Probabilidades, Variáveis Aleatórias e Procesos Estocásticos

– revisão –

Profa GABRIELA W GABRIEL

IEE-S / ITA – Sala 195 – Ramal 5991

São José dos Campos, 26 de Agosto de 2021

Determinístico X Aleatório

Experimento - Toda e qualquer realização que produza algum **RESULTADO**

Resultado **CERTO** – **DETERMINÍSTICO**

Determinístico X Aleatório

Experimento - Toda e qualquer realização que produza algum **RESULTADO**

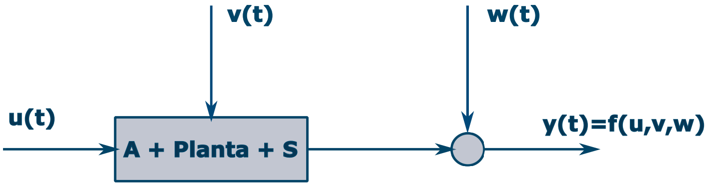
Resultado **CERTO** – **DETERMINÍSTICO**

Resultado **INCERTO** – **ALEATÓRIO**

Determinístico X Aleatório

- Planta Determinística

- Planta Estocástica



Espaço de Probabilidades

- Espaço Amostral / Evento Certo

> **ESPAÇO AMOSTRAL** ($\mathcal{S}$) é o conjunto de todos os possíveis resultados de um experimento.

- Evento

> **EVENTOS** (ξ_1, ξ_2, $\ldots$) são todos os possíveis resultados ou conjunto de resultados que um experimento pode gerar.

A todo EVENTO está associada uma medida de PROBABILIDADE.

Espaço de Probabilidades

- Sigma-álgebra

σ-álgebra ($\mathcal{F}$) associada a um conjunto $\mathcal{S}$ é uma classe **não vazia** de subconjuntos de $\mathcal{S}$ que satisfaz

1. $\emptyset \in \mathcal{F}$
2. $\xi_1, \xi_2, \ldots \in \mathcal{F} \implies \bigcup_{k=1}^{\infty} \xi_k \in \mathcal{F}$
3. $\xi \in \mathcal{F} \Rightarrow \overline{\xi} \in \mathcal{F}$

- Sempre é possível encontrar uma classe σ-álgebra que contenha os subconjuntos (eventos) de interesse.

- $\mathcal{F}$ define uma medida sobre $\mathcal{S}$.

Notação e Operações de Conjuntos

- Operações com Conjuntos
  1. **Igualdade**: $\mathcal{A} = \mathcal{B} \Leftrightarrow \mathcal{A} \subset \mathcal{B}$ e $\mathcal{B} \subset \mathcal{A}$, simultaneamente.
  2. **Diferença de Conjuntos** ($\mathcal{A} - \mathcal{B}$): elementos que estão em $\mathcal{A}$ e não estão em $\mathcal{B}$.
  3. **União de Conjuntos** ($\mathcal{A} \cup \mathcal{B}$): conjunto de elementos que estão somente em $\mathcal{A}$, somente em $\mathcal{B}$ ou em ambos os conjuntos.
     (Esta operação é associativa, $(\mathcal{A} \cup \mathcal{B}) \cup \mathcal{C} = \mathcal{A} \cup (\mathcal{B} \cup \mathcal{C})$, e comutativa, $\mathcal{A} \cup \mathcal{B} = \mathcal{B} \cup \mathcal{A}$.)
     (O evento $\bigcup_{k=1}^{\infty} \xi_k$ é o evento que ocorre quando pelo menos um dos ξ_k ocorre, $k = 1, 2, \cdots$.)
  4. **Intersecção de Conjuntos** ($\mathcal{A} \cap \mathcal{B}$): conjunto dos elementos que estão em $\mathcal{A}$ e em $\mathcal{B}$ simultaneamente.
     (Esta operação é associativa, $(\mathcal{A} \cap \mathcal{B}) \cap \mathcal{C} = \mathcal{A} \cap (\mathcal{B} \cap \mathcal{C})$, comutativa, $\mathcal{A} \cap \mathcal{B} = \mathcal{B} \cap \mathcal{A}$.)
     (A interseção e união satisfazem ainda a propriedade distributiva, $\mathcal{A} \cap (\mathcal{B} \cup \mathcal{C}) = (\mathcal{A} \cap \mathcal{B}) \cup (\mathcal{A} \cap \mathcal{C})$.)
     (O evento $\bigcap_{k=1}^{\infty} \xi_k$ é o evento que ocorre quando todos os ξ_k ocorrem simultaneamente, $k = 1, 2, \cdots$.)

Notação e Operações de Conjuntos

- Operações com Conjuntos (Continuação)
  5. **Complemento de Conjunto** ($\bar{\mathcal{A}} = \mathcal{S} - \mathcal{A}$): elementos de $\mathcal{S}$ que não estão em $\mathcal{A}$.
  6. **Mutuamente Exclusivos**: Dois conjuntos $\mathcal{A}$ e $\mathcal{B}$ são mutuamente exclusivos, ou disjuntos, se $\mathcal{A} \cap \mathcal{B} = \emptyset$.
  7. **Partição de um Conjunto**: Uma partição $\mathcal{X}$ de $\mathcal{S}$ é um conjunto de N subconjuntos não vazios e mutuamente exclusivos $\mathcal{X}_i$ de $\mathcal{S}$, tal que $\bigcup_{i=1}^{N} \mathcal{X}_i = \mathcal{S}$.

- Propriedades de Conjuntos
  1. **Conjunto Vazio** ($\emptyset$): Conjunto que não contém qualquer elemento.
  2. $\emptyset \subset \mathcal{A} \subset \mathcal{A} \subset \mathcal{S}$
  3. **Transitividades**: $\mathcal{C} \subset \mathcal{B}$ e $\mathcal{B} \subset \mathcal{A} \Rightarrow \mathcal{C} \subset \mathcal{A}$
  4. **Lei de De Morgan**: $\overline{\mathcal{A} \cup \mathcal{B}} = \overline{\mathcal{A}} \cap \overline{\mathcal{B}}$ e $\overline{\mathcal{A} \cap \mathcal{B}} = \overline{\mathcal{A}} \cup \overline{\mathcal{B}}$

Espaço de Probabilidades

- Espaço de Probabilidades

O **Espaço de Probabilidades** associado a um experimento corresponde a terna $(\mathcal{S}, \mathcal{F}, \mathcal{P})$, onde

$\mathcal{S}$ é o espaço amostral das possíveis combinações de resultados de um experimento.

$\mathcal{F}$ é a σ-álgebra associada ao experimento formada pelos eventos ξ, que são subconjuntos de $\mathcal{S}$.

