

RESUMO DE PROBABILIDADE

Aula 1 - Continuidade da probabilidade

Definição 1.1. Um espaço de probabilidade é uma tripla (Ω, \mathcal{F}, P) , onde

- Ω é um conjunto qualquer contendo todos os possíveis resultados do experimento aleatório;
- \mathcal{F} é uma σ -álgebra de subconjuntos de Ω , isto é, \mathcal{F} satisfaz:
 - $\emptyset \in \mathcal{F}$;
 - se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$;
 - se $A_1, \dots, A_n, \dots \in \mathcal{F}$, então $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$;
- P é uma função $P : \mathcal{F} \rightarrow \mathbb{R}_+$, chamada **medida de probabilidade** (ou, apenas, probabilidade) satisfazendo:
 - $P(\emptyset) = 0$;
 - Se $A_1, \dots, A_n, \dots \in \mathcal{F}$ são dois a dois disjuntos (ou seja, $A_i \cap A_j = \emptyset$, para todo $i \neq j$), então $P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$;
 - $P(\Omega) = 1$.

Proposição 1.1.

1. Sejam $A, B \in \mathcal{F}$, com $A \subset B$, então $P(A) \leq P(B)$.
2. $P(A^c) = 1 - P(A)$.
3. $P(A) \in [0, 1], \forall A \in \mathcal{F}$.
4. (Continuidade no vazio) Sejam $(A_n) \in \mathcal{F}$, com $A_1 \supset A_2 \supset \dots \supset A_n \supset A_{n+1} \supset \dots$ e $\cap_{n=1}^{\infty} A_n = \emptyset$. Então $\lim_{n \rightarrow \infty} P(A_n) = 0$.

Teorema 1.1 (Continuidade da probabilidade).

- a) Se $(A_n) \in \mathcal{F}$ e $A_n \uparrow A$ (isto é, $A_n \subset A_{n+1}$ e $A = \cup_{n=1}^{\infty} A_n$), então $P(A) = \lim_{n \rightarrow \infty} P(A_n)$;
- b) Se $(A_n) \in \mathcal{F}$ e $A_n \downarrow A$ (isto é, $A_n \supset A_{n+1}$ e $A = \cap_{n=1}^{\infty} A_n$), então $P(A) = \lim_{n \rightarrow \infty} P(A_n)$;

Exercício 1.1.

1. Problema da agulha de Buffon.

2. Verdadeiro ou falso. Sejam \mathcal{A} e \mathcal{F} σ -álgebras.
 - a) $\mathcal{A} \cap \mathcal{F}$ é σ -álgebra.
 - b) $\mathcal{A} \cup \mathcal{F}$ é σ -álgebra.
3. Seja \mathcal{C} uma coleção qualquer de subconjuntos de Ω . Então existe $\sigma(\mathcal{C})$ a menor σ -álgebra que contém \mathcal{C} .
4. Mostre que se uma função for finitamente aditiva e se vale a continuidade no vazio, então essa função é σ -aditiva.

Aula 2 - Probabilidade condicional

Definição 2.1. Sejam (Ω, \mathcal{F}, P) espaço de probabilidade e $A, B \in \mathcal{F}$, com $P(B) \neq 0$. Definimos a **probabilidade condicional de A e B**, como $P(A | B) = \frac{P(A \cap B)}{P(B)}$.

Teorema 2.1 (Teorema da multiplicação). Sejam A_1, A_2, \dots, A_n eventos, então

$$P(\cap_{i=1}^n A_i) = P(A_1)P(A_2 | A_1) \dots P(A_n | \cap_{i=1}^{n-1} A_i).$$

Definição 2.2. Dizemos que uma família $(A_n)_{n=1}^\infty$ de eventos é uma **partição de Ω** , se

- i) $\cup_{n=1}^\infty A_n = \Omega$;
- ii) $A_i \cap A_j = \emptyset, \forall i \neq j$.

Teorema 2.2 (Teorema da probabilidade total). Sejam $B \in \mathcal{F}$ e (A_n) partição de Ω . Então $P(B) = \sum_{n=1}^\infty P(A_n)P(B | A_n)$.

Corolário 2.1 (Fórmula de Bayes). Sejam $B \in \mathcal{F}$ e (A_n) partição de Ω . Então, para todo i , $P(A_i | B) = \frac{P(A_i)P(B | A_i)}{\sum_{n=1}^\infty P(A_n)P(B | A_n)}$.

Definição 2.3. Sejam A, B eventos. Dizemos que A e B são **eventos independentes**, se $P(A \cap B) = P(A)P(B)$.

Teorema 2.3. Se A e B são eventos independentes, então A e B^c também são eventos independentes. Vale o mesmo para A^c e B , e A^c e B^c .

Teorema 2.4. Um evento A é independente de si mesmo se, e somente se, $P(A) = 0$ ou 1.

Definição 2.4. Seja $(A_i)_{i \in I}$ coleção de eventos. Dizemos que:

- a) a coleção $(A_i)_{i \in I}$ é **2 a 2 independente**, se $\forall i, j \in I, i \neq j$, temos que A_i e A_j são independentes;
- b) a coleção $(A_i)_{i \in I}$ é **coletivamente independente**, se para todo $k \in \mathbb{N}$ e para todo $i_1, i_2, \dots, i_k \in I, i_\ell \neq i_j$, vale $P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \prod_{j=1}^k P(A_{i_j})$.

Exercício 2.1.

1. Fixado $B \in \mathcal{F}$, prove que $P(\cdot \mid B) : \mathcal{F} \rightarrow [0, 1]$ é medida de probabilidade.
2. Exercícios de 1 a 23 do capítulo 1 do Barry James.

Aula 3 - Lema de Borel-Cantelli

Definição 3.1. Sejam (A_n) sequência de eventos. Definimos o **limite superior da sequência** (A_n) , denotado por $\limsup A_n$, $\overline{\lim} A_n$ ou $(A_n, i.v.)$, como sendo $\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$. Analogamente, definimos o **limite inferior da sequência** (A_n) , denotado por $\liminf A_n$ ou $\underline{\lim} A_n$, como sendo $\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$.

Proposição 3.1.

1. Seja $(A_n) \in \mathcal{F}$, então $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$.
2. a) Se $P(A_n) = 0, \forall n$, então $P(\bigcup_{n=1}^{\infty} A_n) = 0$;
b) Se $P(A_n) = 1, \forall n$, então $P(\bigcap_{n=1}^{\infty} A_n) = 1$.

Teorema 3.1 (Lema de Borel-Cantelli). Seja (A_n) uma sequência de eventos. Então

- Se $\sum P(A_n) < \infty$, então $P(\overline{\lim} A_n) = 0$.
- Se (A_n) for coletivamente independente e $\sum P(A_n) = \infty$, então $P(\overline{\lim} A_n) = 1$.

Corolário 3.1 (Lema 0 – 1 de Cantelli). Seja (A_n) uma sequência de eventos coletivamente independentes. Então $P(A_n, i.v.) = 0$ ou 1 .

Exercício 3.1.

1. a) Seja (A_n) tal que $\liminf A_n = \limsup A_n$. Defina $A = \liminf A_n = \limsup A_n$, e o denote por $\lim A_n$. Então prove que $\lim P(A_n) = P(A)$.
b) Verifique que se $A_n \uparrow A$, então $\overline{\lim} A_n = \underline{\lim} A_n = \bigcup_{n=1}^{\infty} A_n$. Se $A_n \downarrow A$, então $\overline{\lim} A_n = \underline{\lim} A_n = \bigcap_{n=1}^{\infty} A_n$.
2. Exercícios de 5, 6 e 7 do capítulo 5 do Barry James.

Aula 4 - Processo de Poisson

Definição 4.1. Seja (Ω, \mathcal{F}, P) um espaço de probabilidade. Uma função $X : \Omega \rightarrow \mathbb{R}$ (ou $\overline{\mathbb{R}}$) é chamada de **variável aleatória** (no contexto da teoria da medida é dita função mensurável), se, para todo $t \in \mathbb{R}$, temos que $\{\omega \in \Omega \mid X(\omega) \in (-\infty, t]\} \in \mathcal{F}$. De modo mais geral, se (Ξ, \mathcal{A}) é um espaço mensurável, dizemos que $X : \Omega \rightarrow \Xi$ é uma variável aleatória (ou função mensurável), se, para todo $A \in \mathcal{A}$, temos que $X^{-1}(A) \in \mathcal{F}$. Chamamos uma sequência de variáveis aleatórias (X_n) de **processo estocástico**.

