \sum_{k=1}^{\infty}\big(F_k(s,t) - F(s)F(t)\big) + \sum_{k=1}^{\infty}\big(F_k(t,s) - F(s)F(t)\big), \quad (1)\end{aligned}$$

onde F_k representa a função de distribuição de (X_1, X_{k+1}).

Devido à sua complexidade, esta função necessita de ser aproximada. Os seus valores próprios, em particular o seu maior valor próprio, revelam-se fundamentais para a caracterização daquelas probabilidades

de cauda. Assim, e numa primeira fase, estudamos um estimador do tipo núcleo para Γ, partindo do estudo de estimadores para as funções de distribuições bivariadas dos pares associados (X_1, X_{k+1}), com $k \geq 1$, que aparecem na função de covariância. Este estudo envolve as propriedades habituais de consistência e de convergência das distribuições de dimensão finita. Todos os resultados relativos a convergências quase certas ou em média quadrática são acompanhados das respectivas velocidades. Obtemos, ainda, a ordem óptima de convergência da sucessão de comprimentos de janelas que surge na definição do estimador, generalizando as caracterizações unidimensionais. Relativamente ao estimador para a função Γ fazemos, também, um estudo da sua consistência e da ordem óptima do número de termos a somar na estimação da série envolvida naquela função, com base nas velocidades de convergência que foram sendo obtidas. O facto de, como referimos anteriormente, as probabilidades de cauda dependerem do maior valor próprio desta função, leva-nos à implementação de um programa numérico para o cálculo do maior valor próprio dos estimadores que obtemos com base em simulações. É levado a cabo, ainda, um estudo dos erros cometidos quando aproximamos os valores próprios da função de covariância pelos valores próprios dos estimadores.

Abstract

With this work we intend to give a contribution for the computation of critical limits in tests depending on statistics which are functions of the empirical process based on sequences of random variables with some kind of positive dependence. Most of this work was carried having in mind the Cramèr-von Mises statistic, although most of the early results should be applicable to others statistics. The Cramèr-von Mises statistic is, as it is well known, the L^2-norm of the empirical process. Under some regularity conditions on the associated and strictly stationary sequence of real random variables $\{X_n\}_{n\in\mathbb{N}}$, with distribution function F, the weak $L^2[0,1]$ convergence proved in 1998 by Oliveira and Suquet, allows us to study those critical limits. This limit process is a zero-mean Gaussian process on $[0,1]$ with covariance Γ defined by

$$\begin{aligned}\Gamma(s,t) &= F(\min\{s,t\}) - F(s)F(t) + \\ &+ \sum_{k=1}^{\infty}\big(F_k(s,t) - F(s)F(t)\big) + \sum_{k=1}^{\infty}\big(F_k(t,s) - F(s)F(t)\big), \quad (2)\end{aligned}$$

where F_k is the distribution function of (X_1, X_{k+1}). The functions appearing in (2) being unknown, the need to approximate this function arises. Moreover, its eigenvalues, and in particular the largest one, are fundamental to characterize upper bounds for the critical limits. We obtain these upper bounds via a Zolatarev's result and by using classical results for covariance operators. Firstly, we define a kernel-type estimator for Γ,

by studying estimators, of the same type, for the bivariate distribution function F_k. This study concerns the usual properties of consistency and finite dimensional functions. In each type of convergence we also characterize the convergence rates. The optimal order of convergence of the bandwidth sequence is also obtained, therefore generalizing unidimensional results. In what concerns the estimator of Γ we obtain the optimal number of terms that we have to sum for its construction. Further interest on the magnitude's order of the largest eigenvalue of the operator defined by the covariance (2) suggested a computational approach for the calculation of this eigenvalue for the estimators obtained with simulations. The behaviour of the distance between approximated eigenvalues and the true ones is also addressed.

Introdução

O tratamento do processo empírico ou de suas funcionais tem sido um campo importante em estatística, quer do ponto de vista teórico quer no que respeita a aplicações. De facto, muitos dos testes de hipóteses que são conhecidos dependem de uma estatística de decisão definida a partir do processo empírico. São exemplos a estatística de Cramèr-von Mises, a estatística de Watson, a estatística de Mallows "pesada", a estatística de Shepp e algumas estatísticas de Anderson e Darling. Assim, o estudo de propriedades assimptóticas do processo empírico é de grande importância para a justificação de boas propriedades para este tipo de testes. É bem conhecido, na literatura, o comportamento do processo empírico quando este é construído a partir de variáveis aleatórias independentes. No entanto, esta hipótese de independência, por ser pouco realista nas aplicações, foi sendo deixada cair, começando a surgir primeiramente resultados supondo condições de mistura, que mereceram a atenção desde a década de 60; mais recentemente, as condições de dependência positiva ou negativa e de associação em particular, têm merecido a atenção de diversos autores. No entanto, quando o processo empírico é construído a partir de variáveis aleatórias dependentes, a função de covariância do processo estocástico limite é bastante mais complexa, precisando também ela de ser aproximada. De facto, o conhecimento de boas aproximações desta função de covariância, permitirá iniciar o estudo das probabilidades de cauda que aparecem na teoria dos testes de hipóteses, caracterizando

o nível de significância do teste. Estas probabilidades dependem essencialmente dos valores próprios do operador integral associado à função de covariância do processo limite. Com este trabalho de investigação pretendemos prioritariamente contribuir para a obtenção de majorações das probabilidades de cauda da estatística de Cramèr-von Mises. Para tal, efectuamos o estudo da estimação da função de covariância do processo limite quando as variáveis aleatórias de base são associadas, usando famílias de estimadores do tipo núcleo. Derivamos as propriedades habituais de consistência e de caracterização dos limites em distribuição de leis de dimensão finita. A obtenção de velocidades para estas convergências permite-nos controlar os erros de aproximação entre os valores próprios da função de covariância estimada e da verdadeira função de covariância do processo limite. A majoração para as probabilidades de cauda da estatística de Cramèr-von Mises depende fundamentalmente do maior valor próprio do operador associado à covariância do processo limite, pelo que desenvolvemos um programa numérico de cálculo do maior valor próprio para as estimativas que obtemos com base em amostras simuladas de sucessões associadas de variáveis aleatórias.

Devemos ainda referir que efectuamos um esforço de redacção, incluindo vários resultados genéricos, no sentido de tornar o texto o mais autocontido possível.

Capítulo 1

Preliminares

Consideremos $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão de variáveis aleatórias reais num espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. Dizemos que as variáveis aleatórias $X_1, \ldots, X_n$, são associadas se a covariância entre pares de funções crescentes dessas variáveis é não negativa.

O conceito de associação foi introduzido em primeiro lugar no âmbito da Estatística, por Esary, Proschan e Walkup [13], 1967, como uma generalização da definição de dependência positiva entre duas variáveis aleatórias reais (Lehmann [25], 1966) a uma colecção de variáveis, tendo como objectivo a aplicação em Teoria de Fiabilidade. Num contexto completamente diferente, no âmbito da Física Matemática, este conceito surge com as desigualdades FKG (Fortuin, Kasteleyn e Ginibre [14], 1971).

O interesse por esta forma de dependência positiva, motivada em grande parte pela Mecânica Estatística, tem levado a inúmeros resultados probabilísticos. De facto, certas sucessões de variáveis aleatórias associadas, sujeitas a condições especiais na estrutura de covariância entre elas, comportam-se assimptoticamente como variáveis aleatórias independentes. Desta forma têm sido obtidos teoremas limites relativos a somas parciais de sucessões de tais variáveis tal como a Lei dos Grandes Núme-

ros (Newman [31], 1980, Birkel [5], 1984), o Lema de Glivenko - Cantelli (Yu [53], 1994), Teoremas do Limite Central (Birkel [6], 1983, Newman [30], 1984, Oliveira [33], 2000) sendo posteriormente acompanhados por resultados relativos a velocidades de convergência e princípios de invariância, tais como de Newman e Wright [32], 1981, Cai e Roussas [7], 1996 e Oliveira e Suquet [34], 1996.

No que diz respeito à convergência do processo empírico, peça fundamental no desenvolvimento do nosso trabalho, Yu [53], 1993, foi o primeiro a apresentar um Teorema limite acerca da convergência em distribuição da sucessão de processos empíricos construídos à custa de variáveis associadas vistos como elementos aleatórios de $D[0,1]$. Mais tarde, Shao e Yu [43], 1996, enfraquecem a condição suficiente obtida por Yu. Muito recentemente, Louhichi [27], 2000, desenvolve um critério de fineza para o processo empírico construído a partir de qualquer sucessão estacionária de variáveis aleatórias associadas, obtendo, também um Teorema limite para a convergência fraca da sucessão de processos empíricos em $D(-\infty,+\infty)$. Por outro lado, Oliveira e Suquet [35, 36], 1995, 1998, apresentam condições suficientes para a sua convergência em $L^2[0,1]$. Este tipo de convergência é particularmente importante no nosso trabalho, na medida em que pretendemos contribuir para o estabelecimento de limites críticos em Estatísticas que apenas envolvem a norma em L^2 do processo empírico, como foi já referido na introdução.

Este primeiro capítulo é essencialmente baseado nos artigos de Newman [31], 1984, Suquet [48], 1994 e Oliveira e Suquet [36], 1998, e tem como objectivo a apresentação de alguns resultados importantes desta forma de dependência positiva, a associação, pondo em evidência a analogia com a independência. Seleccionamos apenas os resultados que, duma maneira ou de outra, constituem os alicerces do trabalho que desenvolvemos nos capítulos seguintes.

1.1 Definições e Propriedades

As definições gerais de associação positiva e negativa são duas generalizações possíveis ao caso multivariado da definição de Lehmann [25], 1966.

Definição 1.1.1 (Lehmann [25], 1966) *Duas variáveis aleatórias X e Y dizem-se positivamente dependentes se, para todo x e $y \in \mathbb{R}$,*

$$P\big(X > x, Y > y\big) - P\big(X > x\big)P\big(Y > y\big) \geq 0,$$

sendo negativamente dependentes se, para todo x e $y \in \mathbb{R}$,

$$P\big(X > x, Y > y\big) - P\big(X > x\big)P\big(Y > y\big) \leq 0.$$

No entanto, a dependência positiva ou negativa, formulada apenas para um par de variáveis aleatórias revela-se insuficiente nas aplicações. Assim, tal como referimos anteriormente, Esary, Proschan e Walkup [13], 1967, estendem a definição de dependência positiva, ou associação positiva, para uma mais exigente e Joag-Dev e Proschan [22], 1983, introduzem a associação negativa. Esary, Proschan e Walkup começam por mostrar, no seu Teorema 4, que a condição na definição de Lehmann é equivalente a $\mathrm{Cov}\Big(f(X), g(Y)\Big) \geq 0$, para todo o par (f, g) de funções crescentes, generalizando seguidamente a definição a vectores de variáveis aleatórias. Em todo este trabalho usamos apenas associação positiva, pelo que nos referimos apenas ao termo associação.

Definição 1.1.2 *Uma aplicação f de $\mathbb{R}^n$ em $\mathbb{R}$ diz-se crescente se para todo $i = 1, \ldots, n$, a aplicação de $\mathbb{R}$ em $\mathbb{R}$:*

$$t \longmapsto f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n) \quad \textit{é crescente.}$$

Definição 1.1.3 (Esary, Proschan e Walkup [13], 1967) *Um vector aleatório $\mathbf{X} = (X_1, \ldots, X_n)$ diz-se associado se*

$$\mathrm{Cov}\,\big(f(\mathbf{X}), g(\mathbf{X})\big) \geq 0$$

para todo o par de funções crescentes f e g de $\mathbb{R}^n$ em $\mathbb{R}$ sempre que esta covariância estiver definida.

Nota 1.1.4 *Na definição de associação podemos substituir cada par de funções crescentes por um par de funções decrescentes.*

Passamos agora a apresentar algumas propriedades construtivas e fundamentais da associação que derivam exclusivamente da definição (Esary, Proschan e Walkup [13], 1967).

Propriedade 1.1.5 *Toda a subsucessão de uma sucessão finita de variáveis aleatórias associadas é associada;*

Esta propriedade permite estender a definição de associação a sucessões infinitas de variáveis aleatórias.

Definição 1.1.6 *Uma família infinita de variáveis aleatórias é associada se toda a sua subfamília finita o for.*

Propriedade 1.1.7 *Se dois conjuntos de variáveis aleatórias associadas são independentes um do outro, então a reunião dos dois é um conjunto associado;*

Demonstração: Sejam $\mathbf{X} = (X_1, \ldots, X_n)$ e $\mathbf{Y} = (Y_1, \ldots, Y_k)$ dois vectores associados tais que $\mathbf{X}$ e $\mathbf{Y}$ são independentes. Sejam, ainda, f e g duas funções de $\mathbb{R}^{n+k}$ para $\mathbb{R}$, crescentes. Denote-se por $\mathbf{Z}$ o vector $(\mathbf{X}, \mathbf{Y})$ e por $\mathbf{U}$ e $\mathbf{V}$ as imagens de $\mathbf{Z}$ através das funções f e g, respectivamente. Pretende-se mostrar que $\mathrm{Cov}\big(f(\mathbf{Z}), \mathbf{g}(\mathbf{Z})\big) \geq \mathbf{0}$, ou seja, $\mathrm{Cov}\big(\mathbf{U}, \mathbf{V}\big) \geq 0$. Ora,

$$\mathrm{Cov}\,(\mathbf{U}, \mathbf{V}) = \mathbb{E}\Big(\mathrm{Cov}\,\big((\mathbf{U}, \mathbf{V})\big|\mathbf{X}\big)\Big) + \mathrm{Cov}\,\Big(\mathbb{E}(\mathbf{U}\big|\mathbf{X}), \mathbb{E}(\mathbf{V}\big|\mathbf{X})\Big).$$

Como os vectores $\mathbf{X}$ e $\mathbf{Y}$ são independentes,

$$\mathbb{E}\Big(\mathrm{Cov}\,\big((\mathbf{U}, \mathbf{V})\big|\mathbf{X}\big)\Big) = \int \mathrm{Cov}\,\big(f(\mathbf{x}, \mathbf{Y}), g(\mathbf{x}, \mathbf{Y})\big) P_{\mathbf{X}}(\mathbf{dx}).$$

Uma vez que o vector $\mathbf{Y}$ é associado, $\mathrm{Cov}\big(f(\mathbf{x},\mathbf{Y}),g(\mathbf{x},\mathbf{Y})\big)\geq 0$ e, consequentemente, $\mathbb{E}\Big(\mathrm{Cov}\big((\mathbf{U},\mathbf{V})|\mathbf{X}\big)\Big)\geq 0$.
Por outro lado, sendo as aplicações que a cada $\mathbf{x}\in\mathbb{R}^n$ fazem corresponder $\mathbb{E}(f(\mathbf{x},\mathbf{Y})|\mathbf{X})$ e $\mathbb{E}(g(\mathbf{x},\mathbf{Y})|\mathbf{X})$, crescentes e uma vez que $\mathbf{X}$ é associado, temos,

$$\mathrm{Cov}\Big(\mathbb{E}(\mathbf{U}|\mathbf{X}),\mathbb{E}(\mathbf{V}|\mathbf{X})\Big)\geq 0.$$

Pelo que $\mathrm{Cov}\,(\mathbf{U},\mathbf{V})\geq 0$, como pretendíamos. ■

Propriedade 1.1.8 *Se um vector $\mathbf{X}=(X_1,\ldots,X_n)$ é associado e se $f_1,\ldots,f_k$ são funções crescentes (ou decrescentes) de $\mathbb{R}^n$ para $\mathbb{R}$, então o vector $\mathbf{Y}=\big(f_1(\mathbf{X}),\ldots,f_k(\mathbf{X})\big)$ é também associado.*

Demonstração: Seja $\mathbf{X}=(X_1,\ldots,X_n)$ um vector associado e sejam $f_1,\ldots,f_k$ funções crescentes de $\mathbb{R}^n$ para $\mathbb{R}$. Considere-se $\mathbf{Y}=\big(f_1(\mathbf{X}),\ldots,f_k(\mathbf{X})\big)$.
Sejam f e g duas funções crescentes de $\mathbb{R}^k$ para $\mathbb{R}$. Pretende-se mostrar que $\mathrm{Cov}\big(f(\mathbf{Y}),g(\mathbf{Y})\big)\geq 0$. Considere-se $F_k=(f_1,\ldots,f_k)$. Como $(f\circ F_k)$ e $(g\circ F_k)$ são funções crescentes de $\mathbb{R}^n$ para $\mathbb{R}$, sendo $\mathbf{X}$ associado, $\mathrm{Cov}\big((f\circ F_k)(\mathbf{X}),(g\circ F_k)(\mathbf{X})\big)\geq 0$. Ora $(f\circ F_k)(\mathbf{X})=f(\mathbf{Y})$ e $(g\circ F_k)(\mathbf{X})=g(\mathbf{Y})$, pelo que $\mathrm{Cov}\big(f(\mathbf{Y}),g(\mathbf{Y})\big)\geq 0$. Logo $\mathbf{Y}$ é um vector associado. ■

Nota 1.1.9 *Esta propriedade fornece um método construtivo para a obtenção de vectores associados e é a base dos métodos de simulação que utilizamos no último capítulo deste trabalho.*

1.2 O Lema de Hoeffding

Com o objectivo de apresentarmos mais propriedades importantes deste tipo de dependência entre variáveis aleatórias reais, as quais permitirão apresentar um conjunto alargado de exemplos de associação, apresentamos uma importante igualdade atribuída a Hoeffding mas evidenciada

pela primeira vez em Lehmann [25], 1966, que estabelece uma relação entre a covariância entre duas variáveis e a covariância entre funções indicatrizes das mesmas. Este resultado é de uma importância extrema na medida em que abre caminho para uma série de resultados fundamentais, que procuramos evidenciar neste texto.

Lema 1.2.1 (Lema de Hoeffding) *Sejam X e Y variáveis aleatórias reais de quadrado integrável. Considere-se $F_{X,Y}$ a função de distribuição do par aleatório (X,Y) e F_X, F_Y as respectivas funções de distribuição marginais. Seja $H_{X,Y}$ a função que para todo $x, y \in \mathbb{R}$, é definida por*

$$H_{X,Y}(x,y) = F_{X,Y}(x,y) - F_X(x)F_Y(y) = \mathrm{Cov}\left(\mathbb{1}_{(-\infty,x]}(X), \mathbb{1}_{(-\infty,y]}(Y)\right).$$

Então,

$$\mathrm{Cov}\,(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \int_{\mathbb{R}^2} H_{X,Y}(x,y)\,dx\,dy,$$

Demonstração: Em primeiro lugar notemos que

$$\begin{aligned} H_{X,Y}(x,y) &= P(X \le x, Y \le y) - P(X \le x)P(Y \le y) = \\ &= \mathrm{Cov}\left(\mathbb{1}_{(-\infty,x]}(X), \mathbb{1}_{(-\infty,y]}(Y)\right) = \\ &= \mathrm{Cov}\left(1 - \mathbb{1}_{(-\infty,x]}(X), 1 - \mathbb{1}_{(-\infty,y]}(Y)\right) = \\ &= \mathrm{Cov}\left(\mathbb{1}_{(x,+\infty)}(X), \mathbb{1}_{(y,+\infty)}(Y)\right) = \\ &= P(X > x, Y > y) - P(X > x)P(Y > y). \end{aligned}$$

Consideremos (X_1, Y_1) e (X_2, Y_2) duas cópias independentes de (X,Y). Seja $A = \mathbb{E}\Big((X_1 - X_2)(Y_1 - Y_2)\Big)$. Ora

$$\begin{aligned} A &= \mathbb{E}(X_1Y_1 - X_1Y_2 - X_2Y_1 + X_2Y_2) = \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(XY) = \\ &= 2\Big(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)\Big). \end{aligned}$$

Por outro lado, interpretando $(X_1-X_2)(Y_1-Y_2)$ como a área do rectângulo definido pelos pontos $(X_1,Y_1),(X_1,Y_2),(X_2,Y_1)$ e (Y_1,Y_2), temos

$$A = \mathbb{E}\big(X_1Y_1 - X_1Y_2 - X_2Y_1 + X_2Y_2\big) =$$
$$= \mathbb{E}\left(\int_{\mathbb{R}^2}\Big(\mathbb{1}_{[x,+\infty)}(X_1) - \mathbb{1}_{[x,+\infty)}(X_2)\Big)\Big(\mathbb{1}_{[y,+\infty)}(Y_1) - \mathbb{1}_{[y,+\infty)}(Y_2)\Big)dxdy\right).$$

Ora, e uma vez que o valor absoluto deste integral é finito, usando o Teorema de Fubini,

$$\begin{aligned}
A = \int_{\mathbb{R}^2}\Big(&P\big(X_1 \geq x, Y_1 \geq y\big) - P\big(X_1 \geq x, Y_2 \geq y\big) - \\
&- P\big(X_2 \geq x, Y_1 \geq y\big) + P\big(X_2 \geq x, Y_2 \geq y\big)\Big)dxdy = \\
= \int_{\mathbb{R}^2}\Big(&P\big(X_1 \geq x, Y_1 \geq y\big) - P\big(X_1 \geq x\big)P\big(Y_2 \geq y\big) - \\
&- P\big(X_2 \geq x\big)P\big(Y_1 \geq y\big) + P\big(X_2 \geq x, Y_2 \geq y\big)\Big)dxdy = \\
= \int_{\mathbb{R}^2}\Big(&P\big(X \geq x, Y \geq y\big) + P\big(X \geq x, Y \geq y\big) - \\
&- P\big(X \geq x\big)P\big(Y \geq y\big) - P\big(X \geq x\big)P\big(Y \geq y\big)\Big)dxdy = \\
= \int_{\mathbb{R}^2} 2\Big(&F_{X,Y}(x,y) - F_X(x)F_Y(y)\Big)dxdy.
\end{aligned}$$

Temos finalmente,

$$\frac{1}{2}A = \Big(\mathbb{E}\big(XY\big) - \mathbb{E}\big(X\big)\mathbb{E}\big(Y\big)\Big) = \int_{\mathbb{R}^2}\left(F_{X,Y}(x,y) - F_X(x)F_Y(y)\right)dxdy,$$

ou seja

$$\mathrm{Cov}\,(X,Y) = \int_{\mathbb{R}^2} H_{X,Y}(x,y)dxdy. \quad \blacksquare$$

Lehmann [25], 1966, estabelece uma consequência importante deste resultado que evidencia o facto de a estrutura da covariância entre duas variáveis aleatórias determinar a independência relativa entre elas.

Propriedade 1.2.2 *Se o par de variáveis aleatórias X e Y é associado, então X e Y são independentes se e só se* $\mathrm{Cov}\,(X,Y) = 0$.

Demonstração: Seja (X, Y) um par de variáveis aleatórias associadas tais que $\mathrm{Cov}\,(X, Y) = 0$. De acordo com o *Lema de Hoeffding*,

$$\int_{\mathbb{R}^2} \mathrm{Cov}\left(\mathbb{1}_{(-\infty,x]}(X), \mathbb{1}_{(-\infty,y]}(Y)\right) dxdy = 0.$$

Uma vez que as funções $g(u, v) = \mathbb{1}_{(-\infty,x]}(u)$ e $h(u, v) = \mathbb{1}_{(-\infty,y]}(v)$ são não crescentes em cada variável, as variáveis aleatórias $g(X, Y)$ e $h(X, Y)$ são associadas. O resultado é, agora, imediato. ■

1.2.1 Independência e Associação. A desigualdade de Newman.

Newman [31], 1984, discute dois tipos de generalização da propriedade 1.2.2. Em primeiro lugar obtém a generalização directa, isto é, estabelece-a para mais do que duas variáveis e, seguidamente, obtém uma generalização indirecta, na qual a não correlação aproximada implica a independência aproximada, obtendo uma desigualdade que tem sido uma ferramenta usual na prova de propriedades que envolvam normalidade assimptótica de sucessões de variáveis associadas.

Para o primeiro tipo de generalização, há que introduzir o Teorema de Lebowitz [24], 1972, cuja importância não se esgota nesta aplicação, como veremos no desenvolvimento deste trabalho.

Teorema 1.2.3 (Lebowitz [24], 1972). *Sejam A e B subconjuntos de $\{1, \ldots, n\}$ e considerem-se os números reais $x_i, i \in A \cup B$. Defina-se $H_{A,B} = P\big(X_i > x_i,\, i \in A \cup B\big) - P\big(X_j > x_j,\, j \in A\big) P\big(X_k > x_k,\, k \in B\big)$. Seja $H_{i,j} = H_{\{i\},\{j\}}$. Se $\mathbf{X} = (X_1, \ldots, X_n)$ é associado então,*

$$0 \leq H_{A,B} \leq \sum_{i \in A, j \in B} H_{i,j}.$$

Demonstração: Comecemos por notar que

$$\begin{aligned} H_{i,j} &= P\big(X_k > x_k,\, k \in \{i,j\}\big) - P\big(X_i > x_i\big)P\big(X_j > x_j\big) = \\ &= P\big(X_i > x_i, X_j > x_j\big) - P\big(X_i > x_i\big)P\big(X_j > x_j\big) = \\ &= \mathrm{Cov}\left(\mathbb{1}_{(x_i,+\infty)}(X_i), \mathbb{1}_{(x_j,+\infty)}(X_j)\right). \end{aligned}$$

Sejam agora $i \in A \cup B$ e x_i fixos e defina-se $\rho_i(X_i) = \mathbb{1}_{(x_i,+\infty)}(X_i)$.
Para $C \subset \{1,\ldots,n\}$, seja $\rho_C(\mathbf{X}) = \prod_{i\in C} \rho_i(X_i)$ e $S_C(\mathbf{X}) = \sum_{i\in C} \rho_i(X_i)$.
Vamos, em primeiro lugar, mostrar que:

- $H_{A,B} = \mathrm{Cov}\left(\rho_A(\mathbf{X}), \rho_B(\mathbf{X})\right)$;
- $\sum_{i\in A, j\in B} H_{i,j} = \mathrm{Cov}\left(S_A(\mathbf{X}), S_B(\mathbf{X})\right)$.

Relativamente ao primeiro ponto,

$$\begin{aligned} \mathrm{Cov}\left(\rho_A(\mathbf{X}), \rho_B(\mathbf{X})\right) &= \mathrm{Cov}\left(\prod_{i\in A} \mathbb{1}_{(x_i,+\infty)}(X_i), \prod_{j\in B} \mathbb{1}_{(x_j,+\infty)}(X_j)\right) = \\ &= \mathbb{E}\left(\prod_{i\in A, j\in B} \mathbb{1}_{(x_i,+\infty)}(X_i)\ \mathbb{1}_{(x_j,+\infty)}(X_j)\right) - \\ &\quad - \mathbb{E}\left(\prod_{i\in A} \mathbb{1}_{(x_i,+\infty)}(X_i)\right)\ \mathbb{E}\left(\prod_{j\in B} \mathbb{1}_{(x_j,+\infty)}(X_j)\right) = \\ &= P\big(X_k > x_k, k \in A\cup B\big) - P\big(X_i > x_i\big)P\big(X_j > x_j\big) = \\ &= H_{A,B}. \end{aligned}$$

Quanto ao segundo ponto,

$$\begin{aligned} \mathrm{Cov}\left(S_A(\mathbf{X}), S_B(\mathbf{X})\right) &= \sum_{i\in A, j\in B} \mathrm{Cov}\left(\rho_i(X_i), \rho_j(X_j)\right) = \\ &= \sum_{i\in A, j\in B} \mathrm{Cov}\left(\mathbb{1}_{(x_i,+\infty)}(X_i), \mathbb{1}_{(x_j,+\infty)}(X_j)\right) = \\ &= \sum_{i\in A, j\in B} H_{i,j}. \end{aligned}$$

Pretendemos então mostrar que, para $\mathbf{X}$ vector associado, se tem

$$0 \leq \mathrm{Cov}\left(\rho_A(\mathbf{X}), \rho_B(\mathbf{X})\right) \leq \mathrm{Cov}\left(S_A(\mathbf{X}), S_B(\mathbf{X})\right).$$

De acordo com as respectivas definições, as funções das variáveis X_i, ρ_A, ρ_B, S_A, e S_B são crescentes. O mesmo é verdade relativamente às funções $S_A - \rho_A$ e $S_B - \rho_B$, como facilmente se pode verificar.
Podemos assim escrever,

$$\begin{aligned}\mathrm{Cov}\left(S_A(\mathbf{X}), S_B(\mathbf{X})\right) = \mathrm{Cov}\left((S_A - \rho_A)(\mathbf{X}), S_B(\mathbf{X})\right) + \\ + \mathrm{Cov}\left(\rho_A(\mathbf{X}), (S_B - \rho_B)(\mathbf{X})\right) + \mathrm{Cov}\left(\rho_A(\mathbf{X}), \rho_B(\mathbf{X})\right).\end{aligned}$$

Uma vez que, por associação, $\mathrm{Cov}\left((S_A - \rho_A)(\mathbf{X}), S_B(\mathbf{X})\right) \geq 0$ e $\mathrm{Cov}\left(\rho_A(\mathbf{X}), (S_B - \rho_B)(\mathbf{X})\right) \geq 0$ obtém-se, como se pretendia,

$$0 \leq \mathrm{Cov}\left(\rho_A(\mathbf{X}), \rho_B(\mathbf{X})\right) \leq \mathrm{Cov}\left(S_A(\mathbf{X}), S_B(\mathbf{X})\right). \blacksquare$$

Tem-se, então, como consequência imediata do *Lema de Hoeffding* e do *Teorema de Lebowitz*, a generalização pretendida que enunciamos seguidamente.

Corolário 1.2.4 (Newman [31], 1984). *Suponhamos que a sucessão* $\{X_n\}_{n\in\mathbb{N}}$ *de variáveis aleatórias reais é associada. Então* $\{X_k\}_{k\in A}$ *é independente de* $\{X_l\}_{l\in B}$ *se e só se para todo* $k \in A$, $l \in B$ *se tem* $\mathrm{Cov}\,(X_k, X_l) = 0$.
Do mesmo modo, a sucessão $\{X_n\}_{n\in\mathbb{N}}$ *é constituída por variáveis conjuntamente independentes se e só se para todo* $k \neq l$, $\mathrm{Cov}\,(X_k, X_l) = 0$.

Apresentamos de seguida a *Desigualdade de Newman*, atrás mencionada como a generalização indirecta do *Lema de Hoeffding*.

Teorema 1.2.5 (Newman [30], 1980). *Seja* $\mathbf{X} = (X_1, \ldots, X_n)$ *um vector associado e* ϕ *a função característica de* $\mathbf{X}$. *Designe-se por* ϕ_j, *para*

cada $j = 1, \ldots, n$, *a função característica de* X_j.
Para todo $(t_1, \ldots, t_n) \in \mathbb{R}^n$ *tem-se que,*

$$\left| \phi(t_1, \ldots, t_n) - \prod_{j=1}^{n} \phi_j(t_j) \right| \leq \frac{1}{2} \sum_{j,k=1,\, j \neq k}^{n} \left| t_j t_k \right| \operatorname{Cov}(X_j, X_k).$$

Demonstração: A demonstração deste Teorema é efectuada por indução sobre n.
Seja $n = 2$. Para $\mathbf{X} = (X_1, X_2)$ a desigualdade que se pretende demonstrar escreve-se,

$$|\phi(t_1, t_2) - \phi_1(t_1)\phi_2(t_2)| \leq 2 \times \frac{1}{2} \left| t_1 t_2 \right| \operatorname{Cov}(X_1, X_2),$$

ou seja,

$$\left| \mathbb{E}\left(\mathrm{e}^{it_1X_1 + it_2X_2}\right) - \mathbb{E}\left(\mathrm{e}^{it_1X_1}\right) \mathbb{E}\left(\mathrm{e}^{it_2X_2}\right) \right| \leq \left| t_1 t_2 \right| \operatorname{Cov}(X_1, X_2).$$

Ora,

$$\mathbb{E}\left(\mathrm{e}^{it_1X_1 + it_2X_2}\right) - \mathbb{E}\left(\mathrm{e}^{it_1X_1}\right) \mathbb{E}\left(\mathrm{e}^{it_2X_2}\right) = \operatorname{Cov}\left(\mathrm{e}^{it_1X_1}, \mathrm{e}^{it_2X_2}\right)$$

que, por integração por partes leva a

$$\begin{aligned} \operatorname{Cov}\left(\mathrm{e}^{it_1X_1}, \mathrm{e}^{it_2X_2}\right) = & \int_{\mathbb{R}^2} it_1 \mathrm{e}^{it_1x_1}\, it_2 \mathrm{e}^{it_2x_2} \times \\ & \times \left[P(X_1 > x_1, X_2 > x_2) - P(X_1 > x_1)\, P(X_2 > x_2) \right] dx_1 dx_2. \end{aligned} \tag{1.2}$$

Atendendo à associação e ao *Lema de Hoeffding*, temos para todo $t_1, t_2 \in \mathbb{R}$,

$$\left| \operatorname{Cov}\left(\mathrm{e}^{it_1X_1}, \mathrm{e}^{it_2X_2}\right) \right| \leq |t_1\, t_2| \operatorname{Cov}(X_1, X_2). \tag{1.3}$$

O Teorema está verificado para $n = 2$.

Suponha-se agora que a *Desigualdade de Newman* é válida para $n - 1$ variáveis. Vamos provar que também é válida para n variáveis.

Efectue-se uma partição do conjunto $\{1, \ldots, n\}$ em dois subconjuntos não vazios A e B definidos por:

(1) $A = \{1, \ldots, n-1\}$ e $B = \{n\}$ se os $t_j, 1 \leq j \leq n$ têm o mesmo sinal;

(2) $A = \{j \in \{1, \ldots, n\} : t_j < 0\}$ e $B = \{j \in \{1, \ldots, n\} : t_j \geq 0\}$.

Para $C \subset \{1, \ldots, n\}$, defina-se $\phi_C(t_1, \ldots, t_n) = \mathbb{E}(\mathrm{e}^{i \sum_{j \in C} t_j X_j})$.
Tem-se que,

$$\left| \phi - \prod_{j=1}^{n} \phi_j \right| = \left| \phi - \phi_A \phi_B + \phi_A \phi_B - \phi_A \prod_{j \in B} \phi_j + \phi_A \prod_{j \in B} \phi_j - \prod_{j=1}^{n} \phi_j \right|,$$

uma vez que $\prod_{j=1}^{n} \phi_j = \prod_{j \in A} \phi_j \prod_{j \in B} \phi_j$. Segue-se então,

$$\begin{aligned} \left| \phi - \prod_{j=1}^{n} \phi_j \right| &\leq |\phi - \phi_A \phi_B| + |\phi_A||\phi_B - \prod_{j \in B} \phi_j| + |\prod_{j \in B} \phi_j||\phi_A - \prod_{j \in A} \phi_j| \leq \\ &\leq |\phi - \phi_A \phi_B| + |\phi_B - \prod_{j \in B} \phi_j| + |\phi_A - \prod_{j \in A} \phi_j|. \end{aligned}$$

De acordo com a hipótese de indução,

$$|\phi_B - \prod_{j \in B} \phi_j| \leq \frac{1}{2} \sum_{j,k \in B,\, j \neq k} |t_j t_k| \mathrm{Cov}\,(X_j, X_k)$$

e

$$|\phi_A - \prod_{j \in A} \phi_j| \leq \frac{1}{2} \sum_{j,k \in A,\, j \neq k} |t_j t_k| \mathrm{Cov}\,(X_j, X_k).$$

No que diz respeito a $|\phi - \phi_A \phi_B|$ considerem-se as variáveis $U = \sum_{j \in A} |t_j| X_j$ e $V = \sum_{j \in B} |t_j| X_j$. O vector (U, V) é obviamente associado. Por outro lado, se os $t_i, i = 1, \ldots, n$ tiverem todos o mesmo sinal $\phi(t_1, \ldots, t_n) = \mathbb{E}(\mathrm{e}^{i(U+V)})$, $\phi_A(t_1, \ldots, t_n) = \mathbb{E}(\mathrm{e}^{iU})$ e $\phi_B(t_1, \ldots, t_n) = \mathbb{E}(\mathrm{e}^{iV})$ ou $\phi(t_1, \ldots, t_n) = \mathbb{E}(\mathrm{e}^{-i(U+V)})$, $\phi_A(t_1, \ldots, t_n) = \mathbb{E}(\mathrm{e}^{-iU})$ e $\phi_B(t_1, \ldots, t_n) = \mathbb{E}(\mathrm{e}^{-iV})$. Enquanto que se os $t_i, i = 1, \ldots, n$ não

tiverem todos o mesmo sinal,
$\phi(t_1,\ldots,t_n)=\mathbb{E}(\mathrm{e}^{i(-U+V)})$, $\phi_A(t_1,\ldots,t_n)=\mathbb{E}(\mathrm{e}^{-iU})$ e
$\phi_B(t_1,\ldots,t_n)=\mathbb{E}(\mathrm{e}^{iV})$ ou $\phi(t_1,\ldots,t_n)=\mathbb{E}(\mathrm{e}^{i(U-V)})$, $\phi_A(t_1,\ldots,t_n)=\mathbb{E}(\mathrm{e}^{iU})$ e $\phi_B(t_1,\ldots,t_n)=\mathbb{E}(\mathrm{e}^{-iV})$. De qualquer modo, escrevendo,

$$\left|\phi-\phi_A\phi_B\right|=\left|\mathbb{E}(\mathrm{e}^{i(\pm U\pm V)})-\mathbb{E}(\mathrm{e}^{i(\pm U)})\mathbb{E}(\mathrm{e}^{i(\pm V)})\right|$$

e usando a associação do vector (U,V) juntamente com a desigualdade (1.3), fazendo $t_1=\pm1$ e $t_2=\pm1$, tem-se que,

$$\begin{aligned}\left|\phi-\phi_A\phi_B\right| &\leq \mathrm{Cov}\,(U,V)=\mathrm{Cov}\,\big(\sum_{j\in A}|t_j|X_j,\sum_{j\in B}|t_j|X_j\big)=\\ &= \sum_{j\in A,k\in B}|t_j||t_k|\mathrm{Cov}\,(X_j,X_k).\end{aligned}$$

O resultado obtém-se agora facilmente.■

Nota 1.2.6 *A desigualdade (1.3) traduz já uma extensão do Lema de Hoeffding. É apresentada em Newman [31], 1984, na sua Proposição* 6.

1.2.2 A generalização de Yu.

Yu [53], 1993, obtém um resultado mais geral que inclui o *Lema de Hoeffding.*

Teorema 1.2.7 (Yu [53], 1993) *Para todo o inteiro positivo k, seja $\{f_i\}_{i=1,\ldots,2k}$, uma família de funções absolutamente contínuas definidas num intervalo de $\mathbb{R}$. Então, para todo o conjunto de variáveis aleatórias $Y_1,\ldots,Y_{2k}$, tem-se que*

$$K\big(f_1(Y_1),f_2(Y_2),\ldots,f_{2k}(Y_{2k})\big)=$$
$$\int_{\mathbb{R}^{2k}}\left(\prod_{i=1}^{2k}f_i'(x_i)\right)K\Big(\mathbb{1}_{(-\infty,x_1]}(Y_1),\ldots,\mathbb{1}_{(-\infty,x_{2k}]}(Y_{2k})\Big)dx_1\ldots dx_{2k},$$

se o segundo membro desta equação for finito, sendo

$$K(\xi_1,\ldots,\xi_{2k}) = \sum_{card\{\pi\}<k} (-1)^{card\{\pi\}} \mathbb{E}\Big(\prod_{j\in\pi}\xi_j\Big)\mathbb{E}\Big(\prod_{j\notin\pi}\xi_j\Big) +$$
$$+\frac{1}{2}\sum_{card\{\pi\}=k} (-1)^{card\{\pi\}} \mathbb{E}\Big(\prod_{j\in\pi}\xi_j\Big)\mathbb{E}\Big(\prod_{j\notin\pi}\xi_j\Big),$$

onde $\pi \subseteq \{1,\ldots,2k\}$.

Nota 1.2.8 *A definição de K atrás apresentada não é a usada no artigo de Yu. De facto, Yu não introduz o coeficiente* $\frac{1}{2}$ *na última parcela. Este erro foi encontrado e corrigido por Jacquemin [20], 1997.*

Nota 1.2.9 *Fazendo* $k = 1$, *o teorema estabelece que :*

$$K\big(f_1(Y_1), f_2(Y_2)\big) = \int_{\mathbb{R}^2} f_1'(x_1) f_2'(x_2) K\big(\mathbb{1}_{(-\infty,x_1]}(Y_1), \mathbb{1}_{(-\infty,x_2]}(Y_2)\big) dx_1 dx_2.$$

Tendo em atenção a definição de K,

$$K\big(f_1(Y_1), f_2(Y_2)\big) = \mathrm{Cov}\left(f_1(Y1), f_2(Y_2)\right) \quad e$$
$$K\big(\mathbb{1}_{(-\infty,x_1]}(Y_1), \mathbb{1}_{(-\infty,x_2]}(Y_2)\big) = \mathrm{Cov}\left(\mathbb{1}_{(-\infty,x_1]}(Y_1), \mathbb{1}_{(-\infty,x_2]}(Y_2)\right).$$

Consequentemente, se f_1 *e* f_2 *satisfazem as condições do teorema anterior,*

$$\mathrm{Cov}\left(f_1(Y1), f_2(Y_2)\right) = \\ = \int_{\mathbb{R}^2} f_1'(x_1) f_2'(x_2) \mathrm{Cov}\left(\mathbb{1}_{(-\infty,x_1]}(Y_1), \mathbb{1}_{(-\infty,x_2]}(Y_2)\right) dx_1\, dx_2. \tag{1.7}$$

Este resultado é, como referimos na Nota 1.2.6, de Newman, [31], 1984, e, mais uma vez, está incluído no resultado mais geral de Yu.

Passamos agora a apresentar a demonstração do teorema 1.2.7, cujo argumento consiste numa generalização da demonstração do Lema de Hoeffding.

Demonstração: (do teorema 1.2.7) Seja $(X_1, \ldots, X_{2k})$ uma cópia independente de $(Y_1, \ldots, Y_{2k})$. O primeiro passo da prova consiste em mostrar que

$$K\big(f_1(Y_1), \ldots, f_{2k}(Y_{2k})\big) = \frac{1}{2}\mathbb{E}\left(\prod_{i=1}^{2k}\big(f_i(Y_i) - f_i(X_i)\big)\right). \tag{1.8}$$

Ora,

$$\begin{aligned}\prod_{i=1}^{2k}\big(f_i(Y_i) - f_i(X_i)\big) &= \sum_{\pi}\left(\prod_{j\in\pi} f_j(Y_j)\right)\left(\prod_{j\notin\pi}\big(-f_j(X_j)\big)\right) = \\ &= \sum_{\pi}(-1)^{card\{\bar{\pi}\}}\left(\prod_{j\in\pi} f_j(Y_j)\right)\left(\prod_{j\notin\pi} f_j(X_j)\right),\end{aligned} \tag{1.9}$$

onde $\bar{\pi}$ denota o complementar do conjunto π. Como $(-1)^{card\{\bar{\pi}\}} = (-1)^{card\{\pi\}}$, já que $card\{\bar{\pi}\}$ e $card\{\pi\}$ têm a mesma paridade por se tratar de uma partição do conjunto $\{1, \ldots, 2k\}$, temos

$$\mathbb{E}\left(\prod_{i=1}^{2k}\big(f_i(Y_i) - f_i(X_i)\big)\right) = \sum_{\pi}(-1)^{card\{\pi\}}\mathbb{E}\left(\Big(\prod_{j\in\pi} f_j(Y_j)\Big)\Big(\prod_{j\notin\pi} f_j(X_j)\Big)\right).$$

Uma vez que $(X_1, \ldots, X_{2k})$ e $(Y_1, \ldots, Y_{2k})$ são independentes,

$$\mathbb{E}\left(\prod_{i=1}^{2k}\big(f_i(Y_i) - f_i(X_i)\big)\right) = \sum_{\pi}(-1)^{card\{\pi\}}\mathbb{E}\left(\prod_{j\in\pi} f_j(Y_j)\right)\mathbb{E}\left(\prod_{j\notin\pi} f_j(X_j)\right).$$

Como $(X_1,\ldots,X_{2k})$ e $(Y_1,\ldots,Y_{2k})$ são identicamente distribuídos,

$$\mathbb{E}\left(\prod_{i=1}^{2k}\big(f_i(Y_i)-f_i(X_i)\big)\right)=\sum_{\pi}(-1)^{card\{\pi\}}\mathbb{E}\left(\prod_{j\in\pi}f_j(Y_j)\right)\mathbb{E}\left(\prod_{j\notin\pi}f_j(Y_j)\right)=$$

$$=\sum_{\pi:card\{\pi\}=k}(-1)^{card\{\pi\}}\mathbb{E}\left(\prod_{j\in\pi}f_j(Y_j)\right)\mathbb{E}\left(\prod_{j\notin\pi}f_j(Y_j)\right)+$$

$$+2\sum_{\pi:card\{\pi\}<k}(-1)^{card\{\pi\}}\mathbb{E}\left(\prod_{j\in\pi}f_j(Y_j)\right)\mathbb{E}\left(\prod_{j\notin\pi}f_j(Y_j)\right)=$$

$$=2K\big(f_1(Y_1),\ldots,f_{2k}(Y_{2k})\big).$$

O segundo passo da prova consiste em mostrar que, para cada $i=1,\ldots,2k$,

$$\big(f_i(Y_i)-f_i(X_i)\big)=\int_{\mathbb{R}}f_i'(x_i)\big(\mathbb{1}_{(-\infty,x_i]}(X_i)-\mathbb{1}_{(-\infty,x_i]}(Y_i)\big)dx_i,$$

o que é consequência imediata do facto de

$$\big(f_i(Y_i)-f_i(X_i)\big)=\int_{X_i}^{Y_i}f_i'(x_i)dx_i=$$

$$=\int_{\mathbb{R}}f_i'(x_i)\big(\mathbb{1}_{(-\infty,x_i]}(X_i)-\mathbb{1}_{(-\infty,x_i]}(Y_i)\big)dx_i.$$

Tem-se então,

$$K\big(f_1(Y_1),\ldots,f_{2k}(Y_{2k})\big)=$$

$$=\frac{1}{2}\mathbb{E}\left(\prod_{i=1}^{2k}\Big(\int_{\mathbb{R}}f_i'(x_i)\big(\mathbb{1}_{(-\infty,x_i]}(X_i)-\mathbb{1}_{(-\infty,x_{i(Y_i)}]}\big)dx_i\Big)\right).$$

Como a função integranda é absolutamente integrável segue-se que, aplicando o Teorema de Fubini,

$$K\big(f_1(Y_1),\ldots,f_{2k}(Y_{2k})\big)=\int_{\mathbb{R}^{2k}}\left(\prod_{i=1}^{2k}f_i'(x_i)\right)\times$$

$$\times\frac{1}{2}\mathbb{E}\Big(\prod_{i=1}^{2k}\big(\mathbb{1}_{(-\infty,x_i]}(X_i)-\mathbb{1}_{(-\infty,x_i]}(Y_i)\big)\Big)dx_1\ldots dx_{2k}.$$

Finalmente, note-se que

$$\begin{aligned}\frac{1}{2}\times \mathbb{E}\Big(\prod_{i=1}^{2k}\big(\mathbb{1}_{(-\infty,x_i]}(X_i)-\mathbb{1}_{(-\infty,x_i]}(Y_i)\big)\Big)&=\\ &= K\big(\mathbb{1}_{(-\infty,x_1]}(X_1),\ldots,\mathbb{1}_{(-\infty,x_{2k}]}(X_{2k})\big)=\\ &= K\big(\mathbb{1}_{(-\infty,x_1]}(Y_1),\ldots,\mathbb{1}_{(-\infty,x_{2k}]}(Y_{2k})\big),\end{aligned}$$

seguindo-se agora imediatamente o resultado. ■

1.3 Outras definições de associação

É conveniente dispor de caracterizações alternativas de associação que envolvam funções pertencentes a classes de funções particulares. Entre elas estão a classe das funções crescentes binárias e a classe das funções crescentes limitadas e contínuas. Estas definições alternativas de associação vão permitir a justificação de algumas propriedades e exemplos úteis no desenvolvimento deste trabalho. A referência aos resultados aqui apresentados é, uma vez mais, Esary, Proschan e Walkup [13], 1967.

Teorema 1.3.1 *O vector* $\mathbf{X}=(X_1,\ldots,X_n)$ *é associado se e só se, para todo o par de funções crescentes* f *e* g *de* $\mathbb{R}^n$ *para* $\{0,1\}$, *se tem*

$$\mathrm{Cov}\,(f(\mathbf{X}),g(\mathbf{X}))\geq 0.$$

Demonstração: Basta mostrar que se $\mathrm{Cov}\,(f(\mathbf{X}),g(\mathbf{X}))\geq 0$, então $\mathbf{X}$ é associado. Sejam $\tilde{f}$ e $\tilde{g}$ duas funções crescentes, quaisquer, de $\mathbb{R}^n$ para $\mathbb{R}$. De acordo *Lema de Hoeffding*,

$$\begin{aligned}\mathrm{Cov}\,\big(\tilde{f}(\mathbf{X}),\tilde{g}(\mathbf{X})\big)&=\int_{\mathbb{R}^2}\mathrm{Cov}\,\Big(\mathbb{1}_{(-\infty,s]}\big(\tilde{f}(\mathbf{X})\big),\mathbb{1}_{(-\infty,t]}\big(\tilde{g}(\mathbf{X})\big)\Big)\,ds\,dt=\\ &=\int_{\mathbb{R}^2}\mathrm{Cov}\,\Big(\mathbb{1}_{[s,+\infty)}\big(\tilde{f}(\mathbf{X})\big),\mathbb{1}_{[t,+\infty)}\big(\tilde{g}(\mathbf{X})\big)\Big)\,ds\,dt.\end{aligned}$$

Sejam $s,t\in\mathbb{R}$ fixos. Considerem-se as aplicações f_s de $\mathbb{R}^n$ para $\{0,1\}$ que a cada $\mathbf{x}\in\mathbb{R}^n$ faz corresponder $\mathbb{1}_{[s,+\infty[}\big(\tilde{f}(\mathbf{x})\big)$ e g_t também de $\mathbb{R}^n$

para $\{0,1\}$ e que a cada $\mathbf{x} \in \mathbb{R}^n$ faz corresponder $\mathbb{1}_{[t,+\infty]}\big(\tilde{g}(\mathbf{x})\big)$. Estas funções são crescentes uma vez que quer $\tilde{f}$ e $\tilde{g}$, quer a função indicatriz, o são. Assim,

$$\mathrm{Cov}\left(\mathbb{1}_{[s,+\infty)}\big(\tilde{f}(\mathbf{X})\big), \mathbb{1}_{[t,+\infty)}\big(\tilde{g}(\mathbf{X})\big)\right) = \mathrm{Cov}\left(f_s(\mathbf{X}), g_t(\mathbf{X})\right) \geq 0$$

e, então

$$\mathrm{Cov}\left(f(\mathbf{X}), g(\mathbf{X})\right) = \int_{\mathbb{R}^2} \mathrm{Cov}\left(\mathbb{1}_{[s,+\infty)}\big(f(\mathbf{X})\big), \mathbb{1}_{[t,+\infty)}\big(g(\mathbf{X})\big)\right) ds\, dt \geq 0.$$

■

Vejamos agora uma outra condição necessária e suficiente de associação.

Teorema 1.3.2 *Sejam u e v duas funções quaisquer de $\mathbb{R}^n$ para $\mathbb{R}$, crescentes, limitadas e contínuas. Então o vector $\mathbf{X} = (X_1, \ldots, X_n)$ é associado se e só se* $\mathrm{Cov}\,(u(\mathbf{X}), v(\mathbf{X})) \geq 0$.

A prova deste teorema passa por um resultado intermédio que passamos a enunciar e cuja demonstração se encontra em Suquet [48], 1994.

Lema 1.3.3 *Se se verificar que para todo o par de funções u e v nas condições do teorema 1.3.2,* $\mathrm{Cov}\,(u(\mathbf{X}), v(\mathbf{X})) \geq 0$, *também se verifica que* $\mathrm{Cov}\,(\phi(\mathbf{X}), \psi(\mathbf{X})) \geq 0$, *para todo o par de funções ϕ e ψ de $\mathbb{R}^n$ para $\{0,1\}$ crescentes e contínuas à direita.*

Vejamos agora a demonstração do teorema 1.3.2.