$\mathcal{P}$ é a medida de probabilidade sobre $\mathcal{S}$, para a qual três axiomas devem ser satisfeitos:

$$\mathcal{P}(\xi) \geq 0, \ \forall \xi \in \mathcal{F};$$

$$\mathcal{P}(\mathcal{S}) = 1;$$

$$\xi_1 \cap \xi_2 = \emptyset \Rightarrow \mathcal{P}(\xi_1 \cup \xi_2) = \mathcal{P}(\xi_1) + \mathcal{P}(\xi_2), \ \forall \xi_1, \xi_2 \in \mathcal{F}.$$

Definições de Probabilidades

- Definição Clássica

DEFINIÇÃO CLÁSSICA (A Priori). Seja $\mathcal{S}$ o espaço amostral resultado de um experimento que gera o conjunto de eventos $\mathcal{F} = \{\emptyset, \mathcal{S}, \xi_1, \xi_2, \cdots, \xi_N, \{\xi_1, \xi_2\}, \cdots\}$, em que cada evento simples ξ_i, $i = \{1, 2, \cdots, N\}$, de $\mathcal{S}$ é igualmente provável. Se $\mathcal{X}$ é um subconjunto de $\mathcal{S}$ contendo $N_{\mathcal{X}}$ elementos ξ_i de $\mathcal{S}$ determinados por $\mathcal{F}$, então

$$\mathcal{P}(\mathcal{X}) = \frac{N_{\mathcal{X}}}{N}$$

- Definição Frequencial

DEFINIÇÃO FREQUENCIAL (Experimental). Se um experimento é realizado n vezes sob mesmas condições, e ocorrem $n_{\mathcal{X}}$ resultados que favorecem $\mathcal{X}$, um subconjunto de $\mathcal{S}$ contendo os eventos ξ_i determinados por $\mathcal{F}$, então

$$\mathcal{P}(\mathcal{X}) = \lim_{n\to\infty} \frac{n_{\mathcal{X}}}{n}$$

Probabilidade Condicional

- Probabilidade Condicional

PROBABILIDADE CONDICIONAL é a probabilidade de ocorrer o evento ξ_i dado que ocorreu o evento ξ_j e é definida como

$$\mathcal{P}(\xi_i|\xi_j) = \frac{\mathcal{P}(\xi_i \cap \xi_j)}{\mathcal{P}(\xi_j)}$$

em que assumimos que $\mathcal{P}(\xi_j)$ é não nula.

- Propriedades decorrentes

1. $\xi_i \subset \xi_j \Rightarrow \mathcal{P}(\xi_j|\xi_i) = 1$;
2. $\xi_i \subset \xi_j \Rightarrow \mathcal{P}(\xi_i|\xi_j) \geq \mathcal{P}(\xi_i)$.

Probabilidade Condicional

- Lei da Probabilidade Total

LEI DA PROBABILIDADE TOTAL
Seja $\mathcal{X} = \{\mathcal{X}_1, \cdots, \mathcal{X}_N\}$ uma partição de $\mathcal{S}$, e ξ_Z um evento qualquer de $\mathcal{S}$. Então,

$$\mathcal{P}(\xi_Z) = \mathcal{P}(\xi_Z|\mathcal{S}) = \mathcal{P}(\xi_Z \cap \mathcal{S}) = \mathcal{P}(\xi_Z \cap (\mathcal{X}_1 \cup \cdots \cup \mathcal{X}_N))$$

como os eventos ξ_i são mutuamente exclusivos,

$$\mathcal{P}(\xi_Z) = \mathcal{P}(\xi_Z|\mathcal{X}_1)\mathcal{P}(\mathcal{X}_1) + \cdots + \mathcal{P}(\xi_Z|\mathcal{X}_N)\mathcal{P}(\mathcal{X}_N)$$

- Regra de Bayes

REGRA DE BAYES é definida como

$$\mathcal{P}(\xi_i|\xi_j) = \frac{\mathcal{P}(\xi_j|\xi_i)\mathcal{P}(\xi_i)}{\mathcal{P}(\xi_j)}$$

em que assumimos que $\mathcal{P}(\xi_j)$ é não nula.

Independência de Eventos

- Eventos Independentes

Dois eventos ξ_1 e ξ_2 são **INDEPENDENTES** se a probabilidade de ξ_1 condicionada a ξ_2 não depende de ξ_2, ou seja,

$$\mathcal{P}(\xi_1|\xi_2) = \mathcal{P}(\xi_1)$$

Neste caso, temos ainda que

$$\mathcal{P}(\xi_1 \cap \xi_2) = \mathcal{P}(\xi_1|\xi_2)\mathcal{P}(\xi_2) = \mathcal{P}(\xi_1)\mathcal{P}(\xi_2)$$

Paradoxo de Simpson

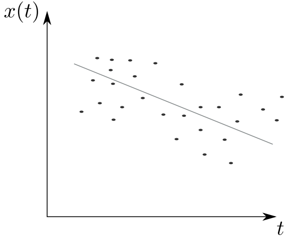
- Para diferentes grupos analisados, verifica-se tendências diferentes.
- Importância da escolha correta do conjunto de dados.

EXEMPLO 1: Dados de aprovação em um Concurso X

	Masculino			Feminino		
	Inscritos	Aprovados	Taxa	Inscritos	Aprovados	Taxa
CURSO A	900	480	53%	100	52	52%
CURSO B	100	36	36%	900	350	38%
TOTAL	1000	516	52%	1000	402	40%

Paradoxo de Simpson

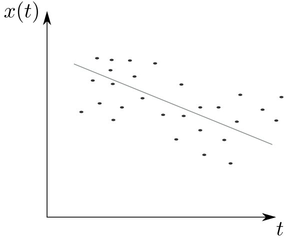
EXEMPLO 2: Análise Gráfica



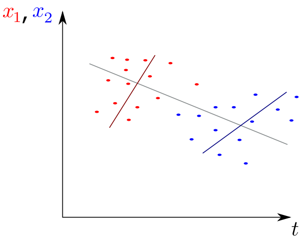
(A) Análise dos dados totais.

Paradoxo de Simpson

EXEMPLO 2: Análise Gráfica



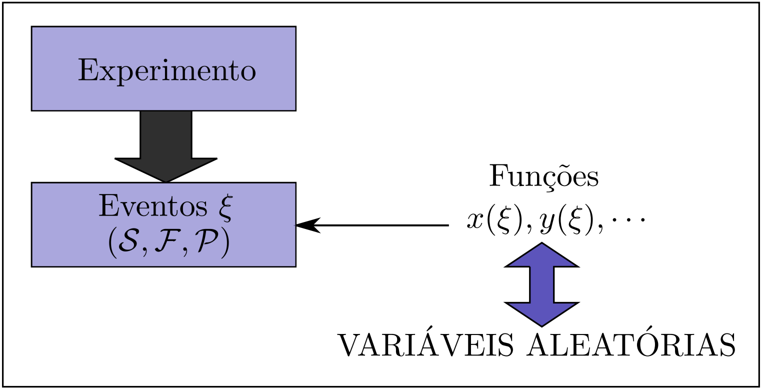
(A) Análise dos dados totais.



(B) Análise por grupos.

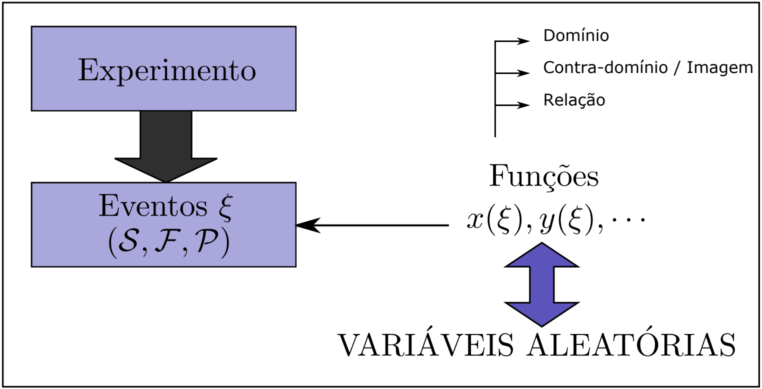
Introdução

Variável Aleatória é uma função $\mathsf{x} : \mathcal{S} \to \mathbb{R}^n$ que, para cada evento $\xi \in \mathcal{S}$, atribui um vetor (de números) $\mathsf{x}(\xi) \in \mathbb{R}^n$.