Definição 4.2. Seja (Ω, \mathcal{F}, P) um espaço de probabilidade, onde

- $\Omega = \{\omega : \mathbb{R}^+ \rightarrow \mathbb{N} \cup \{0\} \mid \exists (t_k) \text{ sequência crescente tal que } \omega(t) = k - 1, \forall t \in (t_{k-1}, t_k]\}$, ou seja, Ω é o conjunto de funções degraus;
- $\mathcal{F} = \sigma(A_{s,t}^k)$, isto é, \mathcal{F} é a menor σ -álgebra que contém os conjuntos $A_{s,t}^k$ definidos como $\{\omega \in \Omega \mid \omega(t+s) - \omega(s) = k\}$, para todo $k \in \mathbb{N}$ e para todo $t, s \in \mathbb{R}^+$;
- P é a função de \mathcal{F} para $[0, 1]$ que satisfaz as seguintes hipóteses:

– **Incrementos estacionários:** $P(A_{s,t}^k)$ não depende de s , isto é,

$$P(A_{s,t}^k) = P(A_{0,t}^k) \quad \forall k \in \mathbb{N} \quad \forall s, t \in \mathbb{R}^+$$

Denotamos $P(A_{0,t}^k)$ por $P_k(t)$;

- **Incrementos independentes:** para todo $s, t, u, v \in \mathbb{R}^+$ e para todo $k, \ell \in \mathbb{N}$ tais que $(s, s+t] \cap (u, u+v] = \emptyset$, temos que $A_{s,t}^k$ e $A_{u,v}^\ell$ são independentes;
- **Não ocorrência simultânea:** $P(\text{pelo menos 2 ocorrências em } [0, t] \mid \text{pelo menos 1 ocorrências em } [0, t]) \xrightarrow{t \rightarrow 0^+} 0$, ou seja,

$$\frac{1 - P_0(t) - P_1(t)}{1 - P_0(t)} \xrightarrow{t \rightarrow 0^+} 0$$

que equivale a $\frac{P_1(t)}{1 - P_0(t)} \xrightarrow{t \rightarrow 0^+} 1$

Teorema 4.1. Para o espaço de probabilidade definido acima (definição 4.2), temos que $P_k(t) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$, onde $\lambda = -\ln(P_0(1))$

Definição 4.3. Considere (Ω, \mathcal{F}, P) o espaço de probabilidade definido acima (definição 4.2). Seja $X : \Omega \rightarrow \mathbb{N}$ a variável aleatória que conta o número de ocorrências em um intervalo de duração 1. Assim, $(X = k) = \{\omega \in \Omega \mid \omega(1+s) - \omega(s) = k\} = A_{s,1}^k, \forall s \in \mathbb{R}^+$, logo $P(X = k) = P_k(1) = \frac{\lambda^k e^{-\lambda}}{k!}$. Essa variável aleatória é chamada **variável de Poisson com parâmetro λ** . O processo estocástico definido como sendo a sequência de variáveis de Poisson $(X_t)_{t \in \mathbb{R}^+}$ com parâmetro λt é dito **processo de Poisson de intensidade λ** .

Aula 5 - Variáveis aleatórias

Definição 5.1. Seja (Ω, τ) espaço topológico. Então $\sigma(\tau)$, a menor σ -álgebra que contém todos os abertos de τ , é chamada **σ -álgebra de Borel** (ou borelianos).

Definição 5.2. Seja $X : \Omega \rightarrow \mathbb{R}$ variável aleatória. Definimos a **função de distribuição de X** como sendo a função $F_X : \mathbb{R} \rightarrow [0, 1]$, $z \mapsto P(X \leq z)$, onde $(X \leq z) = \{\omega \in \Omega \mid X(\omega) \leq z\}$.

Teorema 5.1 (Propriedades da função de distribuição).

- a) F_X é não decrescente.
- b) F_X é contínua à direita.
- c) $\lim_{z \rightarrow -\infty} F_X(z) = 0$ e $\lim_{z \rightarrow \infty} F_X(z) = 1$

Proposição 5.1.

1. $F_X(z) - F_X(z^-) = P(X = z)$, onde $F_X(z^-)$ é o limite lateral à esquerda de F_X . Este resultado diz que $P(X = z)$ é o salto de F_X no ponto z .
2. Seja $D = \{z \in \mathbb{R} \mid F_X \text{ é descontínua em } z\}$. Então D é enumerável. Em outros termos, F_X só admite uma quantidade enumerável de descontinuidades.
3. Se $F : \mathbb{R} \rightarrow [0, 1]$ é uma função que satisfaz as três propriedades do teorema 5.1, então F é função distribuição de alguma variável aleatória.

Definição 5.3. Seja $X : \Omega \rightarrow \mathbb{R}$ uma variável aleatória. Dizemos que X é:

- **discreta**, se $X(\Omega)$ for enumerável;
- **absolutamente contínua**, se existir uma função $f : \mathbb{R} \rightarrow \mathbb{R}^+$, dita **função densidade de probabilidade**, tal que $F_X(z) = \int_{-\infty}^z f(t)dt$.

Exemplo 5.1. Abaixo seguem exemplos de funções de distribuição para variáveis aleatórias discretas e absolutamente contínuas.

- Variáveis aleatórias discretas.

1. **Distribuição de Poisson com parâmetro λ** , $X \sim \pi(\lambda)$.

$$X : \Omega \rightarrow \mathbb{N} \cup \{0\} \text{ e } P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}$$

2. **Distribuição de Bernoulli com parâmetro p** , $X \sim \text{Ber}(p)$.

$$X : \Omega \rightarrow \{0, 1\} \text{ e } p = P(X = 1) = 1 - P(X = 0)$$

3. Distribuição binomial com parâmetros n e p , $X \sim b(n, p)$.

Conta o número de sucesso em n realizações independentes de variáveis de Bernoulli de parâmetro p .

$$X : \Omega \rightarrow \{0, 1, \dots, n\} \text{ e } P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}$$

4. Distribuição geométrica com parâmetro p , $X \sim \text{Geom}(n, p)$.

Em uma repetição independente de sucessivas variáveis aleatórias de Bernoulli de parâmetro p , X indica em qual tentativa obtemos o primeiro sucesso.

$$X : \Omega \rightarrow \mathbb{N} \text{ e } P(X = n) = (1 - p)^{n-1} p$$

5. Distribuição uniforme em $X(\Omega)$.

$$P(X = z) = \frac{1}{\#X(\Omega)}, \forall z \in X(\Omega)$$

- Variáveis aleatórias absolutamente contínuas.

1. Distribuição exponencial com parâmetro λ , $X \sim \exp(\lambda)$.

Conta o tempo decorrido até a primeira ocorrência do processo de Poisson com intensidade λ .

$$F_X(z) = \begin{cases} 1 - e^{-\lambda z}, & z > 0 \\ 0, & z \leq 0 \end{cases} \text{ e } f_X(z) = \begin{cases} \lambda e^{-\lambda z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

2. Distribuição uniforme no intervalo $[a, b]$, $X \sim \mathcal{U}[a, b]$.

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

3. Distribuição normal com parâmetros μ e σ^2 , $X \sim \mathcal{N}(\mu, \sigma^2)$.

$$f_X(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right]$$

Exercício 5.1.

1. X é variável aleatória se, e somente se, $X^{-1}(I) \in \mathcal{F}$, para todo I intervalo.

Aula 6 - Distribuição de uma variável aleatória e vetor aleatório

Exemplo 6.1. Seja $([1, 4], \Lambda, P)$ espaço de probabilidade, onde Λ é a σ -álgebra de Lebesgue e P é a medida de Lebesgue (muitas vezes denotada por λ). Considere as variáveis aleatórias

$$\begin{array}{ccc} Y : [1, 4] & \rightarrow & \mathbb{R} \\ \omega & \mapsto & \pi \end{array} \quad \text{e} \quad \begin{array}{ccc} Z : [1, 4] & \rightarrow & \mathbb{R} \\ \omega & \mapsto & \omega \end{array}$$

Seja $X : [1, 4] \rightarrow \mathbb{R}$ a variável aleatória definida como $X(\omega) = \max\{Y(\omega), Z(\omega)\}$. Então

$$(X \leq z) = \begin{cases} \emptyset, & z < \pi \\ [1, \pi], & z = \pi \\ [\pi, z], & z \in (\pi, 4) \\ [1, 4], & z \geq 4 \end{cases} \Rightarrow F_X(z) = P(X \leq z) = \begin{cases} 0, & z < \pi \\ \frac{\pi-1}{3}, & z = \pi \\ \frac{z-1}{3}, & z \in (\pi, 4) \\ 1, & z \geq 4 \end{cases}$$

Daí, X não é variável aleatória discreta, nem absolutamente contínua.

Definição 6.1. Seja $F(z)$ uma função de distribuição.

- Definimos a **função parte discreta de $F(z)$** , como sendo $F_d(z) = \sum_{x \in I} F(x) - F(x^-)$, onde $I = \{x \in \mathbb{R} \mid x \leq z \text{ e } x \text{ é ponto de descontinuidade de } F\}$ é o conjunto dos elementos.
- Definimos a **função parte absolutamente contínua de $F(z)$** , como sendo $F_{ac}(z) = \int_{-\infty}^z f dt$.
- Definimos a **função parte singular de $F(z)$** , como sendo $F_s(z) = F(x) - F_d(x) - F_{ac}(x)$.

Teorema 6.1. Se X é variável aleatória, então $(X \in B) \in \mathcal{F}$, $\forall B \in \mathcal{B}(\mathbb{R})$. Além disso, se B é intervalo, podemos calcular $P(X \in B)$ conhecendo F_X , a função de distribuição de X .

Definição 6.2. Definimos a **distribuição da variável aleatória X** , como sendo a medida P_X , definida em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ da seguinte maneira $P_X(B) = P(X \in B) = P(\{\omega \in \Omega \mid X(\omega) \in B\}) = P(X^{-1}(B))$.

Definição 6.3. Seja (Ω, \mathcal{F}, P) espaço de probabilidade. Dizemos que $(X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ é um **vetor aleatório**, se $X_i : \Omega \rightarrow \mathbb{R}$ for variável aleatória, para todo $i = 1, \dots, n$.

