Probabilidade Avançada, Aula 01

25/02/2014

- Durret
- Bilingsley
- Feller (não tanto)
- Royden (alguma coisa)

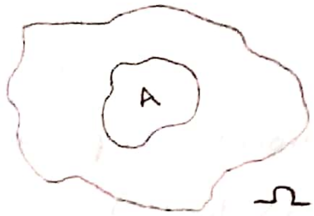
$$\frac{\max\{P_1,P_2\} + \max\{\min\{P_1,P_2\}, P_3\}}{2}$$

Trio : $(\Omega, \mathcal{F}, P)$, Ω : espaço amostral : "conj. não vazio"

$\mathcal{F}$: "classe de subconjuntos de Ω" "eventos"

P : "probabilidade em $\mathcal{F}$"

$P : \mathcal{F} \to [0,1]$



Exemplo: "lançamento de uma moeda honesta um número arbitrário de vezes". Sendo

$$0 = \text{"cara"} \quad e \quad 1 = \text{"coroa"},$$

segue que $\Omega = \{0,1\}^{\mathbb{N}}$ (conjunto de funções de $\mathbb{N}$ em $\{0,1\}$).

Obs. A^B = conjunto de funções de B em A.

$\mathbb{N}_+^{\mathbb{R}}$ = conjunto de sequências, por exemplo.

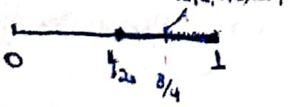
Temos Ω não-enumerável, pois $[0,1]$ não o é.

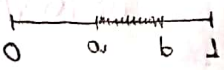
Basta notar que existe uma associação entre cada $x \in [0,1]$ à mais de um elemento de Ω.

$$x_{[0,1]} = \sum_{i \geq 1} a_i \left(\frac{1}{2}\right)^i \quad , \quad x = \frac{1}{2} = \begin{cases} (1,0,0,\dots) \\ (0,1,1,\dots) \end{cases}$$

$x = (a_1, a_2, \dots)$, $a_i = 1 \vee 0$.

Ω =





$A = (a,b) \in [0,1]$

$P(A) = b - a$

Questão: Posso estender essa "medida de probabilidade" para o conjunto das partes de Ω, $\mathcal{P}(\Omega)$.

Suponha que quero definir P em $\Omega = (0,1]$ tal que

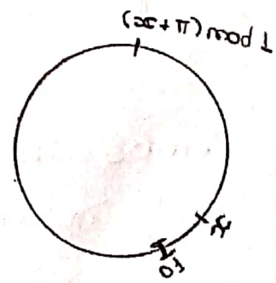
1. Se $A = (a,b] \in \Omega \Rightarrow P(A) = b - a$
2. Se A tem $P(A) \Rightarrow P(A \oplus x) = P(A)$, com
   $\{A \oplus x\} = \{\omega \in \Omega : \omega = (y + x) \bmod 1,\ y \in A\}$, "A transladado por x"
3. P está definida em $\mathcal{P}(\Omega)$ (todos os subconjuntos de Ω)
4. Se $A_1, A_2, \dots$ são disjuntos
   $$P\left(\bigcup_n A_n\right) = \sum_{n \geq 1} P(A_n)$$

Não consigo as 4 declarações. Assim, vamos enfraquecer a condição 3, mas não de qualquer modo. Vamos tomar a maior classe possível (σ-álgebra).

Exemplo: $\Omega = [0,1]$, $x \in \Omega$

Defino "Família de x" $= F(x) \subset \Omega$

$F(x) = \{y \in \Omega : y = (x + n\pi) \bmod 1,\ n \in \mathbb{Z}\}$



gira na circunferência de comprimento 1.

Para cada família, escolho o "chefe". Após a escolha, tenho E.

$P(E) = ?$ $\qquad E_k \cap E_l = \phi$

$E \to E_k = (E \oplus k\pi) \bmod 1,\ k \in \mathbb{Z}$

$\bigcup_k E_k = (0,1] = \Omega$

- Axioma da Escolha
- Paradoxo de Banach-Tarskis
  ↳ Uma maçã corto-a em vários pedaços e recupero duas maçãs.

$(\Omega, \mathcal{F}, P)$

Ω = "conjuntos não vazio"

$\mathcal{F}$ = "classe de conjuntos de Ω" tal que

1. $\Omega \in \mathcal{F}$, $\emptyset \in \mathcal{F}$. A é um evento

"evento certo" "evento impossível"

2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

3. Se $A_1, A_2, \dots$ estão em $\mathcal{F}$

$\bigcup_i A_i$: "acontecer algum deles" está em $\mathcal{F}$

$A_1, A_2, \dots$ em $\mathcal{F} \Rightarrow \bigcup_i A_i \in \mathcal{F}$

$\mathcal{F}$ é σ-álgebra

Obs. Se 3 for válida apenas para uma sequência finita : álgebra.

Exemplo: $\mathcal{F} = \{\emptyset, \Omega\}$

$\mathcal{F}_A = \{\emptyset, A, A^c, \Omega\}$

$\mathcal{F} = \mathcal{P}(\Omega)$: "todos os subconjuntos de Ω"

Obs. Se $\mathcal{A}$ é uma coleção qualquer. Defino

$\sigma(\mathcal{A})$ = menor σ-álgebra contendo $\mathcal{A}$

Exercício. Seja $\mathcal{F}_\theta$, $\theta \in \Theta$, Θ um espaço de índices, σ-álgebras, então

$$\mathcal{F} = \bigcap_{\theta \in \Theta} \mathcal{F}_\theta \text{ é } \sigma\text{-álgebra}$$

Obs. Vale p/ união? Não! Contra-exemplo.

$\mathcal{A}_1 = \{\{a\}, \{b,c\}, \emptyset, \{a,b,c\}\}$ $\mathcal{A}_1 \cup \mathcal{A}_2$ não é σ-álgebra pois

$\mathcal{A}_2 = \{\{a,b\}, \{c\}, \emptyset, \{a,b,c\}\}$ $\{b\}^* = \{b,c\} \cap \{a,b\}$ não estão em $\mathcal{A}_1 \cup \mathcal{A}_2$

* Obs. Se $A, B \in \mathcal{F}$,

$$A \cup B \in \mathcal{F} \text{ e } A \cap B \in \mathcal{F} \text{ pois}$$

$$(A \cap B) = \underbrace{(\underbrace{\underbrace{A^c}_{\in \mathcal{F}} \cup \underbrace{B^c}_{\in \mathcal{F}}}_{\in \mathcal{F}})^c}_{\in \mathcal{F}}$$

$(\Omega, \mathcal{F}, P)$

$\mathcal{F}$. σ-álgebra

P é uma função à valores reais definida em $\mathcal{F}$

1. $P(\Omega) = 1$; $P(\emptyset) = 0$
2. $A \in \mathcal{F} \Rightarrow 0 \leq P(A) \leq 1$
3. Se $A_1, A_2, \ldots$ é uma coleção disjunta de eventos, então

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \quad (\sigma\text{-aditividade})$$

Exercício.

. Se $A, B \in \mathcal{F}$, então

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

. Se $A_1, A_2, \ldots, A_n$ estão em $\mathcal{F}$,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) +$$

$$\ldots + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n)$$

Exemplo: "Baralho embaralhado"

Baralho tem N cartas, $N \geq 1$

$(\Omega, \mathcal{F}, P)$

$\Omega =$ conj. das permutações de N cartas $=$

$= \{x = (a_1, a_2, \dots, a_N), a_i \neq a_j$ se $i \neq j$, $a_i \in \{1, 2, \dots, N\}\}$

Obs. $|A| = \#$ de elementos de A

Seja $A_N =$ "nenhuma carta está na posição inicial".

Queremos $P(A_N)$. Para tal, consideremos

$A_N^c =$ "pelo menos uma carta está na posição inicial"

e

$A_i =$ "a ~~carta~~ i-ésima carta esta na posição i".

Logo,

$$A_N^c = \bigcup_{i=1}^{N} A_i \quad e \quad P(A_N^c) = P\left(\bigcup_{i=1}^{N} A_i\right) \overset{?}{=} 1 - e^{-1}$$

Assim,

$$P(A_N) = 1 - P(A_N^c) = e^{-1}.$$

Probabilidade Avançada I, Aula 02 27/02/2014

Na aula passada,

- Espaço de probabilidade: $(\Omega, \mathcal{F}, P)$

Ω: espaço amostral

$\mathcal{F}$: classe de subconjuntos de Ω (σ-álgebra)

- $\Omega, \emptyset \in \mathcal{F}$
- Se $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
- Se $A_1, A_2, \dots$ em $\mathcal{F}$, $\bigcup_i A_i \in \mathcal{F}$

P: medida de probabilidade em $\mathcal{F}$ $(P: \mathcal{F} \to \mathbb{R})$

- $P(\Omega) = 1 = 1 - P(\emptyset)$
- Se $A_1, A_2, \dots$ em $\mathcal{F}$ disjuntos $(A_i \cap A_j = \emptyset, i \neq j)$, então $P(\bigcup_i A_i) = \sum_i P(A_i)$ (σ-aditividade)

Agora,

Propriedades:

a) Monotocidade. Se $A \subset B$, então $P(A) \leq P(B)$

Se $A \subset B$, então $A \cap B = A$. Note agora que

$$B = (A \cap B) \cup (A^c \cap B)$$

é uma forma de reescrever B como união de dois conjuntos disjuntos, pois $(A \cap B) \cap (A^c \cap B) = (A \cap (A^c \cap B)) \cap (B \cap A^c \cap B)) = \emptyset \cap (B \cap A^c) = \emptyset$.

Logo,

$$P(B) = P(A \cap B) + P(A^c \cap B) = P(A) + \underbrace{P(B \cap A^c)}_{\geq 0} \Leftrightarrow$$

$$P(B) \geq P(A)$$

b) Sub-σ-aditividade

Se $A_1, A_2, \dots$ estão em $\mathcal{F}$, então

$$P(\bigcup_i A_i) \leqslant \sum_i P(A_i)$$

Se $A_1 = B_1$ e $B_k = A_k \cap A_1^c \cap \dots \cap A_{k-1}^c$, então os B_k's são disjuntos e $\bigcup_k A_k = \bigcup_k B_k$, de modo que

$$P(\bigcup_k A_k) = P(\bigcup_k B_k) = \sum_{k=1}^{n} P(B_k)$$

Como $P(B_k) \leqslant P(A_k)$ pela monotocidade, temos que

$$P(\bigcup_k A_k) = \sum_k P(B_k) \leqslant \sum_k P(A_k), \text{ i.e.}$$

$$P(\bigcup_k A_k) \leqslant \sum_k P(A_k) \qquad \square$$

c) Continuidade por sequências crescentes.

Se $A_n \uparrow A$, (i.e., $A_1 \subset A_2 \subset \dots$ e $\bigcup A_i = A$)

$$P(A_n) \xrightarrow[n \uparrow \infty]{} P(A)$$

Uma vez que $A_1 \subset A_2 \subset \dots$, segue que $P(A_i)$ é não decrescente em i pela monotocidade. Como também é limitada, então $\lim_{n \to \infty} P(A_n)$ existe.

Considerando agora $B_1 = A_1$, $B_k = A_k \cap A_1^c \cap A_2^c \cap \dots \cap A_{k-1}^c$, temos

$$P(A) = P(\bigcup_{i=1}^{\infty} A_i) = P(\bigcup_{i=1}^{\infty} B_i) = P(B_1) + P(B_2) \dots =$$

$P(B_1) = P(A_1)$, $P(B_k) = P(A_k) - P(A_{k-1})$

$$= P(A_1) + [P(A_2) - P(A_1)] + [P(A_3) - P(A_2)] + \dots,$$

Daí, pela definição de conv. de séries infinita:

$$P(A) = \lim_{n \to \infty} [P(A_1) + [P(A_2) - P(A_1)] + \dots + [P(A_n) - P(A_{n-1})]] = \lim_{n \to \infty} P(A_n).$$

d) Continuidade por sequências decrescentes.

$A_1 \supset A_2 \supset A_3 \supset \dots \qquad \bigcap_{k=1}^{\infty} A_k = A$, $P(A_i)$ é não-cresc. em i

Nesse caso, note que

$A_1^c \subset A_2^c \subset A_3^c \subset \dots$ é uma seq. crescente com

$$\bigcup_{k=1}^{\infty} A_k^c = A^c$$

Daí, pelo item (c),

$$P(A^c) = \lim_{n\to\infty} P(A_n^c) \iff$$

$$1 - P(A) = \lim_{n\to\infty} [1 - P(A_n)] \iff$$

$$P(A) = \lim_{n\to\infty} P(A_n) \Rightarrow P(A_n) \downarrow P(A)$$
$$\scriptstyle n\uparrow\infty$$

- Medida de Lebesgue $(\mathbb{R}, \mathcal{B})$ espaço mensurável
- Existência e unicidade da extensão da medida

Idéia: Quero definir μ, medida:

1. Defino μ numa classe pequena de sub. de Ω (ou $\mathbb{R}$)
2. Estendo μ para uma álgebra
3. Estendo para σ-álgebra

$(\mathbb{R}, \mathcal{B})$

$\mathcal{B}$ = sigma álgebra de Borel = menor σ-álgebra contendo os intervalos abertos de $\mathbb{R}$

Defino : μ = "comprimento" para o intervalo $(a,b]$

$\mu((a,b]) = b - a$

Isto é, defino μ em

$\mathcal{J}$ = classe de int. da forma $(a,b]$, $a<b$ em $\mathbb{R}$

Definição. Uma coleção $\mathcal{S}$ de conjuntos é dita uma semi-álgebra se

1. É fechada por intersecções

   Se $A, B \in \mathcal{S} \Rightarrow A \cap B \in \mathcal{S}$

2. Se $A \subset \mathcal{S}$, então A^c é uma união finita de elementos de $\mathcal{S}$.

Seja $\mathcal{A} = \emptyset$ mais todas as uniões finitas de elementos disjuntos de $\mathcal{J}$

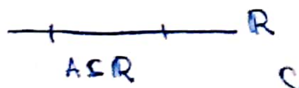
Obs: $\mathcal{A}_d = \emptyset$ mais uniões finitas de int. disjuntos em $\mathbb{R}^d$

$$(a_1, b_1] \times (a_2, b_2] \times \dots \times (a_d, b_d]$$

Exercício. $\mathcal{A}$ é uma álgebra.

1. $\Omega = \mathbb{R} \in \mathcal{A}$.

Exercício. Escreva a fórmula $\mu\left(\bigcup_{i=1}^{n} (a_i, b_i]\right)$.



Seja $A_n, n \geq 1$ em $\mathcal{A}$ tal que $A \subset \bigcup_n A_n$

$$\mu^*(A) = \inf_{\substack{A_n \\ \bigcup A_n \supset A}} \sum_n \mu(A_n)$$

: aproximação por cima da medida de A.

"medida exterior"

No entanto, $\mu^*(\cdot)$ não satisfaz a σ-aditividade.

Caratheodory: $\mathcal{M}$: classe dos conjuntos mensuráveis.

Digo que A é mensurável se

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c), \forall E \subset \mathbb{R}$$

Em M, μ^* é sub-σ-aditiva.

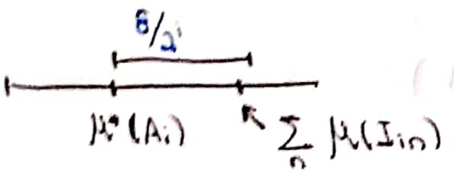
Lema: Se $A_1, A_2, \ldots$ subconjuntos de $\mathbb{R}$,

$$\mu^*\left(\bigcup_i A_i\right) \leq \sum_i \mu^*(A_i)$$

Demonstração:

$$\mu^*(A_i) = \inf_{\{I_{in}\}_{n\geq 1}} \sum_n \mu(I_{in})$$

A_i: cobertura I_{in}, $n \geq 1$



prop. do ínfimo

Para cada i, existe $\{I_{in}\}_{n\geq 1}$ tal que

$$\sum_n \mu(I_{in}) - \mu^*(A_i) < \frac{\varepsilon}{2^i}$$

$$A_i \subset \bigcup_n I_{in} \Leftrightarrow \bigcup_i A_i \subset \bigcup_i \bigcup_n I_{in} \Leftrightarrow$$

$$\mu^*\left(\bigcup_i A_i\right) \leq \sum_i \sum_n \mu(I_{in})$$

$$\leq \sum_i \left(\mu^*(A_i) + \frac{\varepsilon}{2^i}\right) =$$

$$\sum_i \mu^*(A_i) + \varepsilon, \quad \forall \varepsilon > 0$$

Logo, $\mu^*\left(\bigcup_i A_i\right) \leq \sum_i \mu^*(A_i)$

Lema. $\mathcal{M}$ é uma álgebra.

1. $\Omega \in \mathcal{M}$.

$$\mu^*(E) = \mu^*(E \cap \Omega) + \mu^*(E \cap \underbrace{\Omega^c}_{\emptyset}) =$$

$$= \mu^*(E \cap \Omega) + \mu^*(E \cap \emptyset) =$$

$$= \mu^*(E)$$

...

Lema. μ^* é aditiva em $\mathcal{M}$.

Demonstração.

Se A, B disjuntos em $\mathcal{M}$,

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$$

$A \in \mathcal{M}$, então $\forall E$

$$\mu^*(E) = \mu^*(E \cap A^c) + \mu^*(E \cap A)$$

Tomando $E = A \cup B$, segue que

$$\mu^*(A \cup B) = \mu^*((A \cup B) \cap A) + \mu^*((A \cup B) \cap A^c) =$$

$$= \mu^*(A) + \mu^*(B) \qquad \blacksquare$$

Lema. $\mathcal{M}$ é σ-álgebra.

μ^* é sub-σ-aditiva

$$\mu^*(\bigcup_i A_i) \leq \sum_i P(A_i)$$

Exercício. Se $A_1, A_2, \ldots$ estão em $\mathcal{M}$, então $\bigcup_i A_i \in \mathcal{M}$

$$\bigcup_i A_i \in \mathcal{M} \Leftrightarrow \forall E \in \mathbb{R}$$

$$\mu^*(E) = \mu^*(E \cap (\bigcup_i A_i)) + \mu^*(E \cap (\bigcup_i A_i)^c)$$

Unicidade: Usamos o Teorema de Dinkin.

Definição: Uma classe de conjuntos $\mathcal{P}$ é dita um sistema-π se for fechada por intersecções.

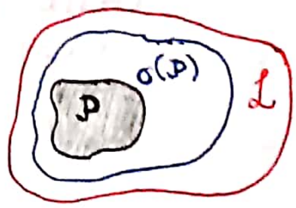
Definição: Uma classe $\mathcal{L}$ de conjuntos é dita um sistema-λ se

1. $\Omega \in \mathcal{L}$
2. Se A e B estão em $\mathcal{L}$ e $A \subset B$, então $B - A \in \mathcal{L}$,
3. Se $A_1 \subset A_2 \subset \dots$ é uma coleção de conjuntos em $\mathcal{L}$, então

$$\bigcup_i A_i \in \mathcal{L}$$

Teorema de Dynkin: Se $\mathcal{P}$ é um sistema-π contido num sistema-λ $\mathcal{L}$, então

$$\mathcal{P} \subset \sigma(\mathcal{P}) \subset \mathcal{L}$$



Demonstração: $\mathcal{P}$: sistema-π

Def. $\ell(\mathcal{P})$: menor sistema-λ contendo $\mathcal{P}$

(*) Lema: $\ell(\mathcal{P})$ é uma σ-álgebra $\Rightarrow \mathcal{P} \subset \sigma(\mathcal{P}) \subset \ell(\mathcal{P})$

Logo, basta provar (*).

Lema. Se $\mathcal{L}$ é um sistema-λ que é um sistema-π, então $\mathcal{L}$ é uma σ-álgebra (Exercício)

Portanto, basta provar que $\ell(\mathcal{P})$ é fechado por intersecções.
Se $A \in \ell(\mathcal{P})$, defino $\mathcal{G}_A = \{B : A \cap B \in \ell(\mathcal{P})\}$

Lema. Se $A \in \ell(\mathcal{P})$ então $\mathcal{G}_A$ é um sistema-λ.

Dem. a) $\Omega \in \mathcal{G}_A$

$$\Omega \cap A = A \Rightarrow \Omega \in \mathcal{G}_A$$

b) Se B, C em $\mathcal{G}_A$ e $B \supset C \overset{?}{\Rightarrow} B - C \in \mathcal{G}_A$

$$A \cap (B - C) = \underbrace{\underbrace{(A \cap B)}_{\in \ell(\mathcal{P})} - \underbrace{(A \cap C)}_{\in \ell(\mathcal{P})}}_{\in \ell(\mathcal{P})}$$

pois $\ell(\mathcal{P})$ é sist.-λ e $(A\cap B) \supset (A \cap C)$.

c) Se $B_n \in \mathcal{G}_A$ com $B_n \uparrow B \overset{?}{\Rightarrow} B \in \mathcal{G}_A$

$$A \cap B_n \in \mathcal{G}_A \ \forall n$$

$$\underset{\in \ell(\mathcal{P})}{(A \cap B_n)} \uparrow \underset{\in \ell(\mathcal{P})}{(A \cap B)} \Rightarrow B \in \mathcal{G}_A$$

pois $\ell(\mathcal{P})$ é um sistema-λ, que é fechado por união de seq. crescentes

• Agora: Se $A \subset \mathcal{P}$ então $\boxed{\mathcal{P} \subset \mathcal{G}_A} = \{B : A \cap B \in \ell(\mathcal{P})\}$

Dem: Se $C \in \mathcal{P} \Rightarrow A \cap C \in \mathcal{P} \in \ell(\mathcal{P})$

$$\ell(\mathcal{P}) \subset \mathcal{G}_A$$

menor sistema-λ

+ Se $A \subset \mathcal{P}$, então $\mathcal{P} \subset \mathcal{G}_A$ pois $A \cap \mathcal{P} = A \in \ell(\mathcal{P}) \Rightarrow \mathcal{P} \subset \mathcal{G}_A \Rightarrow$ $\mathcal{G}_A$ é um sistema-λ que contém $\mathcal{P}$. Como $\ell(\mathcal{P})$ é o menor sistema-λ que contém $\mathcal{P}$, segue que $\ell(\mathcal{P}) \subset \mathcal{G}_A$

Probabilidade Avançada I, Aula 03 06/03/2014

$(\Omega, \mathcal{F}, P)$

$(\mathbb{R}, \mathcal{B}, \mu)$ μ Lebesgue

1. Defino P (ou μ) numa semiálgebra $\mathcal{S}$

a) Fechada por intersecções

b) $A \in \mathcal{S}$ então A^c = união finita de elementos de $\mathcal{S}$

Exemplo: $\mathcal{S} = \emptyset$ mais intervalos $(a,b]$

$$\overset{P}{\mathcal{S}} \rightarrow \mathcal{A} \xrightarrow{*} \sigma(\mathcal{A})$$

álgebra gerada por $\mathcal{A}$: "conj. de todas as uniões disjuntas de elementos de $\mathcal{S}$".

- Teorema da Extensão de Caratheodory
  - Unicidade
  - Existência

Unicidade: Teorema $\lambda - \pi$

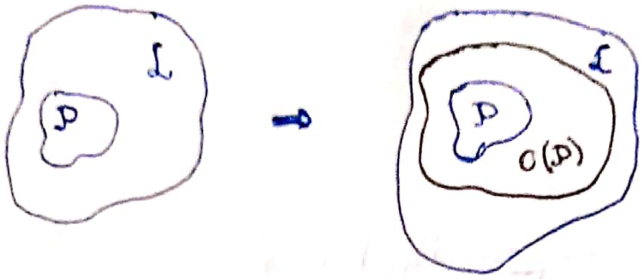
Sistema π $\mathcal{P}$ fechado por intersecções

Sistema-λ $\mathcal{L}$: a) $\Omega \in \mathcal{L}$

b) Se $A \subset B$ em $\mathcal{L}$, $B - A \subset \mathcal{L}$

c) Se $A_n \uparrow A$, $A_n \in \mathcal{L}$, $n \geq 1$ então $A \in \mathcal{L}$

Teorema λ-π: Se $\mathcal{P}$ é sistema π e



$\mathcal{L}$ é um sistema λ com $\mathcal{L} \supset \mathcal{P} \Rightarrow \mathcal{P} \subset \sigma(\mathcal{P}) \subset \mathcal{L}$

Corolário. Seja $\mathcal{P}$ um sistema π e ν_1 e ν_2 duas medidas em duas σ-álgebras $\mathcal{F}_1$ e $\mathcal{F}_2$, contendo $\mathcal{P}$, que concordam em $\mathcal{P}$

$$A \in \mathcal{P} \Rightarrow \nu_1(A) = \nu_2(A)$$

Suponha que existe $A_n \uparrow \Omega$ tal que $\nu_i(A_n) < \infty$, $\forall n \geq 1$, $i=1,2$. Então ν_1 e ν_2 concordam em $\sigma(\mathcal{P})$.

Demonstração: Seja $A \in \mathcal{P}$ tal que

$$\nu_1(A) = \nu_2(A) < \infty$$

$$\mathcal{G}_A = \{B \in \sigma(\mathcal{P}) : \nu_1(A \cap B) = \nu_2(A \cap B)\}$$

Mostro que $\mathcal{G}_A$ é sistema λ.

1. $\Omega \in \mathcal{G}_A$ ✓

   $\Omega \in \sigma(\mathcal{P})$ e $\nu_1(A \cap \Omega) = \nu_2(A \cap \Omega) = \nu_i(A)$, $i=1,2$, pois $A \in \mathcal{P}$.

2. Se $B \subset C$ em $\mathcal{G}_A \Rightarrow C - B$ em $\mathcal{G}_A$

   $(C-B) \cap A = (C \cap A) - (B \cap A)$

   $\nu_1((C-B)\cap A) = \nu_1(C \cap A) - \nu_1(B \cap A) \overset{C,B \in \mathcal{G}_A}{=} \nu_2(C \cap A) - \nu_2(B \cap A) =$
   $= \nu_2((C-B) \cap A)$

3. Se $B_n \uparrow B$, $B_n \in \mathcal{G}_A \Rightarrow B \in \mathcal{G}_A$

$B_n \uparrow B$

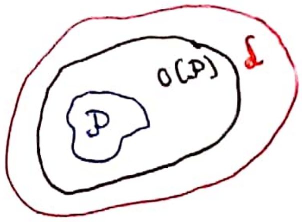
$$\nu_1(A \cap B) \overset{\text{cont. por sequências crescentes}}{=} \lim_{n \uparrow \infty} \nu_1(A \cap B_n) \underset{B_n \in \mathcal{G}_A}{=}$$

$$= \lim_{n \uparrow \infty} \nu_2(A \cap B_n) = \nu_2(A \cap B)$$

$\mathcal{P}$ sistema π

$\mathcal{L} := \mathcal{G}_A$ sistema λ

Teorema λ-π $\Rightarrow \sigma(\mathcal{P}) \subset \mathcal{G}_A$



Então se $A \in \mathcal{P}$ com $\nu_1(A) = \nu_2(A) < \infty$ e $B \in \sigma(\mathcal{P})$

$$\nu_1(A \cap B) = \nu_2(A \cap B)$$

$A \to A_n \uparrow \Omega$, $A_n \in \mathcal{P}$ com $\nu_1(A_n) = \nu_2(A_n) < \infty$

* continuidade

$$\lim_{n \uparrow \infty} \nu_1(A_n \cap B) = \lim_{n \uparrow \infty} \nu_2(A_n \cap B) \overset{*}{\Rightarrow}$$

$$\nu_1(\Omega \cap B) = \nu_2(\Omega \cap B) \Rightarrow$$

$$\nu_1(B) = \nu_2(B), \ \forall B \in \sigma(\mathcal{P}) \quad \blacksquare$$

Pergunta: $\mathcal{G}_{A_n}$ é σ-álgebra? (?)

1. $A_n \in \mathcal{P}$, então se $A \in \mathcal{G}_A$

$\nu_1(A_n \cap A) = \nu_2(A_n \cap A)$, $A^c \in \mathcal{G}_A$?

$\nu_1(A_n^c \cup A) = \nu_1(A_n^c) + \nu_1(A) - \nu_1(A_n^c \cap A)$

Contra-exemplo: $\Omega = \{1,2,3,4\}$ e $\mathcal{F} = \mathcal{P}(\Omega)$

e $C_1 = \{1,2\}$, $C_2 = \{2,3\}$, $C_3 = \{3,4\}$

$\mathcal{C} = \{C_1, C_2, C_3\}$

	ν_1	ν_2
1.	1/4	1/8
2.	1/4	3/8
3.	1/4	1/8
4.	1/4	3/8

ν_1 e ν_2 concordam em $\mathcal{C}$ mas...

$\mathcal{C}$ não é sistema-π

Existência

μ em A-álgebra

medida exterior μ^*

$A \subset \Omega$

$$\mu^*(A) = \inf_{\substack{A \subset \bigcup A_n \\ A_n \in A}} \sum \mu(A_n)$$

Propriedades:

1. Monotonicidade

$$E \subset F \Rightarrow \mu^*(E) \leq \mu^*(F)$$

2. sub-aditividade

Se $F \subset \bigcup F_i$, $\mu^*(F) \leq \sum_i \mu^*(F_i)$

Definição: Conjuntos Mensuráveis

$\mathcal{M}$ conjuntos dos E t.q. para todo $F \subset \Omega$,

$$\mu^*(F) = \mu^*(F \cap E) + \mu^*(F \cap E^c)$$

Obs. Dado E,

$F \subset \bigcup F_i$, $F_1 = F \cap E$, $F_2 = F \cap E^c$, $F_i = \emptyset$, $i \geq 3$

Por 2)

$$\mu^*(F) \leq \mu^*(F\cap E) + \mu^*(F\cap E^c)$$

Então a igualdade segue da desigualdade oposta.

$\cup F_i \subset F = \mu^*(\cup F_i) \leq \mu^*(F)$ por 1.

Lema.

a) Se $A \in \mathcal{A}$, então

$$\mu^*(A) = \mu(A)$$

b) $\mathcal{A} \subset \mathcal{M}$

Todos os elementos da álgebra são mensuráveis

Demonstração: Se $A \in \mathcal{A} \subset \cup A_i$, $A_i \in \mathcal{A}$

Em $\mathcal{A}$, μ é subaditiva

$$\mu(A) \leq \sum_i \mu(A_i)$$

$$\Downarrow$$

$$\mu(A) \leq \inf_{\substack{A_i \in \mathcal{A} \\ \cup A_i \supset A}} \sum \mu(A_i) = \mu^*(A)$$

- Preciso mostrar que $\mu^*(A) \leq \mu(A)$.

$\mu^*(A) = \inf_{\substack{A_i \in \mathcal{A} \\ \cup A_i \supset A}} \sum \mu(A_i)$. Tomando $A_1 = A$, $A_i = \emptyset$, $i \geq 2$, temos que

$$\mu^*(A) \leq \mu(A). \qquad \text{Daí, } \mu(A) = \mu^*(A).$$

- $\mu^*(A) \leq \sum_i \mu(A_i)$ para uma $\{A_i\}_{i\geq 1}$ com $A \subset \cup A_i$. Em particular, para $A_1 = A$, $A_i = \emptyset$, $i \geq 2$, $\mu^*(A) \leq \mu(A)$.

Agora, verifico que

$$A \in \mathcal{A} \Rightarrow \mu^*(F) \geq \mu^*(F\cap A) + \mu^*(F\cap A^c), \quad \forall F \subset \Omega$$

Se $\mu^*(F) = \infty$, ok

Basta ~~verificar~~ considerar $F \subset \Omega$ t.q. $\mu^*(F) < \infty$.

Dado $\varepsilon > 0$, $\exists \{B_i\}_{i \geq 0}$ em $\mathcal{A}$ t.q. $F \subset \bigcup_i B_i$

$$\mu^*(F) + \varepsilon \geq \sum_i \mu(B_i) \geq \mu^*(F) \Leftrightarrow$$

$$\mu^*(F) + \varepsilon \geq \sum_i [\underbrace{\mu(B_i \cap A) + \mu(B_i \cap A^c)}_{\sigma\text{-aditivida na álgebra}}] = \sum_i \mu(B_i \cap A) + \sum_i \mu(B_i \cap A^c)$$

$$\geq \mu^*\left(\bigcup_i B_i \cap A\right) + \mu^*\left(\bigcup_i B_i \cap A^c\right)$$

$$\geq \mu^*(F \cap A) + \mu^*(F \cap A^c)$$

Como ε é arbitrário, segue $(*)$ $\square$

Lema. M é uma σ-álgebra.