INTRODUÇÃO

Variável Aleatória é uma função x : $\mathcal{S} \to \mathbb{R}^n$ que, para cada evento $\xi \in \mathcal{S}$, atribui um vetor (de números) $x(\xi) \in \mathbb{R}^n$.



Notação Utilizada

NOTAÇÃO DE VA

- $x(\cdot)$, x denotam a variável aleatória.
- $x(\xi)$ denota a VA calculada para o evento $\xi \in \mathcal{S}$.
- $\{x \leq \alpha\} = \{\xi \in \mathcal{S} \mid x(\xi) \leq \alpha\}$.
- $\{x \in \mathcal{A}\} = \{\xi \in \mathcal{S} \mid x(\xi) \in \mathcal{A}\}$.

Observe que as duas últimas notações não são subconjuntos de $\mathbb{R}$ e sim de $\mathcal{S}$.

FORMALIZAÇÃO DA DEFINIÇÃO DE VA

Uma **variável aleatória** definida no conjunto dos números reais e associada ao espaço de probabilidades $\{\mathcal{S}, \mathcal{F}, \mathcal{P}\}$ é uma função x que mapeia o espaço amostral $\mathcal{S}$ em $\mathbb{R}$, ou seja, atribui um número real $x(\xi)$ a cada possível resultado de um experimento definido por $\xi \in \mathcal{S}$ de forma que:

[1.] o conjunto $\{x \leq \alpha\}$ é um evento de $\mathcal{F}$, $\forall \alpha \in \mathbb{R}$; e,

[2.] $\mathcal{P}(x = +\infty) = \mathcal{P}(x = -\infty) = 0$.

Qualquer ponto da reta real e qualquer intervalo definido ou semi-definido ou qualquer união ou interseção contável de conjuntos semi-definidos pode ser decomposto em intervalos do tipo $\{x \leq \alpha\}$.

O argumento das VA é um conjunto formado por resultados de um experimento e, portanto, de eventos de $\mathcal{S}$. Através do item [1.], estes são também elementos de $\mathcal{F}$ e a eles atribuímos uma medida de probabilidade!

Função Distribuição de Probabilidade

Associada à Definição de VA, temos a definição da **função distribuição** que define a medida da probabilidade associada à variável aleatória x quando esta assume valores dentro do intervalo $\{x \leq \alpha\}$.

A função **DISTRIBUIÇÃO DE PROBABILIDADES** $F_x(\alpha)$ é definida como

$$F_x(\alpha) = \mathcal{P}(x \leq \alpha), \ \forall \alpha \in \mathbb{R}$$

Propriedades da Função Distribuição de Probabilidade

PROPRIEDADES

(i) $F_x(+\infty) = 1$

(ii) $F_x(-\infty) = 0$

(iii) $F_x(\alpha_1) \leq F_x(\alpha_2)$ para $\alpha_1 \leq \alpha_2$ (função não decrescente)

(iv) $\lim_{\alpha \downarrow \alpha^*} F_x(\alpha) = F_x(\alpha^*_+) = F_x(\alpha^*)$ (função contínua pela direita)

- Concluímos que a FUNÇÃO DISTRIBUIÇÃO DE PROBABILIDADES
  - é **monótona não-decrescente**; e,
  - $\mathbb{I}\text{m}(F_x(\alpha)) = [0, 1]$.

Tipos de Variáveis Aleatórias

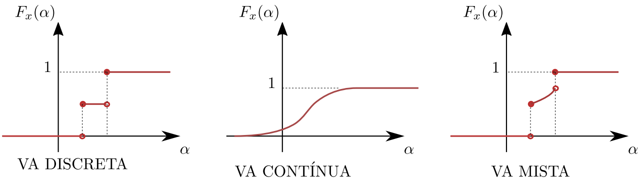


Figura: Tipo de Variáveis Aleatórias

VA Discreta:

$$F_x(\alpha) = \sum_{k=1}^{n} a_k u(\alpha - \alpha_k), \text{ onde } a_k = F_x(\alpha_k^+) - F_x(\alpha_k^-)$$

VA Contínua:

$$\mathcal{P}(x = \alpha) = F_x(\alpha_k^+) - F_x(\alpha_k^-) = 0$$

Variáveis Aleatórias Discretas

- Função Densidade de Probabilidade

FUNÇÃO DENSIDADE DE PROBABILIDADE (pmf) de uma variável aleatória **discreta** x tendo uma função distribuição F_x é qualquer função f_x definida em $\mathbb{R}$ como

$$f_x(\alpha_k) = \mathcal{P}(x = \alpha_k)$$

para todo $\alpha_k \in \mathbb{R}$, com $f_x(\alpha_k) \geq 0$, $\forall \alpha_k \in \mathbb{R}$.

onde x assume valores não nulos em α_k, ou seja,

$$\begin{cases} f_x(\alpha) > 0 & , \quad \forall \alpha = \alpha_k, \; k = 1, 2, \cdots \\ f_x(\alpha) = 0 & , \quad \forall \alpha \neq \alpha_k, \; k = 1, 2, \cdots \end{cases}$$

VARIÁVEIS ALEATÓRIAS DISCRETAS

- VA Discretas Especiais

  - **VA UNIFORME DISCRETA**

$$f_x(\alpha) = \frac{1}{n}\sum_{i=1}^{n} \delta_D(\alpha - \alpha_i)$$

  - **VA DE BERNOULLI**

$$f_x(\alpha = 0) = \mathcal{P}(x = 0) = 1 - p$$
$$f_x(\alpha = 1) = \mathcal{P}(x = 1) = p$$

onde p é tal que $0 \leq p \leq 1$

VARIÁVEIS ALEATÓRIAS DISCRETAS

- VA Discretas Especiais (CONT.)

  - **VA BINOMIAL**

$$f_x(\alpha) = \left(\begin{array}{c} n \\ \alpha \end{array} \right) p^{\alpha}(1-p)^{n-\alpha}, \; \alpha = 0, 1, \cdots, n$$

onde

$$\left(\begin{array}{c} n \\ \alpha \end{array} \right) = \frac{n!}{(n-\alpha)! \; \alpha!}$$

  - **VA GEOMÉTRICA**

$$f_x(\alpha) = \mathcal{P}(x = \alpha) = (1-p)^{\alpha-1} p, \; \alpha = 1, 2, \cdots$$

Variáveis Aleatórias Contínuas

- Função Densidade de Probabilidade

FUNÇÃO DENSIDADE DE PROBABILIDADES (pdf) de uma variável aleatória contínua x tendo uma função distribuição F_x é qualquer função não negativa f_x definida em $\mathbb{R}$ com a propriedade de que

$$F_x(\alpha) = \mathcal{P}(x \in (-\infty, \alpha]) = \int_{-\infty}^{\alpha} f_x(\gamma) d\gamma$$

para todo $\alpha \in \mathbb{R}$, com $f_x(\alpha) \geq 0$, $\forall \alpha \in \mathbb{R}$.