Definição 6.4. Seja $X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ vetor aleatório. Definimos a **função de distribuição do vetor X** , como $F_X : \mathbb{R}^n \rightarrow [0, 1]$, $(z_1, \dots, z_n) \mapsto P(\cap_{i=1}^n (X_i \leq z_i))$.

Teorema 6.2.

1. F_X é não decrescente em cada uma das variáveis, isto é, para todo $j \in \{1, \dots, n\}$ e para todo $z = (z_1, \dots, z_n) \in \mathbb{R}^n$, se $z_j \leq z'_j$, então $F(z_1, \dots, z_j, \dots, z_n) \leq F(z_1, \dots, z'_j, \dots, z_n)$.
2. F é contínua à direita em cada uma de suas variáveis.
3. (a) $\lim_{z_i \rightarrow -\infty} F(z_1, \dots, z_n) = 0$, para cada $i \in \{1, \dots, n\}$.
(b) $\lim_{(z_1, \dots, z_n) \rightarrow (\infty, \dots, \infty)} F(z_1, \dots, z_n) = 1$.

Observação 6.1. Nem toda função que satisfaz as propriedades do teorema 6.2 é função de distribuição de algum vetor aleatório. Contudo, se uma função F , além de satisfazer tais propriedades, também satisfizer a condição

$$\Delta_{1,I_1} \Delta_{2,I_2} \dots \Delta_{n,I_n} F \geq 0$$

para todo $I_i = (a_i, b_i]$, com $i = 1, \dots, n$, onde $\Delta_{k,I_k} F(x_1, \dots, x_n) = F(x_1, \dots, b_k, \dots, x_n) - F(x_1, \dots, a_k, \dots, x_n)$, então existe um vetor aleatório tal que F é função de distribuição.

Aula 7 - Distribuição de funções de vetor aleatório

Definição 7.1. Sejam $X = (X_1, \dots, X_n)$ vetor aleatório e $F_X(z_1, \dots, z_n)$ sua função de distribuição. Definimos a **distribuição marginal de X_i** como sendo a função $F_{X_i} : \mathbb{R} \rightarrow [0, 1]$, $z \mapsto \lim_{k \rightarrow \infty} F(k, \dots, k, z, k, \dots, k)$.

Definição 7.2. Dizemos que uma sequência de variáveis aleatórias, (X_n) , é **(coletivamente) independentes**, se $\forall n \in \mathbb{N}$ e $\forall B_1, B_2, \dots, B_n \in \mathcal{B}(\mathbb{R})$, vale que $P(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^n P(X_i \in B_i)$.

Teorema 7.1 (Critério de independência).

- a) Sejam X_1, \dots, X_n variáveis aleatórias independentes. Então a função distribuição de $X = (X_1, \dots, X_n)$ é $F_X(z_1, \dots, z_n) = \prod_{i=1}^n F_{X_i}(z_i)$.
- b) Se $F_X(z_1, \dots, z_n) = \prod_{i=1}^n g_i(z_i)$ e $\lim_{z_i \rightarrow \infty} g_i(z_i) = 1$, $\forall 1, \dots, n$, então $g_i = F_{X_i}$, $\forall i$ e X_1, \dots, X_n são independentes.

Corolário 7.1.

- a) Se X_1, \dots, X_n são variáveis aleatórias independentes, com densidades f_{X_i} , então a densidade do vetor $X = (X_1, \dots, X_n)$ é $f_X(z_1, \dots, z_n) = \prod_{i=1}^n f_{X_i}(z_i)$.

b) Se $X = (X_1, \dots, X_n)$ é absolutamente contínua e sua função densidade é

$$f_X(z_1, \dots, z_n) = \prod_{i=1}^n g_i(z_i)$$

com $\int_{-\infty}^{\infty} g_i(z_i) = 1$, então X_i é absolutamente contínua, com densidade $f_{X_i} = g_i$ e X_1, \dots, X_n são independentes.

Exemplo 7.1. Dizemos que o vetor (X_1, X_2) tem **distribuição normal bivariada com parâmetros $\rho, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2$** , se possui densidade conjunta da forma

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1} \right)^2 + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) \right] \right\}$$

Aula 8 - Esperança de uma variável aleatória

Definição 8.1. Sejam $g, F : \mathbb{R} \rightarrow \mathbb{R}$, onde g é contínua e F é monótona. Definimos a **integral de Riemann-Stieltjes** como sendo

$$\int_a^b g dF = \lim_{\|\pi\| \rightarrow 0} \sum_i g(y_i) (F(x_i) - F(x_{i-1}))$$

se tal limite existir e independer dos pontos y_i e da partição do intervalo $[a, b]$, $\pi = \{x_i\}$.

Proposição 8.1.

1. Se F for discreta, então $\int_a^b g dF = \sum_i g(x_i) (F(x_i) - F(x_i^-))$, onde F são os pontos de descontinuidade de F
2. Se F for absolutamente contínua, então $\int_a^b g dF = \int_a^b g(x) \varphi(x) dx$, onde φ é tal que $F(x) = \int_{-\infty}^x \varphi(t) dt$.

Definição 8.2. Seja $X : \Omega \rightarrow \mathbb{R}$ uma variável aleatória. Definimos a **esperança de X** , como $E(X) = \int_{\mathbb{R}} x dF_X(x)$, onde F_X é a função de distribuição de X .

Observação 8.1.

$$EX = \int_{\mathbb{R}} x dF_X(x) = \int_{\Omega} X dP = \int_{\mathbb{R}} x dP_X(x)$$

onde as últimas duas integrais são integrais de Lebesgue.

Definição 8.3. Definimos $X^+ = \max(X, 0)$, $X^- = \max(-X, 0)$. Assim, $X = X^+ - X^-$ e $|X| = X^+ + X^-$.

Definição 8.4. Dizemos que X é **integrável** se $EX^+ < \infty$ e $EX^- < \infty$. Neste caso, $EX = EX^+ - EX^-$. Se $EX^+ = \infty$ e $EX^- < \infty$, definimos $EX = \infty$; se $EX^+ < \infty$ e $EX^- = \infty$, definimos $EX = -\infty$; se $EX^+ = EX^- = \infty$, então EX não está bem definida.

Proposição 8.2.

$$EX = \int_{-\infty}^{\infty} x dF_X(x) = \int_0^{\infty} (1 - F(x)) dx - \int_{-\infty}^0 F(x) dx.$$

Corolário 8.1. Se X é não negativo, então $EX = \int_0^{\infty} P(X > x) dx$. Em particular, se $X : \Omega \rightarrow \mathbb{N}$, então $EX = \sum_{n=1}^{\infty} P(X \geq n)$.

Exercício 8.1.

1. Prove que

- (a) $\int_a^b g dF = \int_a^c g dF + \int_c^b g dF$;
- (b) $\int_a^b (\alpha g + \beta h) dF = \alpha \int_a^b g dF + \beta \int_a^b h dF$;
- (c) $\int_a^b g d(\alpha F + \beta H) = \alpha \int_a^b g dF + \beta \int_a^b g dH$.

2. Calcule a esperança de

- (a) $X \sim \text{Ber}(p)$
- (b) $X \sim b(n, p)$
- (c) $X \sim \text{Geom}(n, p)$
- (d) $X \sim \pi(\lambda)$
- (e) $X \sim \mathcal{U}\{x_1, \dots, x_n\}$
- (f) $X \sim \exp(\lambda)$
- (g) $X \sim \mathcal{U}[a, b]$
- (h) $X \sim \mathcal{N}(\mu, \sigma^2)$
- (i) $X \sim \max(\mathcal{U}[0, 2], 1)$

Aula 9 - Propriedades de esperança e variância

Definição 9.1. Definimos o espaço $\mathcal{L}^p(\Omega, \mathcal{F}, P)$ como o conjunto das variáveis aleatórias X tais que $E|X|^p < \infty$.

Proposição 9.1 (Propriedades da esperança).

- 1. Se $X = c$ qc, onde c é constante, então $EX = c$.
- 2. Se $X \leq Y$ qc, então $EX \leq EY$.
- 3. (Linearidade) $E(aX + bY) = aEX + bEY$, $\forall a, b \in \mathbb{R}$.

4. (Desigualdade de Jensen) Seja $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ convexa, então $E\varphi(X) \geq \varphi(EX)$.
5. (Desigualdade de Lyapounov) Seja $X \geq 0$ variável aleatória, então $\varphi(t) = (EX^t)^{\frac{1}{t}}$ é não decrescente, isto é, $(EX^t)^{\frac{1}{t}} \geq (EX^s)^{\frac{1}{s}}, \forall t \geq s$.
6. (Critério de integrabilidade) X é integrável se, e somente se, $\sum_{n=1}^{\infty} P(|X| \geq n) < \infty$
7. Se $Y \geq 0$ é integrável e $0 \leq |X| \leq Y$, então X é integrável. Em particular, se $|X| \leq c$, com c constante, então X é integrável.

Observação 9.1. Se X é variável aleatória e $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ é uma função mensurável, então $E\varphi(X) = \int_{\mathbb{R}} \varphi(x) dF_X(x) = \int_{\mathbb{R}} \varphi(x) dF_{\varphi(X)}(x)$.

Definição 9.2. Sejam X uma variável aleatória, $b \in \mathbb{R}$ e $k \in \mathbb{N}$. Definimos o **k-ésimo momento de X em torno de b** como $E|X - b|^k$. Quando $b = EX$, dizemos **k-ésimo momento central**.