Demonstração.

a) Se $E \in M \Rightarrow E^c \in M$ (imediato pela simetria)

Dado $F \subset \Omega$, $E \in M \Leftrightarrow$

$$\mu^*(F) = \mu^*(F \cap E) + \mu^*(F \cap E^c) =$$
$$= \mu^*(F \cap A^c) + \mu^*(F \cap A), \text{ com } A = E^c.$$

Daí, $A = E^c \in M$.

* 1º, mostramos que é álgebra

b) E_1, E_2 em $M \Rightarrow E_1 \cup E_2 \in M$

Se $G \subset \Omega$,

$$\mu^*(G) \overset{E_1 \in M}{=} \mu^*(G \cap E_1) + \mu^*(G \cap E_1^c)$$

$$\overset{E_2 \in M}{=} \mu^*(G \cap E_1) + \mu^*(G \cap E_1^c \cap E_2) + \mu^*(\underbrace{G \cap E_1^c \cap E_2^c}_{G \cap (\cup E_i)^c})$$

$G \cap [E_1 \cup (E_2 \cap E_1^c)] = G \cap (E_1 \cup E_2)$

$\mu^* [G \cap (E_1 \cup E_2)] \leq \mu^* (G \cap E_1) + \mu^* (G \cap E_2 \cap E_1^c) \Rightarrow$

$\mu^* (G) \geq \mu^* (G \cap (E_1 \cup E_2)) + \mu^* (G \cap (E_1 \cup E_2)^c) \Rightarrow$

$E_1 \cup E_2 \in \mathcal{M}$ ∎

de (a) + Morgan, segue que se

E_1, E_2 em $\mathcal{M} \Rightarrow E_1 \cap E_2 \in \mathcal{M}$

Quero mostrar que uniões enumeráveis de elementos de $\mathcal{M}$ estão em $\mathcal{M}$.

c) Se $\{E_i\}_{i \geq 1}$ em $\mathcal{M}$, então $\bigcup_i E_i \in \mathcal{M}$.

Posso tomar $E_1' = E_1$,

$E_2' = E_2 - E_1 = E_2 \cap E_1^c$

$\vdots$

$E_n' = E_n \cap (\bigcap_{i<n} E_i^c)$

Posso assumir $\{E_i\}_{i \geq 1}$ disjuntos por (b).

Se $F_n = \bigcup_{i \leq n} E_i \in \mathcal{M}$ por b

Mas

d) Se $G \subset \Omega$ e $E_1, E_2, \dots, E_n$ estão em $\mathcal{M}$ disjuntos

$\mu^* (G \cap (\bigcup_{i=1}^{n} E_i)) = \sum_i \mu^* (G \cap E_i)$

Demonstração: $F_n = \bigcup_{i \leq n} E_i \supset E_n$

$F_{n-1} \cap E_n \neq \emptyset$

$$\mu^*(G\cap F_n) = \mu^*(G\cap \overbrace{F_n\cap E_n}^{E_n}) + \mu^*(G\cap \overbrace{F_n\cap E_n^c}^{F_{n-1}}) =$$

$$= \mu^*(G\cap E_n) + \mu^*(G\cap F_{n-1})$$

item (*)

$$\mu^*(G\cap(\bigcup_{i=1}^{n} E_i)) = \sum_{i=1}^{n} \mu^*(G\cap E_i)$$

Então, usando (d) na demonstração de (c),

$$\mu^*(G) = \mu^*(G\cap F_n) + \mu^*(G\cap F_n^c)$$

$$F_n^c \supset E^c$$

$$\geqslant \mu^*(G\cap F_n) + \mu^*(G\cap E^c)$$

$$\overset{(d)}{\geqslant} \sum_{i=1}^{n} \mu^*(G\cap E_i) + \mu^*(G\cap E^c)$$

$$\downarrow n\to\infty$$

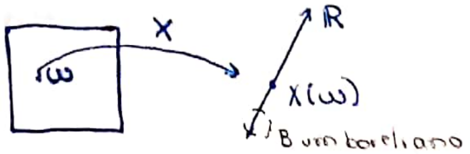
$$\geqslant \sum_i \mu^*(G\cap E_i) + \mu^*(G\cap E^c)$$

$$\geqslant \mu^*(\underbrace{G\cap(\cup E_i)}_{G\cap E}) + \mu^*(G\cap E^c) \Rightarrow E\in M$$

Aula 04 , Prob. Avançada 11/03

$(\Omega, \mathcal{F})$: espaço mensurável

$X: (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$



"Qual a prob. de ocorrer alguma coisa, tal que X assume valores em B?"

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$$

Definição: $X: (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ é dita variável aleatória se
$$X^{-1}(B) \in \mathcal{F} \text{ para todo } B \in \mathcal{B}$$

"X é mensurável com respeito à $\mathcal{F}$". "$X \in \mathcal{F}$".

Se $X: (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$ é uma v.a. tenho uma medida induzida em $(\mathbb{R}, \mathcal{B})$.

$$\mu(B) = P(X^{-1}(B)) \ , \ B \in \mathcal{B}$$

Exercício. Mostre que μ é uma medida em $(\mathbb{R}, \mathcal{B})$.

Obs. Verificar

- $X^{-1}(A^c) = [X^{-1}(A)]^c$
- $X^{-1}(\cap A_i) = \cap X^{-1}(A_i)$
- $X^{-1}(\cup A_i) = \cup X^{-1}(A_i)$

Definição: Função distribuição de X

$$X: (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$$

$$F_X: \mathbb{R} \to \mathbb{R} \ , \ F_X(x) = P(X \leq x)$$

Propriedades de F_X

1) F_X é não-decrescente

2) $\lim_{x \downarrow -\infty} F_X(x) = 0$

$\lim_{x \uparrow \infty} F_X(x) = 1$

3) F_X é contínua à direita

$$\lim_{y \downarrow x} F_X(y) = F_X(x)$$

4) F_X tem limite à esquerda

$$F_X(x^-) = \lim_{y \uparrow x} F_X(y) \underset{\uparrow}{=} P(X < x)$$

5) $P(X = x) = F_X(x) - F_X(x^-)$

Exercício. Dada F satisfazendo 1), 2), 3) e 4), construa

$$X : ([0,1], \mathcal{B}_{[0,1]}) \to (\mathbb{R}, \mathcal{B})$$

com esta distribuição.

Exemplo: X com $F(x) = \begin{cases} 0 & , \text{ se } x < 0 \\ 1 - e^{-\lambda x} & , \text{ se } x \geq 0 \end{cases}$

$X = -\frac{1}{\lambda} \log U$ com $U \sim U(0,1)$

(iii) Mostre que

$$X(\omega) = \sup \{ y . F(y) < \omega \}$$

$$X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$$
$$(S, \mathcal{S})$$

Num contexto um pouco mais geral, se

$$X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$$

X é dito mensurável de $(\Omega, \mathcal{F})$ em $(S, \mathcal{S})$ se

$$X^{-1}(B) \in \mathcal{F}, \forall B \in \mathcal{S}$$

Obs: Se $S = \mathbb{R}^d$, X é dito um vetor aleatório.

Lema: Se $X^{-1}(A) \in \mathcal{F}$ para todo $A \in \mathcal{A}$ tal que $\sigma(\mathcal{A}) = \mathcal{S}$, então X é uma v.a. $(X \in \mathcal{F})$.

Dem. (Exercício)

Lema: Se $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ e $f : (S, \mathcal{S}) \to (T, \mathcal{T})$ são mensuráveis então

$$f(X) : (\Omega, \mathcal{F}) \to (T, \mathcal{T})$$

é mensurável.

Dem. (Exercício)

Lema. Se $X_1, X_2, \ldots, X_n$ são v.a. e

$$f : (\mathbb{R}^n, \mathcal{B}_n) \to (\mathbb{R}, \mathcal{B})$$

é mensurável, então $f(X_1, \dots, X_n)$

$$f : (\mathbb{R}^n, \mathcal{B}_n) \to (\mathbb{R}, \mathcal{B})$$

é mensurável. (é uma v.a.)

Demonstração. (Exercício).

Obs. Se $X_1, X_2, \dots$ são v.a. então

$$\left.\begin{array}{l} \inf_n X_n \\ \sup_n X_n \\ \underline{\lim}_n X_n \\ \overline{\lim}_n X_n \end{array}\right\} \text{ são variáveis aleatórias,}$$

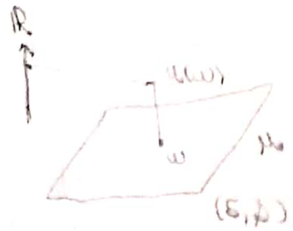
$$\underline{\lim}_n X_n = \sup_n \left(\inf_{m \geq n} X_m\right)$$

Def. Seja $\Omega_0 = \{\omega \in \Omega : \overline{\lim} X_n = \underline{\lim} X_n\} =$

$= \{\omega : \lim X_n \text{ existe}\} \in \mathcal{F}$

Se $P(\Omega_0) = 1$ dizemos que $\{X_n\}_{n \geq 1}$ converge quase certamente (q.c.)

Integração



Suponha μ seja uma medida σ-finita em $(\Omega, \mathcal{F})$, ou seja, existe $A_n \uparrow \Omega$ tal que $\mu(A_n) < \infty \ \forall n$.

Se $f : (\Omega, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B}, dx)$

medida de <u>Lebesgue</u>

Queremos definir

$$\int f \, d\mu$$

Obs. $\int_A f \, d\mu = \int f \cdot \mathbb{1}_A \, d\mu$, $A \in \mathcal{F}$ (notação)

Em 4 passos:

1) funções simples
2) funções limitadas
3) funções não-negativas
4) funções integráveis

Passo 1: φ é dita uma função simples se

$$\varphi(\omega) = \sum_{\substack{i=1 \\ a_i \in \mathbb{R}}}^{n} a_i \, \mathbb{1}_{A_i}(\omega)$$

com $A_1, A_2, \dots, A_n$ disjuntos e $\mu(A_i) < \infty$.

Obs: $\mathbb{1}_A : \Omega \to \mathbb{R}$, $\mathbb{1}_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \text{c.c.} \end{cases}$

Def: $\int \varphi \, d\mu = \sum_{i=1}^{n} a_i \, \mu(A_i)$

Obs: verifique que não depende da representação (a_i não necessariamente distintos)

Def: $\varphi \geq \psi$ se e somente se

$$\varphi(x) \geq \psi(x) \quad \text{q.c. } \mu$$

$$\mu\{x : \varphi(x) < \psi(x)\} = 0$$

Lema. Se φ e ψ são funções simples.

1. Se $\varphi \geq 0$ então $\int \varphi \, d\mu \geq 0$

2. $\int a\varphi \, d\mu = a \int \varphi \, d\mu$, $a \in \mathbb{R}$

3. $\int (\varphi + \psi) \, d\mu = \int \varphi \, d\mu + \int \psi \, d\mu$

Dem: 1) e 2) imediatos.

3. $\varphi = \sum_{i=1}^{m} a_i \mathbb{1}_{A_i}$, $A = \bigcup A_i$

$\psi = \sum_{j=1}^{n} b_j \mathbb{1}_{B_j}$, $B = \bigcup B_j$

$a_0 = 0$ em $B - A = A_0$

$b_0 = 0$ em $A - B = B_0$

$$\varphi + \psi = \sum_{i=0}^{m} \sum_{j=0}^{n} (a_i + b_j) \mathbb{1}_{A_i \cap B_j} \quad \text{é simples}$$

$A_i \cap B_j$ disjuntos

$$\int (\varphi + \psi)\, d\mu = \sum_{i=0}^{m} \sum_{j=0}^{n} (a_i + b_j)\, \mu(A_i \cap B_j) =$$

$$= \sum_{i=0}^{m} a_i \sum_{j=0}^{n} \mu(A_i \cap B_j) + \sum_{j=0}^{n} b_j \sum_{i=0}^{m} \mu(A_i \cap B_j) =$$

$$= \int \varphi\, d\mu + \int \psi\, d\mu. \qquad \square$$

Lema. Se ①, ② e ③ são verdadeiros segue que

4) $\varphi \leq \psi$ q.c. $\Rightarrow \int \varphi\, d\mu \leq \int \psi\, d\mu$

5) $\varphi = \psi$ q.c. $\Rightarrow \int \varphi\, d\mu = \int \psi\, d\mu$

6) $\left| \int \varphi\, d\mu \right| \leq \int |\varphi|\, d\mu$

Dem: $\psi - \varphi \geq 0$ q.c.

$$\int (\psi - \varphi)\, d\mu \geq 0$$

$$\psi = \varphi + \underbrace{(\psi - \varphi)}_{\geq 0}$$

③ $\rightarrow \int \psi\, d\mu = \int \varphi\, d\mu + \overbrace{\int (\psi - \varphi)\, d\mu}^{\geq 0} \geq \int \varphi\, d\mu \qquad (4)$

(5) Se $\varphi = \psi$, então $\varphi \geq \psi$ q.c. e $\psi \geq \varphi$ q.c. Daí uso (4) duas vezes

6) $|x| = \max\{x, -x\}$

$$\varphi \leq |\varphi| \Rightarrow \int \varphi \, d\mu \leq \int |\varphi| \, d\mu$$

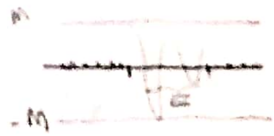
$$-\varphi \leq |\varphi| \Rightarrow \underbrace{\int (-\varphi) \, d\mu}_{-\int \varphi \, d\mu} \leq \int |\varphi| \, d\mu$$

$$\left| \int \varphi \, d\mu \right| = \max\left\{ \int \varphi \, d\mu, -\int \varphi \, d\mu \right\} \leq \int |\varphi| \, d\mu.$$

<u>Passo 2</u>: Funções limitadas

Seja f uma função limitada que se anula fora de um conjunto E com $\mu(E) < \infty$.

Obs: limitada quer dizer que existe M tal que $|f| \leq M$ q.c. μ



<u>Def</u>:

$$\int_1 f \, d\mu = \sup_{\substack{\varphi \leq f \\ \varphi \text{ simples}}} \int \varphi \, d\mu \qquad (1)$$

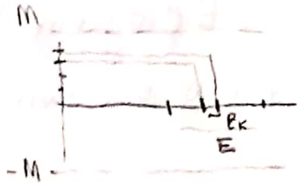
ou

$$\int_2 f \, d\mu = \inf_{\psi \geq f} \int \psi \, d\mu$$

Lema. $\int_1 f \, d\mu = \int_2 f \, d\mu$.

Tenho $\int_1 f \, d\mu \leq \int_2 f \, d\mu$ $\qquad \left.\begin{matrix} \varphi \leq f \\ \psi \geq f \end{matrix}\right\} \Rightarrow \varphi \leq \psi$ q.c.

Basta provar a outra desigualdade



$$E_k = \left\{ x \in E : \frac{(k-1)M}{n} \leq f(x) \leq \frac{kM}{n} \right\}$$

$-n \leq k \leq n$, disjuntos

- Aproximação por cima: $\psi_n(x) = \sum_{k=-n}^{n} \frac{kM}{n} \mathbb{1}_{E_k}$
- " " baixo: $\varphi_n(x) = \sum_{k=-n}^{n} \frac{(k-1)M}{n} \mathbb{1}_{E_k}$

$$\psi_n(x) - \varphi_n(x) = \frac{M}{n} \mathbb{1}_E$$

$$\int (\psi_n - \varphi_n)\, d\mu = \frac{M}{n} \mu(E)$$

$$\varphi_n \leq f \leq \psi_n$$

$$\sup_{\varphi \leq f} \int \varphi\, d\mu \geq \int \varphi_n\, d\mu =$$

$$\varphi_n = \psi_n - \frac{M}{n} \mathbb{1}_E$$

$$= \int \psi_n\, d\mu - \frac{M}{n} \mu(E) \geq \inf_{\psi \geq f} \int \psi\, d\mu - \frac{M}{n} \mu(E)$$ (segue o resultado, basta tomar n adequado.)

Então as duas definições concordam

$$\int f\, d\mu = \sup_{\substack{\varphi \leq f \\ \varphi \text{ simples}}} \int \varphi\, d\mu = \inf_{\substack{\psi \geq f \\ \psi \text{ simples}}} \int \psi\, d\mu$$

Lema. Se $E \subset \mathcal{F}$ com $\mu(E) < \infty$, μ σ-finita

Se f e g são funções limitadas que anulam fora de E, então

1. $f \geq 0$ q.c. $\Rightarrow \int f \, d\mu \geq 0$

2. $\int af \, d\mu = a \int f \, d\mu$

3. $\int (f+g) \, d\mu = \int f \, d\mu + \int g \, d\mu$

4. se $g \leq f$ ~~q.c.~~ $\overset{q.c.}{\Rightarrow} \int g \, d\mu \leq \int f \, d\mu$
5. Se $g = f$ q.c $\Rightarrow \int g \, d\mu = \int f \, d\mu$
6. $\left| \int f \, d\mu \right| \leq \int |f| \, d\mu$

Dem.

1) OK

2) $a\varphi \leq af \Leftrightarrow \varphi \leq f$ se $a > 0$

$a\varphi \leq af \Leftrightarrow \varphi \geq f$ se $a < 0$

3) Se $\varphi_1 \leq f$ e $\varphi_2 \leq g \Rightarrow \varphi_1 + \varphi_2 \leq f+g$

$$\int (f+g) \, d\mu = \sup_{\varphi \leq f+g} \int \varphi \, d\mu$$

$$\geq \sup_{\substack{\varphi_1 \leq f \\ \varphi_2 \leq g}} \int (\varphi_1 + \varphi_2) \, d\mu = \int f$$

(outro lado) desigualdade oposta

0, 13, 16, 23

Durret

$f \to -f \quad g \to -g$

$\int (-f-g) \, d\mu \geq \int (-f) \, d\mu + \int (-g) \, d\mu$

$\int (f+g) \, d\mu \leq \int f \, d\mu + \int g \, d\mu$

Aula 05, Probabilidade Avançada I 18/03

Funções mensuráveis

$$f : (\Omega, \mathcal{F}, P) \to (S, \mathcal{S}, \mu)$$

mensurável se

$$f^{-1}(B) \in \mathcal{F}, \quad \forall B \in \mathcal{S}$$

$\{\omega : f(\omega) \in B\}$

μ induzida por P.

$\int f d\mu$

Integração.

1. Funções simples
2. " limitadas
3. " não-negativas
4. " integráveis

1. $\varphi = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$. A_i disjuntos

$$\int \varphi \, d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$$

2. 

$\mu(E) < \infty$ e $f(x) = 0$ em E^c

f limitada (q.c. μ) $[\exists M \; |f(x)| \leq M]$

$$\int f d\mu = \sup_{\varphi \leq f} \int \varphi \, d\mu = \inf_{\psi \geq f} \int \psi \, d\mu$$

Lema. Se $E \in \mathcal{F}$ tal que $\mu(E) < \infty$, f e g limitadas que se anulam fora de E

1. Se $f \geq 0 \Rightarrow \int f\,d\mu \geq 0$

2. $\int af\,d\mu = a\int f\,d\mu$

3. $\int (f+g)\,d\mu = \int f\,d\mu + \int g\,d\mu$

4. Se $g \leq f \overset{q.c.}{\Rightarrow} \int g\,d\mu \leq \int f\,d\mu$

5. Se $g = f \overset{q.c.}{\Rightarrow} \int g\,d\mu = \int f\,d\mu$

6. $\left|\int f\,d\mu\right| \leq \int |f|\,d\mu$

3. Funções Não-Negativas Quaisquer

$f \geq 0$,

$$\int f\,d\mu = \sup\left\{\int h\,d\mu \,;\, h \geq 0 \text{ limitada com } \mu\{\omega: h(\omega) > 0\} < \infty\right\}$$

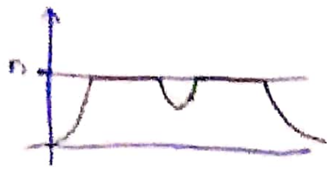
Lema.

E_n

Seja $E_n \uparrow \Omega$ com $\mu(E_n) < \infty$

Notação: $a \wedge b = \min\{a,b\}$
$a \vee b = \max\{a,b\}$

$f_n = f \wedge n$



então

$$\int_{E_n} f_n \, d\mu = \int_{\Omega} f_n \, \mathbb{1}_{E_n} \, d\mu \to \int f \, d\mu$$

Dem. Seq. não decrescente

$$f_n \leq f_{n+1}$$

g limitada ($\exists M$, $|g(x)| \leq M$ q.c.) $\quad \mu(\overbrace{\omega : g(\omega) > 0}^{A}) < \infty$

$$g = g \wedge n \leq f \wedge n$$

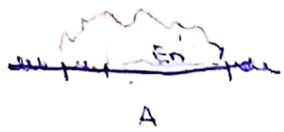
$$\int_{E_n} (f \wedge n) \, d\mu \geq \int_{E_n} (g \wedge n) \, d\mu$$

se tomo $n > M$

$$\int_{E_n} (f \wedge n) \, d\mu \geq \int_{E_n} (g \wedge n) \, d\mu \Rightarrow$$

$$\int_{E_n} (f \wedge n) \, d\mu \geq \int_{E_n} g \, d\mu = \int g \, d\mu - \int_{E_n^c} g \, d\mu$$

$$g \, \mathbb{1}_{E_n} = g - g \, \mathbb{1}_{E_n^c}$$

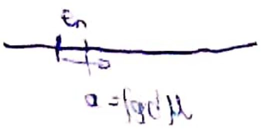


$$0 \leq \int g \, \mathbb{1}_{E_n^c} \, d\mu \leq M \, \mu(A \cap E_n^c) \downarrow \emptyset_{\; n \to \infty}$$

$E_n \uparrow \Omega \Rightarrow E_n^c \downarrow \emptyset$

cont. por baixo de μ

$$\underline{\lim}_{n} \int_{E_n} f \wedge n \, d\mu \geq \int g \, d\mu$$



$a_n \geq a - \varepsilon_n$

$\varepsilon_n = \int_{E_n^c} g_n \, d\mu \downarrow 0$

$a_n = \int_{E_n} f_n \, d\mu$

$\int f \, d\mu = \sup \{ \int h \, d\mu$ h lim. com sup. finito $\}$

$a_n = \int_{E_n} f_n \, d\mu \uparrow \int f \, d\mu$

$$\underline{\lim_n} \int_{E_n} f_n \, d\mu \geq \int g \, d\mu \Rightarrow \underline{\lim_n} \int_{E_n} f_n \, d\mu \geq \int f \, d\mu$$

Da def.

$$\overline{\lim_n} \int_{E_n} f_n \, d\mu \leq \int f \, d\mu$$

$a_n \leq b$

Propriedades.

1.

2.

3.

4.

5.

6.

Dem.

1. V

2. $\int af \, d\mu = a \int f \, d\mu$

$ah \leq af \Leftrightarrow h \leq f$ $\quad a > 0$

$\geq$ $\quad a < 0$ (Não conta pois estamos com f não neg)

3. $\int (f+g) \, d\mu = \sup \{ \int h \, d\mu$, "$h$ limitada e suporte finito", $h \leq f+g \}$

$\geq \sup_{\substack{h \leq f \\ m \leq g}} \{ \int (h+m) \, d\mu$, h, m limitados e sup. finito $\}$. "ang + infinito"

Então $\int (f+g)\, d\mu \geq \int f\, d\mu + \int g\, d\mu$

Des. oposta.

<u>Obs.</u> $(a+b)\wedge n \leq a\wedge n + b\wedge n$

(verifique!)

$$\left\{ \begin{array}{l} (a+b) \geq n \text{ e } a \geq n,\ b \geq n \Rightarrow \\ \quad (a+b)\wedge n = n \text{ e } a\wedge n + b\wedge n = a+b \geq n \end{array} \right.$$

$E_n \uparrow \Omega$

$\mu(E_n) < \infty$

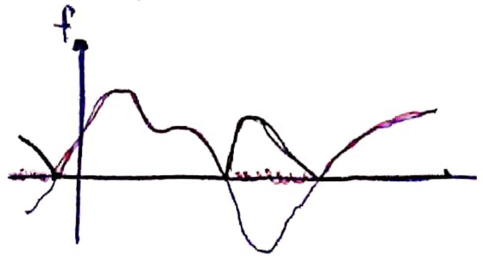
$$\int_{E_n} (f+g)\wedge n\, d\mu \leq \int_{E_n} f\wedge n\, d\mu + \int_{E_n} g\wedge n\, d\mu$$

lado esquerdo $\Rightarrow \int (f+g)\, d\mu$ (lema anterior)

lado direito $\Rightarrow \int f\, d\mu + \int g\, d\mu$

Passo 4: Funções Quaisquer (integráveis)

Def. Dada f qualquer



$f^+ = \max\{0, f\} = f \vee 0$

$f^- = \max\{0, -f\} = (-f)\vee 0$

$$f = f^+ - f^-$$

$$|f| = f^+ + f^-$$

Def. Se f é uma função qualquer, f é dita integrável se

$$\int |f|\, d\mu < \infty$$

~~Lema~~ **Def.** Se f é uma função integrável qualquer

$$\int f\,d\mu = \int f^+ d\mu - \int f^- d\mu$$

Lema. Se f e g são não-negativas com $\int f\,d\mu < \infty$ e $\int g\,d\mu < \infty$

$$\int (f-g)\,d\mu = \int f\,d\mu - \int g\,d\mu$$

Dem: Use

$$f + (f-g)^- = g + (f-g)^+$$

$f \geq g \rightarrow$

$f \leq g \rightarrow$

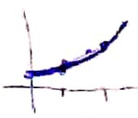
Teorema. valem as 6 propriedades

Dem: Exercícios.

Algumas Propriedades da Integral

Desigualdade de Jensen

Def. Uma função φ é dita convexa se para todo $\alpha \in (0,1)$



Pego 2 ptos da função, ~~toda~~ média ponderada deles está acima da função [illegible]

$$\alpha\,\varphi(x) + (1-\alpha)\,\varphi(y) \geq \varphi(\alpha x + (1-\alpha) y)$$

Des. de Jensen: se f e $\varphi(f)$ são integráveis e φ convexa, então

$$\varphi\left(\int f\,d\mu\right) \leq \int \varphi(f)\,d\mu$$

Dem. se $c = \int f\,d\mu$ tomo $l(x) = ax + b$ com $l(c) = \varphi(c)$

$$\varphi(x) \geq l(x) \quad \forall x$$

$$\int \varphi(x)\,d\mu \geq \int l(f)\,d\mu = \int af + b\,d\mu =$$

$$= a\int f\,d\mu + b = l(c) = \varphi(c)$$

Probabilidade Avançada, Aula [redacted] 06

P1 - 22/04 www.vision.ime.usp.br

P2 - 12/06

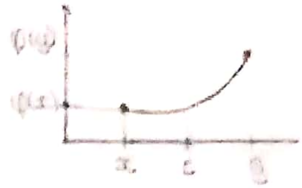
P3 - seg. ou terça seguinte de junho

Estudar algumas desigualdades e propriedades de convergência

Propriedades

. Desigualdade de Jensen

φ convexa, $\alpha \in (0,1)$



$$\varphi(\alpha x + (1-\alpha) y) \leq \alpha \varphi(x) + (1-\alpha)\varphi(y)$$

cresce mais rápido conforme anda

$$\lim_{h \downarrow 0} \frac{\varphi(c) - \varphi(c-h)}{h} \leq \lim_{h \downarrow 0} \frac{\varphi(c+h) - \varphi(c)}{h}$$

$$\varphi(c-h) + \varphi(c+h) \geq 2\varphi(c) \Leftrightarrow$$

$$\frac{\varphi(c+h) - \varphi(c)}{h} \geq \frac{\varphi(c) - \varphi(c-h)}{h}$$

$$\int \varphi(f)\, d\mu \geq \int l(f)\, d\mu$$

$$\varphi(x) \geq l(x) = \int (af + b)\, d\mu = a \underbrace{\int f\, d\mu}_{c} + b = l(c) = \varphi(c) = \varphi\left(\int f\, d\mu\right) \quad \square$$

Seja: $\|f\|_p = \left(\int |f|^p \, d\mu \right)^{1/p}$ $\quad 1 \leq p < \infty \quad$ "Norma p"

Desigualdade de Hölder:

Se p e q satisfazem $\frac{1}{p} + \frac{1}{q} = 1$ e f, g são funções quaisquer

$$\int |fg| \, d\mu \leq \|f\|_p \|g\|_q \qquad ②$$

Assumimos que f e g são integráveis

Obs: Caso Particular: Cauchy-Schwarz $(p = q = 2)$

$$\int |fg| \, d\mu \leq \left(\int |f|^2 \, d\mu \right)^{1/2} \left(\int |g|^2 \, d\mu \right)^{1/2}$$

Demonstração:

Se $y > 0$ fixado,

$$\varphi(x) = \varphi_y(x) = \frac{x^p}{p} + \frac{y^q}{q} - xy$$

$$\varphi'(x) = x^{p-1} - y = 0 \Rightarrow x_0 = y^{\frac{1}{p-1}}$$

$$\varphi''(x) = (p-1) x^{p-2}$$

$$\varphi(x_0) = \frac{y^{\frac{p}{p-1}}}{p} + \frac{y^q}{q} - y^{1 + \frac{1}{p-1}} = y^q \Big[\underbrace{\frac{1}{p} + \frac{1}{q}}_{1} - 1 \Big] = 0$$

$$\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p}, \quad q = \frac{p}{p-1}$$

$\varphi(x) \geq 0 \qquad \frac{x^p}{p} + \frac{y^q}{q} \geq xy$

$f \to \bar{f} = \frac{|f|}{\|f\|_p} \qquad \|\bar{f}\|_p = 1$

$g \to \bar{g} = \frac{|g|}{\|g\|_p} \qquad \|\bar{g}\|_q = 1$

- $\int |fg| \, d\mu \leq 1$ se $\|f\|_p = \|g\|_q = 1$

$|fg| \leq \frac{|f|^p}{p} + \frac{|g|^q}{q}$

$$\int |fg| \, d\mu \leq \int \frac{|f|^p}{p} d\mu + \int \frac{|g|^q}{q} d\mu = \frac{1}{p} + \frac{1}{q} = 1$$

$\|f\|_p = 1, \ \|g\|_q = 1$

se $f : \|f\|_p < \infty$ $\quad fg$ é integrável
e $g . \|g\|_q < \infty$

Quando vale

$$f_n \to f \Rightarrow \int f_n \, d\mu \to \int f \, d\mu \qquad (?)$$

Def. $f_n \to f$ q.o. μ

$\mu(\omega : f_n(\omega) \to f) = 1$

Def. $f_n \to f$ em medida. Dado $\varepsilon > 0$

$$\lim_{n\uparrow\infty} \mu(\omega : |f_n(\omega) - f(\omega)| > \varepsilon) = 0$$

Teorema da Convergência Limitada.

$f_n \to f$ em medida

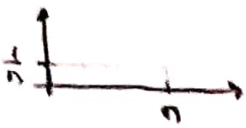
tal que f_n limitada e se anula fora de um conjunto E com $\mu(E) < \infty$.

Então: $\lim_{n\uparrow\infty} \int f_n \, d\mu = \int \lim_{n\uparrow\infty} f_n \, d\mu = \int f \, d\mu$.

Obs: "Contra-exemplo"

1) $(\mathbb{R}, \mathcal{B}, d_x)$

med. de Lebesgue



$f_n \to 0$ em medida

2) f_n

O problema aqui é que $\mu(E)$ não é finito.