Demonstração: Pretende-se mostrar que se $\mathbf{X} = (X_1, \ldots, X_n)$ é um vector tal que para qualquer par de funções u e v de $\mathbb{R}^n$ para $\mathbb{R}$, crescentes, uniformemente limitadas e contínuas, se verifica $\mathrm{Cov}\,(u(\mathbf{X}), v(\mathbf{X})) \geq 0$, então $\mathbf{X}$ é associado.
Consideremos duas funções quaisquer crescentes ϕ e ψ de $\mathbb{R}^n$ para $\{0,1\}$. De acordo com o teorema 1.3.1, é suficiente mostrar que
$\mathrm{Cov}\,(\phi(\mathbf{X}), \psi(\mathbf{X})) \geq 0$.

Denote-se por $\mathbb{P}$ a lei de probabilidade do vector $\mathbf{X}$. Consideremos $A = \phi^{-1}(\{1\})$ e $B = \psi^{-1}(\{1\})$. Sabemos então que, para qualquer conjunto A e para todo $\epsilon > 0$, existe um compacto C contido em A tal que $\mathbb{P}(A) \leq \mathbb{P}(C) + \epsilon$.
Denote-se por $\preceq$ a relação de ordem parcial em $\mathbb{R}^n$ definida por

$$(x_1, \ldots, x_n) \preceq (y_1, \ldots, y_n) \quad \text{se e só se} \quad x_i \leq y_i, \forall i = 1, \ldots, n.$$

Defina-se ainda, para $\mathbf{t} \in \mathbb{R}^n$ fixo e tal que $\mathbf{t} \succeq 0$, o conjunto $\widetilde{C} = \{\mathbf{c}+\mathbf{t} : \mathbf{c} \in C, \mathbf{t} \succeq 0\}$. Como ϕ é crescente, $\widetilde{C} \subset A$. Além disso, $\widetilde{C}$ é um conjunto fechado.
Seja φ a função indicatriz de $\widetilde{C}$. Então $\varphi \leq \phi$ e φ é crescente e contínua à direita em todos os pontos de $\widetilde{C}$, uma vez que para todo $\mathbf{x} \in \widetilde{C}$, $\widetilde{C}$ contém todos os pontos $\mathbf{y} \succeq \mathbf{x}$, numa vizinhança de $\mathbf{x}$. Como $\widetilde{C}$ é fechado, φ é também contínua à direita em $\mathbb{R}^n \setminus \widetilde{C}$.
Construa-se, do mesmo modo, uma função binária ϑ a partir de ψ. Seja K um compacto contido em B e $\mathbf{t} \in \mathbb{R}^n$ fixo e tal que $\mathbf{t} \succeq 0$. Considere-se, ainda, o conjunto $\widetilde{K} = \{\mathbf{k}+\mathbf{t} : \mathbf{k} \in K, \mathbf{t} \succeq 0\}$. Fazendo $\vartheta = \mathbb{1}_{\widetilde{K}}$, $\vartheta \leq \psi$ e, além disso, ϑ é contínua à direita em $\mathbb{R}^n$. Ora, de acordo com o lema 1.3.3, $\mathrm{Cov}\,(\varphi(\mathbf{X}), \vartheta(\mathbf{X})) \geq 0$. Vamos agora provar que $\mathrm{Cov}\,(\phi(\mathbf{X}), \psi(\mathbf{X})) \geq 0$.
Em primeiro lugar note-se que de $\varphi \leq \phi$ e $\vartheta \leq \psi$, vem $0 \leq \varphi\,\vartheta \leq \phi\,\psi \leq 1$. Este facto implica que $\mathbb{E}\,(\varphi(\mathbf{X})\,\vartheta(\mathbf{X})) \leq \mathbb{E}\,(\phi(\mathbf{X})\,\psi(\mathbf{X}))$. Temos que $\mathbb{E}\,(\phi) = \mathbb{E}\,(\mathbb{1}_A(\mathbf{X})) = P(A)$ e $\mathbb{E}\,(\varphi) = \mathbb{E}\left(\mathbb{1}_{\widetilde{C}}(\mathbf{X})\right) = P(\widetilde{C})$. Do mesmo modo, $\mathbb{E}\,(\psi) = \mathbb{E}\,(\mathbb{1}_B(\mathbf{X})) = P(B)$ e $\mathbb{E}\,(\vartheta) = \mathbb{E}\left(\mathbb{1}_{\widetilde{K}}(\mathbf{X})\right) = P(\widetilde{K})$. Ora, como vimos, $P(A) \leq P(\widetilde{C}) + \epsilon$ e $P(B) \leq P(\widetilde{K}) + \epsilon$.
Obtemos assim, $\mathbb{E}\,(\phi(\mathbf{X})) \leq \mathbb{E}\,(\varphi) + \epsilon$ e $\mathbb{E}\,(\psi(\mathbf{X})) \leq \mathbb{E}\,(\vartheta) + \epsilon$.
Destas desigualdades vem $-\mathbb{E}\,(\phi(\mathbf{X}))\,\mathbb{E}\,(\psi(\mathbf{X})) \geq -\left(\mathbb{E}\,(\varphi(\mathbf{X}))\right)\left(\mathbb{E}\,(\vartheta(\mathbf{X}))\right)$.
Podemos, então, escrever

$$\begin{aligned}\mathrm{Cov}\,(\phi(\mathbf{X}), \psi(\mathbf{X})) &\geq \mathbb{E}\,(\varphi(\mathbf{X}))\,\mathbb{E}\,(\vartheta(\mathbf{X})) - \left(\mathbb{E}\,(\varphi(\mathbf{X}))\right)\left(\mathbb{E}\,(\vartheta(\mathbf{X}))\right) \geq \\ &\geq \mathrm{Cov}\,(\varphi(\mathbf{X}), \vartheta(\mathbf{X})) - \epsilon^2 - 2\epsilon\end{aligned}$$

Como ϵ é arbitrário, segue-se que Cov $(\phi(\mathbf{X}), \psi(\mathbf{X})) \geq 0$. Consequentemente $\mathbf{X}$ é associado. ■

Vejamos agora outras propriedades importantes de associação.

Propriedade 1.3.4 *Todo o conjunto formado por uma única variável aleatória real é associado.*

Demonstração: De acordo com o teorema 1.3.1, é suficiente mostrar que para todo o par de funções binárias crescentes, ϕ e ψ, de $\mathbb{R}$ para $\{0,1\}$ e para toda a variável aleatória real X, se tem Cov $(\phi(X), \psi(X)) \geq 0$. As funções que verificam a condição precedente são funções indicatrizes de intervalos do tipo $[a, +\infty[$. Tem-se que $\phi < \psi$ ou $\psi < \phi$. Seja $\psi < \phi$. Neste caso, $\phi\psi = \psi^2 = \psi$ e então,

$$\begin{aligned}
&\mathbb{E}\left(\phi(X)\psi(X)\right) - \mathbb{E}\left(\phi(X)\right)\mathbb{E}\left(\psi(X)\right) = \\
&\quad = \mathbb{E}\left(\psi(X)\right) - \mathbb{E}\left(\phi(X)\right)\mathbb{E}\left(\psi(X)\right) = \\
&\quad = \mathbb{E}\left(\psi(X)\right)\left(1 - \mathbb{E}\left(\phi(X)\right)\right) \geq 0. \blacksquare
\end{aligned}$$

Nota 1.3.5 *Esta propriedade, juntamente com a propriedade (1.1.7), leva a que todo o conjunto de variáveis aleatórias independentes é associado.*

Propriedade 1.3.6 *Se o vector* $X^{(k)} = \left(X_1^{(k)}, \ldots, X_n^{(k)}\right)$ *é associado para todo o* k *e se* $X^{(k)}$ *converge em lei para* $\mathbf{X} = (X_1, \ldots, X_n)$, *então* $\mathbf{X}$ *é um vector associado.*

Demonstração: De acordo com o teorema 1.3.2 é suficiente mostrar que para quaisquer u, v funções crescentes, contínuas e limitadas de $\mathbb{R}^n$ para $\mathbb{R}$, Cov $\big(u(\mathbf{X}), v(\mathbf{X})\big) \geq 0$. Como $X^{(k)} = \left(X_1^{(k)}, \ldots, X_n^{(k)}\right)$ é associado, sendo u, v funções crescentes, contínuas e limitadas de $\mathbb{R}^n$ para $\mathbb{R}$, Cov $\big(u(X^{(k)}), v(X^{(k)})\big) \geq 0$, para todo k. Uma vez que u, v, uv são contínuas e limitadas,

$$\begin{aligned}
&\lim_{k\to\infty} \mathbb{E}\big[u(X^{(k)})v(X^{(k)})\big] - \mathbb{E}\big[u(X^{(k)})\big]\mathbb{E}\big[v(X^{(k)})\big] = \\
&\quad = \mathbb{E}\big[u(\mathbf{X})v(\mathbf{X})\big] - \mathbb{E}\big[u(\mathbf{X})\big]\mathbb{E}\big[v(\mathbf{X})\big],
\end{aligned}$$

pelo que Cov $\big(u(\mathbf{X}), v(\mathbf{X})\big) \geq 0$.■

1.4 Alguns exemplos

Nesta secção pretendemos apresentar alguns exemplos de vectores associados.

Exemplo 1.4.1 *Todo o vector aleatório de componentes independentes é associado.*

Este exemplo resulta imediatamente da nota 1.3.5.

Exemplo 1.4.2 *Para qualquer variável aleatória X e para qualquer número natural n, o vector constituído por n variáveis aleatórias todas idênticas a X, $\mathbf{X} = (X, \ldots, X)$ é associado.*

Demonstração: Sejam A e B subconjuntos de $\{1, \ldots, n\}$ com cardinalidade a e b, respectivamente. Considerem-se G e H funções não decrescentes em que $G : \mathbb{R}^a \to \mathbb{R}$ e $H : \mathbb{R}^b \to \mathbb{R}$. Sejam ainda g e h funções de $\mathbb{R}$ para $\mathbb{R}$, tais que $g(X) = G(\mathbf{X})$ e $h(X) = H(\mathbf{X})$. Ora, como é evidente, g e h são funções não decrescentes de X pelo que se tem Cov $(g(X), h(X)) \geq 0$, sendo, consequentemente, o vector $\mathbf{X} = (X, \ldots, X)$ associado. ■

Exemplo 1.4.3 *Seja $\mathbf{X} = (X_1, \ldots, X_n)$ um vector associado. Consideremos, para $i = 1, \ldots, n$, $S_i = \sum_{j=1}^{i} X_j$. O vector das somas parciais $S = (S_1, \ldots, S_n)$ é associado.*

Demonstração: Considerem-se as aplicações f_i de $\mathbb{R}^n$ para $\mathbb{R}$, tais que para cada $\mathbf{x} = (x_1, \ldots, x_n)$, $\quad f_i(\mathbf{x}) = \sum_{j=1}^{i} x_j$. Uma vez que estas aplicações assim definidas são crescentes, aplicando a propriedade 1.1.8, é

imediato que o vector $S = (S_1, \ldots, S_n)$ seja associado. ■
Vejamos agora um resultado que fornece uma condição necessária e suficiente para que um vector multidimensional Gaussiano seja associado. A condição suficiente, apesar de conjecturada durante muitos anos, foi provada apenas em 1982 por Pitt [38].

Teorema 1.4.4 (Pitt [38], 1982) *Seja* $\mathbf{X} = (X_1, \ldots, X_n)$ *um vector Gaussiano, com matriz de covariância* Σ *de termo genérico* σ_{ij}. *É condição necessária e suficiente para que as variáveis aleatórias* $X_i, i = 1, \ldots, n$ *sejam associadas que* $\sigma_{ij} \geq 0$ *para* $1 \leq i, j \leq n$.

Demonstração: Pitt considera, em primeiro lugar, que Σ é não singular e mostra que Cov $(f(\mathbf{X}), g(\mathbf{X})) \geq 0$, para funções f e g continuamente diferenciáveis e com derivadas parciais de primeira ordem limitadas. Em seguida exibe argumentos técnicos necessários para a prova no caso geral. Relativamente à hipótese de as funções em causa serem continuamente diferenciáveis, isto é de classe C^∞, com derivadas parciais de primeira ordem limitadas, refere que esta pode ser removida se se considerar f e g funções crescentes continuas e limitadas. De facto, se f for uma função crescente, continua e limitada e $\{\psi_t\}_{t \in T}$ for uma sucessão de funções de classe C^∞, que converge uniformemente para a função identidade, então, para cada t, $\psi_t * f$ é de classe C^∞, crescente e com derivadas limitadas. Além disso $\psi_t * f \to f$ em $L^p(\mathbb{R}^n)$.
Relativamente à hipótese de a matriz de covariância Σ do vector $\mathbf{X}$ ser não singular, esta não é essencial. De facto, para qualquer covariância Σ e qualquer $\epsilon > 0$, a matriz $\Sigma + \epsilon I$, onde I é a identidade, é não singular. É então suficiente considerar este caso particular.

Considere-se $\mathbf{Z} = (Z_1, \ldots, Z_n)$ uma cópia independente de $\mathbf{X}$. $(\mathbf{X}, \mathbf{Z})$ é um vector Gaussiano.
Defina-se para $\lambda \in [0,1]$ o vector aleatório $\mathbf{Y}(\lambda) = \lambda \mathbf{X} + \sqrt{1-\lambda^2}\,\mathbf{Z}$. Para qualquer $\lambda \in [0,1]$, $\mathbf{Y}(\lambda)$ é, também, um vector Gaussiano.
De maneira a aligeirar a notação, escreve-se, no que se segue, $Y_i(\lambda) =$

$Y_i, i = 1, \ldots, n$. O facto de X e Z serem independentes leva a

$$\begin{aligned}\mathbb{E}\big(Y_iY_j\big) &= \lambda^2\,\mathbb{E}\big(X_iX_j\big)+\lambda\sqrt{1-\lambda^2}\;\mathbb{E}(X_i)\mathbb{E}(Z_j)+\\ &+ \lambda\sqrt{1-\lambda^2}\;\mathbb{E}(X_j)\mathbb{E}(Z_i)+(1-\lambda^2)\,\mathbb{E}(Z_iZ_j).\end{aligned}$$

Por outro lado,

$$\begin{aligned}\mathbb{E}(Y_i)\mathbb{E}(Y_j) = \lambda^2\,\mathbb{E}(X_i)\mathbb{E}(X_j)+\lambda\sqrt{1-\lambda^2}\,\mathbb{E}(X_i)\mathbb{E}(Z_j)+\\ +\lambda\sqrt{1-\lambda^2}\;\mathbb{E}(X_j)\mathbb{E}(Z_i)+(1-\lambda^2)\mathbb{E}(Z_i)\mathbb{E}(Z_j).\end{aligned}$$

Temos então,

$$\begin{aligned}\mathrm{Cov}\,\big(Y_i,Y_j\big)&=\lambda^2\mathbb{E}\big(X_iX_j\big)-\lambda^2\mathbb{E}(X_i)\mathbb{E}(X_j)+\\ &\quad+(1-\lambda^2)\mathbb{E}(Z_iZ_j)-(1-\lambda^2)\mathbb{E}(Z_i)\mathbb{E}(Z_j)=\\ &=\lambda^2\mathrm{Cov}\,\big(X_i,X_j\big)+(1-\lambda^2)\mathrm{Cov}\,\big(Z_i,Z_j\big)=\\ &=\lambda^2\mathrm{Cov}\,\big(X_i,X_j\big)+(1-\lambda^2)\mathrm{Cov}\,\big(X_i,X_j\big)=\\ &=\mathrm{Cov}\,\big(X_i,X_j\big).\end{aligned}$$

Conclui-se que, para qualquer $\lambda \in [0,1]$, $\mathbf{Y}(\lambda)$ tem a mesma matriz de covariância que $\mathbf{X}$. Considere-se agora,

$$\begin{aligned}&\mathrm{Cov}\;(X_i,Y_j) = \mathbb{E}\big(X_iY_j\big)-\mathbb{E}(X_i)\mathbb{E}(Y_j)=\\ &=\mathbb{E}\Big(X_i\big(\lambda X_j+\sqrt{1-\lambda^2}\,Z_j\big)\Big)-\mathbb{E}(X_i)\Big(\lambda\mathbb{E}(X_j)+\sqrt{1-\lambda^2}\,\mathbb{E}(Z_j)\Big)=\\ &\text{porque } \mathbf{X} \text{ e } \mathbf{Z} \text{ são independentes,}\\ &=\lambda\mathbb{E}\big(X_iY_j\big)+\sqrt{1-\lambda^2}\,\mathbb{E}(X_i)\mathbb{E}(Z_j)-\lambda\mathbb{E}(X_i)\mathbb{E}(X_j)-\sqrt{1-\lambda^2}\,\mathbb{E}(X_i)\mathbb{E}(Z_j)=\\ &=\lambda\,\mathrm{Cov}\,\big(X_i,X_j\big)+\sqrt{1-\lambda^2}\,\mathrm{Cov}\,\big(X_i,Z_j\big)=\lambda\,\mathrm{Cov}\,\big(X_i,X_j\big).\end{aligned}$$

Conclui-se, então que

$$\mathrm{Cov}\,\big(X_i,Y_j\big)=\mathrm{Cov}\,\big(X_i,Y_j(\lambda)\big)=\lambda\,\mathrm{Cov}\,\big(X_i,X_j\big)=\lambda\;\sigma_{ij}.$$

Seja agora $F(\lambda)=\mathbb{E}\big(f(\mathbf{X})g(\mathbf{Y}(\lambda))\big)$. Observe-se que F é continua.
Uma vez que $\mathbf{Y}(0)=\mathbf{Z}$, $F(0)=\mathbb{E}\big(f(\mathbf{X})g(\mathbf{Z})\big)$, que por $\mathbf{X}$ e $\mathbf{Z}$ serem

independentes, é igual a $\mathbb{E}(f(\mathbf{X}))\mathbb{E}(g(\mathbf{Z})) = \mathbb{E}(f(\mathbf{X}))\mathbb{E}(g(\mathbf{X}))$, porque $\mathbf{X}$ e $\mathbf{Z}$ são identicamente distribuídos.
Sabemos também que, por definição de $\mathbf{Y}(\lambda)$, $\mathbf{Y}(1) = \mathbf{X}$, pelo que $F(1) = \mathbb{E}\big(f(\mathbf{X})g(\mathbf{X})\big)$. Resta agora mostrar que $F'(\lambda)$ existe e é tal que $F'(\lambda) > 0$, para todo $\lambda \in [0,1]$. Nesse caso,

$$F(0) \leq F(1) \Longleftrightarrow \mathbb{E}(f(\mathbf{X}))\mathbb{E}(g(\mathbf{X})) \leq \mathbb{E}\big(f(\mathbf{X})g(\mathbf{X})\big),$$

o que conclui a demonstração. A prova da diferenciabilidade de F usa o facto de $\mathbf{X}$ ser um vector Gaussiano e de f e g serem de classe C^∞, com derivadas parciais de primeira ordem limitadas. Seja $C = \Sigma^{-1}$ com entradas c_{ij}. A densidade do vector $\mathbf{X}$ pode escrever-se como

$$\phi(\mathbf{x}) = (2\pi)^{-\frac{k}{2}} (\det C)^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} \sum_{i,j=1}^{k} c_{ij} x_i x_j \right\}$$

e a distribuição condicional de $\mathbf{Y}(\lambda)$ quando $\mathbf{X} = \mathbf{x}$ é dada por

$$p(\lambda; x, y) = (1-\lambda^2)^{-\frac{k}{2}} \phi\left((1-\lambda^2)^{-\frac{1}{2}} (\lambda \mathbf{x} - \mathbf{y}) \right).$$

Assim, sendo $g(\lambda, \mathbf{x}) = \displaystyle\int_{\mathbb{R}^k} p(\lambda; x, y) g(\mathbf{y}) d\mathbf{y}$, temos,

$$F(\lambda) = \int_{\mathbb{R}^k} \phi(\mathbf{x}) f(\mathbf{x}) g(\lambda, \mathbf{x})\, d\mathbf{x}.$$

Seja agora $\phi_\lambda(\mathbf{x}) = (1-\lambda^2)^{-\frac{k}{2}} \phi\left((1-\lambda^2)^{-\frac{1}{2}} \mathbf{x} \right)$. Então,

$$g(\lambda, \mathbf{x}) = (\phi_\lambda * g)(\lambda \mathbf{x}) = \int_{\mathbb{R}^k} g(\lambda \mathbf{x} - \mathbf{y}) \phi_\lambda(\mathbf{y}) d\mathbf{y}.$$

Nestas condições as derivadas parciais $\frac{\partial g(\lambda, \mathbf{x})}{\delta x_i}$ existem e são limitadas. Além disso, como g é crescente e $\lambda > 0$, temos $\dfrac{\partial g}{\partial x_i}(\lambda, \mathbf{x}) \geq 0$. A partir

das equações de Plackett [37], 1954, obtém-se uma expressão conveniente para $\frac{\partial p}{\partial \lambda}$, de modo que,

$$F'(\lambda) = -\frac{1}{\lambda}\int_{\mathbb{R}^n} \phi(\mathbf{x}) f(\mathbf{x}) \left\{ \sum_{i,j} \sigma_{ij} \frac{\partial^2 g(\lambda, \mathbf{x})}{\partial x_i \partial x_j} - \sum_i x_i \frac{\partial g(\lambda, \mathbf{x})}{\partial x_i} \right\} d\mathbf{x}.$$

Integrando, agora, por partes

$$F'(\lambda) = \frac{1}{\lambda}\int_{\mathbb{R}^n} \phi(\mathbf{x}) \left\{ \sum_{i,j} \sigma_{ij} \frac{\partial f(\mathbf{x})}{\partial x_i} \frac{\partial g(\lambda, \mathbf{x})}{\partial x_j} \right\} d\mathbf{x}.$$

Uma vez que $f(\mathbf{x})$ é crescente, temos $\frac{\partial f}{\partial x_i} \geq 0$. Combinando este facto com $\sigma_{ij} \geq 0$ e com $\frac{\partial g}{\partial x_i}(\lambda, \mathbf{x}) \geq 0$, obtém-se $F'(\lambda) \geq 0$, como se pretendia mostrar. ■

O exemplo que apresentamos a seguir é uma consequência do resultado anterior.

Exemplo 1.4.5 *Sejam $X_1, \ldots, X_n$ variáveis aleatórias independentes com distribuição Normal de média zero e variância* 1.

Defina-se, para cada $i > 1$, $Y_i = \alpha X_i + \beta X_{i-1}$. Cada Y_i tem distribuição $N(0, \alpha^2 + \beta^2)$ e para $|i - j| = 1$, $\mathrm{Cov}\,(Y_i, Y_j) = \alpha\beta$. Se $\alpha\beta > 0$, as variáveis aleatórias $Y_2, \ldots, Y_n$ são associadas.

Antes de apresentarmos o exemplo seguinte, consideremos a definição de $m-$associação (Oliveira [33], 2000).

Definição 1.4.6 *As variáveis aleatórias $X_1, \ldots, X_n$ dizem-se $m-$associadas se, para $|k-i| \leq m$, X_i e X_k são associadas e se para $|k-i| > m$, X_i e X_k são independentes.*

Nota 1.4.7 *De acordo com esta definição, podemos dizer que as variáveis do exemplo anterior são $1-$associadas.*

Exemplo 1.4.8 *Sejam $X_1, \ldots, X_n$ variáveis aleatórias independentes, identicamente distribuídas, com função de distribuição F. Para $m \in \mathbb{R}$ fixo, considere-se*

$$Y_i = X_i + X_{i+1} + \cdots + X_{i+m}.$$

A sucessão de variáveis aleatórias assim obtida é $(m+1)-$associada. Se, tomarmos a função de distribuição F das variáveis $X_i, i = 1, \ldots, n$, Gaussiana centrada e com variância um, então cada Y_i tem distribuição Gaussiana e

$$\operatorname{Cov}(Y_i, Y_j) = \begin{cases} m - |i-j| + 1 & \text{se } |i-j| \leq m \\ 0 & \text{caso contrário.} \end{cases}$$

Vejamos ainda mais um exemplo de associação, onde a técnica utilizada consiste em tomar o mínimo de m variáveis aleatórias, independentes e identicamente distribuídas, sucessivas.

Exemplo 1.4.9 *Sejam $X_1, \ldots, X_n$ variáveis aleatórias independentes com função de distribuição contínua F. Seja $m \in \mathbb{N}$, fixo.*
Considere-se $Y_i = \min\{X_i, \ldots, X_{i+m}\}$. Cada uma destas variáveis aleatórias é identicamente distribuída com função de distribuição $G_Y(x) = 1 - (1 - F(x))^{m+1}$.
Vejamos, agora, qual a expressão para a função de distribuição conjunta do par (Y_1, Y_{k+1}) para $k = 1, \ldots, n-1$, fixo. Ora, como facilmente se verifica, se $k > m$, as variáveis aleatórias Y_1, Y_{k+1} são independentes, pelo que, neste caso,

$$P(Y_1 \leq y_1, Y_{k+1} \leq y_{k+1}) = \left(1 - (1 - F(y_1))^{m+1}\right)\left(1 - (1 - F(y_{k+1}))^{m+1}\right).$$

Seja, agora, $k \leq m$.

$$\begin{aligned} P(Y_1 \leq y_1, Y_{k+1} \leq y_{k+1}) = 1 - (1 - F(y_1))^{m+1} - (1 - F(y_{k+1}))^{m+1} + \\ + (1 - F(y_1))^k (1 - F(y_{k+1}))^k (1 - F(\max\{y_1, y_{k+1}\}))^{m+1-k}. \end{aligned}$$

A sucessão $\{Y_n\}_{n\in\mathbb{N}}$ assim obtida é $m-$associada.

Nota 1.4.10 *De uma maneira perfeitamente análoga, poder-se-ia obter uma sucessão associada de variáveis aleatórias identicamente distribuídas, construída à custa do máximo de m variáveis independentes e identicamente distribuídas sucessivas.*

1.5 Máximo das somas parciais.

Newman e Wright [32], 1981, estabelecem a versão, sob associação, de um resultado clássico sob independência, devido a Kolmogorov.

Teorema 1.5.1 (Newman e Wright [32], 1981) *Seja* $\mathbf{X} = (X_1, \ldots, X_n)$ *um vector associado de variáveis aleatórias reais centradas e de quadrado integráveis.*
Sendo $S_i = \sum_{j=1}^{i} X_j$ *e* $M_m = \max(S_1, \ldots, S_m)$, *tem-se que:*

$$\mathbb{E}\big(M_m^2\big) \leq \mathrm{Var}\,(S_m).$$

Demonstração: A demonstração é levada a cabo por indução sobre m.

- $m = 1$. Neste caso, $M_1 = \max(S_1) = X_1$. Pelo que o resultado é trivialmente verdadeiro.

- Suponhamos que a desigualdade é válida para $m-1$ variáveis aleatórias.

Considerem-se, agora, m variáveis $X_1, \ldots, X_m$ nas condições do Teorema. Definam-se:
$K_m = \min\{X_2 + \cdots + X_m, X_3 + \cdots + X_m, \ldots, X_m, 0\}$,
$L_m = \max\{X_2, X_2 + X_3, \ldots, X_2 + \cdots + X_m\}$ e
$J_m = \max\{0, L_m\}$.
Ora, $K_m = X_2 + \cdots + X_m - J_m$, é uma função não decrescente das variáveis aleatórias $X_i, i = 1, \ldots, m$, pelo que $\mathrm{Cov}\,(X_1, K_m) \geq 0$. Além

disso, $J_m^2 \leq L_m^2$ e $M_m = X_1 + J_m$. Temos, então,

$$\begin{aligned}\mathbb{E}(M_m^2) &= \mathbb{E}\left((X_1 + J_m)^2\right) = \mathrm{Var}\,(X_1) + 2\,\mathrm{Cov}\,(X_1, J_m) + \mathbb{E}(J_m^2) = \\ &= \mathrm{Var}\,(X_1) + 2\,\mathrm{Cov}\,(X_1, X_2 + \cdots + X_m) - 2\,\mathrm{Cov}\,(X_1, K_m) + \\ &\quad + \mathbb{E}(J_m^2).\end{aligned}$$

Ora $-2\,\mathrm{Cov}\,(X_1, K_m) \leq 0$ e, além disso, $\mathbb{E}(J_m^2) \leq \mathbb{E}(L_m^2)$. Assim,

$$\mathbb{E}(M_m^2) \leq \mathrm{Var}\,(X_1) + 2\,\mathrm{Cov}\,(X_1, X_2 + \cdots + X_m) + \mathbb{E}(L_m^2).$$

Usando agora a hipótese de indução, $\mathbb{E}(L_m^2) \leq \mathrm{Var}\,(X_2 + \cdots + X_m)$, e uma vez que

$$\begin{aligned}\mathrm{Var}\,(X_1) &+ 2\,\mathrm{Cov}\,(X_1, X_2 + \cdots + X_m) + \mathrm{Var}\,(X_2 + \cdots + X_m) = \\ &= \mathbb{E}\left((X_1 + (X_2 + \cdots + X_m))^2\right) = \\ &= \mathrm{Var}\,(S_m),\end{aligned}$$

o resultado segue imediatamente. ■

Como corolário deste teorema obtém-se a *Desigualdade de Kolmogorov*, a menos de um factor multiplicativo.

Corolário 1.5.2 *Sob as mesmas condições do Teorema anterior,*

$$P\big(\max_{1\leq i\leq m} |S_i| \geq \alpha\big) \leq \frac{2}{\alpha^2}\mathrm{Var}\,(S_m).$$

O teorema anterior, assim como o seu corolário dão-nos uma ferramenta, tal como no caso independente, para a obtenção de razões de convergência optimizadas para a *Lei Forte dos Grandes Números* como veremos no próximo capítulo.

1.6 A Lei Forte dos Grandes Números.

Nesta secção vamos referir dois resultados importantes, para sucessões associadas de variáveis aleatórias, relativos à *Lei Forte dos Grandes Números*. O primeiro é de Newman [31], 1984 e estabelece que *para uma*

sucessão estritamente estacionária e associada de variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$ *com variância finita, a Lei Forte dos Grandes Números verifica-se se*

$$\lim_{n\to\infty} n^{-1}\sum_{j=1}^{n}\mathrm{Cov}\,(X_1, X_j) = 0.$$

Em 1988, Birkel [5] obtém uma *Lei Forte dos Grandes Números*, para sucessões associadas não necessariamente estacionárias, sob uma condição muito próxima da de Kolmogorov no caso independente. Passamos a enunciar este resultado sob a forma de teorema. O resultado de Newman é um corolário deste.

Teorema 1.6.1 (Birkel [5], 1988) *Seja* $\{X_n\}_{n\in\mathbb{N}}$ *uma sucessão associada de variáveis aleatórias de quadrado integráveis. Para* $n \geq 1$, *seja* $S_n = \sum_{j=1}^{n} X_j$. *Suponha-se que*

$$\sum_{j=1}^{+\infty}\frac{1}{j^2}\mathrm{Cov}\,(X_j, S_j) < \infty.$$

Então, $\dfrac{S_n - \mathbb{E}(S_n)}{n} \underset{n\to\infty}{\longrightarrow} 0, \quad q.\ c..$

Demonstração: Suponha-se, sem perda de generalidade, que as variáveis aleatórias $X_1, \ldots, X_n$ são centradas. Pretende-se provar que

$$\frac{S_k}{k} \underset{k\to\infty}{\longrightarrow} 0, \quad \text{q. c..} \tag{1.13}$$

A demonstração é levada a cabo em dois passos:

- a subsucessão $\{2^{-n}S_{2^n}\}_{n\in\mathbb{N}}$ converge quase certamente para 0;
- $\left\{2^{-n}\max_{2^n<k\leq 2^{n+1}}|S_k - S_{2^n}|\right\}_{n\in\mathbb{N}}$ converge quase certamente para 0.

De facto, a prova da convergência quase certa desta duas subsucessões é suficiente para se concluir (1.13), já que se $k \in \mathbb{N}$ e $n_k \in \mathbb{N}$ são tais que $2^{n_k} < k \leq 2^{n_k+1}$, então

$$\left|\frac{S_k}{k}\right| \leq \left|\frac{S_k - S_{2^{n_k}}}{2^{n_k}}\right| + 2 \times \frac{\left|S_{2^{n_k}}\right|}{2^{n_k}},$$

pelo que

$$|\frac{S_k}{k}|\leq \max_{2^{n_k}<j\leq 2^{n_k+1}}\left|\frac{S_k-S_{2^{n_k}}}{2^{n_k}}\right|+2\times\frac{\left|S_{2^{n_k}}\right|}{2^{n_k}}.$$

Verifica-se assim que as duas etapas referidas anteriormente permitem demonstrar o resultado pretendido.

Para provar o primeiro passo recorre-se ao *Lema de Borel - Cantelli* que garante esta convergência desde que:

$$\forall_{\epsilon>0}\,,\ \sum_{n=1}^{\infty}P\Big(2^{-n}S_{2^n}\geq\epsilon\Big)<\infty.$$

De acordo com a *desigualdade de Chebyshev*,

$$P\Big(2^{-n}S_{2^n}\geq\epsilon\Big)=P\Big(\left|S_{2^n}\right|\geq 2^n\times\epsilon\Big)\leq\frac{\mathrm{Var}\left(S_{2^n}\right)}{2^{2n}\epsilon^2}=\frac{4^{-n}\mathrm{Var}\left(S_{2^n}\right)}{\epsilon^2}.$$

Mas,

$$\mathrm{Var}\left(S_{2^n}\right)=\sum_{j,k=1}^{2^n}\mathrm{Cov}\left(X_j,X_k\right)\leq 2\sum_{j=1}^{2^n}\mathrm{Cov}\left(X_j,S_j\right).$$

Temos então,

$$\begin{aligned}\sum_{n=1}^{\infty}P\Big(2^{-n}S_{2^n}\geq\epsilon\Big) &\leq \frac{2}{\epsilon^2}\sum_{n=1}^{\infty}\sum_{j=1}^{2^n}4^{-n}\mathrm{Cov}\left(X_j,S_j\right)=\\ &= \frac{2}{\epsilon^2}\sum_{j=1}^{\infty}\Big(\sum_{n\in\mathbb{N}:2^n\geq j}4^{-n}\Big)\mathrm{Cov}\left(X_j,S_j\right).\end{aligned}$$

Ora, $\displaystyle\sum_{k\geq k_0}4^{-k}=\frac{4}{3}\times 4^{-k_0}$, pelo que

$$\sum_{n\in\mathbb{N}:2^n\geq j}4^{-n}=\sum_{n\in\mathbb{N}:n\geq\frac{\log j}{\log 2}}=\frac{4}{3}\times 4^{-\frac{\log j}{\log 2}}=\frac{4}{3\,j^2}.$$

Em conclusão,

$$\sum_{n=1}^{\infty}P\Big(2^{-n}S_{2^n}\geq\epsilon\Big)\leq\frac{8}{3\epsilon^2}\sum_{j=1}^{\infty}\frac{1}{j^2}\mathrm{Cov}\left(X_j,S_j\right)<\infty.$$

Relativamente ao segundo passo, usa-se, novamente, o *Lema de Borel - Cantelli* e recorre-se à desigualdade do tipo Kolmogorov referida anteriormente para majorar o termo geral da série de interesse.

$$P\Big(2^{-n}\max_{2^n<k\leq 2^{n+1}}\big|S_k-S_{2^n}\big|\geq\epsilon\Big)\leq\frac{2}{\epsilon^2\times 4^n}\mathrm{Var}\left(S_{2^{n+1}}-S_{2^n}\right).$$

Ora,

$$\begin{aligned}\mathrm{Var}\left(S_{2^{n+1}}-S_{2^n}\right) &= \mathrm{Var}\left(\sum_{j=2^n+1}^{2^{n+1}}X_j\right)=\sum_{j,k=2^n+1}^{2^{n+1}}\mathrm{Cov}\left(X_j,X_k\right)\leq\\ &\leq \sum_{j,k=1}^{2^{n+1}}\mathrm{Cov}\left(X_j,X_k\right)=\mathrm{Var}\left(S_{2^{n+1}}\right).\end{aligned}$$

Temos então,

$$\sum_{n=1}^{\infty}P\Big(2^{-n}\max_{2^n<k\leq 2^{n+1}}\big|S_k-S_{2^n}\big|\geq\epsilon\Big)\leq\frac{8}{\epsilon^2}\sum_{n=1}^{\infty}\frac{1}{4^{n+1}}\mathrm{Var}\left(S_{2^{n+1}}\right)<\infty,$$

como foi visto no primeiro passo. ■

Estes resultados de Newman e Birkel requerem a existência de momentos de segunda ordem, uma vez que a estrutura de covariância é usada para a descrição das propriedades assimptóticas do processo. Com o objectivo de dar resposta a situações em que os momentos de segunda ordem não são finitos, Louhichi [28], 2000, prova uma *Lei Forte dos Grandes Números* do tipo Marcinkiewicz - Zygmund de ordem p, $p\in[1,2)$, isto é, estuda o problema da convergência quase certa de $n^{-\frac{1}{p}}S_n$, para sucessões associadas não necessariamente estacionárias.

Seja $\{X_n\}$ uma sucessão de variáveis aleatórias reais com variância infinita. A estratégia usada em Louhichi [28],2000, consiste em obter um controle entre as covariâncias de funções apropriadas das variáveis aleatórias $X_i, i=1,\ldots,n$, que no limite correspondem exactamente às covariâncias entre as variáveis. Considere-se, então, para $i,r\in\mathbb{N}$ fixos,

$v \in \mathbb{R}^+$, também fixo, $x, y, u \in \mathbb{R}$,

$$\begin{aligned} \Delta_{i,i+r}(x,y) &= P(X_i \geq x, X_{i+r} \geq y) - P(X_i \geq x)P(X_{i+r} \geq y); \\ g_v(u) &= \max\{\min\{u, v\}, -v\}; \\ G_{i,i+r}(v) &= \text{Cov}\,(g_v(X_i), g_v(X_{i+r})) = \int_{|x|\leq v}\int_{|y|\leq v} \Delta_{i,i+r}(x,y)dxdy. \end{aligned}$$

Note-se que da associação da sucessão $\{X_n\}_{n\in\mathbb{N}}$ segue-se $\Delta_{i,i+r}(x,y) \geq 0$ para todo $x, y \in \mathbb{R}$, e que $G_{i,i+r}(v)$ é positiva. Além disso esta função está limitada por v^2. Note-se ainda que se as variáveis $\{X_i\}_{i\in\mathbb{N}}$ tiverem variância finita, então $G_{i,i+r}(\infty) = \text{Cov}\,(X_i, X_{i+r})$.

Teorema 1.6.2 (Louhichi [28], 2000) *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão associada de variáveis aleatórias reais e $p \in [1,2[$ fixo. Suponhamos que existe uma variável aleatória positiva X tal que $\mathbb{E}(X^p) < \infty$ e que, para qualquer $x \in \mathbb{R}^+$, se tem $\sup_i P(|X_i| > x) \leq P(X > x)$. Se*

$$\sum_{1\leq i<j\leq\infty} \int_{j^{\frac{1}{p}}}^{+\infty} v^{-3} G_{i,j}(v)dv < \infty, \tag{1.14}$$

então

$$\lim_{n\to\infty} n^{-\frac{1}{p}} S_n = 0, \quad \textit{quase certamente.}$$

A prova deste Teorema faz-se via *Lema de Borel-Cantelli* e através da Proposição que enunciamos de seguida, aplicada à sucessão associada $\{X_n\}_{n\in\mathbb{N}}$.

Proposição 1.6.3 *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão associada de variáveis aleatórias reais. Suponhamos que são verificadas todas as condições do Teorema 1.6.2 para algum $p \in [1,2[$. Então, para qualquer $\epsilon > 0$, tem-se*

$$\sum_{n=1}^{\infty} n^{-1} P\left(M_n > \epsilon\, n^{\frac{1}{p}}\right) < \infty,$$

onde $M_n = \max\{S_1, \ldots, S_n\}$.

Vejamos, agora, como Louhichi [28],2000, obtém um majorante para $P\left(M_n > x\right), x > 0$, sem que as variáveis tenham, necessariamente, variância finita.

Lema 1.6.4 *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão associada de variáveis aleatórias reais. Suponhamos que existe uma variável aleatória positiva X tal que para qualquer $x > 0$, se tem*

$$\sup_{i>0} P(|X_i| > x) \leq P(X > x).$$

Então para quaisquer números reais positivos x e A,

$$\begin{aligned} P\left(M_n > x\right) \quad \leq \quad & \frac{4n}{x^2}\mathbb{E}\left(X^2 \mathbb{1}_{(0,A]}(X)\right) + \frac{4n}{x}\mathbb{E}\left(X\,\mathbb{1}_{(A,\infty)}(X)\right) + \\ & + \frac{4nA^2}{x^2}P(X > A) + \frac{8}{x^2}\sum_{1\leq i<j\leq n} G_{i,j}(A). \end{aligned}$$

Demonstração: Definam-se, para $A \in \mathbb{R}^+$, as variáveis aleatórias

$$\bar{X}_i = \bar{X}_{i,A} = \max\{\min\{X_i, A\}, -A\}, \quad \text{e} \quad \widetilde{X}_i = \widetilde{X}_{i,A} = X_i - \bar{X}_i.$$

Seja, ainda, $\bar{S}_n = \sum_{i=1}^{n}(\bar{X}_i - \mathbb{E}(\bar{X}_i))$ e $\bar{M}_n = \max\{\bar{S}_1, \ldots, \bar{S}_n\}$. É claro que, $M_n \leq \bar{M}_n + \sum_{i=1}^{n}(|\widetilde{X}_i| + \mathbb{E}|\widetilde{X}_i|)$. Então, a desigualdade de Markov garante que

$$P(M_n > x) \leq P(\bar{M}_n > \frac{x}{2}) + \frac{4}{x}\sum_{i=1}^{n}\mathbb{E}|X_i| \leq \frac{4}{x^2}\mathbb{E}(\bar{M}_n^2) + \frac{4}{x}\sum_{i=1}^{n}\mathbb{E}|\widetilde{X}_i|.$$

Ora, as variáveis $\bar{X}_i - \mathbb{E}(\bar{X}_i)$ têm variância finita além de serem, obviamente, centradas. Além disso, a sucessão $\{\bar{X}_n\}_{n\in\mathbb{N}}$ é associada, por ser uma função não decrescente de $\{X_n\}_{n\in\mathbb{N}}$. Podemos, assim, aplicar à sucessão $\{X_n - \mathbb{E}(\bar{X}_n)\}_{n\in\mathbb{N}}$ a desigualdade de Newman e Wright, Teorema 1.5.1, obtendo,

$$P(M_n > x) \leq \frac{4}{x^2}\sum_{i=1}^{n}\mathbb{E}(\bar{X}_i^2) + \frac{8}{x^2}\sum_{1\leq i<j\leq n}\mathrm{Cov}\left(\bar{X}_i, \bar{X}_j\right) + \frac{4}{x}\sum_{i=1}^{n}\mathbb{E}|\widetilde{X}_i|.$$

Basta, agora, notar que

$$\mathbb{E}(\bar{X}_i^2) \leq \mathbb{E}\left(X^2 \mathbb{1}_{(0,A]}(X)\right) + A^2 P(X > A)$$

e também,

$$\mathbb{E}|\widetilde{X}_i| \leq \mathbb{E}\left((|X_i| - A)\, \mathbb{1}_{(A,+\infty]}(X)\right) \leq \mathbb{E}\left(X\, \mathbb{1}_{(A,+\infty]}(X)\right).$$

O resultado do Lema é, agora, imediato. ■

A aplicação deste Lema a $n^{-1}P(M_n > \epsilon\, n^{\frac{1}{p}})$, fazendo $x = \epsilon\, n^{\frac{1}{p}}$ e $A = n^{\frac{1}{p}}$, conduz a

$$\begin{aligned} n^{-1}P(M_n > \epsilon\, n^{\frac{1}{p}}) \quad \leq \quad & \frac{4}{\epsilon n^{\frac{1}{p}}}\mathbb{E}\left(X\, \mathbb{1}_{(n,+\infty)}(X^p)\right) + \frac{4}{\epsilon^2 n^{\frac{2}{p}}}\mathbb{E}\left(X^2 \mathbb{1}_{(0,n]}(X^p)\right) + \\ & + \frac{1}{\epsilon^2}P(X > n^{\frac{1}{p}}) + \frac{8\, n^{-1-\frac{2}{p}}}{\epsilon^2} \sum_{1 \leq i < j \leq n} G_{i,j}(n^{\frac{1}{p}}). \end{aligned}$$

Usando o Teorema de Fubini, podemos escrever,

$$\sum_{n=1}^{+\infty} n^{-\frac{1}{p}}\mathbb{E}\left(X\, \mathbb{1}_{(n,+\infty)}(X^p)\right) = \mathbb{E}\left(X \sum_{n=1}^{+\infty} n^{-\frac{1}{p}} \mathbb{1}_{(n,+\infty)}(X^p)\right) \leq c_p \mathbb{E}\left(X^p\right),$$

com c_p uma constante que depende apenas de p.

Quanto a $\sum_{n=1}^{+\infty} n^{-\frac{2}{p}}\, \mathbb{E}\left(X^2 \mathbb{1}_{(0,n]}(X^p)\right) + \sum_{n=1}^{+\infty} P(X > n^{\frac{1}{p}})$ também pode ser majorada por $c_p\, \mathbb{E}\left(X^p\right)$, novamente por utilização do Teorema de Fubini.

Vejamos agora $\sum_{n=1}^{+\infty} n^{-1-\frac{2}{p}} \sum_{1 \leq i < j \leq n} G_{i,j}(n^{\frac{1}{p}})$.

Para $p \in [1, 2[$, $s \geq 1$ fixo e $x, y \in \mathbb{R}$, tem-se $\left(\max\{|x|, |y|, s^{1/p}\}\right)^{-2} = \int_0^1 \mathbb{1}_{(0,u^{-1/2}]}(|x|) \mathbb{1}_{(0,u^{-1/2}]}(|y|) \mathbb{1}_{[u,1]}(s^{-2/p}) du.$

Assim, por aplicação do Teorema de Fubini ,

$$\int_{\mathbb{R}^2} \left(\max\{|x|, |y|, s^{1/p}\}\right)^{-2} \Delta_{i,j}(x,y)\, dxdy = \int_0^1 \mathbb{1}_{[u,1]}(s^{-2/p}) G_{i,j}(u^{-1/2}) du.$$

De acordo com estes resultados,

$$\sum_{n=1}^{+\infty} n^{-1-\frac{2}{p}} \sum_{1\le i<j\le n} G_{i,j}(n^{\frac{1}{p}}) =$$

$$= \sum_{1\le i<j\le +\infty} \int_{\mathbb{R}^2} \sum_{n=1}^{+\infty} n^{-1-2/p} \mathbb{1}_{[n,+\infty)} \left(\max\{|x|^p, |y|^p, j\}\right) \Delta_{i,j}(x,y)dxdy \le$$

$$\le c_p \sum_{1\le i<j\le +\infty} \int_{\mathbb{R}^2} \left(\max\{|x|^p, |y|^p, j\}\right)^{-2/p} \Delta_{i,j}(x,y)dxdy =$$

$$= 2\, c_p \sum_{1\le i<j\le +\infty} \int_{j^{1/p}}^{+\infty} v^{-3} G_{i,j}(v)\, dv,$$

que é finita de acordo com a hipótese do Teorema.
Nestas condições, $\sum_{n=1}^{+\infty} n^{-1}\, P\left(M_n > \epsilon\, n^{1/p}\right) < +\infty$, para todo $\epsilon > 0$ e $1 < p < 2$. A aplicação do *Lema de Borel Cantelli* completa a demonstração do Teorema. ■

1.7 Convergência do Processo Empírico

1.7.1 Introdução

Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas com função de distribuição marginal contínua F.

A função de distribuição empírica $\mathbb{F}_n$ de $X_1, \ldots, X_n$ é definida como:

$$\mathbb{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{(-\infty,x]}(X_i),\ x \in \mathbb{R}.$$

O processo empírico β_n construído com base em $X_1, \ldots, X_n$ é definido por,

$$\beta_n(x) = \sqrt{n}\big[\mathbb{F}_n(x) - F(x)\big],\ x \in \mathbb{R}. \tag{1.15}$$

Um dos problemas que surge quando se estuda o processo empírico a partir de variáveis aleatórias associadas, é o da ligação entre as covariâncias de funções indicatrizes das variáveis aleatórias e as covariâncias

entre as variáveis aleatórias de base. Yu [53], 1993, obtém uma relação para variáveis aleatórias U_1, U_2, associadas, com distribuição uniforme no intervalo $[0,1]$, estabelecendo que

$$0 \leq \mathrm{Cov}\left(\mathbb{1}_{[0,s)}(U_1), \mathbb{1}_{[0,t)}(U_2)\right) \leq 4\,\mathrm{Cov}^{1/3}(U_1, U_2).$$

No entanto, Roussas [41], 1998, estabelece uma desigualdade do mesmo tipo, válida para variáveis aleatórias com distribuição contínua qualquer e densidades limitadas, combinando dois resultados, um de Newman [31], 1984 (a equação (1.3)) e outro de Sadikova, (corolário do Teorema 1 em Sadikova [42], 1966). Sadikova fornece uma extensão da desigualdade Berry - Essén ao caso bidimensional, isto é, estabelece um majorante da distância máxima entre duas funções de distribuição bidimensionais, em termos das correspondentes funções características.

Teorema 1.7.1 (Sadikova [42], 1966) *Sejam (X,Y) e (X',Y') dois vectores aleatórios. Considerem-se, respectivamente, F e G as suas funções de distribuição e ϕ e ψ as suas funções características. Tome-se, para cada $s,t \in \mathbb{R}$, $\hat{\phi}(s,t) = \phi(s,t) - \phi(s,0)\phi(0,t)$ e $\hat{\psi}(s,t) = \psi(s,t) - \psi(s,0)\psi(0,t)$. Seja, ainda, $A_1 = \sup_{u,v\in\mathbb{R}} \dfrac{\partial G}{\partial x}(u,v)$, $A_2 = \sup_{u,v\in\mathbb{R}} \dfrac{\partial G}{\partial y}(u,v)$. Então, para qualquer $T > 0$,*

$$\begin{aligned}\sup_{x,y\in\mathbb{R}} |F(x,y) - G(x,y)| \;\leq\; & \frac{2}{(2\pi)^2} \int_{[-T,T]\times[-T,T]} \left| \frac{\hat{\phi}(s,t) - \hat{\psi}(s,t)}{st} \right| ds\,dt + \\ & + 2\sup_{x\in\mathbb{R}} |F(x,+\infty) - G(x,+\infty)| + \\ & + 2\sup_{y\in\mathbb{R}} |F(+\infty,y) - G(+\infty,y)| + \\ & + \frac{2(A_1+A_2)}{T}(3\sqrt{2} + 4\sqrt{3}). \end{aligned} \tag{1.16}$$

Demonstração: A demonstração deste resultado encontra-se dividida em duas partes. A primeira parte é iniciada com a especificação duma determinada densidade, Q_T, usada na prova do Teorema como um artifício de ligação entre as distribuições, arbitrárias, dos vectores aleatórios

(X, Y) e (X', Y').

Considere-se, então, a densidade unidimensional, $q(x) = \frac{3}{8\pi}\left(\frac{\sin(x/4)}{x/4}\right)^4$.

Esta densidade tem como função característica

$$h(t) = \begin{cases} 0 & \text{se } |t| \geq 1, \\ 2(1-|t|)^3 & \text{se } \frac{1}{2} \leq |t| < 1, \\ 1 - 6t^2 + 6|t|^3 & \text{se } 0 \leq |t| < \frac{1}{2}, \end{cases}$$

e satisfaz as condições $\int_{\mathbb{R}} x^2 q(x)dx = 12$ e $\int_{\mathbb{R}} |x| q(x)dx < 3\sqrt{2}$.

Considere-se agora $\lambda \in \mathbb{R}$, tal que

$$\int_{(-\infty,-\lambda]} q(x)dx = \int_{(\lambda,+\infty)} q(x)dx = 1 - \frac{\sqrt{3}}{2}.$$

A *Desigualdade de Chebychev* permite concluir $\lambda \leq 4\sqrt{3}$.

Defina-se, para $T > 0$, a densidade bidimensional $Q_T(x, y) = T\, q(Tx)Tq(Ty)$.