- $\exists \ \frac{dF_x(\alpha)}{d\alpha} \ \Rightarrow \ f_x(\alpha) \ = \ \frac{dF_x(\alpha)}{d\alpha}$
- $\mathcal{P}(x \in \mathbb{R}) = \mathcal{P}(\mathcal{S}) = \int_{-\infty}^{\infty} f_x(\gamma) d\gamma = 1$
- $\mathcal{P}(\alpha \leq x \leq \alpha + d\alpha) = \int_{\alpha}^{\alpha + d\alpha} f_x(\gamma) d\gamma = f_x(\alpha) d\alpha$, para $d\alpha \to 0$
- $\mathcal{P}(x = \alpha) = 0$ e, $\mathcal{P}(x \leq \alpha) = \mathcal{P}(x < \alpha)$

Variáveis Aleatórias Contínuas

- VA Contínuas Especiais

  - **VA UNIFORME CONTÍNUA**

$$f_x(\alpha) = \begin{cases} 0 & \alpha \notin [a, b] \\ \frac{1}{b-a} & \alpha \in [a, b] \end{cases}$$

  - **VA EXPONENCIAL**

$$f_x(\alpha) = \begin{cases} \lambda e^{-\lambda\alpha} & \alpha \geq 0 \\ 0 & \alpha < 0 \end{cases}$$

com parâmetro $\lambda > 0$.

VARIÁVEIS ALEATÓRIAS CONTÍNUAS

- VA Contínuas Especiais (CONT.)

  - **VA NORMAL**

$$f_x(\alpha) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\alpha-\mu)^2}{2\sigma^2}}$$

com parâmetros μ (média) e σ^2 (variância).

x definida por $x \sim N(\mu, \sigma^2)$ é dita gaussiana.

Outras funções de VA

FUNÇÃO DISTRIBUIÇÃO CONJUNTA é a função distribuição decorrente da **interseção de dois eventos distintos**. Assim sejam duas variáveis aleatórias x_1 e x_2, definidas no espaço de probabilidade $(\mathcal{S}, \mathcal{F}, \mathcal{P})$ em que o conjunto

$$\begin{aligned} \{x_1 \leq \alpha_1\} \cap \{x_2 \leq \alpha_2\} &= \{x_1 \leq \alpha_1, x_2 \leq \alpha_2\} \\ &= \{\xi \in \mathcal{S} \ : \ x_1(\xi) \leq \alpha_1 \text{ e } x_2(\xi) \leq \alpha_2\} \end{aligned}$$

Por definição, estes conjuntos são eventos de $\mathcal{F}$ e portanto $\{x_1 \leq \alpha_1, x_2 \leq \alpha_2\}$ também é um evento de $\mathcal{F}$. Logo, a função distribuição conjunta é dada por

$$F_{x_1 x_2}(\alpha_1, \alpha_2) = \mathcal{P}(x_1 \leq \alpha_1, x_2 \leq \alpha_2)$$

Suas propriedades são análogas ao caso unidimensional!

Outras funções de VA

FUNÇÃO DENSIDADE DE PROBABILIDADE CONJUNTA associada às variáveis aleatórias x_1 e x_2 é definida como

$$F_{x_1 x_2}(\alpha_1, \alpha_2) = \int_{-\infty}^{\alpha_1} \int_{-\infty}^{\alpha_2} f_{x_1 x_2}(\beta_1, \beta_2) d\beta_1 d\beta_2$$

tal que $f_{x_1 x_2}(\alpha_1, \alpha_2) \geq 0$, $\forall \alpha_1, \alpha_2 \in \mathbb{R}$.

FUNÇÃO DISTRIBUIÇÃO CONDICIONAL $F_{x_1|x_2}$ de uma variável aleatória x_1 dado o evento $\{x_2 \leq \alpha_2\}$ é definido como

$$F_{x_1|x_2}(\alpha_1|\alpha_2) \triangleq \mathcal{P}\{x_1 \leq \alpha_1 | x_2 \leq \alpha_2\} = \frac{F_{x_1 x_2}(\alpha_1, \alpha_2)}{F_{x_2}(\alpha_2)}$$

para todo $\alpha_1 \in \mathbb{R}^n$ e $\alpha_2 \in \mathbb{R}^m$.

Outras funções de VA

FUNÇÃO DENSIDADE DE PROBABILIDADE CONDICIONAL $f_{x_1|x_2}$ de uma variável aleatória x_1 dado o evento $\{x_2 = \alpha_2\}$ é definido como

$$f_{x_1|x_2}(\alpha_1|\alpha_2) = \frac{f_{x_1 x_2}(\alpha_1, \alpha_2)}{f_{x_2}(\alpha_2)}$$

para todo $\alpha_1 \in \mathbb{R}^n$ e $\alpha_2 \in \mathbb{R}^m$.

VARIÁVEIS ALEATÓRIAS INDEPENDENTES. Duas variáveis aleatórias x_1 e x_2 são independentes se os eventos $\{x_1 \leq \alpha_1\}$ e $\{x_2 \leq \alpha_2\}$ forem independentes para qualquer α_1 e α_2, ou seja,

$$\mathcal{P}(x_1 \leq \alpha_1, x_2 \leq \alpha_2) = \mathcal{P}(x_1 \leq \alpha_1)\mathcal{P}(x_2 \leq \alpha_2)$$

Outras funções de VA

N VARIÁVEIS ALEATÓRIAS INDEPENDENTES. N variáveis aleatórias $x_1, x_2, \cdots, x_N$ são independentes se os eventos $\{x_1 \leq \alpha_1\}, \{x_2 \leq \alpha_2\}, \cdots, \{x_N \leq \alpha_N\}$ forem independentes para quaisquer escolhas de $\alpha_1, \alpha_2, \cdots, \alpha_N$, ou seja,

$$\mathcal{P}(x_1 \leq \alpha_1, \cdots, x_N \leq \alpha_N) = \mathcal{P}(x_1 \leq \alpha_1) \cdots \mathcal{P}(x_N \leq \alpha_N)$$

Outras funções de VA

VALOR ESPERADO (VALOR MÉDIO)(MÉDIA)(ESPERANÇA) denotado por $\mathcal{E}(x)$, μ ou μ_x, de uma v.a. x com densidade de probabilidade f_x é definido por

$$\mathcal{E}(x) = \int_{-\infty}^{\infty} \alpha f_x(\alpha) d\alpha$$

caso a integral exista.

O ***k*-ÉSIMO MOMENTO** de uma v.a., $k = 0, 1, 2, \cdots$, também dito momento de ordem k, é denotado

$$m_{x^k} = \int_{-\infty}^{\infty} \alpha^k f_x(\alpha) d\alpha$$

caso a integral exista.

Notação utilizada: m_{x^k} ou simplesmente m_k.

Outras funções de VA

O k-ésimo **MOMENTO CENTRAL** de uma v.a., $k = 0, 1, 2, \cdots$, é o momento calculado em relação à sua média, ao invés de calculado em relação a zero, portanto, é denotado por

$$\eta_{x^k} = \int_{-\infty}^{\infty} (\alpha - \mu)^k f_x(\alpha) d\alpha$$

caso a integral exista.

Notação utilizada: η_{x^k} ou simplesmente η_k.

A **ESPERANÇA CONDICIONAL** de uma v.a. x, dada uma v.a. $y = \beta$, é

$$\mathcal{E}(x|y = \beta) = \int_{-\infty}^{\infty} \alpha f_{x|y}(\alpha|\beta) d\alpha$$

caso a integral exista.

Outras funções de VA

Duas v.a. x e y são **DESCORRELACIONADAS** se

$$\mathcal{E}(xy) = \mathcal{E}(x)\mathcal{E}(y)$$

caso a integral exista.

