

# MAE 5871: Análise Espectral de Séries Temporais

Pedro A. Morettin

Instituto de Matemática e Estatística  
Universidade de São Paulo  
pam@ime.usp.br  
<http://www.ime.usp.br/~pam>

## Aula 5

1 de setembro de 2021

# Sumário

- 1 Funções generalizadas
- 2 Núcleos e janelas
- 3 Convergência de variáveis aleatórias

# Delta de Dirac

1. Para que  $f(t)$  tenha uma transformada de Fourier, ela deve satisfazer certas condições, por exemplo, ser de quadrado integrável ou absolutamente integrável. Usando a teoria das funções generalizadas é possível estender a teoria para funções que não satisfazem essas condições. Tal teoria está fora do alcance deste curso, de modo que vamos nos limitar a apresentar brevemente, e de modo não rigoroso, a função delta de Dirac.
2. A **função delta** de Dirac (ou **função impulso**) é definida pelas relações

$$\delta(t) = \begin{cases} +\infty, & \text{se } t = 0 \\ 0, & \text{se } t \neq 0, \end{cases} \quad (1)$$

$$\int_{-\infty}^{\infty} \delta(t) dt = 1. \quad (2)$$

3. A rigor,  $\delta(t)$  não é uma função, pois assume o valor  $+\infty$  para  $t = 0$  e a integral (2) deveria ser zero.

## Delta de Dirac

1. Usualmente,  $\delta(t)$  é introduzida pela propriedade

$$\int_{-\infty}^{\infty} \delta(t)f(t)dt = f(0), \quad (3)$$

se  $f(t)$  for uma função real, contínua e que se anula fora de um intervalo limitado. Contudo, como  $\delta(t)$  não é uma função, a integral do primeiro membro de (3) não faz sentido.

2. Para contornar esse problema, vamos considerar uma sequência de funções contínuas  $\delta_n(t)$ , tais que:

(i)  $\delta_n(t) \geq 0$ , para todo  $t$ ;

(ii)  $\int_{-\infty}^{\infty} \delta_n(t)dt = 1$ ;

(iii) para todo  $\epsilon > 0, \eta > 0$ , existe  $n_0$  tal que, para  $n \geq n_0$ ,

$$\int_{|t|>\eta} \delta_n(t)dt < \epsilon.$$

3. Dessa maneira, (3) poderia ser definida por

$$\int_{-\infty}^{\infty} \delta(t)f(t)dt = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \delta_n(t)f(t)dt. \quad (4)$$

# Delta de Dirac

1. Uma sequência de funções  $\delta_n(t)$  como a definida acima é chamada **sequência de núcleos de Dirac**. Pode-se demonstrar, usando (4), que a relação (3) é realmente correta, se impusermos a condição adicional que  $\delta_n(t)$  seja par, para todo  $n$ .
2. A interpretação física da função delta é a de um impulso de energia num sistema. Suponha que esse sistema seja definido pela integral

$$y(t) = \int_{-\infty}^{\infty} h(u)x(t-u)du. \quad (5)$$

3. Aqui,  $x(t)$  é a **entrada** e  $y(t)$  é a **saída** do sistema, o qual é caracterizado pela função  $h(t)$ . Tal sistema diz-se um **filtro linear**. Se a entrada for  $x(t) = \delta(t)$ , então

$$y(t) = \int_{-\infty}^{\infty} h(u)\delta(t-u)du = h(t),$$

ou seja, a resposta do sistema a um impulso é a função  $h(t)$ , donde o nome **função resposta de impulso**.