Definição 9.3. Definimos a **variância** de uma variável aleatória X como $E(X - EX)^2$, isto é, é o 2.º momento central. Notação: VX ou $\text{Var } X$.

Proposição 9.2 (Propriedades da variância).

1. Se $X = c$ qc, onde c é constante, então $VX = 0$.
2. $V(aX + b) = a^2VX$.

Aula 10 - Covariância e correlação

Teorema 10.1 (Desigualdade de Chebyshev). Sejam $X : \Omega \rightarrow \mathbb{R}^+$ variável aleatória e $\lambda \geq 0$. Então $P(X > \lambda) \leq \frac{EX}{\lambda}$.

Corolário 10.1.

- a) Qualquer que seja a variável aleatória X , temos que $P(|X - EX| > \lambda) \leq \frac{VX}{\lambda^2}$.
- b) Se $X \geq 0$ e $EX = 0$, então $X = 0$ qc. Em particular, se $VX = 0$, então $X = EX$ qc.

Observação 10.1. O mínimo da função $f(b) = E(X - b)^2$ é atingido quando $b = EX$ e o mínimo da função $f(b) = E|X - b|$ é atingido quando $b = \sup_{\lambda} (P(X \leq \lambda) \leq \frac{1}{2})$, isto é, quando b é a mediana de X .

Definição 10.1. Sejam $X, Y : \Omega \rightarrow \mathbb{R}$ variáveis aleatórias. Definimos a **covariância entre X e Y** como sendo $\text{cov}(X, Y) = E((X - EX)(Y - EY))$. Note que, $\text{cov}(X, Y) = E(XY) - EXEY$.

Teorema 10.2. Sejam X_1, \dots, X_n variáveis aleatórias. Então $V(X_1 + \dots + X_n) = \sum_{i=1}^n V X_i + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j)$.

Observação 10.2.

- i) $\text{cov}(X, X) = V X$.
- ii) $\text{cov}(X, Y) = \text{cov}(Y, X)$.
- iii) $\text{cov}(\cdot, \cdot)$ é bilinear.
- iv) Se X, Y são independentes, então $\text{cov}(X, Y) = 0$.

Definição 10.2. Sejam X, Y variáveis aleatórias. Definimos o **coeficiente de correlação entre X e Y** como $\rho = \frac{\text{cov}(X, Y)}{\sqrt{V X \cdot V Y}}$. Quando $\rho = 0$, dizemos que X e Y são **não correlacionadas**.

Teorema 10.3. Sejam X, Y variáveis aleatórias.

- a) $\rho(X, Y) \in [-1, 1]$.
- b) $\rho(X, Y) = 1 \Leftrightarrow Y = aX + b$ qc, com $b \in \mathbb{R}$ e $a > 0$.
- c) $\rho(X, Y) = -1 \Leftrightarrow Y = aX + b$ qc, com $b \in \mathbb{R}$ e $a < 0$.

Aula 11 - Prova 1

Foi realizada a primeira prova.

Aula 12 - Teoremas de convergência

Exemplo 12.1.

- Sejam X e (X_n) variáveis aleatórias, onde X tem distribuição de Cauchy e $X_n = X \mathbb{1}_{(|X| \leq n)}$. Então $X_n \rightarrow X$ qc. Note que X_n é integrável, pois é limitada e $EX_n = 0$. Contudo, $EX_n \not\rightarrow EX$, já que EX não está bem definida.
- Sejam $Z \sim \mathcal{U}[0, 1]$ e $X_n = n \mathbb{1}_{(0 \leq Z \leq \frac{1}{n})}$. Então $X_n \rightarrow 0$ qc, porém $EX_n = 1 \neq E0 = 0$.

Teorema 12.1 (Convergência monótona). Sejam X_n, X variáveis aleatórias tais que $X_n \geq 0$, para todo n , e $X_n \uparrow X$. Então $\lim EX_n = EX$.

Teorema 12.2 (Convergência dominada). Sejam $X_n, X, Y : \Omega \rightarrow \mathbb{R}$ variáveis aleatórias tais que $X_n \rightarrow X$ qc, $Y \geq 0$ é integrável e $|X_n| \leq Y$, para todo n . Então $EX_n \rightarrow EX$.

Aula 13 - Modos de convergência e Lei Fraca dos Grandes Números

Definição 13.1. Sejam $(X_n), X : \Omega \rightarrow \mathbb{R}$ variáveis aleatórias.

- a) Dizemos que $\mathbf{X_n}$ converge para \mathbf{X} quase certamente, $X_n \xrightarrow{qc} X$, se $P(\{\omega \in \Omega \mid X_n(\omega) \rightarrow X(\omega)\}) = 1$.
- b) Dizemos que $\mathbf{X_n}$ converge para \mathbf{X} em \mathcal{L}^p , $X_n \xrightarrow{\mathcal{L}^p} X$, se $\lim_{n \rightarrow \infty} E|X_n - X|^p \rightarrow 0$.
- c) Dizemos que $\mathbf{X_n}$ converge para \mathbf{X} em probabilidade, $X_n \xrightarrow{P} X$, se $\forall \varepsilon > 0$, $\lim_{n \rightarrow \infty} P(\{\omega \in \Omega \mid |X_n(\omega) - X(\omega)| \geq \varepsilon\}) = 0$.

Teorema 13.1.

- a) Convergência em $\mathcal{L}^p \Rightarrow$ convergência em P .
- b) Convergência qc \Rightarrow convergência em P .

Exemplo 13.1.

- 1. Convergência qc \nRightarrow convergência em \mathcal{L}^p : sejam $X \sim \mathcal{U}[0, 1]$ e $Y_n = n\mathbb{1}_{(0 \leq X \leq \frac{1}{n})}$. Então $Y_n \rightarrow 0$ qc, porém $E|Y_n - 0| = EY_n = n \frac{1}{n} = 1 \nrightarrow EX = 0$.
- 2. Convergência $\mathcal{L}^p \nRightarrow$ convergência em qc: sejam $A_1 = [0, 1]$, $A_2 = [0, \frac{1}{2}]$, $A_3 = [\frac{1}{2}, 1]$, $A_4 = [0, \frac{1}{3}]$, \dots no espaço $([0, 1], \mathcal{B}([0, 1]), \lambda)$. Tome $X_n = \mathbb{1}_{A_n}$. Então $X_n \xrightarrow{\mathcal{L}^p} 0$, porém $X_n \nrightarrow^{qc} 0$.

Definição 13.2. Sejam (X_n) e $S_n = \sum_{i=1}^n X_i$ variáveis aleatórias, onde $X_n : \Omega \rightarrow \mathbb{R}$. Dizemos que vale uma **Lei Fraca dos Grandes Números para $\mathbf{X_n}$** se $\frac{S_n - ES_n}{n} \xrightarrow{P} 0$. Se vale a convergência qc, dizemos que vale uma **Lei Forte dos Grandes Números para $\mathbf{X_n}$**

Teorema 13.2 (Lei Fraca dos Grandes Números de Chebyshev). Sejam (X_n) sequência de variáveis aleatórias 2 a 2 não correlacionadas e com variância uniformemente limitada (isto é, existe $c > 0$ tal que $VX_n \leq c$, para todo $n \in \mathbb{N}$). Então $\frac{S_n - ES_n}{n} \rightarrow 0$ em probabilidade.

Corolário 13.1.

$$\frac{b(n, p)}{n} \xrightarrow{P} p.$$

Exemplo 13.2.

1. (Jogador compulsivo). Sejam $p = \frac{60}{67}$ a probabilidade de ganhar o jogo, $m = 5$ reais o preço de jogar e $M + m = 10^7$ reais o prêmio. Seja $X_n : \Omega \rightarrow \{-m, M\}$ a variável aleatória tal que $X_n(\omega) = -m$, se $\omega \in \{\text{perde}\}$ e $X_n(\omega) = M$, se $\omega \in \{\text{ganha}\}$. Então $\frac{S_n - \sum_{i=1}^n EX_i}{n} \rightarrow 0$ em probabilidade. Como $EX_n = Mp - m(1-p)$, para todo n , então $\frac{S_n}{n} \xrightarrow{P} Mp - m(1-p) < 0$. Logo, $S_n \xrightarrow{P} -\infty$, ou seja, o capital acumulado do jogador tende a menos infinito.
2. (Monte Carlo). Seja $f : [0, 1] \rightarrow [0, 1]$ um função. Sejam $\xi_1, \eta_1, \xi_2, \eta_2, \dots$ variáveis aleatórias com distribuição $\mathcal{U}[0, 1]$. Defina

$$X_n = \begin{cases} 1, & \eta_n \leq f(\xi_n) \\ 0, & \text{caso contrário} \end{cases}$$

A LGN nos diz que $\frac{S_n}{n} \rightarrow P(\eta_1 \leq f(\xi_1)) = \int_0^1 f(x) dx$.

Aula 14 - Lei 0-1 de Kolmogorov e passeio aleatório

Definição 14.1. Sejam (ξ_n) variáveis aleatórias independentes. Considere a menor σ -álgebra que torna ξ_i mensurável, para todo $i \geq n$, $\mathcal{F}_n^\infty = \sigma(\xi_n, \xi_{n+1}, \dots)$. Definimos a **σ -álgebra caudal** como sendo $\chi = \bigcap_{n \in \mathbb{N}} \mathcal{F}_n^\infty$. Se $A \in \chi$, dizemos que ele é um **evento caudal**.

Exemplo 14.1.