Demonstração:

$$\left| \int f \, d\mu - \int f_n \, d\mu \right| = \left| \int (f - f_n) \, d\mu \right| \leq \int |f - f_n| \, d\mu$$

Dado $\varepsilon > 0$

$E_n = \{x : |f_n(x) - f(x)| < \varepsilon\}$

Então

f_n limitado

$\Rightarrow \exists M$ tal que $|f_n| \leq M$

$$= \int_{E_n \cap E} + \int_{E_n^c \cap E} + \int_{E^c} \overset{0}{}$$

$$\leq \varepsilon . \mu(E) + m \, \underbrace{\mu(E - E_n)}_{\to 0} \qquad \blacksquare$$

Lema de Fatou Se $f_n \geq 0$, então

$$\varliminf_n \int f_n \, d\mu \geq \int \varliminf_n f_n \, d\mu$$

Demonstração: Seja $g_n = \inf_{m \geq n} f_m$. Daí

$$g_n \leq f_n \quad e \quad g_n \uparrow g = \varliminf_n f_n$$

$$\int f_n \, d\mu \geq \int g_n \, d\mu$$

Basta mostrar agora que

$$\varliminf_n \int g_n \, d\mu \geq \int g \, d\mu \qquad (*)$$

pois teríamos

$$\varliminf_n \int f_n \, d\mu \geq \int g_n d\mu \geq \varliminf_n \int g_n \, d\mu \geq \int g \, d\mu = \int \varliminf_n f_n \, d\mu$$

Seja $E_m \uparrow \Omega$, $\mu(E_m) < \infty$

$$(g_n \wedge m) \mathbb{1}_{E_m} \xrightarrow{n \uparrow \infty} (g \wedge m) \mathbb{1}_{E_m}$$

$$g_n \geq (g_n \wedge m) . \mathbb{1}_{E_m} \xrightarrow[n\uparrow\infty]{} (g \wedge m)\, \mathbb{1}_{E_m}$$

$$\underline{\lim_n} \int g_n \, d\mu \geq \underline{\lim_n} \int_{E_m} (g_n \wedge m)\, d\mu =$$

$$\overset{TLC}{=} \int_{E_m} (g \wedge m)\, d\mu \xrightarrow[m\uparrow\infty]{(*)} \int g \, d\mu$$

(*) Usa o Lema

$$\int f \, d\mu = \lim_{n\uparrow\infty} \int_{E_n} (f \wedge n)\, d\mu \quad , E_n \uparrow \Omega \quad , (\mu(E_n) < \infty)$$

Teorema da convergência monótona

Se $f_n \geq 0$ e $f_n \uparrow f$ então

$$\int f_n \, d\mu \to \int f \, d\mu$$

Demonstração:

$$f_n \leq f$$

$$\int f_n \, d\mu \leq \int f \, d\mu$$

$$\overline{\lim_n} \int f_n \, d\mu \leq \int f \, d\mu$$

Desigualdade oposta: Lema de Fatou (direto)

Teorema da Convergência Dominada

$f_n \to f$ q.c. e $|f_n| \leq g$, g integrável

então

$$\int f_n \, d\mu \to \int f \, d\mu$$

Demonstração:

$$-g \leq f_n \leq g \Rightarrow f_n + g \geq 0$$

Lema de Fatou

$$\underline{\lim} \int (f_n + g) \, d\mu \geq \int (f+g) \, d\mu \Rightarrow \text{lineari. do int.}$$

$$\underline{\lim} \int f_n \, d\mu \geq \int f \, d\mu$$

Por outro lado, vejo que

$$g - f_n \geq 0$$

$$\underline{\lim} \int (-f_n + g) \, d\mu \geq \int (-f + g) \, d\mu \Leftrightarrow$$

$$\overline{\lim} \int f_n \, d\mu \leq \int f \, d\mu$$

$\square$

O $\underline{\lim}$ de uma sequência numérica é o limsup do conjunto com os limites dos termos tomados: ...

Desigualdade de Chebyshev (Markov)

Se $\varphi(x) \geq 0$

$i_A = \inf \{\varphi(x), x \in A\}$

$i_A P(X \in A) \leq E(\varphi(X)/A) \leq E\varphi(X)$

Se X é uma v.a. e $a > 0$

$$P(|X| \geq a) \leq \frac{EX^2}{a^2}$$

Notação: $E[X/A] = \int X \mathbb{1}_A \, d\mu = \int_A X \, d\mu$

$$i_A \mathbb{1}_A \leq \varphi(X)\mathbb{1}_{\{X\in A\}} \leq \varphi(X)$$

Tomando a esperança,

$$i_A P(X\in A) \leq E(\varphi(X)/A) \leq E(\varphi(X))$$

Basta tomar $\varphi(x) = x^2$ e temos a desigualdade de Markov

$a^2 P(|X| \geq a) \leq EX^2 \Rightarrow$

$$P(|X| \geq a) \leq \frac{EX^2}{a^2}$$

Probabilidade Avançada, Aula 07 25/03/2014

"Antes de falar de independência..."

Produto de Espaços Mensuráveis

Medidas Produto

Sejam $(\Omega_1, \mathcal{F}_1, \mu_1)$ e $(\Omega_2, \mathcal{F}_2, \mu_2)$ dois espaços de medida.

Defino o espaço produto $(\Omega, \mathcal{F}, \mu)$, onde

1. $\Omega = \Omega_1 \times \Omega_2 = \{\omega = (\omega_1, \omega_2), \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$

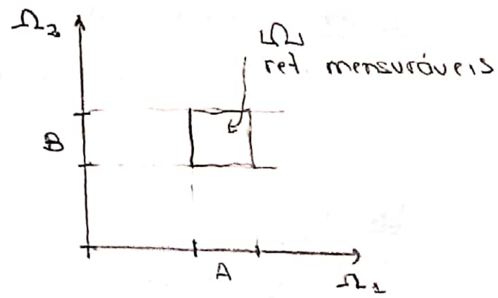
2. $\mathcal{F} = \sigma(\phi) = \mathcal{F}_1 \times \mathcal{F}_2$

   Tomo $\phi = \{E = A \times B;\ A \in \mathcal{F}_1 \text{ e } B \in \mathcal{F}_2\}$.

   ϕ é uma semi-álgebra.

   $\mathcal{F} = \sigma(\phi) \underset{\uparrow}{=} \mathcal{F}_1 \times \mathcal{F}_2$

   notação "cuidado"



3. Em $(\Omega, \mathcal{F})$ defino μ tal que

$$\mu(A \times B) = \mu_1(A)\mu_2(B)$$

μ: extensão de μ em ϕ para $\mathcal{F}$.

Teorema de Fubini. $f : (\Omega, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B})$

$f : (\Omega_1 \times \Omega_2, \mathcal{F}, \mu)$

Se $f \geq 0$ ou $\int |f| d\mu < \infty$

$$\int_\Omega f\,d\mu = \int_{\Omega_1}\left(\int_{\Omega_2} f(x,y)\,\mu_2(dy)\right)\mu_1(dx) =$$

$$= \int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)\,\mu_1(dx)\right)\mu_2(dy)$$

Questão: $h(x) = \int_{\Omega_2} f(x,y)\,\mu_2(dy) \in \mathcal{F}_1$?

$$h : (\Omega_1, \mathcal{F}_1) \to (\mathbb{R}, \mathcal{B})$$

<u>Lema</u> "fatias de $E \in \mathcal{F}$" estão nas σ-álgebras marginais

<u>Lema</u>:

Se $E \in \mathcal{F}$,

$$E_x = \{y \in \Omega_2 : (x,y) \in E\}$$

então $E_x \in \mathcal{F}_2 \ \forall x \in \Omega_1$

Demonstração: fixado $x \in \Omega_1$

$$\mathcal{G}_x = \{E \in \mathcal{F} : E_x \in \mathcal{F}_2\}$$

$E = A \times B$ $\quad$ $\mathcal{G}_x$ contém os retângulos

Se $E \in \mathcal{G}_x \Rightarrow E^c \in \mathcal{G}_x$

• Se $E \in \mathcal{G}_x \Rightarrow E_x \in \mathcal{F}_2 \Rightarrow (E_x)^c \in \mathcal{F}_2 \Rightarrow (E^c)_x \in \mathcal{F}_2 \Rightarrow E^c \in \mathcal{G}_x$

• Se $E_1, E_2, \dots$ em $\mathcal{G}_x$ então $\bigcup E_i \in \mathcal{G}_x$

$$\left(\bigcup E_i\right)_x = \bigcup (E_i)_x$$

$\mathcal{G}_x$ é uma σ-álgebra que contém os retângulos
$\mathcal{F}$ " a menor σ-álgebra " " "

$\mathcal{G}_x \supset \mathcal{F}$. Logo $E \in \mathcal{F} \Rightarrow E \in \mathcal{G}_x \Rightarrow E_x \in \mathcal{F}_2$. ∎

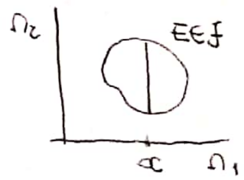
Lema. ~~Se~~ $E \in \mathcal{F}$ seja

$$g(x) = \mu_2(E_x)$$

é mensurável em $\mathcal{F}_1$, e

$$\int_{\Omega_1} g(x)\, \mu_1(dx) = \mu(E)$$

$*$



$h(x) = \int_{\Omega_2} \overset{\mathbb{1}_E}{f(x,y)}\, \mu_2(dy)$

Dem: $\mathcal{G}$ = conj. dos $E \in \mathcal{F}$ tal que $*$ é verdadeiro

- $\mathcal{G}$ é um sistema λ
- $\mathcal{G}$ contém um sistema π, pois $\mathcal{G}$ contém retângulos (∗)
- Se $A \in \mathcal{G} \Rightarrow A^c \in \mathcal{G}$

(•)

Se $E = A \times B$ 

$$g(x) = \begin{cases} \mu_2(B) & \text{se } x \in A \\ 0 & \text{c.c.} \end{cases} =$$

$$= \mu_2(B)\, \mathbb{1}_A \quad \text{mens. em } \mathcal{F}$$

$$\int_{\Omega_1} g(x)\, \mu_1(dx) = \mu_1(A)\, \mu_2(B) = \mu(E)$$

$\mathcal{G}$ é sistema-λ

1) $A \in \mathcal{G} \Rightarrow A^c \in \mathcal{G}$ (ok) (verifique!)

2) Se $A, B \in \mathcal{G}$, $A \supset B \Rightarrow A - B \in \mathcal{G}$

$$g_{A-B}(x) = \mu_2((A-B)_x) = \mu_2(A_x - B_x) =$$

$$= \mu_2(A_x) - \mu_2(B_x)$$

$$\int_{\Omega_1} g_{A-B}(x)\, \mu_1(dx) = \mu(A) - \mu(B) = \mu(A-B).$$

3. Se $A_n \uparrow A$, $A_n \in \mathcal{G} \Rightarrow A \in \mathcal{G}$

$$g_{A_n}(x) = \mu_2(A_n^x)$$

$$\int g_{A_n}(x)\, \mu_1(dx) = \mu(A_n)$$

Se $A_n \uparrow A \Rightarrow (A_n)_x \uparrow A_x$ (verificar)

$$\lim_{n\to\infty} \int g_{A_n}(x)\, \mu_1(dx) = \lim_{n\to\infty} \mu(A_n)$$

$$\downarrow$$

$$\mu(A)$$

converg. monótona

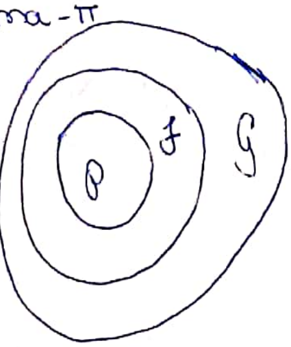
continuidade da med.

$$\lim_{n\uparrow\infty} g_{A_n}(x) = g_A(x)$$

cont. de μ_2

$\mathcal{G}$ é um sistema - λ

$\mathcal{G}$ contém um sistema-π



pelo Teorema λ-π

$\mathcal{G} \supset \mathcal{F}$

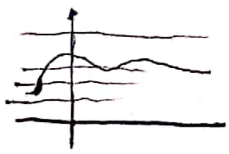
1.

Está mostrado Fubini para $f = \mathbb{1}_E$, $E \in \mathcal{F}$

$$\int f\, d\mu = \mu(E) = \int_{\Omega_1} \left[\int_{\Omega_2} f(x,y)\, \mu_2(dy) \right] \mu_1(dx) = \cdots$$

2. f simples

3. $f \geq 0$



proposta Durret. $f_n(x) = \dfrac{[2^n f(x)]}{2^n} \wedge n$

4. $f = f^+ - f^-$ $\quad |f| = f^+ + f^-$, $\int |f|\, d\mu < \infty$

Independência

Def. Se $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$ são σ-àlgebras, Dizemos que esta coleção é independente se

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = \prod_{i=1}^{n} P(A_i)$$

Para todo $A_1 \in \mathcal{F}_1, \ldots, A_n \in \mathcal{F}_n$.

Def. Se $X_1, X_2, \ldots, X_n$ são v.a. Esta coleção é dita independente se

$$P(X_1 \in B_1, X_2 \in B_2, \ldots, X_n \in B_n) = \prod_{i=1}^{n} P(X_i \in B_i)$$

para todo $B_1 \in \mathcal{B}, B_2 \in \mathcal{B}, \ldots, B_n \in \mathcal{B}$.

Def. Uma coleção de eventos $A_1, A_2, \ldots, A_n$ é dita independente se

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

para todo $I \subset \{1, 2, \ldots, n\}$ = um cjto de índices.

Def. Uma coleção $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ de classes de conjuntos é dito ind. se

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

para todo $I \subset \{1, 2, \ldots, n\}$ e $A_1 \subset \mathcal{A}_1, A_2 \subset \mathcal{A}_2, \ldots, A_n \subset \mathcal{A}_n$

Def. Alternativa. $\mathcal{A}_1, \dots, \mathcal{A}_n$ é ind. se

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$$

$A_i \in \overline{\mathcal{A}}_i \equiv \mathcal{A}_i \cup \{\Omega\}$.

Teorema: Se $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ é uma coleção de classes de conjuntos independentes e cada $\mathcal{A}_i$ é um sistema-π, então

$$\sigma(\mathcal{A}_1), \dots, \sigma(\mathcal{A}_n) \text{ é independente}$$

Exercício: Encontre um contra-exemplo.

Dem: $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ ind.

$$\Downarrow$$

$\sigma(\mathcal{A}_1), \mathcal{A}_2, \dots, \mathcal{A}_n$ ind.

Fixe $A_2 \in \mathcal{A}_2, \dots, A_n \in \mathcal{A}_n$ e seja $F = \bigcap_{i=2}^{n} A_i$

$$\mathcal{G}_F = \{A \subset \Omega : P(A \cap F) = P(A)P(F)\}$$

$\mathcal{G}_F \supset \mathcal{A}_1$ por hipótese

$\mathcal{G}_F$ é um sistema λ

1. $\Omega \in \mathcal{G}_F$ ✓
2. $A, B \in \mathcal{G}_F$ com $A \subset B \Rightarrow B - A \in \mathcal{G}_F$

$$P((B-A) \cap F) = P(B \cap F - A \cap F) = P(B \cap F) - P(A \cap F) = P(B)P(F) - P(A)P(F) =$$
$$= P(B-A)P(F)$$

3. $B_n \uparrow B$, $B_n \in \mathcal{G}_F \Rightarrow B \in \mathcal{G}_F$

Se $B_n \uparrow B \Rightarrow B_n \cap F \uparrow B \cap F$

$$P(B_n \cap F) = P(B_n)\,P(F)$$

cont. da prob. $\downarrow$ $\qquad\qquad \downarrow$

$$P(B \cap F) = P(B)\,P(F)$$

Então $\mathcal{G}_F \supset \sigma(\mathcal{A}_1)$

Então se $A_1 \in \sigma(\mathcal{A}_1) \Rightarrow A_1 \in \mathcal{G}_F$

$$P(A_1 \cap F) = P(A_1)\,P(F)$$

$$P\left(A_1 \cap \left(\bigcap_{i=2}^{n} A_i\right)\right) = \prod_{i=1}^{n} P(A_i) \quad \forall A_1 \in \sigma(\mathcal{A}_1)$$

$$A_i \in \mathcal{A}_i,\ 2 \le i \le n$$

. Itero o argumento

Probabilidade Avançada, Aula 08 27/03

Corolário. Se $X_1, X_2, \dots, X_n$ são v.a. e, para toda coleção $x_1, x_2, \dots, x_n$ em $\mathbb{R} \cup \{\infty\}$ vale

$$P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) = \prod_{i=1}^{n} P(X_i \leq x_i)$$

Então $X_1, X_2, \dots, X_n$ são independentes.

Dem. (Idéia)

- Teorema aula passada e, $\mathcal{A}_i = \{A = (-\infty, x_i], x_i \in \mathbb{R} \cup \{\infty\}\}$
- Uso o exercício

  se $\mathcal{A}$ gera $\mathcal{B}$ $(\sigma(\mathcal{A}) = \mathcal{B})$ então $X^{-1}(\mathcal{A})$ gera $\sigma(X)$.

$\sigma(X) =$ "só sei o valor de x"

$X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$

$\sigma(X) =$ menor σ-álgebra $\mathcal{F}$ tal que X é mensurável.

Obs: $\mathcal{F} =$ "informação" $E[X / \mathcal{F}]$

$\{\emptyset, \Omega\}$ - "total ignorância"

$\{\emptyset, A, A^c, \Omega\}$ = "só sei se A aconteceu ou não"

$P(\Omega)$: "total conhecimento"

Lemas de Borel-Cantelli

Se $A_1, A_2, \dots$ sequência de eventos

$$\overline{\lim}\, A_i = \bigcap_n \bigcup_{m \geq n} A_m$$

i.o.

$\{\omega \in \overline{\lim}_n\, A_i$ se $\omega \in A_i$ infinitas vezes$\}$

$$\underline{\lim}\, A_i = \bigcup_n \bigcap_{m \geq n} A_i$$

$\{\omega \in \underline{\lim}\, A_i$ se $\omega \in A_i$ "eventualmente"$\}$

$\omega \in A_i$ exceto por uma coleção finita de A_i

Obs:

$$\left.\begin{array}{l} \underline{\lim}\, \mathbb{1}_{A_i} = \mathbb{1}_{\underline{\lim} A_i} \\ \\ \overline{\lim}\, \mathbb{1}_{A_i} = \mathbb{1}_{\overline{\lim} A_i} \end{array}\right\} \text{verificar}$$

Notação: $\overline{\lim}\, A_i = \{A_i \text{ i.o.}\}$

Primeiro Lema de Borel-Cantelli

Se $A_1, A_2, \dots$ são tais que $\sum_{i=1}^{\infty} P(A_i) < \infty$ então

$$P(\overline{\lim}\, A_i) = 0$$

Dem. $\overline{\lim}\, A_i = \bigcap_n \underbrace{\bigcup_{m \geq n} A_m}_{B_n}$

$$0 \leq P(\overline{\lim} A_i) \leq P(B_n) = P(\bigcup_{m \geq n} A_m) \overset{\text{subad}}{\leq} \sum_{m \geq n} P(A_m) \downarrow_{n \uparrow \infty} 0$$

$\downarrow$

cauda da série $\sum_{i=1}^{\infty} P(A_i)$

Exemplo de Aplicação

Convergência em prob. e conv. quase certa

Lema: Se $X_n \to X$ em prob., então existe subsequência

$X_{n_1}, X_{n_2}, \dots$ tal que

$X_{n_k} \to X$ q.c.

$\varepsilon > 0$

$$P(|X_n - X| > \varepsilon) \to 0$$

$\varepsilon_k \downarrow 0$ com $k \uparrow \infty$

$\{n_k\}_{k \geq 1}$ $\quad \delta_k = 1/2^k$

n_1 ; $n \geq 1$ tal que

$$P(|X_{n_1} - X| > \varepsilon_1) \leq \overset{\delta_1}{1/2}$$

$n_2 > n_1$

$$P(|X_{n_2} - X| > \varepsilon_2) \leq \overset{\delta_2}{1/2^2}$$

$A_i = \{|X_{n_i} - X| > \varepsilon_i\}$, $\sum P(A_i) < \infty$

$$P(A_i \text{ i.o.}) = 0 \qquad \square$$

2. Versão Fraca da Lei Forte dos Grandes Números

Teorema: Se $X_1, X_2, \dots$, i.i.d. e $EX_i = \mu$ e $EX_i^4 < \infty$, então

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu \quad \text{q.c.}$$

Dem.

$X_i - \mu$ posso tomar $\mu = 0$ (sem perda)

Se $S_n = \sum_{i=1}^{n} X_i$

$$ES_n^4 = E\left(\sum_{i=1}^{n} X_i\right)^4 = E(\ )(\ )(\ )(\ )$$

termos $\quad EX_i^2 X_j X_k \quad EX_i^2 X_j^2$

$E X_i X_j X_k X_l$

$EX_i^3 X_j$

$E X_i^4$

Uso: Se $X_1, \dots, X_n$ são v.a. ind. ($X_i \geq 0$ ou $E|X_i| < \infty$) então

$$EX_1 X_2 \dots X_n = \prod_{i=1}^{n} EX_i$$

Daí

$$ES_n^4 = nEX_i^4 + 3n(n-1)(EX_i^2)^2 \leq n^2 \underset{\uparrow \text{ cte } c<\infty}{C}$$

Logo,

$$P(|S_n| \geq n\varepsilon) \leq \frac{ES_n^4}{(n\varepsilon)^4} \leq \frac{c}{n^2\varepsilon^4} \quad \text{e} \quad \sum_{n\geq 1} P(A_n) < \infty$$

Assim,

$$\frac{S_n}{n} \to 0 \quad q.c.$$

Segundo Lema de Borel-Cantelli

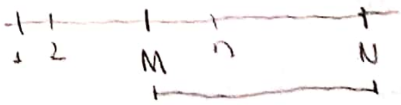
(Borges, biblioteca de 1 livro de areia, macaco digitando)

Se $A_1, A_2, \ldots$ é uma de eventos independentes tal que

$$\sum P(A_i) = \infty \quad , \text{ então}$$

$P(A_i \text{ i.o.}) = 1.$

Dem: $\overline{\lim} A_n = \bigcap_n \bigcup_{m \geq n} A_m$



$$\left(\bigcup_{m=M}^{N} A_m \right)^c = \bigcap_{m=M}^{N} A_m^c$$

$$P\left(\uparrow \right) = P\left(\bigcap_{n=M}^{N} A_m^c \right) = \prod_{m=M}^{N} P(A_m^c) =$$

$$= \prod_{m=M}^{N} [1 - P(A_m)] \overset{*}{\leq} \prod_{m=M}^{N} e^{-P(A_m)} = e^{-\sum_{m=M}^{N} P(A_m)} \underset{N \uparrow \infty}{\downarrow} 0$$

$* \; 1 - x \leq e^{-x}$

Logo $P(\overline{\lim} A_n) = 1$.

Fórmula de Mudança de Variáveis

$X : (\Omega, \mathcal{F}, P) \to (S, \phi)$

μ induzida , $\mu = P \circ X^{-1}$

$f : (S, \phi) \to (\mathbb{R}, \mathcal{B})$

$$\begin{tikzcd}
(\Omega, \mathcal{F}, P) \arrow[rr, "f(X)"] \arrow[dr, "X"'] & & (\mathbb{R}, \mathcal{B}) \\
& (S, \phi) \arrow[ur, "f"'] &
\end{tikzcd}$$

μ

$$\int_\Omega f(X)\,dP = \int_S f(y)\,\mu(dy)$$

Demonstração. Passo-a-passo.

1. $f = \mathbb{1}_E \quad E \in \phi$

$$\int f(X)\,d\mathbb{P}$$

$f(X) : (\Omega, \mathcal{F}) \to (S, \phi)$

$$f(X)(\omega) = \begin{cases} 1 & \text{se } X(\omega) \in E \\ 0 & \text{c.c.} \end{cases}$$

$$f(X) = \mathbb{1}_{X^{-1}(E)}$$

$$\int f(X)\,d\mathbb{P} = \int \mathbb{1}_{X^{-1}(E)}\,d\mathbb{P} = P(X^{-1}(E)) = \mu(E) = \int f\,d\mu$$

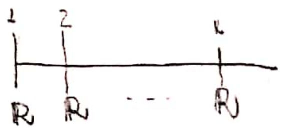
2. Simples
3. $f \geq 0$
4. f integrável

Probabilidade Avançada, Aula 09 (10) 03/04/2013

Teorema de Kolgomorov

Dada μ_K medida em $(\mathbb{R}^K, \mathcal{B}^K)$

$\{\mu_K\}_{K\geq 1}$ consistentes



Problema:

Estender $\{\mu_K\}_{K\geq 1} \to \mu$ em $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}})$

Teo. Kolgomorov

$\{\mu_K\}_{K\geq 1}$ consistente $\Rightarrow \exists$ extensão única

1) $\{\mu_K\}_{K\geq 1} \Rightarrow$ medida em

$\mathcal{S}$ = cilindros = semi-álgebra

Seja $\mathcal{A} = \mathcal{A}(\mathcal{S})$ = cjto de uniões finitas de elementos de $\mathcal{S}$

Lema. Se $S = \bigcup_{i=1}^{n} S_i$, $S, S_i \in \mathcal{S}$, S_i disjuntos

$$\mu(S) = \sum_{i=1}^{n} \mu(S_i)$$

μ_k em $(\mathbb{R}^k, \mathcal{B}^k)$

$$\mu_k([a_1,b_1] \times (a_2,b_2] \times \dots \times (a_k,b_k]) = \prod_{i=1}^{k} [F_{x_i}(b_i) - F_{x_i}(a_i)]$$

$X_1, \dots, X_k$

$$\bar{\mu}_2 \text{ em } (\mathbb{R}^{\mathbb{N}}, \mathcal{A})$$

$$\downarrow$$

$$A(\mathcal{A})$$

$$\downarrow$$

$$\sigma(\mathcal{A})$$

Lema: Se $B_n \in \mathcal{A}$, $B_n \downarrow \emptyset$, então $\mu(B_n) \downarrow 0$

Demonstração:

$$B_n = \bigcup_{k=1}^{k_n} \{\omega = (\omega_1, \omega_2, \dots) \in \mathbb{R}^{\mathbb{N}} : \omega_i \in (a_i^k, b_i^k], 1 \leq i \leq n, -\infty \leq a_i^k < b_i^k \leq \infty\}$$

$$C_n = \bigcup_{k=1}^{k_n} \{\omega : \omega_i \in [\bar{a}_i^k, \bar{b}_i^k], 1 \leq i \leq n\}$$

tal que, para $\delta > 0$

$$\mu(B_n - C_n) \leq \delta / 2^{n+1}$$

* $\bar{a}_i^k$ e $\bar{b}_i^k$ são aproximações por interv. fechados para $(a_i^k, b_i^k]$, que ficam melhores ao interior

$$D_n = \bigcap_{m=1}^{n} C_m$$

$$\mu(B_n - D_n)$$

$$B_n \cap \left(\bigcap_{m=1}^{n} C_m\right)^c = \bigcup_{m=1}^{n} (B_n \cap C_m^c)$$

$1 \leq m \leq n$

$B_m \supset B_n$

$B_n = \bigcap_{m=1}^{n} B_m \Rightarrow B_n \subset B_m$

$$\mu(B_n - D_n) \leq \sum_{m=1}^{n} \mu(B_n - C_m) \leq \sum_{m=1}^{n} \mu(B_m - C_m) \leq \sum_{m=1}^{\infty} \mu(B_m - C_m) \leq$$

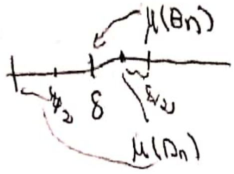
$$\sum_{m=1}^{\infty} \frac{\delta}{2^{m+1}} = \frac{\delta}{2}$$

$B_n \downarrow \emptyset$

Suponha, por absurdo, que

$B_n \downarrow \emptyset$ mas $P(B_n) \downarrow \delta > 0$

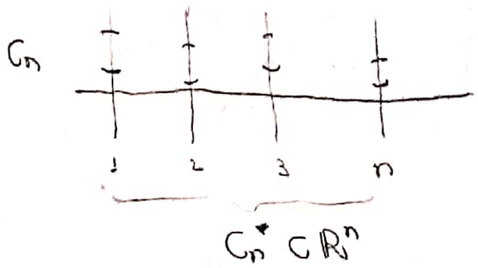
se $\mu(B_n) \to \delta > 0$ então $\underline{\lim}\, \mu(D_n) \geq \frac{\delta}{2}$



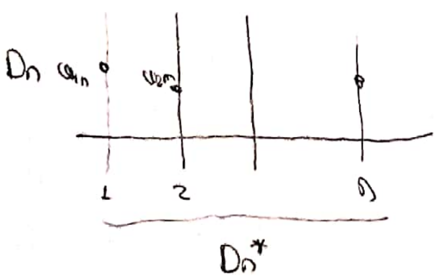
$B_n \downarrow \emptyset$

$D_n \subset B_n$

compacto



$\mathbb{R} \times \mathbb{R} \times$



$C_n = C_n^* \times \mathbb{R} \times \mathbb{R} \times \ldots$

$D_n = \bigcap_{m=1}^{n} C_m$

$C_n^* \subset \mathbb{R}^n \quad D_n^*$

$D_n^* = C_n^* \cap (C_{n-1}^* \times \mathbb{R}) \cap (C_{n-2}^* \times \mathbb{R}^2) \ldots \times (C_1^* \times \mathbb{R}^{n-1})$

$\mu(D_n) \geq \delta/2 \quad D_n \neq \emptyset, \forall n$

$\mu(D_n) = \mu_n(D_n^*)$, $D_n^* \neq \emptyset, \forall n$

Para cada $m \geq 1$ escolho

$\omega_m \in D_m = \bigcap_{m \geq 1}^{n} C_m \subset D_1$, $1 \leq m$

$\omega_2 \in D_2$

$\omega_m = (\omega_{1m}, \omega_{2m}, \dots) \in \mathbb{R}^{\mathbb{N}}$

$\omega_m = (\omega_{1m}, \omega_{2m}, \dots)$

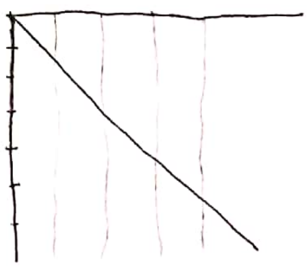
$\omega_{m,1} \in D_1^*$, $\forall m \geq 1$

primeira comp. D_1^* compacto

$\Rightarrow \exists$ uma subseq. convergente

sub seq. $m(1,j) \in \mathbb{N}$

$\omega_{m(1,j),1} \to \theta_1 \in D_1^*$



$\omega_{im} \in D_i^*$, $m \geq i$

$\{m(1,j)\}_{j \geq 1}$ sub-seq do $\mathbb{N}$

$\{m(2,j)\}_{j \geq 1}$ sub-seq de $\{m(1,j)\}_{j \geq 1}$

$(\omega_{m(2,j),1}, \omega_{m(2,j),2}) \in D_2^*$

compacto de $\mathbb{R}^2$

(θ_1, θ_2)

Escolho $\{m(2,j)\}_{j \geq 1}$ t.q. $\omega_{m(2,j),2} \to \theta_2$

Itero, para determinar

$$\omega_{m(k,j),k} \xrightarrow[j\uparrow\infty]{} \theta_k \quad \{m(k,j)\}_{j\geq 1} = \text{subseq de } \{m(k-1,j)\}_{j\geq 1}$$

$$(\omega_{m(i,j),N} \to \theta_i, \ i \leq k)$$

$$\omega_{m(2,j),1} \to \theta_1$$

$$\omega_{m(1,j),1} \to \theta_1$$

$$\{\omega_m\}_{m\geq 1}$$

$$\omega'_k = \omega_{m(k,k)} \qquad \omega'_k \in \mathbb{R}^{\mathbb{N}}$$

$$(\omega'_k)_i = \text{componente } i$$

$$\omega'_{k,i} \to \theta_i$$

$$\theta_1 \in D_1^*$$

$$(\theta_1, \theta_2) \in D_2^*$$

$$\vdots$$

$$\bar{\theta}_n = (\theta_1, \theta_2, \dots, \theta_n) \in D_n^* \subset \mathbb{R}^n$$

$$\theta_n = (\theta_1, \theta_2, \dots, \theta_n, x_{n+1}, x_{n+2}, \dots) \in D_n \subset \mathbb{R}^{\mathbb{N}}$$

$$\theta_n \to \theta = (\theta_1, \theta_2, \dots) \in \mathbb{R}^{\mathbb{N}}$$

$$\theta_n \in D_n \ \forall n$$

$$\Downarrow$$

$$\theta \in \underline{\lim} D_n \ \therefore \ \underline{\lim} D_n \neq \emptyset \Rightarrow \underline{\lim} B_n \neq \emptyset$$

Lei Fraca em L^2

Obs:

1. Conv. em prob.

$$X_n \to X \text{ em prob.}$$

se $\forall \varepsilon > 0$,

$$P(|X_n - X| > \varepsilon) \xrightarrow[n \uparrow \infty]{} 0$$

2. Converg. em L^1

$$X_n \to X \text{ em } L^1 \text{ se } X_n, X \quad E|X_n| < \infty, \; E|X| < \infty$$

se

$$E(|X_n - X|) \xrightarrow[n \uparrow \infty]{} 0$$

L^p $\mathcal{L}^p$

$\left(\int |\varphi|^p \, d\mu\right)^{1/p}$ pode ser 0 quando [illegible]

3. Em L^2

$$X_n \xrightarrow{L^2} X \text{ se } X_n, X \in L^2 \quad EX_n^2 < \infty, \; EX^2 < \infty$$

$$E(X_n - X)^2 \xrightarrow[n \uparrow \infty]{} 0$$

4. Convergência em Distribuição (em Lei)

$$X_n \to X \text{ em dist.}$$

se $F_{X_n}(x) \to F_X(x)$, $\forall x \in \mathbb{R}$ tal que F_X é contínua em x (x = pto de continuidade de F_X)

Algumas Relações

- Conv. Quase Certa $\Rightarrow$ prob.
  em lei

- Cov. em probabilidade $\Rightarrow$ em Lei

- Conv. em L^1 $\Rightarrow$ prob. (Chebyshev)
  em Lei

- Conv. em L^2 $\Rightarrow$ L^1, lei, prob. (Schwarz)

- Se $p > 0$ e $E|X_n|^p \to 0$ então $X_n \to 0$ em prob.