A sua correspondente função característica, $h(s/T)h(t/T)$ é igual a zero fora do quadrado $\max(|s|, |t|) \leq T$. Além disso,

$$\begin{aligned} \int_{\mathbb{R}^2} |x|\, Q_T\left(x \pm \frac{\lambda}{T}, y \pm \frac{\lambda}{T}\right) dxdy &= \int_{\mathbb{R}^2} |y|\, Q_T\left(x \pm \frac{\lambda}{T}, y \pm \frac{\lambda}{T}\right) dxdy \leq \\ &\leq \frac{3\sqrt{2}}{T} + \frac{\lambda}{T} \end{aligned} \tag{1.17}$$

e

$$\begin{aligned} &\int_{(-\infty,0]\times(-\infty,0]} Q_T\left(x \pm \frac{\lambda}{T}, y \pm \frac{\lambda}{T}\right) dxdy = \\ &\quad = \int_{[0,+\infty)\times[0,+\infty)} Q_T\left(x \pm \frac{\lambda}{T}, y \pm \frac{\lambda}{T}\right) dxdy = \\ &\quad = \left(\frac{\sqrt{3}}{2}\right)^2 + \frac{3}{4} \end{aligned} \tag{1.18}$$

Sejam (X, Y) e (X', Y') vectores aleatórios com densidades de probabilidade P e Q, respectivamente e seja (X'', Y'') o vector aleatório com

densidade de probabilidade Q_T, anteriormente definida, independente de (X,Y) e de (X',Y'). Tome-se $R(u,v) = P(X < u, Y < v) - P(X' < u, Y' < v)$. Existem duas possibilidades relativamente ao supremo desta diferença:

(i) $\gamma = \sup_{u,v\in\mathbb{R}} |R(u,v)| = \sup_{u,v\in\mathbb{R}} R(u,v) = \gamma_1$;

ou

(ii) $\gamma = \sup_{u,v\in\mathbb{R}} |R(u,v)| = \sup_{u,v\in\mathbb{R}} (-R(u,v)) = \gamma_2$.

Estabeleça-se

$$R_1(x,y) = P\left(X + X'' < x, Y + Y'' < y\right) - P\left(X' + X'' < x, Y' + Y'' < y\right).$$

Seja $\theta = 1$, no caso (i) e $\theta = -1$, no caso (ii). Então, de acordo com a fórmula para a convolução de duas distribuições, tem-se que

$$R_1\left(x + \frac{\theta\lambda}{T}, y + \frac{\theta\lambda}{T}\right) = \int_{\mathbb{R}^2} R(x-u, y-v) Q_T\left(u + \frac{\theta\lambda}{T}, v + \frac{\theta\lambda}{T}\right) du\, dv.$$

Concentremo-nos no caso (i), já que o caso (ii) é análogo. Em primeiro lugar note-se que $\gamma_1 \leq \gamma$. Vejamos uma estimativa para a diferença entre γ_1 e γ. Seja $\epsilon > 0$ e x_ϵ, y_ϵ números reais tais que $R(x_\epsilon, y_\epsilon) \geq \gamma - \epsilon$. Para $u \leq 0$, $v \leq 0$, temos

$$\begin{aligned} R(x_\epsilon, y_\epsilon) &= F(x_\epsilon - u, y_\epsilon - v) - G(x_\epsilon - u, y_\epsilon - v) \geq \\ &\geq F(x_\epsilon, y_\epsilon) - G(x_\epsilon, y_\epsilon) - (G(x_\epsilon - u, y_\epsilon - v) - G(x_\epsilon, y_\epsilon)) \geq \\ &\geq R(x_\epsilon, y_\epsilon) - A_1|u| - A_2|v| \geq \gamma - \epsilon - A_1|u| - A_2|v|. \end{aligned}$$

Multiplicando a última desigualdade por $Q_T(u + \lambda/T, v + \lambda/T)$, e integrando sobre a região $(-\infty, 0] \times (-\infty, 0]$, obtém-se, de acordo com (1.17) e (1.18),

$$\begin{aligned} J_1 &= \int_{(-\infty,0]\times(-\infty,0]} R(x_\epsilon - u, y_\epsilon - v)\, Q_T\left(u + \frac{\lambda}{T}, v + \frac{\lambda}{T}\right) du\, dv \geq \\ &\geq \frac{3}{\lambda}(\gamma - \epsilon) - \frac{(A_1 + A_2)}{T}(3\sqrt{2} + \lambda). \end{aligned}$$

Por outro lado, fazendo $D = [0, +\infty) \times \mathbb{R}$ e $D' = \mathbb{R} \times [0, +\infty)$, temos, ainda, a seguinte majoração

$$|J_2| = \left| \int_D R(x_\epsilon - u, y_\epsilon - v)\, Q_T \left(u + \frac{\lambda}{T}, v + \frac{\lambda}{T} \right) du\, dv \right| \leq \frac{\gamma}{4},$$

que permanece a mesma se a integração for levada a cabo ao longo de D'. Assim,

$$\begin{aligned} \gamma_1 &\geq R_1 \left(x_\epsilon + \frac{\lambda}{T}, y_\epsilon + \frac{\lambda}{T} \right) \geq |J_1| - |J_2| \geq \\ &\geq \frac{1}{2}\gamma - \frac{3}{4}\epsilon - \frac{A_1 + A_2}{T}(3\sqrt{2} + \lambda). \end{aligned}$$

Uma vez que ϵ é qualquer, obtém-se

$$2\gamma_1 + \frac{2(A_1 + A_2)}{T}(3\sqrt{2} + \lambda) \geq \gamma. \tag{1.19}$$

Obtém-se exactamente a mesma desigualdade para (ii).
A segunda parte da demonstração diz respeito à majoração de γ_1.
A fórmula de inversão dá origem à seguinte expressão,

$$\begin{aligned} &\int_{\mathbb{R}^2} \left(F(x-u, y-v) - F(x-u, +\infty) F(+\infty, y-v) \right) Q_T(u, v)\, dudv = \\ &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \frac{\phi(s,t) - \phi(s,0)\phi(0,t)}{-st} \quad \mathrm{e}^{-i(sx+ty)} h\left(\frac{s}{T}\right) h\left(\frac{t}{T}\right) dsdt. \end{aligned}$$

Obtém-se, do mesmo modo, uma expressão similar para G. Nestas condições vem

$$\begin{aligned} &\int_{\mathbb{R}^2} R(x-u, y-v) Q_T(u, v) dudv - \int_{\mathbb{R}^2} \left(F(x-u, +\infty) F(+\infty, y-v) - \right. \\ &\qquad \left. - G(x-u, +\infty) G(+\infty, y-v) \right) Q_T(u, v) dudv = \\ &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \frac{\hat{f}(s,t) - \hat{g}(s,t)}{-st} \quad \mathrm{e}^{-i(sx+ty)} h\left(\frac{s}{T}\right) h\left(\frac{t}{T}\right) dsdt. \end{aligned} \tag{1.21}$$

Note-se, ainda, que

$$|F(x-u,+\infty)F(+\infty,y-v)-G(x-u,+\infty)G(+\infty,y-v)| \leq$$
$$\leq |F(x-u,+\infty)-G(x-u,+\infty)|+|F(+\infty,y-v)-G(+\infty,y-v)|\,.$$

Finalmente (1.19) e (1.21) implicam o resultado pretendido, equação (1.16).
■

Enunciamos, de seguida, como corolário deste Teorema, a desigualdade que fornece uma estimativa da distância máxima entre a função de distribuição conjunta de duas variáveis aleatórias e as respectivas funções de distribuição marginais.

Corolário 1.7.2 *Para quaisquer duas variáveis U e V, sejam $F_{U,V}, F_U$ e F_V as suas funções de distribuição conjunta e marginais, respectivamente. Suponhamos que F_U e F_V têm funções de densidade de probabilidade f_U e f_V limitadas por M_0. Sejam ainda $\phi_{U,V}, \phi_U$ e ϕ_V as funções características conjunta e marginais, respectivamente, das variáveis U e V. Nestas condições,*

$$\sup_{u,v\in\mathbb{R}}\left|F_{U,V}(u,v)-F_U(u)F_V(v)\right| \leq$$
$$\leq \frac{1}{2\pi^2}\int_{[-T,T]\times[-T,T]}\left|\frac{\phi_{U,V}(s,t)-\phi_U(s)\phi_V(t)}{st}\right| ds\,dt+\frac{45M_0}{T}, \tag{1.23}$$

onde T é um número positivo qualquer.

Demonstração: A aplicação da desigualdade (1.16) às funções $F_{U,V}$ e $F_U F_V$ conduz a

$$\sup_{u,v\in\mathbb{R}}\left|F_{U,V}(u,v)-F_U(u)F_V(v)\right| \leq$$
$$\leq \frac{1}{2\pi^2}\int_{[-T,T]\times[-T,T]}\left|\frac{\phi_{U,V}(s,t)-\phi_U(s)\phi_V(t)}{st}\right| ds\,dt+\frac{4M_0}{T}(3\sqrt{2}+4\sqrt{3}),$$

uma vez que se tem
$\sup_{x\in\mathbb{R}}|F_{U,V}(x,+\infty)-F_U(x)| = \sup_{y\in\mathbb{R}}|F_{U,V}(+\infty,y)-F_V(y)| = 0$, sendo

$M_0 = \sup_u F_U(u) = \sup_v F_V(v)$. Tomando agora $4 \times (3\sqrt{2} + 4\sqrt{3}) \cong 45$, o resultado é imediato. ■

Vejamos, agora, como usar esta desigualdade para controlar a independência relativa de pares de variáveis aleatórias, em termos da covariância entre elas.

A *desigualdade de Newman* (1.3), pode-se escrever, utilizando a notação dos resultados precedentes

$$|\phi_{U,V}(s,t) - \phi_U(s)\,\phi_V(t)| \leq |s\,t|\mathrm{Cov}\,(U,V),$$

para quaisquer $s, t \in \mathbb{R}$. Temos assim um majorante para a função integranda em (1.23). O resultado que se segue é, agora, imediato.

Lema 1.7.3 (Roussas [41], 1998) *Com a notação do corolário anterior e sob as mesmas condições, se as variáveis aleatórias U e V são positivamente associadas, então*

(*i*) $\sup_{u,v\in\mathbb{R}} \big|F_{U,V}(u,v) - F_U(u)F_V(v)\big| \leq M_1 \left(T^2\mathrm{Cov}\,(U,V) + \dfrac{1}{T}\right)$, *onde* $M_1 = \max\{2/\pi^2, 45M_0\}$;

(*ii*) *A expressão do lado direito na desigualdade anterior é minimizada para* $T = \big(2\mathrm{Cov}\,(U,V)\big)^{-1/3}$, *desde que* $\mathrm{Cov}\,(U,V) \neq 0$, *pelo que se tem, fazendo* $M = \dfrac{1}{2M_1}$,

$$\sup_{u,v\in\mathbb{R}} \big|F_{U,V}(u,v) - F_U(u)F_V(v)\big| \leq M\mathrm{Cov}^{\,1/3}\big(U,V\big). \qquad (1.25)$$

Nota 1.7.4 *Este resultado, que estabelece um majorante para a covariância entre funções indicatrizes de variáveis associadas em termos da covariância entre as variáveis de base, é usado inúmeras vezes em todo este trabalho.*

Posto isto, seja $D[-\infty, +\infty]$ o espaço das funções de domínio $\mathbb{R}$ que são contínuas à direita e com limites laterais à esquerda finitos e, além disso,

com limites finitos em $\pm\infty$. Suponhamos o espaço $D[-\infty,+\infty]$ munido da topologia de Skohorod. Esta é a ferramenta clássica para o estudo da convergência fraca do processo empírico $\{\beta_n(x), x \in \mathbb{R}\}$ para um processo Gaussiano $\mathbb{B}$ com média zero e covariância $\mathrm{Cov}\,(\mathbb{B}(x),\mathbb{B}(y)) = \sum_{k\in\mathbb{Z}} \mathrm{Cov}\,\left(\mathbb{1}_{(-\infty,x]}(X_0), \mathbb{1}_{[-\infty,y]}(X_k)\right)$. Este estudo origina resultados bastante úteis acerca do comportamento assimptótico de uma ampla classe de funcionais. A prova deste tipo de teoremas requer dois passos:

- estabelecer a convergência das distribuições de dimensão finita da sucessão de processos;
- estabelecer uma condição de fineza para a dita sucessão.

A primeira parte prova-se, em geral, usando um Teorema do Limite Central adequado. O segundo passo é aquele que é normalmente mais difícil de provar.

Note-se que, se se admitir que a função de distribuição F, das variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$, é contínua em $\mathbb{R}$, basta mostrar a convergência do processo empírico uniforme $\{\alpha_n(x), x \in [0,1]\}$(Billingsley [4], 1968).

Assumindo que X_1 tem densidade limitada f, tanto quanto sabemos, até agora, a convergência das distribuições de dimensão finita de $\{\alpha_n\}_{n\in\mathbb{N}}$ foi provada sob a condição de $\sum_{n\geq 1} \mathrm{Cov}^{1/3}(X_1, X_n) < \infty$, por Yu [53], 1993, onde o expoente $1/3$ é justificado pela equação (1.25). Assim, a condição $\mathrm{Cov}\,(X_1, X_n) = O(n^{-3})$ parece ser necessária para a prova da convergência do processo empírico.

Para a prova da fineza da sucessão de processos empíricos $\{\alpha_n\}_{n\in\mathbb{N}}$, esta condição não basta e Yu [53], 1993, foi o primeiro autor a obter uma condição suficiente para sucessões estacionárias e associadas de variáveis aleatórias reais. Ele supôs

$$\mathrm{Cov}\,\big(X_1, X_n\big) = O(n^{-b}), \quad \text{para algum } b > 15/2,$$

tendo esta hipótese sido enfraquecida posteriormente por Shao e Yu [43], 1996, provando-se que seria suficiente tomar $b > (3+\sqrt{33})/2$. Recente-

mente, em 2000, Louhichi [27] estabelece um novo critério para a prova daquele segundo item, conseguindo uma condição suficiente mais fraca, a saber,

$$\mathrm{Cov}\left(X_1, X_n\right) = O(n^{-b}), \quad \text{para algum } b > 4,$$

provando, ainda, que com o método utilizado esta condição é também necessária.

Como já referimos anteriormente, este tipo de convergência é bastante útil para a prova de resultados acerca de algumas funcionais de trajectórias, tais como aquelas que envolvem o supremo. A estatística de Kolmogorov Smirnov é um exemplo.
No entanto, se pretendemos estudar estatísticas como a Cramér-von Mises, W_n^2, usada em testes de ajustamento,

$$W_n^2 = \int_{\mathbb{R}} n\left(\mathbb{F}_n(x) - F(x)\right)^2 dF(x) = \int_{\mathbb{R}} \beta_n^2(x)\, dF(x) = \int_{[0,1]} \alpha_n^2(t)\, dt,$$

é suficiente considerar as trajectórias do processo como elementos de $L^2[0,1]$, espaço das funções de $[0,1]$ para $\mathbb{R}$ de quadrado integráveis. De facto, esta estatística é a norma $L^2[0,1]$ do processo empírico.
Sendo a topologia em $L^2[0,1]$ mais fraca que a de Skohorod, podemos esperar obter condições suficientes para a convergência do processo empírico naquele espaço que sejam também mais fracas. O estudo desta convergência foi, em primeiro lugar, levado a cabo por Oliveira e Suquet [35, 36], 1995 e 1998, obtendo como condição suficiente $\sum_{n\geq 1} \mathrm{Cov}^{1/3}(X_1, X_n) < \infty$. Mais tarde, Morel [29], 1999, prova uma condição necessária e suficiente para a convergência do processo empírico uniforme, sob estacionaridade, em $L^2[0,1]$, a saber, $\sum_n \left(\frac{2}{3} - \mathbb{E}\left(\max\left(X_0, X_n\right)\right)\right) < +\infty$.

Um dos principais objectivos deste trabalho é precisamente contribuir para a especificação de limites críticos para testes que envolvam a estatística de Cramér-von Mises, pelo que o estudo da convergência do processo empírico em $L^2[0,1]$ é suficiente.

1.7.2 Convergência em $L^2[0,1]$

Consideremos o processo empírico $\{\beta_n(t), t \in [0,1]\}$ construído com base numa sucessão $\{X_n\}_{n\in\mathbb{N}}$ associada, de variáveis aleatórias reais, com função de distribuição marginal F, contínua:

$$\beta_n(t) = \frac{1}{\sqrt{n}} \left(\sum_{i=1}^{n} \mathbb{1}_{(0,t]} X_i - F(t) \right).$$

Oliveira e Suquet [36], 1998, estabelecem o seguinte resultado genérico para a fineza de uma sucessão de elementos aleatórios em $L^p[0,1], p \geq 2$, que pretendemos particularizar para $L^2[0,1]$ e para a sucessão de elementos aleatórios $\{\beta_n(t)\}_{n\in\mathbb{N}}$.

Teorema 1.7.5 (Oliveira, Suquet [36], 1998) *Seja $\{\xi_n\}_{n\in\mathbb{N}}$ uma sucessão de elementos aleatórios em $L^p[0,1]$ tal que:*

(*i*) *para algum $\gamma > 1$,* $\sup_{n\geq 1} \mathbb{E}\|\xi_n\|_1^\gamma < \infty$;

(*ii*) $\lim_{h\to 0} \sup_{n\geq 1} \mathbb{E}\,\|\xi_n(\cdot + h) - \xi_n\|_p^p = 0.$

Então a sucessão $\{\xi_n\}_{n\in\mathbb{N}}$ é fina em $L^p[0,1]$.

Demonstração: A demonstração deste teorema usa resultados clássicos de Análise Funcional, em particular a aproximação uniforme de funções $f \in L^p$, por produtos de convolução $f * g$ onde g é uma densidade de probabilidade.
Consideremos a sucessão de funções aleatórias ξ_n em $L^p[0,1]$ e estabeleça-se que, para cada t fora de $[0,1]$, $\xi_n(t) = 0$. Seja κ uma densidade com suporte $[-1,1]$. Suponha-se, ainda, que κ é Lipschitziana, isto é, $\|\kappa\|_{\text{Lip}} = \sup_{s\neq t} \dfrac{\kappa(t) - \kappa(s)}{|t-s|} < +\infty$ e, para cada inteiro positivo, j, seja $\kappa_j(t) = j\kappa(jt)$. Esta sucessão de densidades $\{\kappa_j\}_{j\in\mathbb{N}}$ é uma aproximação

da identidade, já que, como é fácil verificar, $\kappa_j \circ f \to f$ em $L^p(\mathbb{R})$.
Notando agora que

$$(\kappa_j * \xi_n)(x) - \xi_n(x) = \int_{|t|\leq \frac{1}{j}} (\xi_n(x-t) - \xi_n(x))\, \kappa_j(t)dt,$$

e usando a desigualdade de Jensen relativamente à medida de probabilidade $\kappa_j(t)dt$, obtém-se

$$\begin{aligned}
&\mathbb{E}\left(\int_{\mathbb{R}} |(\kappa_j * \xi_n)(x) - \xi_n(x)|^p\, dx\right) \leq \\
&\quad \leq \mathbb{E}\left(\int_{|t|\leq \frac{1}{j}} \kappa_j(t)dt \int_{\mathbb{R}} |\xi_n(x-t) - \xi_n(t)|^p\, dx\right) \leq \\
&\quad \leq \sup_{|t|\leq \frac{1}{j}} \mathbb{E}\|\xi_n(\cdot) - \xi_n(\cdot + t)\|_p^p.
\end{aligned}$$

Temos, então, por (ii),

$$\lim_{j\to\infty} \sup_{n\geq 1} \mathbb{E}\|(\kappa_j * \xi_n) - \xi_n\|_p^p = 0. \tag{1.26}$$

Temos também, $\mathbb{E}\|\kappa_j * \xi_n\|_\infty \leq \|\kappa_j\|_\infty \sup_{i\geq 1} \mathbb{E}\|\xi_i\|_1$ e, para qualquer $0 \leq s < t \leq 1$ e para algum $\gamma > 1$,

$$\mathbb{E}\,|(\kappa_j * \xi_n)(t) - (\kappa_j * \xi_n)(s)|^\gamma \leq \|\kappa\|_{\text{Lip}}^\gamma \sup_{i\geq 1} \mathbb{E}\|\xi_i\|_1^\gamma |t-s|^\gamma.$$

Nestas condições, para cada j fixo, a sucessão $\{\kappa_j * \xi_n\}_{n\in\mathbb{N}}$ é fina em $C[0,1]$, de acordo com Billingsley [4], 1968, Teorema 12.3, pelo que também é fina em $L^p[0,1]$.
Seja, agora, $\eta > 0$ fixo. Defina-se, para cada $i \geq 1$, $\eta_i = 2^{-i}\eta$ e escolha-se uma sucessão ϵ_i de números positivos convergente para zero. De acordo com a equação (1.26) e a desigualdade de Markov, pode ser extraída uma subsucessão de índices j_i tal que para $n \geq 1$ e $i \geq 1$, se tem

$$P(\|\xi_n - \kappa_{j_i} * \xi_n\|_p > \epsilon_i) > \eta_i. \tag{1.27}$$

Uma vez que a sucessão $\{\kappa_{j_i} * \xi_n\}_{n\in\mathbb{N}}$ é fina em $L^p[0,1]$, existe algum compacto $C_i \in L^p[0,1]$ tal que para $n \geq 1$ e $i \geq 1$, se verifica

$$P(\kappa_{j_i} * \xi_n \notin C_i) < \eta_i. \tag{1.28}$$

Seja $A = \{f \in L^p[0,1] : \kappa_{j_i} * f \in C_i \text{ e } \|\kappa_{j_i} * f - f\|_p \leq \epsilon_i, i \geq 1\}$. A é compacto uma vez que o espaço $L^p[0,1]$ é completo e, para $\delta > 0$ fixo, A pode ser coberto por um número finito de bolas de raio 2δ. Além disso, das desigualdades (1.27) e (1.28) deduz-se que para $n \geq 1$, $P(\xi_n \in A) > 1 - 2\eta$. Logo a sucessão $\{\xi_n\}_{n\in\mathbb{N}}$ é fina em $L^p[0,1]$. ■

Vejamos agora o resultado que estabelece a fineza da sucessão de processos empíricos $\{\beta_n(t), t \in [0,1]\}_{n\in\mathbb{N}}$ vistos como elementos aleatórios de $L^2[0,1]$.

Teorema 1.7.6 *Defina-se para todo* $s, t \in [0,1]$,

$$L_n(s,t) = \frac{1}{n} \sum_{j,k=1}^{n} \left(P\big(X_j \leq s, X_k \leq t\big) - P(X_j \leq s) P(X_k \leq t) \right).$$

Se a sucessão $\{L_n\}_{n\in\mathbb{N}}$ *converge uniformemente sobre* $[0,1]^2$, *então a sucessão* $\{\beta_n\}_{n\in\mathbb{N}}$ *é fina em* $L^2[0,1]$.

Demonstração: Sejam s e t tais que $0 \leq s < t \leq 1$. Comecemos por estabelecer uma fórmula, em termos da sucessão de funções $\{L_n\}_{n\in\mathbb{N}}$, para $\mathbb{E}\,|\beta_n(t) - \beta_n(s)|^2$.

$$\mathbb{E}\,|\beta_n(t) - \beta_n(s)|^2 = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{Cov}\,\left(\mathbb{1}_{(s,t]}(X_i), \mathbb{1}_{(s,t]}(X_j)\right) =$$

$$= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(P(s < X_i \leq t, s < X_j \leq t) - P(s < X_i \leq t)\, P(s < X_j \leq t) \right).$$

Observemos que a função integranda pode ser reescrita do seguinte modo:

$$\begin{aligned} &P(s < X_i \leq t, s < X_j \leq t) - P(s < X_i \leq t)\, P(s < X_j \leq t) = \\ &\quad = P(X_i \leq t, X_j \leq t) - P(X_i \leq t, X_j \leq s) - P(X_i \leq s, X_j \leq t) + \\ &\qquad + P(X_i \leq s, X_j \leq s) - (P(X_i \leq t) - P(X_i \leq s)) \\ &\qquad\quad (P(X_j \leq t) - P(X_j \leq s)) \end{aligned}$$

Assim, de acordo com a definição de L_n, podemos escrever

$$\mathbb{E}\,|\beta_n(t) - \beta_n(s)|^2 = L_n(t,t) + L_n(s,s) - L_n(s,t) - L_n(t,s).$$

Como $\{L_n\}_{n\in\mathbb{N}}$ converge uniformemente para L, digamos, temos, substituindo t por $s+h$, para qualquer $\epsilon > 0$ e com n suficientemente grande,

$$\begin{aligned} &\mathbb{E}\,|\beta_n(s+h) - \beta_n(s)|^2 \leq \\ &\quad \leq (L(s+h, s+h) + L(s,s) - L(s, s+h) - L(s+h, s)) + \epsilon. \end{aligned}$$

Sendo, ainda, $\{L_n\}_{n\in\mathbb{N}}$ uma sucessão de funções contínuas, L é obviamente contínua e, temos

$$\begin{aligned} &\mathbb{E}\,|\beta_n(s+h) - \beta_n(s)|^2 \leq \\ &\quad \leq \sup_{0\leq t\leq 1} 2 \times L(t, t+h) \leq \epsilon,\ s \in [0,1],\ n \geq 1. \end{aligned}$$

A condição (ii) do teorema 1.7.2 é obtida por integração com respeito a s. Quanto à condição (i) do mesmo teorema, obtém-se da majoração atrás obtida, fazendo $t = 0$ e $\gamma = 2$. ■

Nota 1.7.7 *Oliveira e Suquet [35], 1995, provam este resultado para a convergência do processo empírico em* $L^2[0,1]$, *seguindo uma abordagem menos clássica. De facto, usam a condição suficiente de Prokhorov modificada, [47], 1993, para a prova da compacidade relativa da sucessão de elementos aleatórios, usando a base de Haar para* $L^2[0,1]$.

No caso da sucessão de variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$ ser estacionária, a condição suficiente no teorema 1.7.6 é equivalente à convergência uniforme em $[0,1]^2$ de

$$\sum_{n=1}^{\infty}\Big|P\big(X_1\leq s, X_n\leq t\big)-P\big(X_1\leq s\big)P\big(X_n\leq t\big)\Big|. \tag{1.30}$$

De facto, sob estacionaridade da sucessão $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias reais, para todo $s,t\in[0,1]$, tem-se

$$\begin{aligned}L_n(s,t) = {} & \operatorname{Cov}\left(\mathbb{1}_{(0,s]}(X_1),\mathbb{1}_{(0,t]}(X_1)\right)+\\ & +\frac{1}{n}\sum_{i=1}^{n-1}(n-i)\left(\operatorname{Cov}\left(\mathbb{1}_{(0,s]}(X_1),\mathbb{1}_{(0,t]}(X_i)\right)+\right.\\ & \left.+\operatorname{Cov}\left(\mathbb{1}_{(0,t]}(X_1),\mathbb{1}_{(0,s]}(X_i)\right)\right).\end{aligned}$$

Para que (1.30) seja suficiente para a fineza da sucessão de processos $\{\beta_n\}_{n\in\mathbb{N}}$ basta, de acordo com o Teorema 1.7.6, provar que a sucessão $\{L_n\}_{n\in\mathbb{N}}$ é uniformemente convergente.
Considere-se $Q_{jk}(s,t)=\operatorname{Cov}\left(\mathbb{1}_{(0,s]}(X_j),\mathbb{1}_{(0,t]}(X_k)\right)$. A condição do Teorema 1.7.6, sob estacionaridade, requer a convergência uniforme em $[0,1]^2$ de $\sum_{n=1}^{\infty}|Q_{1n}(s,t)|$.
Ora, $Q_{1i}(s,t)=Q_{i1}(t,s)$, pelo que se $\sum_{n=1}^{\infty}|Q_{1n}|$ converge uniformemente sobre $[0,1]^2$, o mesmo também é verdade para $\sum_{n=1}^{\infty}|Q_{n1}|$ e, consequentemente para $\sum_{n=1}^{\infty}|Q_{1n}+Q_{n1}|$. Então

$$\begin{aligned}\left|L_n-Q_{11}-\sum_{i=1}^{n-1}(Q_{1i}+Q_{i1})\right| &= \left|-\sum_{i=1}^{n-1}\frac{i}{n}(Q_{1i}+Q_{i1})\right|\leq\\ &\leq\sum_{i=1}^{n-1}\frac{i}{n}\left|Q_{1i}+Q_{i1}\right|\\ &\leq\sum_{i=1}^{\sqrt{n}-1}\frac{i}{n}\left|Q_{1i}+Q_{i1}\right|+\sum_{i=\sqrt{n}}^{n-1}\frac{i}{n}\left|Q_{1i}+Q_{i1}\right|.\end{aligned}$$

Uma vez que para $i < \sqrt{n}$, $\frac{i}{n} < \frac{1}{\sqrt{n}}$ e para $\sqrt{n} < i < n-1$, $\frac{i}{n} < \frac{n-1}{n}$, tem-se,

$$\begin{aligned}\left|L_n - Q_{11} - \sum_{i=1}^{n-1}(Q_{1i}+Q_{i1})\right| &\leq \frac{1}{\sqrt{n}}\sum_{i=1}^{\sqrt{n}-1}|Q_{1i}+Q_{i1}| + \\ &+ \sum_{i=\sqrt{n}}^{n-1}\frac{i}{n}|Q_{1i}+Q_{i1}| \leq \\ &\leq \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{\infty}|Q_{1i}(s,t)| + \sum_{i=1}^{\infty}|Q_{i1}(s,t)|\right) + \\ &+ \sum_{i=\sqrt{n}}^{\infty}|Q_{1i}(s,t)| + \sum_{i=\sqrt{n}}^{\infty}|Q_{i1}(s,t)| \\ &\leq \frac{2}{\sqrt{n}}\sup_{s,t\in[0,1]}\sum_{i=1}^{+\infty}|Q_{i1}(s,t)| + 2\sup_{s,t\in[0,1]}\sum_{i=\sqrt{n}}^{+\infty}|Q_{i1}(s,t)|\,.\end{aligned}$$

Ora a função $\sum_{i=1}^{+\infty}|Q_{i1}(s,t)|$ é contínua por ser o limite uniforme de uma sucessão de funções contínuas. Como está definida num compacto, é limitada. Consequentemente,

$$\frac{2}{\sqrt{n}}\sum_{i=1}^{\infty}|Q_{i1}(s,t)| \underset{n\to\infty}{\longrightarrow} 0.$$

Por outro lado $\sum_{i=\sqrt{n}}^{+\infty}|Q_{i1}(s,t)| \to 0$, quando $n \to \infty$, por ser o resto de uma série convergente. Nestas condições, para todo $\epsilon > 0$, existe $n_0 \in \mathbb{N}$, tal que para $n \geq n_0$, se tem

$$\left|L_n(s,t) - Q_{11}(s,t) - \sum_{i=1}^{n-1}(Q_{1i}(s,t)+Q_{i1}(s,t)\right| \leq \epsilon,$$

para todo $s,t \in [0,1]$.
Obtém-se finalmente, para cada $\epsilon > 0$,

$$\lim_{n\to\infty}\sup_{s,t\in[0,1]}|L_n(s,t)| \leq |Q_{11}(s,t)| \lim_{n\to\infty}\sup_{s,t\in[0,1]}\sum_{i=1}^{\infty}|Q_{i1}(s,t)| + \epsilon,$$

o que completa a demonstração da convergência uniforme de $\{L_n\}_{n\in\mathbb{N}}$.

Com a fineza da sucessão de processos empíricos estabelecida, para provar a convergência em $L^2[0,1]$ de $\{\beta_n\}_{n\in\mathbb{N}}$ para o processo $\mathbb{B}$, basta verificar a convergência em distribuição de $\displaystyle\int_0^1 g(t)\beta_n(t)dt$ para $\displaystyle\int_0^1 g(t)\mathbb{B}(t)dt$, para cada $g \in L^2[0,1]$.

Uma vez que

$$\int_0^1 g(t)\beta_n(t)dt = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\int_0^1 g(t)\mathbb{1}_{(0,t]}(X_i)dt - \int_0^1 g(t)\mathbb{E}\left(\mathbb{1}_{(0,t]}(X_i)\right)dt\right),$$

então,

$$\int_0^1 g(t)\beta_n(t)dt = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\int_{X_i}^1 g(t)dt - \mathbb{E}\int_{X_i}^1 g(t)dt\right).$$

De facto, com as condições do Teorema 1.7.6 esta convergência é estabelecida graças a um resultado de Newman [31], 1984, que é um Teorema do Limite Central para as imagens de variáveis associadas por funções absolutamente contínuas.

Teorema 1.7.8 (Newman [31], 1984) *Seja $(Y_n)_{n\in\mathbb{N}}$ uma sucessão de variáveis aleatórias associadas, estritamente estacionária com função de distribuição marginal F. Seja $X_n = G(Y_n)$ onde G é uma função absolutamente contínua.*
Seja ainda $\bar{X}_n = \bar{G}(Y_n)$ onde $\bar{G}(t) = \displaystyle\int_{[0,t]} \left|G'(u)\right| du$.
Defina-se

$$\begin{aligned}\Gamma(x,y) = {} & F(\min\{x,y\}) - F(x)F(y) + \\ & + \sum_{k=2}^{+\infty}\big(P(Y_1 \le x, Y_k \le y) - P(Y_1 \le x)P(Y_k \le y)\big) + \\ & + \sum_{k=2}^{+\infty}\big(P(Y_1 \le y, Y_k \le x) - P(Y_1 \le y)P(Y_k \le x)\big).\end{aligned} \tag{1.31}$$

Se

$$\int_{[0,1]^2} G'(x)G'(y)\Gamma(x,y)\,dxdy < \infty,$$

então $\bar{X}_1$ *é de quadrado integrável e*

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\big[X_j - \mathbb{E}(X_j)\big] = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\big[G\big(Y_j\big) - \mathbb{E}\big(G\big(Y_j\big)\big)\big]$$

converge em distribuição para uma variável aleatória com distribuição normal de média 0 e variância

$$\sigma^2 = \int_{[0,1]^2} G'(x)G'(y)\Gamma(x,y)\,dxdy.$$

Este teorema permite demonstrar o resultado pretendido.

Teorema 1.7.9 *Seja* $\{X_n\}_{n\in\mathbb{N}}$ *uma sucessão de variáveis aleatórias associadas, estritamente estacionária, com distribuição contínua em* $[0,1]$. *Suponhamos que a série em (1.31) converge uniformemente em* $[0,1]^2$. *Então o processo empírico converge fracamente em* $L^2[0,1]$ *para um processo gaussiano centrado com função de covariância dada por (1.31).*

Demonstração: Em primeiro lugar refira-se que a compacidade relativa do processo empírico decorre da convergência uniforme da série $\sum_{k=2}^{+\infty} |P(X_1 \le y, X_k \le x) - P(X_1 \le y)P(X_k \le x)|$, que mais não é que a série envolvida em (1.31) uma vez que, por associação da sucessão $\{X_n\}_{n\in\mathbb{N}}$, o termo geral desta série é não negativo.
Aplique-se, agora, o Teorema 1.7.8 com $X_n = G(Y_n) = \int_{[0,X_n]} g(u)du$. Evidentemente que a função G é absolutamente contínua em $[0,1]$. Vejamos que a condição

$$\int_{[0,1]^2} G'(x)G'(y)\Gamma(x,y)\,dxdy < \infty$$

é verificada.
A função Γ interveniente nesta condição é contínua, por ser a soma de

funções contínuas, sendo, obviamente, limitada em $[0,1]^2$. Além disso, como $G' \in L^2[0,1]$ temos,

$$\int_{[0,1]^2} G'(x)G'(y)\Gamma(x,y)\,dxdy \leq c \int_{[0,1]^2} G'(x)G'(y)dxdy < \infty. \blacksquare$$

Capítulo 2

A função de distribuição bivariada

Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão associada de variáveis aleatórias identicamente distribuídas, com função de distribuição marginal F desconhecida, definidas num espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$.
Consideremos a amostra aleatória $\widetilde{X} = (X_1, \ldots, X_n)$. Se a sucessão $\{X_n\}_{n\in\mathbb{N}}$ for estritamente estacionária, a função de distribuição conjunta do par (X_i, X_{k+i}), com $k = 1, \ldots, n-1$, fixo é, para cada $i = 1, \ldots, n-k$, a função de distribuição F_k de (X_1, X_{k+1}). Neste capítulo iremos estudar um estimador para a função de distribuição conjunta F_k, com k fixo.
No âmbito deste trabalho, o interesse em aproximar funções de distribuição de pares aleatórios surge das caracterizações da distribuição limite do processo empírico. De facto, o termo geral da série que aparece na especificação da função de covariância do processo limite do processo

empírico, equação (1.31) do capítulo anterior,

$$\begin{aligned}\Gamma(s,t) = F(\min\{s,t\}) - F(s)F(t) + \\ + \sum_{k=2}^{+\infty} (P(X_1 \leq s, X_k \leq t) - P(X_1 \leq s)P(X_k \leq t)) + \\ + \sum_{k=2}^{+\infty} (P(X_1 \leq t, X_k \leq s) - P(X_1 \leq t)P(X_k \leq s))\end{aligned}$$

é, precisamente, a diferença entre a função de distribuição bidimensional do par (X_1, X_k) e o produto das respectivas funções de distribuição marginais. No entanto, além desta aplicação, temos no âmbito da inferência estatística a necessidade de conhecer o comportamento de estatísticas de teste, que envolvam estimadores deste tipo, para o conhecimento dos limites críticos dos mesmos e até mesmo para o estabelecimento de intervalos de confiança.

Na primeira secção deste capítulo definimos o estimador para F_k que iremos usar em todo este trabalho e estabelecemos algumas condições de regularidade relativas à sucessão $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias de base.

A segunda secção é dedicada ao estudo do estimador, nomeadamente no que diz respeito à consistência. Provamos que este é assimptoticamente centrado, que converge pontual e uniformemente. Em todos os casos em que se prova a convergência quase certa, é obtida, ainda, a velocidade desta convergência.

Com o objectivo de obtermos um majorante para a velocidade da convergência forte do estimador, recorremos a uma desigualdade exponencial do tipo Hoeffding-Bernestein apresentada por Ioannides e Roussas [18], 1999, que, juntamente com uma desigualdade do tipo Kolmogorov, provada por Newman e Wright [32], 1981, incluída no Capítulo I, Teorema 1.5.1, torna possível uma majoração desta velocidade.

Na terceira secção estudamos o comportamento do Erro Quadrático Médio do estimador, obtendo, por minimização da sua expressão, a ordem

de convergência óptima da sucessão de comprimentos $\{h_n\}_{n\in\mathbb{N}}$ que intervêm na definição do estimador. Além disso, com uma condição um pouco mais forte na sucessão de covariâncias entre as variáveis de base, obtemos a mesma ordem de convergência que se obtém sob independência, tal como Cai e Roussas [8], 1998, haviam obtido para $p = 1$, sendo este resultado generalizado para estimadores do tipo núcleo p - dimensionais com $p > 2$.
A quarta e última secção é dedicada ao estudo do comportamento assimptótico das leis de dimensão finita.

2.1 Definições e Hipóteses Prévias.

Um estimador natural para F_k é a função de distribuição empírica bivariada

$$\widehat{\varphi}_{n,k}(x,y) = \frac{1}{n-k}\sum_{i=1}^{n-k} \mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_i, X_{k+i}); \qquad (2.1)$$

onde $\mathbb{1}_A(x,y)$ representa a função indicatriz do conjunto A em (x,y). O comportamento deste estimador foi já estudado por Henriques e Oliveira [17], *em publicação*.
Neste trabalho vamos considerar um estimador no qual se consideram funções de distribuição contínuas, conhecidas, das variáveis aleatórias que constituem a amostra, as funções núcleo, em vez de se considerarem funções indicatrizes das mesmas. Este estimador é, portanto, uma função contínua.

Definição 2.1.1 *Seja U uma função de distribuição conhecida e $\{h_n\}_{n\in\mathbb{N}}$ uma sucessão de números reais positivos que converge para 0. A função $\widehat{F}_{n,k}$ definida para cada $(x,y) \in \mathbb{R}^2$ por*

$$\begin{aligned}\widehat{F}_{n,k}(x,y) &= \frac{1}{n-k}\sum_{i=1}^{n-k} U\Big(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\Big) = \\ &= \int_{\mathbb{R}^2} U\Big(\frac{x-s}{h_n}, \frac{y-t}{h_n}\Big)\, d\widehat{\varphi}_{n,k}(s,t)\end{aligned}$$

é um estimador para a função de distribuição F_k do par aleatório (X_1, X_{k+1}).

Com o propósito de estudarmos este estimador, vamos apresentar algumas condições que serão utilizadas ao longo deste capítulo relativas à sucessão associada $\{X_n\}_{n\in\mathbb{N}}$ de variáveis reais, à função núcleo U e à sucessão $\{h_n\}_{n\in\mathbb{N}}$.

Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão associada, estritamente estacionária de variáveis aleatórias reais que verificam as seguintes condições:

(C_1) As variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$ são identicamente distribuídas com função distribuição marginal desconhecida F, contínua e função densidade de probabilidade f limitada;

(C_2) A função densidade de probabilidade f admite derivada contínua e limitada;

(C_3) Para cada inteiro fixo k, F_k é a função de distribuição conjunta do par (X_1, X_{k+1}); F_k tem derivadas parciais de primeira e segunda ordens uniformemente limitadas e contínuas;

(C_4) Para cada inteiro positivo j, $F_{k,j}$ é a função de distribuição conjunta de $(X_1, X_{k+1}, X_j, X_{k+j})$; $F_{k,j}$ tem, para cada j derivadas parciais de primeira e segunda ordens uniformemente limitadas e contínuas;

(C_5) A função núcleo U é duas vezes diferenciável. A derivada de segunda ordem de U, $u = \dfrac{\partial^2 U}{\partial x \partial y}$ verifica as seguintes condições:

$$\int_{\mathbb{R}^2} w\, u(w,v)\, dw\, dv = \int_{\mathbb{R}^2} v\, u(w,v)\, dw\, dv = 0,$$

$$\int_{\mathbb{R}^2} w^2 u(w,v) dw\, dv < \infty \text{ e } \int_{\mathbb{R}^2} v^2 u(w,v) dw\, dv < \infty;$$

(C_6) A sucessão $\{h_n\}_{n\in\mathbb{N}}$ é tal que $n\,h_n^2 \to 0$;

(C_7) $\displaystyle\sum_{n=1}^{\infty} n\,Cov^{1/3}(X_1, X_n) < \infty$;

$(C_7)'$ $\displaystyle\sum_{n=1}^{\infty} Cov^{1/3}(X_1, X_n) < \infty$;

(C_8) A função $V = \dfrac{\partial^2 \mathrm{U}^2}{\partial \mathrm{x} \partial \mathrm{y}}$ satisfaz

$$\int_{\mathbb{R}^2} w^2 V(w,v)dwdv < \infty \qquad \text{e} \qquad \int_{\mathbb{R}^2} v^2 V(w,v)dwdv < \infty.$$

As condições $(C_1), (C_2), (C_3), (C_5)$ e (C_7) foram usadas em Cai, Roussas [8], 1998, para o tratamento do caso univariado. As condições $(C_1), (C_2)$ e (C_5) são usadas para a prova da convergência da média, com especificação de velocidades de convergência do estimador, para a função de distribuição marginal das variáveis aleatórias de base, enquanto que a condição (C_7) é usada especificamente para a prova da sua consistência. A condição $(C_7)'$ foi usada por Oliveira e Suquet [35, 36], 1995, 1998, na prova da convergência fraca do processo empírico em $L^2[0,1]$.

2.2 Consistência do Estimador

Nesta secção vamos, em primeiro lugar, provar que $\widehat{F}_{n,k}$ é assimptoticamente centrado, caracterizando ainda a velocidade de convergência de $\mathbb{E}\left(\widehat{F}_{n,k}(x,y)\right)$. Seguidamente provamos a convergência quase certa do estimador, aplicando uma lei forte dos grandes números às variáveis aleatórias

$$U\Big(\frac{x - X_i}{h_n}, \frac{y - X_{k+i}}{h_n}\Big), \quad \text{para} \quad i = 1, ..., n-k.$$

Com o objectivo de especificar a velocidade desta convergência, impomos um decrescimento geométrico nas covariâncias entre as variáveis aleatórias X_i, $i = 1, \ldots, n$, e aplicamos a desigualdade exponencial de Ioannides

e Roussas [18], 1999, juntamente com uma desigualdade de Newman e Wright [32], 1981, Teorema 1.5.1, como já referimos anteriormente.

Será ainda provada a convergência uniforme do estimador, também com especificação de um majorante para a velocidade desta convergência.

Teorema 2.2.1 *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão associada de variáveis aleatórias reais identicamente distribuídas. Seja F_k a função de distribuição conjunta do par (X_1, X_{k+1}) com k fixo. Então, para cada $x, y \in \mathbb{R}$,*

$$\mathbb{E}\left(\widehat{F}_{n,k}(x,y)\right) \underset{n\to\infty}{\longrightarrow} F_k(x,y).$$

Demonstração: De acordo com a definição de $\widehat{F}_{n,k}$,

$$\mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big) = \mathbb{E}\Big(\int_{\mathbb{R}^2} U\big(\frac{x-s}{h_n}, \frac{y-t}{h_n}\big)\, d\,\widehat{\varphi}_{n,k}(s,t)\Big).$$

Uma vez que $\mathbb{E}\left(\hat{\varphi}_{n,k}(x,y)\right) = F_k(x,y)$, e tendo a função integranda esperança finita, por aplicação do Teorema da Convergência Dominada, vem

$$\mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big) = \int_{\mathbb{R}^2} U\left(\frac{x-s}{h_n}, \frac{y-t}{h_n}\right)\, d\,F_k(s,t).$$

Como

$$U(x,y) = \int_{(-\infty,x]} \int_{(-\infty,y]} u(w,v)dwdv,$$

podemos escrever

$$\mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big) = \int_{\mathbb{R}^2} \left(\int_{(-\infty, \frac{x-s}{h_n}]} \int_{(-\infty, \frac{y-t}{h_n}]} u(w,v)dwdv \right) dF_k(s,t).$$

Aplicando novamente o Teorema de Fubini, dado que u é integrável relativamente à medida de Lesbegue em $\mathbb{R}^2$, e invertendo a ordem de integração, obtém-se

$$\mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big) = \int_{\mathbb{R}^2} u(w,v) F_k(x - wh_n, y - vh_n)dvdw. \qquad (2.4)$$

Aplicando agora o Teorema da Convergência Dominada segue-se imediatamente que

$$\mathbb{E}\left(\widehat{F}_{n,k}(x,y)\right) \underset{n\to\infty}{\longrightarrow} F_k(x,y).\blacksquare$$

Vejamos agora a velocidade desta convergência. Para isso desenvolva-se $F_k(x-wh_n, y-vh_n)$ em polinómio de Taylor, até à segunda ordem, em torno de (x,y),

$$\begin{aligned} F_k(x-wh_n, y-vh_n) &= F_k(x,y) - wh_n\frac{\partial F_k}{\partial x}(x,y) - \\ &-vh_n\frac{\partial F_k}{\partial y}(x,y) + \frac{{h_n}^2}{2}\Big(w^2\frac{\partial^2 F_k}{\partial x^2}((x,y)-\theta h_n(w,v)) + \\ &+v^2\frac{\partial^2 F_k}{\partial y^2}\big((x,y)-\theta h_n(w,v)\big) + \\ &+wv\frac{\partial^2 F_k}{\partial x\partial y}\big((x,y)-\theta\, h_n(w,v)\big)\Big),\ 0\le\theta\le 1. \end{aligned}$$

Substituindo esta expressão em (2.4), usando o Teorema da Convergência Dominada e as condições (C_3) e (C_5) , obtemos,

$$\begin{aligned} \mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big) &= F_k(x,y) + \\ &+\frac{h_n^2}{2}\Big(\frac{\partial^2 F_k}{\partial x^2}(x,y)\int_{\mathbb{R}^2} s^2 u(s,t)\,ds\,dt + \\ &+\frac{\partial^2 F_k}{\partial x\partial y}(x,y)\int_{\mathbb{R}^2} s\,t\,u(s,t)\,ds\,dt + \\ &+\frac{\partial^2 F_k}{\partial y^2}(x,y)\int_{\mathbb{R}^2} t^2 u(s,t)\,ds\,dt\Big) \quad + \quad o(h_n^2). \end{aligned}$$

O resultado para a velocidade de convergência a seguir enunciado é, agora, imediato.

Teorema 2.2.2 *Suponhamos que a sucessão $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias reais satisfaz as condições $(C_1), (C_3)$ e (C_5). Então para cada $x, y \in \mathbb{R}$,*

$$\mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big) = F_k(x,y) + O(h_n^2).$$

O próximo objectivo é estabelecer a convergência quase certa de $\widehat{F}_{n,k}$. Para tal aplicamos a lei forte dos grandes números provada por Newman [31], 1984, apresentada no capítulo anterior na secção 1.6, às funções de variáveis aleatórias

$$U\Big(\frac{x-X_j}{h_n},\frac{y-X_{k+j}}{h_n}\Big),\ j=1,\ldots,n-k,$$

que, por serem funções não decrescentes de variáveis associadas, são também associadas (Propriedade 1.1.8).

Em primeiro lugar vamos relacionar as covariâncias entre estas variáveis aleatórias e as covariâncias entre as funções indicatrizes das variáveis aleatórias iniciais.

Lema 2.2.3 *Suponhamos que a sucessão $\{X_n\}_{n\in\mathbb{N}}$, de variáveis aleatórias reais satisfaz as condições (C_1), (C_3), (C_4) e (C_5). Então, para cada $j\geq 1$ e para cada $x,y\in\mathbb{R}$,*

$$\begin{aligned}\operatorname{Cov}&\left(U\left(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}\right),U\left(\frac{x-X_j}{h_n},\frac{y-X_{k+j}}{h_n}\right)\right)=\\&=\operatorname{Cov}\left\{\mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_1,X_{k+1}),\mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_j,X_{k+j})\right\}+\\&\quad+O({h_n}^2)=\\&=F_{k,j}(x,y,x,y)-F_k^2(x,y)+O(h_n^2).\end{aligned}$$

Demonstração: Escrevam-se as covariâncias como

$$\begin{aligned}\operatorname{Cov}&\Big(U(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}),U(\frac{x-X_j}{h_n},\frac{y-X_{k+j}}{h_n})\Big)=\\&=\int_{\mathbb{R}^4}U(\frac{x-s}{h_n},\frac{y-t}{h_n})U(\frac{x-w}{h_n},\frac{y-v}{h_n})dF_{k,j}(s,t,w,v)-\\&\quad-\left(\int_{\mathbb{R}^2}U(\frac{x-s}{h_n},\frac{y-t}{h_n})dF_k(s,t)\right)^2.\end{aligned}$$

A segunda parcela do segundo membro desta equação é $\Big(\mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big)\Big)^2$ e o seu comportamento está descrito no Teorema 2.2.2. Quanto à pri-

meira parcela, uma vez que,

$$U(x,y) = \int_{(-\infty,x]} \int_{(-\infty,y]} u(w,v) dw dv,$$

usando o Teorema de Fubini e invertendo a ordem de integração tem-se,

$$\int_{\mathbb{R}^4} U\big(\frac{x-s}{h_n}, \frac{y-t}{h_n}\big) U\big(\frac{x-w}{h_n}, \frac{y-v}{h_n}\big)\, dF_{k,j}(s,t,w,v) =$$
$$= \int_{\mathbb{R}^4} u(a,b)\, u(c,d)\, F_{k,j}(x-ah_n, y-bh_n, x-ch_n, y-dh_n)\, dd\, dc\, db\, da.$$

Desenvolvendo $F_{k,j}(x-ah_n, y-bh_n, x-ch_n, y-dh_n)$ em polinómio de Taylor, de segunda ordem, em torno de (x,y,x,y), e usando as condições (C_4) e (C_5), este integral é igual a $F_{k,j}(x,y,x,y) + O(h_n^2)$.
Como $\mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big) = F_k(x,y) + O(h_n^2)$, obtemos o resultado pretendido.

■

Vejamos agora como majorar o resultado do lema anterior em termos de covariâncias entre as variáveis aleatórias X_i, pertencentes à sucessão associada $\{X_n\}_{n\in\mathbb{N}}$.