## Heaviside e pente de Dirac

1. Duas outras funções úteis são a **função de Heaviside** e o **pente de Dirac**, que denotaremos por  $H(t)$  e  $\eta(t)$ , respectivamente. Essas são definidas por:

2.

$$H(t) = \begin{cases} 0, & \text{se } t < 0 \\ 1/2, & \text{se } t = 0 \\ 1, & \text{se } t > 0. \end{cases} \quad (6)$$

3. e

$$\eta(t) = \sum_{j=-\infty}^{\infty} \delta(t - 2\pi j). \quad (7)$$

4. A função  $H(t)$  pode ser encarada como o limite de funções  $H_n(t)$  (para  $n \rightarrow \infty$ ) que têm limite zero, para  $t < 0$ , e um, para  $t > 0$ . Por exemplo, considere

$$H_n(t) = \begin{cases} e^{-nt}/2, & \text{se } t < 0 \\ 1 - e^{-nt}/2, & \text{se } t \geq 0. \end{cases}$$

5. A função  $\eta(t)$  tem a propriedade

$$\int_{-\infty}^{\infty} f(t)\eta(t)dt = \sum_{j=-\infty}^{\infty} f(2\pi j). \quad (8)$$

# Núcleos

## 1. Considere

$$f_N(t) = \sum_{n=-N}^N c_n e^{i\lambda_n t}, \quad (9)$$

onde  $\lambda_n$  são as frequências de Fourier. Substituindo o valor de  $c_n$ , obtemos

$$f_N(t) = \int_{-T/2}^{T/2} D_N\left[\frac{2\pi}{T}(t-u)\right] f(u) du, \quad (10)$$

onde

$$D_N(t) = \frac{1}{T} \frac{\text{sen}[(N+1/2)t]}{\text{sen}(t/2)} \quad (11)$$

é o **núcleo (kernel) de Dirichlet**, que tem as propriedades:

2. (i)  $D_N(t)$  é uma função par;
- (ii)  $\int_{-T/2}^{T/2} D_N(t) dt = 1$ ;
- (iii)  $D_N\left(\frac{2\pi t}{T}\right)$  é uma função periódica de período  $T$ ;
- (iv)  $D_N(0) = \frac{2N+1}{T}$ .

## Dirichlet

1. A equação (10) mostra que  $f_N(t)$  é **soma** de valores de  $f(t)$  ponderada pelo núcleo  $D_N(\cdot)$ . Dizemos que  $D_N(\cdot)$  é uma janela pela qual  $f_N(t)$  “vê”  $f(t)$ .
2. De modo geral, se  $f(t)$  e  $g(t)$  são funções periódicas, de período  $T$ , tal que

$$g(t) = \int_{-T/2}^{T/2} K(t-u)f(u)du, \quad (12)$$

dizemos que  $K(\cdot)$  é uma **janela pela qual  $g(t)$  “vê”  $f(t)$** .

3. Propriedades de  $K(\cdot)$  geralmente necessárias são:

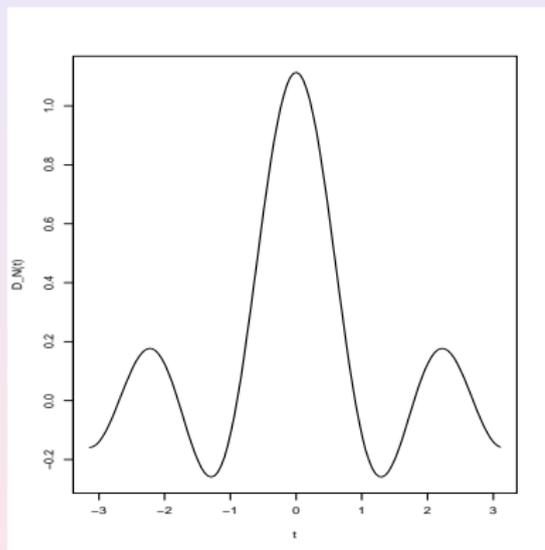
$$(i) \int_{-T/2}^{T/2} K(t)dt = 1;$$

$$(ii) K(-t) = K(t);$$

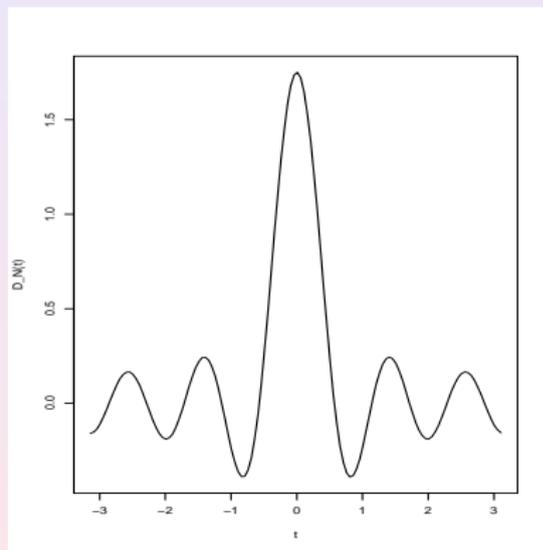
$$(iii) |K(t)| \leq K(0), \text{ para todo } t.$$