Teorema. (Lei Fraca L^2)

Se $X_1, X_2, \dots$ não correlacionadas (isto é, $EX_iY_j = EX_i\, EY_j$, se $i \neq j$) e $EX_i = \mu$ e $Var(X_i) \leq C < \infty$

então

$$\frac{1}{n} \sum X_i \to \mu \quad \text{em } L^2 \text{ (em prob.)}$$

Dem:

$$E\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right)^2 \xrightarrow[n\uparrow\infty]{} 0$$

Ex. $var(\sum X_i) = \sum var(X_i)$

08.04.2014 PROBABILIDADE AVANÇADA I Aula 10

Lei fraca dos Grandes Números

• Arranjos Triangulares:

$$\left.\begin{matrix} X_{11} & & & \\ X_{21} & X_{22} & & \\ \vdots & & \ddots & \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{matrix}\right\} \quad X_{nk},\ 1 \leq k \leq n$$

$S_n = \sum_{i=1}^{k} X_{ni}$; $\frac{S_n}{n} \longrightarrow ?$

Lei fraca "mais simples": $\{X_{nk},\ 1 \leq k \leq n\}_{n \geq 1}$ com $\sigma_n^2 = Var(S_n)$ finita e $\mu_n = ES_n$.

Se $\sigma_n^2 / b_n^2 \to 0$ $(n \uparrow +\infty)$, então

• $\frac{S_n - \mu_n}{b_n} \xrightarrow{P} 0$

Demonstração: Convergência em L^2

$E\left(\frac{S_n-\mu_n}{b_n}\right) = \frac{Var(S_n)}{b_n^2} \longrightarrow 0$ (condição).

Então, converge em prob.

Exemplo: colecionador de figurinhas

N figurinhas $\{1, 2, \ldots, N\}$

$X_1, X_2, \ldots$ i.i.d com $X_i \sim$ Uniforme em $\{1, 2, \ldots, N\}$

T_k^n: "instante" no qual encontra a k-ésima figurinha "nova"

$\left.\begin{cases} T_0^n \equiv 0 \\ T_1^n = 1 \\ T_2^n = 2 \ldots \end{cases}\right\}$ tomar $X_{nk} = T_k^n - T_{k-1}^n \sim$ Geométrica(p_k) (indep)

onde $p_k = 1 - \frac{k-1}{N}$

Exercício: Se $X \sim Geom(p)$, então $EX = \frac{1}{p}$ e $Var(X) = \frac{1-p}{p^2}$.

$$\mu_n = E T_n^n = \sum_{k=1}^{N} \frac{N}{N-(k-1)} = N \underbrace{\sum_{k=1}^{N} \frac{1}{N-(k-1)}}_{\frac{1}{N}+\frac{1}{N-1}+\dots+\frac{1}{1} \to \infty \text{ quando } N\uparrow\infty} \sim N \log(N)$$

$O(\log(N))$

(comparar c/ $\log(N)$)

$$\sigma_n^2 = Var(T_n^n) \leq \sum_{k=1}^{N} \frac{1}{\left(1-\frac{k-1}{N}\right)^2} = N^2 \sum_{m=1}^{N} \frac{1}{m^2} \leq cN^2$$

(portanto $\frac{\sigma_n^2}{b_n^2} \to 0$)

Tomando $b_n = n\log(n)$: $\frac{T_N^N - N\sum_{m=1}^{N} 1/m}{N \log N} \xrightarrow{P} 0$

- $\frac{N\sum_{m=1}^{N} 1/m}{N\log N} \longrightarrow 1$ (exercício)

- $\frac{T_N^N}{N\log N} \xrightarrow{P} 1$

Lei fraca p/ Arranjos triangulares

Para cada $n \geq 1$, seja $X_{n1}, X_{n2}, \dots, X_{nn}$ independentes

Se $b_n \geq 0$ com $b_n \uparrow \infty$, $\tilde{X}_{nk} = X_{nk} \mathbb{1}\{|X_{nk}| \leq b_n\}$.

v.a. truncada

Suponha: 1) $\sum_{k=1}^{n} P(|X_{nk}| > b_n) \xrightarrow[n\uparrow\infty]{} 0$

2) $\frac{\sum_{k=1}^{n} E\tilde{X}_{nk}^2}{b_n^2} \xrightarrow[n\uparrow\infty]{} 0$.

Se $S_n = \sum_{k=1}^{n} X_{nk}$ e $a_n = \sum_{k=1}^{n} E\tilde{X}_{nk}$, então $\boxed{\frac{S_n - a_n}{b_n} \xrightarrow{P} 0}$

Demonstração: $\tilde{S}_n = \sum_{k=1}^{n} \tilde{X}_{nk}$

$P(A \cap B) \leq P(B)$

$$P\Big(\underbrace{\Big|\frac{S_n - a_n}{b_n}\Big| > \varepsilon}_{A}\Big) = P\big(A \cap \{S_n = \tilde{S}_n\}\big) + P\big(A \cap \{S_n \neq \tilde{S}_n\}\big) \leq$$

$P(A \cap B) \leq P(B)$

$$\leq P\Big(\Big|\frac{\tilde{S}_n - a_n}{b_n}\Big| > \varepsilon\Big) + P(S_n \neq \tilde{S}_n)$$

(II) (I)

(I) $P(S_n \neq \tilde{S}_n) \leq P\Big(\bigcup_{k=1}^{n} \{X_{nk} \neq \tilde{X}_{nk}\}\Big) \leq$

$$\leq \sum_{i=1}^{n} P\{X_{nk} \neq \tilde{X}_{nk}\} \xrightarrow[n \uparrow \infty]{} 0 \; //$$

(II) $P\Big(\frac{|\tilde{S}_n - a_n|}{b_n} > \varepsilon\Big) \leq \frac{E(\tilde{S}_n - a_n)^2}{\varepsilon^2 b_n^2} \leq$

$$\leq \frac{Var(\tilde{S}_n^2)}{\varepsilon^2 b^2} = \frac{\sum_{i=1}^{n} Var(\tilde{X}_{nk}^2)}{\varepsilon^2 b^2} \leq$$

$$\leq \frac{1}{\varepsilon^2 b^2} \sum_{i=1}^{n} E\tilde{X}_{nk}^2 \xrightarrow[n \uparrow \infty]{} 0 \; //$$

$\therefore P(A_n) \xrightarrow[n \uparrow \infty]{} 0 \Rightarrow \frac{S_n - a_n}{b_n} \xrightarrow{P} 0$ ∎

Teorema (Lei Fraca dos Grandes Números)

Se $X_1, X_2 \ldots$ são v.as i.i.d. com $x P(|X_1| > x) \xrightarrow[x \uparrow \infty]{} 0$,

$S_n = \sum_{i=1}^{n} X_i$ e $\mu_n = E(X_1 \mathbb{1}_{\{|X_1| \leq n\}})$.

Então, $\boxed{\frac{S_n}{n} - \mu_n \to 0 \text{ em prob.}}$

Demonstração: $X_{nk} = X_k$, $\forall n$

$b_n = n$

$$① \sum_{k=1}^{n} P(|X_{nk}| > b_n) = nP(|X_1| > n) \xrightarrow[n\to\infty]{} 0$$

X_1 iid ... por hipótese

$$② \frac{1}{b_n^2} \sum_{k=1}^{n} E\tilde{X}_{nk}^2 \xrightarrow[?]{} 0 \Rightarrow \frac{1}{n^2} nE X_1^2 \mathbb{1}\{|X_1| < n\} \longrightarrow 0 \quad ?$$

Lema auxiliar: Seja $Y \geq 0$ e $p > 0$.

- $EY^p = \int_0^\infty p y^{p-1} P(Y > y)\, dy$

↳ Obs: $EY = \int_0^\infty P(Y > y)\, dy \quad (p = 1)$

Demonstração: $\int_0^\infty p y^{p-1} P(Y > y)\, dy =$

$$= \int_0^\infty p y^{p-1} \int_\Omega \mathbb{1}\{Y > y\}\, dP\, dy \quad (*)$$

Z

$$\Rightarrow Z = \int_0^\infty p y^{p-1} \mathbb{1}\{Y(\omega) > y\}\, dy = \int_0^{Y(\omega)} p y^{p-1}\, dy = (Y(\omega))^p$$

$$(*) = \int y^p\, dP = EY^p$$

Voltando a ②

$$\Rightarrow E|\tilde{X}_{n1}|^2 = \int_0^\infty 2y\, P(|\tilde{X}_{n1}| > y)\, dy \leq \quad \text{provar!}$$

$$\leq \int_0^n 2y\, P(|X_1| > y)\, dy$$

- $\frac{1}{n} E|\tilde{X}_{n1}|^2 \leq \int_0^\infty \frac{2y}{n} P(\tilde{X}_{n1} > y)\, dy$

$f(y)$

$f(y) \xrightarrow[y \uparrow \infty]{} 0$. (pois $y P(|X_1| > y) \to 0$ por hipótese)

Seja $M = \sup_{y \in \mathbb{R}^+} \{ f(y) \}$.

Dado $K > 0$, $\varepsilon_K = \sup_{y \geq K} \{ f(y) \}$

$\Rightarrow \frac{1}{n} \int_0^n f(y)\,dy \leq \frac{M \cdot K}{n} + \varepsilon_K \left(1 - \frac{K}{n} \right)$ $\quad \frac{\varepsilon_K (n-K)}{n}$

• $\limsup \left(\frac{1}{n} \int_0^n f(y)\,dy \right) \leq \varepsilon_{K_j}, \ \forall j$.

Se $K_j \uparrow \infty$ com $j \uparrow \infty \Rightarrow \varepsilon_{K_j} \xrightarrow[j \uparrow \infty]{} 0$. ■

Probabilidade Avançada, Aula 11

Teo. (Lei Fraca na Forma Usual)

$X_1, X_2, \dots$ i.i.d. com $E|X_i| < \infty$ e $EX_i = \mu$ então

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu \quad \text{em prob.}$$

Dem.

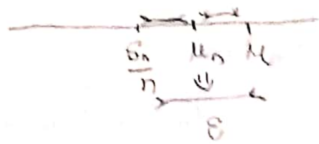
$$\underbrace{E\left[|X_1| \mathbb{1}_{\{|X_1|>x\}}\right]}_{\downarrow \, 0} \geq x P(|X_1| > x) \qquad \cdot\ |X_1| \mathbb{1}_{\{|X_1|>x\}} \geq x \mathbb{1}_{\{|X_1|>x\}}$$

$$\mu_n = E[X_1 \mathbb{1}_{\{|X_1| \leq n\}}] \to \mu$$

Pela lei fraca anterior ("lei fraca truncada")

$$P\left(\left|\frac{S_n}{n} - \mu_n\right| > \delta\right) \to 0 \ \forall \delta > 0 \quad \text{e então}$$

$$P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \xrightarrow[n \uparrow \infty]{} 0$$



Obs: $X_1, X_2, \dots$ i.i.d. com $x P(|X_1| > x) \xrightarrow[x \uparrow \infty]{} 0$.

Tomando n tal que

$$\left\{\left|\frac{S_n}{n} - \mu_n\right| < \frac{\varepsilon}{2}\right\} \cap \left\{|\mu - \mu_n| < \varepsilon/2\right\} \to \left\{\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right\}$$

Paradoxo de Petesburg

$Y \sim$ geométrica $(1/2)$

martingal - cobresto, que

$P(X = 2^i) = \frac{1}{2^i}$ Cobro 2 reais a cada lançamento.

$X = 2^Y$

Daí, $EX = \sum_{i \geq 1} 2^i \frac{1}{2^i} = \infty$. O habitual seria pagar inf. pro jogar, o que não é razoável do ponto de vista fair-play, já que o jogo event. acaba. (Mostra que EX não é tão razoável)

Logo, vamos considerar a compra do número de lançamentos. no caso, n.

Usamos a lei fraca pra arranjos triangulares.

$$X_{nk}, \quad 1 \leq k \leq n \qquad \begin{array}{cccc} X_{11} & & & \\ X_{21} & X_{22} & & \\ \vdots & & & \\ X_{n1} & \cdots & & X_{nn} \end{array} \quad \text{ind.}$$

$b_n \uparrow \infty$

$$\tilde{X}_{nk} = X_{nk} \mathbb{1}_{\{|X_{nk}| \leq b_n\}} \qquad a_n = \sum_{k=1}^{n} E\tilde{X}_{nk}$$

Se

1) $\sum_{k=1}^{n} P(|X_{nk}| > b_n) \to 0$

2) $\frac{1}{b_n^2} \sum_{k=1}^{n} E\tilde{X}_{nk}^2 \to 0$

Então

$$\frac{\sum_{k=1}^{n} X_{nk} - a_n}{b_n} \longrightarrow 0 \quad \text{em prob}$$

~~Tomo~~

$$P(X \geq 2^m) = \sum_{k \geq m} \frac{1}{2^k} = 2^{-m} \sum_{k \geq 0} 2^{k} = 2^{1-m}$$

Tomo

$$b_n = 2^{m(n)} = 2^{\log_2(n) + k(n)} = n\, 2^{k(n)} \quad , \; k(n) \uparrow \infty \text{ com } n \uparrow \infty$$

Cond. 1.

$$n\, P(|X_1| > b_n) = n\, 2^{1-m(n)}$$

$$n\, P(|X_1| > n 2^{k(n)}) \quad \frac{2n}{2^{m(n)}} = \frac{n\, 2}{n\, 2^{k(n)}} = \frac{2}{2^{k(n)}} \xrightarrow[n \uparrow \infty]{} 0$$

Cond. 2.

$$\tilde{X}_{nk} = X_{nk} \mathbb{1}_{\{|X_{nk}| \leq b_n\}}$$

$$\tilde{X}_{nk}^2 = \begin{cases} X_{nk}^2 & \text{se } |X_{nk}| \leq 2^{m(n)} \\ 0 & \text{c.c.} \end{cases}$$

$$E\tilde{X}_{nk}^2 = \sum_{j=1}^{m(n)} (2^j)^2 \frac{1}{2^j} = \sum_{j=1}^{m(n)} 2^j = 2^{m(n)} \left[\frac{1}{2^{m(n)-1}} + \frac{1}{2^{m(n)-2}} + \dots + \frac{1}{2} + 1 \right]$$

$$\leq 2^{m(n)+1} = 2 b_n$$

$$\frac{1}{b_n^2} \sum_{k=1}^{n} E\tilde{X}_{nk}^2 \leq \frac{n\, 2 b_n}{b_n^2} = \frac{2n}{b_n} = \frac{2n}{2^{m(n)}} = \frac{2n}{n\, 2^{k(n)}} \xrightarrow[n \uparrow \infty]{} 0$$

b_n OK

$$a_n = \sum_{k=1}^{n} E\tilde{X}_{nk} \Rightarrow a_n = n\,m(n)$$

$$E\tilde{X}_{nk} = \sum_{j=1}^{m(n)} 2^j \frac{1}{2^j} = m(n)$$

$$\tilde{X}_{nk} = \begin{cases} X_{nk} & \text{se } X_{nk} \leq 2^{m(n)} \\ 0 & \end{cases}$$

$$a_n = n\,m(n) = n[\log n + k(n)] = n\log n + n k(n)$$

$$b_n = n 2^{k(n)}$$

$$\frac{S_n - a_n}{b_n} \to 0 \text{ em prob}$$

$$\frac{a_n}{b_n} = \frac{n\log n + n k(n)}{n 2^{k(n)}}$$

Tomo $k(n) \uparrow \infty$ mais lentamente que $\log(n)$

$$\frac{k(n)}{\log(n)} \xrightarrow[n\uparrow\infty]{} 0 \quad , \quad k(n) = \log\log n$$

$$\frac{S_n - a_n}{n 2^{\log\log n}} = \frac{S_n - a_n}{n\log n} \to 0 \text{ em prob.}$$

$$\frac{a_n}{n 2^{\log\log n}} = \frac{a_n}{n\log n} = \frac{n\log n + n\log\log n}{n\log n} \to 1$$

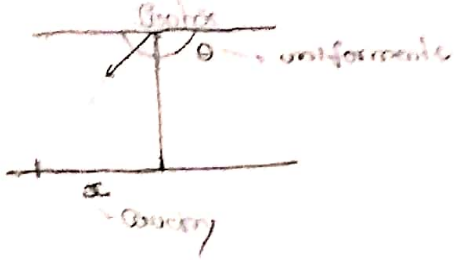
paga 10 p/ jogo 1024 vezes

$$\frac{S_n}{n\log n} \to 1 \text{ em prob.}$$

"Custo para jogar n vezes é $\log n$ por jogada"

Teorema. Sejam $X_1, X_2, \ldots$ v.a. i.i.d. com $E|X_1| = \infty$, então

$$P(|X_n| \geq n \ \text{ i.o.}) = 1$$



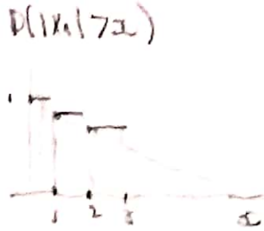
Corolário.

$$A = \left\{ \omega : \text{existe } a \text{ com } \lim_{n \to \infty} \frac{S_n(\omega)}{n} = a \in \mathbb{R} \right\}, \quad P(A) = 0$$

$E|X_1| = \infty$

$$E|X_1| = \int_0^\infty P(|X_1| > x)\, dx \leq$$

$$\sum_{n=0}^{\infty} P(|X_n| > n)$$



Se $E[X_1] = \infty \Rightarrow$

$$\sum P(|X_n| > n) = \infty \quad + \quad \text{Borel + Cantelli} \Rightarrow$$

$$P(|X_n| \geq n \ \text{ i.o.}) = 1$$

Aula 13. Probabilidade Avançada I 06/05/2014

Teorema Limite Central

"Tomar médias → Padronizá-las → vari. aleatória ≈ em geral, $N(0,1)$"

Convergência Fraca.

Dizemos que $\{F_n\}_{n\geq 1}$, sequência de funções distribuições, converge (fracamente) para F (notação: $F_n \Rightarrow F$), se F for distribuição e

$$F_n(x) \to F(x) \text{ para todo ponto de continuidade de } F.$$

Dizemos que $X_1, X_2, \ldots$, seq. de v.a., converge fracamente para X se as correspondentes distribuições convergirem (como no caso anterior)

$F_n(x) = P(X_n \leq x)$ e $F_n \to F$

Obs: Seja X uma v.a.

$$X_n = X + \frac{1}{n}$$

$$F_n(x) = P(X_n \leq x) = P(X + 1/n \leq x) = P(X \leq x - 1/n) = F(x - 1/n)$$

$F_n(x) \to$ limite a esquerda

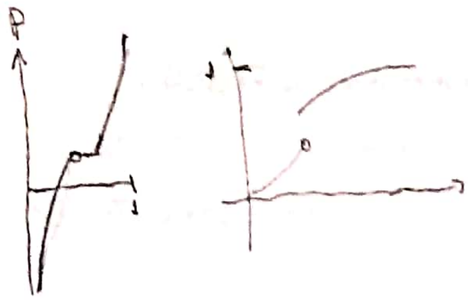
"Random walk insights"
↳ Livro com curiosidades

Obs. Se $F_n \to F$ então posso contruir uma sequencia $Y_n \Rightarrow Y$ com $F_n(x) = P(Y_n \leq x)$ e $F(x) = P(Y \leq x)$.

Dado $F \Rightarrow Y$ com $\Omega = (0,1)$

$\mathcal{F} = \mathcal{B}(0,1)$

$F_n \Rightarrow F$ $\quad$ P = lebesgue



$Y_n(\omega) = \sup\{y : F(y) \leq \omega\}$

Se $a_x = \sup\{y : F(y) < x\}$ e

$b_x = \inf\{y : F(y) > x\}$

$\Omega_0 = \{x : (a_x, b_x) = \emptyset\}$

$\Omega - \Omega_0$ é enumerável. (Exercicio: complete).

$\{F_n\}_{n \geq 1}$ f. dist. de prob.

Se $F_n \to F$ pontualmente, como garantir que F é f. distribuição.

Exemplo.

↑ não é f. d. de prob.

tight (Não podemos deixar a massa escapar)

Definição. Dizemos que uma coleção de funções dist. $\{F_n\}_{n\geq 1}$ é "tight" se, para dado $\varepsilon>0$, existe M_ε tal que

$$1-[F_n(M_\varepsilon)-F_n(-M_\varepsilon)] < \varepsilon \quad \forall\, n\geq 1$$

Teorema. Se $\{F_n\}_{n\geq 0}$ é tight e $F_n \to F$ pontualmente com F não decrescente e contínua à direita, então

F é função distribuição.

Definição: Se X é uma v.a. defino

$$\varphi_X(t) = E\, e^{itX} = E(\cos tX) + i\, E(\operatorname{sen} tX)$$

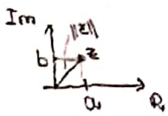
$\uparrow$

função característica de X

Obs: $e^{ix} = \cos x + i \operatorname{sen} x$.

Obs:

$z \in \mathbb{C}$, $z = a+bi$



. complexo conjugado: $z = a+bi \Rightarrow \bar{z} = a-bi$

$\|z\|^2 = a^2+b^2 = z\bar{z}$

Teorema. Propriedades de $\varphi_X(t) = \varphi(t)$

a) $\varphi(0) = 1$

b) $\varphi(t) = \overline{\varphi(-t)}$

c) $|\varphi(t)| = |E\, e^{itX}| \leq E|e^{itX}| = 1$.

d) $|\varphi(t+h)-\varphi(t)| \leq E(e^{itX}-1)$

e) $E e^{it(aX+b)} = e^{itb} \varphi_X(ta)$, i.e., $\varphi_{aX+b}(t) = e^{itb} \varphi_X(ta)$

Demonstração.

1. Imediato

2. $\varphi(-t) = E e^{-itX} = E[\cos(-tX)] + i E[\text{sen}(-tX)] \overset{*}{=}$
$= E[\cos(tX)] - i E[\text{sen}(tX)] =$
$= \overline{\varphi(t)}$

* cos e sen são f. impar e par, respectivamente.

3. Jensen em $\mathbb{R}^2$ com a função convexa $\varphi(x,y) = \sqrt{x^2+y^2}$
e $|e^{ix}| = \cos^2 x + \text{sen}^2 x = 1$

4. $|\varphi(t+h) - \varphi(t)| = |E(e^{i(t+h)X} - e^{itX})| \leq E|e^{i(t+h)X} - e^{itX}| \overset{*}{=}$
$= E(e^{ihX} - 1)$

* $|e^{i(t+h)X} - e^{itX}| = |e^{itX}||e^{ihX} - 1|$

5. Imediato.

Motivação para considerar $\varphi(t)$.

Lema: Se X_1 e X_2 são v.a. independentes, então

$$\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\, \varphi_{X_2}(t).$$

Dem: imediato (ex.)

Exemplos.

a) X tal que $X \in \{-1, 1\}$ com $P(X=1) = p$

$$\varphi_X(t) = E e^{itX} = e^{it} p + (1-p) e^{-it} =$$

$$= e^{-it} + p(e^{it} - e^{-it}) =$$

$$= \cos t - i \operatorname{sen} t + p[\cos t + i \operatorname{sen} t - (\cos t - i \operatorname{sen} t)] =$$

$$= \cos t - i \operatorname{sen} t + 2pi \operatorname{sen} t =$$

$$= \cos t + (2p-1) i \operatorname{sen} t \;//$$

Obs: Se $p = 1/2$, $\varphi(t) \in \mathbb{R}$.

b) Poisson, $X \sim \text{Poisson}(\lambda)$

$$\varphi_X(t) = E e^{itX} = \sum_{k \geq 0} e^{itk} \frac{\lambda^k e^{-\lambda}}{k!} =$$

$$= e^{-\lambda} \sum_{k \geq 0} \frac{[\lambda e^{it}]^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it} - 1)} \;//$$

c) $X \sim N(0,1)$

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \, dx = e^{-t^2/2}$$ verifique!

4) $X \sim U(a,b)$

$$\varphi_X(t) = \frac{e^{itb} - e^{ita}}{it\,(b-a)}$$

5) $X \sim \exp(\lambda)$

$$\varphi_X(t) = \frac{1}{1-it}$$

Teorema da Inversão. Seja

$$\varphi(t) = E\,e^{itX} = \int e^{itx}\,\mu(dx)$$

μ = medida que caracteriza a dist. de X.

então:

$$\mu(a,b) + \frac{\mu(\{a,b\})}{2} = \lim_{T\uparrow\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it}\,\varphi(t)\,dt$$

Dem.

Seja

$$I_T = \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it}\,\varphi(t)\,dt.$$

Então,

$$I_T = \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \int e^{itx}\,\mu(dx)\,dt =$$

$$= \int_{-T}^{T} \int \boxed{\frac{e^{-ita} - e^{-itb}}{it} \, e^{itx}} \, \mu(dx) \, dt$$

$$\left| \quad ? \quad \right| \leq b-a$$

$$\frac{e^{-ita} - e^{-itb}}{it} = \int_a^b e^{-ity} \, dy \qquad \text{verifique}$$

$$\left| \quad ? \quad \right| = \left| \quad ? \quad \right| \leq \int_a^b \underbrace{|e^{-ity}|}_{=1} \, dy = b-a$$

$$I_T = \int \underbrace{\int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \, e^{itx} \, dt}_{} \, \mu(dx)$$

$$\int_{-T}^{T} \frac{e^{it(x-a)} - e^{it(x-b)}}{it} \, dt$$

$$\frac{\cos t(x-a) + i\,\mathrm{sen}\, t(x-a) - \cos t(x-b) - i\,\mathrm{sen}\, t(x-b)}{it} =$$

$$= \underbrace{\frac{\cos t(x-a) - \cos t(x-b)}{it}}_{\text{f. impar}} + \frac{\mathrm{sen}\, t(x-a) - \mathrm{sen}\, t(x-b)}{t}$$

$$I_T = \int \left[\underbrace{\int_{-T}^{T} \frac{\mathrm{sen}\, t(x-a)}{t} dt}_{R(x-a,T)} - \underbrace{\int_{-T}^{T} \frac{\mathrm{sen}\, t(x-b)}{t} dt}_{R(x-b,T)} \right] \mu(dx)$$

Sejam $R(\theta,T) = \int_{-T}^{T} \frac{\mathrm{sen}\, \theta t}{t} \, dt$

$$S(T) = \int_0^T \frac{\mathrm{sen}\, x}{x} \, dx \xrightarrow[T \uparrow \infty]{} \frac{\pi}{2}$$

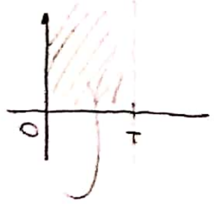
$$R(\theta,T) = \int_{-T}^{T} \frac{\operatorname{sen}\theta t}{t}\, dt \qquad s = \theta t \quad ds = \theta dt$$

$$= \int_{-\theta T}^{+\theta T} \frac{\operatorname{sen} s}{s}\, ds = \begin{cases} 2S(\theta T) & \text{se } \theta > 0 \\ -2S(-\theta T) & \text{se } \theta < 0 \end{cases} , \ 0 \text{ se } \theta = 0.$$

função par → simetria

Obs: $S(T) \underset{T\uparrow\infty}{\rightarrow} \frac{\pi}{2}$

Idéia:



$f(x,y) = e^{-xy} \operatorname{sen} x \qquad y > 0, \ 0 \leq x \leq T$

$\int f(x,y)\, dx\, dy$ Fubini ok ✓

$$\int_0^T \int_0^\infty e^{-xy} \operatorname{sen} x\, dy\, dx = \int_0^T \frac{\operatorname{sen} x}{x}\, dx$$

$$\int_0^\infty e^{-xy} \operatorname{sen} x\, dy = \left(\frac{\operatorname{sen} x\, e^{-xy}}{-x} \right) \Bigg|_0^\infty = \frac{\operatorname{sen} x}{x}$$

$$\int_0^\infty \underbrace{\boxed{\int_0^T e^{-xy} \operatorname{sen} x\, dx}}_{I}\, dy$$

$$I = \int_0^T e^{-xy} \operatorname{sen} x\, dx = (-e^{-xy} \cos x)\Big|_0^T - \int_0^T y\, e^{-xy} \cos x\, dx$$

$$-(e^{-xy} \cos x)' = y\, e^{-xy} \cos x + e^{-xy} \operatorname{sen} x$$

$$= -e^{-Ty}\cos T + 1 - \overbrace{\int_0^T y e^{-xy}\cos x\,dx}^{I_2}$$

$$(y e^{-xy}\,\mathrm{sen}\,x)' = -y^2 e^{-xy}\,\mathrm{sen}\,x + y e^{-xy}\cos x$$

$$I_2 = (y e^{-xy}\,\mathrm{sen}\,x)\Big|_0^T + \int_0^T y^2 e^{-xy}\,\mathrm{sen}\,x\,dx =$$

$$= y e^{-Ty}\,\mathrm{sen}\,T + y^2 I$$

$$I = 1 - e^{-Ty}\cos T - y^2 I - y e^{-Ty}\,\mathrm{sen}\,T$$

$$I(1+y^2) = 1 - e^{-Ty}\cos T - y e^{-Ty}\,\mathrm{sen}\,T$$

$$I = \frac{1}{1+y^2} - \frac{e^{-Ty}\cos T}{1+y^2} - \frac{y e^{-Ty}\,\mathrm{sen}\,T}{1+y^2}$$

$$\int_0^{\frac{\pi}{2}} d\theta \leftarrow \int_0^{\infty} \frac{1}{1+y^2}\,dy \qquad y = tg\,\theta = \frac{\mathrm{sen}\,\theta}{\cos\theta}$$

$$dy = (1+tg^2\theta)\,d\theta = (1+y^2)\,d\theta$$

Aula 14, Probabilidade Avançada I, 08/05/2014

...

Teorema de Inversão

Dado $\varphi(t) \to \mu \quad (F)$

Fórmula de Inversão. Se $\varphi(t) = \int e^{itx} \mu(dx)$

$$\mu(a,b) + \frac{\mu(\{a,b\})}{2} =$$

$$\lim_{T\uparrow\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-iat} - e^{ibt}}{it} \varphi(t)\, dt$$

- Fubini

$$J_T = \int_{-T}^{T} \int (\quad) \, dt \, \mu(dx)$$

- $e^{ix} = \cos x + i \operatorname{sen} x$ + fato de $\cos x$ / x sen ímpar

$$I_T = \int \Big(\underbrace{\int_{-T}^{T} \frac{\operatorname{sen} t(x-a)}{t} dt}_{R(x-a,t)} - \underbrace{\int_{-T}^{T} \frac{\operatorname{sen} t(x-b)}{t} dt}_{R(x-b)} \Big) \mu(dx)$$

- $R(\theta,T) \to 2\,\text{sinal}(\theta)\, S(|\theta| T)$

$$S(T) = \int_0^T \frac{\operatorname{sen} x}{x} dx \to \frac{\pi}{2}$$

- $R(\theta,T) \to \pi \,\text{sinal}(\theta)$

$$I_T = \int (R(x-a,T) - R(x-b,T))\, \mu(dx)$$

$$\xrightarrow[T\uparrow\infty]{} \begin{cases} 0 & \text{se } x \notin [a,b] \\ 2\pi & \text{se } x \in (a,b) \\ \pi & \text{se } x \in \{a,b\} \end{cases}$$

Daí,

$$I_T \to 2\pi\, \mu(a,b) + \pi\, \mu(\{a,b\})$$

$$\frac{1}{2\pi} I_T \to \mu(a,b) + \frac{1}{2}\mu(\{a,b\})$$

Posso usar tal fato para escrever $P(X \in (a,b) \cup \{a,b\}) = \lim_{T\uparrow\infty} \frac{1}{2\pi} I_T$

Brinco com a e b para descobrir se $\mu\{a,b\}$ é significante.