Duas v.a. x e y são **ORTOGONAIS** se

$$\mathcal{E}(xy) = 0$$

caso a integral exista.

Vetor de Variáveis Aleatórias

Um **vetor aleatório** $\mathrm{x} = \{x_1, \cdots, x_n\}$ é uma coleção de v.a. x_i, definidas para $\alpha = \{\alpha_1, \cdots, \alpha_n\}$. Assim,

$$\mathcal{P}(x \in D) = \int_D f_{\mathrm{x}}(\alpha) d\alpha$$

A **densidade de probabilidade conjunta** das v.a. x_i:

$$f_{\mathrm{x}}(\alpha) = f_{x_1, \cdots, x_n}(\alpha_1, \cdots, \alpha_n) = \frac{\partial^n F_{x_1, \cdots, x_n}(\alpha_1, \cdots, \alpha_n)}{\partial \alpha_1 \cdots \partial \alpha_n}$$

onde a **função distribuição conjunta** é dada por:

$$F_{\mathrm{x}}(\alpha) = F_{x_1, \cdots, x_n}(\alpha_1, \cdots, \alpha_n) = \mathcal{P}(x_1 \leq \alpha_1, \cdots, x_n \leq \alpha_n)$$

Vetor de Variáveis Aleatórias

A **MÉDIA DE UMA FUNÇÃO** $g(x)$ é dada por

$$\mathcal{E}(g(x)) = \int_{-\infty}^{\infty} g(\alpha_1, \cdots, \alpha_n) f_{x_1, \cdots, x_n}(\alpha_1, \cdots, \alpha_n) d\alpha$$

Chamaremos de

- V_{ij}, a **matriz de covariância** entre x_i e x_j
- R_{ij} é a **matriz de correlação** entre x_i e x_j.

Assim,

$$V_{ij} = \mathcal{E}((x_i - \mu_i)(x_j - \mu_j)) = \mathcal{E}(x_i x_j) - \mu_i \mu_j = R_{ij} - \mu_i \mu_j$$

$$\sigma_i^2 = V_{ii} = \mathcal{E}((x_i - \mu_i)^2) = \mathcal{E}(x_i^2) - (\mathcal{E}(x_i))^2$$

Duas v.a. são **MUTUAMENTE DESCORRELACIONADAS** se $V_{ij} = 0$ para todo $i \neq j$.

Sequência de Medidas

Seja $\mathrm{x} = \{x_1, \cdots, x_n\}$ uma **sequência de medidas** de χ.

As componentes de x

- são **independentes e identicamente distribuídos (iid)**
- possuem mesma distribuição de χ.

Definiremos a **média das amostras** M_n

$$M_n = \frac{1}{n}\sum_{j=1}^{n} x_j = \frac{S_n}{n}$$

SEQUÊNCIA DE MEDIDAS

Um bom **estimador para** χ deve satisfazer:

1. Na média, o estimador $\mu = \mathcal{E}(\chi) = \mathcal{E}(M_n)$;
2. $\mathcal{E}((M_n - \mu)^2)$ deve ser pequena.

De fato,

$$\mathcal{E}(M_n) = \mathcal{E}\left(\frac{1}{n}\sum_{j=1}^{n} x_j\right) = \frac{1}{n}\sum_{j=1}^{n}\mathcal{E}(x_j) = \mu$$

com $\mathcal{E}(\chi) = \mathcal{E}(x_j) = \mu$, $\forall j \in \{1, \cdots, n\}$.

A média das amostras é um **ESTIMADOR** **não polarizado** de μ.

Sequência de Medidas

Por outro lado,

$$\mathcal{E}((M_n - \mu)^2) = \mathcal{E}((M_n - \mathcal{E}(M_n))^2) = var(M_n)$$

e que

$$\begin{aligned} var(S_n) &= var(x_1 + x_2 + \cdots + x_n) \\ &= var(x_1) + var(x_2) + \cdots + var(x_n) = n \times var(\chi) = n\sigma^2 \end{aligned}$$

Como consequência,

$$var(M_n) = 1/n^2 \times var(S_n) = \sigma^2/n$$

Assim,

$$\boxed{\mathcal{P}\Big(\mathcal{E}(M_n) = \mu\Big) \to 1 \text{ quando } n \to \infty}$$

Lei dos Grandes Números

LEI DOS GRANDES NÚMEROS (Fraca): sejam $x_1, x_2, \cdots$ uma sequencia de iid variáveis aleatórias com média finita $\mathcal{E}(x) = \mu$, então para $\varepsilon > 0$,

$$\lim_{n \to \infty} \mathcal{P}(|M_n - \mu| < \varepsilon) = 1$$

que depende da média de χ. E,

LEI DOS GRANDES NÚMEROS (Forte): sejam $x_1, x_2, \cdots$ uma sequencia de iid variáveis aleatórias com média finita $\mathcal{E}(x) = \mu$ e variância finita, então

$$\mathcal{P}\left(\lim_{n \to \infty} M_n = \mu\right) = 1$$

As duas definições embutem o conceito de **convergência** de uma **sequência** de VA.

Sequências de V.A.

Uma **SEQUÊNCIA DE VA** é uma função que atribui uma quantidade infinita, porém contável, de valores reais ao evento ξ associado a cada resultado de um determinado espaço amostral $\mathcal{S}$. Representaremos esta sequência como

$$\{\mathsf{x}_n\} = \{\mathsf{x}_n(\xi)\} = \{x_1(\xi), \cdots, x_n(\xi), \cdots\}$$

Todo elemento x_i, $i = 1, \cdots, n$, é definido para o mesmo evento ξ.

Definições de Convergência de Sequências

A **sequência** $\{x_n\}$ **CONVERGE** para x se, dado qualquer $\varepsilon > 0$, nós podemos especificar um inteiro N, tal que, para todos os valores de n maiores que N, podemos garantir que $|x_n - x| < \varepsilon$.

Esta definição pressupõe o conhecimento do valor verdadeiro x.

A **sequência** $\{x_n\}$ **CONVERGE se e somente se**, dado qualquer $\varepsilon > 0$, nós podemos especificar um inteiro N', tal que, para n e m maiores que N', teremos que $|x_n - x_m| < \varepsilon$.

Das definições acima podemos estabelecer as seguintes definições de convergência de v.a.

Definições de Convergência de Sequências de V.A.

CONVERGÊNCIA NO SENTIDO CERTO. Neste caso, dizemos que a variável aleatória $\{x_n(\xi)\}$ converge no sentido certo para a variável aleatória $x(\xi)$, se a sequência de funções $x_n(\xi)$, calculada no ponto ξ, convergem para a função $x(\xi)$ quando $n \to \infty$ para todo evento $\xi \in \mathcal{S}$, em outras palavras,

$$X_n(\xi) \to X(\xi), \text{ quando } n \to \infty \ \forall \xi \in \mathcal{S}$$

Em cada ponto ξ a sequência converge para um valor $x(\xi)$.

Definições de Convergência de Sequências de V.A.

CONVERGÊNCIA NO SENTIDO QUASE CERTO. Uma sequência de v.a. $\{x_n(\xi)\}$ converge quase certamente para outra variável aleatória $x(\xi)$ se a sequência de funções $x_n(\xi)$ converge para a função $x(\xi)$ quando $n \to \infty$, para todo $\xi \in \mathcal{S}$, exceto possivelmente para um conjunto de eventos cuja probabilidade da convergência é nula, ou seja,

$$\mathcal{P}\left(\xi \ : \ X_n(\xi) \to X(\xi), \text{ quando } n \to \infty\right) = 1$$

Nesta última definição, podemos ter algum evento ξ, para o qual sua sequência não converge, porém, este evento deve pertencer ao conjunto de probabilidade nula.