Lema 2.2.4 *Suponhamos que o vector (X_1, X_2, X_3, X_4) é associado com funções densidade de probabilidade marginais limitadas. Então,*

$$\begin{aligned}\mathrm{Cov}\Big(\mathbb{1}_{(-\infty,x_1]\times(-\infty,x_2]}(X_1,X_2), \mathbb{1}_{(-\infty,x_3]\times(-\infty,x_4]}(X_3,X_4)\Big) \leq \\ \leq M\,\big(\mathrm{Cov}^{\,1/3}(X_1,X_3) + \mathrm{Cov}^{\,1/3}(X_1,X_4) + \\ + \mathrm{Cov}^{\,1/3}(X_2,X_3) + \mathrm{Cov}^{\,1/3}(X_2,X_4)\big),\end{aligned}$$

onde M é a constante interveniente na desigualdade (1.25).

Demonstração: Aplicando o Lema de Lebowitz, Teorema 1.2.3, vem,

$$\begin{aligned}&\mathrm{Cov}\Big(\mathbb{1}_{(-\infty,x_1]\times(-\infty,x_2]}(X_1,X_2), \mathbb{1}_{(-\infty,x_3]\times(-\infty,x_4]}(X_3,X_4)\Big) \leq \\ &\leq H_{1,3} + H_{1,4} + H_{2,3} + H_{2,4},\end{aligned}$$

onde

$$H_{i,j} = P(X_i > x_i, X_j > x_j) - P(X_i > x_i)P(X_j > x_j) =$$
$$= P(X_i \leq x_i, X_j \leq x_j) - P(X_i \leq x_i)P(X_j \leq x_j).$$

Ora, de acordo com a desigualdade (1.25), $H_{i,j} \leq M\mathrm{Cov}^{1/3}(X_i, X_j)$, pelo que o resultado pretendido é imediato. ■

Estamos em condições de provar a convergência quase certa do estimador $\widehat{F}_{n,k}$.

Teorema 2.2.5 *Suponhamos que a sucessão $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias reais, satisfaz as condições (C_1), (C_3), (C_4), (C_5) e $(C_7)'$. Então, para todo $x, y \in \mathbb{R}$,*

$$\widehat{F}_{n,k}(x,y) \underset{q.c.}{\longrightarrow} F_k(x,y).$$

Demonstração: No Teorema 2.2.2 foi já provado que $\mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big) \underset{n\to\infty}{\longrightarrow} F_k(x,y)$, pelo que basta agora provar que

$$\widehat{F}_{n,k}(x,y) - \mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big) \underset{q.c.}{\longrightarrow} 0.$$

Para tal iremos usar a Lei Forte dos Grandes Números de Newman, [31], 1984, que garante esta convergência desde que se tenha

$$\frac{1}{n-k}\sum_{i=1}^{n-k} \mathrm{Cov}\ \left(U\Big(\frac{x-X_1}{h_n}, \frac{y-X_{1+k}}{h_n}\Big), U\Big(\frac{x-X_i}{h_n}, \frac{y-X_{i+k}}{h_n}\Big)\right) \underset{n\to\infty}{\longrightarrow} 0, \tag{2.9}$$

uma vez que a sucessão de variáveis aleatórias $\left\{U\left(\frac{x-X_m}{h_n}, \frac{x-X_{m+k}}{h_n}\right)\right\}_{m\geq 1}$ é estacionária e associada (Propriedade 1.1.8).

Como vimos no Lema 2.2.3,

$$\mathrm{Cov}\ \left(U\Big(\frac{x-X_1}{h_n}, \frac{y-X_{k+1}}{h_n}\Big), U\Big(\frac{x-X_j}{h_n}, \frac{y-X_{k+j}}{h_n}\Big)\right) =$$
$$= F_{k,j}(x,y,x,y) - F_k^2(x,y) + O(h_n^2),$$

onde, de acordo com o lema 2.2.4,

$$\begin{aligned} F_{k,j}(x,y,x,y) - F_k^2(x,y) = \\ = \mathrm{Cov}\left(\mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_1, X_{k+1}), \mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_j, X_{k+j})\right) \leq \\ \leq M\big(2 \times \mathrm{Cov}^{1/3}(X_1, X_j) + \mathrm{Cov}^{1/3}(X_1, X_{k+j}) + \mathrm{Cov}^{1/3}(X_{k+1}, X_j)\big). \end{aligned}$$

Uma vez que $\sum_{j=1}^{\infty} \mathrm{Cov}^{1/3}(X_1, X_{j+1}) < \infty$, a condição (2.9) verifica-se imediatamente. Assim, de acordo com a Lei Forte dos Grandes Números de Newman,

$$\frac{1}{n-k}\sum_{i=1}^{n-k}\left[U\Big(\frac{x-X_i}{h_n}, \frac{y-X_{i+k}}{h_n}\Big) - \mathbb{E}\left(U\Big(\frac{x-X_i}{h_n}, \frac{y-X_{i+k}}{h_n}\Big)\right)\right] \underset{n\to\infty}{\longrightarrow} 0.$$

Ou seja,

$$\widehat{F}_{n,k}(x,y) - \mathbb{E}\Big(\widehat{F}_{n,k}(x,y)\Big) \underset{q.c.}{\longrightarrow} 0. \quad \blacksquare$$

Com o mesmo conjunto de hipóteses do Teorema anterior podemos ainda provar a convergência uniforme do estimador.

Teorema 2.2.6 *Suponhamos que a sucessão* $\{X_n\}_{n\in\mathbb{N}}$ *satisfaz as condições* $(C_1), (C_3), (C_4), (C_5)$ *e* $(C_7)'$. *Então,*

$$\sup_{x,y\in\mathbb{R}}\Big|\widehat{F}_{n,k}(x,y) - F_k(x,y)\Big| \underset{q.c.}{\longrightarrow} 0.$$

Demonstração: No Teorema anterior provámos que, para $x, y \in \mathbb{R}$, fixos,

$$\Big|\widehat{F}_{n,k}(x,y) - F_k(x,y)\Big| \underset{q.c.}{\longrightarrow} 0.$$

Pretende-se agora provar que o supremo desta diferença, quando (x,y) varia em $\mathbb{R}^2$, ainda converge para 0.

Consideremos a decomposição usual de $\mathbb{R}^2$ num determinado conjunto de pontos previamente escolhidos. O objectivo é controlar as diferenças

entre a função em estudo nos pontos sobre o reticulado obtido e nos pontos que ficam fora desse reticulado.

Seja $A \geq 1$ um inteiro fixo, $0 \leq i, j \leq A$ e Q a função quantil definida para cada $t \in \mathbb{R}$, por $Q(t) = \inf\{x : F(x) \geq t\}$. Consideremos, para cada $i \in \{1, \ldots, A\}$, $x_{A,i} = Q\left(\frac{i}{A}\right)$. Defina-se, para cada $n \in \mathbb{N}$,

$$\Delta_{A,n} = \max_{0 \leq i,j \leq A} \left| \widehat{F}_{n,k}(x_{A,i}, x_{A,j}) - F_k(x_{A,i}, x_{A,j}) \right|,$$

e

$$\Delta_n = \sup_{x,y \in \mathbb{R}} \left| \widehat{F}_{n,k}(x,y) - F_k(x,y) \right|.$$

Sabemos, pelo Teorema anterior, que $\Delta_{A,n} \underset{q.c.}{\longrightarrow} 0$ quando $n \to \infty$. Além disso, para $i, j \in \{1, \ldots, A\}$, atendendo à escolha dos pontos $x_{A,i}$ efectuada, seque-se imediatamente que

$$|F_k(x_{A,i}, x_{A,j}) - F_k(x_{A,i-1}, x_{A,j-1})| \leq \frac{2}{A}.$$

Consideremos agora x, y tal que, para i, j fixos,

$$x_{A,i-1} \leq x \leq x_{A,i} \qquad \text{e} \qquad x_{A,j-1} \leq y \leq x_{A,j}.$$

Uma vez que $\widehat{F}_{n,k}$ é não decrescente relativamente a cada variável, $\widehat{F}_{n,k}(x,y) \leq \widehat{F}_{n,k}\left(x_{A,i}, x_{A,j}\right)$ e, além disso,

$$\widehat{F}_{n,k}\Big(x_{A,i}, x_{A,j}\Big) \leq \Delta_{A,n} + \frac{2}{A} + F_k\Big(x_{A,i-1}, x_{A,j-1}\Big) \leq \Delta_{A,n} + \frac{2}{A} + F_k(x,y).$$

Por outro lado, e usando o mesmo tipo de argumentos,

$$\widehat{F}_{n,k}(x,y) \geq -\Delta_{A,n} - \frac{2}{A} + F_k(x,y).$$

Temos então,

$$\left| \widehat{F}_{n,k}(x,y) - F_k(x,y) \right| \leq \Delta_{A,n} + \frac{2}{A}.$$

Uma vez que A é arbitrário e $\Delta_{A,n} \underset{q.c.}{\longrightarrow} 0$, então $\Delta_n \underset{q.c.}{\longrightarrow} 0$. ■

Para que sejam especificadas as velocidades destas convergências, vamos apresentar na subsecção seguinte uma nota sobre ordens de convergência na Lei Forte dos Grandes Números para variáveis associadas.

2.2.1 Uma nota sobre velocidades de convergência.

Suponhamos que a sucessão associada $\{X_n\}_{n\in\mathbb{N}}$, é constituída por variáveis aleatórias reais com média zero e identicamente distribuídas. A Lei Forte dos Grandes Números de Newman, no caso da sucessão de variáveis aleatórias ser estritamente estacionária, ou de Birkel, caso em que a sucessão de variáveis aleatórias não necessita ser estritamente estacionária, estabelece que, sob algumas condições para as covariâncias entre as variáveis aleatórias X_i, (vidé *Preliminares*), $\frac{X_1+\cdots+X_n}{n} \to 0$, *quase certamente,* quando $n \to \infty$, ou seja, $|\sum_{i=1}^{n} X_i| = o(n)$ *q.c.*. É natural, tal como no caso independente, perguntar qual a velocidade desta convergência.

Recentemente, Ioannides e Roussas [18], 1999, provaram uma desigualdade exponencial do tipo Hoeffding - Bernstein para variáveis aleatórias associadas que iremos usar para obter um majorante para a velocidade desta convergência. A identificação da ordem exacta de convergência implica um conhecimento de um minorante para aquela velocidade. Ora, para tal torna-se necessário conhecer uma versão do segundo Lema de Borel-Cantelli que, tanto quanto sabemos, ainda não existe na literatura para variáveis associadas.

Para que a desigualdade de Ioannides e Roussas possa ser aplicada, as variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$ devem obedecer às seguintes condições:

(A_1) As variáveis aleatórias são associadas e estacionárias de segunda ordem;

(A_2) As variáveis aleatórias são limitadas, $\big|X_i\big| \leq \dfrac{M_1}{2}, i \geq 1$;

(A_3) A sucessão $\{Cov(X_1, X_{k+1})\}_{k\in\mathbb{N}}$ é não crescente.

Antes de enunciar a desigualdade exponencial de Ioannides e Roussas, introduzimos alguma notação que está directamente relacionada com o enunciado do teorema e com a sua demonstração.

O conjunto de índices das variáveis aleatórias, $\{1, 2, \ldots, n\}$ é dividido em $2r$ grupos sucessivos de p elementos. Tanto o valor de p, como o valor de r dependem de n. Em rigor, deveríamos escrever $p(n)$ e $r(n)$. Não o faremos para não tornar as notações demasiado pesadas. Além disso, r e p tendem para infinito quando n tende para infinito. Seja, então, $r = r(n)$ o maior inteiro tal que:

$$0 < r < n,\ r \to \infty,\ 2pr \leq n, \tag{2.14}$$

ou seja, p e r estão relacionados de modo que $\dfrac{n}{2pr} \to 1$.
Temos, como já referimos, o conjunto $\{1, 2, \ldots, n\}$ partido em $2r$ grupos, cada um com p elementos; o conjunto com os restantes $n - 2pr < p$ elementos pode ser vazio.

Teorema 2.2.7 (Ioannides, Roussas [18], 1999) *Sejam* $\bar{S}_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}X_i)$ *e* ϵ_n *uma sucessão convergente para zero. Consideremos que se verificam as condições* $(A_1), (A_2)$ *e* (A_3). *Suponhamos ainda que as variáveis aleatórias* X_n *obedecem à seguinte relação, com* r *e* p *definidos anteriormente,*

$$\mathrm{Cov}\,(X_1, X_{p+1}) \leq \exp\Big[-\frac{4(M_1+1)}{3M_1^2} r\epsilon_n\Big].$$

Então, para n *suficientemente grande,*

$$P\Big(|\bar{S}_n| \geq \epsilon_n\Big) \leq C_0 \exp(-cr{\epsilon_n}^2),$$

com C_0 *constante e* $c = \dfrac{2M_1^2}{9}$.

A demonstração deste teorema tem como base a partição de $\bar{S}_n$ em três blocos de somas normalizadas de variáveis cujos índices distam entre si pelo menos p unidades. Esta estratégia tem como objectivo controlar as covariâncias entre as variáveis aleatórias de cada bloco, uma vez que é provada uma desigualdade exponencial para cada bloco. De facto,

definindo $Y_i = X_i - \mathbb{E}X_i$, para cada i, $|Y_i| \leq M_1$ e as variáveis Y_i são positivamente associadas. Agora, à custa destas variáveis centradas, definem-se as variáveis U_i e V_i, para cada $i = 1, \ldots, r$, e W_n por:

$$\begin{aligned} U_i &= Y_{2(i-1)p+1} + \cdots + Y_{(2i-1)p}, \\ V_i &= Y_{(2i-1)p+1} + \cdots + Y_{2ip}, \\ W_n &= Y_{2pr+1} + \cdots + Y_n. \end{aligned}$$

Sendo $\bar{U}_n, \bar{V}_n$ e $\bar{W}_n$ dados por

$$\bar{U}_n = \frac{1}{n}\sum_{i=1}^{r} U_i, \;\; \bar{V}_n = \frac{1}{n}\sum_{i=1}^{r} V_i, \;\; \bar{W}_n = \frac{W_n}{n}, \tag{2.16}$$

tem-se $\bar{S}_n = \bar{U}_n + \bar{V}_n + \bar{W}_n$.
É provada uma desigualdade exponencial para $\bar{U}_n$ definida em (2.16) válida também para $\bar{V}_n$, uma vez que as suas construções são idênticas. Estas desigualdades constituem a base da demonstração do Teorema 2.2.7.

Lema 2.2.8 (Ioannides e Roussas [18], 1999) *Seja $\epsilon_n > 0$ e suponhamos que a sucessão $\{X_n\}_{n\in\mathbb{N}}$ verifica as condições $(A_1), (A_2)$ e (A_3). Então para uma determinada constante C_1,*

$$P\Big(\bar{U}_n \geq \epsilon_n\Big) \leq C_1 \exp(-\frac{2}{M_1^2} r {\epsilon_n}^2),$$

desde que $\mathrm{Cov}\,(X_1, X_{p+1}) \leq \exp\Big[-\dfrac{4(M_1+1)}{M_1^2} r\epsilon_n\Big]$.

Quanto a $\bar{W}_n$, no *Lema 3.3* de Ioannides e Roussas [18], 1999, prova-se facilmente que $P\Big(|\bar{W}_n| \geq \epsilon_n\Big) = 0$, para n suficientemente grande. Além disso, as variáveis aleatórias $-Y_i, i = 1, \ldots, n$, têm as mesmas propriedades que as correspondentes Y_i, pelo que, por exemplo relativamente a $\bar{U}_n$, tem-se

$$P\left(\left|\bar{U}_n\right| \geq \epsilon_n\right) = P\left(\bar{U}_n \geq \epsilon_n\right) + P\left(-\bar{U}_n \geq \epsilon_n\right) \leq 2C_1 \;\exp(-2r\epsilon_n^2/M_1^2),$$

o mesmo sendo válido para $\bar{V}_n$.
A demonstração do Teorema 2.2.7 consiste em combinar estes resultados. Assim, para $n > n_0$,

$$\begin{aligned} P\left(\left|\bar{S}_n\right| \geq 3\,\epsilon_n\right) &\leq P\left(\left|\bar{U}_n\right| \geq \epsilon_n\right) + P\left(\left|\bar{V}_n\right| \geq \epsilon_n\right) + P\left(\left|\bar{W}_n\right| \geq \epsilon_n\right) \leq \\ &\leq P\left(\left|\bar{U}_n\right| \geq \epsilon_n\right) + P\left(\left|\bar{V}_n\right| \geq \epsilon_n\right) \leq \\ &\leq 4\,C_1\,\exp\left(-2\,r\,\epsilon_n^2/M_1^2\right). \end{aligned}$$

Fazendo $C_0 = 4C_1$ e substituindo ϵ_n por $\frac{\epsilon_n}{3}$, obtém-se,

$$P\left(\left|\bar{S}_n\right| \geq \epsilon_n\right) \leq C_0 \exp(-cr\epsilon_n^2),\ \ c = \frac{2}{9M_1^2},\ n \geq n_0,$$

desde que

$$\mathrm{Cov}\left(X_1, X_{p+1}\right) \leq \exp\left(-\frac{4(M_1+1)}{3M_1^2} r\epsilon_n\right).$$

Vejamos agora como escolher ϵ_n de forma que $\bar{S}_n$ convirja quase certamente para zero. Considere-se no Lema 2.2.8, $\frac{2}{M_1^2}\,r\,{\epsilon_n}^2 = \log n^\alpha$ para algum $\alpha > 1$. Isto implicaria uma escolha para ϵ_n, dada por

$$\epsilon_n = \left(\frac{\alpha M_1^2}{2}\right)^{1/2}\left(\frac{\log n}{r}\right)^{1/2}. \tag{2.17}$$

Com esta escolha para ϵ_n temos

$$P\Big(\bar{S}_n \geq \epsilon_n\Big) \leq C_0 \exp(-\alpha M_2 \log n) \sim n^{-\alpha} \tag{2.18}$$

onde $M_2 = \frac{1}{9}$. Como $\sum_{n=1}^{\infty} n^{-\alpha}$ converge para $\alpha > 1$, pelo Lema de Borel-Cantelli, podemos concluir que a velocidade de convergência de $\bar{S}_n$ para zero é da ordem de $\left(\frac{\log n}{r}\right)^{1/2}$.
A velocidade desta convergência depende de $r = r(n)$, sendo a velocidade óptima obtida tomando $r = n$, o que não é, obviamente, possível com esta desigualdade exponencial.
Tal como no caso independente, com o objectivo de melhorar esta velocidade, usamos uma desigualdade do tipo Kolmogorov, provada por

Newman e Wright [32], 1981, que enunciámos no primeiro capítulo, Teorema 1.5.1. Note-se ainda que se fizermos $\mathbb{E}\big(S_m^2\big)= s_m^2$, então

$$P\Big(\max\left(|S_1|,\ldots,|S_m|\right) \geq \lambda s_m\Big) \leq 2P\Big(|S_m| \geq (\lambda - \sqrt{2})s_m\Big).$$

Vejamos o resultado relativo ao majorante para a velocidade da convergência para zero de $\bar{S}_n$.

Teorema 2.2.9 *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão associada de variáveis aleatórias reais, com média zero e variância finita. Suponhamos que esta sucessão verifica as condições $(A_1), (A_2)$ e (A_3), que as covariâncias $\mathrm{Cov}\,(X_1, X_n)$ são decrescentes e que $\mathbb{E}(S_n^2) = O(n)$. Suponhamos ainda que as variáveis aleatórias X_n obedecem à seguinte relação:*

$$\mathrm{Cov}\,(X_1, X_{p+1}) \leq \exp(-M_2\, r_n \epsilon_n), \tag{2.19}$$

com M_2 constante. Seja

$$\epsilon_n \sim \Big(\frac{\log\log n}{r_n}\Big)^{1/2}, \tag{2.20}$$

Então, $P\left(\limsup \frac{\bar{S}_n}{\epsilon_n} > 1\right) = 0$.

Demonstração: Façamos $t_n = \sqrt{n}\,\epsilon_n$. De acordo com a desigualdade exponencial de Ioannides e Roussas,

$$P(\bar{S_n} \geq \epsilon_n) = P(S_n \geq n\,\epsilon_n) \leq C_0 \exp(-c\, r_n\, \epsilon_n^2),$$

onde $c = \frac{2M_1^2}{9}$. Com a escolha feita para ϵ_n temos, para algum $c' > 0$,

$$P(S_n \geq n\,\epsilon_n) = P(S_n \geq \sqrt{n}\,t_n) \leq C_0 (\log n)^{-c'}.$$

É condição suficiente para que a conclusão do Teorema se verifique que, para todo o $\delta > 0$, se tenha

$$P\left(\limsup\left\{S_n \geq (1+\delta)\sqrt{n}\,t_n\right\}\right) = 0. \tag{2.23}$$

De $\mathbb{E}(S_n)^2 = O(n)$ segue-se que existem constantes c_1, c_2 tais que $c_1 n \leq \mathbb{E}(S_n)^2 \leq c_2 n$. Sejamos mais precisos quanto à escolha da sucessão ϵ_n. Esta deve ser tal que

$$\epsilon_n^2 = \left(\frac{c_2}{c_1}\right)^2 \frac{1}{c} \frac{\log\log n}{r_n}.$$

Usando agora o Teorema 2.2.7 obtém-se

$$\begin{aligned} \text{Ora} \quad P(S_n \geq (1+\delta)\sqrt{n}\, t_n) &\leq C_0 \exp(-\left(\frac{c_2}{c_1}\right)^2 (1+\delta)^2 \log\log n) \sim \\ &\sim C_0 (\log n)^{-\left(\frac{c_2}{c_1}\right)^2 (1+\delta)^2}, \end{aligned} \tag{2.24}$$

onde C_0 é a constante interveniente na Desigualdade de Ioannides e Roussas. Se escolhermos $n_k = [\theta^k]$, com $\theta > 1$, onde $[x]$ representa o maior inteiro que não excede x, obtemos $(\log n_k)^{-\left(\frac{c_2}{c_1}\right)^2 (1+\delta)^2} \sim k^{-\left(\frac{c_2}{c_1}\right)^2 (1+\delta)^2}$, o que define, em (2.24), uma série convergente, pelo que, neste caso, (2.23) verifica-se.

Vejamos agora o que acontece para $n \neq n_k$. Seja

$$S_{n_k}^* = \max_{n \leq n_k} S_n, \;\; n \in (n_{k-1}, n_k].$$

Então,

$$P(S_n > (1+\delta)\sqrt{n}\, t_n) \leq P(S_{n_k}^* > (1+\delta)\sqrt{n_{k-1}}\, t_{n_{k-1}}). \tag{2.25}$$

Ora, $\sqrt{n_{k-1}}\, t_{n_{k-1}} = n_{k-1}\epsilon_{n_{k-1}} \sim \dfrac{n_k}{\theta}\, \epsilon_{n_k} \sqrt{\dfrac{\log\log\theta^{k-1}}{r_{\theta^{k-1}}} \dfrac{r_{\theta^k}}{\log\log\theta^k}}$.

Fazendo

$$\frac{r_{\theta^k}}{r_{\theta^{k-1}}} \sim \tilde{c}, \tag{2.26}$$

segue-se que $\sqrt{n_{k-1}}\, t_{n_{k-1}} \sim n_k \epsilon_{n_k} \frac{\tilde{c}}{\theta}$. É sempre possível escolher θ tal que $\dfrac{\tilde{c}}{\theta} > 1$. Seja θ tal que $1 < \dfrac{\tilde{c}}{\theta} < 1 + \delta$. Então, $\dfrac{1+\delta}{\theta/\tilde{c}} = 1 + \delta'$, com $\delta' < \delta$. Assim, para k suficientemente grande,

$$P(S_n > (1+\delta)\sqrt{n}\, t_n) \leq P(S_{n_k}^* > (1+\delta')\sqrt{n_k}\, t_{n_k}). \tag{2.27}$$

De acordo com o resultado de Newman e Wright [32], 1981,

$$P(S^*_{n_k} > (1+\delta')\sqrt{n_k}\,t_{n_k}) \leq 2\,P(S_{n_k} > (\frac{1+\delta'}{c_2} - \frac{\sqrt{2}}{t_{n_k}})c_1\sqrt{n_k}\,t_{n_k}).$$

Usando agora a desigualdade exponencial de Ioannides e Roussas, obtemos,

$$\begin{aligned} P(S_n > (1+\delta)\sqrt{n}\,t_n) &\leq 2C_0 \exp\left(-\left(\frac{1+\delta'}{c_2} - \frac{\sqrt{2}}{t_{n_k}}\right)^2 c_1^2 \log\log n_k\right) \leq \\ &\leq 2C_0 (\log n_k)^{-\left(\frac{1+\delta'}{c_2} - \frac{\sqrt{2}}{t_{n_k}}\right)^2 c_1^2} \sim k^{-\left(\frac{1+\delta'}{c_2} - \frac{\sqrt{2}}{t_{n_k}}\right)^2 c_1^2}. \end{aligned} \tag{2.28}$$

Ora, se $t_{n_k} \to +\infty$, ou seja, $\sqrt{n}\,\epsilon_n \to +\infty$, o lado direito da desigualdade anterior é o termo geral de uma série convergente. ■

Vejamos qual a escolha para as sucessões $\{r_n\}_{n\in\mathbb{N}}$ e $\{p_n\}_{n\in\mathbb{N}}$ que verifiquem esta condição. Relativamente a r_n, deverá ser tal que $\frac{n\log\log n}{r_n} \to +\infty$.

Sabemos também que, para que a desigualdade exponencial possa ser aplicada, a sucessão $\{p_n\}$ deve ser tal que

$$\mathrm{Cov}\,(X_1, X_{p_n+1}) \leq \exp(-\frac{4(M_1+1)}{3M_1^2} r_n \epsilon_n) = \exp\Big(-M_2'(r_n \log\log n)^{1/2}\Big),$$

o que, tendo em conta a relação entre r_n e p_n, implica que

$$r_n \leq M_3 \frac{\log^2 \mathrm{Cov}\,(X_1, X_{p+1})}{\log\log n}. \tag{2.29}$$

A escolha mais favorável para r_n ocorre quando se dá a igualdade.
Como consequência deste teorema, podemos obter um majorante para a velocidade de convergência da lei forte dos grandes números, no caso em que as covariâncias são geometricamente decrescentes.

Corolário 2.2.10 *Seja* $\{X_n\}_{n\in\mathbb{N}}$ *uma sucessão estritamente estacionária e associada de variáveis aleatórias reais, com média zero e variância finita. Suponhamos que esta sucessão verifica as condições* $(A_1), (A_2)$ *e* (A_3). *Sejam as covariâncias* $\mathrm{Cov}\,(X_1, X_n)$ *decrescentes e tais que existe* $\rho_0 > 0$ *e* $\rho \in (0,1)$ *tal que* $\mathrm{Cov}\,(X_1, X_n) = \rho_0\rho^n$. *Então,*

$$P\left(\limsup \frac{n^{1/3}}{(\log\log n)^{2/3}}\frac{1}{n}S_n > 1\right) = 0.$$

Demonstração: Se $\mathrm{Cov}\,(X_1, X_{k+1}) = \rho_0\rho^k$, $0 < \rho < 1$, a condição expressa em (2.29) escreve-se

$$r_n \leq M_3\frac{p_n^2\log^2\rho}{\log\log n} + \frac{\log^2\rho_0}{\log\log n}. \tag{2.30}$$

É evidente que a segunda parcela desta soma é irrelevante para o comportamento de r_n, uma vez que converge para zero. Queremos r_n o maior possível pelo que, o caso que interessa é aquele em que se dá a igualdade em (2.30). A forma como p_n e r_n estão definidas implica que $p_n = \dfrac{1}{2x_n}\dfrac{n}{r_n}$, para alguma sucessão x_n convergente para 1. Assim, temos,

$$r_n = O\Big(\frac{n^2}{\log\log n}\Big)^{1/3}$$

obtendo-se, neste caso,

$$p_n = O(n\log\log n)^{1/3}.$$

Notemos que, nesta situação, a condição (2.26) imposta sobre a sucessão $\{r_n\}_{n\in\mathbb{N}}$ é sempre verificada.

Assim, e nestas condições, a velocidade de convergência de $\bar{S}_n$ para zero é da ordem de $\left(\dfrac{(\log\log n)^2}{n}\right)^{1/3}$.

No que respeita à condição do Teorema 2.2.9 relativa a $\mathbb{E}(S_n^2)$, esta verifica-se uma vez que a hipótese $\mathrm{Cov}\,(X_1, X_n) = \rho_0\rho^{n-1}$, com $\rho_0 > 0$

constante e $\rho \in (0,1)$ implica $\rho_0 n \leq \mathbb{E}(S_n^2) \leq \rho_0 \frac{\rho}{1-\rho} n$. ■
Consideremos agora a sucessão estritamente estacionária de variáveis aleatórias associadas

$$\left\{U\Big(\frac{x-X_m}{h_n}, \frac{y-X_{k+m}}{h_n}\Big)\right\}_{m\in\mathbb{N}}.$$

Naturalmente que as variáveis aleatórias são limitadas e se

$$\mathrm{Cov}\,(X_1, X_{j+1}) = \rho_0 \rho^j, \text{para algum } \rho_0 > 0 \text{ e } 0 < \rho < 1,$$

tendo em conta o Lema 2.2.3 e o Lema 2.2.4, tem-se que

$$\mathrm{Cov}\,\left\{U\Big(\frac{x-X_1}{h_n}, \frac{y-X_{k+1}}{h_n}\Big), U\Big(\frac{x-X_j}{h_n}, \frac{y-X_{k+j}}{h_n}\Big)\right\} = O\big(\rho_1 \rho^{j/3}\big)$$

com $\rho_1 = \rho_0^{1/3}$. Vejamos, então, um majorante para a velocidade da convergência quase certa de $\widehat{F}_{n,k}$.

Corolário 2.2.11 *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas que verifica as condições $(C_1), (C_3)$, (C_5) e é tal que para $j \geq 1$,*

$$\mathrm{Cov}\,(X_1, X_{j+1}) = \rho_0 \rho^j, \textit{para algum } \rho_0 > 0 \;\; e \;\; 0 < \rho < 1.$$

Então

$$\Big|\widehat{F}_{n,k}(x,y) - F_k(x,y)\Big| = O\Big(\frac{(\log\log n)^2}{n}\Big)^{1/3} + O(h_n^2)$$

Demonstração: Podemos sempre escrever

$$\begin{aligned}\Big|\widehat{F}_{n,k}(x,y) - F_k(x,y)\Big| \leq \\ \leq \Big|\widehat{F}_{n,k}(x,y) - \mathbb{E}\left(\widehat{F}_{n,k}(x,y)\right)\Big| + \Big|\mathbb{E}\left(\widehat{F}_{n,k}(x,y)\right) - F_k(x,y)\Big|.\end{aligned}$$

Quanto à primeira parcela do segundo membro desta desigualdade obtém-se, de um modo análogo ao da demonstração do Corolário 2.2.10,

$$\Big|\widehat{F}_{n,k}(x,y) - \mathbb{E}\left(\widehat{F}_{n,k}(x,y)\right)\Big| = O\Big(\frac{(\log\log n)^2}{n}\Big)^{1/3}.$$

Agora, de acordo com o Teorema 2.2.2, a segunda parcela da desigualdade é da mesma ordem que h_n^2, pelo que o resultado é imediato. ■

Relativamente à convergência uniforme, vimos já no Teorema 2.2.6 que, para $A > 1$ arbitrário mas fixo e $i, j \in \{1, \ldots, A\}$, se tem, com $x_{A,i}$ definido como na demonstração do referido teorema,

$$\sup_{x,y\in\mathbb{R}} \left|\widehat{F}_{n,k}(x,y) - F_k(x,y)\right| \leq$$
$$\leq \max_{0\leq i,\, j\leq A} \left|\widehat{F}_{n,k}\Big(x_{A,i}, x_{A,j}\Big) - F_k\Big(x_{A,i}, x_{A,j}\Big)\right| + \frac{2}{A}.$$

Além disso, temos também, de acordo com o Teorema 2.2.2, que a ordem de convergência de $\left|F_k(x_{A,i}, x_{A,j}) - \mathbb{E}\left(\widehat{F}_{n,k}(x_{A,i}, x_{A,j})\right)\right|$ é, para cada A fixo da $O(h_n^2)$. Assim,

$$\begin{aligned}
&\sup_{x,y\in\mathbb{R}} \left|\widehat{F}_{n,k}(x,y) - F_k(x,y)\right| \leq \\
&\quad \leq \max_{0\leq i,\, j\leq A} \left|\widehat{F}_{n,k}\Big(x_{A,i}, x_{A,j}\Big) - \mathbb{E}\left(\widehat{F}_{n,k}\Big(x_{A,i}, x_{A,j}\Big)\right)\right| + \\
&\qquad + \max_{0\leq i,\, j\leq A} \left|F_k\Big(x_{A,i}, x_{A,j}\Big) - \mathbb{E}\left(\widehat{F}_{n,k}\Big(x_{A,i}, x_{A,j}\Big)\right)\right| + \frac{2}{A} \leq \\
&\quad \leq \max_{0\leq i,\, j\leq A} \left|\widehat{F}_{n,k}\Big(x_{A,i}, x_{A,j}\Big) - \mathbb{E}\left(\widehat{F}_{n,k}\Big(x_{A,i}, x_{A,j}\Big)\right)\right| + A^2 h_n^2 + \frac{2}{A}.
\end{aligned} \tag{2.31}$$

Assim a convergência uniforme depende da velocidade de convergência de A para $+\infty$, quando este deixa de ser fixo. Uma primeira análise à desigualdade (2.31) leva a que esta velocidade deverá ser, pelo menos, de ordem inferior à de h_n. Vejamos, então, a primeira parcela desta expressão.

Suponhamos que $\{\delta_n\}_{n\in\mathbb{N}}$ é uma sucessão de números positivos conver-

gente para zero.

$$P\left(\max_{0\leq i,\,j\leq A}\left|\widehat{F}_{n,k}\Big(x_{A,i},x_{A,j}\Big)-\mathbb{E}\left(\widehat{F}_{n,k}\Big(x_{A,i},x_{A,j}\Big)\right)\right|>(1+\delta_n)\,\epsilon_n\right)\leq$$

$$\leq\sum_{i,j=1}^{A}P\left(\left|\widehat{F}_{n,k}\Big(x_{A,i},x_{A,j}\Big)-\mathbb{E}\left(\widehat{F}_{n,k}\Big(x_{A,i},x_{A,j}\Big)\right)\right|>(1+\delta_n)\epsilon_n\right)\sim$$

$$\sim A^2(\log n_k)^{-\left(\frac{1+\delta_{n_k}}{c_2}-\frac{\sqrt{2}}{\sqrt{n_k}\epsilon_{n_k}}\right)^2c_1^2}, \tag{2.32}$$

em que $n_k=\left[\theta^k\right]$, com $\theta>1$ e levando a cabo um raciocínio análogo ao da demonstração do Teorema 2.2.9, que nos conduz a uma expressão do tipo da (2.28). Ora se A^2 for da ordem de $(\log n)^{\beta_n}$, com $\beta_n\to 0$ então, uma vez que $t_{n_k}\to\infty$ e $\delta_{n_k}\to 0$, o lado direito da desigualdade (2.32) é o termo geral de uma série convergente.

Como vimos, fazendo $\mathrm{Cov}\,(X_1,X_{j+1})=\rho_0\rho^j, 0<\rho<1$ e $\rho_0>0$, as escolhas para r_n e p_n levam a $\epsilon_n=O\Big(\frac{(\log\log n)^2}{n}\Big)^{1/3}$, pelo que a velocidade da convergência uniforme do estimador é sempre $\min\left\{O(\epsilon_n)+\frac{1}{(\log n)^{\beta_n}},o(h_n)\right\}$.

É, agora, imediato, o resultado relativo à ordem de convergência uniforme do estimador.

Corolário 2.2.12 *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas tal que para $j\geq 1$ e $\rho_0>0$,*

$$\mathrm{Cov}\,(X_1,X_{j+1})=\rho_0\rho^j,\ 0<\rho<1.$$

Seja, ainda, $\{\beta_n\}_{n\in\mathbb{N}}$ uma sucessão de inteiros tal que $\beta_n\to 0$. Então

$$\sup_{x,y\in\mathbb{N}}\left|\widehat{F}_{n,k}(x,y)-F_k(x,y)\right|=O\left(\left(\frac{(\log\log n)^2}{n}\right)^{1/3}+\frac{1}{(\log n)^{\beta_n}}\right)+o(h_n).$$

2.3 O Erro Quadrático Médio do Estimador

Nesta secção fazemos a análise do Erro Quadrático Médio (MSE) do estimador $\widehat{F}_{n,k}$ e, com base na expressão obtida, por minimização rela-

tivamente a h_n, obtemos a velocidade óptima da convergência para zero da sucessão $\{h_n\}_{n\in\mathbb{N}}$.
Esta velocidade de convergência depende, como veremos, da condição imposta para a sucessão $\{n\,\mathrm{Cov}\,(X_1, X_n)\}_{n\in\mathbb{N}}$, nomeadamente a condição (C_7). Com esta condição, obtemos uma velocidade da ordem de $O(n^{-1})$. Sob uma condição mais restritiva, usada também por Cai e Roussas [8], a velocidade de convergência optima é $O(n^{-1/3})$, a mesma que no caso independente e a mesma obtida por Cai, Roussas no caso univariado.
Terminamos esta secção mostrando que se considerarmos um estimador $p-$dimensional, com $p > 2$, para a função de distribuição multivariada $\mathbf{F}_p$ do vector $\mathbf{X} = (X_1, \ldots, X_p)$, obtemos, ainda, uma velocidade de convergência óptima da $O(n^{-1/3})$. Refira-se que, para o caso independente, este resultado encontra-se demonstrado em Jin, Shao [23], 1999.

Vamos, em primeiro lugar, provar um resultado intermédio com o fim de obter a especificação da variância do estimador $\widehat{F}_{n,k}$. Começamos por definir as funções auxiliares, V_1, V_2 e V_3 de $\mathbb{R}^2$ em $\mathbb{R}$, do seguinte modo:

$$\begin{aligned}
V_1(x,y) &= \frac{\partial^2 F_k}{\partial x^2}(x,y)\int_{\mathbb{R}^2} s^2 u(s,t)\,ds\,dt + \frac{\partial^2 F_k}{\partial x\,\partial y}(x,y)\int_{\mathbb{R}^2} s\,t\,u(s,t)\,ds\,dt + \\
&\quad + \frac{\partial^2 F_k}{\partial y^2}(x,y)\int_{\mathbb{R}^2} t^2 u(s,t)\,ds\,dt; \\
V_2(x,y) &= \frac{\partial F_k}{\partial x}(x,y)\int_{\mathbb{R}^2} s\,V(s,t)\,ds\,dt + \frac{\partial F_k}{\partial y}(x,y)\int_{\mathbb{R}^2} t\,V(s,t)\,ds\,dt; \\
V_3(x,y) &= \frac{\partial^2 F_k}{\partial x^2}(x,y)\int_{\mathbb{R}^2} s^2 V(s,t)\,ds\,dt + \frac{\partial^2 F_k}{\partial x\,\partial y}(x,y)\int_{\mathbb{R}^2} s\,t\,V(s,t)\,ds\,dt + \\
&\quad + \frac{\partial^2 F_k}{\partial y^2}(x,y)\int_{\mathbb{R}^2} t^2 V(s,t)\,ds\,dt.
\end{aligned}$$

onde V é a função definida na condição (C_8) da primeira secção deste capítulo.

Lema 2.3.1 *Suponhamos que a sucessão de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$ *satisfaz as condições* $(C_1), (C_3), (C_4), (C_5)$ *e* (C_8). *Então,*

$$(i)\mathbb{E}\left(U^2\left(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}\right)\right) = F_k(x,y) - h_n\,V_2(x,y) + \frac{h_n^2}{2}\,V_3(x,y) + \\ +o(h_n^2);$$

$$(ii)\left|\mathrm{Var}\left(U\left(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}\right)\right) - (F_k(x,y)(1-F_k(x,y)) + h_n\,V_2(x,y))\right| = \\ = \tfrac{h_n^2}{2}\,(V_3(x,y) - F_k(x,y)V_1(x,y)) + o(h_n^2).$$

Demonstração: Relativamente à condição (i), por definição de esperança matemática,

$$\mathbb{E}\left(U^2\left(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}\right)\right) = \\ = \int_{\mathbb{R}^2} U^2\left(\frac{x-s}{h_n},\frac{y-t}{h_n}\right)\,dF_k(s,t) = \\ = \int_{\mathbb{R}^2}\left(\int_{-\infty}^{\frac{x-s}{h_n}}\int_{-\infty}^{\frac{y-t}{h_n}} V(a,b)\,db\,da\right)\,dF_k(s,t).$$

Utilizando o Teorema de Fubini e efectuando mudança de variável, obtemos,

$$\mathbb{E}\left(U^2\left(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}\right)\right) = \int_{\mathbb{R}^2} V(a,b)F_k(x-ah_n,y-bh_n)dbda.$$

Por desenvolvimento de $F_k(x-ah_n,y-bh_n)$ num polinómio de segundo grau, em torno de (x,y), utilizando a condição (C_5) e tendo em conta as definições das funções V_2 e V_3, vem,

$$\mathbb{E}\left(U^2\left(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}\right)\right) = F_k(x,y) - V_2(x,y)\,h_n + \frac{h_n^2}{2}\,V_3(x,y) + o(h_n^2).$$

Quanto à variância de $U\left(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}\right)$ sabemos que

$$\mathrm{Var}\left(U\left(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}\right)\right) = \\ = \mathbb{E}\left(U^2\left(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}\right)\right) - \left(\mathbb{E}\left(U\left(\frac{x-X_1}{h_n},\frac{y-X_{k+1}}{h_n}\right)\right)\right)^2.$$

Ora, de acordo com a definição de V_1, a demonstração do Teorema 2.2.2, permite escrever,

$$\mathbb{E}\left(U\left(\frac{x-X_1}{h_n},\frac{x-X_{k+1}}{h_n}\right)\right)=F_k(x,y)+\frac{h_n^2}{2}V_1(x,y)+o(h_n^2).$$

Substituindo agora a expressão obtida em (i) e o quadrado desta última na expressão para a variância, o resultado obtém-se imediatamente. ■

Nota 2.3.2 *No Lema 2.2.3 obtivemos um resultado para a covariância entre as funções de variáveis aleatórias* $Y_{i,j}=U(\frac{x-X_i}{h_n},\frac{y-X_j}{h_n})$, *a saber,*

$$\begin{aligned}&\mathrm{Cov}\left(U(\tfrac{x-X_1}{h_n},\tfrac{y-X_{k+1}}{h_n}),U(\tfrac{x-X_j}{h_n},\tfrac{y-X_{k+j}}{h_n})\right)=\\&=\mathrm{Cov}\left\{\mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_1,X_{k+1}),\mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_j,X_{k+j})\right\}+O({h_n}^2)=\\&=F_{k,j}(x,y,x,y)-F_k^2(x,y)+O(h_n^2).\end{aligned}$$

O objectivo é, agora, tentar especificar o termo $O(h_n^2)$ *interveniente nesse resultado. Para tal definimos uma nova função auxiliar* V_4 *de* $\mathbb{R}^2$ *em* $\mathbb{R}$ *por*

$$\begin{aligned}V_4(x,y)=&\frac{\partial^2F_{k,j}}{\partial x_1^2}(x,y,x,y)\int_{\mathbb{R}^2}s^2\,u(s,t)\,ds\,dt+\\&+2\frac{\partial^2F_{k,j}}{\partial x_1\,\partial x_2}(x,y,x,y)\int_{\mathbb{R}^2}s\,t\,u(s,t)\,ds\,dt+\\&+\frac{\partial^2F_{k,j}}{\partial x_2^2}(x,y,x,y)\int_{\mathbb{R}^2}t^2\,u(s,t)\,ds\,dt+\\&+\frac{\partial^2F_{k,j}}{\partial x_3^2}(x,y,x,y)\int_{\mathbb{R}^2}s^2\,u(s,t)\,ds\,dt+\\&+2\frac{\partial^2F_{k,j}}{\partial x_3\,\partial x_4}(x,y,x,y)\int_{\mathbb{R}^2}s\,t\,u(s,t)\,ds\,dt+\\&+\frac{\partial^2F_{k,j}}{\partial x_4^2}(x,y,x,y)\int_{\mathbb{R}^2}t^2\,u(s,t)\,ds\,dt.\end{aligned}$$

Retomando a demonstração do Teorema 2.2.2, ao considerarmos o desenvolvimento de $F_{k,j}(x-ah_n,y-bh_n,x-ch_n,y-dh_n)$ *em série de*

polinómios de segunda ordem em torno de (x,y,x,y), *usando as condições* (C_4) *e* (C_8), *de acordo com a definição de* V_1 *e* V_4, *o integral*

$$\int_{\mathbb{R}^4} u(a,b)\,u(c,d)\,F_{k,j}(x-ah_n, y-bh_n, x-ch_n, y-dh_n)\,da\,db\,dc\,dd$$

é igual a

$$F_{k,j}(x,y,x,y) + \frac{h_n^2}{2}\,V_4(x,y) + o(h_n^2).$$

Nestas condições, para $j>1$,

$$\begin{aligned}
&\mathrm{Cov}\left(U\left(\frac{x-X_1}{h_n}, \frac{y-X_{k+1}}{h_n}\right), U\left(\frac{x-X_j}{h_n}, \frac{y-X_{k+j}}{h_n}\right)\right) = \\
&\quad = \mathrm{Cov}\left(\mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_1, X_{k+1}), \mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_j, X_{k+j})\right) + \\
&\qquad + \frac{h_n^2}{2}\left(V_4(x,y) - 2F_k(x,y)\,V_1(x,y)\right) + o(h_n^2).
\end{aligned}$$

Podemos agora enunciar o resultado relativo à variância do estimador.

Teorema 2.3.3 *Suponhamos que a sucessão* $\{X_n\}_{n\in\mathbb{N}}$ *de variáveis aleatórias reais verifica as condições* $(C_1), (C_3), (C_4), (C_5), (C_6), (C_7)$ *e* (C_8). *Então,*

$$\begin{aligned}
&(n-k)\mathrm{Var}\left(\widehat{F}_{n,k}(x,y)\right) = \\
&\quad = \sigma^2(x,y) - h_n\,V_2(x,y) + (n-k-1)h_n^2\left(V_4(x,y) - F_k(x,y)\,V_1(x,y)\right) + \\
&\qquad + O(h_n^2) - c_n(x,y),
\end{aligned}$$

onde

$$\sigma^2(x,y) = F_k(x,y) - F_k^2(x,y) + 2\sum_{j=2}^{\infty}\left(F_{k,j}(x,y,x,y) - F_k^2(x,y)\right) \quad e$$

$$\begin{aligned}
c_n(x,y) \;&=\; 2\sum_{j=n-k+1}^{+\infty}\left(F_{k,j}(x,y,x,y) - F^2(x,y)\right) + \\
&\quad + \frac{1}{n-k}\sum_{j=2}^{n-k}(j-1)\left(F_{k,j}(x,y,x,y) - F^2(x,y)\right). \qquad (2.34)
\end{aligned}$$

Nota 2.3.4 *O lema 2.2.4 juntamente com as condições* $(C_7)'$ *e* (C_7) *são suficientes para garantir que as funções* σ^2 *e* c_n *estejam bem definidas.*

Demonstração: Por definição de variância,

$$\text{Var}\left(\widehat{F}_{n,k}(x,y)\right) = \frac{1}{(n-k)^2}\times$$
$$\times \sum_{i,j=1}^{n-k} \text{Cov}\left(U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right), \left(U\left(\frac{x-X_j}{h_n}, \frac{y-X_{k+j}}{h_n}\right)\right)\right) =$$

(usando a estacionaridade da sucessão $\{X_n\}_{n\in\mathbb{N}}$,)

$$= \frac{1}{(n-k)}\text{Var}\left(U\left(\frac{x-X_1}{h_n}, \frac{y-X_{k+1}}{h_n}\right)\right) + \frac{2}{(n-k)^2}\sum_{j=2}^{n-k}(n-k-j+1)\times$$
$$\times\left(\text{Cov}\left(U\left(\frac{x-X_1}{h_n}, \frac{y-X_{k+1}}{h_n}\right), U\left(\frac{x-X_j}{h_n}, \frac{y-X_{k+j}}{h_n}\right)\right)\right).$$

Utilizando agora os resultados do lema e nota anteriores, assim como o resultado do Lema 2.2.4,

$$(n-k)\text{Var}\left(\widehat{F}_{n,k}(x,y)\right) = F_k(x,y) - F_k^2(x,y) - V_2(x,y)\,h_n +$$
$$+(V_3(x,y) - F_k(x,y)V_1(x,y))\,h_n^2 + \frac{2}{n-k}\sum_{j=2}^{n-k}(n-k-j+1)\times$$
$$\times\left(F_{k,j}(x,y,x,y) - F_k^2(x,y) + \frac{h_n^2}{2}(V_4(x,y) - F_k(x,y)V_1(x,y)\right).$$

Temos, então,

$$(n-k)\text{Var}\left(\widehat{F}_{n,k}(x,y)\right) = F_k(x,y) - F_k^2(x,y) - V_2(x,y)\,h_n +$$
$$+(V_3(x,y) - F_k(x,y)V_1(x,y))\,h_n^2 + \sum_{j=2}^{n-k}\left(F_{k,j}(x,y,x,y) - F_k^2(x,y)\right) +$$
$$+(n-k-1)\,h_n^2(V_4(x,y) - F(x,y)V_1(x,y)) -$$
$$-\frac{2}{n-k}\sum_{j=2}^{n-k}(j-1)\left(F_{k,j}(x,y,x,y) - F^2(x,y)\right) + O(h_n^2).$$

Substituindo, agora, $\sum_{j=2}^{n-k}\left(F_{k,j}(x,y,x,y)-F_k^2(x,y)\right)$ por $\sum_{j=2}^{\infty}\left(F_{k,j}(x,y,x,y)-F_k^2(x,y)\right)$ e subtraindo ao resultado anterior, $\sum_{j=n-k+1}^{\infty}\left(F_{k,j}(x,y,x,y)-F_k^2(x,y)\right)$ obtemos a expressão que pretendemos para a variância de $\widehat{F}_{n,k}$. ■

Vejamos, então, a expressão do Erro Quadrático Médio de $\widehat{F}_{n,k}$.

Teorema 2.3.5 *Suponhamos que a sucessão de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$ *satisfaz as condições* $(C_1),(C_2),(C_3),(C_4),(C_5),(C_6),(C_7)$ *e* (C_8). *Então,*

$$(n-k)\text{MSE}\left(\widehat{F}_{n,k}(x,y)\right)=\sigma^2(x,y)-V_2(x,y)\,h_n+$$
$$+(n-k-1)\,h_n^2\,\left(V_4(x,y)-F_k(x,y)\,V_1(x,y)\right)+o(h_n^2).$$

Demonstração: O resultado deste teorema é imediato de acordo com a definição de Erro Quadrático Médio e usando os resultados anteriores relativos à média e variância de $\widehat{F}_{n,k}$. ■

Como corolário obtemos a ordem óptima de convergência para zero da sucessão $\{h_n\}_{n\in\mathbb{N}}$.

Corolário 2.3.6 *Suponhamos que a sucessão de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$ *satisfaz as condições* $(C_1),(C_2),(C_3),(C_4),(C_5),(C_6),(C_7)$ *e* (C_8). *Então, a escolha óptima da sucessão* $\{h_n\}_{n\in\mathbb{N}}$, *no sentido de minimizar a expressão para o Erro Quadrático Médio obtida no teorema anterior, é*

$$h_n(x,y)=\frac{V_2(x,y)}{2(n-k-1)\left(V_4(x,y)-F_k(x,y)\,V_1(x,y)\right)}.$$

Conforme já referimos, com o objectivo de obtermos, tal como no caso independente, uma velocidade de convergência para a sucessão $\{h_n\}_{n\in\mathbb{N}}$ da ordem de $n^{-1/3}$, substituímos as hipóteses (C_6) e (C_7) por duas hipóteses correspondentes a estas, utilizadas por Cai, Roussas [8], 1998, para a prova daquela velocidade no caso univariado, sob associação. Vamos supor, ainda, um decrescimento na sucessão de covariâncias cujo termo geral é $\text{Cov}\,(X_1,X_{k+1})$.