1. $\{\omega \in \Omega \mid \sum \xi_n(\omega) < \infty\} \in \chi$.
2. $\{\omega \in \Omega \mid \exists \lim \xi_n(\omega)\} \in \chi$, porém $\{\omega \in \Omega \mid \exists \lim (\xi_1(\omega) + \dots + \xi_n(\omega)) \geq c\} \notin \chi$.

Lema 14.1 (Teorema da aproximação). Seja (Ω, \mathcal{A}, P) espaço de probabilidade e \mathcal{F} um álgebra de Ω tal que $\sigma(\mathcal{F}) = \mathcal{A}$. Então $\forall A \in \mathcal{A}$ e $\forall \varepsilon > 0$ existe $F \in \mathcal{F}$ tal que $P(A \Delta F) < \varepsilon$.

Teorema 14.1 (Lei 0 – 1 de Kolmogorov). Sejam (ξ_n) variáveis aleatórias independentes e A um evento caudal, isto é, $A \in \chi$. Então $P(A) = 0$ ou $P(A) = 1$.

Corolário 14.1. Se X é χ -mensurável, ou seja, $(X \in B) \in \chi$, para todo $B \in \mathcal{B}(\mathbb{R})$, então X é degenerada, isto é, $\exists c \in \mathbb{R}$ tal que $P(X = c) = 1$.

Corolário 14.2 (Lei 0 – 1 de Borel). Seja (A_n) sequência de eventos independentes. Então $P(\limsup A_n) = 0$ ou $P(\limsup A_n) = 1$.

Definição 14.2. Sejam (ξ_n) variáveis aleatórias independentes e identicamente distribuídas, onde $P(\xi_i = 1) = p = 1 - P(\xi_i = -1)$, para todo i . O processo estocástico (S_n) , onde $S_n = \sum_{i=1}^n \xi_i$ é dito **passeio aleatório em \mathbb{Z}** . Dizemos que o passeio aleatório (S_n) é **recorrente** se $P(S_n = 0, i.v.) = 1$. Caso contrário, dizemos que (S_n) é **transiente**.

Teorema 14.2. Seja (S_n) o passeio aleatório em \mathbb{Z} .

- a) Se $p = 1/2$, então $P(S_n = 0, i.v.) = 1$.
- b) Se $p \neq 1/2$, então $P(S_n = 0, i.v.) = 0$.

Aula 15 - Lei 0-1 de Hewitt-Savage e teorema da uma série

Definição 15.1. Dizemos que $\pi : \mathbb{N} \rightarrow \mathbb{N}$ é uma **permutação finita**, se π é bijeção e $\pi_n \neq n$ no máximo finita vezes.

Definição 15.2. Sejam $B \in \mathcal{B}(\mathbb{R}^\infty)$, $A = \{\omega \in \Omega \mid (\xi_1, \dots, \xi_n, \dots) \in B\}$ e $\pi : \mathbb{N} \rightarrow \mathbb{N}$ permutação finita. Defina $\pi(A) = \{\omega \in \Omega \mid (\xi_{\pi(1)}, \dots, \xi_{\pi(n)}, \dots) \in B\}$. Dizemos que A é **simétrico com respeito a ξ** , se $A = \pi(A)$, para toda permutação finita π .

Teorema 15.1 (Lei 0 – 1 de Hewitt-Savage). Sejam (ξ_n) variáveis aleatórias independentes e identicamente distribuídas e A evento simétrico. Então $P(A) = 0$ ou $P(A) = 1$.

Exemplo 15.1. Se ξ_1 é variável aleatória que sorteia 0 ou 1 e $\xi_n = \xi_1, \forall n \geq 1$, então $\sum \xi_n = 0$ ou $\sum \xi_n = \infty$, logo $P(\sum \xi_n < \infty) = \frac{1}{2}$.

Teorema 15.2 (Desigualdade maximal de Kolmogorov). Sejam (ξ_n) variáveis aleatórias independentes com $E\xi_n = 0$ e $E\xi_n^2 < \infty$, para todo $n \in \mathbb{N}$, e $S_n = \sum_{k=1}^n \xi_k$. Então, para todo ε ,

- a) $P(\max_{k \leq n} |S_k| \geq \varepsilon) \leq \frac{\sum_{k=1}^n E\xi_k^2}{\varepsilon^2}$.
- b) Se (ξ_n) for uniformemente limitada, isto é, existe $c > 0$ tal que $P(|\xi_n| \leq c) = 1, \forall n$, então $P(\max_{k \leq n} |S_k| \geq \varepsilon) \geq 1 - \frac{(c + \varepsilon)^2}{\sum_{k=1}^n E\xi_k^2}$.

Teorema 15.3 (Teorema da uma série). Sejam (ξ_n) variáveis aleatórias independentes, com $E\xi_n = 0$ e $E\xi_n^2 < \infty, \forall n \in \mathbb{N}$. Então,

- a) se $\sum_{n=1}^{\infty} E\xi_n^2 < \infty$, então $\sum_{n=1}^{\infty} \xi_n < \infty$ qc.
- b) Se (ξ_n) for uniformemente limitada e $\sum_{n=1}^{\infty} E\xi_n^2 < \infty$, então $\sum_{n=1}^{\infty} \xi_n < \infty$ qc.

Exemplo 15.2. Sejam (ξ_n) variáveis aleatórias independentes e identicamente distribuídas com $P(\xi_n = 1) = P(\xi_n = -1) = \frac{1}{2}$ e $(a_n) \in \mathbb{R}$ tais que $|a_n| \leq c$, então $\sum \xi_n a_n < \infty$ qc $\Leftrightarrow \sum a_n^2 < \infty$.

Exercício 15.1.

1. Se (ξ_n) são independentes e identicamente distribuídas, então todo evento caudal é simétrico.

Aula 16 - Aula 16 - Teorema das 2 e 3 séries e lei forte dos grandes números

Teorema 16.1 (Teorema das duas séries). Seja (ξ_n) variáveis aleatórias independentes, então

- a) $\sum V\xi_n < \infty$ e $\sum E\xi_n < \infty \Rightarrow \sum \xi_n < \infty$ qc.
- b) Se (ξ_n) é uniformemente limitada e $\sum \xi_n < \infty$ qc, então $\sum E\xi_n, \sum V\xi_n < \infty$.

Definição 16.1. Sejam ξ variável aleatória e $c \in \mathbb{R}$. Então $\xi^c = \xi \mathbb{1}_{(|\xi| \leq c)}$.

Teorema 16.2 (Teorema das três séries). Seja (ξ_n) variáveis aleatórias independentes.

- a) Se $\exists c > 0$ tal que $\sum V\xi_n^c, \sum E\xi_n^c, \sum P(|\xi_n| > c) < \infty$, então $\sum \xi_n < \infty$.
- b) Se $\sum \xi_n < \infty$ qc, então $\sum V\xi_n^c, \sum E\xi_n^c, \sum P(|\xi_n| > c) < \infty, \forall c > 0$.

Teorema 16.3 (Lei forte dos grandes números de Cantelli). Sejam (ξ_n) variáveis aleatórias independentes e identicamente distribuídas, com $E\xi^4 < \infty$. Então $\frac{S_n - ES_n}{n} \rightarrow 0$ qc, onde $S_n = \sum_{j=1}^n \xi_j$.

Lema 16.1 (Lema de Kronecker). Sejam $(b_n) > 0$ sequência com $b_n \rightarrow \infty$ e (a_n) tal que $\sum_{n=1}^{\infty} a_n = a$. Então $\frac{1}{b_n} \sum_{k=1}^n b_k a_k \xrightarrow{n \rightarrow \infty} 0$.

Teorema 16.4 (1.^ª Lei forte de Kolmogorov). Sejam (ξ_n) variáveis aleatórias independentes e $(b_n) > 0$ sequência com $b_n \rightarrow \infty$. Se $\sum \frac{V\xi_n}{b_n^2} < \infty$, então $\frac{S_n - ES_n}{b_n} \rightarrow 0$ qc.

Aula 17 - 2.^a lei forte de Kolmogorov

Teorema 17.1 (2.^a Lei forte de Kolmogorov). Sejam (ξ_n) variáveis aleatórias independentes e identicamente distribuídas, com $E|\xi_1| < \infty$. Então $\frac{S_n}{n} \rightarrow E\xi_1$ qc.

Observação 17.1.

1. Existe um resultado mais forte, chamado de **lei forte dos grandes números de Etemadi**: seja (ξ_n) sequência identicamente distribuída, com $E|\xi_1| < \infty$ e 2 a 2 independentes. Então $\frac{S_n}{n} \rightarrow E\xi_1$.
2. Vale a recíproca para a 2.^a lei forte de Kolmogorov, isto é, sejam (ξ_n) variáveis aleatórias independentes e identicamente distribuídas, com $\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow c < \infty$ qc. Então ξ_n é integrável e $E\xi_n = c, \forall n$.
3. A 2.^a lei forte de Kolmogorov ainda é válida nos casos onde ξ é não integrável, mas é bem definida, isto é, $E\xi = \infty$ ou $E\xi = -\infty$.

Aula 18 - Funções características

Definição 18.1. Dizemos que $Z : \Omega \rightarrow \mathbb{C}$ é **variável aleatória**, se $\operatorname{Re} Z$ e $\operatorname{Im} Z$ são variáveis aleatórias reais. Neste caso, $EZ = E\operatorname{Re} Z + iE\operatorname{Im} Z$.