Definições de Convergência de Sequências de V.A.

CASOS ESPECIAIS:

CONVERGÊNCIA EM MÉDIA QUADRÁTICA. A sequência de v.a. $\{x_n(\xi)\}$ converge em média quadrática para uma v.a. $x(\xi)$ se

$$\mathcal{E}((X_n(\xi) - X(\xi))^2) \to 0, \text{ quando } n \to \infty$$

CONVERGÊNCIA EM PROBABILIDADE. A sequência de v.a. $\{x_n(\xi)\}$ converge em probabilidade para uma v.a. $x(\xi)$ se, para todo $\varepsilon > 0$,

$$\mathcal{P}(|X_n(\xi) - X(\xi)| > \varepsilon) \to 0, \text{ quando } n \to \infty$$

Convergência de V.A.

CONVERGÊNCIA EM DISTRIBUIÇÃO. A sequência de v.a. $\{x_n(\xi)\}$ com funções distribuição de probabilidades $\{F_n(x)\}$ converge em distribuição para uma v.a. $x(\xi)$ com função distribuição $F(x)$, se

$$F_n(x) \to F(x), \text{ quando } n \to \infty$$

para todo x para o qual $F(x)$ é contínua.

Processos Estocásticos

PROCESSO ESTOCÁSTICO é uma família de v.a. $\{x_n(\xi)\}$ indexadas por um parâmetro $t \in \mathcal{T}$.

$$\{x(\xi, t), t \in \mathcal{T}\}$$

Notação: $x(\xi, t)$ ou $x(t)$.

A depender do parâmetro $t \in \mathcal{T}$, podem ser

- *a Tempo Contínuo*: $\mathcal{T}$ é um conjunto não enumerável contínuo.
- *a Tempo Discreto*: $\mathcal{T}$ é um conjunto enumerável (finito ou infinito), discreto.

Processos Estocásticos

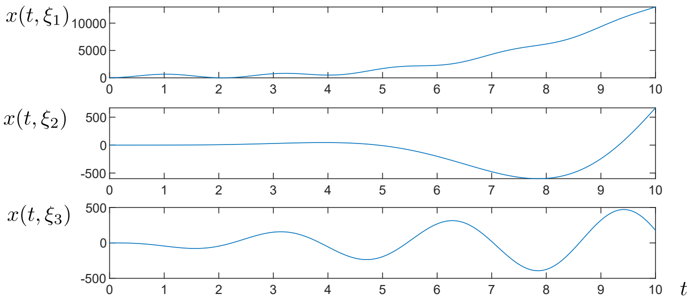


Figura: Exemplo de processo estocástico.

- $x(\xi_i, t_i)$, é um número real;
- $x(\xi_i, t)$, é uma função determinística do tempo;
- $x(\xi, t_i)$, é uma variável aleatória;
- $x(\xi, t)$, é um processo estocástico.

Probabilidades de P.E.

- Notação: $\{x(t) \leq \alpha\} = \{\xi \in \mathcal{S} \;:\; x(\xi, t) \leq \alpha \;\; \forall t \in \mathcal{T}\}$
- Expressões de Probabilidades:

FUNÇÃO DISTRIBUIÇÃO DE PROBABILIDADE de um processo estocástico $x(\xi, t)$ é definida como

$$F_{x(t)}(\alpha, t) = \mathcal{P}(x(t) \leq \alpha)$$

para todo $t \in \mathcal{T}$ e todo $\alpha \in \mathbb{R}^n$.

FUNÇÃO DENSIDADE DE PROBABILIDADE de um processo estocástico $x(\xi, t)$ é qualquer função $f_{x(t)}$ definida em $\mathbb{R}^n$ que satisfaça

$$F_{x(t)}(\alpha, t) = \int_{-\infty}^{\alpha} f_{x(t)}(\beta, t) d\beta$$

para todo $t \in \mathcal{T}$ e todo $\alpha \in \mathbb{R}^n$, e $f_{x(t)}(\beta, t) \geq 0$ para todo $\beta \in \mathbb{R}^n$ e todo $t \in \mathcal{T}$.

PROBABILIDADES CONJUNTAS DE P.E.

– Para $t_1, t_2 \in \mathcal{T}$ tais que $x(t_1)$ e $x(t_2)$ são duas v.a., definimos o evento

$$\{x(t_1) \leq \alpha_1, x(t_2) \leq \alpha_2\} = \{\xi \ : \ x(t_1) \leq \alpha_1 \text{ e } x(t_2) \leq \alpha_2\}$$

– Para este conjunto de eventos definimos

FUNÇÃO DISTRIBUIÇÃO DE SEGUNDA ORDEM (uma vez que relacional o comportamento de $x(t)$ em dois instantes distintos de tempo) associada a um processo estocástico $x(\xi, t)$ como

$$F_{x(t_1)x(t_2)}(\alpha_1, \alpha_2, t_1, t_2) = \mathcal{P}(x(t_1) \leq \alpha_1, x(t_2) \leq \alpha_2)$$

para todo $t_1, t_2 \in \mathcal{T}$ e todo $\alpha_1, \alpha_2 \in \mathbb{R}^n$.

Probabilidades Conjuntas de P.E.

FUNÇÃO DENSIDADE DE SEGUNDA ORDEM (uma vez que relacional o comportamento de $x(t)$ em dois instantes distintos de tempo) associada a um processo estocástico $x(\xi, t)$ como qualquer função $f_{x(t_1)x(t_2)}$ definida em $\mathbb{R}^n$ que satisfaça

$$F_{x(t_1)x(t_2)}(\alpha_1, \alpha_2, t_1, t_2) = \int_{-\infty}^{\alpha_1} \int_{-\infty}^{\alpha_2} f_{x(t_1)x(t_2)}(\beta_1, \beta_2, t_1, t_2) d\beta_2 d\beta_1$$

para todo $t_1, t_2 \in \mathcal{T}$ e todo $\alpha_1, \alpha_2 \in \mathbb{R}^n$ e $f_{x(t_1)x(t_2)}(\beta_1, \beta_2, t_1, t_2) \geq 0$ para todo $\beta_1, \beta_2 \in \mathbb{R}^n$ e todo $t_1, t_2 \in \mathcal{T}$.