Relação entre convergência de funções características e convergência fraca

Teorema: Sejam $\{\mu_n\}_{n\geq 1}$, $\mu = \mu_\infty$ e $\{\varphi_n\}_{n\geq 1}$ e $\varphi = \varphi_\infty$

$$\begin{array}{ccc} \mu_n & \longrightarrow & \mu \\ \updownarrow & & \updownarrow \\ \varphi_n & \longrightarrow & \varphi \end{array}$$

Problema: $\varphi_n(t) = e^{-n\frac{t^2}{2}}$

$\mu_n \sim N(0,n)$

$\mu_n((-\infty, x]) \to 1/2 \quad \forall x$

$$\geq 2\int_{|x|\geq \frac{2}{U}} \left(1-\frac{\operatorname{sen} Ux}{x}\right)\mu_n(x)$$

$$\frac{\operatorname{sen} Ux}{Ux} \leq \frac{1}{|Ux|}$$

$\leq 1/2$

$$\geq 2\int_{|x|\geq \frac{2}{U}} \left(1-\frac{1}{|Ux|}\right)\mu_n(dx) \geq \mu_n(x : |x|\geq 2/U)$$

Continuidade de $\varphi(t)$ em $t=0$

$$\frac{1}{U}\int_{-U}^{U}(1-\varphi_n(t)) \xrightarrow[U\downarrow 0]{} 0$$

Dado $\varepsilon>0$ escolho U tal que

$$\frac{1}{U_\varepsilon}\int_{-U_\varepsilon}^{U_\varepsilon}(1-\varphi(t))\,dt < \varepsilon$$

$$\varphi_n(t)\to\varphi(t)$$

$n\geq N \quad |\varphi_n(t)-\varphi(t)|<\varepsilon \quad \forall t$

Teorema da Convergencia Limitada

$$\frac{1}{U}\int_{-U}^{U}(1-\varphi_n(t)) \leq 2\varepsilon \Rightarrow \text{tight}$$

$\mu_n(x : |x|\geq 2/U)$

, $M_\varepsilon = \frac{1}{U_\varepsilon}$

1) Se $\mu_n \Rightarrow \mu$ (con. fraca) então

$$\varphi_n(t) \to \varphi(t)$$

2) Se $\varphi_n(t) \to \varphi(t)$ e $\varphi(t)$ é contínua na origem então $\{\mu_n\}$ é tight e

$$\mu_n \Rightarrow \mu$$

Dem:

① imediato (exercício)

② "$\{\varphi_n(t) \to \varphi(t)$ + cont. em $0\} \Rightarrow \{\mu_n \Rightarrow \mu\}$"

Dado $U > 0$

$$\int_{-U}^{U} (1 - e^{itx})\, dt = 2U - 2\frac{\operatorname{sen} Ux}{x}$$

$$\int \left(\frac{1}{U} \int_{-U}^{U} (\quad) \, dt \right) \mu_n(dx) \overset{\text{Fubini}}{=}$$

$$\frac{1}{U} \int_{-U}^{U} 1 - \varphi_n(t)\, dt =$$

$$= 2 \int \left(1 - \frac{\operatorname{sen} Ux}{Ux} \right) \mu_n(dx)$$

$$1 - \frac{\operatorname{sen} Ux}{Ux} \geq 0 \qquad |\operatorname{sen} x| \leq |x|$$

$$\left| \frac{\operatorname{sen} x}{x} \right| \leq 1 \qquad -1 \leq \frac{\operatorname{sen} x}{x} \leq 1$$

TLC: Se $X_1, X_2, \dots$ i.i.d. com $E X_1 = \mu$ e $var(X_i) = \sigma^2 \in (0,\infty)$ então

$$P\left(\underbrace{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma}} \leq x\right) \rightarrow \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-s^2/2} \, ds$$

$$W_n \Rightarrow W, \quad W \sim N(0,1)$$

Aula 15 - Probabilidade Avançada I — 15/05/2014

TLC: $X_1, X_2, \dots$ v.a. i.i.d. com $EX_i = \mu$ e $var(X_i) = \sigma^2 \in (0,\infty)$

Então

$$\frac{\frac{1}{n}\sum X_i - \mu}{\sigma/\sqrt{n}} \Rightarrow N(0,1)$$

Demonstração:

Posso tomar $\mu = 0$.

$$X_i \to X_i - \mu$$

(A) $\phi(t) = 1 - \frac{\sigma^2 t^2}{2} + o(t^2)$

Dem. A:

Lema:

$$\left| e^{ix} - \sum_{m=0}^{n} \frac{(ix)^m}{m!} \right| \leq \min\left\{ \frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!} \right\}$$

Dem: Lema:

$$\int_0^x (x-s)^n e^{is}\, ds = I$$

$$\left[\frac{(x-s)^{n+1}}{n+1} e^{is} \right]' = -(x-s)^n e^{is} + i\frac{(x-s)^{n+1}}{n+1} e^{is}$$

$$I = \frac{x^{n+1}}{n+1} + i\int_0^x \frac{(x-s)^{n+1}}{(n+1)} e^{is}\, ds$$

Para $n=0$,

$$\int_0^x e^{is}\,ds = x + i\int_0^x (x-s)\,e^{is}\,ds$$

$$\parallel$$

$$\frac{e^{ix}-1}{i}$$

$$e^{ix} = 1 + xi + i^2\int_0^x (x-s)\,e^{is}\,ds$$

<u>Itero</u>:

$$e^{ix} - \sum_{m=0}^{n} \frac{(ix)^m}{m!} = \frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is}\,ds$$

$$\uparrow$$

$$\Big| \qquad\qquad \Big| \leq \frac{1}{n!}\int_0^x (x-s)^n\,ds = \frac{|x|^{n+1}}{(n+1)!}$$ (boa p/ $|x|$ pequeno)

$$\mathrm{II} = \frac{1}{n}\int_0^x (x-s)^n e^{is}\,ds = \frac{x^n}{n} + \int_0^x (x-s)^{n-1} e^{is}\,ds$$

$$\left[\frac{(x-s)^n}{n}e^{is}\right]' = -(x-s)^{n-1}e^{is} + i\,\frac{(x-s)^n}{n}e^{is}$$

$$\frac{x^n}{n} = \int_0^x (x-s)^{n-1}\,ds$$

$$\mathrm{II} = \int_0^x (x-s)^{n-1}\left[1-e^{is}\right]ds \quad \frac{i^n}{(n-1)!}$$

Então,

$$\frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is}\, ds =$$

$$\frac{i^n}{(n-1)!}\int_0^x (x-s)^{n-1}\left[1-e^{is}\right] ds$$

$$|1-e^{is}| = \sqrt{(1-e^{is})(1+e^{is})} = \sqrt{\underbrace{2(1-\cos s)}_{\leq 4}} \leq 2$$

$$\left| e^{ix} - \sum_{m=0}^{n} \frac{(ix)^m}{m!} \right| \leq \frac{1}{(n-1)!}\, 2 \int_0^x (x-s)^{n-1}\, ds$$

$$\frac{2|x|^n}{n!}$$

Tomando o valor esperado

$$\left| E e^{itx} - \sum_{m=0}^{n} \frac{E(itX)^m}{m!} \right| \overset{\text{Jensen}}{\leq} E\left(\min\left\{ \frac{|tX|^{n+1}}{(n+1)!}, \frac{2|tX|^n}{n!} \right\}\right)$$

Ⓐ Lema

Se $E|X|^2 < \infty$ então

$$\varphi(t) = 1 + it\,EX - \frac{t^2 EX^2}{2} + o(t^2)$$

Dem A.

$$\text{erro} \leq \frac{t^2}{2} E\left[\frac{|t||X|^3}{3} \wedge \overset{\leq |X|^2}{2|X|^2} \right]$$

$$o(t^3)$$

Voltando TLC.

$$\varphi_{X_i}(t) = 1 - \frac{\sigma^2 t^2}{2} + o(t^2)$$

$$\varphi_{\frac{S_n}{\sigma\sqrt{n}}}(t) = E\left[e^{it\frac{S_n}{\sigma\sqrt{n}}} \right] = \left[\varphi_{X_1}\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n =$$

$$= \left[1 - \frac{\sigma^2 t^2}{2\sigma^2 n} + o\left(\frac{t^2}{\sigma^2 n}\right)\right]^n \xrightarrow{?} e^{-\frac{t^2}{2}}$$

Teo: Se $c_n \in \mathbb{C}$, $c_n \to c \in \mathbb{C}$,

$$\left(1 + \frac{c_n}{n}\right)^n \to e^c$$

Dem. Se $z_1, z_2, \ldots, z_n \in \mathbb{C}$

$w_1, w_2, \ldots, w_n$

todos com norma $\leq \theta$ então

$$\left| \prod_{m=1}^{n} z_m - \prod_{m=1}^{n} w_m \right| \leq \theta^{n-1} \sum_{m=1}^{n} |z_m - w_m| \qquad \text{Indução}$$

Lema 2: $b \in \mathbb{C}$, $|b| \leq 1$ então

$$|e^b - (1+b)| \leq b^2$$

$\uparrow$

expando (Durret)

Lema 1 + Lema 2 $\Rightarrow$ Teorema.

$$z_m = \left(1 + \frac{c_n}{n}\right) \qquad m = 1, 2, \ldots, n$$

$$w_n = e^{\frac{c_n}{n}} \qquad c_n \to c \in \mathbb{C}$$

Seja r, $r > |c|$. então, como $c_n \to c$, existe N,

$$|c_n| < r, \; n \geq N$$

$$1 + \frac{r}{n} \leq e^{r/n}$$

$$\left| \left(1 + \frac{c_n}{n}\right) - e^{c_n} \right| \leq e^{r\frac{n-1}{n}} \, n \left(\frac{c_n}{n}\right)^2$$

$$\leq e^{r\frac{(n-1)}{n}} \frac{r^2}{n} \downarrow 0$$

Teorema Limite Central para Arranjos Triangulares

Para cada n, X_{nm} $\quad 1 \leq m \leq n$

Independentes com $EX_{nm} = 0$

X_{11}

X_{21} X_{22}

$\vdots$

$$S_n = \sum_{m=1}^{n} X_{nm}$$

Suponha que

1. $\sum_{m=1}^{n} E X_{nm}^2 \to \sigma^2 \in (0,\infty)$

2. Para todo $\varepsilon > 0$

$$\sum_{m=1}^{n} E[X_{nm}^2 ; |X_{nm}| > \varepsilon] \to 0$$

Então

$$\frac{S_n}{\sigma} \Rightarrow N(0,1)$$

Obs: Esse teo. é mais geral que o anterior
Se $Y_1, Y_2, \dots$ i.i.d. com $EY_i = 0$
$Var(Y_i) = \sigma^2$
tomo $X_{nm} = \frac{Y_m}{\sqrt{n}}$

Então

$$\sum_{m=1}^{n} E X_{nm}^2 = \sigma^2$$

$$\sum_{m=1}^{n} E[X_{nm}^2 ; |X_{nm}| > \varepsilon] =$$

$$= \sum_{m=1}^{n} \frac{1}{n} E[Y_m^2 ; |Y_m| > \varepsilon\sqrt{n}] =$$

$$= E[Y_1^2 ; |Y_1| > \varepsilon\sqrt{n}] \to 0$$

Dem:

$$\varphi_{nm}(t) = E\, e^{itX_{nm}}$$

$$\sigma^2_{nm} = E X^2_{nm}$$

Basta provar que

$$\prod_{m=1}^{n} \varphi_{n,m}(t) \to e^{-\frac{t^2\sigma^2}{2}}$$

Notação:

$$Z_{nm} = \varphi_{nm}(t)$$
$$W_{nm} = \left(1 - \frac{t^2\sigma^2_{nm}}{2}\right)$$

$$|Z_{nm} - W_{nm}| \leq E\left[|tX_{nm}|^3 \wedge 2|tX_{nm}|^2\right]$$

Fixo $\varepsilon > 0$

$$\leq E\left[|tX_{nm}|^3 ; |X_{nm}| \leq \varepsilon\right] + E\left[2|tX_{nm}|^2 ; |X_{nm}| > \varepsilon\right]$$

$$\leq t^3 \varepsilon\, E\left[|X_{nm}|^2 ; |X_{nm}| \leq \varepsilon\right] + 2t^2\, E\left[|X_{nm}|^2 ; |X_{nm}| > \varepsilon\right]$$

$$\sum_{m=1}^{n} |Z_{n,m} - W_{nm}| \leq t^3 \varepsilon \underbrace{\sum_{m=1}^{n} E X^2_{nm}}_{\downarrow\ \sigma^2} + 2t^2 \underbrace{\sum_{m=1}^{n} E\left[|X_{nm}|^2 ; |X_{nm}| > \varepsilon\right]}_{\downarrow\ 0}$$

Daí,

$$0 \leq \overline{\lim_n} \sum_{m=1}^{n} |Z_{nm} - W_{nm}| \leq t^3 \varepsilon \sigma^2 \quad \forall \varepsilon .$$ Então

$$\lim_n \sum_{m=1}^{n} |z_{nm} - w_{nm}| = 0$$

Uso os 2 lemas

$$\left| \prod_{m=1}^{n} \varphi_{nm}(t) - \prod_{m=1}^{n} \left(1 - \frac{t^2 \sigma^2_{nm}}{2}\right) \right|$$

$$\sigma^2_{nm} \leq \varepsilon^2 + \underbrace{E[X^2_{nm} ; |X_{nm}| > \varepsilon]}_{\to 0}$$ pois $\sum_{m=1}^{n} a \to 0$

$$\overline{\lim_n} \, \sigma^2_{nm} \leq \varepsilon \quad \forall \varepsilon$$

$$-1 \leq 1 - \frac{t^2 \sigma^2_{nm}}{2} \leq 1$$

Esperança Condicional

X

$E[X / \mathcal{F}]$ – melhor chute p/ X dada a inf. $\mathcal{F}$

se X for $\mathcal{F}$-mensurável, temos toda a inf. sobre $X \Rightarrow E[X/\mathcal{F}] = X$

– Se $\mathcal{F} = \{\Omega, \emptyset\}$, $E(X/\mathcal{F}) = E(X)$.

Seja $(\Omega, \mathcal{F}_0, P)$ um esp de probabilidade e $\mathcal{F} \subset \mathcal{F}_0$ X é uma v.a.

e $X \in \mathcal{F}_0$ com $E|X| < \infty$

Defino $E[X|\mathcal{F}]$ como uma v.a. Y tal que

1) $Y \in \mathcal{F}$
2) Se $A \in \mathcal{F}$ então $\int_A Y \, dP = \int_A X \, dP$

Y é dita uma versão da esperança condicional $E[X|\mathcal{F}]$.
Para provar existência, precisamos do Teorema de Radon-Nikodyn

Def. Dizemos que μ é uma medida com sinal (signed measure) se

1. $\mu : \mathcal{F} \to (-\infty, +\infty]$
2. $\mu(\emptyset) = 0$
3. Se $E = \bigcup_n E_n$, E_n disjuntos

então

$$\mu(E) = \sum_n \mu(E_n)$$

se $\mu(E_n) < \infty$ série absolutamente somável
se $\mu(E) = \infty$ $\sum \mu^+(E_n) = \infty$ e $\sum \mu^-(E_n) > -\infty$

$$f : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$$

1. Exemplo: Se f é uma função tal que

$$\int f^- \, dP < \infty$$

defino $\mu(A) = \int_A f \, dP$.

2. Se μ_1 e μ_2 são duas medidas então

$$\mu = \mu_1 - \mu_2 \qquad \text{com } \mu_2(\Omega) < \infty$$

<u>Def</u>: Dado $(\Omega, \mathcal{F}, \mu)$ μ medida com sinal.

Dizemos que $A \in \mathcal{F}$ é positivo se $B \subset A$, $B \in \mathcal{F}$ então $\mu(B) \geq 0$

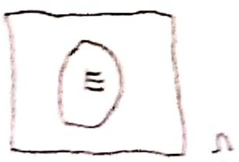
Analogamente p/ negativo.

Obs.

1. Se A é positivo e $B \subset A$ com $B \in \mathcal{F}$ então B é positivo

2. Se A_n é uma col. de conj. positivos então $\bigcup A_n$ é positivo.

Análogo p/ negativo.



$\mu(E) < 0$

Lema: Se $E \in \mathcal{F}$ com $\mu(E) < 0$ então existe $F \subset E$, F <u>negativo</u> com $\mu(F) < 0$.

Dem:

Se E negativo. OK

Suponha que não.

Então existe E' com $\mu(E') > 0$.

Seja n_1 = menor inteiro positivo tal que existe $E_1 \subset E$ tal que

$$\mu(E_1) \geq \frac{1}{n_1}$$

$$F_1 = E - E_1 = E \cap E_1^c$$

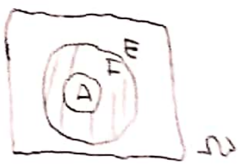
Para $k \geq 0$

n_k = menor inteiro tal que existe $E_{k+1} \subset F_k$

tal que $\mu(E_k) \geq \frac{1}{n_k}$ $\qquad F_k = E - \bigcup_{i=1}^{k} E_i$

E_k disjuntos

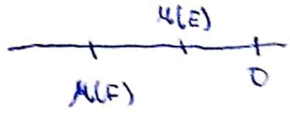
$$F = E - \underbrace{\bigcup E_k}_{A}$$



$$F = E - \underbrace{\bigcup E_k}_{A}$$

$$F \cup A = E$$

$$\mu(E) = \mu(F) + \mu(A) = \mu(F) + \underbrace{\sum_k \mu(E_k)}_{<\infty} \geq \mu(F)$$



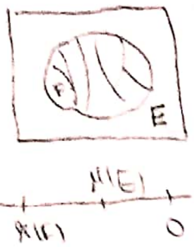
F não tem $G \subset F$ com $\mu(G) > 0$. Então F é negativo e $\mu(F) < 0$.

Aula 16 - Probabilidade Avançada

20/05/2014

Lema: Se $E \in \mathcal{F}$ com $\mu(E) < 0$ então existe F negativo, $F \subset E$, com $\mu(F) < 0$.

Dem.



E_1 "maior pedaço de E" com med. positiva

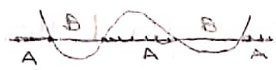
menor n_1 : $\mu(E_1) \geq \frac{1}{n_1}$ $\qquad E_n \downarrow 0$

$$F_k = E - \bigcup_{1}^{k} E_i \to F$$

Teorema: Decomposição de Hahn

Se μ é uma medida com sinal em $(\Omega, \mathcal{F})$ então existem A positivo e B negativo com $\Omega = A \cup B$ e $A \cap B = \emptyset$.

Dem:



Seja $c = \inf\{\mu(B) ; B \text{ negativo}\} \leq 0$

B_n conj. negativos com $\mu(B_n) \downarrow c$

tomo $B = \bigcup_n B_n \to \mu(B) = c$

. $\mu(B) \geq c$, por definição.

União de neg. é negativo. Como c é o inf, segue a afirmação.

$$B = \Big(\bigcup_{n \neq i} B_n \Big) \cup B_i$$

$$\mu(B) = \mu(B_i) + \mu(B \setminus B_i)$$

$$\leq \mu(B_i) \qquad \forall i \qquad B \cap B_i^c \subseteq B \text{ negativo}$$

$$\downarrow c$$

i.e. $\mu(B) \leq \mu(B_i) \to c$

$\Rightarrow \mu(B) = c > -\infty$.

• $A = B^c$ é positivo (?)

(se existe $C \subset B^c$ com medida < 0 então $B \cup C$ é mais negativa $\mu(B) + \mu(C) < c$ contradição!) $\square$

Definição: Duas medidas μ_1 e μ_2 em $(\Omega, \mathcal{F})$ são ditas mutuamente singulares se existe $A \in \mathcal{F}$ tal que

$$\mu_1(A) = 0 \quad e \quad \mu_2(A^c) = 0$$

Notação: $\mu_1 \perp \mu_2$

Teo. Decomposição de Jordan

Se μ é uma medida com sinal então existem duas medidas μ_+ e μ_- com

$$\mu_+ \perp \mu_- \quad e \quad \mu = \mu_+ - \mu_-$$

Dem: Seja A e B uma decomposição de Hahn

Se $E \subset \mathcal{F}$

$$\mu_+(E) = \mu(E \cap A)$$

$$\mu_-(E) = -\mu(E \cap B)$$

Exercício: são medidas. μ_+ e μ_- são únicas.

Teorema. Decomposição de Lebesgue

Se μ e ν são duas medidas (σ-finitas) então existem medidas $\underset{\text{regular}}{\nu_r}$ e $\underset{\text{singular}}{\nu_s}$ tal que

$$\nu = \nu_r + \nu_s \ , \quad \nu_s \perp \mu$$

e existe g tal

$$\nu_r(A) = \int_A g \, d\mu \qquad \forall A \in \mathcal{F}$$

Dem: μ e ν finitas

$$\mathcal{G}_E = \{ g : \nu(E) \geq \int_E g \, d\mu \ , \forall E \in \mathcal{F} \}$$

1) Se g e h estão em $\mathcal{G}_E$ então

$$g \vee h \in \mathcal{G}_E$$

$g \vee h = \max\{g, h\}$

Dem (1),

Sejam

$A = \{g > h\}$

$B = \{g \leq h\}$

$$\int_E g \vee h \, d\mu = \int_{E \cap A} g \, d\mu + \int_{E \cap B} h \, d\mu \leq \nu(E \cap A) + \nu(E \cap B) = \nu(E).$$

$$k_i = \sup_{g \in \mathcal{G}_i} \left\{ \int_E g \, d\mu \right\} \leq \nu(\Omega) < \infty$$

Defino g_n, $n \geq 1$, tal que

$$\int g_n \, d\mu > k_i - \frac{1}{n}$$

$\{g_n\}_{n \geq 1}$

$h_n = g_1 \vee g_2 \vee \dots \vee g_n = \max\{g_1, \dots, g_n\} \in \mathcal{G}_i$ por (I)

Seja $h = \lim_n h_n$, $h_n \uparrow h$ então $\underset{TCM}{}$

$$k_i \geq \int_i h \, d\mu = \int_i \lim_n h_n \, d\mu \overset{TCM}{=} \lim_n \int_i h_n \, d\mu \geq$$

$$\lim_n \int_i g_n \, d\mu = k_i.$$

$$\nu_r(A) = \int_A h \, d\mu$$

ν_r é uma medida (exercício.)

$\nu_s = \nu - \nu_r$

Preciso mostrar que ν_s é singular com respeito a μ.

Para cada $\varepsilon > 0$ seja
$A_\varepsilon, B_\varepsilon$ a decomposição de Hahn para

$$\mu_\varepsilon \equiv \nu_s - \varepsilon\mu$$

$$0 \leq \mu_\varepsilon(E \cap A_\varepsilon) = \nu_s(E \cap A_\varepsilon) - \varepsilon\mu(E \cap A_\varepsilon)$$

$$\varepsilon\mu(E \cap A_\varepsilon) \leq \nu_s(E \cap A_\varepsilon)$$

$$\int_E (h + \varepsilon \mathbb{1}_{A_\varepsilon}) \, d\mu = \int_E h \, d\mu + \varepsilon\mu(E \cap A_\varepsilon) \leq$$

$$\nu_r(E) + \nu_s(E \cap A_\varepsilon) \leq \nu_s(E) \leq \nu(E)$$

Então $r = h + \varepsilon \mathbb{1}_{A_\varepsilon} \in \mathcal{G}$.

$$\int r \, d\mu \leq K$$
$$\|$$
$$\int h \, d\mu + \varepsilon\mu(A_\varepsilon) \leq K$$
$$\downarrow$$
$$K + \varepsilon\mu(A_\varepsilon) \leq K \Rightarrow \mu(A_\varepsilon) = 0 \ \forall \varepsilon > 0$$

$\varepsilon_n = 1/n \downarrow 0$

$A = \bigcup_n A_{\frac{1}{n}}$ $\quad \mu(A) = 0$

$\nu_s(A^c) =$

mas se $\nu_s(A^c) > 0$

$\nu_s(A^c) - \varepsilon\mu(A^c) > 0$ para ε suf. pequeno

$\mu_\varepsilon(A^c) > 0$

Aula 17 - Probabilidade Avançada I

Teo. (Dec. de Lebesgue)

μ, ν medidas

$\nu = \nu_r + \nu_s$

$\nu_s \perp \mu$ e $\nu_r = \int g \, d\mu$ p/ algum g

Dem:

$\mathcal{G} =$ "aprox. por baixo de ν"

h realiza $K = \sup \{ \int g \, d\mu, g \in \mathcal{G} \}$

$r = h + \varepsilon \mathbb{1}_{A_\varepsilon}$ $\quad \int_E r \, d\mu \leq \nu(E)$

$\mu_\varepsilon = \nu_s - \varepsilon \mu \Rightarrow$ dec. de Hahn $A_\varepsilon, B_\varepsilon$

$r \in \mathcal{G}$

$\mu(A_\varepsilon) = 0 \ \forall \varepsilon > 0$. $\varepsilon_n = 1/n \downarrow 0$

$A_n = \bigcup_n A_{1/n}$ $\quad \mu(A) = 0$ $\quad A^c \in B_\varepsilon \ \forall \varepsilon > 0$

Quero verificar que $\nu_s(A^c) = 0$

Por obs: suponha $\nu_s(A^c) \geq \delta > 0$ então

$\mu_\varepsilon(A^c) = \nu_s(A^c)^{\geq\delta} - \varepsilon \mu(A^c) > 0$ para $\varepsilon > 0$ suf. peq., mas $A^c \subset B_\varepsilon$ (...)

ν, μ duas medidas

Def. ν é dita absolutamente contínua com respeito a μ (notação: $\nu \ll \mu$) se para todo $A \in \mathcal{F}$ com $\mu(A) = 0$ então $\nu(A) = 0$.

Teorema de Radon-Nikodym

Se μ e ν são medidas σ-finitas e $\nu \ll \mu$ então $\exists\, g \geq 0$ tal que $\nu(E) = \int_E g \, d\mu \quad \forall E \in \mathcal{F}$. (*)

Se h tbm satisfaz (*) então $h = g$ q.c. μ.

Dem:

$\nu, \mu \Rightarrow \nu = \nu_r + \nu_s$, $\nu_r = \int g \, d\mu$ e $\nu_s \perp \mu$

Seja $\nu = \nu_r + \nu_s$.

Seja A tal que

$$\nu_s(A^c) = 0 \text{ e } \mu(A) = 0$$

com $\nu \ll \mu$.

$$\mu(A) = 0 \Rightarrow \nu(A) = 0 = \nu_s(A) + \nu_r(A) \geq \nu_s(A) \Rightarrow \nu_s(A) = 0.$$

<u>Unicidade</u> (Exercício)

Sugestão: $E \in \{g > h, g \leq n\} = A_n$ se h e g satisfazem (*) $\square$

Notação: $\nu(E) = \int_E g \, d\mu$

$\nu(E) = \nu_r(E) + \nu_s(E)$ "$g = \frac{d\nu}{d\mu}$", = derivada de Radon-Nikodym

Esperança Condicional

Pelo T.R.N.. Se μ e ν são duas medidas com ~~[illegible]~~ $\nu << \mu$, existe $g = \frac{d\nu}{d\mu}$, tal que

$$\nu(E) = \int_E g \, d\mu$$

Existe $E[X/\mathcal{F}]$.

Def: X v.a. em $(\Omega, \mathcal{F}_0, P)$

$\mathcal{F} \subset \mathcal{F}_0$, $E|X| < \infty$

$h = E[X/\mathcal{F}]$ é qualquer função

Existência Segue de Radon-Nikodym

+ Se $X \geq 0$ $E|X| < \infty$

P, $\nu(E) = \int_E X \, dP$ ex: é uma medida $\nu << P$

RN $\nu(E) = \int_E g \, dP$ $E \in \mathcal{F}$ $g = \frac{d\nu}{d\mu} = E[X/\mathcal{F}]$

$$\nu(E) = \int_E X \, dP \overset{R.N}{=} \int_E \frac{d\nu}{dP} \, dP$$

$$\frac{d\nu}{dP} = \cdot$$

então $\frac{d\nu}{dP}$ é uma versão de $E[X/\mathcal{F}]$.

Para X geral

$$X = X^+ - X^- \Rightarrow E[X/\mathcal{F}] = E[X^+/\mathcal{F}] - E[X^-/\mathcal{F}]$$

seja

$Y_1 = E[X^+/\mathcal{F}]$ e $Y_2 = E[X^-/\mathcal{F}]$ ok por $(X \geq 0)$

$Y_1 - Y_2 \in \mathcal{F}$

Prop. 2

$$\int_E X \, dP = \int_E X^+ dP - \int_E X^- dP =$$

$$= \int_E E[X^+/\mathcal{F}] \, dP - \int_E E[X^-/\mathcal{F}] \, dP$$

$$= \int_E (Y_1 - Y_2) \, dP = \int_E E[X/\mathcal{F}] \, dP$$

Aula 18 - Probabilidade Avançada

27/05/2014

Teorema. Se $\mathcal{F}_1 \subset \mathcal{F}_2$

1. $E[E[X/\mathcal{F}_1]/\mathcal{F}_2] = E[X/\mathcal{F}_1]$

2. $E[E[X/\mathcal{F}_2]/\mathcal{F}_1] = E[X/\mathcal{F}_1]$

Dem:

1. OK

$Y \in \mathcal{F}_1 \subset \mathcal{F}_2$

$Y \in \mathcal{F}_2$

2. Se $A \in \mathcal{F}_1$,

$$\int_A E[X/\mathcal{F}]\,dP = \int_A X\,dP = \int_A E[X/\mathcal{F}_2]\,dP$$

$\rightarrow$ "$E[X/\mathcal{F}_2]$ é uma versão de $E[E[X/\mathcal{F}_2]/\mathcal{F}_1]$"

Teorema. Se $X \in \mathcal{F}$, Y: $E|Y| < \infty$ e $E|XY| < \infty$

$$E[XY/\mathcal{F}] = X\,E[Y/\mathcal{F}]$$

Dem. 4 passos

1. $X = \mathbb{1}_B$, $B \in \mathcal{F}$

$A \in \mathcal{F}$,

$$\int_A X\,E[Y/\mathcal{F}]\,dP = \int_A \mathbb{1}_B\,E[Y/\mathcal{F}]\,dP = \int_{A\cap B} E[Y/\mathcal{F}]\,dP = \int_{A\cap B} Y\,dP = \int_A \mathbb{1}_B\,Y\,dP = \int_A XY\,dP$$

2. $X = \sum_i b_i \mathbb{1}_{B_i}$ "funções simples"

3. $X \geq 0$ X_n simples $\uparrow X$

4. $X = X^+ - X^-$ $\square$

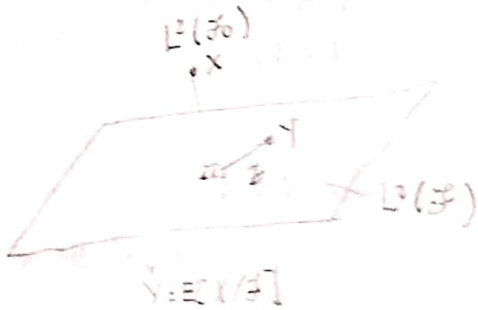
"$E[X/\mathcal{F}]$ é uma projeção"

$X: (\Omega, \mathcal{F}_0, P) \to (\mathbb{R}, \mathcal{B})$

Teorema: Se $EX^2 < \infty$ então $Y = E[X/\mathcal{F}]$ é a v.a. que minimiza $E(Y-X)^2$.

Def. $L^2(\mathcal{F}_0) = \{W \in \mathcal{F}_0 : EW^2 < \infty\}$

$L^2(\mathcal{F}) \subset L^2(\mathcal{F}_0)$



Seja $Z \in L^2(\mathcal{F})$

$$E[ZX/\mathcal{F}] = Z E[X/\mathcal{F}]$$

$$E[E[ZX/\mathcal{F}]] = E[Z E[X/\mathcal{F}]] = EZX$$

$Y \in L^2(\mathcal{F})$

$Z = E[X/\mathcal{F}] - Y$

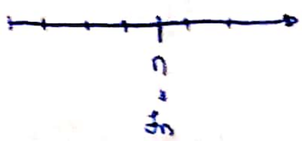
$$E[X-Y]^2 = E(X - E[X/\mathcal{F}] + Z)^2 = E(X - E(X/\mathcal{F}))^2 + EZ^2 + 2E[Z(X - E(X/\mathcal{F}))]$$

$EZX - E[ZE(X/\mathcal{F})]$

Mínimo para $Z = 0 \Rightarrow Y = E[X/\mathcal{F}]$

Martingais:

$\{X_n\}_{n\geq 0}$



Definição: $\{\mathcal{F}_n\}_{n\geq 1}$ é dita <u>filtragem</u> se for uma sequência não-decrescente de σ-álgebras.

<u>Def</u>: Dizemos que $\{X_n\}_{n\geq 1}$ é adaptada com respeito a filtragem $\{\mathcal{F}_n\}_{n\geq 1}$ se $X_n \in \mathcal{F}_n$ $\forall n \geq 1$.