Definição 18.2. Seja X uma variável aleatória (ou F_X uma função de distribuição). Definimos a **função característica de X** (ou de F_X), como $\varphi_X : \mathbb{R} \rightarrow \mathbb{C}, t \mapsto Ee^{itX} = \int_{\Omega} e^{itX} dP = \int_{\mathbb{R}} e^{itx} dF_X(x)$.

Observação 18.1. Se X for absolutamente contínua com densidade $f(x)$, então $\varphi_X(t) = \int_{\mathbb{R}} e^{itx} f(x) dx$.

Proposição 18.1. Propriedades das funções características.

1. $|\varphi_X(t)| \leq 1$.
2. $\varphi_X(0) = 1$.
3. $\varphi_X(-t) = \overline{\varphi_X(t)}$.
4. $\varphi_X(t)$ é uniformemente contínua.
5. Se X e Y são independentes, então $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$. Mais geralmente, se X_1, \dots, X_n são coletivamente independentes, então $\varphi_{\sum_i X_i}(t) = \prod_i \varphi_{X_i}(t)$.
6. Sejam $a, b \in \mathbb{R}$. Então $\varphi_{aX+b}(t) = e^{itb} \varphi_X(at)$.

7. X é simétrica em torno do zero (isto é, $P(X \leq x) = P(X \geq -x)$) $\Leftrightarrow \varphi_X(t)$ é real.
8. Se $E|X|^n < \infty$, então $\varphi_X(t)$ é de classe C^n . Além disso, $\forall k \leq n$, temos que $\varphi_X^{(k)}(0) = i^k EX^k$.
9. (Fórmula de inversão). Seja $\varphi(t)$ função característica da distribuição F . Então

$$\frac{F(y) + F(y^-)}{2} - \frac{F(x) + F(x^-)}{2} = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{e^{-itx} - e^{-ity}}{it} \varphi(t) dt.$$

Corolário (unicidade): a função característica determina de modo único a função de distribuição associada.

$$F(z) = \lim_{y \rightarrow z^+} \lim_{x \rightarrow -\infty} \frac{1}{2\pi} \int_{\mathbb{R}} \frac{e^{-itx} - e^{-ity}}{it} \varphi(t) dt$$

Aula 19 - Convergência em distribuição

Definição 19.1. Sejam $(X_n), X$ variáveis aleatórias e $(F_n), F$ suas respectivas funções de distribuição. Dizemos que **X_n converge para X em distribuição**, $X_n \xrightarrow{d} X$, se $F_n(x) \rightarrow F(x)$, para todo x ponto de continuidade de F .

Observação 19.1.

1. (X_n) não precisam estar definidas no mesmo espaço de probabilidade.
2. Se $X_n = \frac{1}{n}$ e $X = 0$, então $X_n \rightarrow X$ qc, \mathcal{L}^p e P , porém $F_{X_n}(0) \not\rightarrow F_X(0) = 1$.

Teorema 19.1. Se $X_n \xrightarrow{P} X$, então $X_n \xrightarrow{d} X$.

Observação 19.2. A recíproca é falsa. Contraexemplo: Considere as variáveis aleatórias $X, X_1, X_2, \dots, X_n, \dots \sim \mathcal{N}(0, 1)$. Então $X_n \xrightarrow{d} X$, pois todas têm a mesma distribuição, porém $P(|X_n - X| \geq \varepsilon) = c > 0$ constante que não depende de n , pois $X_n - X \sim \mathcal{N}(0, 2)$.

Teorema 19.2. Se $X_n \xrightarrow{d} X$ e X é variável aleatória degenerada (isto é, $X = c$ qc), então $X_n \xrightarrow{P} X$.

Teorema 19.3 (Helly-Bray). Se $X_n \xrightarrow{d} X$, então $Eg(X_n) \rightarrow Eg(X)$, para toda $g: \mathbb{R} \rightarrow \mathbb{R}$ contínua e limitada.

Aula 20 - Teorema de Paul Lévy

Teorema 20.1 (Teorema de Paul Lévy). Sejam (F_n) funções de distribuição e (φ_n) suas respectivas funções características. Se existe $\lim_{n \rightarrow \infty} \varphi_n(t) = \varphi(t)$, para todo $t \in \mathbb{R}$, e $\varphi(t)$ seja contínua em $t = 0$, então

- a) existe F função de distribuição tal que $F_n \xrightarrow{d} F$;
- b) φ é função característica de F .

Corolário 20.1. $X_n \xrightarrow{d} X \Leftrightarrow \varphi_{X_n}(t) \rightarrow \varphi_X(t), \forall t$.

Definição 20.1. Sejam X_n variáveis aleatórias, com $VX_n < \infty$ e $VX_k > 0$, para algum k . Dizemos que vale o **teorema central do limite para a sequência X_n** , se $\frac{S_n - ES_n}{\sqrt{VS_n}} \xrightarrow{d} \mathcal{N}(0, 1)$.

Observação 20.1. Aplicações do teorema de Paul Lévy.

1. (*Lei fraca dos grandes números*). Sejam (X_n) variáveis aleatórias independentes e identicamente distribuídas com $E|X| < \infty$. Então $\frac{S_n - ES_n}{n} \xrightarrow{P} 0$.
2. (*Teorema de Poisson*). $b\left(n, \frac{\lambda}{n}\right) \xrightarrow{d} \pi(\lambda)$.
3. (*Teorema central do limite*). Sejam (X_n) variáveis aleatórias independentes e identicamente distribuídas. Então $\frac{S_n - ES_n}{\sqrt{VS_n}} \xrightarrow{d} \mathcal{N}(0, 1)$.

Corolário (*teorema de de Moivre-Laplace*). Sejam (X_n) variáveis aleatórias independentes e identicamente distribuídas, com $X_n \sim \text{Ber}(p)$. Então

$$P\left(a < \frac{b(n, p) - np}{\sqrt{npq}} < b\right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

Aula 21 - Teorema de Lindeberg e teorema de Liapounov

Definição 21.1. Sejam (X_n) variáveis aleatórias tais que $VX_n < \infty$ e $VX_k > 0$, para algum k . Dizemos que (X_n) satisfaz a **condição de Lindeberg**, se $\forall \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^n \int_{|x - \mu_k| > \varepsilon s_n} (x - \mu_k)^2 dF_k(x) = 0$$

onde $\mu_k = EX_k$, $F_k = F_{X_k}$, $S_n = \sum_{i=1}^n X_i$ e $s_n^2 = VS_n$.

Teorema 21.1 (Teorema central do limite de Lindeberg). Sejam (X_n) variáveis aleatórias independentes. Se (X_n) satisfaz a condição de Lindeberg, então $\frac{S_n - ES_n}{\sqrt{VS_n}} \xrightarrow{d} \mathcal{N}(0, 1)$.

Teorema 21.2. Se vale a condição de Lindeberg, então $\frac{\max_{1 \leq k \leq n} \sigma_k^2}{s_n^2} \rightarrow 0$, onde $\sigma_k^2 = VX_k$.

Corolário 21.1. Vale o teorema central do limite para o caso onde as variáveis aleatórias são independentes e identicamente distribuídas.

Observação 21.1. Vale uma “quase” recíproca do teorema de Lindeberg, o chamado **teorema de Feller**: seja (X_n) variáveis aleatórias independentes; se vale $\lim_{n \rightarrow \infty} \max_{k \leq n} \frac{\sigma_k^2}{s_n^2} = 0$ e vale o teorema central do limite, então a condição de Lindeberg é satisfeita.

Definição 21.2. Sejam (X_n) variáveis aleatórias tais que $VX_n < \infty$ e $VX_k > 0$, para algum k . Dizemos que (X_n) satisfaz a **condição de Liapounov**, se $\exists \delta > 0$ tal que

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\delta}} \sum_{k=1}^n E|X_k - \mu_k|^{2+\delta} = 0$$

Teorema 21.3 (Teorema central do limite de Liapounov). Sejam (X_n) variáveis aleatórias independentes. Se (X_n) satisfaz a condição de Liapounov, então $\frac{S_n - ES_n}{\sqrt{VS_n}} \xrightarrow{d} \mathcal{N}(0, 1)$.

Aula 22 - Prova 2

Foi realizada a segunda prova.

Aula 23 - Teorema central do limite de Lindeberg

Teorema 23.1 (Teorema central do limite de Lindeberg). Sejam $(\xi_{n,k})$, $n \in \mathbb{N}$ e $k = 1, \dots, k_n$, variáveis aleatórias com $(\xi_{n,k})_k$ independentes, para todo $n \in \mathbb{N}$. Sejam $\mu_{n,k} = E\xi_{n,k}$, $\sigma_{n,k}^2 = V\xi_{n,k}$, $S_n = \sum_{k=1}^{k_n} \xi_{n,k}$ e $s_n^2 = VS_n$. Se para todo $\varepsilon > 0$ vale

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^{k_n} \int_{|\xi_{n,k} - \mu_{n,k}| > \varepsilon s_n} (\xi_{n,k} - \mu_{n,k})^2 dP = 0$$

então $\frac{S_n - ES_n}{\sqrt{VS_n}} \xrightarrow{d} \mathcal{N}(0, 1)$.

Observação 23.1. Pode valer a teorema central do limite, sem que a condição de Lindeberg seja satisfeita. Nesses casos, dizemos que vale o teorema central do limite para **condição não clássicas**. Além disso, pelo teorema de Feller, não vale que $\lim_{n \rightarrow \infty} \max_{k \leq n} \frac{\sigma_k^2}{s_n^2} = 0$.