Probabilidades de Segunda Ordem

– Funções marginais:

$$F_{x(t_1)}(\alpha_1, t_1) = F_{x(t_1)x(t_2)}(\alpha_1, \infty, t_1, t_2)$$
$$F_{x(t_2)}(\alpha_2, t_2) = F_{x(t_1)x(t_2)}(\infty, \alpha_2, t_1, t_2)$$
$$f_{x(t_1)}(\alpha_1, t_1) = \int_{-\infty}^{\infty} f_{x(t_1)x(t_2)}(\alpha_1, \alpha_2, t_1, t_2)d\alpha_2$$
$$f_{x(t_2)}(\alpha_2, t_2) = \int_{-\infty}^{\infty} f_{x(t_1)x(t_2)}(\alpha_1, \alpha_2, t_1, t_2)d\alpha_1$$

– Regra de Bayes:

$$f_{x(t_1)|x(t_2)}(\alpha_1, t_1 \mid \alpha_2, t_2) = \frac{f_{x(t_1)x(t_2)}(\alpha_1, \alpha_2, t_1, t_2)}{f_{x(t_2)}(\alpha_2, t_2)}$$
$$f_{x(t_2)|x(t_1)}(\alpha_2, t_2 \mid \alpha_1, t_1) = \frac{f_{x(t_1)x(t_2)}(\alpha_1, \alpha_2, t_1, t_2)}{f_{x(t_1)}(\alpha_1, t_1)}$$
$$f_{x(t_1)|x(t_2)}(\alpha_1, t_1 \mid \alpha_2, t_2) = \frac{f_{x(t_2)|x(t_1)}(\alpha_2, t_2 \mid \alpha_1, t_1)f_{x(t_1)}(\alpha_1, t_1)}{f_{x(t_2)}(\alpha_2, t_2)}$$

Probabilidades de Ordem k

FUNÇÃO DISTRIBUIÇÃO DE ORDEM k (uma vez que relacional o comportamento de $x(t)$ em dois instantes distintos de tempo) associada a um processo estocástico $x(\xi, t)$ como

$$F_{x(t_1)\cdots x(t_k)}(\alpha_1, \cdots, \alpha_k, t_1, \cdots, t_k) = \mathcal{P}(x(t_1) \leq \alpha_1, \cdots, x(t_k) \leq \alpha_k)$$

para todo $t_1, \cdots, t_k \in \mathcal{T}$ e todo $\alpha_1, \cdots, \alpha_k \in \mathbb{R}^n$.

FUNÇÃO DENSIDADE DE ORDEM k (uma vez que relacional o comportamento de $x(t)$ em k instantes distintos de tempo) associada a um processo estocástico $x(\xi, t)$ como qualquer função $f_{x(t_1)\cdots x(t_k)}$ definida em $\mathbb{R}^n$ que satisfaça

$$F_{x(t_1)\cdots x(t_k)}(\alpha_1, \cdots, \alpha_k, t_1, \cdots, t_k) = \int_{-\infty}^{\alpha_1} \cdots \int_{-\infty}^{\alpha_k} f_{x(t_1)\cdots x(t_k)}(\beta_1, \cdots, \beta_k, t_1, \cdots, t_k) d\beta_k \cdots d\beta_1$$

para todo $t_1, \cdots, t_k \in \mathcal{T}$ e todo $\alpha_1, \cdots, \alpha_k \in \mathbb{R}^n$ e
$f_{x(t_1)\cdots x(t_k)}(\beta_1, \cdots, \beta_k, t_1, \cdots, t_k) \geq 0$.

Esperança e Momentos de P.E.

A **MÉDIA** de um processo estocástico $x(\xi, t)$ é dado por

$$\mu_x(t) = \mathcal{E}(x(t)) = \int_{-\infty}^{\infty} \alpha f_{x(t)}(\alpha, t) d\alpha$$

que, no geral, são funções do tempo.

O **VALOR MÉDIO QUADRÁTICO** de um processo estocástico $x(\xi, t)$ é, para o caso vetorial,

$$P_x(t) = \mathcal{E}(x(t)x(t)') = \int_{-\infty}^{\infty} \alpha\alpha' f_{x(t)}(\alpha, t) d\alpha$$

Esperança e Momentos de P.E.

A (matriz de) **VARIÂNCIA** do processo $x(\xi, t)$ é dada por

$$\begin{aligned} V_x(t) &= \mathcal{E}\left((x(t) - \mu_x(t))(x(t) - \mu_x(t))'\right) \\ &= \int_{-\infty}^{\infty} (\alpha - \mu_x(t))(\alpha - \mu_x(t))' f_{x(t)}(\alpha, t) d\alpha \end{aligned}$$

Permanecendo válidas as relações:

$$V_x(t) = P_x(t) - \mu_x(t)\mu_x(t)'$$

e o Teorema Fundamental do Valor esperado

$$\mathcal{E}(g(x(t))) = \int_{-\infty}^{\infty} g(\alpha) f_{x(t)}(\alpha, t) d\alpha$$

Esperança e Momentos de P.E.

A (matriz de) **AUTOCORRELAÇÃO** de um p.e. $x(t)$ é a matriz $R_x(t_1, t_2)$ definida como

$$R_x(t_1, t_2) = \mathcal{E}\left(x(t_1)x(t_2)'\right)$$

A (matriz de) **AUTOCOVARIÂNCIA** de um p.e. $x(t)$ é a matriz $C_x(t_1, t_2)$ definida como

$$V_x(t_1, t_2) = \mathcal{E}\left((x(t_1) - \mu_x(t_1))(x(t_2) - \mu_x(t_2))'\right)$$

sendo $\mu(t) = \mathcal{E}(x(t))$.

- Observe que as definições acima, análogas às de v.a. foram formalizadas para p.e. vetoriais. Salienta-se que tanto as definições de v.a. quanto de p.e. são válidos para o caso $x \in \mathbb{R}^n$ ou $x(t) \in \mathbb{R}^n$, sendo $n \geq 1$, o que inclui o caso unidimensional.

Esperança e Momentos de P.E.

O **COEFICIENTE DE CORRELAÇÃO** de um p.e. $x(t) \in \mathbb{R}$ é a relação

$$r(t_1, t_2) = \frac{V_x(t_1, t_2)}{\sqrt{V_x(t_1) V_x(t_2)}}$$

sendo $\mu(t) = \mathcal{E}(x(t))$.

A (matriz de) **CROSS-CORRELAÇÃO** entre dois processos $x(t)$ e $y(t)$ é dada por

$$R_{xy}(t_1, t_2) = \mathcal{E}\left(x(t_1) y(t_2)'\right) = R_{yx}(t_2, t_1)'$$

A matriz de **CROSS-COVARIÂNCIA** entre dois processos $x(t)$ e $y(t)$ é dada por

$$V_{xy}(t_1, t_2) = \mathcal{E}\left((x(t_1) - \mu_x(t_1))(y(t_2) - \mu_y(t_2))'\right)$$

sendo $V_{xy}(t_1, t_2) = R_{xy}(t_1, t_2) - \mu_x(t_1)\mu_y(t_2)'$.

Esperança e Momentos de P.E.

Dois p.e. $x(t)$ e $y(t)$ são ditos **MUTUAMENTE ORTOGONAIS** se

$$R_{xy}(t_1, t_2) = 0, \quad \text{para toda escolha de } t_1, t_2$$

e um p.e. é **DESCORRELACIONADO** se

$$V_{xy}(t_1, t_2) = 0, \quad \text{para toda escolha de } t_1, t_2$$

Dizemos que um p.e. $x(t)$ é um **RUÍDO BRANCO** se os valores de $x(t_i)$ e de $x(t_j)$ são descorrelacionadas para todo t_i e $t_j \neq t_i$.

$$V_{xx}(t_1, t_2) = 0, \quad \text{para toda escolha de } t_1, t_2$$

Processos Estocásticos Estacionários

Dizemos que um p.e. $x(t)$ é **ESTACIONÁRIO NO SENTIDO ESTRITO** se $x(t)$ e $x(t+\tau)$ possuem as mesmas estatísticas para todo $\tau \ : \ t, t+\tau \in \mathcal{T}$.