Def: Martingal: coleção $\{X_n\}_{n\geq 1}$ adaptada com respeito (acr) $\{\mathcal{F}_n\}_{n\geq 1}$.

1. $E|X_n| < \infty$
2. $E[X_{n+1} / \mathcal{F}_n] = X_n$

Jogo justo

Então $\{X_n, \mathcal{F}_n\}_{n\geq 1}$ é um martingal. Na verd. $\{X_n\}_{n\geq 1}$ é um martingal com resp. à $\{\mathcal{F}_n\}_{n\geq 1}$

Def. Submartingal.

1. $E|X_n| < \infty$
2. $E[X_{n+1} / \mathcal{F}_n] \geq X_n$

Vai ficando maior

Análogo a uma seq. que cresce

Def. Supermartingal.

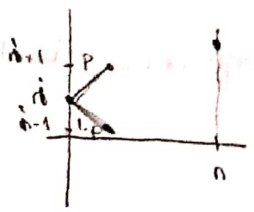
1. $E|X_n| < \infty$
2. $E[X_{n+1} / \mathcal{F}_n] \leq X_n$

Vai ficando menor

Jogo de Casino, por exemplo

* Se tenho um submartingal que é limitado por cima, então ele converge q.c.

Ex: Jogo da roleta

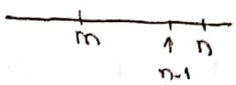


$p = \frac{18}{38}$ $\qquad p_i = p\,p_{i+1} + (1-p)\,p_{i-1}$

Teorema: Se $\{X_n\}$ é um supermartingal com respeito a $\{\mathcal{F}_n\}_{n\geq 1}$ e $n > m$

$$E[X_n / \mathcal{F}_m] \leq X_m$$

Def: $E[X_n | \mathcal{F}_m] = E[E[X_n | \mathcal{F}_{n-1}] | \mathcal{F}_m] \leq E[X_{n-1} / \mathcal{F}_m]$



Teorema: Se $\{X_n\}_{n\geq 1}$ é mart. com resp. a $\{\mathcal{F}_n\}_{n\geq 1}$ e φ é uma função convexa com $E|\varphi(X_n)| < \infty$ $\forall n \geq 1$ então

$\{\varphi(X_n)\}_{n\geq 1}$ é um submartingal c.r.a $\{\mathcal{F}_n\}_{n\geq 1}$.

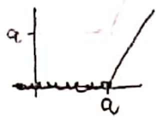
Demonstração:

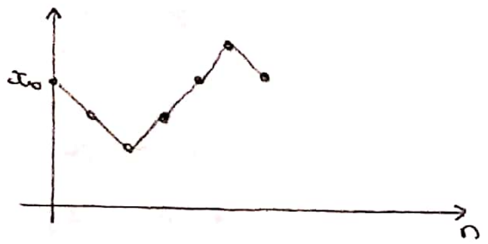
$E[\varphi(X_{n+1}) / \mathcal{F}] \geq \varphi[E[X_{n+1} | \mathcal{F}]) = \varphi(X_n)$ $\quad\square$

Ex. $X_n \to |X_n|^p$, $p \geq 1$.

Teorema: Se X_n é <u>submartingal</u> c.r.a. $\mathcal{F}_n$ e φ é convexa crescente então $\{\varphi(X_n)\}_n$ é submartingal.

Exemplo: $\varphi(x) = (x-a)^+ = \max\{x-a, 0\}$.





$\{\xi_n\}_{n\geq 1}$ i.i.d. com $P(\xi_n = 1) = 1 - P(\xi_n = -1)$

$$X_n = X_0 + \sum_{i=1}^{n} \xi_i$$

Def. Diremos que $\{H_n\}_{n\geq 1}$ é uma sequência de funções previsíveis se se

$$H_n \in \mathcal{F}_{n-1} \quad \forall n$$

se "aposto H_n" na jogada n.

Fortuna no instante n =

$$(H \cdot X)_n = \sum_{m=1}^{n} \underbrace{H_m}_{\text{qte apostada}} \underbrace{(X_m - X_{m-1})}_{\text{dif. } \begin{smallmatrix} +1 \text{ ganhou} \\ -1 \text{ perdeu} \end{smallmatrix}}$$

Teorema. Se $\{X_n\}_{n\geq 1}$ é um supermartingal c.r.a. $\{\mathcal{F}_n\}_{n\geq 1}$.
Se $\{H_n\}_{n\geq 1}$ previsível c.r.a. $\{\mathcal{F}_n\}_{n\geq 1}$ e $H_n \geq 0$.
H_n limitado.

Então

$\{(H.X)_n\}_{n\geq 1}$ é um supermartingal.

Dem.

$$E[(H.X)_{n+1} | \mathcal{F}_n] = E[(H.X)_n + H(X_{n+1} - X_n) | \mathcal{F}_n] \overset{H_{n+1} \in \mathcal{F}_n}{=}$$

$$= (H.X)_n + \underbrace{\underbrace{H_{n+1}}_{\geq 0} \, E[\underbrace{X_{n+1} - X_n | \mathcal{F}_n}_{\leq 0}]}_{\leq 0}$$

$$\leq (H.X)_n \qquad \square$$

Ex. correspondente p/ sub e martingal

Def. $N \in \mathbb{N}$ é dita um tempo de parada c.r.a. $\mathcal{F}_n$ se

$$\{N = n\} \in \mathcal{F}_n \quad \forall n \geq 1$$

"Parar no inst. n"

$$\{N \geq n\} = \{N \leq n-1\}^c \in \mathcal{F}_{n-1}$$

Obs.

$H_n = \mathbb{1}_{\{N \geq n\}}$ é previsível

Aula 19 - Probabilidade Avançada I

Def. (Upcrossing) "Cruzamento para $\{X_n\}_{n\geq 0}$ submartingal (c.r.a. $\{\mathcal{F}_n\}_{n\geq 0}$) cima"

Fixa $a < b$

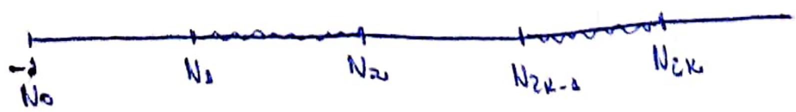
Seja $N_0 = -1$
para $k \geq 1$

$$N_{2k-1} = \inf\{m \geq N_{2k-2} : X_n \leq a\}$$

$$N_{2k} = \inf\{m \geq N_{2k-1} : X_n \geq b\}$$

Obs.
$N_1 = \inf\{m \geq 1 : X_n \leq a\}$
$N_2 = \dots$

N_k tempos de parada



$\{N_{2k-1} < m \leq N_{2k}\} = \{N_{2k-1} \leq m-1\} \cap \{N_{2k} \leq m-1\}^c \in \mathcal{F}_{m-1}$

Sequência previsível

$$H_m = \begin{cases} 1 & \text{se } \{N_{2k-1} < m \leq N_{2k}\} \\ 0 & \text{c.c.} \end{cases}$$

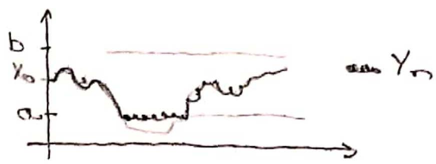
Seja:

$$U_n = \sup\{k : N_{2k} \leq n\} = \#\text{ de upcrossings até } n$$

Teorema.

$$EU_n \leq \frac{E(X_n - a)^+ - E(X_0 - a)^+}{(b-a)}$$

Dem.: Seja $Y_n = a + (X_n - a)^+$



Y_n é uma f. convexa crescente de X_n.
$\therefore$ Y é submartingal

\# de upcrossings de X_n e Y_n é o mesmo

$$(b-a)\,U_n \leq (H \cdot Y)_n$$
$$\overset{\shortparallel}{\sum_{m=1}^{n} H_m (Y_m - Y_{m-1})}$$

Pois

1) cada upcrossing gera "lucro" $\geq b-a$
2) o último, incompleto, upcrossing gera lucro ≥ 0 (p/ Y_n)

$K_n = 1 - H_n$

Obs. (Teo) Se X_n supermartingal e $\{H_n\}_{n\geq 1}$ previsível, então $\{(H\cdot X)_n\}_{n\geq 1}$ é supermartingal

$$(H\cdot Y)_n + (K\cdot Y)_n = \sum_{m=1}^{n} H_m (Y_m - Y_{m-1}) + \sum_{m=1}^{n} (1-H_m)(Y_m - Y_{m-1})$$

$$= \sum_{m=1}^{n} (Y_m - Y_{m-1}) = Y_n - Y_0$$

$$E(K\cdot Y)_n \geq E(K\cdot Y)_0 = 0$$

$$E(H\cdot Y)_n \leq E(H\cdot Y)_n + E(K\cdot Y)_n = E(Y_n - Y_0)$$

$$= E[a + (X_n - a)^+ - a - (X_0 - a)^+]$$

$$(b-a)\,EU_n \leq E(X_n - a)^+ - E(X_0 - a)^+ \qquad (*)$$

Teorema de Convergência de Martingais

$\{X_n\}_{n\geq 0}$ é um submartingal com $\sup_n E X_n^+ < \infty$ então

existe X, $E|X| < \infty$ t.q. $X_n \to X$ q.c.

Uso $(*)$ + $(X-a)^+ \leq X^+ + |a|$ (verifique)

Então

$$(b-a)\,EU_n \leq E[X_n^+ + |a| - E(X_0 - a)^+] \leq E[X_n^+ + |a|]$$

$\to U_n \uparrow U$

$\Rightarrow EU < \infty \Rightarrow U < \infty$ q.c. para todos $a < b$

Então

$$P(\{\underline{\lim}\, X_n < a < b < \overline{\lim}\, X_n\}) = 0 \quad \forall\, a, b$$

$\therefore$

$$\underline{\lim}\, X_n = \overline{\lim}\, X_n = \lim X_n \qquad \square$$

Análogo para supermartingal.

Teo. $\{X_n\}_{n \geq 0}$ supermartingal $X_n \geq 0$. Então, existe X, $X_n \to X$ q.c. e $EX \leq EX_0$. (Exercício)

Teorema da Decomposição de Doob.

Todo submartingal $\{X_n\}_{n \geq 0}$ pode ser decomposto, de forma única, como

$$X_n = M_n + A_n$$

onde $\{M_n\}_{n \geq 0}$ é um martingal e $\{A_n\}_{n \geq 0}$ é previsível e crescente ("Drift")

Dem: Quero $X_n = M_n + A_n$ com

$$E[M_n | \mathcal{F}_{n-1}] = M_{n-1} \quad e \quad A_n \in \mathcal{F}_{n-1}$$

$$E[X_n \mid \mathcal{F}_{n-1}] = E[M_n \mid \mathcal{F}_{n-1}] + E[A_n \mid \mathcal{F}_{n-1}] =$$

$$= M_{n-1} + A_n + A_{n-1} - A_{n-1} =$$

$$= X_{n-1} + (A_n - A_{n-1}) \Rightarrow$$

$(A_n - A_{n-1}) = E[X_n \mid \mathcal{F}_{n-1}] - X_{n-1}$ que é previsível

$A_0 = 0$

A_n por indução. $(A_n \in \mathcal{F}_{n-1})$.

Urna de Polya

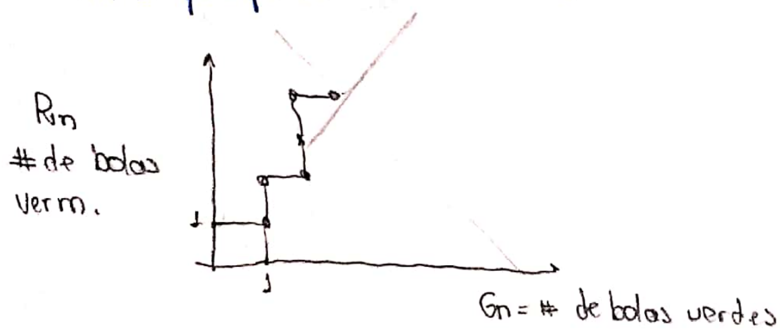
Urna com r bolas vermelhas

g " verdes

sorteio uma, devolvo junto com c bolas da mesma cor.

<u>Caso particular</u>: $r = g = c = 1$

Y_n = prop. de bolas verdes



Exercício: $P(G_n = m) = \frac{1}{n}$.

X_n é martingal

$E[X_n | \mathcal{F}_{n-1}] = X_{n-1}$ $\qquad X_n \geq 0$

$(n-1)$ $\left[\begin{array}{l} i \text{ verm.} \\ j \text{ verdes} \end{array}\right.$ $\qquad X_{n-1} = \frac{1}{i+j}$

$\begin{array}{l} i+c \\ j \end{array}$ se vermelha com prob $\frac{i}{i+j}$ $\quad \searrow \frac{j}{i+j+c}$

$\begin{array}{l} i \\ j+c \end{array}$ se verde com prob. $\frac{j}{i+j}$ $\quad \searrow \frac{j+c}{i+j+c}$

$$\frac{j+c}{i+j+c} \cdot \frac{j}{i+j} + \frac{j}{i+j+c} \cdot \frac{i}{i+j} = \frac{j}{i+j}\left[\frac{j+c}{i+j+c} + \frac{i}{i+j+c}\right]$$

Probabilida Avançada, Aula 1, Tarefas 26/02

Logo,

$$A_N^c = A_1 \cup A_2 \cup \ \ \cup A_N$$, onde

A_i = "a carta i está na posição i"

Assim,

$$\mathbb{P}(A_N^c) = \sum_{i=1}^{N} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) + \dots +$$

$$+ (-1)^{N+1} \mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_N)$$

Mas

$$\mathbb{P}(A_i) = \frac{(N-1)!}{N!} = \frac{1}{N}, \quad i = 1, \dots, N$$

$$\mathbb{P}(A_i \cap A_j) = \frac{(N-2)!}{N!} = \frac{1}{N(N-1)}, \quad i, j \in \{1, \dots, N\}, i < j$$

$\vdots$

$$\mathbb{P}(A_1 \cap \dots \cap A_N) = \frac{1}{N!}$$

• O nº de termos na soma

$\sum_{i \in D} \mathbb{P}(\bigcap_{i \in D} A_i)$ é $\binom{N}{N-|D|}$

D o conjunto de índices.

Daí,

$$\mathbb{P}(A_N^c) = \sum_{i=1}^{N} \frac{1}{N} - \sum_{i<j} \frac{(N-2)!}{N!} + \sum_{i<j<k} \frac{(N-3)!}{N!} + \dots + (-1)^{N+1} \cdot \frac{1}{N!}$$

$$= N \cdot \frac{1}{N} - \frac{(N-1)N}{2} \cdot \frac{(N-2)!}{N!} + \frac{N!}{(N-3)!\,3!} \frac{(N-3)!}{N!} + \dots + =$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \dots + (-1)^{N+1} \frac{1}{N!}$$

Assim,

$$P(A) = 1 - P(A^c) = 1 - \left(1 - \frac{1}{2!} + \frac{1}{3!} - \dots + (-1)^{N+1}\frac{1}{N!}\right)$$

$$= \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots + (-1)^{N+1}\frac{1}{N!}$$

Conforme $N \to \infty$, $P(A) \to e^{-1}$, pois $e^{-1} = 1 - 1 + \frac{(-1)^2}{2!} + \frac{(-1)^3}{3!} + \dots$

$$= 1/2! - 1/3! + 1/4! - \dots .$$

Agora, qual o valor esperado de N: v.a. representando o número de cartas em suas respectivas posições

Devemos notar que

$$N = X_1 + \dots + X_N, \quad X_i = \begin{cases} 1 & \text{se a carta } i \text{ está na p. } i \\ 0 & \text{, c.c.} \end{cases}$$

Pela linearidade da esperança,

$$E[N] = \sum_{i=1}^{N} E(X_i) = \sum_{i=1}^{N} \left(1 \cdot \frac{(N-1)!}{N!}\right) = \frac{N\,(N-1)!}{N!} = 1 _{\blacksquare}$$

- Axioma da Escolha
- ~~Axioma~~ Paradoxo de Banach-Tarkis

Probabilidade Avançada I, Aula 1, Tarefas

Mostre que toda σ-álgebra é uma álgebra.

Seja $\mathcal{F}$ uma σ-álgebra dado. Pela definição, sabemos que

1. $\Omega \in \mathcal{F}$, $\emptyset \in \mathcal{F}$
2. Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$
3. Se $A_1, A_2, \ldots$ estão em $\mathcal{F}$, então

$$\bigcup_i A_i \in \mathcal{F}$$

Para que $\mathcal{F}$ seja uma álgebra, basta garantir que dados A_1, A_2, A_n em $\mathcal{F}$, então $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$.

Para tal, tomamos $A_1, A_2, \ldots, A_n$ e $A_i = \emptyset$ para $i > n$. Da propriedade 3, temos que

$$\bigcup_i A_i = \bigcup_{i=1}^{n} A_i \in \mathcal{F} \qquad \blacksquare$$

A união de σ-álgebras é uma σ-álgebra?

Não! Basta considerar o seguinte contra-exemplo

$\Omega = \{a, b, c, d\}$

$\mathcal{A}_1 = \{\{a, b\}, \{c, d\}, \emptyset, \{a, b, c, d\}\}$ e $\mathcal{A}_2 = \{\{a, c\}, \{b, d\}, \emptyset, \{a, b, c, d\}\}$

Note que

$\mathcal{F} = \mathcal{A}_1 \cup \mathcal{A}_2 = \{\{a, b\}, \{a, c\}, \{b, d\}, \{c, d\}, \emptyset, \{a, b, c, d\}\}$

Tomemos os eventos $A_1 = \{a, b\}$, $A_2 = \{b, d\}$. Temos $A_1^c = \{c, d\}$, $A_2^c = \{a, c\}$, que continuam pertencendo a $\mathcal{F}$. $A_1^c \cup A_2^c = \{a, c, d\} \in \mathcal{F}$. Mas,

$(A_1^c \cup A_2^c)^c = \{b\} \notin \mathcal{F}$. Logo, não é σ-álgebra.

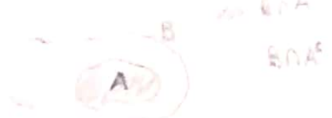
Probabilidade Avançada, Aula 2, Tarefas

Propriedades da medida de probabilidade.

a) monotonicidade Se $A \subset B$, então $P(A) \leq \mathbb{P}(B)$

Podemos escrever B como a união disjunta de $(B \cap A)$ e $(B \cap A^c)$, i.e.,

$$B = (B \cap A) \cup (B \cap A^c)$$



Daí, pela aditividade da prob,

$$P(B) = P(B \cap A) + P(B \cap A^c) \overset{B \cap A = A}{=}$$
$$= P(A) + \underbrace{P(B \cap A^c)}_{\geq 0} \implies P(B) \geq P(A) \quad \square$$

b) Sub-σ-aditividade. Se $A_1, A_2, \ldots$ estão em $\mathcal{F}$, então

$$P(\bigcup_i A_i) \leq \sum_i P(A_i)$$

Se $B_1 = A_1$ e $B_k = A_k \cap A_1^c \cap \ldots \cap A_{k-1}^c$, $k \geq 2$, então os B_k são disjuntos e $\bigcup_k A_k = \bigcup_k B_k$, de modo que

$$P(\bigcup_k A_k) = P(\bigcup_k B_k) = \sum_k P(B_k)$$

Dados l e $m = l+k$,
$B_l = A_l \cap A_1^c \cap \ldots \cap A_{l-1}^c$
$B_m = A_{l+k} \cap A_1^c \ldots \cap A_{l+k-1}^c$
$B_l \cap B_m = A_{l+k} \cap A_l^c \cap \ldots \cap A_{l+k-1}^c \cap \emptyset = \emptyset$

Como $P(B_k) \leq P(A_k)$ pela monotonicidade, temos que

$$P(\bigcup_k A_k) = \sum_k P(B_k) \leq \sum_k P(A_k)$$

c) continuidade por seq. crescentes. Se $A_n \uparrow A$ (isto é, $A_1 \subset A_2 \subset \ldots$ e $\bigcup A_i = A$)

$$P(A_n) \xrightarrow[n \uparrow \infty]{} P(A)$$

Uma vez que $A_1 \subset A_2 \subset \ldots$, segue que $P(A_i)$ é não decrescente em i pela monotonicidade. Como também é limitada, então $\lim_{n \uparrow \infty} P(A_n)$ existe.

d

Considerando agora $B_1 = A_1$, $B_k = A_k \cap A_{k-1}^c$, temos que os B_k são disjuntos ($B_k \cap B_\ell = B_k \cap B_{k+p} = A_k \cap A_{k-1}^c \cap A_{k+p} \cap A_{k+p-1}^c = A_k \cap A_{k+p-1}^c = \emptyset$), e $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$. Temos

$$P(A) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = P(B_1) + P(B_2) + \quad =$$

$$= [P(A_1) + [P(A_2) - P(A_1)] + \cdots \quad]$$

Daí, pela definição de convergência de séries infinitas,

$$P(A) = \lim_{n\to\infty} [P(A_1) + P(A_2) - P(A_1) + \cdots + P(A_n) - P(A_{n-1})] =$$

$$= \lim_{n\to\infty} P(A_n)$$

Resolução Durret.

Seja $B_n = A_n - A_{n-1}$. Então B_n são disjuntos e tem $\bigcup_{m=1}^{\infty} B_n = A$, $\bigcup_{m=1}^{n} B_m = A_n$, de modo que

$$\mu(A) = \sum_{m=1}^{\infty} \mu(B_m) = \lim_{n\to\infty} \sum_{m=1}^{n} \mu(B_m) = \lim_{n\to\infty} \mu(A_n)$$

$\downarrow$

série converge pro limite do n-ésimo termo quando $n\to\infty$.

$A_n \downarrow A$

d) Continuidade por sequências decrescentes. $A_1 \supset A_2 \supset \ldots$ $\bigcap_{k=1}^{\infty} A_k = A$, $P(A) = \lim P(A_n)$

Note que $A_1^c \subset A_2^c \subset \ldots$, i.e., $A_n^c \uparrow A^c$. Daí, pelo item (c)

$$P(A^c) = \lim_{n\to\infty} P(A_n^c) \Leftrightarrow$$

$$1 - P(A) = \lim_{n\to\infty} 1 - P(A_n) \Leftrightarrow$$

$$P(A) = \lim_{n\to\infty} P(A_n) \Rightarrow P(A_n) \xrightarrow[n\to\infty]{} P(A)$$

Probabilidade Avançada, Aula 02, Tarefas

Exercício. $\mathcal{A} = \emptyset$ mais todas as uniões finitas de elementos disjuntos de $\mathcal{I}$, onde

$\mathcal{I} =$ classe de intervalos da forma $(a,b]$, $a \leq b$ em $\mathbb{R}$.

$\mathcal{A}$ é uma álgebra. (Mostre)

Pela definição, $\mathcal{A}$ será uma álgebra se satisfazer as três seguintes propriedades

1. $\Omega \in \mathcal{A}$, $\emptyset \in \mathcal{A}$
2. Se $A \in \mathcal{A}$, então $A^c \in \mathcal{A}$
3. Se $A, B \in \mathcal{A}$, então $A \cup B \in \mathcal{A}$

Antes de provar isso, vejamos que $\mathcal{I}$ é uma semi-álgebra, pois se $(a,b] \in \mathcal{I}$ e $(c,d] \in \mathcal{I}$, então

$$(a,b] \cap (c,d] = \begin{cases} \emptyset = (a,a] \\ (c,b] \text{ se } a \leq c \leq b \leq d \\ (c,d] \text{ se } a \leq c \leq d \leq b \\ (a,b] \text{ se } c \leq a \leq b \leq d \\ (a,d] \text{ se } c \leq a \leq d \leq b \end{cases}$$

E $(a,b]^c = \underbrace{(-\infty,a]}_{\in \mathcal{I}} \cup \underbrace{(b,+\infty]}_{\in \mathcal{I}}$. Usualmente dizem que se $A \subset B$, $A, B \in \mathcal{I}$, então $B - A$ é união finita de elementos de $\mathcal{I}$.

Aí, toma $A = (a,b]$ e $B = (a-p, b+q]$, $p, q \geq 0$

$B - A = (a-p, a] \cup (b, b+q]$

e

3. Se $A = \bigcup_{k=1}^{m} (a_k, b_k]$ e $B = \bigcup_{j=1}^{n} (c_j, d_j]$, então

$$A \cup B = \left(\bigcup_{k=1}^{m} (a_k, b_k]\right) \cup \left(\bigcup_{j=1}^{n} (c_j, d_j]\right) =$$

$$= \bigcup_{l=1}^{m+n} \left(a_l \mathbb{1}_{\{l \leq m\}} + c_{l-m} \mathbb{1}_{\{l > m\}}, b_l \mathbb{1}_{\{l \leq m\}} + d_{l-m} \mathbb{1}_{\{l > m\}}\right]$$

que é união finita de intervalos da forma $(a, b]$, $a < b$ $\Rightarrow$ $A \cup B \in \mathcal{A}$.

2. $A = \bigcup_{i=1}^{n} (a_i, b_i]$, então

$$A^c = \bigcap_{i=1}^{n} \left[(-\infty, a_i] \cup (b_i, +\infty]\right] = (-\infty, \min\{a_i, i=1,\dots,n\}] \cup (\max\{b_i, i=1,\dots,n\}, \infty]$$

que união disjuntas de intervalos em $\mathcal{J}$. Logo $A^c \in \mathcal{A}$

1. Se $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ e, por 3, $A \cup A^c = \Omega \in \mathcal{S}$. Na verdade, poderíamos ter tomado $\Omega = (-\infty, \infty] = \mathbb{R}$, e $\emptyset = (a, a]$.

Mostre que $\mathcal{M}$ (classe dos conjuntos mensuráveis é uma σ-álgebra)

1. $\Omega \in \mathcal{M}$ ($\mu^*(E \cap \Omega) + \mu^*(E \cap \Omega^c) = \mu^*(E) + \mu^*(\emptyset) = \mu^*(E)$ ✓

2. Se $A \in \mathcal{M}$, então para qualquer $D \subset \Omega$, temos

$$\mu^*(D) = \mu^*(D \cap A) + \mu^*(D \cap A^c) =$$
$$= \mu^*(D \cap A^c) + \mu^*(D \cap (A^c)^c)$$

o que implica que $A^c \in \mathcal{M}$. Assim, $\mathcal{M}$ é fechada pela complementação.

3. Resta provar que $\mathcal{M}$ é fechada sob uniões enumeráveis

Lista 1. 24/03

1.1. Feito nos exercícios da Aula 02.

1.3. Seja $\mathbb{R}^n = \{(x_1, \dots, x_n) : x_i \in \mathbb{R}\}$. $\mathcal{R}^n$ = os conjuntos de Borel de $\mathbb{R}^n$ é definido como a σ-álgebra gerada pelos subconjuntos abertos de $\mathbb{R}^n$. Prove que esta é a mesma $\mathcal{R} \times \mathcal{R} \times \dots \times \mathcal{R}$ = a σ-algebra gerada pelos cjtos da forma $A_1 \times \dots \times A_n$. Dica: Mostre que ambas coincidem com a gerada por $(a_1, b_1) \times \dots \times (a_n, b_n)$.

É suficiente mostrar que se $\mathcal{F}$ é a σ-álgebra gerada por $(a_1, b_1) \times \dots \times (a_n, b_n)$, então $\mathcal{F}$ contém

(i) os conjuntos abertos e,

(ii) todos os conjuntos da forma $A_1 \times \dots \times A_n$, onde $A_i \in \mathcal{R}$

Para (i), note que se G é um conjunto aberto e $x \in G$ então há um conjunto da forma $(a_1, b_1) \times \dots \times (a_n, b_n)$ com $a_i, b_i \in \mathbb{Q}$ que contém x e cai em G, de modo que qualquer conjunto aberto é uma união enumerável dos nosso conjuntos basicos.

Para (ii), fixe $A_2, \dots, A_n$ e seja $\mathcal{G} = \{A : A \times A_2 \times \dots \times A_n \in \mathcal{F}\}$. Desde que $\mathcal{F}$ é uma σ-álgebra, é fácil ver que se $\Omega \in \mathcal{G}$ então $\mathcal{G}$ é uma σ-algebra de modo que se $\mathcal{G} \subset \mathcal{A}$ então $\mathcal{G} \subset \sigma(\mathcal{A})$. Do último resultado, segue que $A_1 \in \mathcal{R}$, $A_1 \times (a_2, b_2) \times \dots \times (a_n, b_n) \in \mathcal{F}$. Repetindo o último argumento mais $n-1$ vezes, provamos (ii).

1.5 Uma σ-algebra $\mathcal{F}$ é dita countably generated se há uma coleção enumerável $\mathcal{C} \subset \mathcal{F}$ de modo que $\sigma(\mathcal{C}) = \mathcal{F}$. Mostre que $\mathcal{R}^d$ é countably generated.

Os conjuntos da forma $(a_1, b_1) \times \dots (a_d, b_d)$ onde $a_i, b_i \in \mathbb{Q}$ é uma coleção enumerável que gera $\mathcal{R}^d$.

2.1 Mostre que se $\mathcal{A}$ gera $\mathcal{S}$, então $X^{-1}(\mathcal{A}) \equiv \{\{X \in A\} \; A \in \mathcal{A}\}$ gera $\sigma(X) = \{\{X \in B\} : B \in \mathcal{S}\}$

Seja $\mathcal{G}$ a menor σ-álgebra contendo $X^{-1}(\mathcal{A})$. Desde que $\sigma(X)$ é uma σ-álgebra contendo $X^{-1}(\mathcal{A})$, devemos ter $\mathcal{G} \subset \sigma(X)$ e portanto $\mathcal{G} = \{\{X \in B\} : B \in \mathcal{F}\}$ para algum $\mathcal{S} \supset \mathcal{F} \supset \mathcal{A}$. No entanto, se $\mathcal{G}$ e uma σ-álgebra então podemos assumir que $\mathcal{F}$ é. Uma vez que $\mathcal{A}$ gera $\mathcal{S}$, segue que $\mathcal{F} = \mathcal{S}$.

2.3 Mostre que se f é contínua e $X_n \to X$ q.c. então

$$f(X_n) \xrightarrow{q.c.} f(X)$$

Seja $\Omega_0 = \{\omega : X_n(\omega) \to X(\omega)\}$. Se $\omega \in \Omega_0$ segue da def. de continuidade que $f(X_n(\omega)) \to f(X(\omega))$. Desde que $P(\Omega_0) = 1$, o resultado desejado segue.

24/08

2.4

(i) Mostre que uma função contínua de $\mathbb{R}^d \to \mathbb{R}$ é um mapa mensurável de $(\mathbb{R}^d, \mathcal{R}^d)$ em $(\mathbb{R}, \mathcal{R})$.

(ii) Mostre que $\mathcal{R}^d$ é a menor σ-álgebra que faz todas as funções contínuas mensuráveis.

i) Se G é um cjto aberto então $f^{-1}(G)$ é aberto e portanto mensurável. Agora use $\mathcal{A}$ = a coleção de conjuntos abertos em (2.1)

ii) Seja G um cjto aberto e seja $f(x)$ a distância de x ao complemento de G, i.e., $\inf\{|x-y| : y \in G^c\}$. f é contínua e $\{f > 0\} = G$, assim precisamos todos os cjtos abertos pra fazer todas as funções contínuas mensuráveis.

3.3 Suponha que φ é estritamente convexa, i.e., $>$ vale para $\lambda \in (0,1)$. Mostre que, sob as suposições de (3.2),

$$\varphi(EX) = E\varphi(X) \overset{q.c.}{\Rightarrow} X = EX$$

Relembre da prova de (5.2) no apêndice. Deixamos $\ell(x) \leq \varphi(x)$ sendo uma função linear com $\ell(EX) = \varphi(EX)$ e notamos que $E\varphi(X) \geq E\ell(X) = \ell(EX)$. Se a igualdade permanece então Ex. 3.1 implica que $\varphi(X) = \ell(X)$ q.c. Quando φ é estritamente convexa temos $\varphi(x) > \ell(x)$ para $x \neq EX$ assim devemos ter $X = EX$ q.c.

3.7 Two non existent lower bounds.

Mostre que:

(i) Se $\varepsilon > 0$, $\inf\{P(|X| > \varepsilon) : EX = 0, var(X) = 1\} = 0$.