Teorema 2.3.7 *Suponhamos que a sucessão* $\{X_n\}_{n\in\mathbb{N}}$ *de variáveis aleatórias reais satisfaz as condições* $(C_1), (C_3), (C_4), (C_5)$ *e* (C_8). *Consideremos que* $\{h_n\}_{n\in\mathbb{N}}$ *é tal que* $n\,h_n^4 \to 0$ *e que existe* τ, *com* $0 < \tau < 1$, *tal que*

$$\sum_{j=1}^{\infty} \left(\mathrm{Cov}\left(X_1, X_{j+1}\right)\right)^{\frac{1-\tau}{3}} < \infty. \tag{2.35}$$

Suponhamos, ainda, que a sucessão de covariâncias $\{\mathrm{Cov}\left(X_1, X_{j+1}\right)\}_{j\in\mathbb{N}}$ *é decrescente. Nestas condições,*

$$(n-k)\mathrm{MSE}\left(\widehat{F}_{n,k}(x,y)\right) = \sigma^2(x,y) - h_n V_2(x,y) + \\ + \frac{(V_1(x,y))^2}{4}(n-k)\,h_n^4 + o(h_n + n\,h_n^4) + c_n,$$

com c_n definida em (2.34).

Nota 2.3.8 *Tal como já foi notado por Cai, Roussas [8], para* $0 < \tau < 1/2$, *a condição* (C_7) *implica que (2.35) se verifique.*

Demonstração: A prova deste teorema segue exactamente os mesmos passos da do teorema anterior. Usamos, contudo, alguns artifícios que, juntamente com as hipóteses apresentadas e resultados anteriores, nos permitem obter o resultado enunciado. Definimos, primeiramente, duas funções auxiliares $I_{n,j}$ e I_j com o objectivo de simplificar a notação envolvida nesta demonstração. Consideremos, então, para $(x,y) \in \mathbb{R}^2$,

$$I_{nj}(x,y) = \mathrm{Cov}\left(U\left(\frac{x-X_1}{h_n}, \frac{x-X_{k+1}}{h_n}\right), U\left(\frac{x-X_j}{h_n}, \frac{x-X_{k+j}}{h_n}\right)\right)$$

e

$$I_j(x,y) = \mathrm{Cov}\left(\mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_1, X_{k+1}), \mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_j, X_{k+j})\right).$$

Como vimos no Lema 2.2.2 e na Nota 2.3.2,

$$|I_{nj}(x,y) - I_j(x,y)| \leq c\,h_n^2,$$

com

$$c = \sup_{(x,y)\in\mathbb{R}^2} \{V_4(x,y) - F_k(x,y)\, V_1(x,y)\}.$$

Por outro lado, de acordo com o Lema 2.2.4, usando a estacionaridade da sucessão $\{X_n\}_{n\in\mathbb{N}}$ e, também, o facto de a sucessão de covariâncias $\{\mathrm{Cov}\,(X_1, X_{j+1})\}_{j\in\mathbb{N}}$ ser decrescente, obtemos para $k \leq j$ fixo,

$$\sup_{(x,y)\in\mathbb{R}^2} I_j(x,y) \leq 4\, M \left(\mathrm{Cov}^{\,1/3}(X_1, X_{j-k})\right),$$

assim como,

$$\sup_{(x,y)\in\mathbb{R}^2} I_{nj}(x,y) \leq 4\, M \left(\mathrm{Cov}^{\,1/3}(X_1, X_{j-k})\right).$$

Temos, então, para cada $(x,y) \in \mathbb{R}^2$,

$$|I_{nj}(x,y) - I_j(x,y)|^{1-\tau} \leq \left(8\, M \mathrm{Cov}^{\,1/3}(X_1, X_{j-k})\right)^{1-\tau}.$$

Ora, podemos sempre escrever,

$$|I_{nj}(x,y) - I_j(x,y)| = |I_{nj}(x,y) - I_j(x,y)|^{\tau}\, |I_{nj}(x,y) - I_j(x,y)|^{1-\tau},$$

obtendo,

$$\begin{aligned} |I_{nj}(x,y) - I_j(x,y)| &\leq c^{\tau}\, h_n^{2\tau} (8M)^{1-\tau} \left(\mathrm{Cov}^{\,1/3}(X_1, X_{j-k})\right)^{1-\tau} = \\ &= \tilde{c}\, h_n^{2\tau} \left(\mathrm{Cov}^{\,1/3}(X_1, X_{j-k})\right)^{1-\tau}. \end{aligned}$$

Considere-se, agora, a seguinte expressão para a variância de $\widehat{F}_{n,k}$,

$$\begin{aligned} (n-k)\mathrm{Var}\left(\widehat{F}_{n,k}(x,y)\right) &= \mathrm{Var}\left(U\left(\frac{x - X_1}{h_n}, \frac{y - X_{k+1}}{h_n}\right)\right) + \\ &+ \frac{2}{n-k} \sum_{j=2}^{n-k} (n-k-j+1)\, |I_{n,j}(x,y) - I_j(x,y)| + \\ &+ \frac{2}{n-k} \sum_{j=2}^{n-k} (n-k-j+1)\, I_j(x,y). \end{aligned}$$

De acordo com o que foi anteriormente estabelecido, temos

$$\begin{aligned}\frac{1}{n-k}\sum_{j=2}^{n-k}(n-k-j+1)\,|I_{n,j}(x,y)-I_j(x,y)| \leq \\ \leq \sum_{j=2}^{n-k}|I_{n,j}(x,y)-I_j(x,y)| \leq \\ \leq \tilde{c}h_n^{2\tau}\sum_{j=1}^{\infty}\left(\mathrm{Cov}^{\,1/3}(X_1,X_{j+1})\right)^{1-\tau}=O(h_n^{2\tau}),\end{aligned}$$

utilizando (2.35).
Assim, e neste caso,

$$(n-k)\mathrm{Var}\left(\widehat{F}_{n,k}(x,y)\right)=\sigma^2(x,y)-h_nV_2(x,y)+O(h_n^{2\tau}).$$

É agora imediato que

$$\begin{aligned}(n-k)\mathrm{MSE}\left(\widehat{F}_{n,k}(x,y)\right) &= \sigma^2(x,y)-h_nV_2(x,y)+ \\ &+ \frac{(V_1(x,y))^2}{4}(n-k)\,h_n^4+o(h_n+n\,h_n^4)+c_n,\end{aligned}$$

uma vez que

$$(n-k)\left(F_k(x,y)-\mathbb{E}\left(\widehat{F}_{n,k}(x,y)\right)\right)^2=\frac{(V_1(x,y))^2}{4}(n-k)\,h_n^4+o(h_n^4).\blacksquare$$

Por minimização relativamente a h_n da expressão para o Erro Quadrático Médio do estimador, obtemos como ordem optima de convergência da sucessão $\{h_n\}_{n\in\mathbb{N}},\ n^{-1/3}$ tal como no caso independente e tal como sob associação no caso univariado. Temos, então, o seguinte resultado.

Corolário 2.3.9 *Suponhamos que a sucessão de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$ *satisfaz satisfaz as condições* $(C_1),(C_3),(C_4),(C_5)$ *e* (C_8). *Consideremos que* $\{h_n\}_{n\in\mathbb{N}}$ *é tal que* $n\,h_n^4\to 0$ *e que existe* τ, *com* $0<\tau<1$, *tal que (2.35) se verifica. Suponhamos, ainda, que a sucessão de covariâncias* $C(X_1,X_{j+1})$ *é decrescente. Então, a escolha optima da sucessão*

$\{h_n\}_{n\in\mathbb{N}}$, *no sentido de minimizar a expressão para o Erro Quadrático Médio de* $\widehat{F}_{n,k}$ *é* $h_n = \left(\dfrac{V_2}{V_1^2(n-k)}\right)^{1/3}$.

A extensão dos resultados relativos à velocidade optima de convergência da sucessão $\{h_n\}_{n\in\mathbb{N}}$ para o espaço $\mathbb{R}^p$, é possível, não sendo alterada a velocidade desta convergência qualquer que seja a dimensão em causa.
Dada uma amostra de dimensão n da sucessão associada e estacionária de variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$, um estimador para a função de distribuição $\mathbf{F}_p$ do vector $\mathbf{X} = (X_1, \ldots, X_p)$, com $p > 2$, $\widehat{F}_{n,p}$ é dado para cada $\mathbf{x} = (x_1, \ldots, x_p) \in \mathbb{R}^p$, por

$$\widehat{F}_{n,p}(\mathbf{x}) = \frac{1}{n-p}\sum_{i=1}^{n-p}\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{i,p}}{h_n}\right),$$

onde $\mathbf{U}$ é uma função de distribuição $p-$variada conhecida, a função núcleo, e para cada p fixo e $i = 1, \ldots, n-p$, consideramos $\mathbf{X}_{i,p} = (X_{i+1}, \ldots, X_{i+p})$.
Com um conjunto de condições do mesmo tipo que o apresentado na primeira secção deste capítulo, obtêm-se resultados análogos aos obtidos, sob associação, no caso univariado por Cai, Roussas [8], 1999, e no caso bivariado aos apresentados na segunda secção, assim como aos obtidos sob independência, no caso multivariado, por Jin, Shao [23], 1999. Refira-se que este conjunto de condições é essencialmente o mesmo que o usado por Jin, Shao no caso independente. A excepção é, como parece evidente, relativa às condições (C_7) e $(C_7)'$, que são inerentes ao facto de estarmos a trabalhar sob associação.
Assim, as condições $(C_3), (C_4)$, (C_5) e (C_8) são substituídas por

$(C_3)'$ A função de distribuição $\mathbf{F}_p$ do vector aleatório $\mathbf{X} = (X_1, \ldots, X_p)$ tem derivadas parciais limitadas e contínuas até à ordem dois;

$(C_4)'$ Para cada inteiro positivo j, a função de distribuição do vector $\mathbf{X_{p,j}} = (X_1, \ldots, X_p, X_{j+1}, \ldots, X_{j+p})$, $\mathbf{F}_{p,j}$, tem derivadas parciais de primeira e segunda ordens limitadas e contínuas;

$(C_5)'$ A função núcleo $\mathbf{U}$, função de distribuição absolutamente contínua de $\mathbb{R}^p$ para $[0,1]$ é $p-$diferenciável e é tal que $\mathbf{u} = \dfrac{\partial^p \mathbf{U}}{\partial x_1 \ldots \partial x_p}$ satisfaz as seguintes condições:

(i) $\displaystyle\int_{\mathbb{R}^p} \mathbf{u}(\mathbf{x}) d\mathbf{x} = 1;$

(ii) $\displaystyle\int_{\mathbb{R}^p} \mathbf{x}\mathbf{u}(\mathbf{x}) d\mathbf{x} = 0;$

(iii) $\displaystyle\int_{\mathbb{R}^p} \mathbf{x}\,\mathbf{x}^T\,\mathbf{u}(\mathbf{x}) d\mathbf{x} < \infty;$

$(C_8)'$ A função $\mathbf{V} = \dfrac{\partial^p \mathbf{U}^2}{\partial x_1 \ldots \partial x_p}$, satisfaz a seguinte condição:

$$\int_{\mathbb{R}^p} \mathbf{x}\,\mathbf{x}^T\,\mathbf{V}(\mathbf{x}) d\mathbf{x} < \infty.$$

Nota 2.3.10 *Convém referir que* $\displaystyle\int_{\mathbb{R}^p} \mathbf{V}(\mathbf{x}) d\mathbf{x} = \mathbf{U^2}(+\infty, \ldots, +\infty) = 1.$

Se definirmos, de um modo análogo ao caso bivariado, funções auxiliares $\mathbf{V_1}, \mathbf{V_2}$ e $\mathbf{V_3}$ de $\mathbb{R}^p$ para $\mathbb{R}$, em que para cada $\mathbf{x} = (x_1, \ldots, x_p)$, se tem

- $\mathbf{V_1}(\mathbf{x}) = \displaystyle\sum_{i=1}^{p} \frac{\partial^2 \mathbf{F}_p}{\partial x_i^2} \int_{\mathbb{R}^p} a_i^2 \mathbf{u}(\mathbf{a}) d\mathbf{a} + 2 \sum_{j=1}^{p-1} \sum_{i=j+1}^{p} \frac{\partial^2 \mathbf{F}_p}{\partial x_j\, \partial x_i} \int_{\mathbb{R}^p} a_i a_j \mathbf{u}(\mathbf{a}) d\mathbf{a};$
- $\mathbf{V_2}(\mathbf{x}) = \displaystyle\sum_{i=1}^{p} \frac{\partial \mathbf{F}_p}{\partial x_i} \int_{\mathbb{R}^p} a_i \mathbf{V}(\mathbf{a}) d\mathbf{a};$
- $\mathbf{V_3}(\mathbf{x}) = \displaystyle\sum_{i=1}^{p} \frac{\partial^2 \mathbf{F}_p}{\partial x_i^2} \int_{\mathbb{R}^p} a_i^2 \mathbf{V}(\mathbf{a}) d\mathbf{a} + 2 \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \frac{\partial^2 \mathbf{F}_p}{\partial x_j\, \partial x_i} \int_{\mathbb{R}^p} a_i a_j \mathbf{V}(\mathbf{a}) d\mathbf{a},$

obtemos, também de um modo análogo ao caso bivariado, os resultados que apresentamos em seguida.

Lema 2.3.11 *Suponhamos que a sucessão de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$ *satisfaz as condições* $(C_1), (C_3)'$ *e* $(C_5)'$. *Então para todo* $\mathbf{x} \in \mathbb{R}^p$, *tem-se* $\mathbb{E}\left(\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}}{h_n}\right)\right) = \mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) + \frac{V_1(\mathbf{x})}{2}h_n^2 + o(h_n^2)$.

Lema 2.3.12 *Suponhamos que a sucessão de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$ *satisfaz as condições* $(C_1), (C_3)', (C_4)', (C_5)'$ *e* (C_8). *Então,*

(i) $\mathbb{E}\left(\mathbf{U}^2(\mathbf{x})\right) = \mathbf{F}_p(\mathbf{x}) - h_n\,\mathbf{V}_2(\mathbf{x}) + \frac{h_n^2}{2}\mathbf{V}_3(\mathbf{x})$;

(ii) $|\text{Var}\,(\mathbf{U}(\mathbf{x})) - (\mathbf{F}_p(\mathbf{x})(1-\mathbf{F}_p(\mathbf{x})) + h_n\,\mathbf{V}_2(\mathbf{x}))\,| = O(h_n^2)$.

Usemos, como anteriormente, as seguintes funções

- $\mathbf{I}_{nj}(\mathbf{x}) = \text{Cov}\left(\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{1,p}}{h_n}\right), \mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{j,p}}{h_n}\right)\right)$;
- $\mathbf{I}_j(\mathbf{x}) = \text{Cov}\left(\mathbb{1}_{(-\infty,\mathbf{x}]}(\mathbf{X}_{1,p}), \mathbb{1}_{(-\infty,\mathbf{x}]}(\mathbf{X}_{j,p})\right)$.

Lema 2.3.13 *Suponhamos que a sucessão estacionária de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$ *satisfaz as condições* $(C_1), (C_3)', (C_4)'$ *e* $(C_5)'$. *Então, para* $j > 1$, *e* $\mathbf{x} \in \mathbb{R}^p$, *tem-se*

(i) $\mathbf{I}_{nj}(\mathbf{x}) = \mathbf{I}_j(\mathbf{x}) + O(h_n^2) = \mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x}) + O(h_n^2)$;

(ii) *Para* $j > p-1$, $\mathbf{I}_j(\mathbf{x}) \leq \sum_{k=1}^{p}(p-k+1)\text{Cov}^{\,1/3}(X_1, X_{j+k}) +$
$+ \sum_{k=1}^{p-1}(p-k)\text{Cov}^{\,1/3}(X_1, X_{j-k+1})$.

Demonstração: A condição (ii) do enunciado do lema obtém-se aplicando o Lema de Lebowitz, Lema 1.2.3, apelando seguidamente à desigualdade (1.25) e, finalmente, usando o facto de a sucessão $\{X_n\}_{n\in\mathbb{N}}$ ser estacionária. De facto, de acordo com o Lema de Lebowitz

$$\text{Cov}\left(\mathbb{1}_{(-\infty,\mathbf{x}]}(\mathbf{X}_{1,p}), \mathbb{1}_{(-\infty,\mathbf{x}]}(\mathbf{X}_{j,p})\right) \leq$$
$$\leq \sum_{k=1}^{p}\sum_{i=1}^{p}\text{Cov}\left(\mathbb{1}_{(-\infty,x_k]}(X_k), \mathbb{1}_{(-\infty,x_{j+i}]}(X_{j+i})\right).$$

Se aplicarmos agora a desigualdade (1.25), obtemos

$$\mathrm{Cov}\left(\mathbb{1}_{(-\infty,x_k]}(X_k), \mathbb{1}_{(-\infty,x_{j+i}]}(X_{j+i})\right) \leq M\,\mathrm{Cov}^{1/3}(X_k, X_{j+i}),$$

pelo que

$$\mathbf{I}_j(\mathbf{x}) \leq M \sum_{k=1}^{p}\sum_{i=1}^{p} \mathrm{Cov}^{1/3}(X_k, X_{j+i}).$$

Como a sucessão $\{X_n\}_{n\in\mathbb{N}}$ é estacionária, vem, como pretendemos,

$$\mathbf{I}_j(\mathbf{x}) \leq M\sum_{k=1}^{p}(p-k+1)\mathrm{Cov}^{1/3}(X_1, X_{j+k})+\sum_{k=1}^{p-1}(p-k)\mathrm{Cov}^{1/3}(X_1, X_{j-k+1}). \blacksquare$$

Nota 2.3.14 *Se admitirmos que a sucessão de covariâncias dada por* $\mathrm{Cov}\,(X_1, X_{j+1})$ *é decrescente,* $\mathbf{I}_j(\mathbf{x}) \leq p^2\mathrm{Cov}^{1/3}(X_1, X_{j+1})$.

O estudo do Erro Quadrático Médio deste estimador e a minimização da sua expressão relativamente a h_n, conduz, tal como no caso bivariado, à obtenção da velocidade óptima de convergência para zero, da $O(n^{-1})$, da sucessão $\{h_n\}$ se considerarmos as condições (C_6) e (C_7).
Sejam, tal como anteriormente, as funções

$$\sigma^{\mathbf{2}}(\mathbf{x}) = \mathbf{F}_p(\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x}) + 2\sum_{j=2}^{\infty}\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right)$$

e

$$\mathbf{c}_n(\mathbf{x}) = 2\left(\sum_{j=n-p+1}^{\infty}\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right) + \right.$$
$$\left. +\frac{1}{n-p}\sum_{j=2}^{n-p}(j-1)\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right)\right).$$

O resultado que apresentamos em seguida, sem demonstração, dá-nos o comportamento do Erro Quadrático Médio do estimador.

Teorema 2.3.15 *Suponhamos que a sucessão associada de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$, *satisfaz as condições* $(C_1), (C_3)', (C_4)', (C_5)', (C_6), (C_7)$ *e* $(C_8)'$. *Então,*

$$(n-p)\text{MSE}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \sigma^{\mathbf{2}}(\mathbf{x}) - h_n\,\mathbf{V_2}(\mathbf{x}) + O(n\,h_n^2) + o(h_n + n\,h_n^2) - \mathbf{c}_n(\mathbf{x}).$$

Como corolário, apresentamos a especificação da velocidade de convergência optima da sucessão $\{h_n\}_{n\in\mathbb{N}}$.

Corolário 2.3.16 *Suponhamos que a sucessão associada de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$, *satisfaz as condições* $(C_1), (C_3)', (C_4)', (C_5)', (C_6)$, (C_7) *e* $(C_8)'$. *Nestas condições a escolha da velocidade de convergência optima, no sentido da minimização da expressão obtida para o Erro Quadrático Médio no teorema anterior, é da ordem de* n^{-1}.

Vejamos agora que, tal como para os casos univariado e bivariado sob associação e independente multivariado, esta velocidade de convergência é da ordem de $n^{-1/3}$ se forem usadas as condições do Teorema 2.3.7 relativas à sucessão de covariâncias $\text{Cov}\,(X_1, X_{j+1})$, $j \geq 1$.

Teorema 2.3.17 *Suponhamos que a sucessão associada de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$, *satisfaz as condições* $(C_1), (C_3)', (C_4)', (C_5)'$ *e* $(C_8)'$. *Consideremos que* $\{h_n\}_{n\in\mathbb{N}}$ *é tal que* $n\,h_n^4 \to 0$ *e que existe* τ, *com* $0 < \tau < 1$, *tal que a relação (2.35) se verifica. Suponhamos, ainda, que a sucessão de covariâncias* $\text{Cov}\,(X_1, X_{j+1})$ *é decrescente. Nestas condições,*

$$(n-p)\text{MSE}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \sigma^2(\mathbf{x}) - h_n\,\mathbf{V_2}(\mathbf{x}) + O(n\,h_n^4) + o(h_n + nh_n^4) - \mathbf{c}_n(\mathbf{x}).$$

Demonstração: A demonstração deste teorema usa essencialmente o facto de, como vimos no Lema 2.3.13, $\mathbf{I}_{nj}(\mathbf{x}) = \mathbf{I}_j(\mathbf{x}) + O(h_n^2)$. Ora, como foi referido na Nota 2.3.14, obtida uma majoração para $\mathbf{I}_j$ e, consequentemente, para $\mathbf{I}_{nj}$ é-nos permitido usar, tal como no caso bivariado, o

seguinte artifício:

$$\begin{aligned} |\mathbf{I}_{nj}(\mathbf{x}) - \mathbf{I}_j(\mathbf{x})| &= |\mathbf{I}_{nj}(\mathbf{x}) - \mathbf{I}_j(\mathbf{x})|^{\tau} |\mathbf{I}_{nj}(\mathbf{x}) - \mathbf{I}_j(\mathbf{x})|^{1-\tau} \leq \\ &\leq c^{\tau}\, h_n^{2\tau}\, p^{2(1-\tau)} | \left(\mathrm{Cov}^{\,1/3}(X_1, X_{j+1}) \right)^{1-\tau} | = \\ &= \tilde{c}\, h_n^{2\tau} \left| \left(\mathrm{Cov}^{\,1/3}(X_1, X_{j+1}) \right)^{1-\tau} \right|, \end{aligned}$$

onde $\tilde{c} = c^{\tau}\, p^{2(1-\tau)}$ é constante. O resto da demonstração segue usando o mesmo tipo de argumentos da demonstração do Teorema 2.3.7. ■
Mais uma vez, minimizando a expressão obtida para o Erro Quadrático Médio do estimador relativamente a h_n obtemos o seguinte resultado.

Corolário 2.3.18 *Suponhamos que a sucessão associada de variáveis aleatórias* $\{X_n\}_{n\in\mathbb{N}}$, *satisfaz as condições* $(C_1), (C_3)', (C_4)', (C_5)'$ *e* $(C_8)'$. *Consideremos que* $\{h_n\}_{n\in\mathbb{N}}$ *é tal que* $n\, h_n^4 \to 0$ *e que existe* τ, *com* $0 < \tau < 1$, *tal que a condição (2.35) se verifica. Suponhamos, ainda, que a sucessão de covariâncias* $\mathrm{Cov}\,(X_1, X_{j+1})$ *é decrescente. Nestas condições, a velocidade optima de convergência para zero da sucessão* $\{h_n\}_{n\in\mathbb{N}}$ *no sentido da minimização do Erro Quadrático Médio é* $O(n^{-1/3})$.

2.4 Convergência em distribuição de $\widehat{F}_{n,k}$

Consideremos, novamente, o estimador para a função de distribuição bivariada F_k, com $k = 1, \ldots, n-1$ fixo. O estimador $\widehat{F}_{n,k}$ introduzido e estudado nas secções anteriores gera, para cada $k = 1, ..., n-1$ fixo, um processo estocástico $\left\{\alpha_n(x,y), (x,y) \in \mathbb{R}^2\right\}$, definido por

$$\alpha_n(x,y) = \sqrt{n-k}\left[\widehat{F}_{n,k}(x,y) - \mathbb{E}\left(\widehat{F}_{n,k}(x,y)\right)\right], \quad (x,y) \in \mathbb{R}^2.$$

É um problema importante obter a convergência fraca deste processo. De facto, sempre que pretendemos fazer inferência estatística em que se utilizam estatísticas que envolvam estes estimadores é essencial conhecer a distribuição dos vectores de dimensão finita, úteis por si só, uma vez

que podem ser usados na construção de intervalos de confiança.
Pretendemos, então, caracterizar a lei limite dos vectores aleatórios de dimensão s, fixa, cujas entradas são

$$\sqrt{n-k}\left[\widehat{F}_{n,k}(x_i,y_i) - \mathbb{E}\left(\widehat{F}_{n,k}(x_i,y_i)\right)\right],\ i=1,\ldots,s,$$

que, por simplicidade, denotamos por $\alpha_n(x_i,y_i)$.
O teorema principal desta secção fornece condições suficientes para a convergência de um vector aleatório $(\alpha_n(x_1,y_1),\ldots,\alpha_n(x_s,y_s))$, para um vector Gaussiano centrado com matriz de covariância especificada. Porém, antes de passarmos ao seu enunciado, defina-se uma função que servirá para descrever o limite a obter.

Definição 2.4.1 *Seja ς^2 uma função de domínio $\mathbb{R}^2\times\mathbb{R}^2$, tal que para todo $x,y,w,v\in\mathbb{R}$,*

$$\begin{aligned}\varsigma^2(x,y,w,v) &= F_k\Big(\min\{x,w\},\min\{y,v\}\Big) - F_k(x,y)F_k(w,v) + \\ &\quad +2\times\sum_{j=2}^{\infty}\big(F_{k,j}(x,y,w,v) - F_k(x,y)F_k(w,v)\big).\end{aligned}$$

As condições que garantem a convergência dos vectores de dimensão finita são as que garantem a consistência de $\widehat{F}_{n,k}$.

Teorema 2.4.2 *Suponhamos que a sucessão associada e estritamente estacionária $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias reais satisfaz as condições $(C_1),(C_3),(C_4),(C_5),(C_6),(C_7)$ e (C_8).*
Então, dados $s\in\mathbb{N}$, e $x_1,\ldots,x_s,y_1,\ldots,y_s\in\mathbb{R}$, o vector aleatório $\big(\alpha_n(x_1,y_1),\ldots,\alpha_n(x_s,y_s)\big)$ converge em distribuição para um vector aleatório Gaussiano centrado com matriz de covariância dada por

$$\Sigma = \begin{bmatrix} \varsigma^2(x_1,y_1,x_1,y_1) & \varsigma^2(x_1,y_1,x_2,y_2) & \ldots & \varsigma^2(x_1,y_1,x_s,y_s) \\ \varsigma^2(x_2,y_2,x_1,y_1) & \varsigma^2(x_2,y_2,x_2,y_2) & \ldots & \varsigma^2(x_2,y_2,x_s,y_s) \\ \ldots & \ldots & \ldots & \ldots \\ \varsigma^2(x_s,y_s,x_1,y_1) & \varsigma^2(x_s,y_s,x_2,y_2) & \ldots & \varsigma^2(x_s,y_s,x_s,y_s) \end{bmatrix}.$$

Nota 2.4.3 *Note-se que esta matriz de covariância é simétrica. Temos, de facto,* $\varsigma^2(x_i, y_i, x_j, y_j) = \varsigma^2(x_j, y_j, x_i, y_i)$, *para todo* $i, j = 1, \ldots, s$.

A demonstração deste resultado utiliza o Teorema de Cramér-Wold (ver James [21], por exemplo) no qual se garante que a convergência em distribuição de um vector aleatório equivale à convergência em distribuição de todas as combinações lineares das coordenadas do vector.
Assim, segundo este teorema basta provar que para todo $(c_1, \ldots, c_s) \in \mathbb{R}^s$ se tem,

$$\sum_{q=1}^{s} c_q \, \alpha_n(x_q, y_q) \underset{D}{\longrightarrow} \sum_{q=1}^{s} c_q \, \mathbb{B}(x_q, y_q), \tag{2.36}$$

onde $\underset{D}{\rightarrow}$ significa convergência em distribuição e em que $(\mathbb{B}(x_1, y_1), \ldots, \mathbb{B}(x_q, y_q))$ é um vector Gaussiano centrado com matriz de covariância Σ. Se denotarmos por $Z_n = \sum_{q=1}^{s} c_q \, \alpha_n(x_q, y_q)$ e por $\mathbb{Z} = \sum_{q=1}^{s} c_q \, \mathbb{B}(x_q, y_q)$, provar a condição (2.36) é equivalente a provar que

$$\left| \mathbb{E}\left(\mathrm{e}^{itZ_n}\right) - \mathrm{e}^{-\frac{t^2 a}{2}} \right| \to 0, \tag{2.37}$$

onde $a = \mathrm{Var}\,(\mathbb{Z})$.
Com o objectivo de provar (2.37) vamos usar uma técnica, já utilizada por Jacob e Oliveira [19], 1999, que consiste em aproximar a soma $Z_n = \sum_{q=1}^{s} c_q \, \alpha_n(x_q, y_q)$, por uma soma de $m+1$ blocos e, via *Desigualdade de Newman*, tratá-los como variáveis independentes.
A demonstração do teorema é bastante longa pelo que, para que seja de mais fácil leitura, será dividida em diversos lemas.
Consideremos, então, $1 < r < n - k$ e $m = \left[\frac{n-k}{r}\right]$, onde $[x]$ significa o maior inteiro menor ou igual a x. Nestas condições, $n - k = mr + b$,

com $0 < b < r$. Podemos escrever,

$$\begin{aligned}\alpha_n(x,y) = \frac{1}{\sqrt{n-k}} \times \\ \times \left\{ \sum_{i=1}^{r} \left(U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right) - \mathbb{E}\left(U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right)\right)\right) + \right. \\ + \sum_{i=r+1}^{2r} \left(U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right) - \mathbb{E}\left(U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right)\right)\right) + \\ + \cdots + \\ + \sum_{i=(j-1)r}^{mr} \left(U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right) - \mathbb{E}\left(U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right)\right)\right) + \\ \left. + \sum_{i=mr+1}^{n-k} \left(U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right) - \mathbb{E}\left(U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right)\right)\right) \right\}.\end{aligned}$$

Se definirmos,

$$T_{n,i}(x,y) = U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right) - \mathbb{E}\left(U\left(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\right)\right),$$

uma vez que,

$$Z_n = \sum_{q=1}^{s} c_q \alpha_n(x_q, y_q),$$

podemos escrever,

$$Z_n = \frac{1}{\sqrt{n-k}} \sum_{q=1}^{s} c_q \sum_{i=1}^{n-k} T_{n,i}(x_q, y_q).$$

Desta forma, Z_n escreve-se como uma combinação linear de m blocos do tipo

$$Y_j^r(x,y) = \frac{1}{\sqrt{r}} \sum_{i=(j-1)r+1}^{jr} T_{n,i}(x,y), \qquad \text{para} \quad j = 1, \ldots, m,$$

de comprimento r mais um bloco de comprimento inferior a r. Começamos por mostrar que podemos considerar apenas a soma dos m blocos

de comprimento r. Este será o primeiro passo na demonstração do Teorema 2.4.2. Para tal, necessitamos de mais alguma notação. Definam-se, ainda,

$$W_j^r = \sum_{q=1}^{s} c_q Y_j^r(x_q, y_q), \quad Z_{n,i} = \sum_{q=1}^{s} c_q T_{n,i}(x_q, y_q).$$

Se considerarmos,

$$Z_{mr}^* = \frac{1}{\sqrt{m}} \sum_{j=1}^{m} c_q \, Y_j^r(x_q, y_q) = \frac{1}{\sqrt{m}} \sum_{j=1}^{m} W_j^r,$$

então

$$Z_{mr}^* = \frac{1}{\sqrt{mr}} \sum_{i=1}^{mr} Z_{n,i}$$

que substitui a soma até $n-k$ por uma soma com um número de termos múltiplo de r.

Antes de iniciar a demonstração do Teorema, vamos descrever o comportamento assimptótico das covariâncias entre as funções α_n em pontos diferentes.

Lema 2.4.4 *Suponhamos que a sucessão $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias reais satisfaz as condições $(C_1), (C_3), (C_4), (C_5), (C_6), (C_7)$ e (C_8). Então para todo $x, y, w, v \in \mathbb{R}$,*

$$\text{Cov } (\alpha_n(x,y), \alpha_n(w,v)) \longrightarrow \varsigma^2(x,y,w,v).$$

Demonstração: De acordo com a definição de α_n e usando a estacionaridade das variáveis aleatórias, podemos escrever,

$$\begin{aligned}
&\text{Cov } (\alpha_n(x,y), \alpha_n(w,v)) = \\
&\quad \text{Cov } \left(U\left(\frac{x-X_1}{h_n}, \frac{y-X_{k+1}}{h_n}\right), U\left(\frac{w-X_1}{h_n}, \frac{v-X_{k+1}}{h_n}\right)\right) + \\
&\quad + \frac{2}{n-k} \sum_{j=2}^{n-k} (n-k-j+1) \times \qquad (2.38) \\
&\quad \times \text{Cov } \left(U\left(\frac{x-X_1}{h_n}, \frac{y-X_{k+1}}{h_n}\right), U\left(\frac{w-X_1}{h_n}, \frac{v-X_{k+j+1}}{h_n}\right)\right)
\end{aligned}$$

Como já vimos, para $j = 1, \ldots, n-k$,

$$\begin{aligned}&\text{Cov}\left(U\left(\frac{x-X_1}{h_n}, \frac{y-X_{k+j+1}}{h_n}\right), U\left(\frac{w-X_1}{h_n}, \frac{v-X_{k+1}}{h_n}\right)\right) = \\ &\quad = F_{k,j}(x,y,w,v) - F_k(x,y)F_k(w,v) + O(h_n^2).\end{aligned}$$

Usando esta identidade em (2.38) obtemos,

$$\begin{aligned}\text{Cov}\,(\alpha_n(x,y), \alpha_n(w,v)) &= F_{k,j}(x,y,w,v) - F_k(x,y)F_k(w,v) + O(h_n^2) + \\ &+ \frac{2}{n-k}\sum_{j=2}^{n-k}(n-k+j+1)\big(F_{k,j}(x,y,w,v) - \\ &\quad -F_k(x,y)F_k(w,v) + O(h_n^2)\big).\end{aligned}$$

Ora, o somatório envolvido nesta última expressão pode ser desdobrado em

$$\begin{aligned}&\sum_{j=2}^{n-k}\big(F_{k,j}(x,y,w,v) - F_k(x,y)F_k(w,v)\big) - \\ &\quad -\frac{1}{n-k}\sum_{j=2}^{n-k}(j-1)\big(F_{k,j}(x,y,w,v)F_k(x,y)F_k(w,v)\big) + O(nh_n^2).\end{aligned}$$

Podemos escrever

$$\begin{aligned}&F_{k,j}(x,y,w,v) - F_k(x,y)F_k(w,v) = \\ &\quad = \text{Cov}\left(1\!\!1_{(-\infty,x]\times(-\infty,y]}(X_1, X_{k+1}), 1\!\!1_{(-\infty,w]\times(-\infty,v]}(X_j, X_{k+j})\right).\end{aligned}$$

Como já vimos antes, de acordo com Lebowitz (1972), obtemos,

$$\begin{aligned}&\text{Cov}\left(1\!\!1_{(-\infty,x]\times(-\infty,y]}(X_1, X_{k+1}), 1\!\!1_{(-\infty,w]\times(-\infty,v]}(X_j, X_{k+j})\right) \leq \\ &\leq \text{Cov}\left(1\!\!1_{(-\infty,x]}(X_1), 1\!\!1_{(-\infty,w]}(X_j)\right) + \text{Cov}\left(1\!\!1_{(-\infty,x]}(X_1), 1\!\!1_{(-\infty,v]}(X_{k+j})\right) + \\ &\quad + \text{Cov}\left(1\!\!1_{(-\infty,y]}(X_{k+1}), 1\!\!1_{(-\infty,w]}(X_j)\right) + \text{Cov}\left(1\!\!1_{(-\infty,y]}(X_{k+1}), 1\!\!1_{(-\infty,v]}(X_{k+j})\right).\end{aligned}$$

Finalmente, uma vez que,

$$\text{Cov}\left(1\!\!1_{(-\infty,x]}(X_i), 1\!\!1_{(-\infty,y]}(X_j)\right) \leq M\text{Cov}^{\,1/3}(X_i, X_j),$$

onde M é a constante especificada na condição (1.25), obtemos

$$\begin{aligned}
&\frac{1}{n-k}\sum_{j=2}^{n-k}(j-1)\big(F_{k,j}(x,y,w,v)-F_k(x,y)F_k(w,v)\big)\leq \\
&\quad\leq \frac{1}{n-k}\sum_{j=2}^{n-k} j\,\mathrm{Cov}\left(\mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_1,X_{k+1}),\mathbb{1}_{(-\infty,w]\times(-\infty,v]}(X_j,X_{k+j})\right)\leq \\
&\quad\leq \frac{M}{n-k}\sum_{j=2}^{n-k} j\left(2\,\mathrm{Cov}^{1/3}(X_1,X_j)+Cov^{1/3}(X_1,X_{k+j})+\right. \\
&\quad\quad \left.+\mathrm{Cov}^{1/3}(X_1,X_j)+Cov^{1/3}(X_{k+1},X_j)\right)
\end{aligned}$$

que, de acordo com a condição (C_7), converge para 0, o que demonstra o lema. ■

Estamos agora em condições de estudar o comportamento assimptótico da variância de Z_n.

Lema 2.4.5 *Suponhamos que a sucessão estritamente estacionária e associada* $\{X_n\}_{n\in\mathbb{N}}$ *de variáveis aleatórias reais satisfaz as condições* $(C_1),(C_3),(C_4),(C_5),(C_6),(C_7)$ *e* (C_8). *Seja*

$$\varsigma^2=\sum_{q=1}^{s}c_q^2\,\varsigma^2(x_q,y_q,x_q,y_q)+2\sum_{q=1}^{s-1}\sum_{l=q+1}^{s}c_q\,c_l\,\varsigma^2(x_q,y_q,x_q,y_q). \quad (2.42)$$

Então

$$\mathrm{Var}\,(Z_n)\longrightarrow\varsigma^2.$$

Demonstração: Escrevamos Z_n de acordo com a notação introduzida. Por definição de variância,

$$\begin{aligned}
\mathrm{Var}\left(Z_n\right) &= \frac{1}{n-k}\mathrm{Var}\left(\sum_{q=1}^{s} c_q \sum_{i=1}^{n-k} T_{n,i}(x_q, y_q)\right) = \\
&= \frac{1}{n-k}\sum_{q=1}^{s}\sum_{l=1}^{s} c_q\, c_l\, \mathrm{Cov}\left(\sum_{i=1}^{n} T_{n,i}(x_q, y_q),\, \sum_{i=1}^{n-k} T_{n,i}(x_l, y_l)\right) = \\
&= \frac{1}{n-k}\left[\sum_{q=1}^{s} c_q^2\, \mathrm{Cov}\left(\sum_{i=1}^{n-k} T_{n,i}(x_q, y_q), \sum_{i=1}^{n-k} T_{n,i}(x_q, y_q)\right) + \right. \\
&\quad \left. + 2\sum_{q=1}^{s-1}\sum_{l=q+1}^{s} c_q\, c_l\ \mathrm{Cov}\left(\sum_{i=1}^{n-k} T_{n,i}(x_q, y_q), \sum_{i=1}^{n-k} T_{n,i}(x_l, y_l)\right)\right].
\end{aligned}$$

Mas

$$\frac{1}{n-k}\mathrm{Cov}\left(\sum_{i=1}^{n-k} T_{n,i}(x_q, y_q),\, \sum_{i=1}^{n-k} T_{n,i}(x_l, y_l)\right) = \mathrm{Cov}\left(\alpha_n(x_q, y_q), \alpha_n(x_l, y_l)\right).$$

De acordo com o Lema 2.4.4,

$$\mathrm{Cov}\left(\alpha_n(x_q, y_q), \alpha_n(x_l, y_l)\right) \longrightarrow \varsigma^2(x_q, y_q, x_l, y_l),$$

sendo, agora, o resultado imediato. ■

Vejamos, a seguir, a variância de cada bloco Y_j^r, com r fixo e, também a variância das combinações lineares entre os elementos dentro de cada bloco W_j^r. Note-se que, por estacionaridade da sucessão $\{X_n\}_{n\in\mathbb{N}}$, $\mathrm{Var}\,(Y_j^r) = \mathrm{Var}\,(Y_1^r)$, para cada $j = 1, \ldots, mr$.

$$\begin{aligned}
&\mathrm{Var}\ (Y_1^r(x, y)) = \mathrm{Cov}\left(\frac{1}{\sqrt{r}}\sum_{i=1}^{r} T_{n,i}(x, y), \frac{1}{\sqrt{r}}\sum_{i'=1}^{r} T_{n,i'}(x, y)\right) = \\
&= \frac{1}{r}\sum_{i,i'=1}^{r} \mathrm{Cov}\left(U\left(\frac{x - X_i}{h_n}, \frac{x - X_{k+i}}{h_n}\right), U\left(\frac{x - X_{i'}}{h_n}, \frac{x - X_{k+i'}}{h_n}\right)\right),
\end{aligned}$$

cujo comportamento está descrito no Lema 2.2.3. Temos então,

$$\text{Var}\,(Y_1^r(x,y)) = \frac{1}{r}\sum_{i,i'=1}^{r}\Big(F_{k,|i'-i+1|}(x,y,x,y) - F_k^2(x,y)\Big) + O(rh_n^2). \tag{2.44}$$

Quanto à variância de W_j^r, tendo em conta a sua expressão,

$$\begin{aligned}\text{Var}\,(W_j^r) &= \sum_{q,q'=1}^{s} c_q c_{q'} \text{Cov}\left[Y_j^r(x_q,y_q), Y_j^r(x_q',y_q')\right] = \\ &= \sum_{q,q'=1}^{s} c_q c_{q'} \frac{1}{r}\sum_{i,i'=1}^{r}\left(F_{k,|i'-i+1|}(x_q,y_q,x_{q'},y_{q'}) - F_k(x_q,y_q)F_k(x_{q'},y_{q'})\right) + \\ &\quad + O(rh_n^2).\end{aligned}$$

Estamos agora em condições de demonstrar o que foi atrás referido como sendo o primeiro passo para a demonstração do teorema principal desta secção. Mostramos que, para efeitos de prova de convergência em distribuição, podemos substituir a soma dos $n-k$ termos em Z_n pela soma Z_{mr}^* que envolve apenas a soma dos m blocos W_j^r.

Lema 2.4.6 *Suponhamos que a sucessão estritamente estacionária e associada $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias reais satisfaz as condições $(C_1), (C_3), (C_4), (C_5), (C_6), (C_7)$ e (C_8). Consideremos r fixo. Então*

$$\left|\mathbb{E}e^{itZ_n} - \mathbb{E}e^{itZ_{mr}^*}\right| \longrightarrow 0.$$

Demonstração: Usemos a desigualdade de Holder e o facto de, para cada $x, y \in \mathbb{R}$, $\left|e^{ix} - e^{iy}\right| \leq |x-y|$, para majorar o valor absoluto da diferença entre a função característica de Z_n e a função característica de Z_{mr}^*.

$$\left|\mathbb{E}e^{itZ_n} - \mathbb{E}e^{itZ_{mr}^*}\right| \leq |t|\mathbb{E}\,|Z_n - Z_{mr}^*| \leq |t|\,\text{Var}^{1/2}(Z_n - Z_{mr}^*). \tag{2.45}$$

Vejamos a majoração da variância desta diferença.

$$\begin{aligned}
\mathrm{Var}\,(Z_n - Z^*_{mr}) &= \mathbb{E}\Big(\frac{1}{\sqrt{mr}}\sum_{i=1}^{mr} Z_{n,i} - \frac{1}{\sqrt{n-k}}\sum_{i=1}^{n-k} Z_{n,i}\Big)^2 = \\
&= \mathbb{E}\Big(\frac{1}{\sqrt{mr}}\sum_{i=1}^{n-k} Z_{n,i} - \frac{1}{\sqrt{mr}}\sum_{i=mr+1}^{n-k} Z_{n,i} - \frac{1}{\sqrt{n-k}}\sum_{i=1}^{n-k} Z_{n,i}\Big)^2 = \\
&= \mathbb{E}\Big((\frac{1}{\sqrt{mr}}\sum_{i=mr+1}^{n-k} Z_{n,i}) + \Big(\frac{1}{\sqrt{n-k}} - \frac{1}{\sqrt{mr}}\Big)\sum_{i=1}^{n-k} Z_{n,i}\Big)^2 \leq \\
&\leq \frac{2}{mr}\mathbb{E}\left(\sum_{i=mr+1}^{n-k} Z_{n,i}\right)^2 + 2\left(\frac{1}{\sqrt{n-k}} - \frac{1}{\sqrt{mr}}\right)^2 \mathbb{E}\left(\sum_{i=1}^{n-k} Z_{n,i}\right)^2.
\end{aligned}$$

Atendendo a que U é uma função de distribuição, segue-se que as variáveis $Z_{n,i}$ são limitadas. De facto,

$$|Z_{n,i}| = \left|\sum_{q=1}^{s} c_q T_{n,i}(x_q, y_q)\right| \leq \sum_{q=1}^{s} |c_q T_{n,i}(x_q, y_q)| \leq 2\sum_{q=1}^{s} |c_q|.$$

Temos então,

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{\sqrt{mr}}\sum_{i=1}^{mr} Z_{n,i} - \frac{1}{\sqrt{n-k}}\sum_{i=1}^{n-k} Z_{n,i}\right)^2 \leq \\
&\quad \leq \frac{8}{mr}(n-k-mr)^2\left(\sum_{q=1}^{s}|c_q|\right)^2 + 2\left(1 - \sqrt{\frac{n-k}{mr}}\right)^2 \mathrm{Var}\,(Z_n).
\end{aligned}$$

De (2.45), obtemos, utilizando esta desigualdade,

$$\begin{aligned}
&\left|\mathbb{E}e^{itZ_n} - \mathbb{E}e^{itZ^*_{mr}}\right| \leq 2\sqrt{2}\,|t| \times \\
&\qquad \times \left[\frac{8r^2}{n-k}\left(\sum_{q=1}^{s}|c_q|\right)^2 + 2\left(1 - \sqrt{\frac{n-k}{mr}}\right)^2 \mathrm{Var}\,(Z_n)\right]^{\frac{1}{2}},
\end{aligned}$$

uma vez que $0 \leq n-k-mr < r$. Ora, quando $n \to \infty$, tem-se $\mathrm{Var}(Z_n) \to \varsigma^2 (< \infty)$ e, além disso, $\frac{mr}{n-k} \to 1$. Resulta, assim, da desigualdade anterior que

$$\left| \mathbb{E} e^{itZ_n} - \mathbb{E} e^{itZ^*_{mr}} \right| \longrightarrow 0. \blacksquare$$

De acordo com este resultado, no que diz respeito à convergência em distribuição, Z_n pode ser substituída por Z^*_{mr}, que é a soma dos m blocos $W^r_1, W^r_2, \ldots, W^r_m$. Estas variáveis aleatórias são dependentes e cada uma delas é uma combinação linear das variáveis Y^r_j, que são funções decrescentes das variáveis aleatórias originais, sendo, por isso, associadas.
Antes de prosseguirmos vejamos uma versão da *Desigualdade de Newman*, adaptada ao tipo de variáveis que necessitamos de tratar.

Lema 2.4.7 (Jacob, Oliveira [19], 1999) *Seja* $\{Y_n\}_{n\in\mathbb{N}}$ *uma sucessão associada de variáveis aleatórias reais,* $r \in \mathbb{N}$ *e* $\alpha_0, \ldots, \alpha_r \in \mathbb{R}$. *Para cada* $n \in \mathbb{N}$, *defina-se*

$$X_n = \sum_{k=0}^{r} \alpha_k Y_{k+n} \qquad e \ \bar{X}_n = \sum_{k=0}^{r} |\alpha_k| Y_{k+n}.$$

Então, para todo $u_1, \ldots, u_r \in \mathbb{R}$, *tem-se que,*

$$\left| \mathbb{E}\left(\mathrm{e}^{\sum_{j=1}^{m} u_j X_j} \right) - \prod_{j=1}^{m} \mathbb{E}\left(\mathrm{e}^{iu_j X_j} \right) \right| \leq 2 \sum_{k\neq l} \left| u_k u_l \mathrm{Cov}\left(\bar{X}_k, \bar{X}_l \right) \right|.$$

Demonstração: Para cada $n \in \mathbb{N}$ definam-se as funções $f_n(y_1, y_2, \ldots) = \sum_{k=0}^{r} \alpha_k y_{k+n}$ e $\bar{f}_n(y_1, y_2, \ldots) = \sum_{k=0}^{r} |\alpha_k| y_{k+n}$. Então,

$$f_n(y_1, y_2, \ldots) + \bar{f}_n(y_1, y_2, \ldots) = \sum_{k=0}^{r} (\alpha_k + |\alpha_k|) y_{k+n}$$

e

$$f_n(y_1, y_2, \ldots) - \bar{f}_n(y_1, y_2, \ldots) = \sum_{k=0}^{r} (|\alpha_k| - \alpha_k) y_{k+n}.$$

Note-se que f_n não é uma função monótona (coordenada a coordenada) das variáveis $y_n, \ldots, y_{r+n}$, uma vez que os coeficientes α_k intervenientes na respectiva combinação linear não têm todos o mesmo sinal. No entanto, a soma e a diferença de f_n com $\bar{f}_n$, conduz a combinações lineares com coeficientes positivos das variáveis $y_n, \ldots, y_{r+n}$. Assim, pode-se aplicar a *Desigualdade de Newman* a estas funções, sendo o resultado do lema, agora, imediato. ■

Posto isto, estamos em condições de aplicar esta desigualdade aos blocos $W_1^r, W_2^r, \ldots, W_m^r$, com o objectivo de controlar o erro cometido quando os substituímos por outros idênticos, quanto à distribuição, mas independentes.