## Dirichlet

Na Figura 1, temos gráficos de  $D_N(\cdot)$  para  $T = 2\pi$  e dois valores de  $N$ .



(a)



(b)

Figura 1: Núcleo de Dirichlet. (a)  $N = 3$  (b)  $N = 5$ .

## Féjer

1. A relação (10) mostra que a aproximação de  $f(t)$  por meio de  $f_N(t)$  será boa se o núcleo  $D_N(\cdot)$  for concentrado ao redor do ponto  $t$ . A Figura 1, por outro lado, nos mostra que o núcleo de Dirichlet não é satisfatório, pois não está suficientemente concentrado ao redor da origem. Mesmo para  $N = 10$ , há muitos picos laterais, com magnitudes razoáveis.
2. Para que  $f_N(t)$  convirja para  $f(t)$ , quando  $N \rightarrow \infty$ , em algum sentido,  $f(t)$  deve satisfazer certas condições de regularidade. No entanto, o **teorema de aproximação de Weierstrass** afirma que se  $f(t)$  for uma função real, contínua, definida no intervalo  $[a, b]$ , existe uma sucessão de polinômios que converge uniformemente para  $f(t)$  em  $[a, b]$ .
3. Uma maneira de construir polinômios harmônicos convergentes para a função em questão foi proposta por Fejér em 1904. Seja  $f(t)$  dada pela sua série de Fourier e considere a média

$$\sigma_{N+1}(t) = \frac{1}{N+1} \{f_0(t) + f_1(t) + \dots + f_N(t)\}. \quad (13)$$

1. Então, pode-se provar que:

- (i)  $\sigma_{N+1}(t)$  converge para  $f(t)$  uniformemente, se  $f(t)$  for contínua e periódica;
- (ii)  $\sigma_{N+1}(t) = \sum_{n=-N}^N (1 - \frac{|n|}{N+1}) c_n e^{i\lambda_n t}$ ;
- (iii)  $\sigma_{N+1}(t) = \int_{-T/2}^{T/2} F_{N+1}[\frac{2\pi}{T}(t-u)] f(u) du$ , onde

$$F_{N+1}(t) = \frac{1}{T(N+1)} \left[ \frac{\text{sen}[(N+1/2)t]}{\text{sen}(t/2)} \right]^2, \quad (14)$$

para  $t \neq 0, \pm T, \pm 2T, \dots$  é o **núcleo de Fejér**.

2. Este tem as propriedades: (a)  $F_{N+1}(t)$  é uma função par;

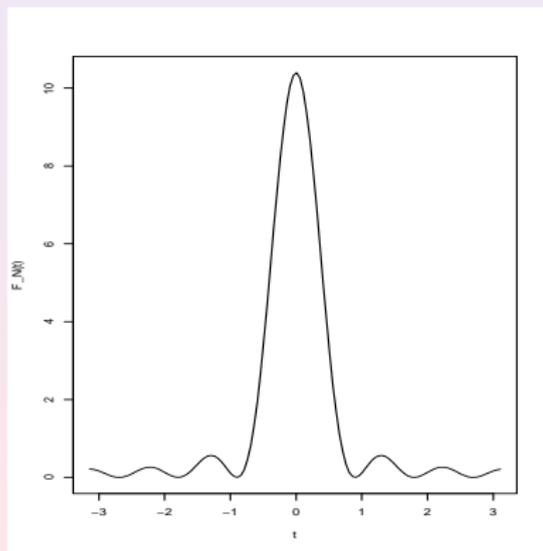
(b)  $\int_{-T/2}^{T/2} F_{N+1}(t) dt = 1$ ;