(ii) se $y \geq 1$, $\sigma^2 \in (0, \infty)$, $\inf\{P(|X| > y) : EX = 1, var(X) = \sigma^2\} = 0$.

i. Seja $P(X = n) = P(X = -n) = 1/2n^2$, $P(X = 0) = 1 - 1/n^2$ para $n \geq 1$.

ii. Seja $P(X = 1 - \varepsilon) = 1 - 1/n$ e $P(X = 1 + b) = 1/n$ para $n \geq 2$. Para termos $EX = 1$, $var(X) = \sigma^2$ precisamos que

$$-\varepsilon(1 - 1/n) + b(1/n) = 0 \qquad \varepsilon^2(1 - 1/n) + b^2(1/n) = \sigma^2$$

A primeira eq. implica $\varepsilon = b/(n-1)$. Usando esta na segunda, temos

$$\sigma^2 = b^2 \frac{1}{n(n-1)} + b^2 \frac{1}{n} = \frac{b^2}{n-1}$$

4.1.

i. Mostre que se X e Y são independentes então $\sigma(X)$ e $\sigma(Y)$ são.

ii. Inversamente, se $\mathcal{F}$ e $\mathcal{G}$ são independentes, $X \in \mathcal{F}$ e $Y \in \mathcal{G}$, então X e Y são independentes.

i. Se $A \in \sigma(X)$ então segue da def. de $\sigma(X)$ que $A = \{X \in C\}$ para algum $C \in \mathcal{R}$. De forma análoga, se $B \in \sigma(Y)$ então $B = \{Y \in D\}$ para algum $D \in \mathcal{R}$, assim usando estes fatos e a ind. de X e Y,

$$P(A \cap B) = P(X \in C, Y \in D) = P(X \in C)P(Y \in D) = P(A)P(B)$$

24/03

(ii) Pelo outro lado, se $X \in \mathcal{F}$, $Y \in \mathcal{G}$ e $C, D \in \mathcal{R}$ segue da def. de mensurabilidade que $\{X \in C\} \in \mathcal{F}$, $\{Y \in D\} \in \mathcal{G}$. Uma vez que $\mathcal{F}$ e $\mathcal{G}$ são ind., segue que $P(X \in C, Y \in D) = P(X \in C) P(Y \in D)$.

5.1. Seja $X_1, X_2, \dots$ v.a. não-correlacionadas com $EX_i = \mu_i$ e $var(X_i)/i \to 0$ conforme $i \to \infty$. Seja $S_n = X_1 + \dots + X_n$ e $\nu_n = ES_n/n$ então qdo $n \to \infty$, $S_n/n - \nu_n \to 0$ no L^2 e em probabilidade.

Primeiro notemos que

$$\frac{var(X_m)}{m} \to 0 \Rightarrow \text{para qq}^{er}\ \varepsilon > 0\ \exists A < \infty$$

de modo que $var(X_m) \leq A + \varepsilon m$. Usando esta estimativa e o fato de que $\sum_{m=1}^n m \leq \sum_{m=1}^n 2m - 1 = n^2$

$$E(S_n/n - \nu_n)^2 = \frac{1}{n^2} \sum_{m=1}^{n} var(X_m) \leq A/n + \varepsilon$$

Dado que ε é arbitrário, isto mostra a conv. em L^2 de $S_n/n - \nu_n$ à 0, e a conv. em prob. segue de (5.3).

5.3. Monte Carlo Integration.

(i) Seja f uma função mensurável no $[0,1]$ com $\int_0^1 |f(x)| dx < \infty$. Seja $U_1, U_2, \dots$ i.i.d on $[0,1]$, e seja

$$I_n = n^{-1} (f(U_1) + \dots + f(U_n))$$

Mostre que $I_n \to I \equiv \int_0^1 f\,dx$ em probabilidade.

(ii) Suponha que $\int_0^1 |f(x)|^2 dx < \infty$. Use a des. de Chebyshev para estimar

$P(|I_n - I| > a/n^{1/2})$.

i) Desde que $f(U_1), f(U_2), \ldots$ são independentes e têm $E|f(U_i)| < \infty$ segue da Lei fraca dos grandes números. (5.8)

ii) $P(|I_n - I| > a/n^{1/2}) \leq (n/a^2)\,Var(I_n) = \sigma^2/a^2$ onde $\sigma^2 = \int f^2 - \left(\int f\right)^2$.

1.1.1 Intersecção arbitrária de σ-álgebras é σ-álgebra. Menor σ-álgebra cont. A. ✓

• 2 Mostrar que $(\Omega, \mathcal{F}, P)$ é esp. de prob., $\mathcal{F}$ = sub. t.q. A ou A^c é contável ✓

• 3 $\sigma(\mathcal{S}_d) = \mathcal{R}^d$ ✓ (?)

• 4 $\mathcal{R}^d$ é enumeravelmente gerada ✓ (?)

• 5 Se $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots$, $\bigcup \mathcal{F}_i$ é álgebra. Mas não σ-álgebra - dar exemplo. ✓

• 6 $A \subset \{1, 2, \dots\}$ tem densidade assintótica θ ... ✓ (?)

1.2.1 $Z = X \mathbb{1}_A + Y \mathbb{1}_{A^c}$, $A \in \mathcal{F}$ é uma variável aleatória ✓

• 3 Mostrar que uma f.d. tem no máximo um nº de disc. enumerável. ✓

• 4 Se $F(x) = P(X \leq x)$ é contínua, $Y = F(X) \sim \mathcal{U}(0,1)$ ✓

1.3.1 Se $\mathcal{A}$ gera $\mathcal{S}$, então $X^{-1}(\mathcal{A}) = \{\{X \in A\} : A \in \mathcal{A}\}$ gera $\sigma(X) = \{\{X \in B\}, B \in \mathcal{S}\}$ ✓

• 3 Se f é contínua e $X_n \xrightarrow{a.s} X$ então $f(X_n) \xrightarrow{a.s} f(X)$ ✓

• 4 f. contínua de $\mathbb{R}^d \to \mathbb{R}$ é uma f. mens. de $(\mathbb{R}^d, \mathcal{R}^d)$ em $(\mathbb{R}, \mathcal{R})$. ✓ Falta $\mathbb{1}_{A}$ (ii)

• → $\mathcal{I}$ de funções mens. é a menor cont. f. simples e fechada sob limites pontuais ✓

• 8 Y é mens. com resp. a $\sigma(X) \Leftrightarrow Y = f(X)$ onde $f: \mathbb{R} \to \mathbb{R}$ é mens. ✓

1.4.1 $f \geq 0$ e $\int f \, d\mu = 0 \Rightarrow f = 0$ a.e. ✓

• 2 $f \geq 0$ e $E_{n,m} = \{x : m/2^n \leq f(x) \leq (m+1)/2^n\}$. $\sum m/2^n \, \mu(E_{n,m}) \uparrow_{n \to \infty} \int f \, d\mu$

1.5.1 $\int |fg| \, d\mu \leq \|f\|_1 \|g\|_\infty$ ✓

• 2 $\|f\|_\infty = \lim_{p \to \infty} \|f\|_p$ se μ é med. de prob. ✓ (Usar liminf = limsup)

• 3 Desigualdade de Minkowski ✓

• 5 $g_n \uparrow g$ e $\int g_1^- \, d\mu < \infty \Rightarrow \int g_n \, d\mu \uparrow \int g \, d\mu$ ✓ (Usar TCM)

Soluções.

Exercício 1.

item 1. Monotonicidade

Note que $B = (B \cap A) \cup (B \cap A^c)$ é uma forma de reescrever B como união disjunta de dois conjuntos, uma vez que $(B \cap A) \cap (B \cap A^c) = \emptyset$. Daí, pela σ-aditividade de P, que implica na aditividade de P, segue que

$$P(B) = P(B \cap A) + P(B \cap A^c) \qquad (1)$$

Como $A \subset B$, então $A \cap B = A$ e $P(B \cap A) = P(A)$. Sub. em (1), temos

$$P(B) = P(A) + P(B \cap A^c)$$

Uma vez que P é uma medida, segue $P(C) \geq P(\emptyset) = 0, \forall C \in \mathcal{F}$. Daí, temos $P(B \cap A^c) \geq 0$, o que leva à

$$P(B) \geq P(A).$$

item 2. Subaditividade

Para tal, seja $A'_k = A_k \cap A$ para todo inteiro $k \geq 1$. Considerando agora $B_1 = A'_1$ e $B_k = A'_k - \bigcup_{m=1}^{k-1} (A'_m)$, temos que os B_k's são disjuntos (B_k é o conjunto dos elementos que pertencem a A e A_k, mas não aos $A_{k'}$'s para $k' < k$), e sua união é o conjunto A. Logo, temos que

$$P(A) = \sum_{m=1}^{\infty} P(B_m) \leq \sum_{m=1}^{\infty} P(A_m)$$

pois $A = +_{m \geq 1} B_m$ e $B_m \subseteq A_m$, $\forall m \geq 1$.

3. Continuidade por baixo.

Seja $B_n = A_n - A_{n-1} = A_n \cap A_{n-1}^c$. Logo, os B_ns são disjuntos (dados n_1 e n_2 inteiros $B_{n_1} \cap B_{n_2} = A_{n_2} \cap A_{n_2-1}^c \cap A_{n_1} \cap A_{n_1-1}^c = A_{min(n_1,n_2)} \cap A^c_{min(n_1,n_2)} = \phi$) e $\bigcup_{m=1}^{\infty} B_m = A$, $\bigcup_{m=1}^{n} B_m = A_n$. Assim,

$$P(A) = \sum_{m=1}^{\infty} P(B_m) = \lim_{n \to \infty} \sum_{m=1}^{n} P(B_n) = \lim_{n \to \infty} P(A_n)$$

4. Continuidade por cima.

Seja $B_n = A_1 - A_n$. Nesse caso, $B_n \uparrow A_1 - A$, i.e., $A_1 - A_n \uparrow A_1 - A_n$. Logo, pela continuidade por baixo de P, segue que

$$P(A_1 - A_n) \uparrow P(A_1 - A)$$

Mas $A_n \subset A_1$ e $A \subset A_1$, de modo que $P(A_1 - A_n) = P(A_1) - P(A_n)$ e $P(A_1 - A) = P(A_1) - P(A)$. Daí

$$P(A_1) - P(A) = \lim_{n \to \infty} P(A_1) - P(A_n) \Leftrightarrow$$

$$- P(A) = - \lim_{n \to \infty} P(A_n)$$

$$P(A) = \lim_{n \to \infty} P(A_n)$$

Como $P(A_n)$ decresce pela monotonicidade de P, segue que $P(A_n) \downarrow P(A)$

Questão 2.

i. Uma coleção $\mathcal{A}$ de subconjuntos de um espaço Ω é dita uma σ-álgebra de Ω se satisfaz as seg. condições

1. $\Omega \in \mathcal{A}$; $\emptyset \in \mathcal{A}$;
2. Se $A \in \mathcal{A}$, então $A^c \in \mathcal{A}$;
3. Se $A_1, A_2, \dots$ estão em $\mathcal{A}$, então $\bigcup_i A_i \in \mathcal{A}$.

Logo, basta verificar que $\bigcap_{i \geq 1} \mathcal{F}_i$ satisfaz as três condições acima, onde $\mathcal{F}_i$ é uma σ-álgebra de Ω para todo $i \geq 1$.

1. $\Omega \in \bigcap_{i\geq 1} \mathcal{F}_i$, $\emptyset \in \bigcap_{i \geq 1} \mathcal{F}_i$.
Note que $\Omega \in \mathcal{F}_i$ para todo i. Logo $\Omega \in \bigcap_{i \geq} \mathcal{F}_i$.
Analogamente, $\emptyset \in \bigcap_{i\geq 1} \mathcal{F}_i$.

2. Se $A \in \bigcap_{i\geq 1} \mathcal{F}_i$, então $A^c \in \bigcap_{i \geq 1} \mathcal{F}_i$:
$A \in \bigcap_{i \geq 1} \mathcal{F}_i \Rightarrow A \in \mathcal{F}_i$ para todo $i \geq 1 \Rightarrow A^c \in \mathcal{F}_i \; \forall i \geq 1 \Rightarrow$
$A^c \in \bigcap_{i\geq 1} \mathcal{F}_i$. Logo, essa condição também é verificada.

3. Se $A_1, A_2, \dots \in \bigcap_{i\geq 1} \mathcal{F}_i$, então $\bigcup_j A_j \in \bigcap_{i \geq 1} \mathcal{F}_i$
$A_1, A_2, \dots \in \bigcap_{i\geq 1} \mathcal{F}_i \Rightarrow A_1, A_2, \dots \in \mathcal{F}_i$ para todo $i \geq 1 \Rightarrow$
$\bigcup_j A_j \in \mathcal{F}_i$ para todo $i \geq 1 \Rightarrow \bigcup_j A_j \in \bigcap_{i\geq 1} \mathcal{F}_i$.

Logo $\bigcap_{i\geq 1} \mathcal{F}_i$ é uma σ-álgebra de Ω.

ii. Seja $F(\mathcal{A}) = \{\mathcal{F} : \mathcal{F}$ é σ-álgebra que contem $\mathcal{A}\}$. Logo,
$\bigcap_{\mathcal{F} \in F(\mathcal{A})} \mathcal{F}$ é σ-álgebra (pelo item i) e que contem $\mathcal{A}$. Ademais

$\bigcap_{\mathcal{F} \in \mathbb{F}(\mathcal{A})} \mathcal{F} \subset \mathcal{F} \; \forall \mathcal{F} \in \mathbb{F}(\mathcal{A})$. Logo $\bigcap_{\mathcal{F} \in \mathbb{F}(\mathcal{A})} \mathcal{F}$ é a menor σ-álgebra ("em termos de cardinalidade") que contém $\mathcal{A}$, e que chamamos de σ-álgebra gerada por $\mathcal{A}$, denotada por $\sigma(\mathcal{A})$.

Questão 3.

Seja $\mathbb{R}^d = \{(x_1, \ldots, x_d) : x_i \in \mathbb{R}\}$ e $\mathcal{R}^d$ o subconjunto de Borel, a σ-álgebra contendo os conjuntos abertos de $\mathbb{R}^d$.

Temos ϕd = conjunto vazio mais todos os conjuntos da forma

$$(a_1, b_1] \times \ldots \times (a_d, b_d] \subset \mathbb{R}^d \quad \text{onde } -\infty \leq a_i < b_i \leq \infty$$

Queremos provar que $\sigma(\phi d) = \mathcal{R}^d$.

Como $\sigma(\phi d)$ é a menor σ-álgebra que contém ϕd, para mostrar que $\sigma(\phi d) = \mathcal{R}^d$, basta mostrar que $\mathcal{R}^d$ contém ϕd e que está contido em $\sigma(\phi d)$.

De fato, a notarmos que

$$(a_i, b_i] = \bigcap_{n=1}^{\infty} (a_i, b_i + 1/n)$$, temos que

$$\underbrace{(a_1, b_1] \times \ldots \times (a_d, b_d]}_{A} = \bigcap_{n=1}^{\infty} (a_1, b_1 + 1/n) \times \ldots \times \bigcap_{n=1}^{\infty} (a_d, b_d + 1/n) =$$

$x \in A \Leftrightarrow x_i \in (a_i, b_i + 1/n) \; \forall n \Leftrightarrow x \in (a_1, b_1 + 1/n) \times \ldots \times (a_d, b_d + 1/n)$ para todo n

$\Leftrightarrow x \in \bigcap_{n=1}^{\infty} (a_1, b_1 + 1/n) \times \ldots \times (a_d, b_d + 1/n) \in \mathcal{R}$

Se $A \in \sigma(\phi d)$, então $A = \bigcup (a_1, b_1] \times$ (...)

$= \bigcup \bigcap ($ ... $) \in \mathcal{R}$. Logo $\sigma(\phi d) \subset \mathcal{R}$

Questão 4.

Do enunciado, temos

$$P(A) = \begin{cases} 0 & \text{se } A \text{ é contável} \\ 1 & \text{se } A^c \text{ é contável} \end{cases}$$

Um espaço de probabilidade é uma tripla $(\Omega, \mathcal{F}, P)$ em que Ω é um conjunto de "realizações", $\mathcal{F}$ é um conjunto de "eventos" e $P: \mathcal{F} \to [0,1]$ é uma função que atribui probabilidades a eventos. Supomos que $\mathcal{F}$ é uma σ-álgebra, i.e., uma coleção não-vazia de subconjuntos de Ω que satisfaz

i) se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$; e
ii) se $A_i \in \mathcal{F}$ é uma seq. contável de conjuntos, então $\bigcup_i A_i \in \mathcal{F}$

Vamos verificar as propriedades acima.
i) Se $A \in \mathcal{F}$, então
(*) A é contável, ou
(**) A^c é contável
Se vale (*) então $A^c \in \mathcal{F}$, pois $(A^c)^c = A$ é contável. Por outro lado, se vale (**), $A^c \in \mathcal{F}$ pois A^c é contável. Assim, (i) é válida.

ii) Se $A_i \in \mathcal{F}$ é uma sequência contável de conjuntos, então
(*) todos os A_i's são contáveis, ou
(**) pelo menos um A_i^c é contável

Se (*) é verdadeira, então $\bigcup_i A_i$ é contável e, portanto, $\bigcup_i A_i \in \mathcal{F}$.
Se (**) é verdadeira, $(\bigcup_i A_i)^c = \bigcap_i A_i^c$ é contável e, logo, $\bigcup_i A_i \in \mathcal{F}$.

Agora, vamos mostrar que $P: \mathcal{F} \to [0,1]$ é uma medida de probabilidade. Para isso, precisamos verificar que

i) $P(A) \geq P(\emptyset) = 0, \ \forall A \in \mathcal{F}$

ii) Se $A_i \in \mathcal{F}$ é uma seq. contável de cjtos disjtos, então
$$P(\bigcup_i A_i) = \sum_i P(A_i)$$

iii) $P(\Omega) = 1$.

Se $A \in \mathcal{F}$, então A ou A^c é contável. No primeiro caso, $P(A) = 0$. No segundo, A não é contável e, por isso, $P(A) = 1 \geq 0$. Quanto à $\emptyset$, este é contável e, portanto, $P(\emptyset) = 0$. Logo, vale (i).

Se $A_i \in \mathcal{F}$ uma seq. contável de cjtos disjuntos, podemos ter uma das duas alternativas:

1. todos os A_i's são contáveis ou
2. um dos A_i's não é contável (Não se pode ter mais de um A_i não-contável - digamos que A_{i_1} e A_{i_2} - porque, nesse caso, também $A_{i_1}^c$ e $A_{i_2}^c$ seriam não contáveis, já que, como A_{i_1} e A_{i_2} são disjuntos, teríamos $A_{i_2}^c \supset A_{i_1}$, e, sendo A_{i_1} incontável, $A_{i_2}^c$ também precisaria sê-lo. Analogamente, temos que $A_{i_1}^c$ tbm necessitaria ser incontável. Portanto, pode existir no máximo um conjunto incontável na seq. de eventos disjuntos $A_i \in \mathcal{F}$).

Caso 1. seja verdadeira, $P(\bigcup_i A_i) = 0 = \sum_i P(A_i)$, pois a união de conj. contáveis tbm é contável.

Caso valha 2., denotemos por i^* o índice do cjto incontável.

de modo que $P(\bigcup A_i) = 1 = \underbrace{P(A_{i^*})}_{} + \sum_{i \neq i^*} \underbrace{P(A_i)}_{0}$

(iii) $\mathbb{P}(\mathbb{R}) = 1$, já que $\mathbb{R}$ é incontável.

Portanto, de fato $\mathbb{P}: \mathcal{F} \to [0,1]$ é uma medida de probabilidade e, juntamente com o fato de $\mathcal{F}$ ser uma σ-álgebra de conjuntos de $\Omega = \mathbb{R}$, segue que $(\Omega, \mathcal{F}, \mathbb{P})$ é um espaço de probabilidade.

Questão 5.

De acordo com a definição, $\mathcal{R}^d$ será enumeravelmente gerada se existe uma coleção enumeráve $C \subset \mathcal{R}^d$ tal que $\sigma(C) = \mathcal{R}^d$. Logo, basta exibir C tal que $\sigma(C) = \mathcal{R}^d$. Como sabemos, $\mathcal{R}^d$ é a σ-álgebra gerada pelos conjuntos abertos do $\mathbb{R}^d$. Note agora que para qualquer aberto A do $\mathbb{R}^d$,

$$x \in A \iff x \in \bigtimes_{i=1}^{d} (q_{i1}, q_{i2}) \text{ onde } q_{i1} \text{ e } q_{i2} \text{ são racionais}$$

Logo, como cada $x \in A$ tem uma relação

1.1.5.

a) 1. Se $A \in \bigcup_i \mathcal{F}_i$, então $A^c \in \bigcup \mathcal{F}_i$?

$A \in \bigcup_i \mathcal{F}_i \Leftrightarrow A \in \mathcal{F}_i$ para algum i $\Leftrightarrow A^c \in \mathcal{F}_i$ para algum i

$\Leftrightarrow A^c \in \bigcup_i \mathcal{F}_i$

2. Se A e $B \in \bigcup_i \mathcal{F}_i$, $A \cup B \in \bigcup_i \mathcal{F}_i$?

Temos:

$A \in \mathcal{F}_i$ para algum i. Suponha $A \in \mathcal{F}_{i_1}$
$B \in \mathcal{F}_i$ para algum i. Suponha $B \in \mathcal{F}_{i_2}$

Pelo fato de $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots$, segue que $\mathcal{F}_{i_1} \subset \mathcal{F}_{\max\{i_1, i_2\}}$ e $\mathcal{F}_{i_2} \subset \mathcal{F}_{\max\{i_1, i_2\}}$, de modo que $A, B \in \mathcal{F}_{\max\{i_1, i_2\}} \Rightarrow$

$A \cup B \in \mathcal{F}_{\max\{i_1, i_2\}}$. Como exibimos um índice $\max\{i_1, i_2\}$, tal que $A \cup B$ está em $\mathcal{F}_{\max\{i_1, i_2\}}$, segue que

$A \cup B \in \bigcup_i \mathcal{F}_i$. "Na verdade, $A \cup B \in \mathcal{F}_i \ \forall i \geq \max\{i_1, i_2\}$"

b) $A_1, A_2, \ldots \in \bigcup \mathcal{F}_i \Rightarrow \bigcup_j A_j \in \bigcup_i \mathcal{F}_i$?
Suponha $A_i \in \mathcal{F}_{i+1}$, daí $\bigcup_{j=1}^{k} A_j \in \bigcup_{i=1}^{k+1} \mathcal{F}_i$

Não é possível

$\Omega_i = \{j\}_{j=1}^{i}$

$\Omega_1 = \{1\}$ $\{1, \emptyset\}$

$\Omega_2 = \{1, 2\}$ $\{1, 2, \{1,2\}, \emptyset\}$

Seja $\mathcal{F}_i$ uma σ-álgebra de conjuntos de Ω_i. Temos que $\{i\} \in \mathcal{F}_i$ para todo i. Daí, $\mathbb{N} = \bigcup_i \{i\} \in \bigcup_i \mathcal{F}_i \Rightarrow \mathbb{N} \in \mathcal{F}_i$ para algum i $\Rightarrow \mathbb{N} \subset \Omega_i$, o que é uma contradição.

6-

Nesse caso, $\Omega = \mathbb{N}$. Daí

$$\lim_{n \to \infty} \frac{|\mathbb{N} \cap \{1, \dots, n\}|}{n} = \lim_{n \to \infty} \frac{n}{n} = 1$$

Se $\lim_{n \to \infty} \frac{|A \cap \{1, \dots, n\}|}{n} = 0$, então

$$\lim_{n \to \infty} \frac{|A^c \cap \{1, \dots, n\}|}{n} = \lim_{n \to \infty} \frac{|\mathbb{N} \cap \{1, \dots, n\}|}{n} - \frac{|A \cap \{1, \dots, n\}|}{n} =$$

$$= 1 - 0$$

$A_n = \{$

1.1.6. Um conjunto $A \subset \{1,2,\dots\}$ é dito to have asymptotic density θ se

$$\lim_{n\to\infty} |A \cap \{1,2,\dots,n\}| / n = \theta$$

Seja $\mathcal{A}$ a coleção de conjuntos para os quais a densidade assintótica existe. $\mathcal{A}$ é uma σ-álgebra? Uma álgebra?

Seja a coleção de conjuntos $A \subset \mathbb{N}$ para o qual a dens. assintótica

$$\theta(A) = \lim_{n\to\infty} \frac{|A \cap \{1,\dots,n\}|}{n}$$

existe. Seja $E = \{2n : n \geq 1\}$ o conjunto dos inteiros pares positivos e observe que $E \in \mathcal{A}$ e que $\theta(E) = 1/2$. Da mesma forma, seja $O = \{2n+1 : n \geq 0\}$ o conjunto dos inteiros positivos ímpares. Para cada $m \geq 1$, seja

$$K_m = \begin{cases} E \cap \{3^m, \dots, 3^{m+1}-1\} & \text{se } m \text{ é par} \\ O \cap \{3^m, \dots, 3^{m+1}-1\} & \text{se } m \text{ é ímpar} \end{cases}$$

e definir $A = \bigcup_{m \geq 0} K_m$. Então $A \in \mathcal{A}$ com $\theta(A) = 1/2$. No entanto, não é o caso de que $E \cap A \in \mathcal{A}$. Para provar esta afirmação, observe que, se $n = 3^{m+1} - 1$ para m par, então $E \cap K_m = K_m$ e

$$\frac{1}{n} |(E \cap A) \cap \{1,\dots,n\}| \geq \frac{1}{n} |K_m| \geq \frac{3^m}{3^{m+1}} = \frac{1}{3}$$

1.2.4

Se F é estritamente não-decrescente e contínua então

$\exists F^{-1} : [0,1] \to \mathbb{R}$ $\qquad 0 < y < 1$

$$P(Y \leq y) = P(F(X) \leq y) = P(F^{-1}(F(X)) \leq F^{-1}(y)) =$$
$$= P(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y$$

$$f_Y(y) = \begin{cases} 1, & \text{se } y \in (0,1) \\ 0, & \text{c.c} \end{cases} \qquad Y = F(X) \sim U(0,1)$$

1.2.1

Basta notar que para todo $B \subset \mathcal{R}$

$$\{\omega : Z(\omega) \in B\} = (\{\omega : Z(\omega) \in B\} \cap A) \cup (\{\omega : Z(\omega) \in B\} \cap A^c) =$$
$$= (\underbrace{\{\omega : X(\omega) \in B\}}_{\in \mathcal{F}} \cap \underbrace{A}_{\in \mathcal{F}}) \cup (\underbrace{\{\omega : Y(\omega) \in B\}}_{\in \mathcal{F}} \cap \underbrace{A^c}_{\in \mathcal{F}}) \in \mathcal{F}$$

Logo Z é uma variavel aleatória.

Mostre que uma função distribuição tem at most countably many discontinuites.

Solução Ivan:

[Baseado em math.stackexchange.com/questions/147612/discontinuity-points-of-a-distribution-function]

Solução. 1

Seja D o conjunto dos pontos de descontinuidade de F. Para cada $x \in D$, temos $F(x^-) < F(x^+)$, de modo que podemos escolher um número racional q_x com $F(x^-) < q_x < F(x^+)$.

Uma vez que F é crescente, podemos verificar que se $x \neq y$ então $q_x \neq q_y$. Assim, $x \mapsto q_x$ é uma função um-pra-um de D em $\mathbb{Q}$, e $\mathbb{Q}$ é contável. Logo, D também o é.

Solução 2. [Teorema 3.23 de "Real Analysis" (FOLLAND), p.101]

F é crescente. Logo, os intervalos $(F(x^-), F(x^+))$, $x \in \mathbb{R}$, são mutuamente disjuntos

Para $|x| < N$, o intervalo $(F(x^-), F(x^+))$ encontra-se em $(F(-N), F(N))$. $\underline{I}$

Os dois primeiros itens implicam que

$$\sum_{|x|<N} [F(x^+) - F(x^-)] \leq F(+N) - F(-N) < \infty$$

Essa soma num conjunto incontável é definida na p.11 (Essencialmente, é o supremo da soma sobre todas as suas somas parciais).
Ser finita implica que o nº de termos $\neq 0$ da soma é contável.

. Isso implica que o cjto $D_N := \{x \in (-N,N) : F(x^-) \neq F(x^+)\}$ é contável.

O conjunto comple. de descont. é $\bigcup_{N \in \mathbb{N}} D_N$, que é contável (uma vez que é a união contável de conjuntos contáveis)

O truque interessante aqui é que vc mapeia as descontinuidades para uma coleção de intervalos disjuntos e então limita o comprimento total desses intervalos. (Muito útil em geral para provar que um conjunto é contável)

Reprodução.

1.

Seja D o cjto dos pontos de descontinuidade

$\forall x \in D$, temos $P(X < x) < P(X \leq x)$, i.e,

$$F(x-) < F(x)$$

Como entre dois reais sempre há um racional, podemos encontrar um racional q_x t.q. $F(x-) < q_x < F(x)$.

Uma vez que F é crescente, temos que se $x \neq y$, $x, y \in D$ então $q_x \neq q_y$. Logo $x \mapsto q_x$ define uma função um-a-um de D em $\mathbb{Q}$. Como $\mathbb{Q}$ é enumerável, segue que D tbm o é.

2.

1.3.1. Mostre que se $\mathcal{A}$ gera $\mathcal{B}$, então $X^{-1}(\mathcal{A}) = \{\{X \in A\} : A \in \mathcal{A}\}$ gera $\sigma(X) = \{\{X \in B\} : B \in \mathcal{B}\}$

Obs: $\{X \in A\}$ é uma forma abreviada de $\{\omega : X(\omega) \in A\}$

Para mostrar que se $\mathcal{A}$ gera $\mathcal{B}$, então

$$X^{-1}(\mathcal{A}) = \{\{X \in A\} : A \in \mathcal{A}\} \text{ gera } \sigma(X) = \{\{X \in B\} : B \in \mathcal{B}\}$$

devemos verificar que

1. $\sigma(X)$ é uma σ-álgebra;
2. se $\{X \in A\} \in X^{-1}(\mathcal{A})$, então $\{X \in A\} \in \sigma(X)$;
3. $\{X \in B\} \in \sigma(X) \Rightarrow \{X \in B\} \in \mathcal{F}$ $\forall$ $\mathcal{F}$ σ-álgebra de subconjuntos de $X^{-1}(\mathcal{A})$

1 e 2 garantem que $\sigma(X)$ é σ-álgebra e que contém $X^{-1}(\mathcal{A})$ e 3 garante que $\sigma(X)$ é a menor σ-álgebra que contém $X^{-1}(\mathcal{A})$.

1. Para comprovar (1), devemos mostrar que

a) se $\{X \in B\} \in \sigma(X)$, então $\{X \in B\}^c \in \sigma(X)$

b) se $\{X \in B_i\} \in \sigma(X)$ é uma coleção contável de subcjtos, então
$$\bigcup_i \{X \in B_i\} \in \sigma(X)$$

a) $\{X \in B\}^c = \{\omega : X(\omega) \in B\}^c = \{\omega : X(\omega) \in B^c\} = \{X \in B^c\} \in \sigma(X)$
pois $B^c \in \mathcal{B}$, já que $B \in \mathcal{B}$ e $\mathcal{B}$ é σ-álgebra.

b) $\bigcup_i \{X \in B_i\} = \bigcup_i \{\omega : X(\omega) \in B_i\} = \{\omega : X(\omega) \in \bigcup_i B_i\} = \{X \in \bigcup_i B_i\}$
pois $B_i \in \mathcal{B}$, $\forall i$, e $\mathcal{B}$ é σ-álgebra.

2)

Se $\{X\in A\}\in X^{-1}(\mathcal{A})$ Tal subconjto pertence a $\sigma(X)$ pois $A\in\mathcal{A} \Rightarrow A\in\mathcal{B}$ pois $\mathcal{B}$ é σ-álgebra de subconjuntos de $\mathcal{A}$, e $A\in\mathcal{A}$.

3) Seja $\{X\in B\}\in\sigma(X)$. Como $B\in\mathcal{B}$ (pois $\{X\in B\}\in\sigma(X)$), então B pode ser escrito como a união de subconjuntos de $\mathcal{A}$, digamos $B=\bigcup_i A_i$. Assim,

$$\{X\in B\}=\{X\in \bigcup_i A_i\}=\bigcup_i\{X\in A_i\}\in\mathcal{F}$$

pois $\mathcal{F}$ é uma σ-álgebra de subconjuntos de $X^{-1}(\mathcal{A})$ e $\{X\in A_i\}\in X^{-1}(\mathcal{A}), \forall i$.

Portanto, podemos afirmar que $X^{-1}(\mathcal{A})$ gera $\sigma(X)$.