Lema 2.4.8 *Suponhamos que a sucessão $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias reais satisfaz as condições $(C_1), (C_2), (C_3), (C_4), (C_5), (C_6), (C_7)$ e (C_8). Então*

$$\left| \mathbb{E}e^{itZ_{mr}^*} - \prod_{j=1}^{m} \mathbb{E}e^{\frac{it}{\sqrt{m}}W_j^r} \right| \leq$$

$$\leq 2t^2 \left[\mathrm{Var}\left(\frac{1}{\sqrt{mr}} \sum_{j=1}^{mr} T_{n,i} \right) - \mathrm{Var}\left(Y_1^r\right) \right] \sum_{q,q'=1}^{s} c_q c_q'.$$

Demonstração: Começamos por aplicar a desigualdade do lema 2.4.7 às variáveis aleatórias $W_j^r = \sum_{q=1}^{s} c_q Y_j^r(x_q, y_q)$.

$$\left| \mathbb{E}e^{itZ_{mr}^*} - \prod_{j=1}^{m} \mathbb{E}e^{\frac{it}{\sqrt{m}}W_j^r} \right| \leq$$

$$\leq 2 \sum_{\substack{i,j=1 \\ i\neq j}}^{m} \frac{t^2}{m} \left| \mathrm{Cov} \left\{ \sum_{q=1}^{s} |c_q| Y_i^r(x_q, y_q), \sum_{q=1}^{s} |c_q| Y_j^r(x_{q'}, y_{q'}) \right\} \right| =$$

$$= \frac{2\,t^2}{m} \sum_{i,j=1_{i\neq j}}^{m} \sum_{q,q'=1}^{s} |c_q c_{q'}| \mathrm{Cov}\left\{Y_i^r(x_q,y_q), Y_j^r(x_{q'},y_{q'})\right\} =$$

$$= 2\,t^2 \sum_{q,q'=1}^{s} |c_q c_{q'}| \sum_{i,j=1_{i\neq j}}^{m} \mathrm{Cov}\left\{\frac{1}{\sqrt{m}} Y_i^r(x_q,y_q), \frac{1}{\sqrt{m}} Y_j^r(x_{q'},y_{q'})\right\} =$$

$$= 2\,t^2 \sum_{q,q'=1}^{s} |c_q c_{q'}| \left[\mathrm{Var}\left(\frac{1}{\sqrt{mr}} \sum_{j=1}^{mr} T_{n,i}\right) - \sum_{j=1}^{m} \mathrm{Var}\left(\frac{1}{\sqrt{m}} Y_j^r\right)\right] =$$

$$= 2\,t^2 \sum_{q,q'=1}^{s} |c_q c_{q'}| \left(\mathrm{Var}\left(\frac{1}{\sqrt{mr}} \sum_{j=1}^{mr} T_{n,i}\right) - \mathrm{Var}\left(Y_1^r\right)\right),$$

devido à estacionaridade da sucessão das variáveis aleatórias.■

O próximo passo é a demonstração de um Teorema do Limite Central para a cópia independente das variáveis W_j^r. Com o objectivo de não complicar mais a notação, estas variáveis aleatórias são denotadas também por W_j^r. Portanto, no próximo lema e apenas neste lema, supomos as variáveis aleatórias independentes. Em primeiro lugar definamos

$$\varsigma_r^2 = \sum_{q,q'=1}^{s} c_q c_{q'} \frac{1}{r} \sum_{i,i'=1}^{r} \left(F_{k,|i'-i+1|}(x_q,y_q,x_{q'},y_{q'}) - F_k(x_q,y_q)F_k(x_{q'},y_{q'})\right).$$

Lema 2.4.9 *Suponhamos que a sucessão $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias reais satisfaz as condições $(C_1),(C_2),(C_3),(C_4),(C_5),(C_6),(C_7)$ e (C_8). Consideremos r fixo. Nestas condições,*

$$\left|\prod_{j=1}^{m} \mathbb{E}e^{\frac{it}{\sqrt{m}} W_j^r} - e^{-\frac{t^2\varsigma_r^2}{2}}\right| \longrightarrow 0.$$

Demonstração: Vamos aplicar a condição de Lindberg às variáveis aleatórias $\frac{1}{\sqrt{m}} W_j^r, j = 1,\ldots,m$, isto é, vamos provar que,

$$\frac{1}{\mathrm{Var}\left(\frac{1}{\sqrt{m}} \sum_{j=1}^{m} W_j^r\right)} \sum_{j=1}^{m} \int_{\left\{\left|\frac{1}{\sqrt{m}} W_j^r\right| \geq \epsilon\varsigma_r^2\right\}} \frac{1}{m} \left(W_j^r\right)^2 d\mathcal{P} \longrightarrow 0.$$

Como $\mathrm{Var}\left(\frac{1}{\sqrt{m}}\sum_{j=1}^{m} W_j^r\right)$ converge conforme se verificou atrás, basta provar que

$$\sum_{j=1}^{m} \int_{\left\{\left|\frac{1}{\sqrt{m}}W_j^r\right| \geq \epsilon\varsigma_r^2\right\}} \frac{1}{m}\left(W_j^r\right)^2 d\mathcal{P} \longrightarrow 0.$$

Mas, de acordo com a definição de W_j^r, esta expressão é igual a

$$\sum_{j=1}^{m} \int_{\left\{\left|\frac{1}{\sqrt{mr}}\sum_{i=(j-1)r+1}^{jr} T_n,i\right| \geq \epsilon\varsigma_r^2\right\}} \frac{1}{mr}\left(\sum_{i=(j-1)r+1}^{jr} T_n, i\right)^2 d\mathcal{P}.$$

Usando agora o *Lema 4* em Utev [50], 1990, o integral anterior pode ser majorado por

$$\sum_{j=1}^{m} \frac{1}{m} \sum_{i=(j-1)r+1}^{jr} \int_{\left\{\left|T_n,i\right| \geq \epsilon\frac{\varsigma_r^2\sqrt{mr}}{r}\right\}} T_{n,i}^2 \, d\mathcal{P}.$$

Ora, para cada $i = 1, \ldots, mr$ e $n \in \mathbb{N}$, $T_{n,i} \leq 2\sum_{q=1}^{s} c_q$ que é finito. Assim, desde que n seja suficientemente grande, o conjunto de integração é vazio. Logo, para estes valores de n cada um dos integrais no último somatório é igual a zero, o que prova a convergência para zero pretendida. ■

Apresentamos, finalmente, a demonstração do Teorema 2.4.2 que, conforme foi anteriormente referido, é equivalente a provar a condição (2.37).

Demonstração: (do Teorema 2.4.2) A prova do Teorema consiste em verificar que $\left|\mathbb{E}e^{itZ_n} - e^{-\frac{t^2 a}{2}}\right| \longrightarrow 0$, onde $a = \sum_{q,q'=1}^{s} c_q c_{q'} \varsigma^2(x_q, y_q, x_{q'}, y_{q'})$. A desigualdade triangular garante que,

$$\begin{aligned}\left|\mathbb{E}e^{itZ_n} - e^{-\frac{t^2 a}{2}}\right| \quad &\leq \quad \left|\mathbb{E}e^{itZ_n} - \mathbb{E}e^{itZ_{mr}^*}\right| + \left|\mathbb{E}e^{itZ_{mr}^*} - \prod_{j=1}^{m} \mathbb{E}e^{\frac{it}{\sqrt{m}}W_j^r}\right| + \\ &+ \quad \left|\prod_{j=1}^{m} \mathbb{E}e^{\frac{it}{\sqrt{m}}W_j^r} - e^{-\frac{t^2\varsigma_r^2}{2}}\right| + \left|e^{-\frac{t^2\varsigma_r^2}{2}} - e^{-\frac{t^2 a}{2}}\right|.\end{aligned}$$

Supondo, por agora, que r está fixo, segue-se, de acordo com os lemas anteriores,

$$\limsup_{m\to\infty} \left| \mathbb{E}e^{itZ_n} - e^{-\frac{t^2 a}{2}} \right| \leq$$
$$\leq 2t^2 \left[\mathrm{Var}\left(\frac{1}{\sqrt{mr}} \sum_{j=1}^{mr} T_{n,i} \right) - \mathrm{Var}\,(Y_1^r) \right] \sum_{q,q'=1}^{s} c_q c_{q'} + \left| e^{-\frac{t^2 \varsigma_r^2}{2}} - e^{-\frac{t^2 a}{2}} \right|.$$

Fazendo agora $r \to +\infty$, este majorante converge para zero, tendo em conta (2.44) e a estacionaridade da sucessão de variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$.

Assim, $\limsup_{m\to\infty} \left| \mathbb{E}e^{itZ_n} - e^{-\frac{t^2 a}{2}} \right| = 0$, quando r tende para infinito. ■

Capítulo 3

O operador de covariância

Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão de variáveis aleatórias num espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. Seja F, contínua, a função de distribuição marginal das variáveis $X_n, n \geq 1$.

Considerando, como no primeiro capítulo, $\mathbb{F}_n$ a função de distribuição empírica de $X_1, \ldots, X_n$,

$$\int_{\mathbb{R}} n\big(\mathbb{F}_n(t) - F(t)\big)^2 \, dF(t),$$

define uma variável aleatória, função da amostra $(X_1, \ldots, X_n)$, denotada por W_n^2 e denominada *estatística de Cramèr-von Mises* (ver, por exemplo Shorack e Wellner [44]). Esta estatística, usada muito frequentemente em testes de ajustamento, não é mais que a norma $L^2(\mathbb{R})$ do processo empírico. Ora, considerando F concentrada em $[0,1]$, sendo $\big\{\beta_n(t), t \in [0,1]\big\}$ o n-ésimo processo empírico construído com base no vector associado $(X_1, \ldots, X_n)$, definido, para cada $t \in [0,1]$, por $\beta_n(t) = \sqrt{n}\Big(\mathbb{F}_n(t) - F(t)\Big)$, designando por $\|\cdot\|_2$ a norma em $L^2([0,1])$, temos $W_n^2 = \big\|\beta_n\big\|_2^2$.

Quando estamos interessados em caracterizar o nível de significância, α, de um teste não paramétrico que envolva a estatística de Cramèr-von

Mises, necessitamos de conhecer um valor aproximado para $P(W_n^2 > \epsilon)$, de modo a ser possível obter o valor crítico do teste, $\epsilon = \epsilon(\alpha, n)$. Temos então, de acordo com o que foi dito anteriormente, $P(W_n^2 > \epsilon) = P\big(\|\beta_n\|_2^2 > \epsilon\big)$.

Sendo $F_k(x,y) = P(X_1 \leq x, X_{k+1} \leq y)$ e supondo que $\mathrm{Cov}\,(X_1, X_n) = O(n^{-b})$, $b > 3$, condição suficiente para a convergência do processo empírico em $L^2[0,1]$ conforme Oliveira e Suquet [35], 1995, vimos já que a sucessão de funções aleatórias, $\big\{\beta_n(t)\big\}_{n\in\mathbb{N}}$, converge em lei para um processo Gaussiano $\big\{\mathbb{B}(t), t \in [0,1]\big\}$, centrado e com função de covariância Γ dada por

$$\begin{aligned}\Gamma(s,t) &= F(\min\{s,t\}) - F(s)F(t) + \\ &\quad + \sum_{k=1}^{\infty}\big(F_k(s,t) - F(s)F(t)\big) + \sum_{k=1}^{\infty}\big(F_k(t,s) - F(s)F(t)\big).\end{aligned}$$

Assim, $P\big(W_n^2 > \epsilon\big)$ é próxima da $P\big(\|\mathbb{B}\|^2 > \epsilon\big)$.

Definição 3.0.10 *Seja Γ o operador integral associado à covariância $\Gamma(s,t)$. Este operador, denominado operador de covariância,*
$\Gamma : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ é tal que para toda a função $f \in L^2(\mathbb{R})$,

$$\Gamma\big(f(t)\big) = \int_{\mathbb{R}} \Gamma(s,t) f(s) ds.$$

Vejamos algumas propriedades importantes dos operadores de covariância.

Teorema 3.0.11 *O operador Γ associado à função de covariância $\Gamma(s,t)$, função de quadrado integrável em $[a,b] \in \mathbb{R}$, é, no espaço $L^2([a,b])$ um operador linear compacto cuja norma satisfaz a desigualdade*

$$\|\Gamma\| \leq \sqrt{\int_{[a,b]\times[a,b]} \big|\Gamma(s,t)\big|^2 ds\, dt} \quad . \tag{3.1}$$

Demonstração:
Vejamos, em primeiro lugar, que Γ satisfaz a condição (3.1).
Seja $g(t) = \int_{[a,b]} \Gamma(s,t) f(s)\, ds$. Então,

$$\begin{aligned} |g(t)|^2 &= \left| \int_{[a,b]} \Gamma(s,t) f(s)\, ds \right|^2 \leq \int_{[a,b]} |\Gamma(s,t)|^2\, ds \int_{[a,b]} |f(s)|^2 ds = \\ &= \|f\|^2 \int_{[a,b]} |\Gamma(s,t)|^2\, ds. \end{aligned}$$

Assim,

$$\|\Gamma(f(t))\|^2 = \int_{[a,b]} |g(t)|^2 dt \leq \|f\|^2 \int_{[a,b]^2} |\Gamma(s,t)|^2\, ds dt.$$

Uma vez que Γ pertence a $L^2([a,b]^2)$, concluímos que (3.1) se verifica.
Consideremos agora $\{\psi_n\}$ um sistema ortogonal completo em $L^2([a,b])$. Então os produtos $\psi_n(s)\psi_m(t)$ de todos os possíveis pares de funções deste sistema constituem uma família completa em $L^2([a,b]^2)$, o que implica que

$$\Gamma(s,t) = \sum_{m,n=1}^{\infty} a_{mn} \psi_m(s) \psi_n(t).$$

Denotemos por $\Gamma_N(s,t) = \sum_{m,n=1}^{N} a_{mn} \psi_m(s) \psi_n(t)$, e consideremos o operador Γ_N associado a este núcleo. Este operador é compacto uma vez que transforma todo o espaço $L^2([a,b])$ num espaço de dimensão finita. De facto, se $\phi \in L^2([a,b])$, então

$$\Gamma_N\big(\phi(s)\big) = \int_{[a,b]} \Gamma_N(s,t)\phi(t)\, dt = \sum_{m,n=1}^{N} a_{mn} \psi_m(s) \int_{\mathbb{R}} \phi(t)\psi_n(t)\, dt =$$

$$= \sum_{m=1}^{N} \psi_m(s) \sum_{n=1}^{N} a_{mn}\, b_n, \quad \text{onde} \quad b_n = \int_{[a,b]} \phi(t)\psi_n(t)\, dt,$$

ou seja, cada elemento ϕ de $L^2([a,b])$ é transformado, pelo operador Γ_N, num elemento do espaço de dimensão finita gerado pelas funções

$\psi_1, \ldots, \psi_N$. Além disso, sendo $\Gamma_N(s,t)$ uma soma parcial da série de Fourier da função $\Gamma(s,t)$, tem-se

$$\int_{[a,b]\times[a,b]} \big(\Gamma(s,t) - \Gamma_N(s,t)\big)^2 ds\, dt \underset{N\to\infty}{\longrightarrow} 0.$$

Usando agora a desigualdade (3.1) ao operador $\Gamma - \Gamma_N$, obtemos, $\big\|\Gamma - \Gamma_N\big\| \underset{N\to\infty}{\longrightarrow} 0$. Assim, e uma vez que o limite de uma sucessão de operadores compactos é compacto, Γ, operador integral associado à covariância $\Gamma(s,t)$ é compacto. ■

Definição 3.0.12 *$\lambda \in \mathbb{R}$ é valor próprio do operador Γ se existir f em $L^2[a,b]$ tal que*

$$\int_{[a,b]} f(s)\Gamma(s,t)ds = \lambda f(t),$$

e f diz-se uma função própria à qual o valor próprio λ está associado.

Apresentamos agora, sem demonstração, uma série de propriedades dos operadores de covariância, cuja justificação pode ser encontrada em, por exemplo, Zimmer [54], 1990.

Propriedade 3.0.13 *Os valores próprios dos operadores de covariância são, no máximo, em quantidade numerável.*

Propriedade 3.0.14 *As funções próprias correspondentes a valores próprios distintos são ortogonais, isto é, $\int_{[a,b]} f_1(t)f_2(t)dt = 0$ para funções próprias correspondentes a valores próprios $\lambda_1 \neq \lambda_2$.*

Propriedade 3.0.15 *Existe no máximo um número finito de funções próprias linearmente independentes associadas a qualquer valor próprio não nulo; ao máximo de tal número dá-se o nome de multiplicidade geométrica do valor próprio.*

Propriedade 3.0.16 *Seja $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ uma ordenação do conjunto dos valores próprios de Γ, operador de covariância, onde cada valor*

próprio aparece tantas vezes quanto a sua multiplicidade. Então podemos escolher funções próprias correspondentes $f_1, f_2, \ldots, f_n, \ldots$ *de modo a que estas sejam ortonormais .*

Propriedade 3.0.17 *Tem-se necessariamente*

$$\sum_i \lambda_i^2 \leq \int_{[a,b]^2} \Gamma^2(s,t) ds dt < \infty,$$

para Γ *operador de covariância, o que implica que, se existe um número infinito de valores próprios* λ_j*, então* $\lambda_j \to 0$ *e* 0 *é o único ponto limite dos* λ_j*.*

Propriedade 3.0.18 *O valor próprio não nulo* λ *com maior valor absoluto é tal que*

$$|\lambda| = \sup_{g \in L^2[a,b]: \|g\|=1} \left\{ \int_{[a,b]^2} g(s)\Gamma(s,t)g(t) ds dt \right\}.$$

Propriedade 3.0.19 *O operador de covariância* Γ *é semidefinido positivo se e só se todos os seus valores próprios são maiores ou iguais a zero.*

Propriedade 3.0.20 *O operador* Γ *é definido positivo se e só se todos os seus valores próprios* λ_j *são maiores que zero e um conjunto de correspondentes funções próprias ortonormais* f_j *é um sistema ortonormal completo.*

O teorema que enunciamos de seguida, *Teorema de Mercer*, estabelece a diagonalização da função de covariância.

Teorema 3.0.21 *Seja* Γ *o operador associado à função de covariância* $\Gamma(s,t)$*. Seja ainda* $\{\lambda_1, \lambda_2, \ldots\}$ *o conjunto dos valores próprios de* Γ *e* $f_1, f_2, \ldots$ *funções próprias ortonormais correspondentes. Então a função de covariância pode ser representada por*

$$\Gamma(s,t) = \sum_{j=1}^{\infty} \lambda_j f_j(s) f_j(t), \forall_{s,t \in [a,b]},$$

onde a série converge uniforme e absolutamente. Além disso,

$$\sum_{j=1}^{\infty} \lambda_j < \infty.$$

Vejamos agora como se obtém a decomposição do processo $\mathbb{B}(t)$. Como referência tomamos Kac e Siegert (1947).

Proposição 3.0.22 *Seja $\Gamma(s,t)$ a função de covariância de um processo $\{\mathbb{B}(t), t \in [a,b]\}$ de média zero e cujas trajectórias estão em $L^2[a,b]$. Se $\Gamma(s,t)$ é tal que $\int_{[a,b]} \Gamma(t,t)dt < \infty$, então $\Gamma(s,t)$ é semidefinida positiva. Além disso,*

(**1**) *se (λ_1, f_1), (λ_2, f_2), ... são soluções de $\int_{[a,b]} f(s)\Gamma(s,t)ds = \lambda f(t)$, com $f_1, f_2, \ldots$ ortonormais então,*
$Z_j = \int_{[a,b]} \mathbb{B}(t) f_j(t)dt$, *para $j = 1, 2, \ldots$, são variáveis aleatórias independentes de média zero e variância λ_j;*

(**2**) $\sum_{j=1}^{m} Z_j f_j \underset{m\to\infty}{\longrightarrow} \mathbb{B}$, *quase certamente;*

(**3**) $\int_{[a,b]} \mathbb{B}^2(t)dt = \sum_{j=1}^{\infty} Z_j^2 = \sum_{j=1}^{\infty} \lambda_j Z_j^{*2}$, *onde* $Z_j^* \equiv \dfrac{Z_j}{\sqrt{\lambda_j}}$;

(**4**) *Para cada t,* $\mathbb{E}\left[\mathbb{B}(t) - \sum_{j=1}^{m} Z_j f_j(t)\right]^2 \underset{m\to\infty}{\longrightarrow} 0$;

(**5**) $\mathbb{B} = \sum_{j=1}^{\infty} \sqrt{\lambda_j} Z_j^* f_j$, *com* $Z_j^* = \dfrac{Z_j}{\sqrt{\lambda_j}}$ *variáveis aleatórias independentes gaussianas de média 0 e variância 1.*

De acordo com esta decomposição,

$$P\big(\|\mathbb{B}(t)\|_2^2 > \epsilon\big) = P\Big(\int_{[a,b]} \mathbb{B}^2(t)dt > \epsilon\Big) = P\Big(\sum_{j=1}^{\infty} \lambda_j Z_j^{*2} > \epsilon\Big).$$

Designando por m a multiplicidade do maior valor próprio do operador de covariância, λ_1, que sabemos ser finita, por γ a função Gama de Euler e por Φ a função de distribuição Gaussiana reduzida, Zolotarev [55], 1958, mostrou que esta última probabilidade se comporta assimptoticamente como,

$$2^{\frac{3-m}{2}}\lambda_1^{\frac{1-m}{2}}\pi^{1/2}\frac{1}{\gamma\left(\frac{m}{2}\right)}\epsilon^{\frac{1-m}{2}}\Big(1-\Phi\big(\frac{\epsilon^{1/2}}{\sqrt{\lambda_1}}\big)\Big)\prod_{i=n+1}^{\infty}\big(1-\frac{\lambda_i}{\lambda_1}\big)^{-1/2}. \qquad (3.3)$$

Ora esta quantidade pode ainda ser majorada. Sabemos que, para $x \in \mathbb{R}$ numa vizinhança de 0, podemos escolher $b > 0$, tal que $(1-x)^{-1/2} \leq 1 + bx$. Assim,

$$\prod_{i=n+1}^{\infty}\big(1-\frac{\lambda_i}{\lambda_1}\big)^{-1/2} \leq \prod_{i=n+1}^{\infty}\Big(1+b\big(\frac{\lambda_i}{\lambda_1}\big)\Big).$$

Por outro lado, para todo $x \in \mathbb{R}$, $\quad 1 + x \leq e^x$ pelo que,

$$\prod_{i=n+1}^{\infty}\Big(1+b\big(\frac{\lambda_i}{\lambda_1}\big)\Big) \leq \prod_{i=n+1}^{\infty} e^{b\left(\frac{\lambda_i}{\lambda_1}\right)} = e^{\frac{b}{\lambda_1}\sum_{i=n+1}^{\infty}\lambda_i}. \qquad (3.4)$$

Uma vez que, como vimos antes, $\sum_{i=1}^{\infty}\lambda_i < \infty$, aquela probabilidade depende essencialmente do maior valor próprio do operador de covariância Γ. Consequentemente, o nosso objectivo principal é obter a ordem de grandeza deste valor próprio: o maior valor próprio do operador associado à covariância do processo limite do processo empírico construído com sucessões associadas de variáveis aleatórias. Evidentemente que essa função de covariância é desconhecida, pois depende das funções de distribuição marginais e bivariadas das variáveis aleatórias de base, também estas desconhecidas.

A estratégia que seguimos na tentativa de resolução deste problema passa por obter um estimador consistente para a função de covariância. Esse estimador é construído a partir dos estimadores estudados e apresentados no capítulo precedente para as funções de distribuição das variáveis

aleatórias iniciais. Através do cálculo numérico do maior valor próprio de estimativas da função de covariância, inferimos acerca da ordem de grandeza do maior valor próprio de Γ. Esta inferência é, evidentemente, acompanhada de resultados relativos aos erros de aproximação.
Assim, este terceiro capítulo é dedicado ao estudo de um estimador para a função de covariância, assim como à obtenção de condições para a convergência dos valores próprios da sucessão de estimadores para os valores próprios do operador Γ.

Como já referimos anteriormente, este estimador incluirá os estimadores para as funções de distribuição marginais e bivariadas das variáveis aleatórias de base, de modo que usaremos e faremos referência ao conjunto de hipóteses introduzido na primeira secção do segundo capítulo desta monografia.

3.1 Definições e resultados preliminares

Seja, como habitualmente, $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas com função de distribuição marginal F, desconhecida, que supomos contínua.

Definição 3.1.1 *Seja $\{q_n\}_{n\in\mathbb{N}}$ uma sucessão de números inteiros cujo limite é $+\infty$. O estimador para a função de covariância do processo estocástico gaussiano, centrado, $\{\mathbb{B}(t), t\in\mathbb{R}\}$,*

$$\begin{aligned}\Gamma(s,t) &= F(\min\{s,t\}) - F(s)F(t) + \sum_{k=1}^{\infty}\big(F_k(s,t) - F(s)F(t)\big) + \\ &\quad + \sum_{k=1}^{\infty}\big(F_k(t,s) - F(s)F(t)\big),\end{aligned}$$

é a função $\widehat{\Gamma}_n$ das variáveis aleatórias $X_1,\ldots,X_n$, definida para cada

$(x,y) \in \mathbb{R}^2$ *por*

$$\begin{aligned}\widehat{\Gamma}_n(x,y) &= \widehat{F}_n(\min\{x,y\}) - \widehat{F}_n(x)\widehat{F}_n(y) + \\ &\quad + \sum_{k=1}^{q_n}\big(\widehat{F}_{n,k}(x,y) - \widehat{F}_n(x)\widehat{F}_n(y)\big) + \sum_{k=1}^{q_n}\big(\widehat{F}_{n,k}(y,x) - \widehat{F}_n(x)\widehat{F}_n(y)\big),\end{aligned}$$

onde

$$\widehat{F}_{n,k}(x,y) = \frac{1}{n-k}\sum_{i=1}^{n-k} U\Big(\frac{x-X_i}{h_n}, \frac{y-X_{k+i}}{h_n}\Big)$$

é o estimador do tipo núcleo estudado no capítulo anterior e

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} U_X\Big(\frac{x-X_i}{h_n}\Big), \ \textit{com}\ U_X(x) = \lim_{y\to+\infty} U(x,y).$$

Nota 3.1.2 *O estimador $\widehat{F}_n$ para a função de distribuição F, foi apresentado e estudado em Cai e Roussas [8], 1998, e usamos os resultados aí obtidos para o estudo de $\widehat{\Gamma}_n$.*

Nota 3.1.3 *A função de distribuição univariada U_X (função núcleo) pode, naturalmente, ser substituída por $U_Y = \lim_{x\to+\infty} U(x,y)$, sem que as propriedades de $\widehat{\Gamma}_n$ sejam modificadas.*

Começamos por apresentar os resultados existentes em Cai, Roussas [8], 1998, para a função $\widehat{F}_n$.

Lema 3.1.4 (Cai, Roussas [8], 1998) *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas, com função densidade de probabilidade marginal, f, limitada. Suponhamos ainda que $\{X_n\}_{n\in\mathbb{N}}$ satisfaz as condições (C_2) e (C_5). Então, para todo $t \in \mathbb{R}$,*

(i) $\Big|\mathbb{E}\widehat{F}_n(t) - F(t)\Big| \le Ah_n^2$, *sendo* $A = \frac{1}{2}\big|f'(t)\big|\int_{\mathbb{R}} s^2\, u_X(s)\, ds$, *onde*

u_X é a função densidade de probabilidade associada a U_X.

$$(ii)\ \left|\mathrm{Var}\left(U_X\Big(\frac{t-X_1}{h_n}\Big)\right) - F(t)\big(1-F(t)\big) + f(t)h_n\int_{\mathbb{R}} s\,u_X(s)\,U_X(s)\,ds\right| \leq$$

$$\leq D\,h_n^2, \quad \textit{sendo } D = \frac{1}{2}\big|f'(t)\big|\int_{\mathbb{R}} s^2 u_X(s) U_X(s)\,ds + 2A.$$

Recordemos as definições das funções I_{nj} e I_j com $j \geq 1$, já introduzidas no segundo capítulo mas, desta feita, para $t \in \mathbb{R}$,

$$\begin{aligned} I_{nj}(t) &= \mathrm{Cov}\left(U_X\Big(\frac{t-X_1}{h_n}\Big), U_X\Big(\frac{t-X_{j+1}}{h_n}\Big)\right) \\ I_j(t) &= \mathrm{Cov}\Big(\mathbb{1}_{(-\infty,t]}\big(X_1\big), \mathbb{1}_{(-\infty,t]}\big(X_{j+1}\big)\Big) = \\ &= F_j(t,t) - F^2(t). \end{aligned}$$

Seja ainda, para $t \in \mathbb{R}$,

$$\sigma^2(t) = F(t)\big(1-F(t)\big) + 2\sum_{j=1}^{\infty} I_j(t).$$

Nota 3.1.5 *Para cada $t \in \mathbb{R}$, a série que define $\sigma^2(t)$ é convergente, desde que a condição $(C_7)'$ se verifique, uma vez que, como já foi visto no primeiro capítulo, a função $I_j(t)$ é majorada por* $\mathrm{Cov}^{1/3}(X_1, X_{j+1})$.

Lema 3.1.6 (Cai, Roussas [8], 1998) *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas, com função densidade de probabilidade marginal, f, limitada. Suponhamos ainda que $\{X_n\}_{n\in\mathbb{N}}$ satisfaz as condições $(C_2), (C_3)$ e (C_5). Então, para todo $t \in \mathbb{R}$,*

$$\big|I_{nj}(t) - I_j(t)\big| \leq \gamma_0 h_n^2, \ \textit{ onde } \gamma_0 \ \textit{ é independente de } n, j \ \textit{ e } t.$$

Se, além das condições anteriores $\{X_n\}_{n\in\mathbb{N}}$ satisfaz (C_7), tem-se que

$$n\mathrm{Var}\left(\widehat{F}_n(t)\right) = \sigma^2(t) + o(h_n).$$

Nota 3.1.7 *Vimos já no capítulo anterior que se a sucessão $\{h_n\}_{n\in\mathbb{N}}$ satisfaz a condição (C_6), a sua ordem de convergência óptima é $O(n^{-1})$. Sublinhe-se ainda que a sucessão $\{h_n\}_{n\in\mathbb{N}}$ envolvida nas definições de $\widehat{F}_{n,k}$ e $\widehat{F}_n$ é a mesma.*

Vamos agora analisar o produto, definido para todo $x, y \in \mathbb{R}$ por,

$$\widehat{F}_n(x)\widehat{F}_n(y) = \frac{1}{n^2}\sum_{i,j=1}^{n} U_X\Big(\frac{x-X_i}{h_n}\Big)U_X\Big(\frac{y-X_j}{h_n}\Big).$$

Lema 3.1.8 *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas que verifica as condições $(C_2), (C_5), (C_6)$ e $(C_7)'$. Então, para todo $x, y \in \mathbb{R}$, $\mathbb{E}\left(\widehat{F}_n(x)\widehat{F}_n(y)\right) = F(x)F(y) + O\left(\frac{1}{n}\right)$.*

Demonstração: Calculemos a esperança matemática do produto de $\widehat{F}_n$ em dois pontos de $\mathbb{R}$ diferentes. Atendendo à sua definição, podemos escrever

$$\mathbb{E}\big(\widehat{F}_n(x)\widehat{F}_n(y)\big) = \frac{1}{n^2}\sum_{i,j=1}^{n} \mathbb{E}\left[U_X\Big(\frac{x-X_i}{h_n}\Big)U_X\Big(\frac{y-X_j}{h_n}\Big)\right] =$$

(por estacionaridade)

$$= \frac{1}{n}\mathbb{E}\left[U_X\Big(\frac{x-X_1}{h_n}\Big)U_X\Big(\frac{y-X_1}{h_n}\Big)\right] + \\ + \frac{2}{n^2}\sum_{j=1}^{n-1}(n-j)\mathbb{E}\left[U_X\Big(\frac{x-X_1}{h_n}\Big)U_X\Big(\frac{y-X_{j+1}}{h_n}\Big)\right].$$

Analisemos agora a esperança do produto da função U_X em pontos diferentes, envolvido na expressão anterior. Para cada $j = 1, \ldots, n$,

$$\mathbb{E}\left[U_X\Big(\frac{x-X_1}{h_n}\Big)U_X\Big(\frac{y-X_{j+1}}{h_n}\Big)\right] = \\ = \mathrm{Cov}\left\{U_X\Big(\frac{x-X_1}{h_n}\Big), U_X\Big(\frac{y-X_{j+1}}{h_n}\Big)\right\} + \\ + \mathbb{E}\Big[U_X\Big(\frac{x-X_1}{h_n}\Big)\Big]\mathbb{E}\Big[U_X\Big(\frac{y-X_{j+1}}{h_n}\Big)\Big].$$

Ora, de acordo com os Lemas 3.1.4 e 3.1.6, temos

$$\mathbb{E}\left[U_X\Big(\frac{x-X_1}{h_n}\Big)U_X\Big(\frac{y-X_{j+1}}{h_n}\Big)\right] = F_j(x,y) + O(h_n^2),$$

onde, para cada $x, y \in \mathbb{R}$, F_j é tal que

$$F_j(x,y) = \begin{cases} F\big(\min\{x,y\}\big) & \text{se } j=0 \\ P\big(X_1 \le x, X_{j+1} \le y\big) & \text{se } j>0 \quad . \end{cases}$$

Podemos então escrever,

$$\mathbb{E}\left(\widehat{F}_n(x)\widehat{F}_n(y)\right) = \frac{1}{n}\left(F\big(\min\{x,y\}\big)+\frac{2}{n}\sum_{j=1}^{n-1}\left((n-j)\,F_j(x,y)+O(h_n^2)\right)\right) =$$

$$= \frac{1}{n}\Big(F\big(\min\{x,y\}\big)+\frac{2}{n}\sum_{j=1}^{n-1}(n-j)\times$$

$$\times\Big(\mathrm{Cov}\left\{1\!\!1_{(-\infty,x]}\big(X_1\big), 1\!\!1_{(-\infty,y]}\big(X_{j+1}\big)\right\}+F(x)F(y)\Big)\Big)+$$

$$+O(\frac{h_n^2}{n})$$

$$= \frac{1}{n}\Big(F\big(\min\{x,y\}\big)\Big)+\frac{2}{n}\sum_{j=1}^{n-1}\mathrm{Cov}\left\{1\!\!1_{(-\infty,x]}\big(X_1\big), 1\!\!1_{(-\infty,y]}\big(X_{j+1}\big)\right\}-$$

$$-\frac{2}{n^2}\sum_{j=1}^{n-1} j\ \mathrm{Cov}\left\{1\!\!1_{(-\infty,x]}\big(X_1\big), 1\!\!1_{(-\infty,y]}\big(X_{j+1}\big)\right\}+$$

$$+\frac{2}{n^2}\sum_{j=1}^{n-1}(n-j)\,F(x)F(y)+O(\frac{h_n^2}{n})$$

Uma vez mais, de acordo com a desigualdade (1.25),

$$\mathrm{Cov}\left\{1\!\!1_{(-\infty,x]}\big(X_1\big), 1\!\!1_{(-\infty,y]}\big(X_{j+1}\big)\right\} \le M\mathrm{Cov}^{\,1/3}(X_1, X_{j+1}),$$

temos então

$$\mathbb{E}\Big(\widehat{F}_n(x)\widehat{F}_n(y)\Big) = \frac{1}{n}\Big(F\big(\min\{x,y\}\big)\Big)+\frac{2M}{n}\sum_{j=1}^{n-1}\mathrm{Cov}^{\,1/3}\big(X_1, X_{j+1}\big)-$$

$$-\frac{2M}{n^2}\sum_{j=1}^{n-1} j\ \mathrm{Cov}^{\,1/3}\big(X_1, X_{j+1}\big)+\frac{n-1}{n}F(x)F(y)+O\left(\frac{h_n^2}{n}\right) =$$

$$= F(x)F(y)+O(n^{-1}).\blacksquare$$

Vejamos agora qual o comportamento assimptótico da variância de $\widehat{F}_n(x)\widehat{F}_n(y)$.

Lema 3.1.9 *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas que verifica as condições (C_2), (C_5) e $(C_7)'$. Suponhamos ainda que a sucessão de covariâncias $\{\mathrm{Cov}\,(X_1, X_{n+1})\}_{n\in\mathbb{N}}$ é decrescente e que $h_n = O(n^{-1})$. Então, para todo $x, y \in \mathbb{R}$,*

$$\mathrm{Var}\left(\widehat{F}_n(x)\widehat{F}_n(y)\right) = O(n^{-1}).$$

Demonstração: Denotemos por $Y_{i,j}(x,y)$ o produto $U_X\left(\frac{x-X_i}{h_n}\right) \times$ $\times U_X\left(\frac{y-X_j}{h_n}\right)$.

A variância de $\widehat{F}_n(x)\widehat{F}_n(y)$ tem, por definição, a seguinte expressão

$$\mathrm{Var}\left(\widehat{F}_n(x)\widehat{F}_n(y)\right) = \frac{1}{n^4}\mathrm{Cov}\left(\sum_{i,j,k,l=1}^{n} Y_{i,j}(x,y)Y_{k,l}(x,y)\right).$$

Apelando à estacionaridade da sucessão de variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$, podemos escrever,

$$\mathrm{Var}\left(\widehat{F}_n(x)\widehat{F}_n(y)\right) = \frac{1}{n^4}\sum_{i,j=1}^{n}\mathrm{Cov}\Big(Y_{i,j}(x,y), Y_{i,j}(x,y)\Big) + \\ + \frac{2}{n^4}\sum_{(i,j)<(k,l)}\mathrm{Cov}\Big(Y_{i,j}(x,y), Y_{k,l}(x,y)\Big),$$

onde se usa a ordem lexicográfica para os pares de índices (i,j).

Denotemos por $F_{ijkl}(x,y,s,t) = P(X_i \le x, X_j \le y, X_k \le s, X_l \le t)$ e

$F_{ij}(x,y) = P(X_i \le x, X_j \le y)$. Notemos que, de acordo com esta notação, $F_{ijij}(x,y,x,y) = F_{ij}(x,y)$. Temos então, para todo $x, y \in \mathbb{R}^2$,

$$\mathrm{Cov}\Big(Y_{i,j}(x,y), Y_{k,l}(x,y)\Big) = F_{ijkl}(x,y,x,y) - F_{ij}(x,y)F_{kl}(x,y) + O(h_n^2).$$

De facto, podemos escrever,

$$\begin{aligned}
&\mathrm{Cov}\Big(Y_{i,j}(x,y), Y_{k,l}(x,y)\Big) = \\
&\quad = \int_{\mathbb{R}^4} U_X\Big(\frac{x-s}{h_n}\Big)\, U_X\Big(\frac{y-t}{h_n}\Big)\, U_X\Big(\frac{x-s'}{h_n}\Big)\, U_X\Big(\frac{y-t'}{h_n}\Big)\, dF_{ijkl}(s,t,s',t') - \\
&\qquad - \Big(\int_{\mathbb{R}^2} U_X\Big(\frac{x-s}{h_n}\Big)\, U_X\Big(\frac{y-t}{h_n}\Big)\, dF_{ij}(s,t)\Big) \times \\
&\qquad\quad \times \Big(\int_{\mathbb{R}^2} U_X\Big(\frac{x-s'}{h_n}\Big)\, U_X\Big(\frac{y-t'}{h_n}\Big)\, dF_{ij}(s',t')\Big)\,.
\end{aligned}$$

A primeira parcela do segundo membro desta igualdade é, tendo em conta a definição de U_X,

$$\int_{\mathbb{R}^4} \int_C u_X(a)\, u_X(b)\, u_X(c)\, u_X(d)\; dF_{ijkl}(s,t,s',t'),$$

onde $C = (-\infty, \frac{x-s}{h_n}] \times (-\infty, \frac{y-t}{h_n}] \times (-\infty, \frac{x-s'}{h_n}] \times (-\infty, \frac{y-t'}{h_n}]$. Usando o Teorema de Fubini e invertendo a ordem de integração,

$$\begin{aligned}
&\int_{\mathbb{R}^4} U_X\Big(\frac{x-s}{h_n}\Big)\, U_X\Big(\frac{y-t}{h_n}\Big)\, U_X\Big(\frac{x-s'}{h_n}\Big)\, U_X\Big(\frac{y-t'}{h_n}\Big)\, dF_{ijkl}(s,t,s',t') = \\
&= \int_{\mathbb{R}^4} u_X(a)\, u_X(b)\, u_X(c)\, u_X(d)\; F_{ijkl}(x-ah_n, y-bh_n, x-ch_n, y-dh_n) da\, db\, dc\, dd \\
&\text{(desenvolvendo } F_{ijkl} \text{ em polinómio de segunda ordem em torno de } (x,y,x,y)) \\
&= F_{ijkl}(x,y,x,y) + O(h_n^2).
\end{aligned}$$

Usando a mesma estratégia para a segunda parcela do segundo membro da equação em causa, obtemos se $(i,j) < (k,l)$, (usando a ordem lexicográfica)

$$\mathrm{Cov}\Big(Y_{ij}(x,y), Y_{kl}(x,y)\Big) \;\; = \;\; F_{ijkl}(x,y,x,y) - F_{ij}(x,y)F_{kl}(x,y) + O(h_n^2)$$

e se $(i,j) = (k,l)$,

$$\mathrm{Cov}\Big(Y_{ij}(x,y), Y_{ij}(x,y)\Big) = F_{ij}(x,y) - F_{ij}^2(x,y) + O(h_n^2).$$

Ora, conforme já foi demonstrado no Lema 2.2.4 da segunda secção do capítulo anterior,

$$\begin{aligned}
&F_{ij}(x,y) - F_{ij}^2(x,y) = \\
&\quad = \mathrm{Cov}\left(\mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_i,X_j), \mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_i,X_j)\right) \leq \\
&\quad\quad \leq M\left(\mathrm{Cov}^{1/3}(X_i,X_i) + \mathrm{Cov}^{1/3}(X_i,X_j) + \right. \\
&\quad\quad\quad \left. + \mathrm{Cov}^{1/3}(X_j,X_i) + \mathrm{Cov}^{1/3}(X_j,X_j)\right).
\end{aligned}$$

Atendendo à estacionaridade segue-se que,

$$F_{ij}(x,y) - F_{ij}^2(x,y) \leq 2M\big(\mathrm{Var}^{1/3}(X_1) + \mathrm{Cov}^{1/3}(X_i,X_j)\big).$$

Temos então,

$$\begin{aligned}
&\frac{1}{n^4}\sum_{i,j=1}^{n} \mathrm{Cov}\left(Y_{ij}(x,y), Y_{ij}(x,y)\right) \leq \\
&\quad \leq \frac{2M}{n^4}\sum_{i,j=1}^{n}\left(\mathrm{Var}^{1/3}(X_1) + \mathrm{Cov}^{1/3}(X_i,X_j) + O(h_n^2)\right) = \\
&\quad = \frac{4M}{n^2}\mathrm{Var}^{1/3}(X_1) + \frac{2M}{n^3}\sum_{j=1}^{n-1}(n-j)\,\mathrm{Cov}^{1/3}(X_1,X_{j+1}) + O(\frac{h_n^2}{n^2}).
\end{aligned}$$

Uma vez que $\mathrm{Var}\,(X_1)$ é finita, de acordo com as condições $(C_7)'$ e (C_7) e usando o facto de $h_n \sim n^{-1}$, vem

$$\frac{1}{n^4}\sum_{i,j=1}^{n} \mathrm{Cov}\left(Y_{ij}(x,y), Y_{ij}(x,y)\right) = O(n^{-1}).$$

Vejamos agora a majoração de

$$\frac{1}{n^4}\sum_{(i,j)<(k,l)} \mathrm{Cov}\left(Y_{ij}(x,y), Y_{kl}(X,y)\right).$$

Usemos novamente o Lema 2.2.4 para majorar, em termos das covariâncias entre as variáveis iniciais X_i, o termo geral da soma anterior,

$$\begin{aligned}
&F_{ijkl}(x,y,x,y) - F_{ij}(x,y)F_{kl}(x,y) = \\
&\quad = \mathrm{Cov}\left(\mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_i,X_j), \mathbb{1}_{(-\infty,x]\times(-\infty,y]}(X_k,X_l)\right) \leq \\
&\quad \leq M\big(\mathrm{Cov}^{1/3}(X_i,X_k) + \mathrm{Cov}^{1/3}(X_i,X_l) + \\
&\qquad + \mathrm{Cov}^{1/3}(X_j,X_k) + \mathrm{Cov}^{1/3}(X_j,X_l)\big),
\end{aligned}$$

onde M é a constante interveniente na desigualdade (1.25). Nestas condições,

$$\begin{aligned}
&\frac{1}{n^4}\sum_{(i,j)<(k,l)} \mathrm{Cov}\left(Y_{ij}(x,y), Y_{kl}(X,y)\right) \leq \\
&\quad \leq \frac{1}{n^4}\sum_{(i,j)<(k,l)} \left(F_{ijkl}(x,y,x,y) - F_{ij}(x,y)F_{kl}(x,y) + O(h_n^2)\right) \leq \\
&\quad \leq O(h_n^2) + \frac{6M}{n^2}\sum_{i<j} \mathrm{Cov}^{1/3}(X_i,X_j) \leq O(h_n^2) + \\
&\qquad + \frac{6M}{n^2}\sum_{j=1}^{n} (n-j)\,\mathrm{Cov}^{1/3}(X_1,X_j).
\end{aligned}$$

Usando novamente as condições (C_7) e $(C_7)'$, obtém-se que esta quantidade é da $O(n^{-1})$. O resultado do lema é agora imediato. ■

3.2 Consistência de $\widehat{\Gamma}_n$

Nesta secção obtemos condições suficientes para a convergência de $\widehat{\Gamma}_n$. A velocidade desta convergência depende naturalmente da sucessão de números positivos $\{q_n\}_{n\in\mathbb{N}}$ que intervêm na definição de $\widehat{\Gamma}_n$, assim como da velocidade de convergência para zero da sucessão cujo termo geral é dado por

$$u(k) = \sum_{j=k}^{\infty} \mathrm{Cov}^{1/3}(X_1, X_{j+1}).$$

Fazemos o controle da velocidade desta sucessão $\{u_k\}_{k\in\mathbb{N}}$ em função do crescimento de q_n.
Comecemos, então, por mostrar que o estimador $\widehat{\Gamma}_n$ é assimptoticamente centrado e vejamos a velocidade desta convergência.

Teorema 3.2.1 *Suponhamos que a sucessão $\{X_n\}_{n\in\mathbb{N}}$ de variáveis aleatórias reais satisfazem as condições $(C_1), (C_2), (C_3), (C_4), (C_5)$ e $(C_7)'$. Seja $h_n = O(n^{-1})$. Então,*

$$\left|\Gamma(x,y) - \mathbb{E}\Big(\widehat{\Gamma}_n(x,y)\Big)\right| = O\left(\frac{q_n}{n} + q_n\, h_n^2 + u(q_n+1)\right).$$

Demonstração: Com base nas definições de Γ e $\widehat{\Gamma}_n$ e usando a desigualdade triangular,

$$\begin{aligned}\left|\Gamma(x,y) - \mathbb{E}\Big(\widehat{\Gamma}_n(x,y)\Big)\right| \leq{}& \left|F(\min\{x,y\}) - \mathbb{E}\Big(\widehat{F}_n(\min\{x,y\})\Big)\right| + \\ &+(2q_n+1)\left|\mathbb{E}\big(\widehat{F}_n(x)\widehat{F}_n(y)\big) - F(x)F(y)\right| + \\ &+\sum_{k=1}^{q_n}\left|F_k(x,y) - \mathbb{E}\big(\widehat{F}_{n,k}(x,y)\big)\right| + \sum_{k=1}^{q_n}\left|F_k(y,x) - \mathbb{E}\big(\widehat{F}_{n,k}(y,x)\big)\right| + \\ &+\sum_{k=q_n+1}^{\infty}\Big(F_k(x,y) - F(x)F(y)\Big) + \sum_{k=q_n+1}^{\infty}\Big(F_k(y,x) - F(x)F(y)\Big).\end{aligned}$$

Usando agora os resultados relativos à convergência dos estimadores $\widehat{F}_n, \widehat{F}_{n,k}$ e do produto $\widehat{F}_n(x)\widehat{F}_n(y)$, tendo em conta a associação da sucessão $\{X_n\}_{n\in\mathbb{N}}$, vem,

$$\left|\Gamma(x,y) - \mathbb{E}\Big(\widehat{\Gamma}_n(x,y)\Big)\right| \leq A\, h_n^2 + B(2q_n+1)n^{-1} + 2Cq_nh_n^2 + 2u(q_n+1),$$

como pretendíamos. ■

Refira-se que A, B e C são constantes positivas que, como sabemos, não dependem de x e y por serem obtidas através majorações uniformes das funções densidade de probabilidade marginais e das derivadas de segunda ordem das funções de distribuição multivariadas das variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$, conforme as condições $(C_1), (C_2), (C_3)$ e (C_4).

Nota 3.2.2 *A melhor velocidade desta convergência nunca é de ordem superior ou mesmo igual à de* n^{-1} *uma vez que a sucessão* $\{q_n\}_{n\in\mathbb{N}}$ *deve tender para infinito. No entanto, é evidente que quanto mais lentamente a sucessão* $\{q_n\}_{n\in\mathbb{N}}$ *tender para infinito, mais rapidamente se dá a convergência do estimador para a função de covariância e, supondo* $u(q_n+1)=O\left(\frac{q_n}{n}\right)$, *mais rapidamente temos a sucessão dos restos* $\{u_k\}_{k\in\mathbb{N}}$ *a tender também para zero.*
Observe-se que se a sucessão $\{q_n\}_{n\in\mathbb{N}}$ *for tal que* $q_n=O(n^\alpha)$, *para algum* $\alpha>0$ *e se* $\{u_k\}_{k\in\mathbb{N}}$ *for da mesma ordem que* $k^{-\rho}$, *com* $\rho>0$, *temos, relativamente ao resultado descrito no Teorema 3.2.1, uma velocidade de convergência da ordem de* $n^{\alpha-1}$, *se* $\alpha>\frac{1}{1+\rho}$ *e de* $O(n^{-\alpha\rho})$, *caso contrário, desde que tomemos* $\alpha\in(0,1)$.

Vejamos em seguida as condições suficientes para a convergência do estimador em média quadrática e qual a velocidade desta convergência.

Teorema 3.2.3 *Suponhamos que a sucessão* $\{X_n\}_{n\in\mathbb{N}}$ *de variáveis aleatórias reais satisfazem as condições* $(C_1),(C_2),(C_3),(C_4),(C_5)$ *e* $(C_7)'$. *Seja* $h_n=O(n^{-1})$. *Então,*

$$\mathbb{E}\Big(\Gamma(x,y)-\widehat{\Gamma}_n(x,y)\Big)^2=O\Big(\frac{q_n^2}{n}+u^2(q_n+1)\Big).$$

Demonstração: Comecemos por obter um majorante para $\mathbb{E}\Big(\Gamma(x,y)-\widehat{\Gamma}_n(x,y)\Big)^2$. Usando as expressões de Γ e de $\widehat{\Gamma}_n$ temos,

$$\begin{aligned}\mathbb{E}\Big(\Gamma(x,y)-\widehat{\Gamma}_n(x,y)\Big)^2=&\\ =\mathbb{E}\Big[F(\min\{x,y\})-\Big(\widehat{F}_n(\min\{x,y\})\Big)+&\\ +(2q_n+1)\Big(\big(\widehat{F}_n(x)\widehat{F}_n(y)\big)-F(x)F(y)\Big)+&\end{aligned}$$

$$
\begin{aligned}
&+\sum_{k=1}^{q_n}\Big(F_k(x,y)-\widehat{F}_{n,k}(x,y)\Big)+\sum_{k=1}^{q_n}\Big(F_k(y,x)-\widehat{F}_{n,k}(y,x)\Big)+\\
&+\sum_{k=q_n+1}^{\infty}\Big(F_k(x,y)-F(x)F(y)\Big)+\sum_{k=q_n+1}^{\infty}\Big(F_k(y,x)-F(x)F(y)\Big)\Big]^2\leq\\
&\leq 2\,\mathbb{E}\Big[F(\min\{x,y\})-\widehat{F}_n(\min\{x,y\})+\\
&\quad+(2q_n+1)\Big(\widehat{F}_n(x)\widehat{F}_n(y)-F(x)F(y)\Big)+\\
&\quad+\sum_{k=1}^{q_n}\big(F_k(x,y)-\widehat{F}_{n,k}(x,y)\big)+\sum_{k=1}^{q_n}\big(F_k(y,x)-\widehat{F}_{n,k}(y,x)\big)\Big]^2+\\
&\quad+2\left(\sum_{k=q_n+1}^{\infty}\Big(F_k(x,y)-F(x)F(y)\Big)+\sum_{k=q_n+1}^{\infty}\Big(F_k(y,x)-F(x)F(y)\Big)\right)^2.
\end{aligned}
$$

A última parcela do segundo membro desta desigualdade é determinística, sendo a sua convergência controlada pela sucessão $\{u_n\}_{n\in\mathbb{N}}$.

Relativamente às restantes parcelas, aplicando sucessivamente a desigualdade $\mathbb{E}(X+Y)^2\leq 2\big(\mathbb{E}(X^2)+\mathbb{E}(Y^2)\big)$, vem,

$$
\begin{aligned}
\mathbb{E}\Big(\Gamma(x,y) &- \widehat{\Gamma}_n(x,y)\Big)^2\leq 8\,\mathbb{E}\Big(F(\min\{x,y\})-\widehat{F}_n(\min\{x,y\})\Big)^2+\\
&+ 8\,\mathbb{E}\Big((2q_n+1)\big(\widehat{F}_n(x)\widehat{F}_n(y)-F(x)F(y)\big)\Big)^2+\\
&+ 4\,\mathbb{E}\left(\sum_{k=1}^{q_n}\big(F_k(x,y)-\widehat{F}_{n,k}(x,y)\big)+\right.\\
&\qquad\left.+\sum_{k=1}^{q_n}\big(F_k(y,x)-\widehat{F}_{n,k}(y,x)\big)\right)^2\\
&+ 4\,M\,\big(u(q_n+1)\big)^2,
\end{aligned}
$$

onde M é, como habitualmente, a constante da desigualdade (1.25). Tendo agora em conta o comportamento do erro quadrático médio dos estimadores $\widehat{F}_n, \widehat{F}_{n,k}$ e do produto $\widehat{F}_n(x)\widehat{F}_n(y)$, temos com c_1, c_2 e c_3

constantes positivas,

$$\mathbb{E}\Big(\Gamma(x,y)-\widehat{\Gamma}_n(x,y)\Big)^2=\frac{8\,c_1}{n}+\frac{8\,c_2\,(2q_n+1)^2}{n}+4\frac{c_3\,q_n^2}{n}+u^2(q_n+1).$$

O resultado é, agora, imediato. ■

As constantes c_1, c_2 e c_3 não dependem das variáveis x e y por, mais uma vez, por serem obtidas através de majorações uniformes das funções densidade de probabilidade marginais e das derivadas de segunda ordem das funções de distribuição multivariadas das variáveis aleatórias $\{X_n\}_{n\in\mathbb{N}}$, conforme as condições $(C_1), (C_2), (C_3)$ e (C_4).