(c)  $F_{N+1}(2\pi t/T)$  é contínua, periódica, de período  $T$ ;

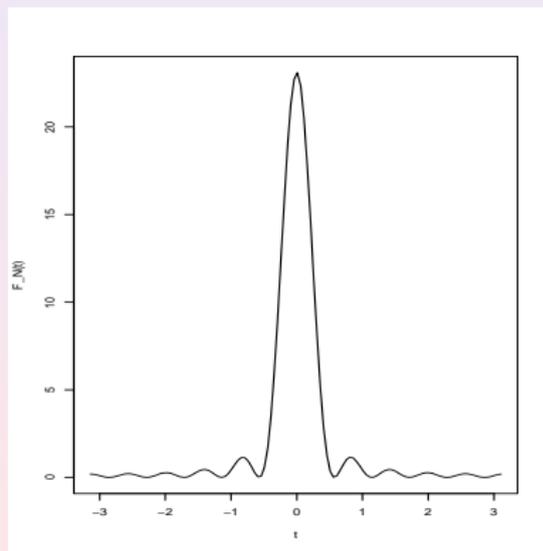
(d)  $F_{N+1}(0) = \frac{N+1}{T}$ .

## Féjer

A Figura 2 ilustra  $F_{N+1}(\cdot)$  para  $T = 2\pi$  e dois valores de  $N$ . Vemos que o núcleo de Féjer é uma janela mais concentrada ao redor da origem do que o núcleo de Dirichlet. Além disso,  $F_{N+1}(t) \geq 0$ , para todo  $t$ , e os picos laterais têm magnitudes pequenas quando comparadas com a do pico central.



(a)



(b)

Figura 2: Núcleo de Féjer

(a)  $N = 3$  (b)  $N = 5$

# Núcleos

1. Podemos considerar, genericamente,

$$f^{(N)}(t) = \sum_n h(n/N) c_n e^{i\lambda_n t}, \quad (15)$$

para alguma função  $h(t)$ , com  $h(0) = 1$ ,  $h(t) = 0$ ,  $|t| > 1$ . A função  $h(n/N)$  é chamada **fator de convergência**, **coeficientes geradores de janelas** ou, ainda, **tapers** e **data windows**

2. Se definirmos

$$H^{(N)}(t) = \frac{1}{T} \sum_n h(n/N) e^{-int}, \quad (16)$$

então podemos escrever (15) como

$$f^{(N)}(t) = \int_{-T/2}^{T/2} H^{(N)}\left[\frac{2\pi}{T}(t-u)\right] f(u) du. \quad (17)$$

3. A função  $H^{(N)}(\cdot)$  é o que chamaremos **janela espectral**, definida no domínio da frequência, enquanto  $h(n/N)$  é definida no domínio do tempo. A forma típica de  $H^{(N)}(\cdot)$  é aquela das figuras 1 e 2.
4. Vários fatores de convergência têm sido sugeridos na literatura. A Tabela 1 ilustra alguns deles.

# Núcleos e janelas

Tabela 1.1 Núcleos e Janelas

$h(n/N), 0 \leq  n  \leq N$	$H^{(N)}(t), -\pi < t < \pi$	Autores
1	$D_N(t)$	Dirichlet
$1 - \frac{ n }{N+1}$	$F_{N+1}(t)$	Fejér, Bartlett
$\frac{1}{2}(1 + \cos(\frac{\pi n}{N}))$	$\frac{1}{2}D_N(t) + \frac{1}{4}D_N(t - \frac{\pi}{N})$ $+ \frac{1}{4}D_N(t + \frac{\pi}{N})$	Hamming, Tukey
$1 - \frac{6n^2}{N^2}(1 - \frac{ n }{N}),  n  \leq \frac{N}{2}$ $2(1 - \frac{ n }{N})^3,  n  \geq \frac{N}{2}$	$\frac{2+\cos(t)}{4\pi N^3} \left[ \frac{\text{sen}(Nt/4)}{\text{sen}(t/4)} \right]^4$	de la Vallée- Poussin, Parzen
$\exp\{-n^2/2N^2\}$	$\approx \frac{N}{\sqrt{2\pi}} \exp\{-N^2 t^2/2\}$	Gauss, Weierstrass
$1 - \frac{n^2}{N^2}$	$D_N(t) + \frac{1}{N^2} \frac{d^2 D_N(t)}{dt^2}$	Riesz, Bochner, Parzen