Dizemos que um p.e. $x(t)$ é **ESTACIONÁRIO DE ORDEM** k se sua função distribuição $F_{x(t)}$ ou densidade $f_{x(t)}$ de ordem k são invariantes sob translação no tempo.

$$
\begin{aligned}
& f_{x(t_1)\cdots x(t_k)}(\alpha_1, \cdots, \alpha_k, t_1, \cdots, t_k) \\
& \qquad = f_{x(t_1+\tau)\cdots x(t_k+\tau)}(\alpha_1, \cdots, \alpha_k, t_1+\tau, \cdots, t_k+\tau)
\end{aligned}
$$

para todo $\tau \ : \ t_1, \cdots, t_k, t_1+\tau \cdots, t_k+\tau \in \mathcal{T}$.

Processos Estocásticos Estacionários

- Nem sempre é possível verificar todas as propriedades estatísticas de determinado processo estocástico. Neste caso, utilizamos um conceito mais abrangente (menos rigoroso) de estacionariedade.

Dizemos que um p.e. $x(t)$ é **ESTACIONÁRIO NO SENTIDO AMPLO** (WSS) se

$$\mu_x(t) = \mu_x \qquad \textbf{e} \qquad R_{x(t_1)x(t_2)}(t_1, t_2) = R_{x(t_1)x(t_2)}(t_1 - t_2)$$

Densidade Espectral de um P.E.

A **POTÊNCIA ESPECTRAL** (ou **DENSIDADE ESPECTRAL**) de um p.e. WSS, $x(t)$, real ou complexo, é a Transformada de Fourier $S(\omega)$ da sua autocorrelação $R_{x(t+\tau)x(t)}(\tau) = \mathcal{E}(x(t+\tau)x^*(t))$

$$S(\omega) = \int_{-\infty}^{\infty} R_{x(t+\tau)x(t)}(\tau) e^{-j\omega\tau} d\tau$$

P.E. GAUSSIANO

Um **P.E. GAUSSIANO** $x(t)$, $t \geq 0$, também chamado de **P.E. NORMAL** se cada uma das v.a. $x(t_i)$ possuir uma distribuição conjunta normal para todo conjunto $i = 1, 2, \cdots, n$, qualquer que seja $n \in \mathbb{N}^*$.

- Um p.e. gaussiano é completamente determinado por sua média e covariância. Representaremos este processo $x(t)$ por

$$x(t) \sim \mathcal{N}(\mu_x, \sigma_x^2)$$

em que σ_x^2 corresponde a sua variância.

Processo Estocástico Ergódico

Dizemos que um p.e. $x(t)$ é **ERGÓTICO** se sua média temporal é idêntica a sua média estocástica, ou seja,

$$\lim_{T\to\infty} \langle g(x(t))\rangle_T = \mathcal{E}\left(g(x(t))\right)$$

com probabilidade 1, para qualquer $g(\cdot)$ para a qual a integral acima exista. O símbolo $\langle\cdot\rangle_T$ denota a média temporal de $(\cdot)$ em T,

$$\langle y(t)\rangle_T = \frac{1}{2T}\int_{-T}^{T} y(t)dt$$

Processo de Markov

Um processo estocástico contínuo é dito **MARKOVIANO** se ele satisfaz a **propriedade de Markov**

$$\mathcal{P}\Big(a < x(t_k) \leq b \Big| x(t_{k-1}) = x_{k-1}, \cdots, x(t_1) = x_1\Big)$$
$$= \mathcal{P}\Big(a < x(t_k) \leq b \Big| x(t_{k-1}) = x_{k-1}\Big)$$

em que t_k são os instantes em que $x(t)$ é observado. Analogamente, também definimos a **CADEIA DE MARKOV** (caso discreto).

Um processo estocástico discreto é dito **MARKOVIANO** se ele satisfaz a **propriedade de Markov**

$$\mathcal{P}\Big(x(t_k) = x_k \Big| x(t_{k-1}) = x_{k-1}, \cdots, x(t_1) = x_1\Big)$$
$$= \mathcal{P}\Big(x(t_k) = x_k \Big| x(t_{k-1}) = x_{k-1}\Big)$$

Cadeias de Markov

- Seja uma cadeia de Markov, com **função densidade de probabilidade**

$$f_{x(k)}(j) = p_j(k), \;\; k = 0, 1, \cdots$$

- De forma que, no instante inicial, podemos estabelecer a seguinte relação

$$p_j(0) \triangleq \mathcal{P}(x_0 = j) \;\; j = 0, 1, 2, 3, \cdots$$

- Note que temos uma **CADEIA DE MARKOV FINITA** se $x_k \in \mathbb{T} \subset \mathbb{N}$, $\mathbb{T}$ finito. Caso contrário, se $\mathbb{T}$ é infinito, temos uma **CADEIA DE MARKOV INFINITA**.

Cadeia de Markov

- Utilizando a definição de probabilidade condicional, podemos prever o estado da cadeia a cada instante segundo

$$\begin{aligned}
&\mathcal{P}\left(x_k = i_k, \cdots, x_0 = i_0\right) \\
&= \mathcal{P}\left(x_k = i_k \middle| x_{k-1} = i_{k-1}\right) \cdots \mathcal{P}\left(x_1 = i_1 \middle| x_0 = i_0\right) \mathcal{P}\left(x_0 = i_0\right)
\end{aligned}$$

- Sendo a cadeia um <u>sistema dinâmico</u>, para prever seu estado futuro devemos saber a **probabilidade inicial** associada a cada um dos estados

$$\mathcal{P}\left(x_0 = i_0\right)$$

e a **probabilidade de transição de estados**

$$\mathcal{P}\left(x_{k+1} = i_{k+1} \middle| x_k = i_k\right) = p_{i_k i_{k+1}}, \ \forall k$$

Cadeia de Markov

- Assim, definimos a **MATRIZ DE TRANSIÇÃO DE PROBABILIDADES** como

$$P = \begin{bmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & p_{11} & \cdots \\ \cdots & & \\ p_{\ell 0} & p_{\ell 1} & \cdots \\ \cdots & & \end{bmatrix}$$

> Se a probabilidade de transição não se altera com o tempo, dizemos que temos uma **CADEIA DE MARKOV HOMOGÊNA**!

- A probabilidade acumulada entre os instantes k e 0 pode ser calculada como

$$\mathcal{P}\left(x_k = i_k, \cdots, x_0 = i_0\right) = p_{i_{k-\mathbf{1}} i_k} \cdots p_{i_\mathbf{1} i_\mathbf{0}} p_{i_\mathbf{0}}(0)$$

Cadeia de Markov

- Podemos escrever ainda o vetor $\mathrm{p}(k) = \{p_i(k)\}$ de **probabilidades do estado da cadeia de Markov** no instante k, onde

$$p_{i_n}(n) = \sum_{\ell} \mathcal{P}\left(x_n = i_n \middle| x_{n-1} = \ell\right) \mathcal{P}\left(x_{n-1} = \ell\right) = \sum_{\ell} p_{\ell i_n} p_{\ell}(n-1)$$

- Assim, a probabilidade dos estados futuros a partir do vetor de estados atual ou inicial são dados, respectivamente, por

$$\mathrm{p}(n) = \mathrm{p}(n-1)P \text{ ou } \mathrm{p}(n) = \mathrm{p}(0)P^n$$

Convergência da Cadeia de Markov

- A cadeia de Markov entra em processo estacionário para $n \to \infty$, quando

$$p_{ij}(n) \to \pi_j, \ \forall i$$

em consequência, P^n convergir equivale ao estado da cadeia de Markov convergir. Neste caso dizemos que a cadeia entrou em **estado permanente** ou em **equilíbrio** com

$$\pi = \pi P$$