Exemplo 23.1. Sejam (ξ_n) variáveis aleatórias com $\xi_n \stackrel{d}{\sim} \mathcal{N}(0, \sigma_n^2)$, onde $\sigma_1^2 = 1$ e $\sigma_n^2 = 2^{n-2}, \forall n \geq 2$. Então vale o teorema central do limite, porém $\lim_{n \rightarrow \infty} \max_{k \leq n} \frac{\sigma_k^2}{s_n^2} = \lim_{n \rightarrow \infty} \frac{\sigma_n^2}{s_n^2} = \lim_{n \rightarrow \infty} \frac{2^{n-2}}{2^{n-1}} = \frac{1}{2}$.

Aula 24 - Esperança condicional dada uma partição

Definição 24.1. Sejam A um evento e $\mathcal{D} = \{D_1, \dots, D_n, \dots\}$ uma partição de Ω . Definimos a **probabilidade condicional de A com respeito à partição \mathcal{D}** , como $P(A | \mathcal{D}) = \sum_i P(A | D_i) \mathbb{1}_{D_i}$.

Definição 24.2. Seja X variável aleatória discreta, isto é, $X = \sum x_i \mathbb{1}_{A_i}$, com x_i distintos. Note que $\mathcal{D}_X = \{A_1, A_2, \dots\}$ é uma partição de Ω . Essa partição é dita **partição induzida por X**. Notação: $P(B | X) = P(B | \mathcal{D}_X)$.

Definição 24.3. Sejam X variável aleatória discreta, isto é, $X = \sum x_i \mathbb{1}_{A_i}$, e $\mathcal{D} = (D_i)_i$ partição. Definimos a **esperança condicional de X dado \mathcal{D}** como $E(X | \mathcal{D}) = \sum_i x_i P(A_i | \mathcal{D})$.

Definição 24.4. Sejam D um evento e X uma variável aleatória. Definimos a **esperança condicional de X dado D** como $E(X | D) = \frac{1}{P(D)} E(X \mathbb{1}_D)$.

Observação 24.1. $E(X | \mathcal{D}) = \sum_j E(X | D_j) \mathbb{1}_{D_j}$.

$$\begin{array}{ccc}
 P(D) & \xrightarrow{\sum_i x_i P(A_i)} & E(X) \\
 \downarrow \frac{P(A \cap D)}{P(D)} & & \downarrow \frac{1}{P(D)} E(X \mathbb{1}_D) \\
 P(A | D) & \xrightarrow{\sum_i x_i P(A_i | D)} & E(X | D) \\
 \downarrow \sum_j P(A | D_j) \mathbb{1}_{D_j} & & \downarrow \sum_j E(X | D_j) \mathbb{1}_{D_j} \\
 P(A | \mathcal{D}) & \xrightarrow{\sum_i x_i P(A_i | \mathcal{D})} & E(X | \mathcal{D})
 \end{array}$$

Proposição 24.1. Propriedades de $P(A | \mathcal{D})$.

1. Se $A \cap B = \emptyset$, então $P(A \cup B \mid \mathcal{D}) = P(A \mid \mathcal{D}) + P(B \mid \mathcal{D})$.
2. $P(A \mid \{\Omega\}) = P(A)$.
3. $E(P(A \mid \mathcal{D})) = P(A)$.

Proposição 24.2. Propriedades de $E(X \mid \mathcal{D})$.

1. $E(aX + bY \mid \mathcal{D}) = aE(X \mid \mathcal{D}) + bE(Y \mid \mathcal{D})$.
2. $E(X \mid \{\Omega\}) = EX$.
3. Se $X = c$ qc, então $E(X \mid \mathcal{D}) = c$.
4. Se $X = \mathbb{1}_A$, então $E(X \mid \mathcal{D}) = P(A \mid \mathcal{D})$.

Definição 24.5. Sejam $\mathcal{D}_1 = \{D_{11}, D_{12}, \dots\}$ e $\mathcal{D}_2 = \{D_{21}, D_{22}, \dots\}$ partições de Ω . Dizemos que **\mathcal{D}_2 é mais fina (ou mais rica) do que \mathcal{D}_1** , se para todo j , existe i tal que $D_{2j} \subset D_{1i}$. Notação: $\mathcal{D}_1 \preceq \mathcal{D}_2$.

Definição 24.6. Dizemos que uma variável aleatória X é **mensurável com respeito a \mathcal{D}** , se $\mathcal{D}_X \preceq \mathcal{D}$.

Teorema 24.1. Sejam X e Y variáveis aleatórias discretas. Se X é \mathcal{D} -mensurável, então $E(XY \mid \mathcal{D}) = XE(Y \mid \mathcal{D})$.

Corolário 24.1. $E(X \mid X) = X$.

Teorema 24.2. Sejam \mathcal{D}_1 e \mathcal{D}_2 , com $\mathcal{D}_1 \preceq \mathcal{D}_2$, e X variável aleatória discreta. Então

- a) $E(E(X \mid \mathcal{D}_1) \mid \mathcal{D}_2) = E(X \mid \mathcal{D}_1)$.
- b) $E(E(X \mid \mathcal{D}_2) \mid \mathcal{D}_1) = E(X \mid \mathcal{D}_1)$.

Corolário 24.2. $E(E(X \mid \mathcal{D})) = E(X)$.

Aula 25 - Esperança condicional dada uma σ -álgebra

Definição 25.1. Sejam X variável aleatória e $\mathcal{D} = \{D_1, D_2, \dots\}$ partição. Dizemos que **X é independente de \mathcal{D}** , se X e $\mathbb{1}_{D_i}$ são independentes, para todo i .

Teorema 25.1. Se X é independente de \mathcal{D} , então $E(X \mid \mathcal{D}) = EX$.

Definição 25.2. Sejam (Ω, \mathcal{A}, P) espaço de probabilidade, $X : \Omega \rightarrow \mathbb{R}$ variável aleatória integrável e \mathcal{F} um sub- σ -álgebra de \mathcal{A} ($\mathcal{F} \preceq \mathcal{A}$).

- a) Se $X \geq 0$, definimos a **esperança condicional de X dada \mathcal{F}** , $E(X \mid \mathcal{F})$, como sendo uma variável aleatória satisfazendo

- i) $E(X | \mathcal{F})$ é \mathcal{F} -mensurável.
- ii) $\int_A E(X | \mathcal{F}) dP = \int_A X dP, \forall A \in \mathcal{F}$.

b) Se X é qualquer, definimos $E(X | \mathcal{F}) = E(X^+ | \mathcal{F}) + E(X^- | \mathcal{F})$.

Teorema 25.2. Se $X \geq 0$, então $E(X | \mathcal{F})$ existe e é única P -qc.

Definição 25.3. Dados $A \in \mathcal{A}$ e $\mathcal{F} \preccurlyeq \mathcal{A}$, definimos a **probabilidade condicional de A dada \mathcal{F}** , como $P(A | \mathcal{F}) = E(\mathbb{1}_A | \mathcal{F})$. Note que

- i) $P(A | \mathcal{F})$ é \mathcal{F} -mensurável.
- ii) $P(A \cap B) = \int_B P(A | \mathcal{F}) dP, \forall B \in \mathcal{F}$.

Teorema 25.3. Sejam $\mathcal{D} = (D_1, D_2, \dots)$ partição de Ω , com $D_i \in \mathcal{A}$, e X variável aleatória. Então $E(X | \mathcal{D}) = E(X | \sigma(\mathcal{D}))$.

Proposição 25.1. Propriedades de $E(X | \mathcal{F})$.

1. Se $X = c$ qc, então $E(X | \mathcal{F}) = c$ qc.
2. Se $X \leq Y$ qc, então $E(X | \mathcal{F}) \leq E(Y | \mathcal{F})$.
3. $|E(X | \mathcal{F})| \leq E(|X| | \mathcal{F})$.
4. Se $a, b \in \mathbb{R}$, então $E(aX + bY | \mathcal{F}) = aE(X | \mathcal{F}) + bE(Y | \mathcal{F})$.
5. Se $\mathcal{F} = \{\emptyset, \Omega\}$, então $E(X | \mathcal{F}) = EX$.
6. $E(X | \mathcal{A}) = X$.
7. $E(E(X | \mathcal{F})) = EX$.
8. Se $\mathcal{F}_1 \preccurlyeq \mathcal{F}_2$, então
 - (a) $E(E(X | \mathcal{F}_2) | \mathcal{F}_1) = E(X | \mathcal{F}_1)$.
 - (b) $E(E(X | \mathcal{F}_1) | \mathcal{F}_2) = E(X | \mathcal{F}_1)$.
9. Se X é independente de \mathcal{F} (isto é, X e $\mathbb{1}_A$ são independentes para todo $A \in \mathcal{F}$), então $E(X | \mathcal{F}) = EX$.
10. Se X é \mathcal{F} -mensurável, então $E(XY | \mathcal{F}) = XE(Y | \mathcal{F})$. Em particular, $E(X | X) = X$.

Exercício 25.1. Sejam ξ_n variáveis aleatórias independentes e identicamente distribuídas e $\tau : \Omega \rightarrow \mathbb{N}$ variável aleatória independente de (ξ_n) . Defina $S_n = \sum_{i=1}^n \xi_i$ e $S_\tau = \sum_n S_n \mathbb{1}_{\{\tau=n\}}$. Mostre que $E(S_\tau | \tau) = (E\xi_1)\tau$. Em particular, $E(S_\tau) = E\xi_1 E\tau$.