# Variável aleatória

1. Seja  $(\Omega, \mathcal{A}, P)$  um espaço de probabilidades, no qual  $\Omega$  é o espaço amostral,  $\mathcal{A}$  é uma  $\sigma$ -álgebra de subconjuntos de  $\Omega$  (os eventos aleatórios) e  $P$  é uma medida de probabilidade sobre  $\mathcal{A}$ . Uma **variável aleatória** (v.a.) é uma função mensurável definida sobre  $\Omega$  com valores reais, isto é, uma função  $X : \Omega \rightarrow \mathbb{R}$  tal que, para todo conjunto de Borel  $B$  de  $\mathbb{R}$ , teremos que  $X^{-1}(B)$  é um evento de  $\mathcal{A}$ . Aqui,  $\mathbb{R}$  denota o conjunto dos números reais. Pode-se provar que  $X$  é uma v.a. se e somente se  $X^{-1}(J) \in \mathcal{A}$ , sendo  $J$  um intervalo da forma  $(-\infty, a]$ .
2. Se  $F$  for a função de distribuição (f.d.) de  $X$ , isto é,  $F(x) = P(X \leq x) = P(\omega : X(\omega) \leq x)$ , então o *valor esperado*, ou a *média* de  $X$  é definida como

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

que se reduz a

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

se  $F(x)$  for absolutamente contínua, isto é,  $dF(x) = f(x)dx$ , com  $f(x)$  sendo a função densidade de probabilidade (f.d.p.) de  $X$ . Se  $X$  for discreta, tomando valores em um conjunto enumerável  $\{x_1, x_2, \dots\}$  com probabilidades  $p(x_j) = P(X = x_j) \geq 0$ , para todo  $j$ , então

$$E(X) = \sum_{i=1}^{\infty} x_i p(x_i).$$

## Variável aleatória

1. Em todos os casos, supõe-se que a integral ou soma exista.  
A **variância** de  $X$  é definida por

$$\text{Var}(X) = E\{[X - E(X)]^2\},$$

e é fácil ver que  $\text{Var}(X) = E(X^2) - [E(X)]^2$ .

2. Se  $X$  e  $Y$  são duas v.a.'s definidas sobre  $(\Omega, \mathcal{A}, P)$ , então a **distribuição conjunta** de  $X$  e  $Y$  é especificada por uma f.d. bivariada  $F(x, y)$ , tal que

$$F(x, y) = P(\omega : X(\omega) \leq x, Y(\omega) \leq y).$$

Se  $F(x, y)$  for absolutamente contínua, existe  $f(x, y)$  tal que  $\partial^2 F(x, y)/\partial x \partial y = f(x, y)$  e  $F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv$ . No caso discreto, quando  $(X, Y)$  toma valores em um conjunto enumerável  $\{(x_i, y_j), i, j = 1, 2, \dots\}$ , existe  $p(x_i, y_j)$  tal que  $F(x, y) = \sum_i \sum_j p(x_i, y_j)$ , onde a soma estende-se sobre todos os pares  $(x_i, y_j)$  tais que  $x_i \leq x, y_j \leq y$ .

3. A **covariância** entre  $X$  e  $Y$  é definida por

$$\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\},$$

que é a  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x - E(X)][y - E(Y)] dF(x, y)$ . Segue-se facilmente que

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

## Variável aleatória

1. As v.a.'s  $X$  e  $Y$  são **independentes** se  $F(x, y) = F_X(x)F_Y(y)$ , para todos os pares  $(x, y)$ , onde  $F_X$  e  $F_Y$  são as f.d.'s (marginais) de  $X$  e  $Y$ , respectivamente. Nesse caso  $E(XY) = E(X)E(Y)$  e  $\text{Cov}(X, Y) = 0$ .
2. Quando  $\text{Cov}(X, Y) = 0$ , dizemos que  $X$  e  $Y$  são *não correlacionadas*. Segue-se que se  $X$  e  $Y$  são independentes, então elas são não correlacionadas, mas a recíproca não é verdadeira.
3. O *coeficiente de correlação*, que é uma medida de associação linear entre  $X$  e  $Y$ , é obtido padronizando-se a expressão para a covariância:

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y},$$

onde  $\sigma_X$  e  $\sigma_Y$  são os desvios padrões de  $X$  e  $Y$ , respectivamente. Não é difícil provar que  $-1 \leq \rho(X, Y) \leq 1$ .