Nota 3.2.4 *A condição para q_n considerada neste Teorema deverá ser mais forte que a do Teorema anterior. De facto, se considerarmos $u(q_n+1)=O\Big(\frac{q_n}{n}\Big)$, a velocidade da convergência em média quadrática será da ordem de $O\Big(\frac{q_n^2}{n}\Big)$, mais lenta que a anterior. Assim, sendo q_n e u_n como na nota anterior, tomando neste caso, $\alpha\in(0,1/2)$ obtemos uma velocidade para a convergência em média quadrática do estimador $\widehat{\Gamma}_n$ da mesma ordem que $n^{2\alpha-1}$, se $\alpha>\frac{1}{2(1+\rho)}$ e da $O(n^{-2\alpha\rho})$, caso contrário.*

Tratemos, agora, da convergência quase certa de $\widehat{\Gamma}_n$. Por uma questão de simplicidade, fazemos esse tratamento analisando o módulo da diferença entre Γ e $\widehat{\Gamma}_n$ por partes. Comecemos por escrever, para cada $x, y\in\mathbb{R}$,

$$\begin{aligned}\Big|\Gamma(x,y)-\widehat{\Gamma}_n(x,y)\Big|\leq&\ \Big|F\big(\min\{x,y\}\big)-\widehat{F}_n\big(\min\{x,y\}\big)\Big|+\\&+(2q_n+1)\Big|\widehat{F}_n(x)\widehat{F}_n(y)-F(x)F(y)\Big|+\sum_{k=1}^{q_n}\Big|F_k(x,y)-\widehat{F}_{n,k}(x,y)\Big|\\&+\sum_{k=1}^{q_n}\Big|F_k(y,x)-\widehat{F}_{n,k}(y,x)\Big|+\\&+\left|\sum_{k=q_n+1}^{\infty}\Big(F_k(x,y)-F(x)F(y)\Big)+\sum_{k=q_n+1}^{\infty}\Big(F_k(y,x)-F(x)F(y)\Big)\right|.\end{aligned}$$

Consideremos, em primeiro lugar, as parcelas

$$\sum_{k=1}^{q_n}\Big|F_k(x,y)-\widehat{F}_{n,k}(x,y)\Big|+\sum_{k=1}^{q_n}\Big|F_k(y,x)-\widehat{F}_{n,k}(y,x)\Big|.$$

Provámos já, na segunda secção do segundo capítulo desta monografia, que se a sucessão de covariâncias definida para cada $j\in\mathbb{N}$ por, $\mathrm{Cov}\,(X_1,X_{j+1})=\rho_0\rho^j,\ 0<\rho<1$, com $\rho_0>0$, então

$$\Big|\widehat{F}_{n,k}(x,y)-F_k(x,y)\Big|=O\left(\frac{(\log\log n)^2}{n}\right)^{1/3}.$$

Assim, para cada $x,y\in\mathbb{R}$,

$$\begin{aligned}\sum_{k=1}^{q_n}\Big|F_k(x,y)-\widehat{F}_{n,k}(x,y)\Big|+\sum_{k=1}^{q_n}\Big|F_k(y,x)-\widehat{F}_{n,k}(y,x)\Big|&\leq\\ \leq 2c\,q_n\left(\frac{(\log\log n)^2}{n}\right)^{1/3},&\qquad(3.8)\end{aligned}$$

com c constante. A convergência para zero desta expressão verifica-se desde que $q_n=o\left(\frac{(\log\log n)^2}{n}\right)^{-1/3}$, tendo-se, neste caso, a menor velocidade de convergência.

O melhoramento desta velocidade pode obter-se considerando q_n a tender para $+\infty$ a uma velocidade não inferior a $(\log\log n)^{\frac{\zeta}{3}}$, com $\zeta>0$, conforme se deduz facilmente analisando a desigualdade (3.8), uma vez que, neste caso, o primeiro membro desta desigualdade é majorado por uma expressão da mesma ordem que $\left(\frac{(\log\log n)^{\zeta+2}}{n}\right)^{1/3}$.

Consideremos agora a parcela

$$\begin{aligned}(2q_n+1)\Big|\widehat{F}_n(x)\widehat{F}_n(y)-F(x)F(y)\Big|\ &\leq\ 2q_n\Big|\widehat{F}_n(x)\widehat{F}_n(y)-F(x)F(y)\Big|+\\ &+\ \Big|\widehat{F}_n(x)\widehat{F}_n(y)-F(x)F(y)\Big|.\end{aligned}$$

Para o tratamento da convergência quase certa desta expressão, precisamos de, em primeiro lugar, estabelecer um resultado de convergência de $\widehat{F}_n(x)\widehat{F}_n(y)$ para $F(x)F(y)$.

Lema 3.2.5 *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas que verificam as condições $(C_2),(C_3)$ e (C_5) e, ainda, as condições $(A_1),(A_2)$ e (A_3). Suponhamos ainda que, para $j\geq 1$, se tem*

$$\mathrm{Cov}\left(X_1,X_{j+1}\right)=\rho_0\rho^j,\ 0<\rho<1,\quad \rho_0>0.$$

Então

$$\left|\widehat{F}_n(x)\widehat{F}_n(y)-F(x)F(y)\right|=O\left(\frac{(\log\log n)^2}{n}\right)^{\frac{1}{3}}+O(h_n^2).$$

Demonstração: Podemos majorar o módulo da diferença $\left|\widehat{F}_n(x)\widehat{F}_n(y)-F(x)F(y)\right|$ apenas em termos dos factores de $\widehat{F}_n(x)\widehat{F}_n(y)$ e de $F(x)F(y)$, do seguinte modo

$$\left|\widehat{F}_n(x)\widehat{F}_n(y)-F(x)F(y)\right|\leq\left|\widehat{F}_n(x)-F(x)\right|+\left|\widehat{F}_n(y)-F(y)\right|.$$

Vejamos as majorações das velocidades de convergência para zero de cada uma das parcelas que constituem o segundo membro da desigualdade anterior. Uma vez que a sucessão $\{X_n\}_{n\in\mathbb{N}}$ satisfaz as condições $(A_1),(A_2)$ e (A_3) da desigualdade exponencial de Ioannides e Roussas [18], o mesmo também se verifica relativamente à sucessão $\left\{U_X\left(\frac{x-X_j}{h_n}\right)\right\}_{j\in\mathbb{N}}$. De facto, tendo em conta o Lema 3.1.6 e a desigualdade 1.25, se $\mathrm{Cov}\left(X_1,X_{j+1}\right)=\rho_0\rho^j,\ 0<\rho<1,\quad \rho_0>0$, então $\mathrm{Cov}\left(U_X\left(\frac{x-X_1}{h_n}\right),U_X\left(\frac{x-X_{j+1}}{h_n}\right)\right)=O(\rho_0\rho^j)^{\frac{1}{3}}$. Assim, com $r_n=O\left(\frac{n^2}{\log\log n}\right)^{\frac{1}{3}}$, estamos nas condições do Corolário 2.2.10, pelo que a ordem de convergência de $\left|\widehat{F}_n(x)-\mathbb{E}\widehat{F}_n(x)\right|$ para zero é de $O\left(\frac{(\log\log n)^2}{n}\right)^{\frac{1}{3}}$. Usando agora o facto de $\left|\mathbb{E}\widehat{F}_n(x)-F(x)\right|=O(h_n^2)$, de acordo com o Lema 3.1.4 (i), o resultado é imediato. ■

Nota 3.2.6 *De acordo com este Teorema, existem constantes positivas A e B tal que,*

$$\left|\widehat{F}_n(x)\widehat{F}_n(y)-F(x)F(y)\right|\leq A\left(\frac{(\log\log n)^2}{n}\right)^{\frac{1}{3}}+\frac{B}{n^2}.$$

Temos, então,

$$\left|\widehat{F}_n(x)\widehat{F}_n(y) - F(x)F(y)\right| = O\left(\frac{(\log\log n)^2}{n}\right)^{\frac{1}{3}}.$$

Assim, a sucessão q_n deve ser escolhida de tal maneira que

$$q_n \left(\frac{(\log\log n)^2}{n}\right)^{\frac{1}{3}} \underset{n\to\infty}{\longrightarrow} 0.$$

Ora, tal como para o termo estudado anteriormente, podemos escolher $q_n = O\big(\log\log n\big)^{\frac{\zeta}{3}}$, com $\zeta > 0$.

Temos que a quantidade $(2q_n+1)\left|\widehat{F}_n(x)\widehat{F}_n(y) - F(x)F(y)\right|$ é da ordem de $O\left(\frac{(\log\log n)^{\zeta+2}}{n}\right)^{\frac{1}{3}}$ com $\zeta > 0$.
Vejamos agora a parcela relativa às séries

$$\sum_{k=q_n+1}^{\infty}\Big(F_k(x,y) - F(x)F(y)\Big) + \sum_{k=q_n+1}^{\infty}\Big(F_k(y,x) - F(x)F(y)\Big).$$

Como foi já referido por diversas vezes, o termo geral destas séries é majorado por $M\mathrm{Cov}^{1/3}(X_1, X_{k+1})$. Tomando, como anteriormente, o termo geral da sucessão $\{u_k\}_{k\in\mathbb{N}}$, da ordem de $\dfrac{q_n}{n}$, temos $\sum_{k=q_n+1}^{\infty}\mathrm{Cov}^{1/3}(X_1, X_{k+1}) = O\left(\frac{q_n}{n}\right)$. Se tomarmos, tal como havemos feito, $q_n = O(\log\log n)^{\frac{\zeta}{3}}, \zeta > 0$, obtemos

$$\Big|\sum_{k=q_n+1}^{\infty}\Big(F_k(x,y) - F(x)F(y)\Big) + \sum_{k=q_n+1}^{\infty}\Big(F_k(y,x) - F(x)F(y)\Big)\Big| = O\Big(\frac{\log\log n}{n}\Big)^{-\frac{\zeta}{3}}.$$

Estamos agora em condições de apresentar o resultado relativo à da velocidade da convergência quase certa do estimador $\widehat{\Gamma}_n$.

Teorema 3.2.7 *Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão estritamente estacionária de variáveis aleatórias associadas que verificam as condições $(C_2), (C_5), (C_6)$ e $(C_7)'$. Se $q_n = O\big(\log\log n\big)^{\frac{\zeta}{3}}$, com $\zeta > 0$, para cada $(x,y) \in \mathbb{R}^2$,* $\left|\Gamma(x,y) - \widehat{\Gamma}_n(x,y)\right| = O\left(\frac{(\log\log n)^{\zeta+2}}{n}\right)^{\frac{1}{3}}$.

Demonstração: Consideremos $\zeta > 0$ e sejam C_1, C_2, C_3 e C_4 constantes positivas. Para cada $x, y \in \mathbb{R}$,

$$\begin{aligned}
\Big|\Gamma(x,y) - \widehat{\Gamma}_n(x,y)\Big| \leq & \Big|F\big(\min\{x,y\}\big) - \widehat{F}_n\big(\min\{x,y\}\big)\Big| + \\
& + (2q_n+1)\Big|\widehat{F}_n(x)\widehat{F}_n(y) - F(x)F(y)\Big| + \\
& + \sum_{k=1}^{q_n}\Big|F_k(x,y) - \widehat{F}_{n,k}(x,y)\Big| + \sum_{k=1}^{q_n}\Big|F_k(y,x) - \widehat{F}_{n,k}(y,x)\Big| + \\
& + \left|\sum_{k=q_n+1}^{\infty}\Big(F_k(x,y) - F(x)F(y)\Big) + \sum_{k=q_n+1}^{\infty}\Big(F_k(y,x) - F(x)F(y)\Big)\right| \leq \\
& \leq C_1\left(\frac{(\log\log n)^2}{n}\right)^{1/3} + C_2\left(\frac{(\log\log n)^{\zeta+2}}{n}\right)^{\frac{1}{3}} + \\
& + 2C_3\left(\frac{(\log\log n)^{\zeta+2}}{n}\right)^{1/3} + 2C_4\big(\frac{\log\log n}{n}\big)^{\frac{\zeta}{3}} = \\
& = O\left(\frac{(\log\log n)^{\zeta+2}}{n}\right)^{\frac{1}{3}}.\blacksquare
\end{aligned}$$

3.3 Os valores próprios de $\widehat{\Gamma}_n$

Uma vez estabelecida a convergência da sucessão $\big\{\widehat{\Gamma}_n(s,t)\big\}_{n\in\mathbb{N}}$, para cada $(s,t) \in \mathbb{R}^2$, devemos obter uma relação entre os valores próprios do estimador e os valores próprios de Γ. Nesta secção dedicamo-nos ao estudo da convergência quase certa dos valores próprios de $\widehat{\Gamma}_n$ para os correspondentes valores próprios de Γ. Usaremos resultados clássicos da Análise Funcional relativos aos valores próprios de operadores lineares compactos dos quais fazem parte, como sabemos, os operadores de covariância. A nossa referência principal para esta secção são Dunford e Schwartz [12].

Em primeiro lugar, notemos que para cada n, $\widehat{\Gamma}_n$ é o operador de covariância associado à função de covariância $\widehat{\Gamma}_n(s,t)$, isto é, para cada

$f \in L^2([a,b])$,

$$\widehat{\Gamma}_n(f(s)) = \int_{[a,b]} \widehat{\Gamma}_n(s,t)\, f(t)dt.$$

Como vimos, a sucessão de estimadores $\left\{\widehat{\Gamma}_n(s,t)\right\}_{n\in\mathbb{N}}$ converge quase certamente e em média quadrática, para cada $(s,t) \in \mathbb{R}^2$, para a função de covariância $\Gamma(s,t)$, associada ao operador integral Γ. Vejamos, agora, que o estudo efectuado da convergência em média quadrática desta sucessão de covariâncias é suficiente para concluir acerca da convergência dos valores próprios do estimador. Em todo o texto que se segue designaremos por $K, K_0, K_1, K_2, \ldots, K_n$ e S operadores de covariância genéricos definidos em $L^2([a,b])$ e denotamos por $\lambda_r(K)$ o $r-$ésimo valor próprio do operador K, supondo uma ordenação não crescente $\lambda_1(K) \geq \lambda_2(K) \geq \ldots$ dos valores próprios do operador em causa (neste caso K). Os resultados que apresentamos a seguir podem ser encontrados em Dunford e Shwartz [12].

Teorema 3.3.1 (Smithies, 1962) *Sejam K_1 e K_2 dois operadores de covariância definidos em $L^2([a,b])$. Considere-se $K_0 = K_1 + K_2$. Então,*

$$\lambda_{p+q-1}(K_0) \leq \lambda_p(K_1) + \lambda_q(K_2).$$

Demonstração: Para $p \geq 1$, seja M_p um subespaço de $L^2([a,b])$ cuja dimensão é, no máximo, $p-1$. Seja N_p a esfera unitária de $M_p^{\perp}$. Designemos por

$$\mu(K, M_p) = \sup_{\Phi \in N_p} \left\langle K(\Phi), \Phi \right\rangle.$$

Sabemos que $\lambda_p(K) = \inf_{M_p} \mu(K, M_p)$, onde o ínfimo é tomado sobre todos os subespaços M_p de dimensão, no máximo, $p-1$.

Sejam $\lambda_1(K_1), \ldots, \lambda_{p-1}(K_1)$ os primeiros $p-1$ valores próprios de K_1 com funções próprias associadas $f_1, \ldots, f_{p-1}$. Seja, ainda, $M_p^{(1)}$ o subespaço gerado por $f_1, \ldots, f_{p-1}$. Considere-se, analogamente, $\lambda_1(K_2), \ldots, \lambda_{q-1}(K_2)$ os primeiros $q-1$ valores próprios de K_2 com funções próprias associadas $g_1, \ldots, g_{q-1}$, sendo $M_p^{(2)}$ o subespaço gerado por

$g_1, \ldots, g_{q-1}$. Ora $M = M_p^{(1)} + M_q^{(2)}$ é o subespaço gerado por $f_1, \ldots, f_{p-1}$ e $g_1, \ldots, g_{q-1}$. Então, M tem, no máximo, dimensão $p+q-2$ pelo que,

$$\lambda_{p+q-1}(K_0) \leq \mu\big(K_0, M_p^{(1)} + M_q^{(2)}\big).$$

Como $K_0 = K_1 + K_2$, vem

$$\lambda_{p+q-1}(K_0) \leq \mu\big(K_1, M_p^{(1)} + M_q^{(2)}\big) + \mu\big(K_2, M_p^{(1)} + M_q^{(2)}\big).$$

Mas $M_p^{(1)} \subset M_p^{(1)} + M_q^{(2)}$ pelo que a esfera unitária N de $M^{\perp}$ está incluída em $N_p^{(1)}$, esfera unitária de $\big(M_p^{(1)}\big)^{\perp}$. Temos então,

$$\mu\big(K_1, M_p^{(1)} + M_q^{(2)}\big) \leq \sup_{\Phi \in N_p^{(1)}} \big\langle K_1(\Phi), \Phi \big\rangle = \mu\big(K_1, M_p^{(1)}\big).$$

De igual modo se tem,

$$\mu\big(K_2, M_p^{(1)} + M_q^{(2)}\big) \leq \mu\big(K_2, M_q^{(1)}\big),$$

pelo que

$$\lambda_{p+q-1}(K_0) \leq \mu\big(K_1, M_p^{(1)}\big) + \mu\big(K_2, M_q^{(1)}\big).$$

Como $\mu\big(K_1, M_p^{(1)}\big) = \lambda_p(K_1)$ e $\mu\big(K_2, M_q^{(1)}\big) = \lambda_q(K_2)$. O resultado segue agora imediatamente. ■

Teorema 3.3.2 *Sejam K e S dois operadores de covariância definidos em $L^2([a,b])$. Então,*

$$\Big|\lambda_r(K) - \lambda_r(S)\Big| \leq \rho\big(K - S\big),$$

onde $\rho(L)$ é o raio espectral do operador L.

Demonstração: Aplicando o resultado anterior com $p = r, q = 1, K_0 = S, K_1 = K$, e $K_2 = S - K$, temos

$$\lambda_r(S) \leq \lambda_r(K) + \lambda_1(S-K) \Leftrightarrow \lambda_r(S) - \lambda_r(K) \leq \lambda_1(K-S) \leq \rho(K-S).$$

Trocando agora os papéis de S e K, o resultado fica provado. ■

O teorema que apresentamos de seguida estabelece ser suficiente a convergência, relativamente à norma de $L^2([a,b])$, da sucessão de operadores de covariância para que os valores próprios correspondentes aos operadores da referida sucessão, convirjam para os respectivos valores próprios do operador de covariância limite. Este é, uma vez mais, consequência de um resultado clássico da teria espectral de operadores compactos, que garante o facto de o raio espectral de um operador compacto ser majorado pela sua norma.

Teorema 3.3.3 *Seja $\{K_n\}_{n\in\mathbb{N}}$ uma sucessão de operadores de covariância definidos em $L^2([a,b])$. Suponhamos que $\{K_n\}_{n\in\mathbb{N}}$ converge relativamente à norma de $L^2\left([a,b]^2\right)$, para o operador de covariância K. Então,*

$$\lim_{n\to\infty}\lambda_r(K_n)=\lambda_r(K) \;\; e \;\; \big|\lambda_r(K)-\lambda_r(K_n)\big|\leq \big\|K-K_n\big\|_2,$$

onde $\|\cdot\|_2$ representa a norma em $L^2\left([a,b]^2\right)$.

Demonstração: Em primeiro lugar, observe-se que dado λ um valor próprio qualquer de K, $|\lambda|\leq\|K\|_2$. De facto, se g é uma função própria associada a λ,

$$\begin{aligned}\big\|\lambda\, g\big\|_2&=|\lambda|\,\big\|g\big\|_2=\\&=\Big\|\int_{[a,b]}K(x,y)g(y)\,dy\Big\|_2\leq\left(\int_{[a,b]^2}\big|K(x,y)\big|^2dx\,dy\right)^{1/2}\|g\|_2.\end{aligned}$$

Consequentemente o raio espectral do operador K é majorado pela sua norma em $L^2\left([a,b]\right)$, isto é, $\rho(K)\leq\big\|K\big\|_2$. Assim, temos imediatamente que

$$\big|\lambda_r(K)-\lambda_r(K_n)\big|\leq\rho\big(K-K_n\big)\leq\big\|K-K_n\big\|_2$$

pelo que, se $\big\|K-K_n\big\|_2\to 0$, também se tem $\big|\lambda_r(K)-\lambda_r(K_n)\big|\to 0$. ■

Enunciamos, por último, o resultado que relaciona os valores próprios,

ordenados por ordem não crescente $\lambda_1, \lambda_2, \ldots$, de Γ com os respectivos valores próprios $\widehat{\lambda}_{1,n}, \widehat{\lambda}_{2,n}, \ldots$, de $\widehat{\Gamma}_n$, que resulta essencialmente da convergência em média quadrática, assim como da respectiva velocidade, da sucessão de funções de covariância anteriormente estudada.

Teorema 3.3.4 *Suponhamos que $\{X_n\}_{n\in\mathbb{N}}$ é uma sucessão estritamente estacionária de variáveis aleatórias associadas com sucessão de covariâncias $\left\{\mathrm{Cov}\,(X_1, X_j)\right\}_{j\in\mathbb{N}}$ decrescente, que satisfaz todas as condições do Teorema 3.2.3. Então,*

$$\mathbb{E}\left|\lambda_i - \widehat{\lambda}_{i,n}\right|^2 = O\left(\frac{q_n^2}{n} + u^2(q_n+1)\right)$$

onde q_n e $u(n)$ são as sucessões definidas nas subsecções anteriores.

Demonstração: Para qualquer $f \in L^2([a,b])$, efectuando cálculos como no início da demonstração do Teorema 3.0.11, temos

$$\left\|\left(\Gamma - \widehat{\Gamma}_n\right)(f(s))\right\|_2^2 \leq \|f\|_2^2 \int_a^b \int_a^b \left|\Gamma(s,t) - \widehat{\Gamma}_n(s,t)\right|^2 dsdt.$$

Nestas condições, uma vez que $f \in L^2([a,b])$,

$$\mathbb{E}\left(\left\|\left(\Gamma - \widehat{\Gamma}_n\right) f(s)\right\|_2^2\right) \leq \int_a^b \int_a^b \mathbb{E}\left(\Gamma(s,t) - \widehat{\Gamma}_n(s,t)\right)^2 dsdt.$$

Assim, e nas condições do Teorema 3.2.3,

$$\mathbb{E}\left(\left\|\left(\Gamma - \widehat{\Gamma}_n\right) f(s)\right\|_2^2\right) = O\left(\frac{q_n^2}{n} + u^2(q_n+1)\right),$$

pelo que, de acordo com o Teorema anterior, se tem

$$\mathbb{E}\left|\lambda_i - \widehat{\lambda}_{i,n}\right|^2 = O\left(\frac{q_n^2}{n} + u^2(q_n+1)\right). \quad \blacksquare$$

Assim, se calcularmos numericamente uma estimativa $\widehat{\lambda}_1$ para o maior valor próprio do estimador $\widehat{\Gamma}_n$, o erro quadrático médio cometido na aproximação de λ_1 por $\widehat{\lambda}_1$ é da ordem de $\frac{q_n^2}{n} + u^2(q_n+1)$.

3.4 Um método numérico

Uma vez controlado o erro cometido quando se calculam os valores próprios de uma estimativa de $\widehat{\Gamma}_n$ para aproximar pelos valores próprios de Γ, importa agora encontrar um método para o cálculo efectivo de tais valores próprios.
Ora, como sabemos, na maior parte dos casos não existe uma via analítica para a resolução do problema. De facto, veremos no capítulo seguinte que as expressões que obtemos para as estimativas das covariâncias $\widehat{\Gamma}_n(s,t)$, apesar de serem funções contínuas e simétricas de $L^2([a,b]^2)$, são muito difíceis de manusear do ponto de vista algébrico. Daí o interesse na obtenção de um método iterativo para o cálculo de, pelo menos, o maior valor próprio.
O nosso objectivo é calcular a ordem de grandeza do maior valor próprio de Γ através do cálculo do maior valor próprio de uma estimativa de Γ. Consideremos então o *Teorema de Hilbert - Schmidt* (em, por exemplo, Dunford e Schwartz [12]).

Teorema 3.4.1 *Dado qualquer operador linear simétrico T, num espaço de Hilbert H, existe uma sucessão ortornormada $\{\phi_n\}_{n\in\mathbb{N}}$ de funções próprias correspondentes aos valores próprios $\{\lambda_n\}$ $(\lambda_n \neq 0)$, tal que todo o elemento $f \in H$ admite uma única decomposição da forma*

$$f(t) = \sum_k c_k \phi_k(t) + f'(t), \quad com \quad f' \in Ker(T).$$

Nestas condições,

$$T(f) = \sum_k \lambda_k \, c_k \, \phi_k(t).$$

Além disso, se a sucessão $\{\phi_n\}_{n\in\mathbb{N}}$ for infinita, $\lim_{n\to\infty} \lambda_n = 0$.

Este resultado dá-nos condições para apresentar um método iterativo para o cálculo numérico do maior valor próprio de $\widehat{\Gamma}_n$, para um dada amostra de dimensão n. De facto, seja T um operador de covariância qualquer

e seja λ o seu maior valor próprio. Temos que $\lambda = \sup_{\|f\|\leq 1} T(f)$, e, além disso, se g é uma função própria associada a λ, $T(g) = \dfrac{\langle Tg, g\rangle}{\|g\|^2} = \lambda$.
Seja $g^{(0)}$ uma função qualquer de $L^2(\mathbb{R})$, tal que $\left\|g^{(0)}\right\| = 1$. De acordo com o Teorema anterior, $g^{(0)} = \sum_{k=1}^{\infty} c_k\phi_k + \left(g^{(0)}\right)'$, com $\left(g^{(0)}\right)' \in Ker(T)$. Assim, e uma vez que $T(\phi_k) = \lambda_k\phi_k$,

$$Tg^{(0)} = \sum_{k=1}^{\infty} c_k\lambda_k\phi_k$$

e, fazendo $g^{(1)} = Tg^{(0)}$, apelando à linearidade de T e ao facto de ϕ_k ser uma função própria associada ao valor próprio λ_k,

$$Tg^{(1)} = T^2g^{(0)} = \sum_{k=1}^{\infty} c_k\lambda_k^2\phi_k.$$

Temos então, para um j qualquer,

$$g^{(j+1)} = \sum_{k=1}^{\infty} c_k\lambda_k^j\phi_k, \quad Tg^{(j+1)} = g^{(j+2)} = \sum_{k=1}^{\infty} c_k\lambda_k^{j+1}\phi_k.$$

Nestas condições,

$$\left\langle g^{(j+2)}, g^{(j+1)}\right\rangle = \sum_{k,k'} c_k\lambda_k^{j+2}\lambda_k^{j+1}c_{k'}\left\langle \phi_k, \phi_{k'}\right\rangle.$$

Uma vez que as funções ϕ_k são ortornormais,

$$\left\langle g^{(j+2)}, g^{(j+1)}\right\rangle = \sum_{k} c_k^2\,\lambda_k^{2j+3}.$$

Além disso, tem-se ainda,

$$\left\langle g^{(j+1)}, g^{(j+1)}\right\rangle = \sum_{k} c_k^2\,\lambda_k^{2j+2}.$$

Vejamos agora que este esquema iterativo converge, quando o número de iterações tende para infinito, para o maior valor próprio de T.

Teorema 3.4.2 *Seja T um operador de covariância e seja $\{\lambda_n\}$ a sucessão dos seus valores próprios com $\lambda_1 > \lambda_2 \geq \lambda_3, \ldots$ O esquema iterativo anterior implica que*

$$\lim_{j\to\infty} \frac{\langle Tg^{j+1}, g^{j+1}\rangle}{\langle g^{j+1}, g^{j+1}\rangle} = \lambda_1.$$

Demonstração: Vejamos que o método iterativo exposto converge para o maior valor próprio de T.

$$\frac{\langle Tg^{j+1}, g^{j+1}\rangle}{\langle g^{j+1}, g^{j+1}\rangle} = \frac{\langle g^{(j+2)}, g^{(j+1)}\rangle}{\langle g^{(j+1)}, g^{(j+1)}\rangle} =$$
$$= \frac{\sum_k c_k^2 \lambda_k^{2j+3}}{\sum_k c_k^2 \lambda_k^{2j+2}} = \frac{\lambda_1^{2j+3}}{\lambda_1^{2j+2}} \frac{c_1^2 + \sum_{k=2}^{\infty} c_k^2 \left(\frac{\lambda_k}{\lambda_1}\right)^{2j+3}}{c_1^2 + \sum_{k=2}^{\infty} c_k^2 \left(\frac{\lambda_k}{\lambda_1}\right)^{2j+2}}$$

Sendo λ_1 o maior valor próprio de T, a sucessão dos valores próprios $\{\lambda_n\}_n$ decrescente e $\sum_k c_k^2 < \infty$, tem-se que,

$$0 \leq \lim_{j\to\infty} \sum_{k=2}^{\infty} c_k^2 \left(\frac{\lambda_k}{\lambda_1}\right)^{2j+2} \leq \sum_{k=2}^{\infty} c_k^2 \lim_{j\to\infty} \left(\frac{\lambda_2}{\lambda_1}\right)^{2j+2} = 0$$

e analogamente

$$0 \leq \lim_{j\to\infty} \sum_{k=2}^{\infty} c_k^2 \left(\frac{\lambda_k}{\lambda_1}\right)^{2j+3} \leq \sum_{k=2}^{\infty} c_k^2 \lim_{j\to\infty} \left(\frac{\lambda_2}{\lambda_1}\right)^{2j+3} = 0.$$

Assim,

$$\lim_{j\to\infty} \frac{\lambda_1^{2j+3}}{\lambda_1^{2j+2}} \frac{c_1^2 + \sum_{k=2}^{\infty} c_k^2 \left(\frac{\lambda_k}{\lambda_1}\right)^{2j+3}}{c_1^2 + \sum_{k=2}^{\infty} c_k^2 \left(\frac{\lambda_k}{\lambda_1}\right)^{2j+2}} = \lambda_1. \blacksquare$$

A justificação deste método e a prova da sua convergência abre caminho para a solução do problema de encontrar limites críticos para a estatística de Cramér-von Mises, conforme justificámos na introdução do presente capítulo. A sua implementação prática traz alguns problemas que abordamos, apresentando, também, soluções possíveis no capítulo seguinte.

Capítulo 4

Ilustração dos resultados

Este capítulo é destinado a ilustrar os resultados dos capítulos precedentes, onde pretendemos essencialmente apresentar aproximações para as funções cujo comportamento estudámos, obtidas com base em simulações de sucessões associadas de variáveis aleatórias. Além disso, é também objectivo obter a ordem de grandeza do maior valor próprio do operador de covariância do processo limite do processo empírico construído com sucessões associadas de variáveis aleatórias, utilizando o método numérico apresentado no terceiro capítulo.

Para a obtenção das aproximações das funções é fundamental a escolha da função núcleo e da sucessão de comprimentos das janelas $\{h_n\}_{n\in\mathbb{N}}$. Relativamente às estimativas para $\widehat{\Gamma}_n$, a escolha do número de termos a somar, isto é, a optimização da sucessão $\{q_n\}_{n\in\mathbb{N}}$ introduzida no capítulo anterior, que verifique as condições suficientes obtidas, também merece uma atenção especial. Justificamos a escolha destas sucessões na primeira secção.

A escolha da sucessão $\{q_n\}_{n\in\mathbb{N}}$ é também justificada tendo como base o estudo teórico dos termos a considerar efectuado no segundo e terceiro capítulos.

Relativamente à obtenção da sucessão óptima $\{h_n\}_{n\in\mathbb{N}}$, o estudo da op-

timização do Erro Quadrático Médio do estimador levado a cabo no segundo capítulo, conduz a uma expressão para o termo geral desta sucessão.
A segunda secção deste capítulo é dedicada à apresentação de quatro exemplos de estimativas para as funções de distribuição marginal e bivariadas de sucessões associadas de variáveis aleatórias, assim como para a correspondente função de covariância.
Na terceira secção, analisamos os resultados que obtemos com o programa para o cálculo do maior valor próprio das aproximações para as funções de covariância.

4.1 Preliminares

Nesta secção justificamos a escolha dos termos gerais das sucessões $\{h_n\}_{n\in\mathbb{N}}$ e $\{q_n\}_{n\in\mathbb{N}}$ tendo como base o estudo efectuado no capítulo segundo e terceiro desta monografia.

4.1.1 Optimização de $\{h_n\}_{n\in\mathbb{N}}$.

Retomemos o estudo do segundo capítulo, aquando da optimização do Erro Quadrático Médio do estimador.
Consideremos, como habitualmente, $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão associada e estritamente estacionária de variáveis aleatórias reais. Para $k = 1, \ldots, n-1$, fixo e $j > 1$, sejam F_k e $F_{k,j}$, respectivamente, as funções de distribuição do par (X_1, X_{k+1}) e do vector $(X_1, X_{k+1}, X_j, X_{k+j})$. Recordemos a definição das funções V_1, V_2 e V_4, de $\mathbb{R}^2$ para $\mathbb{R}$, introduzidas atrás.

$$\begin{aligned} V_1(x,y) &= \frac{\partial^2 F_k}{\partial x^2}(x,y)\int_{\mathbb{R}^2} s^2 u(s,t)\,ds\,dt + \frac{\partial^2 F_k}{\partial x\,\partial y}(x,y)\int_{\mathbb{R}^2} s\,t\,u(s,t)\,ds\,dt + \\ &\quad + \frac{\partial^2 F_k}{\partial y^2}(x,y)\int_{\mathbb{R}^2} t^2 u(s,t)\,ds\,dt \end{aligned}$$

$$V_2(x,y) = \frac{\partial F_k}{\partial x}(x,y)\int_{\mathbb{R}^2} s\,V(s,t)\,ds\,dt + \frac{\partial F_k}{\partial y}(x,y)\int_{\mathbb{R}^2} t\,V(s,t)\,ds\,dt,$$

$$\begin{aligned} V_4(x,y) = {} & \frac{\partial^2 F_{k,j}}{\partial x_1^2}(x,y,x,y)\int_{\mathbb{R}^2} s^2 u(s,t)\,ds\,dt + \\ & + \frac{\partial^2 F_{k,j}}{\partial x_2^2}(x,y,x,y)\int_{\mathbb{R}^2} t^2 u(s,t)\,ds\,dt + \\ & + \frac{\partial^2 F_{k,j}}{\partial x_3^2}(x,y,x,y)\int_{\mathbb{R}^2} s^2 u(s,t)\,ds\,dt + \\ & + \frac{\partial^2 F_{k,j}}{\partial x_4^2}(x,y,x,y)\int_{\mathbb{R}^2} t^2 u(s,t)\,ds\,dt + \\ & + 2\times \frac{\partial^2 F_{k,j}}{\partial x_1\,\partial x_2}(x,y,x,y)\int_{\mathbb{R}^2} s\,t\,u(s,t)\,ds\,dt + \\ & + 2\times \frac{\partial^2 F_{k,j}}{\partial x_3\,\partial x_4}(x,y,x,y)\int_{\mathbb{R}^2} s\,t\,u(s,t)\,ds\,dt, \end{aligned}$$

onde $V = \dfrac{\partial^2 U^2}{\partial x \partial y}$ conforme a condição (C_8) do segundo capítulo. Consideremos ainda, tal como no segundo capítulo,

$$\sigma^2(x,y) = F_k(x,y) - F_k^2(x,y) + 2\sum_{j=1}^{\infty}\big(F_{k,j}(x,y,x,y) - F_k^2(x,y)\big);$$

e $\{c_n\}_{n\in\mathbb{N}}$ sucessão, independente de h_n, convergente para zero, dada por

$$\begin{aligned} c_n = -2 \quad \times \quad & \left(\sum_{j=n-k-1}^{\infty} \big(F_{k,j}(x,y,x,y) - F_k^2(x,y)\big) + \right. \\ + \quad & \left. \frac{1}{n-k}\sum_{j=2}^{n-k}(j-1)\big(F_{k,j}(x,y,x,y) - F_k^2(x,y)\big) \right). \end{aligned}$$

Provámos já que

$$\begin{aligned} & (n-k)\mathrm{MSE}\left(\widehat{F}_n(x,y)\right) = \\ & = \sigma^2(x,y) - h_n\,V_2(x,y) + (n-k-1)h_n^2\big(V_4(x,y) - F(x,y)\,V_1(x,y)\big) + O(h_n^2). \end{aligned}$$

Da minimização do MSE $\left(\widehat{F}_n(x,y)\right)$ em ordem a h_n, vem

$$h_n = \frac{V_2(x,y)}{(n-k-1)(V_4(x,y) - F_k(x,y)\, V_1(x,y))}.$$

Como as funções V_1, V_2 e V_4 dependem de $F_k, F_{k,j}$ e das suas derivadas parciais até à segunda ordem, sendo estas desconhecidas, apenas podemos ter uma estimativa para h_n baseada em estimativas para as funções de distribuição bivariadas e multivariadas das variáveis aleatórias.

Um procedimento usual nestas situações, designado por método de "plug-in" usado, por exemplo em Cai, Roussas [8], 1998, para o caso univariado, é o seguinte: assumir um determinado valor inicial para h_n, seja h_n^0; calcular as estimativas para F_k e $F_{k,j}$ com base neste comprimento; calcular as derivadas parciais até à segunda ordem destas funções e, posteriormente, o valor estimado para o coeficiente de h_n a usar nas simulações.

Gijbels [16], 1998, apresenta também um método prático para a selecção dos comprimentos das janelas. Supõe-se que as funções de distribuição desconhecidas, F_k e $F_{k,j}$ pertencem a uma determinada família de distribuições. São, por exemplo, Gaussianas com média e matriz de covariância não especificadas. Calculam-se estimativas pontuais para estes parâmetros em função das amostras que temos, calculando de seguida o comprimento óptimo com base nestas distribuições.

Note-se que se substituirmos a condição (C_6) sobre os comprimentos das janelas por $n\, h_n^4 \to 0$ e se a condição (C_7) sobre a sucessão $\{n \text{ Cov}\,(X_1, X_n)\}_{n\in\mathbb{N}}$ for substituída por $\sum_{j=1}^{\infty}\Big(\text{Cov}\,(X_1, X_{j+1})\Big)^{1/3(1-\tau)}$, $0 < \tau < 1$, obtemos, conforme Corolário 2.3.9 do segundo capítulo, uma escolha óptima para h_n da mesma ordem que $\left(\frac{V_2(x,y)}{V_1^2(x,y)(n-k-1)}\right)^{1/3}$. Do ponto de vista prático, esta escolha para h_n é mais simples de obter, por envolver apenas as funções V_2 e V_1 que dependem única e exclusivamente da função de distribuição F_k, sendo esta a única a ser previamente estimada ou com uma distribuição que especificamos para iniciar o procedimento.

4.1.2 A escolha para $\{q_n\}_{n\in\mathbb{N}}$.

Retomemos a definição do estimador $\widehat{\Gamma}_n$, de Γ, função de $\mathbb{R}^2$ para $\mathbb{R}$ tal que, para cada (x,y),

$$\widehat{\Gamma}_n(x,y) = \widehat{F}_n(\min\{x,y\}) - (2\,q_n+1)\widehat{F}_n(x)\widehat{F}_n(y) + \sum_{k=1}^{q_n}\big(\widehat{F}_{n,k}(x,y)+\widehat{F}_{n,k}(y,x)\big).$$

Como foi já visto no terceiro capítulo, para além da necessidade de impor condições à sucessão $\{q_n\}_{n\in\mathbb{N}}$, é também imposta uma condição para a sucessão dos restos da série $\sum_n \mathrm{Cov}^{1/3}(X_1,X_n)$. Mas, como oportunamente foi referido, esta sucessão pode ser sempre definida de tal maneira que não interfira no resultado final. Assim, designando por $m=m(n)$ uma sucessão de números positivos tal que $m=o(n)$ e, definindo $u_m=\sum_{j=m+1}^{\infty}\mathrm{Cov}^{1/3}(X_1,X_j)$, se $u_m=O\left(\frac{m}{n}\right)$, o resultado para a velocidade de convergência nunca depende desta sucessão. Podemos, assim, concentrar-nos apenas em q_n.
Vimos, no capítulo anterior, que a velocidade de convergência optima é, no caso da convergência quase certa, da mesma ordem de $(\log\log n)^{\frac{\zeta}{3}}$, para algum $\zeta>0$. De um ponto de vista prático e bastante informal, isto significa que necessitamos de envolver muito poucos termos no somatório relativamente à dimensão da amostra. Todas as constantes intervenientes neste resultado são funções das derivadas parciais, de primeira e segunda ordens, das funções que pretendemos estimar, F_k e $F_{k,j}$. Temos ainda, como referência, a velocidade de convergência da sucessão definida por $u(n)=\sum_{k\geq n}\mathrm{Cov}^{1/3}(X_1,X_k)$, que deve ser ligeiramente superior à de q_n. Assim, no caso de, por exemplo, $m-$associação, o número de termos a somar nunca deverá ser superior ou igual a m. Contrariamente à escolha de h_n não temos, neste caso, uma única condição para a escolha óptima de q_n, mas sim um conjunto de considerações que devemos ter em conta aquando da sua escolha.

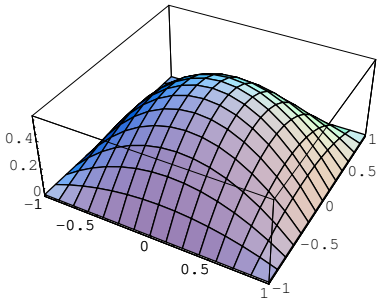
4.1.3 As funções núcleo.

Para a escolha das funções núcleo, tomamos como referência o trabalho efectuado por Gijbels [16], 1998. As funções mais usadas no caso unidimensional são as seguintes:

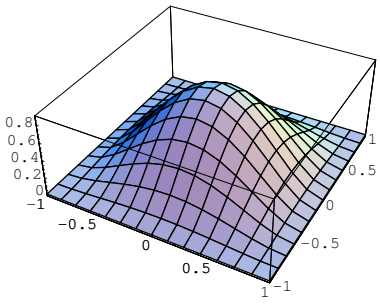
- Função núcleo *Epanechnikov*. Esta função é definida, para cada $x \in [-1,1]$ por $u(x) = \frac{3}{4}(1-u^2)$;
- Função núcleo *Biweight*, definida para cada $x \in [-1,1]$ por $u(x) = \frac{15}{16}(1-u^2)^2$;
- Função núcleo *Triweight*, definida para cada $x \in [-1,1]$ por $u(x) = \frac{35}{32}(1-u^2)^3$;
- Função núcleo *Gaussiana*, definida para cada $x \in \mathbb{R}$, por $u(x) = \frac{1}{\sqrt{2\pi}} \exp\{-1/2u^2\}$;
- Função núcleo *Triangular*, definida para cada $x \in [-1,1]$ por $u(x) = (1-|x|)$;
- Função núcleo *Uniforme*, definida para cada $x \in [-1,1]$ por, $u(x) = \frac{1}{2}$.

A passagem natural de núcleos unidimensionais para multidimensionais, consiste em considerar o produto tensorial entre as respectivas funções. Assim, para as nossas simulações, consideramos o núcleo

- *Epanechnikov bivariado*, definido para cada par $(x,y) \in [-1,1]^2$ por, $u(x,y) = \frac{9}{16}(1-x^2)(1-y^2)$, cuja representação gráfica é



- e ainda o núcleo *Biweight bivariado* definido para cada $x, y \in \mathbb{R}$ por $u(x,y) = \left(\frac{15}{12}\right)^2 \left(1-x^2\right)^2 \left(1-y^2\right)^2$ cujo gráfico é



Estas funções obedecem às condições teóricas estipuladas no segundo capítulo.

4.2 Simulações

Vamos, nesta secção, estudar e relacionar quatro exemplos de obtenção de estimativas para as funções de distribuição marginais e conjuntas de sucessões associadas de variáveis aleatórias identicamente distribuídas, $\{X_n\}_{n\in\mathbb{N}}$, simuladas. Estas sucessões verificam as condições suficientes

impostas no segundo capítulo deste trabalho e são construídas de maneira a que possamos identificar as suas funções de distribuição conjunta e, consequentemente, conhecer a verdadeira função de covariância, com o objectivo de estabelecer comparações.
No primeiro exemplo, geramos variáveis aleatórias com distribuição uniforme, $m-$associadas, a partir de uma sucessão de variáveis aleatórias independentes Gaussianas, reduzidas. O segundo exemplo é uma aplicação do resultado de Pitt [38], 1982, em que a sucessão associada gerada é constituída por variáveis com distribuição de Gauss. No terceiro exemplo geramos também variáveis Gaussianas mais uma vez recorrendo à $m-$associação, mas a escolha para m é de tal forma que a expressão para a função de covariância é bastante mais simples que a do exemplo anterior. Temos, finalmente, um quarto exemplo onde geramos variáveis aleatórias com distribuição exponencial de parâmetro unitário. Todos estes exemplos têm como base um procedimento genérico que se baseia essencialmente em:

- gerar, pseudo-aleatoriamente, n variáveis aleatórias $Y_1, \ldots, Y_n$, independentes, com distribuição especificada F,
- tomar, para cada $i = 1, \ldots, n$, $X_i = f_i(Y_1, \ldots, Y_n)$ com as funções f_i todas crescentes ou todas decrescentes.

Sabemos que esta construção origina sucessões associadas de variáveis aleatórias reais.
Vejamos, então, os exemplos.

Exemplo 4.2.1 *O procedimento utilizado para gerar sucessões associadas de variáveis uniformes, a partir de sucessões independentes de variáveis Gaussianas é o seguinte:*

- *Simulamos n variáveis aleatórias $Y_1, Y_2, \ldots, Y_n$ com distribuição de Gauss reduzida, isto é para cada i, $Y_i \sim \mathcal{N}(0,1)$;*

- *Para um determinado $m < n$, construímos a sucessão associada, definida para cada $j = 1, \ldots, n-m$, por*
  $L_j = \frac{1}{\sqrt{m+1}}\big(Y_j + Y_{j+1} + \cdots + Y_{j+m}\big)$;
- *A distribuição de cada L_j é Gaussiana reduzida. Seja, Φ a sua função de distribuição. Se considerarmos agora uma sucessão de variáveis X_j dada por $X_j = \Phi(L_j)$, esta é constituída por variáveis associadas com distribuição uniforme no intervalo* $[0, 1]$.

Vejamos agora qual a distribuição dos pares de variáveis associadas (X_1, X_{k+1}), com $k = 1, \ldots, n-m-1$.
Sabemos que para $k = 1, \ldots, m$, a distribuição dos pares (L_1, L_{k+1}) é Gaussiana centrada, com matriz de covariância dada por:

$$\begin{pmatrix} 1 & \frac{m-k+1}{m+1} \\ \frac{m-k+1}{m+1} & 1 \end{pmatrix}.$$

Para $k > m$, as variáveis aleatórias L_1 e L_{k+1} são independentes, sendo, neste caso a matriz de covariância a matriz identidade de ordem 2.
Consequentemente a expressão para a função de distribuição do par (X_1, X_{k+1}), com $k = 1, \ldots, m$, é dada por

$$F_k(x, y) = P(X_1 \leq x, X_{k+1} \leq y) = P(L_1 \leq \Phi^{-1}(x), L_{k+1} \leq \Phi^{-1}(y)),$$

sendo, para $k > m$, $F_k(x, y) = P(X_1 \leq x, X_{k+1} \leq y) = xy$.
Nestas condições temos a função de covariância

$$\begin{aligned} \Gamma(x, y) = \min\{x, y\} - xy \quad &+ \quad \sum_{k=1}^{m} \Big(P(X_1 \leq x, X_{k+1} \leq y) - xy \Big) \\ &+ \quad \sum_{k=1}^{m} \Big(P(X_1 \leq y, X_{k+1} \leq x) - xy \Big). \end{aligned}$$

Vejamos o resultado das aproximações obtidas usando uma amostra de $n = 500$, variáveis Gaussianas reduzidas. Fazemos $m = 50$. Simulamos,

assim, um total de 450 variáveis aleatórias uniformes associadas. Utilizamos, nestas simulações o núcleo bivariado *Epanechnikov*, tomando como comprimento da janela $h = 0.1$. As figuras seguintes mostram as funções em estudo assim como as respectivas aproximações. Assim, na figura 4.1.1, temos o gráfico da aproximação da função de distribuição marginal das variáveis $X_i, i = 1, \ldots, 450$ e a respectiva função na figura 4.1.2.

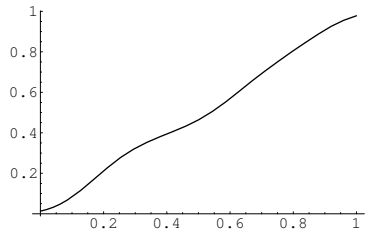


figura 4.1 .1

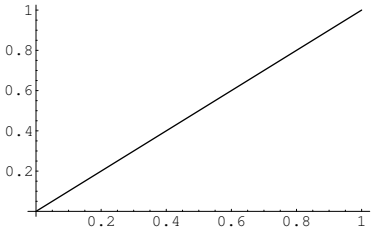


figura 4.1 .2

As figuras seguintes dizem respeito à aproximação para a função de distribuição do par (X_1, X_2), figura 4.1.3, e a figura 4.1.4 representa a função de distribuição de (X_1, X_2).

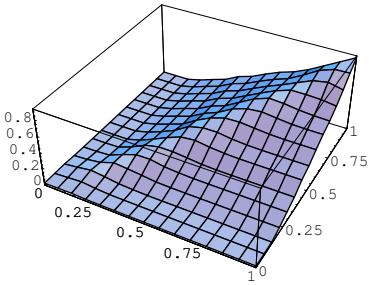


figura 4.1 .3

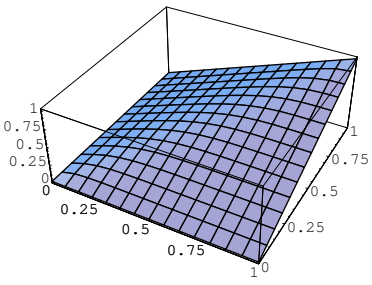


figura 4.1 .4

Vejamos, também, a aproximação para a função de distribuição de (X_1, X_{48}), figura 4.1.5 e a respectiva função de distribuição, não aproximada, na figura 4.1.6.

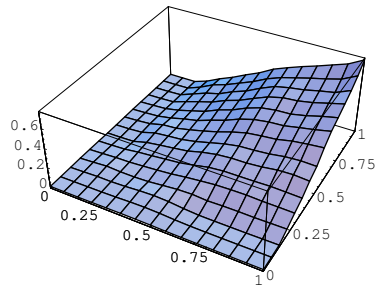


figura 4.1 .5

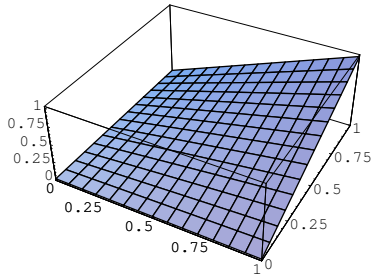


figura 4.1 .6

Por último mostramos os gráficos da aproximação para a função de covariância, obtida com uma escolha de $q = 10$ termos utilizados no somatório, figura 4.1.7, e da verdadeira função de covariância, figura 4.1.8.