Aula 26 - Propriedades de esperança condicional

Teorema 26.1. Propriedades de $E(X | \mathcal{F})$.

1. Sejam X_n, X, Y variáveis aleatórias tais que $X_n \rightarrow X$ qc e $|X_n| \leq Y$ qc, com Y integrável. Então $E(X | \mathcal{F}) = \lim_{n \rightarrow \infty} E(X_n | \mathcal{F})$.
2. Sejam $X_n \uparrow X$, com $X_n \geq Y$ onde $EY > -\infty$. Então $\lim_{n \rightarrow \infty} E(X_n | \mathcal{F}) = E(X | \mathcal{F})$.
3. Sejam $X_n \downarrow X$, com $X_n \leq Y$ onde $EY < \infty$. Então $\lim_{n \rightarrow \infty} E(X_n | \mathcal{F}) = E(X | \mathcal{F})$.
4. Sejam (X_n) e Y variáveis aleatórias tais que $X_n \geq Y$ e $EY > -\infty$. Então $E(\liminf X_n | \mathcal{F}) \leq \liminf E(X_n | \mathcal{F})$.
5. Sejam (X_n) e Y variáveis aleatórias tais que $X_n \leq Y$ e $EY < \infty$. Então $E(\limsup X_n | \mathcal{F}) \geq \limsup E(X_n | \mathcal{F})$.
6. Sejam (X_n) variáveis aleatórias integráveis tais que $S_n = \sum_{i=1}^n X_i$, com $|S_n| < Y$ onde $EY < \infty$. Então $E(\sum_{n=1}^{\infty} X_n | \mathcal{F}) = \sum_{n=1}^{\infty} E(X_n | \mathcal{F})$.

Aula 27 - Probabilidade condicional regular

Exemplo 27.1. Sejam (ξ_n) variáveis aleatórias independentes e identicamente distribuídas, com $P(\xi_n = 1) = p = 1 - P(\xi_n = -1)$. Defina $S_n = \sum_{k=1}^n \xi_k$. Então $E(S_{n+1} | \xi_1, \dots, \xi_n) = S_n + E(\xi_{n+1}) = S_n + 2p - 1$. Lembre que (ξ_1, \dots, ξ_n) é a menor σ -álgebra que torna ξ_i mensuráveis, com $1 \leq i \leq n$. Quando $p = \frac{1}{2}$, então $E(S_{n+1} | \xi_1, \dots, \xi_n) = S_n$. Essa propriedade é dita **propriedade de Martingal**. Jogos honestos têm essa propriedade.

Lema 27.1. Sejam $X : \Omega \rightarrow \mathbb{R}$ variável aleatória e $Y : \Omega \rightarrow \mathbb{R}$ $\sigma(X)$ -mensurável. Então existe $m : \mathbb{R} \rightarrow \mathbb{R}$ Borel mensurável tal que $Y = m \circ X$.

Observação 27.1. Sabemos que $E(Y | X)$ é X -mensurável. Então, pelo lema, existe $m : \mathbb{R} \rightarrow \mathbb{R}$ Borel mensurável tal que $E(Y | X) = m \circ X$. Defina $E(Y | X = x)$ como sendo $m(x)$. Observe que para todo $A \in \sigma(X)$, $\int_{A=X^{-1}(B)} Y dP = \int_B m(x) dP_X(x)$.

Definição 27.1. Seja (Ω, \mathcal{A}, P) espaço de probabilidade.

- a) Sejam X e Y variáveis aleatórias. Definimos a **esperança condicional de Y dado (X = x)** como qualquer função $m : \mathbb{R} \rightarrow \mathbb{R}$ tal que $\forall A \in \sigma(X)$, $\int_{A=X^{-1}(B)} Y dP = \int_B m(x) dP_X(x)$.
- b) Dados $A \in \mathcal{A}$ e X variável aleatória, definimos a **probabilidade condicional de A dado (X = x)**, como sendo $E(\mathbb{1}_A | X = x)$. Notação: $P(A | X = x)$.

Observação 27.2.

1. Defina a medida $Q : \mathcal{B}(\mathbb{R}) \rightarrow \overline{\mathbb{R}}$, onde $Q(B) = \int_{A=X^{-1}(B)} Y dP$. Note que $Q \ll P_X$. Então, por Radón-Nikodým, existe $m : \mathbb{R} \rightarrow \mathbb{R}$ mensurável tal que $Q(B) = \int_{A=X^{-1}(B)} Y dP = \int_B m(x) dP_X(x)$, onde $m = \frac{dQ}{dP_X}$.
2. Conhecendo $E(Y \mid X)$ podemos definir, $E(Y \mid X = x)$, $\forall x \in \mathbb{R}$. A recíproca também é verdadeira. Se conhecemos $E(Y \mid X = x) = m(x)$ tal que $\forall A \in \sigma(X)$, $\int_{A=X^{-1}(B)} Y dP = \int_B m(x) dP_X(x) = \int_A m(X) dP$, como $m(X)$ é σ -mensurável, temos que $m(X) = E(Y \mid X)$ qc.
3. Observe que $P(A \cap (X \in B)) = \int_{X^{-1}(B)} \mathbb{1}_A dP = \int_B P(A \mid X = x) dP_X(x)$.

Exemplo 27.2. Seja (Y, X) vetor aleatório absolutamente contínuo com densidade $f(y, x)$. Definimos a **densidade condicional de Y dado $(X = x)$** , como $f_{Y|X}(y \mid x) : \mathbb{R} \rightarrow \mathbb{R}_+$

$$f_{Y|X} = \begin{cases} \frac{f(y, x)}{f_X(x)}, & \text{se } f_X(x) \neq 0 \\ 0, & \text{se } f_X(x) = 0 \end{cases}$$

Então $P(Y \in C \mid X = x) = \int_C f_{Y|X}(y \mid x) dy$.

Definição 27.2. Seja (Ω, \mathcal{A}, P) espaço de probabilidade. Dados $A \in \mathcal{A}$ e $\mathcal{F} \preccurlyeq \mathcal{A}$. Definimos a **probabilidade condicional regular de A dada \mathcal{F}** , como uma função $P(\omega, A) : \Omega \times \mathcal{A} \rightarrow \mathbb{R}$ satisfazendo

- a) $P(\omega, \cdot) : \mathcal{A} \rightarrow \mathbb{R}$ é medida de probabilidade, para todo $\omega \in \Omega$.
- b) $P(\cdot, A) : \Omega \rightarrow \mathbb{R}$ é uma versão de $P(A \mid \mathcal{F})$, para todo $A \in \mathcal{A}$, isto é, $P(\omega, A) = P(A \mid \mathcal{F})(\omega)$ qc, para todo $A \in \mathcal{A}$.

Teorema 27.1. Sejam X, Y variáveis aleatórias e $P(\omega, A)$ a probabilidade condicional regular de A dado $\sigma(X)$. Então $E(Y \mid X)(\omega) = \int_{\Omega} Y(\tilde{\omega}) P(\omega, d\tilde{\omega})$.

Corolário 27.1. Seja (Y, X) absolutamente contínua. Então $E(g(Y) \mid X = x) = \int_{\mathbb{R}} g(y) f_{Y|X}(y \mid x) dy$, com $g : \mathbb{R} \rightarrow \mathbb{R}$ integrável.

Aula 28 - Distribuição e função de distribuição condicional regular

Definição 28.1. Seja (X_n) sequência de variáveis aleatórias. Dizemos que (X_n) é uma **cadeia de Markov**, se $\forall B \in \mathcal{B}(\mathbb{R})$ e $\forall n \in \mathbb{N}$, temos que $P(X_{n+1} \in B \mid X_1, X_2, \dots, X_n)(\omega) = P(X_{n+1} \in B \mid X_n)(\omega)$.

Definição 28.2. Sejam (Ω, \mathcal{A}, P) espaço de probabilidade (Γ, \mathcal{S}) espaço de medida. Dados $\mathcal{F} \preceq \mathcal{A}$ e $X : \Omega \rightarrow (\Gamma, \mathcal{S})$, definimos a **distribuição condicional regular de X dado \mathcal{F}** , como sendo $Q(\omega, B) : \Omega \times \mathcal{S} \rightarrow \mathbb{R}$ uma função que satisfaz:

- a) $\forall \omega \in \Omega$, $Q(\omega, \cdot)$ é medida de probabilidade em (Γ, \mathcal{S}) .
- b) $\forall B \in \mathcal{S}$, $Q(\cdot, B)$ é uma versão de $P(X \in B \mid \mathcal{F})$.

Definição 28.3. Dados $\mathcal{F} \preceq \mathcal{A}$ e $X : \Omega \rightarrow \mathbb{R}$ variável aleatória, definimos a **função de distribuição condicional regular de X dado \mathcal{F}** , como sendo $F(\omega, x) : \Omega \times \mathbb{R} \rightarrow [0, 1]$ uma função que satisfaz:

- a) $\forall \omega \in \Omega$, $F(\omega, \cdot)$ é uma função de distribuição.
- b) $\forall x \in \mathbb{R}$, $F(\cdot, x)$ é uma versão de $P(X \leq x \mid \mathcal{F})$.

Teorema 28.1. Existem a distribuição e a função de distribuição condicionais regulares de X dado \mathcal{F} .

Aula 29 - Prova 3

Foi realizada a terceira prova.