## Variável aleatória

1. Suponha, agora, que  $X$  seja uma v.a. complexa, ou seja,  $X : \Omega \rightarrow \mathcal{C}$ , sendo que  $\mathcal{C}$  é o conjunto dos números complexos. Então,  $X = Y + iZ$ , com  $Y$  e  $Z$  v.a.'s reais. Dizemos que  $Y$  é a parte real de  $X$  e escrevemos  $Y = \mathcal{R}(X)$ , enquanto  $Z$  é a parte imaginária de  $X$ , e escrevemos  $Z = \mathcal{I}(X)$ .
2. Nesse caso, a média de  $X$  é definida por

$$E(X) = E(Y) + iE(Z),$$

e a variância de  $X$  é definida por

$$\text{Var}(X) = E\{|X - E(X)|^2\}.$$

3. É fácil ver que  $\text{Var}(X) = \text{Var}(Y) + \text{Var}(Z)$  e que, para  $Z = 0$ , obtemos as definições usuais para v.a.'s reais.
4. Sejam, agora,  $X_1$  e  $X_2$  duas v.a.'s complexas. A covariância entre ambas agora é definida como

$$\text{Cov}(X_1, X_2) = E\{[X_1 - E(X_1)][\overline{X_2 - E(X_2)}]\},$$

pois se assim não fosse, colocando  $X_1 = X_2$ , a variância não seria real. Vale a relação

$$\text{Cov}(X_1, X_2) = \text{Cov}(Y_1, Y_2) + \text{Cov}(Z_1, Z_2) + i[\text{Cov}(Z_1, Y_2) - \text{Cov}(Y_1, Z_2)],$$

se  $X_j = Y_j + iZ_j$ ,  $j = 1, 2$ .

# Desigualdades

1. **Desigualdade de Chebyshev.** Se  $\phi$  for uma função positiva e estritamente crescente sobre  $(0, \infty)$ , com  $\phi(u) = \phi(-u)$ , e  $X$  for uma v.a. com média finita, então para cada  $u > 0$ ,

$$P\{|X| \geq u\} \leq \frac{E\{\phi(X)\}}{\phi(u)}.$$

O caso familiar surge quando  $\phi(u) = u^2$  e  $X = Y - E(Y)$ , resultando

$$P\{|Y - E(Y)| \geq u\} \leq \frac{\text{Var}(Y)}{u^2}.$$

2. **Desigualdade de Jensen.** Se  $\phi$  for uma função convexa sobre  $\mathbb{R}$  e  $X$  e  $\phi(X)$  forem v.a.'s integráveis,

$$\phi\{E(X)\} \leq E\{\phi(X)\}.$$

Por exemplo, se  $\phi(x) = x^2$ , então  $[E(X)]^2 \leq E(X^2)$ .

3. **Desigualdade de Schwarz.** Se  $X$  e  $Y$  forem v.a.'s, então

$$|E(XY)| \leq E(|XY|) \leq [E|X|^2 E|Y|^2]^{1/2}.$$

# Convergência de v.a.

1. Seja  $\{X_n, n \geq 1\}$  uma sequência de v.a.'s definidas no mesmo espaço de probabilidades, digamos  $(\Omega, \mathcal{A}, P)$ .
2. A sequência  $\{X_n\}$  **converge em probabilidade** para a v.a.  $X$  se e somente se para todo  $\epsilon > 0$  tivermos

$$\lim_{n \rightarrow \infty} P\{|X_n - X| > \epsilon\} = 0,$$

e indicamos  $X_n \xrightarrow{P} X$ .