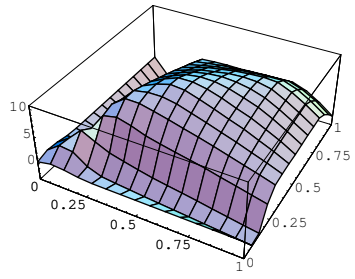


figura 4.1 .7

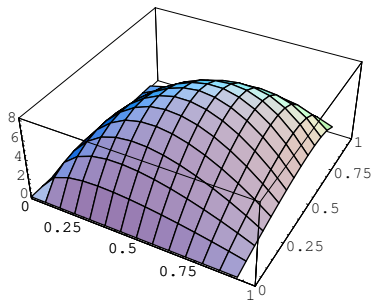


figura 4.1 .8

Convém referir que a escolha do número de termos envolvidos no somatório é importante, tendo sido, para cada caso, obtido um número mínimo de termos a somar. Se tomarmos, por exemplo, $q = 6$ obtemos o gráfico da figura 4.1.9 que, como se pode verificar, tem como máximo um valor inferior ao da verdadeira função assim como ao da aproximação com 10 termos no somatório. Vejamos, também, na figura 4.4.10, a aproximação para a função com $q = 20$ termos no somatório. Como se verifica, as alterações relativamente à figura 4.1.7 são pouco significativas.

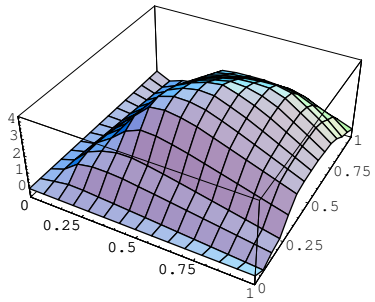


figura 4.1 .9

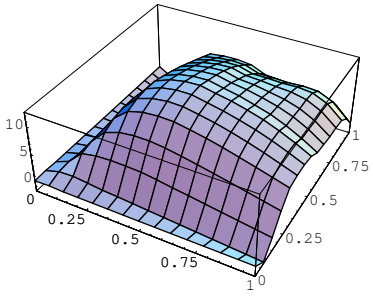


figura 4.1 .10

Nos dois exemplos que apresentamos de seguida geramos sucessões associadas de variáveis aleatórias Gaussianas.

Exemplo 4.2.2 *Consideremos a seguinte construção.*

- *Simulamos um vector* $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ *com distribuição Gaussiana* $n-$*variada, de tal maneira que a matriz de covariância tenha todas as entradas não negativas.*

- *Sabemos que, nessas condições, de acordo com Pitt [38], as variáveis* $X_j, j \geq 1$, *são associadas com distribuição Gaussiana com média e variâncias de acordo com a especificação do vector* $\mathbf{X}$. *Seja* F *a função de distribuição marginal comum das variáveis* $X_j, j \geq 1$.

- *Consideremos que as covariâncias entre as variáveis decrescem à medida que estas se afastam estabelecendo, por exemplo,* $\mathrm{Cov}\,(X_i, X_j) = \dfrac{0.5^{|j-i|}}{1-0.5^2}$ *originando uma matriz obviamente simétrica, com as entradas todas positivas e decrescendo à medida que se afastam da diagonal. Temos, assim, todos os pares aleatórios do tipo* (X_1, X_{k+1}) *com distribuição Gaussiana centrada e matriz de*

covariância dada por

$$\begin{pmatrix} \dfrac{1}{1-0.5^2} & \dfrac{0.5^k}{1-0.5^2} \\ \dfrac{0.5^k}{1-0.5^2} & \dfrac{1}{1-0.5^2}, \end{pmatrix}$$

para cada $k = 1, \ldots, n$.
Designemos, como habitualmente, por F_k *a função de distribuição do par* (X_1, X_{k+1}).

- *A função de covariância é dada por*

$$\Gamma(s,t) = F(\min\{s,t\}) - F(s)F(t) + \sum_{k=1}^{\infty} \big(F_k(s,t) + F_k(t,s) - 2F(s)F(t)\big).$$

Vejamos os resultados das aproximações das funções marginais, da função de distribuição bidimensional (X_1, X_2) e da função de covariância, com base numa amostra simulada de dimensão $n = 500$, usando a função núcleo *Biweight* bivariada. Tomamos, como comprimento para a janela, $h = 0.15$. O número considerado de termos envolvidos no somatório é $q = [\log 500] = 6$. Comecemos por apresentar o gráfico relativo à aproximação das funções de distribuições marginais das variáveis, figura 4.2.1, assim como o gráfico da verdadeira função, figura 4.2.2.

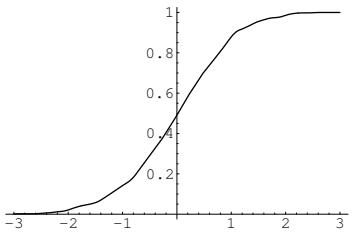


figura 4.2 .1

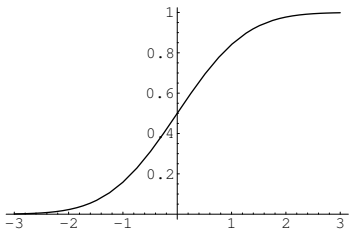


figura 4.2 .2

De seguida apresentamos o gráfico da função de distribuição aproximada

do par (X_1, X_2), figura 4.2.3, lado a lado com a verdadeira função, figura 4.2.4.

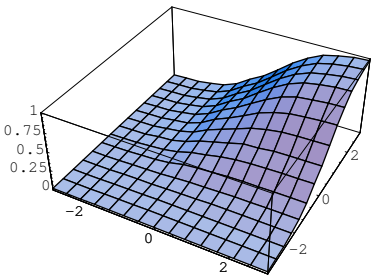


figura 4.2 .3

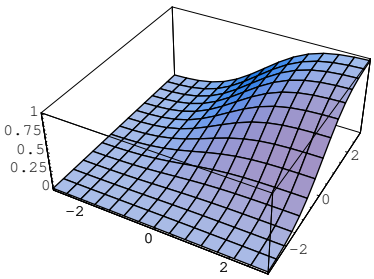


figura 4.2 .4

Vejamos, ainda, a aproximação e a respectiva função de distribuição do par (X_1, X_{50}) nas figuras 4.2.5 e 4.2.6, respectivamente.

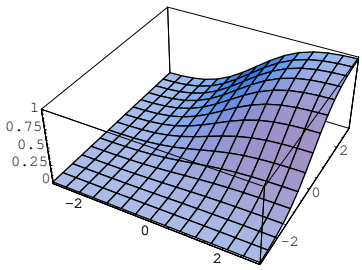


figura 4.2 .5

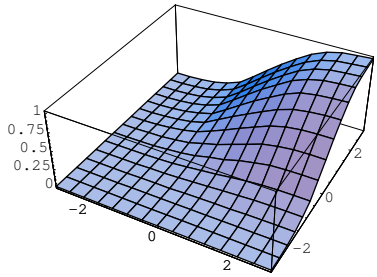


figura 4.2 .6

Por último temos a representação gráfica da função de covariância aproximada, assim como da verdadeira função de covariância, figuras 4.2.7 e 4.2.8, respectivamente.

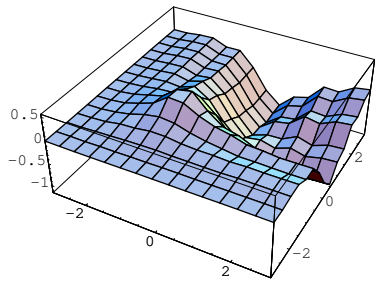


figura 4.2 .7

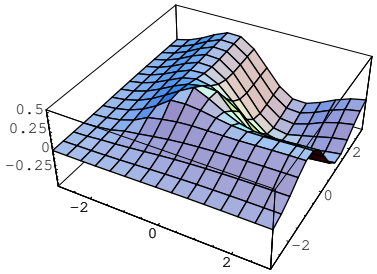


figura 4.2 .8

Vamos agora usar variáveis aleatórias independentes Gaussianas reduzidas para gerar variáveis também Gaussianas, mas com um grau de associação que permite uma expressão bastante simples para a função de covariância. De facto, as variáveis que obtemos e usamos são apenas 1− associadas.

Exemplo 4.2.3 *Consideremos então o seguinte procedimento.*

- *Simulamos n variáveis aleatórias, $Y_1, \ldots, Y_n$ com distribuição de Gauss reduzida;*

- *Consideramos, para cada $j > 1$, $X_j = \alpha Y_j + \beta Y_{j-1}$, com $\alpha\beta > 0$. As variáveis X_j são associadas e Gaussianas de média nula e variância $\alpha^2 + \beta^2$;*

- *Para $|i - j| = 1$, $\mathrm{Cov}(X_i, X_j) = \alpha\beta$, pelo que os pares aleatórios do tipo (X_1, X_{k+1}) têm distribuição Gaussiana bidimensional centrada, com matriz de covariância diagonal para $k > 2$. Seja F_k a função de distribuição dos pares (X_1, X_{k+1});*

- *A expressão para a função de covariância é*

$$\Gamma(s,t) = F(\min\{s,t\}) - F(s)F(t) + \big(F_1(s,t) + F_1(t,s) - 2F(s)F(t)\big),$$

bastante mais simples que a do exemplo anterior.

Vejamos agora qual o resultado das aproximações para uma amostra de dimensão $n = 500$, $\alpha = 0.1$ e $\beta = 0.9$ assim como para as respectivas funções obtidas através das expressões teóricas atrás identificadas. Usamos para as aproximações do operador as funções núcleo *Biweight* e *Epanechnikov* bivariada com $h = 0.11$ e $h = 0.3$. Nas aproximações para as funções de distribuição marginais usamos apenas o núcleo *Biweight*. Vejamos, em primeiro lugar, as representações gráficas da função de distribuição aproximada, figura 4.3.1 e da verdadeira função de distribuição, figura 4.3.2.

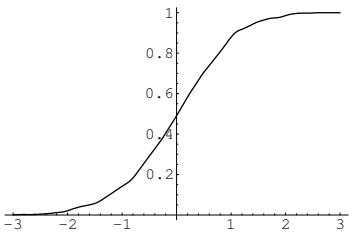


figura 4.3 .1
núcleo Biweight

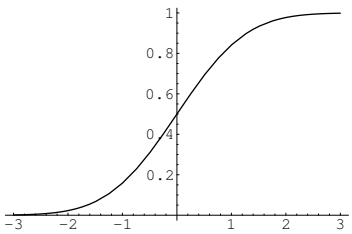


figura 4.3 .2

Para a aproximação da função de distribuição bidimensional do par (X_1, X_2) usamos os dois núcleos atrás referidos com o mesmo comprimento de janela. As figuras 4.3.3, 4.3.4, e 4.3.5 são relativas, respectivamente, à função aproximada com o núcleo Biweight, à verdadeira função e à função aproximada com o núcleo Epanechnikov.

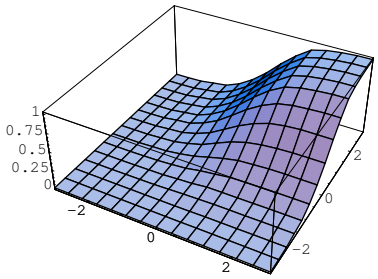


figura 4.3 .3
núcleo Biweight

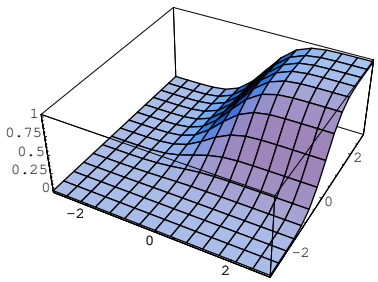


figura 4.3 .4

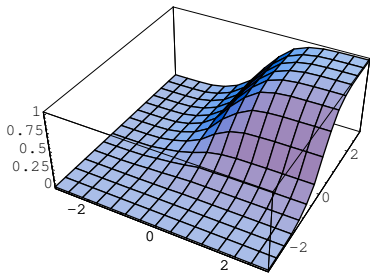


figura 4.3 .5
núcleo Epanechnikov

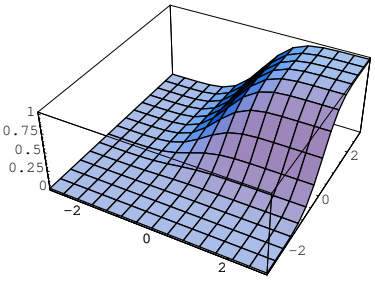


figura 4.3 .4

A análise destes gráficos revela-nos que não há diferenças visíveis entre as aproximações. Vejamos se tal se mantém para a função de covariância. Os gráficos relativos à aproximação para a função de covariância, usando o núcleo Biweight, figura 4.3.7 e Epanechnikov, figura 4.3.8, com o mesmo comprimento de janela, não apresentam entre si diferenças significativas.

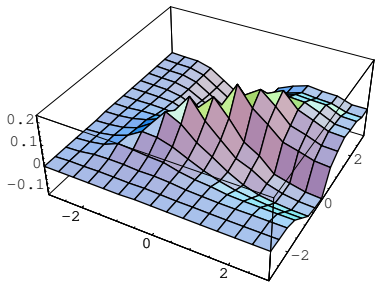


figura 4.3 .7
núcleo Biweight
h = 0.11

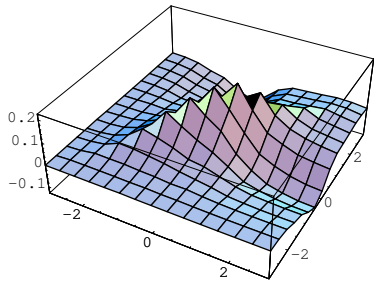


figura 4.3 .8
núcleo Epanechnikov
h = 0.11

Vejamos, agora, o gráfico da verdadeira função de covariância na figura 4.3.9.

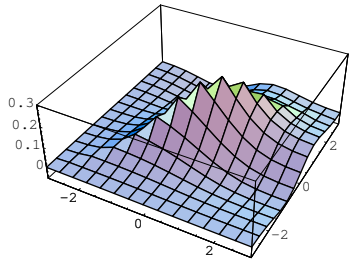


figura 4.3 .9
gráfico da função de covariância

Escolhamos, agora, outro comprimento para a janela. Façamos $h = 0.3$. Apresentamos, na figura 4.3.10, o gráfico da aproximação da função de covariância obtido com utilização do núcleo Biweight que, comparado com a mesma função para $h = 0.11$, revela diferenças bastante significativas.

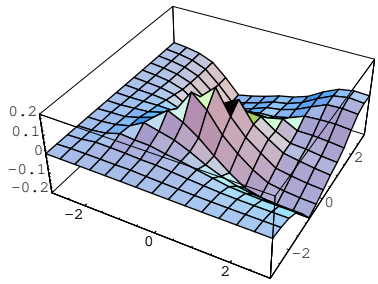


figura 4.3 .10
núcleo Biweight
h = 0.3

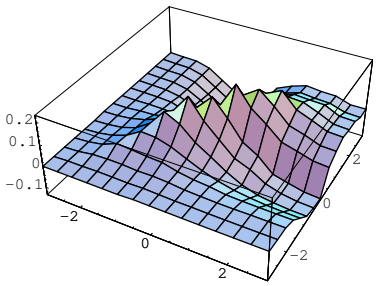


figura 4.3 .7
núcleo Biweight
h = 0.11

Aproximemos, ainda, o operador para $\alpha = 0.5$ e $\beta = 0.5$, com $h = 0.11$, primeiro usando o núcleo *Biweight*, figura 4.3.11 e, em seguida, o núcleo *Epanechnikov* na figura 4.3.12. Neste exemplo, as diferenças entre as aproximações com núcleos diferentes, embora ténues, são mais significativas que anteriormente.

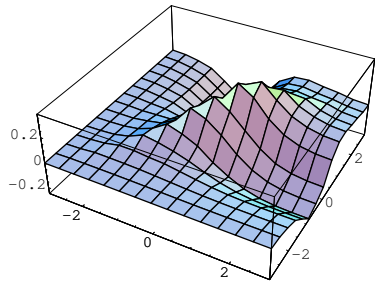


figura 4.3 .12
núcleo Epanechnikov

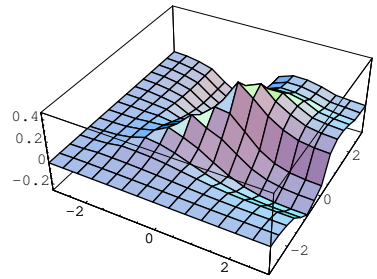


figura 4.3 .11
núcleo Biweight

O gráfico da função de covariância para $\alpha = 0.5$ e $\beta = 0.5$ é apresentado de seguida.

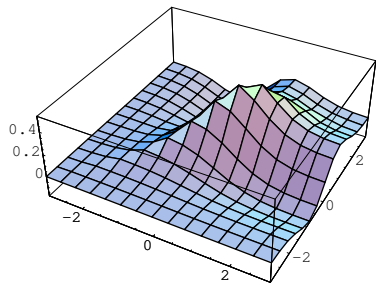


figura 4.3 .13

No exemplo que apresentamos de seguida vamos trabalhar com variáveis aleatórias com distribuição exponencial. Devido à simplicidade e às boas propriedades da função de distribuição exponencial, este exemplo permite-nos um melhor tratamento em termos de obtenção de resultados para o cálculo numérico do maior valor próprio, que discutiremos na secção seguinte.

Exemplo 4.2.4 *Vamos, primeiramente, obter sucessões associadas de variáveis aleatórias exponenciais de parâmetro unitário.*

- *Simulamos n variáveis aleatórias independentes $Y_1, \ldots, Y_n$, com distribuição exponencial de parâmetro um;*

- *Para um $m < n$, fixo, construímos a sucessão associada $\{X_n\}_{n\in\mathbb{N}}$, definida para cada $j = 1, \ldots, n-m$ por,*
  $X_j = (m+1)\min\{Y_j, Y_{j+1}, \ldots, Y_{j+m}\}$. *Estas variáveis aleatórias são identicamente distribuídas com distribuição exponencial de parâmetro um.*

Vejamos qual a função de distribuição do par (X_1, X_{k+1}) para cada $k = 1, \ldots, n-m-1$.
Para $k > m$ as variáveis X_1, X_{k+1} são independentes, pelo que, neste caso, $F_k(x,y) = \left(1 - \mathrm{e}^{-x}\right)\left(1 - \mathrm{e}^{-y}\right)$. Suponhamos agora $k = 1, \ldots, m$.

Neste caso,

$$F_k(x,y) = 1 - \mathrm{e}^{-x} - \mathrm{e}^{-y} + \mathrm{e}^{-\frac{k}{m+1}x}\mathrm{e}^{-\frac{k}{m+1}y}\mathrm{e}^{-\frac{m+1-k}{m+1}\max\{x,y\}}.$$

A função de covariância é, para todo $x, y \in \mathbb{R}_0^+$, definida por,

$$\begin{aligned}\Gamma(x,y) &= \left(1 - \mathrm{e}^{\min\{x,y\}}\right) - \left(1-\mathrm{e}^{-x}\right)\left(1-\mathrm{e}^{-y}\right) + \\ &+ \sum_{k=1}^{m}\left(F_k(x,y) - \left(1-\mathrm{e}^{-x}\right)\left(1-\mathrm{e}^{-y}\right)\right) + \\ &+ \sum_{k=1}^{m}\left(F_k(y,x) - \left(1-\mathrm{e}^{-x}\right)\left(1-\mathrm{e}^{-y}\right)\right)\end{aligned}$$

Vejamos os resultados das simulações obtidas para uma amostra inicial de dimensão $n = 500$, onde se considera $m = 50$, pelo que temos simuladas 450 variáveis associadas com distribuição exponencial de parâmetro unitário. Usamos, neste exemplo, o núcleo *Epanechnikov* bivariado, com um comprimento de janela $h = 0.1$ e o número de termos a somar na estimativa para a função de covariância é $q_n = q = 6$. Comecemos por apresentar, como habitualmente, o gráfico da função que aproxima as funções de distribuição marginais das variáveis aleatórias, figuras 4.4.1 e 4.4.2, respectivamente.

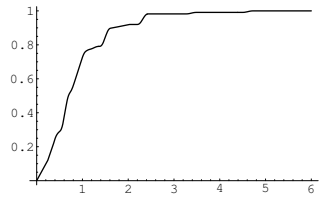


figura 4.4 .1

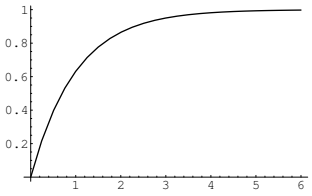


figura 4.4 .2

Vejamos, agora, a representação gráfica da aproximação para a função de distribuição bidimensional do par (X_1, X_2), figura 4.4.3 e a respectiva

função exacta na figura 4.4.4.

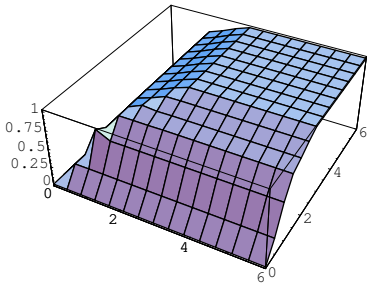


figura 4.4 .3

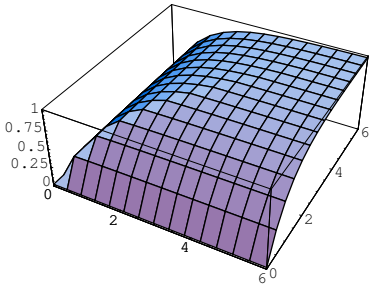


figura 4.4 .4

Consideremos, também as representações gráficas da função de distribuição de (X_1, X_{50}) primeiramente a aproximação, figura 4.4.5 e seguidamente a verdadeira função na figura 4.4.6.

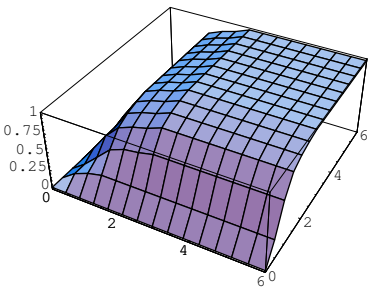


figura 4.4 .5

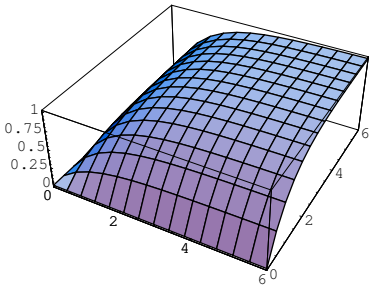


figura 4.4 .6

Por último, vejamos a representação gráfica da aproximação para o operador, figura 4.4.7 assim como a representação gráfica do verdadeiro operador na figura 4.4.8.

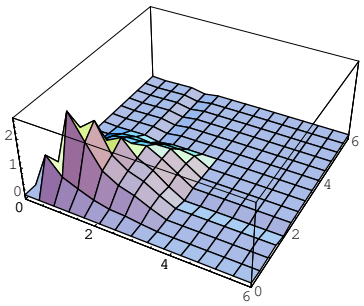


figura 4.4 .7

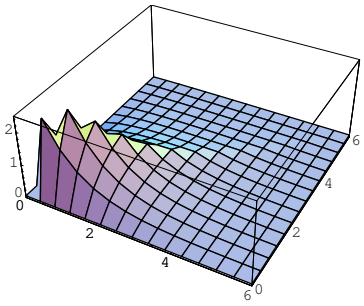


figura 4.4 .8

4.3 O maior valor próprio.

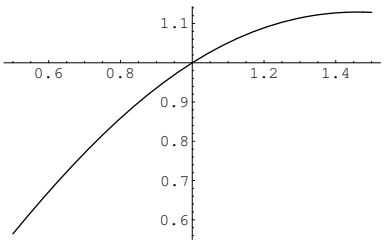
Nesta secção procuramos obter majorações para as probabilidades de cauda da Estatística de Cramèr-von Mises obtidas com base no maior valor próprio do estimador calculado através do programa numérico já estudado. Assim, e a título de ilustração, apresentamos um exemplo onde obtemos os valores críticos $\epsilon = \epsilon(n, \alpha)$ desta estatística para valores de n elevados e para diferentes valores de α, nível de significância do teste. Como vimos no terceiro capítulo, para grandes amostras,

$$P\left(W_n^2 > \epsilon\right) \sim 2^{\frac{3-m}{2}} \lambda_1^{\frac{1-m}{2}} \pi^{1/2} \frac{1}{\gamma(\frac{m}{2})} \epsilon^{\frac{1-m}{2}} \left(1 - \Phi\left(\frac{\epsilon^{1/2}}{\sqrt{\lambda_1}}\right)\right),$$

onde, conforme foi oportunamente referido, λ_1 é o maior valor próprio de Γ, com multiplicidade m, finita, γ é a função Gama de Euler, e Φ é a função de distribuição de Gauss reduzida.

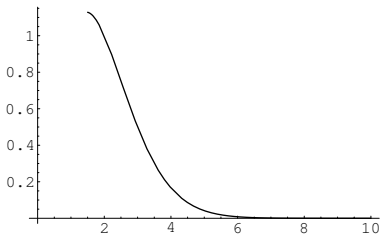
Ora, apesar de conhecermos um valor aproximado para λ_1, nada sabemos sobre a sua multiplicidade. Vejamos qual o comportamento da probabilidade anterior em função de m. Para $1 \leq m \leq 3$, temos o termo que envolve a função Gama limitado por valores próximos de um, a saber,

$0.56 \leq \frac{1}{\gamma(m/2)} \leq 1.128$, como facilmente podemos verificar por análise de $1/\gamma(x)$, para x entre 0.5 e 1.5 :



Por outro lado e para $m = 2, 3, \quad 2^{\frac{3-m}{2}} \lambda_1^{\frac{1-m}{2}} \epsilon^{\frac{1-m}{2}} \leq \frac{2}{\lambda_1 \epsilon}$. Temos então, que para este valor de m, aquela probabilidade próxima de $\frac{2}{\lambda_1 \epsilon} \left(1 - \Phi\left(\frac{\epsilon^{1/2}}{\sqrt{\lambda_1}}\right)\right)$.

Vejamos, agora, qual o comportamento da probabilidade em causa para $m > 3$. Neste caso, o termo que envolve a função Gama de Euler é muito pequeno como se pode verificar através do gráfico de $1/\gamma(x)$, para $x > 3$,



Além disso, para quaisquer valores de λ_1 e de ϵ, $\quad 2^{\frac{3-m}{2}} \lambda_1^{\frac{1-m}{2}} \epsilon^{\frac{1-m}{2}} \leq 1$. Assim, para $m > 3$, a probabilidade tem um comportamento assimptótico majorado por

$$\left(1 - \Phi\left(\frac{\epsilon^{1/2}}{\sqrt{\lambda_1}}\right)\right).$$

Consideremos $m = 1$. Neste caso,

$$P\left(W_n^2 > \epsilon\right) \sim 2\left(1 - \Phi\left(\frac{\epsilon^{1/2}}{\sqrt{\lambda_1}}\right)\right),$$

uma vez que $\pi^{1/2}\frac{1}{\gamma(\frac{1}{2})} = 1$. Tomemos, por exemplo o maior valor próprio calculado numericamente para o operador do último exemplo apresentado na secção anterior, $\lambda_1 = 0.9$. Para este valor próprio e amostras grandes, os valores críticos ϵ da estatística de Cramèr-von Mises são os seguintes para os diversos valores α, nível de significância do teste:

α	ϵ
0.01	5.97141
0.05	3.45731
0.1	2.43491

4.3.1 Conclusões

O programa de cálculo do maior valor próprio de operadores de Hilbert-Schmidt, em particular dos operadores de covariância que aproximamos, apesar de não estar optimizado (consome demasiada memória e tempo) produz resultados satisfatórios na medida em que converge ao fim de, aproximadamente, dez iterações com um erro que não excede 0.001. Como parece natural, quanto mais próxima estiver a aproximação para o operador do operador real, mais próximos estão os respectivos valores próprios. Assim, uma boa aproximação para este operador é crucial. Ora, de todas as experiências que temos feito na questão das simulações e cálculo de valores próprios, existe uma que pensamos merecedora de reflexão particular. Tal diz respeito à escolha da sucessão de comprimentos $\{h_n\}_{n\in\mathbb{N}}$. De facto, esta escolha influencia bastante o resultado das aproximações. Note-se que, se os comprimentos das janelas forem muito pequenos, próximos de zero, temos uma soma de funções delta de Dirac com valores nas observações. Por outro lado, se esses comprimentos forem demasiado grandes perdemos todos os detalhes, pelo que a informação contida na amostra dilui-se. Em contrapartida a escolha da função núcleo não influencia grandemente a qualidade das nossas aproximações, reproduzindo e acentuando até o comportamento já bem

conhecido destes estimadores na estimação da densidade. O caso em que se nota mais a influencia desta função núcleo na aproximação das funções foi apresentado no exemplo 4.2.3 na secção anterior. No entanto, pensamos que a obtenção de núcleos óptimos, no sentido da minimização do Erro Quadrático Médio do Estimador é um trabalho a desenvolver, assim como, pela sua importância, o desenvolvimento de métodos eficientes para a escolha dos comprimentos das janelas. Relativamente ao programa desenvolvido para o cálculo do maior valor próprio dos estimadores, pensamos que a sua optimização é importante, na medida em que, como vimos, a sua eficiência aliada a uma boa aproximação da função de covariância, abre caminho para a caracterização do nível de significância em testes de hipóteses cujas probabilidades de cauda das estatísticas de decisão dependem essencialmente dos valores próprios da função de covariância do limite do processo empírico.

Capítulo 5

Apêndice

5.1 Programas para as aproximações das funções

Nesta secção apresentamos os programas que elaborámos no *Mathematica*® para obter os resultados gráficos das aproximações das funções e das funções apresentados no último capítulo. Para cada um dos exemplos, referimos e apresentamos, em primeiro lugar, o procedimento comum e, em seguida, o procedimento específico de cada um deles.
Nos exemplos 4.2.1, 4.2.2, 4.2.3 e 4.2.4 foram usados pacotes já existentes no *Mathematica*® para gerar as sucessões de variáveis pseudo-aleatórias com a distribuição pretendida. Nos programas em que se obtêm as aproximações são testadas em primeiro lugar as condições exigidas teoricamente para a função núcleo. Assim todos os programas têm o mesmo início que passamos a descrever.

```
<< Statistics'ContinuousDistributions'
Clear[n, m, fdp, h_n, CondicoesFDP, funcaoTchapeu];
Clear[CondicoesFDP, fdp];
CondicoesFDP[fdp_]:= (Integrate[fdp,{x,-1,1},{y,-1,1}] == 1   ∧
Integrate[x fdp,{x,-1,1},{y,-1,1}] == 0   ∧
Integrate[y fdp,{x,-1,1},{y,-1,1}] == 0   ∧
```

Integrate[x^2 fdp,$\{x,-1,1\},\{y,-1,1\}] < \infty \quad \wedge$
Integrate[y^2 fdp,$\{x,-1,1\},\{y,-1,1\}] < \infty$);
Clear[VerificaCondicoes, funU, fdp];
VerificaCondicoes[funU_,fdp_]:=
Integrate$\big[x^2$(funU fdp + D[funU,x]D[funU,y]),
$\{x,-1,1\},\{y,-1,1\}\big] < \infty \wedge$
Integrate$\big[y^2$(funU fdp + D[funU,x]D[funU,y])
$\{x,-1,1\},\{y,-1,1\}] < \infty$);

Clear[funcaoTchapeu, n, m, fdp, hn, funM, funU];
funcaoTchapeu[$n_, m_$, fdp_,hn_]:= Module[$\{varx_i, vary_j, varz_i$,
Fchapeu, Gchapeu, Hchapeu, Tchapeu, TchapeuAux, carLog$n, k\}$,

Neste ponto cada programa é específico relativamente à simulação das variáveis já que pretendemos uma determinada distribuição. Assim, e para o exemplo 4.2.1, temos

$(varz_j = Table[Random[NormalDistribution[0,1]],\{n\}]$;
vary_i = Table [$\frac{1}{\sqrt{m+1}}$(Sum[varz_j[[j]], $\{j, i, i+m\}$]),$\{i, 1, n-m\}$];
varx_i = Table[CDF[NormalDistribution[0,1],vary_i]];
Print[varx_i];

O que se segue é, mais uma vez, comum em todos os programas. Diz respeito à definição das funções núcleo bivariadas e univariadas que intervêm na definição do nosso estimador e à construção, idêntica para todos os casos, das aproximações.

Clear[a,b,funUaux,funU];
funUaux[$a_, b_$] = Integrate[fdp, $\{x,-1,a\},\{y,-1,b\}$];
funU[$a_, b_$]:= funUaux[a, b] = /; $(-1 \leq a \leq 1 \wedge -1 < b \leq 1)$;
funU[$a_, b_$]:= 0 /; $(a < -1)||(-1 < a < 1 \wedge b > 1)||(-1 < a < 1 \wedge b < -1)||(a > 1 \wedge b < 1)$;
funU[$a_, b_$]:= 1 /; $(a > 1 \wedge b > 1)$;
Print["U(a, b) = ", funUaux[a, b]];

```
Clear[a,b,funMaux,funM];
funMaux[a_] = Integrate[funUaux[a, b], {b, −1, 1}];
funM[a_]:= funMaux[a]/; −1 ≤ a ≤ 1;
funM[a_]:= 0 /; a < −1;
funM[a_]:= 1 /; a > 1;
Print["M[a] = ", funMaux[a]];
Clear[s,t];
If[VerificaCondicoes[funUaux[x,y], fdp], (
Fchapeu = Table [((1/(n − m − k)) ∗
Sum[funU[(s − varx_i[[k + j]])/h_n, (t − varx_i[[j]])/h_n, {j, 1, n − m − k}], {k, 1, n−
m − 1}];
Gchapeu = Table [((1/(n − m − k)) ∗
Sum[funU[(t − varx_i[[k + j]])/h_n, (s − varx_i[[j]])/h_n, {j, 1, n − m − k}], {k, 1, n−
m − 1}];
Hchapeu[s_] = ((1/(n − m)) ∗ Sum[funM[(s − varx_i[[k + j]])/h_n], {j, 1, n −
m − k}];

Plot3D[Fchapeu[[1]], {s,0,1}, {t,0,1}, PlotRange→ All];
Plot3D[Fchapeu[[m]], {s,0,1}, {t,0,1}, PlotRange→ All];
Plot[Hchapeu[s], {s,0,1}, PlotRange→ All]; ) ];

Tchapeu = H[Min[{s, t}]-(2carLognn +1)∗ H[s] ∗ H[t] +

Sum[Fchapeu[[k]], {k, 1, carLogn}]+Sum[Gchapeu[[k]], {k, 1, carLogn}];
Plot3D[Tchapeu,{s, 0, 1}, {t, 0, 1}, PlotRange→ All];
Return [Tchapeu]; ) ]
```

Quanto ao programa para a representação gráfica das verdadeiras funções, ele é estruturalmente o mesmo para cada exemplo tendo sido, no caso do exemplo 4.2.1, introduzido o pacote da distribuição normal multivariada.

```
<<Statistics 'MultinormalDistribution'
Clear[operador, funcaoG, n, m, G, InvG];
operador[n_, m_]:=Module[{Gchapeu, Gamma, carLogn, h, f},(Clear[G, x];
```

$f =$ `NormalDistribution[0, 1];`

$G[x_] := CDF[f, x];$

`Clear[funcaoG, x, y, k];`

$\left(r := \left\{\left\{1, \frac{m-(k+1)}{m+1}\right\}, \left\{\frac{m-(k+1)}{m+1}, 1\right\}\right\};\right.$

`ndist=MultinormalDistribution` $[\{0,0\}, r];)$

`funcaoG[k_, x_, y_] = CDF[ndist,` $\{x, y\}$`];`

`Print[Passei1];`

`Clear[InvG, s, t];`

$h = x/.($`Solve`$[G[x] == s, x])[[1,1]];$`Clear`$[s];$`InvG`$[z_] = h/.s \to z;$

`Plot3D[funcaoG[2,InvG[`x`], InvG[`y`]],` $\{x,0,1\},\{y,0,1\}$`,PlotRange→All];`

`Plot3D[funcaoG[`m`,InvG[`x`], InvG[`y`]],` $\{x,0,1\},\{y,0,1\}$`,PlotRange→All];`

`carLogn = IntegerPart[N[Log[`$n-m$`]]];`

`Clear[Gama, s, t];`

`Gama`$[s_, t_] := \text{Min}\,[\{s,t\}] - (2m+1)s*t+$

$\sum_{k=2}^{m+1}$(`funcaoG`$[k,$`InvG`$[s],$`InvG`$[t]] +$ `funcaoG`$[k,$`InvG`$[t],$`InvG`$[s]];$

`Plot3D[Gama`$[s,t]\{x,0,1\},\{y,0,1\},$`PlotRange` $\to$ `All];)]`

Vejamos, agora, como se geram, com o *Mathematica*®, as variáveis pseudo-aleatórias especificadas no exemplo 4.2.2.

vec = Table[0, {n}];

mat = Table[(ex = Join[Table[k, {k, i − 1, 1, −1}], Table[k, {k, 0, n − i}]];

$$\texttt{Table}\left[\left[\frac{\frac{1}{2}^{\texttt{ex}[[j]]}}{\frac{3}{4}}, \{j, 1, n\}\right]\right), \{i, 1, n\}\Bigg];$$

varx_i = Table[Random[MultinormalDistribution[vec, mat]]];

Print[varx_i];

Quanto às instruções necessárias para o gráfico das verdadeiras funções de distribuição marginais e bivariadas pretendidas, assim como da função de covariância, não há nada de especial a registar uma vez que se tratam de funções que envolvem directamente a distribuição de Gauss.

f = NormalDistribution[0, 1];

G[$x_$] :=CDF[f, x];

$$\left(r := \{\{\frac{1}{1-0.5^2}, \frac{0.5^k}{1-0.5^2}\}, \{\frac{0.5^k}{1-0.5^2}, \frac{1}{1-0.5^2}\}\};\right.$$

ndist = MultinormalDistribution[{0, 0}, r];)

Gama[$s_, t_$]:= G[Min[{s, t}]] − $(2(n-1)+1)G[s] * G[t]+$

$$+\sum_{k=2}^{n}(\texttt{funcaoG}[k, s, t] + \texttt{funcaoG}[k, t, s]);$$

Gerar variáveis pseudo-aleatórias para o exemplo 4.2.3 é bastante simples, usando apenas o pacote `Statistics'NormalDistribution'`.

vary_j= Table[Random[NormalDistribution[0, 1]], {n}];

varx_i = Table[$0.5 *$ var$y_j[[j+1]] + 0.5 *$ var$y_j[[j]], \{j, 1, n-1\}$]

Print[varx_i];

No entanto, para conseguirmos os gráficos das verdadeiras funções, necessitamos ainda do pacote `Statistics'MultinormalDistribution'`, para as distribuições bivariadas e para a função de covariância. De resto, o

procedimento é análogo ao referido para o exemplo 1.4.2.

Vejamos o procedimento usado para gerar as variáveis pseudo-aleatórias do exemplo 4.2.4. O pacote utilizado é o `Statistics'ContinuousDistributions'`.

`var`y_i `= Table[Random[ExponentialDistribution[1]],` $\{n\}$`];`

`var`x_i `= Table[`$(m+1)$`Min[Flatten[Table[var`y_i`[[`k`]],` $\{k, i, i+m\}$`]]],` $\{i, 1, n-m\}$`];`

`Print[var`x_i`];`

Apresentamos, por último, a parte do programa referente à representação gráfica das verdadeiras funções, mais especificamente às funções de distribuição bivariadas.

`Clear[funcaoGaux1, funcaoGaux2, funcaoG,` x, y, k`];`

`funcaoGaux1[`$k_, x_, y_$`]:=` $\text{Exp}[-x] * \text{Exp}[-y]$;

`funcaoGaux2[`$k_, x_, y_$`]:=` $\text{Exp}[-\frac{x*k}{m+1}]*\text{Exp}[-\frac{y*k}{m+1}]*\text{Exp}\left[-\frac{\text{Max}[\{x,y\}](m+1-k)}{m+1}\right]$;

`funcaoG[`$k_, x_, y_$`]:=` $1-\text{Exp}[-x]-\text{Exp}[-y]+$`funcaoGaux1`$[k, x, y]/.(m < k)$;

`funcaoG[`$k_, x_, y_$`]:=` $1-\text{Exp}[-x]-\text{Exp}[-y]+$`funcaoGaux2`$[k, x, y]/.(m \geq k)$;

Todos estes programas terminam com o retorno das funções de covariância, pois estas são introduzidas, como veremos na secção seguinte, no cálculo do maior valor próprio da respectiva função de covariância.

5.2 Cálculo do maior valor próprio

Nesta secção descrevemos dois programas perfeitamente análogos para o cálculo do maior valor próprio, sendo o primeiro um programa que usa o cálculo simbólico do *Mathematica*® e o segundo apenas o cálculo numérico. O primeiro programa, o simbólico, apenas pode ser utilizado se

as funções com que trabalhamos têm primitivas que o *Mathematica*® consegue calcular. Ora, na maior parte dos casos isto não é possível. Nos exemplos que analisámos apenas a verdadeira função de covariância do exemplo 4.2.4 pode ser utilizada neste programa. Houve, então, necessidade de incorporar apenas cálculos numéricos no referido programa, o que deu origem ao segundo programa, o numérico. Vamos, então, apresentar o programa simbólico e, seguidamente, referir os comandos que são alterados a fim de se obter o programa numérico.

```
Clear[valorProprio, tol, maxiter, operador, ginicial];
valorProprio[operador_, gInicial_, tol_, maxiter_]:=
Module[{g =gInicial, i = 1, normag, h, miu, erro= 10},
normag=NIntegrate[g*g, {t, 0, 1}];
g = g/√normag;
While [√erro > tol    ∧   i ≤ maxiter,
h = Simplify [∫_0^t ((s + operador) * (g/.{t → s})) ds+
∫_t^1 ((t + operador) * (g/.{t → s})) ds];
miu = NIntegrate[h*h, {t, 0, 1}]; vp = √miu;
If[vp == 0, Print["Escolha outra função g inicial"]; Break[]];
erro = NIntegrate [(g − h/vp) * (g − h/vp), {t, 0, 1}];
Print["erro = ", √Abs[erro]]; Print["Valor próprio = ", vp];
g = Simplify[h/vp];
If[√Abs[erro] < tol, Print["função própria = ", g, "valor próprio
= "vp, "número de iterações:  ", i]; Break[]];
i = i + 1;];
If[i >maxiter, Print["O número máximo de iterações foi excedido.
O processo está completo sem sucesso."]]];
```

Vejamos agora o programa numérico, nomeadamente no que diz respeito ao cálculo de `h` que, em cada iteração, dividido pela norma de `g`, função própria, é a aproximação para o maior valor próprio da função em causa.

Usamos a função de covariância do exemplo 4.2.1.

```
Clear[valorProprio, tol, maxiter, operador, ginicial];
valorProprio[operador_, gInicial_, tol_, maxiter_]:=
Clear[g,h,k];
h[k_]:= NIntegrate[g[k-1]*var[k]-7*g[k-1]*var[k]*var[k+1]+
g[k-1]*(operador[var[k],var[k+1]]),{var[k],0,var[k+1]}] +
NIntegrate[g[k-1]*var[k+1]-7*g[k-1]*var[k]*var[k+1]+g[k-
1]*(operador[var[k],var[k+1]]),{var[k],0,var[k+1]}];
g[0]=gInicial /.  t -> var[1];
g[k_]:=h[k]/lvp[[k]]; lvp = {};
i = 1; Clear[var];
While [Sqrt[erro] > tol   &&  i <= maxiter,
lvp = Append[lvp, Sqrt[NIntegrate[h[i]*h[i],{var[i+1],0,1}]]];
Print[lvp[[i]]];
erro = NIntegrate[(g[i-1]-h[i]/lvp[[i]])^2,{var[i+1],0,1}];]
```

A ordem de paragem é exactamente a mesma que no programa simbólico. Resta-nos referir, uma vez mais, que estes programas, simbólico e numérico, não estão, de maneira nenhuma, optimizados uma vez que consomem demasiada memória e a sua execução é bastante morosa.

Bibliografia

[1] Anderson, T.W. and Darling, D.A; Asymptotic theory of certain "goodness of fit"criteria based on stochastic processes, *Ann. Math. Statis.*, 23:193-212, 1952.

[2] Azevedo, C. and Oliveira, P.E., Kernel-type estimation of bivariate distribution function for associated random variables, *New Trends in Probability and Statistics, Proceedings of the* 6^{th} *Tartu Conference, VSP,* 5:17-25, 2000.

[3] Azevedo, C. and Oliveira, P.E., Covariance Estimation of the limit process of the empirical process under association, *Recent Advances in Applied and Theoretical Mathematics, ed. N.Mastorakis, Actas do WSES International Conference on Applied and Theoretical Mathematics, Vravrona, Atenas, Grécia*, 1-3/12/00, 78–83, 2000.

[4] Billingsley, P.; Convergence of Probability Measures, Willey, New York, 1968.

[5] Birkel, T.; A note on the strong law of large numbers for positively dependent random variables, *Statist. Probab. Letters*, 7:17-20, 1988.

[6] Birkel, T.; A functional central limit theorem for positively dependent random variables, *Multiv. Anal.*, 44:314-320, 1993.

[7] Cai, Z. and Roussas, G.; Berry-Esseen Bounds for smooth estimator of a distribution function under association, *J. Nonparametr. Statist.*, 11:79-106, 1996.

[8] Cai, Z. and Roussas, G.; Efficient Estimation of a Distribution Function under Quadrant Dependence, *Scand. J. Statist.*, 25:211-224, 1998.

[9] Cai, Z. and Roussas, G.; Weak convergence for smooth estimator of a distribution function under negative association, *Stochastic Anal. Appl.*, 17:145-168, 1999.

[10] Devroye, L.; A Course in density estimation, *Birkhauser*, 1987.

[11] Donsker, M.D.; An invariance principle for certain probability limit theorems, *Mem. Amer. Math. Soc.*, 6, 1951.

[12] Dunford, N. and Schwartz, J.; Linear Operators, *Interscience Publishers Inc., New York*, 1957.

[13] Esary, J.D., Proschan, F. and Walkup, D.W; Association of random variables, with applications, *Ann. Math. Statist*, 38:1466-1474, 1967.

[14] Fourtuin, C., Kasteleyn, P. and Ginibre, J.; Correlation inequalities on some partially ordered sets, *Comm. Math. Phys.*, 22:89-103, 1971.

[15] Franche, A. Estimation des valeurs propres de la covariance du processus empririque sous dépendence faible. Application à la statistique de test de Cramér-von Mises, *preprint*, Univ. Scienc. Techn. Lille, 2000.

[16] Gijbels, I.; Density estimation and applications, *On Nonparametric and Semiparametric statistics*, Mendes Lopes, M., Gonçalves, E. eds, CIM, 1998.

[17] Henriques, C. and Oliveira, P.E.; Covariance estimation for associated random variables, *Journal of Stat. Plann. and Inf.*, em publicação.

[18] Ioannides, D.A. and Roussas, G.; Exponential inequality for associated random variables, *Statist. Probab. Lett.*, 42:423-431, 1999.

[19] Jacob, P. and Oliveira, P.E.; Histograms and associated point processes, *Statist. Inf. Stoch. Proc.*, 2:227-321, 1999.

[20] Jacquemin, M. P.; Theoremes limites pour les suites de variables aleatoires associee, *Memoiré DEA*, Universitè Montpellier *II*, 1997.

[21] James, B.; Probabilidade: um curso em nível intermediário, *IMPA*, 1996.

[22] Joag-Dev and Proschan, F.; Negative association of random variables, with applications, *Ann. Statist.* 11:286-295, 1983.

[23] Jin, Z. and Shao, Y.; On kernel estimation of a multivariate distribution function, *Statist. Probab. Lett.*, 41:163-168, 1999.

[24] Lebowitz, J.; Bounds on the correlations and analyticity properties of ferromagnetic Ising spin systems, *Comm. Math. Phys.*, 28:313-321, 1972.

[25] Lehmann, E.L.; Some concepts of dependence, *Ann. Math. Statist.*, 37:1137-1153, 1966.

[26] Loève, M.; Probability theory, *Springer-Verlag*, 1977.

[27] Louhichi, S.; Weak convergence for empirical process of associated sequences, *Ann. Inst. Henri Poincaré, Probabilités et Statistiques*, 36:547-567, 2000.

[28] Louhichi, S.; Convergence rates in the strong law for associated random variables, *PhD Thesis*, Université de Paris-Sud, 2000.

[29] Morel, B.; Weak convergence in Besov spaces with an application to invariance principle for summation processes, *Pub IRMA Lille*, 49-IX, 1999.

[30] Newman, C.M.; Normal fluctuations and the FKG inequalities, *Comm. Math. Phys.*, 75:119-128, 1980.

[31] Newman, C.M.; Asymptotic independence and limit theorems for positively and negatively dependent random variables, *Inequalities in Statistics and Probability*, IMS Lecture Notes-Monograph Series 5:127-140, 1984.

[32] Newman, C.M. and Wright, A.L.; An invariance principle for certain dependent sequences, *Ann. Probab.*, 9:671-675, 1981.

[33] Oliveira, P.E.; A central limit theorem for associated variables, *Actas do VII Congresso Anual da SPE*, 500-516, 2000.

[34] Oliveira, P.E. and Suquet, Ch.; An $L^2[0,1]$ invariance principle for LPQD random variables, *Port. Mathematica*, 53:367-379, 1996.

[35] Oliveira, P.E. and Suquet, Ch.; $L^2[0,1]$ weak convergence of the empirical process for dependent variables, *A. Antoniadis and G. Oppenheim, Eds, Actes des XV Rencontres Franco-Belges de Statisticiens (Ondelettes et Statistic), Lecture Notes in Statistics 103, Wavelets an Statistics*, 331-344, 1995.

[36] Oliveira, P.E. and Suquet, Ch.; Weak convergence in $L^p[0,1]$ of the uniform empirical process under dependence, *Statist. Probab. Lett.*, 39:363-370, 1998.

[37] Plackett, R.L.; A redution formula for normal multivariate integrals, *Biometrika*, 41:351-360, 1954.

[38] Pitt, L.; Positively correlated normal variables are associated, *Ann. Probab.*, 10:496-499, 1982.

[39] Roussas, G.; Smooth estimation of the distribution function under association: Asymptotic Normality, *Tech. Report* #299, *Division of Statistics, University of California, Davis*, 1993.

[40] Roussas, G.; Asymptotic normality of a smooth estimate of a random field distribution function under association, *Statist. Probab. Lett.*, 24:77-90, 1995.

[41] Roussas, G.; Asymptotic normality of the kernel estimate of a probability density function under association, *Statist. Probab. Lett.*, 50:1-12, 2000.

[42] Sadikova, S.M.; Two dimensional analogues of an inequality of Essén with applications to the central limit theorem, *Theory Probab. Appl.*, 11:325-335, 1966.

[43] Shao, Q.M. and Yu, H.; Weak convergence for weighted empirical process of dependent sequences, *Ann. Probab.*, 24:2098-2127, 1996.

[44] Shorack, G. and Wellner, J.; Empirical processes with applications to statistics, *Willey series in probability and mathematical statistics*, 1986.

[45] Scott, D.; Multivariate density estimation, *John Willey and Sons, Inc.*, 1992.

[46] Soares, M.J.; Introdução à teoria das onduletas, *preprint, Departamento de Matemática, Universidade do Minho*, 1997.

[47] Suquet, Ch.; Convergences stochastiques de suites de mesures aléatoires à signe considérées comme variables aléatoires hilbertiennes, *Pub. Inst. Stat. Univ. Paris* $XXXVII$, 1-2:71-99, 1993.

[48] Suquet, Ch.; Introduction à l'association, *Pub. Irma. Lille*, 34:$XIII$, 1994.

[49] Tricomi, F.G.; Integral equations, *Dover Publications, Inc., New York*, 1957.

[50] Utev, S.A.; On the central limit theorem for $\phi-$mixing arrays of random variables, *Th. Probab. Appl.*, 35:131-139, 1990;

[51] Vakhania, N.N., Tarieladze V.I. and Chobanyan, S.A.; Probability Distributions on Banach Spaces, *Mathematics and its applications*, D. Reidel Publishing Company. Sovietic Series, 1987.

[52] Valença, M.R., Notas sobre o *Mathematica*, *preprint, Departamento de Matemática, Universidade do Minho*, 1998.

[53] Yu, H.; A Glivenko-Cantelli lemma and weak convergence for empirical process of associated sequences, *Probab. Theory and Related Fields*, 95:357-370, 1993.

[54] Zimmer, R.J.; Essential results of functional analysis, *Chicago Lectures in Mathematics*, 1990.

[55] Zolotarev, V.M; Concerning a certain probability theorem, *Theor. Probab. Appl.*, 6:201-204, 1961