3. A sequência  $\{X_n\}$  **converge quase certamente** para a v.a.  $X$  se e somente se existe um conjunto  $D$  com  $P(D) = 0$  e tal que

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega) < \infty, \text{ para todo } \omega \in (\Omega - N)$$

e indicamos por  $X_n \xrightarrow{q.c.} X$ .

## Convergência de v.a.

1. O conceito de convergência quase certamente (ou quase em toda a parte) é mais forte do que o conceito de convergência em probabilidade, e o seguinte resultado vale.

**Proposição.** Se  $X_n \xrightarrow{q.c.} X$  valer, então  $X_n \xrightarrow{P} X$  também vale.

2. A sequência  $\{X_n\}$  converge para a v.a.  $X$  em **média quadrática** (m.q.) ou em  $L_2$ , se e somente se

$$\lim_{n \rightarrow \infty} E\{|X_n - X|^2\} = 0.$$

A notação é  $X_n \xrightarrow{m.q.} X$  ou  $X_n \xrightarrow{L_2} X$ .

3. Da desigualdade de Chebyshev segue imediatamente que

**Proposição.** Se  $X_n \xrightarrow{m.q.} X$ , então  $X_n \xrightarrow{P} X$ .

## Convergência de v.a.

1. Convergência em m.q. e em probabilidade não implicam convergência quase certamente, mas o seguinte resultado vale.

**Proposição.** Se a sequência  $\{X_n\}$  convergir em m.q. para  $X$  (e logo em probabilidade), então existe uma subsequência  $n_k$  tal que  $X_{n_k}$  converge quase certamente para  $X$ .

2. Convergência quase certamente não implica convergência em média quadrática.
3. Dizemos que  $X_n$  é **dominada** por  $Y$  se e somente se  $|X_n| \leq Y$  q.c. Então, é verdade que, se  $X_n \xrightarrow{P} X$ , então  $X_n \xrightarrow{m.q.} X$ , desde que  $\{X_n\}$  seja dominada por uma v.a.  $Y$  que seja de quadrado integrável.
4. Seja  $\{X_n\}$  uma sequência de v.a.'s e  $\{F_n\}$  a sequência correspondente de f.d.'s. Dizemos que  $\{X_n\}$  converge para a v.a.  $X$  em distribuição se e somente se a sequência  $F_n$  converge para a f.d.  $F$  de  $X$  nos pontos de continuidade de  $F$ .  
Usaremos a notação  $X_n \xrightarrow{D} X$ .
5. Esse é o conceito de convergência “mais fraco”, pois convergência em probabilidade implica em convergência em distribuição.

## Teorema de Slutsky e TLC

1. Se  $X_n$  converge em probabilidade para  $X$  e  $Y_n$  converge em probabilidade para  $Y$ , então  $X_n + Y_n$  converge em probabilidade para  $X + Y$ , e o mesmo vale para produtos.
2. Essas afirmações não valem para convergência em distribuição.
3. Uma exceção é o Teorema de Slutsky: Se  $Y_n$  converge em probabilidade para uma constante  $c$ , e  $X_n$  converge em distribuição para  $X$ , então  $X_n + Y_n$  converge em distribuição para  $X + c$  e  $X_n Y_n$  converge em distribuição para  $cX$ .
4. (TLC) Suponha que  $\{X_1, \dots, X_n\}$  seja uma sequência i.i.d. de v.a. com  $E[X_i] = \mu$  e  $\text{Var}[X_i] = \sigma^2 < \infty$ . Quando  $n \rightarrow \infty$ , as v.a.  $\sqrt{n}(\bar{X}_n - \mu)$  converge em distribuição para uma v.a. normal  $\mathcal{N}(0, \sigma^2)$ .
5. TLC mais gerais valem para v.a. independentes e suposições sobre momentos: Lyapunov e Lindeberg-Feller.

## Referências

Morettin, P. A. (2014). *Ondas e Ondaletas*. Segunda edição. São Paulo: EDUSP.

Morettin, P. A. and Toloí, C. M. C. (2018). *Análise de Séries Temporais*. Terceira Edição. Blucher.

Percival, D. and Walden, A. (1993). *Spectral Analysis for Physical Applications*. Cambridge.