1.3.3

Se $X_n \to X$ quase certamente, então

$$\Omega_0 = \{\omega : X_n(\omega) \to X(\omega)\} \text{ é tal que}$$

$P(\Omega_0) = 1.$

Agora, para todo $\omega \in \Omega_0$, temos

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \overset{*}{\Rightarrow} f \text{ é contínua}$$

$$\lim_{n\to\infty} f(X_n(\omega)) = f(X(\omega))$$

de modo que se $\omega \in \Omega_0$, então $\omega \in \Omega_1$, onde $\Omega_1 = \{\omega : f(X_n(\omega)) \to f(X(\omega))\}$

i.e., $\Omega_0 \subset \Omega_1$, e portanto

$$P(\Omega_1) \geq P(\Omega_0) = 1 \Rightarrow P(\Omega_1) = 1, \text{ j.q. } P \text{ e uma med. de prob} \Rightarrow f(X_n) \to f(X) \text{ quase certamente.}$$

* Teorema. Sejam f e g duas funções tais que $\operatorname{Im} f \subset D_g$. Se $\lim_{x\to p} f(x) = a$ e g contínua em a, então

$$\lim_{x\to p} g(f(x)) = \lim_{u\to a} g(u) = g(a)$$

Demonst.

Sendo g contínua em a, $\lim_{u\to a} g(u) = g(a)$. Prec. provar que p/ todo $\varepsilon > 0$ dado, existe $\delta > 0$ t.q.

$$0 < |x - p| < \delta \Rightarrow g(a) - \varepsilon < g(f(x)) < g(a) + \varepsilon$$

Como g é cont. em a, dado $\varepsilon > 0$, existe $\delta_1 > 0$ tal que

(2) $$a-\delta_1 < u < a+\delta_1 \Rightarrow g(a)-\varepsilon < g(u) < g(a)+\varepsilon$$

Como $\lim_{x \to p} f(x) = a$, para o $\delta_1 > 0$ acima existe $\delta > 0$ tal que

(3) $$0 < |x-p| < \delta \Rightarrow a-\delta_1 < f(x) < a+\delta_1$$

De (2) e (3), segue

$$0 < |x-p| < \delta \Rightarrow g(a)-\varepsilon < g(f(x)) < g(a)+\varepsilon$$

Obs.

O Teorema acima conta-nos que, se g for contínua em a e $\lim_{x \to p} f(x) = a$ então $\lim_{x \to p} g(f(x)) = g(a) = g(\lim_{x \to p} f(x))$, o que nos mostra que os símbolos $\lim_{x \to p}$ e g podem ser permutados em $\lim_{x \to p} g(f(x))$.

$$\lim_{x \to p} g(f(x)) = g(\lim_{x \to p} f(x))$$

3.4

i) Mostre que uma função contínua de $\mathbb{R}^d \to \mathbb{R}$ é uma função mensurável de $(\mathbb{R}^d, \mathcal{B}^d) \to (\mathbb{R}, \mathcal{B})$

Seja $f: \mathbb{R}^d \to \mathbb{R}$ uma f. contínua.

Quero mostrar que f é mensurável de $(\mathbb{R}^d, \mathcal{B}^d) \to (\mathbb{R}, \mathcal{B})$

Para isso, primeiro mostro que f é mensurável de $(\mathbb{R}^d, \mathcal{B}^d) \to (\mathbb{R}, \mathcal{A})$, onde $\mathcal{A} = \{(-\infty, a) : a \in \mathbb{R}\}$

$$f^{-1}(A) = \{(x_1, \dots, x_d) : f(x_1, \dots, x_d) \in A\}, \ \forall A \in \mathcal{A}$$
$$f^{-1}(A) = \{(x_1, \dots, x_d) : f(x_1, \dots, x_d) < a\}, \ \forall A \in \mathcal{A}$$

Como f é contínua, a imagem inversa de um aberto é um aberto, e segue que

$$f^{-1}(A) = \{(x_1, \dots, x_d) : f(x_1, \dots, x_d) < a\} \in \mathcal{B}^d, \ \forall A \in \mathcal{A}$$

Pelo Teorema 1.3.1 segue que se $f^{-1}(A) \in \mathcal{B}^d, \forall A \in \mathcal{A}$ então f é mensurável em $(\mathbb{R}^d, \mathcal{B}^d) \to (\mathbb{R}, \underbrace{\sigma(\mathcal{A})}_{\mathcal{B}})$

iii) Mostre que $\mathcal{B}^d$ é a menor σ-álgebra que faz todas as funções contínuas mensuráveis

Seja $\mathcal{F}_c$ a classe de funções cont. $f: \mathbb{R}^d \to \mathbb{R}$. Queremos encontrar $\sigma(f) \ \forall f \in \mathcal{F}_c$ e mostrar que $\bigcap_{f \in \mathcal{F}_c} \sigma(f) = \mathcal{B}^d$.

Para f contínua,

$$\sigma(f) = \{\{f \in B\} : B \in \mathcal{B}\}$$

Basta lembrar que

1.3.7. Uma função $\varphi : \Omega \to \mathbb{R}$ é dita simples se

$$\varphi(\omega) = \sum_{m=1}^{n} c_m \mathbb{1}_{A_m}(\omega)$$

onde c_m são ~~valores~~ reais e $A_m \in \mathcal{F}$. Mostre que a classe de funções $\mathcal{F}$-mensuráveis é a menor classe contendo as funções simples e closed under pointwise limits.

Seja $\mathcal{F}_m$ a classe de funções mensuráveis $f : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{R})$. Uma vez que a função indicadora de qualquer conjunto mensurável $A \in \mathcal{F}$ é mensurável em relação à $\mathcal{F}$ ($\{\mathbb{1}_A \in B\} = \{\omega : \mathbb{1}_A(\omega) \in B\} = \Omega$, se $\{0,1\} \in B$; A se $1 \in B$, mas $0 \notin B$; A^c se $0 \in B$, mas $1 \notin B$, $\emptyset$, se nem 0 nem 1 estão em B. Mas todos pertencem à $\mathcal{F}$), temos que cada função simples

$$\varphi(\omega) = \sum_{m=1}^{n} c_m \mathbb{1}_{A_m}(\omega)$$

soma de mens. é mensurável

pertence a $\mathcal{F}_m$. Além disso, pela teoria, sabemos que $\mathcal{F}_m$ é fechada sob limites pontuais. Para mostrar que $\mathcal{F}_m$ é a menor classe de funções $\mathcal{F}$-mensuráveis que contém as funções simples e é fechada por meio de limites pontuais, é suficiente mostrar que cada função $f \in \mathcal{F}_m$ pode ser expressa como limite pontual de uma seq. de funções simples. Uma seq. é dada pela seguinte construção.

Para cada $n \geq 1$ e cada $k = -n^2, \dots, n^2$, seja

$$E_{n,k} = \left\{ \omega \in \Omega : \frac{k-1}{n} < f(\omega) \leq \frac{k}{n} \right\}$$

e definir

$$\varphi_n(\omega) = \sum_{k=-n^2}^{n^2} \left(\frac{k}{n}\right) \mathbb{1}_{E_{n,k}}(\omega)$$

Uma vez que $f \in \mathcal{F}_m$, cada conj. $E_{m,k} \in \mathcal{F}$ e assim φ_n é uma f. simples. Além disso, uma vez que

$$|f(\omega) - \varphi_n(\omega)| < 1/n$$

$\forall n \geq |f(\omega)|$, φ_n converge pontualmente para f.

1.3.8.

($\Rightarrow$) Y é mensurável com respeito a $\sigma(X)$ se
$\forall B \subset \mathbb{R}$, $Y^{-1}(B) = \{\omega : Y(\omega) \in B\} \in \sigma(X)$

Isto é, se $\sigma(Y) \subset \sigma(X)$.

Queremos encontrar a função mensurável f tal que $Y = f(X)$. Para n fixado, considere o conjunto

$$A_{m,n} = \{\omega : m 2^{-n} \leq Y(\omega) \leq (m+1) 2^{-n}\} \text{ para } m = 0, \pm 1, \dots$$

Uma vez que esse conjunto está em $\sigma(Y)$, ele tbm está em $\sigma(X)$, e, portanto, pode ser escrito como $\{\omega : X(\omega) \in B_{m,n}\}$ para algum subconjunto boreliano $B_{m,n}$ da reta real.

Considere a função $f_n(x) = \sum_m m 2^{-n} \mathbb{1}(x \in B_{m,n})$. Claramente essa função está definida de maneira que $f_n(X)$ está perto de Y, e de fato está a $1/2^n$ unidade de Y. A função que buscamos é obtida tomando o limite

$$f(x) = \lim_{n \to \infty} f_n(x)$$

Exigimos dois resultados: primeiro que o limite existe, e segundo que o limite satisfaz a propriedade $f(X) = Y$.

A conv. da seq. segue do fato de que, para todo x, a sequência $f_n(x)$ é monotonicamente crescente.
O fato de $Y = f(X)$ segue facilmente uma vez que, para todo n, $f_n(X) \leq Y \leq f_n(X) + \frac{1}{2^n}$. Tomando limites qdo $n \to \infty$, ficamos $f(X) \leq Y \leq f(X)$.

$\Leftarrow)$

Suponha $Y = f(X)$. Então, para um boreliano B arbitrário, $[Y \in B] = [f(X) \in B] = [X \in f^{-1}(B)] = [X \in B_2]$ p/ todo boreliano $B_2 \in \mathcal{B}$. Isso mostra que $\sigma(Y) \subset \sigma(X)$, isto é, Y é mensurável com respeito a $\sigma(X)$

1.4.1 Mostre que se $f \geq 0$ e $\int f \, d\mu = 0$, então $f = 0$ a.e.

Suponha que $f \geq 0$ é mensurável e seja $E_0 = \{x : f(x) \geq 1\}$ e

$E_n = \{x : f(x) \in (1/n-1, 1/n)]\}$. Então $E = \{x : f(x) > 0\}$ é igual a união disjunta $\bigcup_{n \geq 0} E_n$ e

$$\mu(E) = \sum_{n \geq 0} \mu(E_n)$$

Se $\mu(E) > 0$, então temos que ter $\mu(E_n) > 0$ para algum n, caso em que $f(x) \geq \phi_n(x) \equiv \frac{1}{n+1} \mathbb{1}_{E_n(x)}$. Observando

que ϕ_n é uma f. simples, segue-se que

$$\int f \, d\mu \geq \int \phi_n \, d\mu = \frac{1}{n+1} \mu(E_n)$$

Isto mostra que se $\int f \, d\mu = 0$, então $\mu(E) = 0$

Outro sol.

Quero mostrar que $f \geq 0$ e $\int f \, d\mu = 0$ então $f = 0$ q.c.

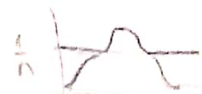
Para que $f = 0$ q.c. preciso mostrar que $\mu(x : f(x) > 0$ ou $f(x) < 0) = 0$

Como μ é medida, temos $\mu(x : f(x) > 0$ ou $f(x) < 0) =$

$$\overset{\text{disj.}}{=} \mu(x : f(x) > 0) + \mu(x : f(x) < 0)^{0}$$

$$\overset{f \geq 0 \text{ q.c.}}{=} \mu(x : f(x) > 0)$$

$\therefore$ basta provar que $\mu(x : f(x) > 0) = 0$



Para isso, defino $A_n = \{x : f(x) \geq \frac{1}{n}\} = \{f \geq \frac{1}{n}\}$

$0 \leq f \mathbb{1}_{A_n} \leq f \Rightarrow 0 \leq \int f \mathbb{1}_{A_n} d\mu \leq \int f \, d\mu = 0 \Rightarrow \int f \mathbb{1}_{A_n} d\mu = 0$

Como $f \mathbb{1}_{A_n} \geq \frac{1}{n} \mathbb{1}_{A_n}$ temos que

$$0 = \int f \mathbb{1}_{A_n} d\mu \geq \int \frac{1}{n} \mathbb{1}_{A_n} d\mu \underset{\text{f. simples}}{=} \frac{1}{n} \mu(A_n) \underset{\text{por ser medida}}{\geq} 0$$

Logo, $\frac{1}{n} \mu(A_n) = 0 \Rightarrow \mu(A_n) = 0$

Note que $A_n \subset A_{n+1}$, $\forall n \geq 1$ e $A_n \uparrow A$, $A = \bigcup_{n=1}^{\infty} A_n =$
$= \bigcup_{i=1}^{\infty} \{f \geq \frac{1}{n}\} = \{f > 0\}$.

Pela prop. da cont por seq. cresc. segue que

$$\mu(A_n) \to \mu(A)$$

Como $\mu(A_n) = 0$, $\forall n \geq 1$, segue que $\mu(A) = \mu(x : f > 0) = 0$

Portanto $f = 0$ q.c.

1.5.1. Tome

$$\|f\|_\infty = \inf(M : \mu(x : |f| > M) = 0)$$

Prove que

$$\int |fg| \, d\mu \leq \|f\|_1 \|g\|_\infty$$

Obs: $\|g\|_\infty < \infty \Leftrightarrow g$ é limitada q.c.

Tome g tal que $\|g\|_\infty = \infty$

Temos duas opções: $|fg|$ é integrável $\Rightarrow \int |fg| \, d\mu < \infty < \|f\|_1 \|g\|_\infty$

$|fg|$ não é integrável $\Rightarrow \int |fg| \, d\mu = \infty = \|f\|_1 \|g\|_\infty$

Logo, quando $\|g\|_\infty = \infty$ vale o resultado. Agora, quero provar para $\|g\| < \infty$. Se g é limitada, então existe M tal que $|g| \leq M$, ou seja, $\mu(x : |g(x)| > M) = 0$.

$$\int |fg| \, d\mu \leq \int |fM| \, d\mu = \int |f| M \, d\mu = M \int |f| \, d\mu = M \|f\|_1$$

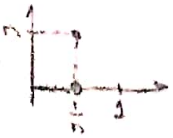
Como vale que $\int |fg| \, d\mu \leq M \|f\|_1$, para todo M tal que $|g| \leq M$ q.c. Então a desigualdade é verdadeira para o inf, ou seja

$$\int |fg| \, d\mu \leq \underbrace{\inf\{M : \mu(x : |g| > M) = 0\}}_{|g| \leq M \text{ q.c.}} \|f\|_1 = \|f\|_1 \|g\|_\infty$$

∎

Obs. Considere a função $f(x) = n \mathbb{1}_{\{x = 1/n\}}$ $(\mathbb{R}, \mathcal{B}, \mu)$

μ med. de Lebesgue



Temos que

$$\|f_n\|_\infty = \inf\{M : \mu(x : |f| > M) = 0\}$$

Obs. que $\mu(x : |f_n| > 0) = \mu(\{\tfrac{1}{n}\}) = 0$, pois μ é a med de Lebesgue

Então, $\|f_n\|_\infty = 0$, mas f_n não é limitada

1.5.2. Mostre que

$$\|f\|_\infty = \lim_{p\to\infty} \|f\|_p$$

Sabemos que

$$\|f\|_p = \left(\int |f|^p \, d\mu\right)^{1/p} \quad , \ 1 \leq p < \infty$$

e

$$\|f\|_\infty = \inf \{M > 0 : \mu\{x : |f| > M\} = 0\}$$

Fixe $\delta > 0$, e $S_\delta := \{x : |f(x)| \geq \|f\|_\infty - \delta\}$ para $\delta < \|f\|_\infty$

Temos

$$\|f\|_p \geq \left(\int_{S_\delta} (\underbrace{\|f\|_\infty - \delta}_{\text{fixo}})^p \, d\mu\right)^{1/p} = \{[\|f\|_\infty - \delta]^p \mu(S_\delta)\}^{1/p} =$$

$$= (\|f\|_\infty - \delta)\, \mu(S_\delta)^{1/p} \quad , \ \forall\, p \geq 1,$$

uma vez que $\mu(S_\delta)$ é finito e positivo. Isso nos dá

$$\liminf_{p\to\infty} \|f\|_p \geq \|f\|_\infty$$ (pois $\frac{1}{p} \xrightarrow[p\uparrow\infty]{} 0$, e δ é arbitrariamente pequeno)

Como $|f(x)| \leq \|f\|_\infty$ para quase todo x, temos, para $p > q$

$$\|f\|_p \leq \left(\int \|f\|_\infty^{p-q} |f(x)|^q \, d\mu\right)^{1/p} = \|f\|_\infty^{\overset{1}{\overbrace{\frac{p-q}{p}}}} \|f\|_q^{\overset{0}{\overbrace{q/p}}} \xrightarrow[p\uparrow\infty]{} \|f\|_\infty$$

resultando a desigualdade oposta. $[\limsup_{p\to\infty} \|f\|_p \leq \|f\|_\infty]$

Portanto, como

$$\liminf_{p\to\infty} \|f\|_p \geq \|f\|_\infty \geq \limsup_{p\to\infty} \|f\|_p$$

e

$$\liminf_{p\to\infty} \|f\|_p \leq \limsup_{p\to\infty} \|f\|_p,$$

chegamos à igualdade desejada

$$\lim_{p\to\infty} \|f\|_p = \|f\|_\infty .$$

1.5.2 Desigualdade de Minkowski

Temos que

$$\|f\|_\infty = \inf\{M > 0 : \mu\{x : |f| > M\} = 0\}.$$

$$\|f+g\|_\infty = \inf\{M > 0 : \mu\{x : |f+g| > M\} = 0\} \leq$$

$$\inf\{M > 0 : \mu\{x : |f| + |g| > M\} = 0\} \leq$$

$$\inf\{M > 0 : \mu\{x : |f| > M\} = 0\} + \inf\{M > 0 : \mu\{x : |g| > M\} = 0\}$$

$$= \|f\|_\infty + \|g\|_\infty$$

Logo,

$$\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$$

1.5.6 Se $g_n \uparrow g$ e $\int g_1^- \, d\mu < \infty$, então $\int g_n \, d\mu \uparrow \int g \, d\mu$

Vamos fazer uso extensivo do Teorema da Convergência Monótona, o qual afirma que se $(f_n)_{n \geq 1}$ é uma sequência de funções não-negativas tal que $f_n \uparrow f$, então

$$\int f_n \, d\mu \uparrow \int f \, d\mu$$

• Relembremos que $\int g_n \, d\mu = \int g_n^+ \, d\mu - \int g_n^- \, d\mu$

Notemos que a sequência $(g_n^+)_{n \geq 1}$ satisfaz as condições do Teorema da Convergência Monótona para $f = g^+$, de modo que

$$\int g_n^+ \, d\mu \uparrow \int g^+ \, d\mu$$

Assim, resta-nos mostrar que

$$-\int g_n^- \, d\mu \uparrow -\int g^- \, d\mu$$

Para isso, consideremos a sequência $(h_n^-)_{n \geq 1}$ e a função h^- definidos como segue

$$h_n^- := g_1^- - g_n^-$$
$$h^- := g_1^- - g^-$$

Observemos que $(h_n^-)_{n \geq 1}$ satisfaz as condições do TCM para $f = h^-$, de maneira que

$$\underbrace{\int (g_1^- - g_n^-) \, d\mu}_{\int g_1^- d\mu - \int g_n^- d\mu} \uparrow \underbrace{\int (g_1^- - g^-) \, d\mu}_{\int g_1^- d\mu - \int g^- d\mu}$$

$\Rightarrow -\int g_n^- \, d\mu \uparrow -\int g^- \, d\mu$

Portanto, como $\int g_n^+ \, d\mu \uparrow \int g^+ \, d\mu$ e $-\int g_n^- \, d\mu \uparrow -\int g^- \, d\mu$,

além de que $\int g_n \, d\mu = \int g_n^+ \, d\mu - \int g_n^- \, d\mu$ e

$\int g \, d\mu = \int g^+ \, d\mu - \int g^- \, d\mu$, concluímos que

$$\int g_n \, d\mu \uparrow \int g \, d\mu$$

Probabilidade Avançada, Aula 02, Tarefas

1º mostramos que é álgebra

- E_1, E_2 em $\mathcal{M} \Rightarrow E_1 \cup E_2 \in \mathcal{M}$

Seja $G \subset \Omega$, então

$$\mu^*(G) = \mu^*(G \cap E_1) + \mu^*(G \cap E_1^c) \quad \text{, pois } E \in \mathcal{M}$$

Como $G \cap E_1^c \subset G$, segue que $G \cap E_1^c \in \Omega$. Daí, uma vez que $E_2 \in \mathcal{M}$,

$$\mu^*(G \cap E_1^c) = \mu^*(G \cap E_1^c \cap E_2) + \mu^*(G \cap E_1^c \cap E_2^c)$$

Daí,

$$\mu^*(G) = \mu^*(G \cap E_1) + \mu^*(G \cap E_1^c \cap E_2) + \mu^*(G \cap E_1^c \cap E_2^c) \quad (1)$$

Note agora que

$$G \cap [E_1 \cup (E_2 \cap E_1^c)] = G \cap (E_1 \cup E_2)$$

Como μ^* é sub-σ-aditiva,

$$\mu^*(G \cap (E_1 \cup E_2)) \leq \mu^*(G \cap E_1) + \mu^*(G \cap E_2 \cap E_1^c) \Rightarrow$$

$$\mu^*(G \cap E_1) \geq \mu^*(G \cap (E_1 \cup E_2)) - \mu^*(G \cap E_2 \cap E_1^c) \quad (2)$$

Substituindo (2) em (1), segue que

$$\mu^*(G) \geq \mu^*(G \cap (E_1 \cup E_2)) - \mu^*(G \cap E_2 \cap E_1^c) + \mu^*(G \cap E_1^c \cap E_2) + \mu^*(G \cap (E_1 \cup E_2)^c) \Rightarrow$$

$$\mu^*(G) \geq \mu^*(G \cap (E_1 \cup E_2)) + \mu^*(G \cap (E_1 \cup E_2)^c) \quad \text{, o que}$$

prova que $E_1 \cup E_2 \in \mathcal{M}$, pois a desigualdade oposta é decorrente da sub-aditividade da medida exterior.

f

Agora, quero mostrar que uniões enumeráveis de elementos de $\mathcal{M}$ estão em $\mathcal{M}$.

a) Se $\{E_i\}_{i\geq 1}$ em $\mathcal{M}$, então $\bigcup_i E_i \in \mathcal{M}$

Tomando $E_1' = E_1$,

$E_2' = E_2 \cap E_1^c$

$\vdots$

$E_n' = E_n \cap E_1^c \cap \dots \cap E_{n-1}^c = E_n \cap \left(\bigcap_{i<n} E_i^c \right)$

temos $\{E_i'\}_{i\geq 1}$ disjuntos

CAPÍTULO 2

Exercício (4ª edição)	*Solutions M.* (2ª edição)	Página no *Solutions M.*	Exercício (4ª edição)	*Solutions M.* (2ª edição)	Página no *Solutions M.*	Exercício (4ª edição)	*Solutions M.* (2ª edição)	Página no *Solutions M.*
2.1.1	4.1	7	**2.2.7**	5.7	13	**2.4.2**	7.2	19
2.1.2	4.2	7	**2.2.8**	5.8	13	**2.4.3**	7.3	19
2.1.3	4.3	8	**2.2.9**	5.9	14	**2.4.4**	7.4	20
2.1.4	4.4	8	**2.3.1**	6.6	16	**2.5.1**	8.1	20
2.1.5	4.5	8	**2.3.2**	6.7	16	**2.5.2**	8.2	20
2.1.6	4.10	10	**2.3.3**	6.8	16	**2.5.3**	8.3	21
2.1.7	4.6	8	**2.3.4**	6.9	16	**2.5.4**	8.4	21
2.1.8	4.7	9	**2.3.5**	6.1	15	**2.5.5**	8.5	21
2.1.9	4.11	10	**2.3.6**	6.2	15	**2.5.6**	8.6	21
2.1.10	4.12	10	**2.3.7**	6.3	15	**2.5.7**	8.7	21
2.1.11	4.13	11	**2.3.8**	6.4	15	**2.5.8**	8.8	22
2.1.12	4.14	11	**2.3.9**	6.5	16	**2.5.9**	8.9	22
2.1.13	4.15	11	**2.3.10**	6.10	17	**2.5.10**	8.10	22
2.1.14	4.16	11	**2.3.11**	6.11	17	**2.5.11**	8.11	22
2.1.15	4.17	11	**2.3.12**	6.12	17	**2.5.12**	8.12	23
2.1.16	4.18	12	**2.3.13**	6.13	17	**2.6.1**	9.1	24
2.1.17	4.19	12	**2.3.14**	6.14	17	**2.6.2**	9.2	24
2.1.18	4.20	12	**2.3.15**	6.15	17	**2.6.3**	9.3	24
2.2.1	5.1	12	**2.3.16**	6.16	18	**2.6.4**	9.4	24
2.2.2	5.2	12	**2.3.17**	6.17	18	**2.6.5**	9.5	24
2.2.3	5.3	13	**2.3.18**	6.18	18	**2.6.6**	9.6	24
2.2.4	5.4	13	**2.3.19**	6.19	18	**2.6.7**	9.7	25
2.2.5	5.5	13	**2.3.20**	6.20	19	**2.6.8**	9.8	25
2.2.6	5.6	13	**2.4.1**	7.1	19	**2.6.9**	9.9	25

1.6.1 φ estrit. convexa. $\varphi(EX) = E\varphi(X) \Rightarrow X = EX$ q.c.

.8 Método do estatístico inconsciente. (Ver da Andressa, tá legal!)

.11 $E|X|^k < \infty \Rightarrow E|X|^j < \infty$, $0 < j \leq k$ e $E|X|^j \leq (E|X|^k)^{j/k}$

.12 Usar des. Jensen p/ mostrar que $\sum_{m=1}^{n} p_m y_m \geq \prod_{m=1}^{n} y_m^{p_m}$ com $p_m = P(X = \log y_m)$

1.7.2 Aplicação do Teorema de Fubini

2.1.1 X, Y ind $\Rightarrow \sigma(X)$ e $\sigma(Y)$ o são. $\mathcal{F}$ e $\mathcal{G}$ ind., $X \in \mathcal{F}$ e $Y \in \mathcal{G}$ tam serão.

.2 A, B ind $\Rightarrow A$ e B^c são, A^c e B são, A^c e B^c são. $\mathbb{1}_A$ e $\mathbb{1}_B$ ind $\Leftrightarrow A$ e B são ind.

.3 $A_1, \ldots, A_n$ ind. $\Rightarrow A_1^c, A_2, \ldots, A_n$ são ind.. Ademais $\mathbb{1}_{A_1}, \mathbb{1}_{A_2}, \ldots, \mathbb{1}_{A_n}$ indep.

.4 $f(x) = g_1(x_1) \cdots g_n(x_n)$, g_m mens. $\geq 0 \Rightarrow X_1, \ldots, X_n$ indep.

.7 $X_n(\omega) = \sin(2\pi n\omega)$, $n = 1, 2, \ldots$ X_n's não correlacionados, mas não independentes.

.10 $X, Y \sim \mathcal{U}\{0, \ldots, k-1\}$, $k \geq 3$ primo. $Z_n = X + nY \bmod k$, $0 \leq n < k$. $Z_0, \ldots, Z_n$ 2-2 ind. $Z_n \sim \mathcal{U}\{0, \ldots, k-1\}$

.13 X, Y ind., $\operatorname{Im} X$, $\operatorname{Im} Y = \mathbb{Z} \Rightarrow P(X+Y=n) = \sum_m P(X=m)P(Y=n-m)$

.14 $X \sim \mathrm{Poi}(\lambda)$, $Y \sim \mathrm{Poi}(\mu)$ ind., $X+Y \sim \mathrm{Poi}(\lambda+\mu)$

.15 $X \sim \mathrm{Bin}(n, p)$, $Y \sim \mathrm{Bin}(m, p)$ ind $\Rightarrow X+Y \sim \mathrm{Bin}(n+m, p)$. $S_n = \sum_{i=1}^{n} X_i \sim \mathrm{Bin}(n, p)$

onde X_i's i.i.d. com dist. $\mathrm{Ber}(p)$.

1.6.1. Suponha que φ é estritamente convexa*, i.e., $>$ é válido para todo $\lambda \in (0,1)$. Mostre que sob as hipóteses do Teorema 1.6.2, $\varphi(EX) = E\varphi(X)$ implica $X = EX$ q.c.

Hipóteses do Teo. 1.6.2

→ φ é convexa (no caso, estritamente convexa)

→ ambas as esperanças $E(\varphi(X))$ e $E(X)$ existem, i.e,

$E|X| < \infty$ e $E|\varphi(X)| < \infty$

Seja $l(x) \leq \varphi(x)$ uma f. linear com $l(EX) = \varphi(EX)$, e note

$\varphi(X) \geq l(X)$ $E\,\varphi(X) \geq E\,l(X) = l(EX) = \varphi(EX)$. Se $\varphi(EX) = E(\varphi X)$, então a última desigualdade ($\geq$) fica substituída por uma igualdade, de modo que $E\varphi(X) = E\,l(X)$. Como $\varphi(X) \geq l(X)$, então $E|\varphi(X) - l(X)| =$

$= E(\varphi(X) - l(X)) = E\varphi(X) - E\,l(X) = 0$. Assim, para todo $\varepsilon > 0$, pela desigualdade de Markov,

$$P(|\varphi(X) - l(X)| \geq \varepsilon) \leq \frac{E|\varphi(X) - l(X)|}{\varepsilon} = 0$$

de modo que $\varphi(X) = l(X)$ q.c.

Por outro lado, quando φ é estritamente convexa, temos que $\varphi(x) > l(x)$ para $x \neq EX$

Como φ é estritamente convexa, é da forma

$$\lambda\varphi(x) + (1-\lambda)\,\varphi(y) > \varphi(\lambda x + (1-\lambda)y) \quad \forall\, \lambda \in (0,1),\ x, y \in \mathbb{R}$$

Consideremos, então, um pto qqer $y \neq EX$.

Seja $z = \lambda EX + (1-\lambda)y$ p/ algum $\lambda \in (0,1)$.

Pela convexidade estrita de φ, temos

$$\lambda \varphi(EX) + (1-\lambda)\varphi(y) > \varphi(\lambda EX + (1-\lambda)y)$$

$$\geq \ell(\lambda EX + (1-\lambda)y) = \lambda \ell(EX) + (1-\lambda)\ell(y)$$

$$= \lambda \varphi(EX) + (1-\lambda)\ell(y) \quad \Leftrightarrow$$

$$(1-\lambda)\varphi(y) > (1-\lambda)\ell(y) \Leftrightarrow$$

$$\varphi(y) > \ell(y)$$

Logo, precisamos ter $X = EX$ q.c.

1.6.8. Suponha que a medida de prob. μ tem $\mu(A) = \int_A f(x)\,dx$ para todo $A \in \mathcal{R}$. Use a prova técnica do Teorema 1.6.9 para mostrar que p/ qualquer $g \geq 0$ ou $\int |g(x)|\mu(dx) < \infty$, temos

$$\int g(x)\,\mu(dx) = \int g(x) f(x)\,dx$$

Suponha-se que μ é uma medida de probabilidade em $(\mathbb{R}, \mathcal{R})$ e que $f: \mathbb{R} \to [0,\infty]$ é uma função Borel-mensurável tal que a identidade

$$\mu(A) = \int_A f(x)\,dx$$

é válida para todos os $A \in \mathcal{A}$. Queremos provar a identidade

$$(*) \quad \int g(x)\,\mu(dx) = \int g(x) f(x)\,dx$$

Vale para todas as funções g com $g \geq 0$ ou $\int |g(x)|\mu(dx) < \infty$. Isto pode ser feito usando o método de prova da mudança das variáveis do teorema (1.6.9). (Este método é chamado a "máquina padrão" por David Williams)

Caso 1: Se $g = 1_A$ é a f. indicadora de um conjunto mensurável, então $(*)$ tem por pressuposto

$$\int_{\mathbb{R}} 1_A(x)\,\mu(dx) = \mu(A) = \int_A f(x)\,dx = \int 1_A(x) f(x)\,dx$$

Caso 2: Se $g = \sum_{i=1}^{n} c_i 1_{A_i}$ é uma f. simples, pelo caso 1 e pela linearidade da integral,

$$\int g(x)\,\mu(dx) = \sum_{i=1}^{n} c_i \int 1_{A_i}(x)\,\mu(dx) = \sum_{i=1}^{n} c_i \int 1_{A_i}(x) f(x)\,dx = \int g(x) f(x)\,dx$$

Caso 3: Se $g \geq 0$ e definimos

$$g_n(x) = (\lfloor 2^n g(x) \rfloor / 2^n) \wedge n$$, em seguida

$(g_n, n \geq 1)$ é uma seq. de f. simples que convergem monotonicamente até g. Como $f \geq 0$, segue-se que $g_n f \uparrow gf$ e assim o Caso 2 combinado com o Teorema da conv. Monotona implica que

$$\int g(x) \mu(dx) = \lim_{n \to \infty} \int g_n(x) \mu(dx) = \lim_{n \to \infty} \int g_n(x) f(x) dx =$$

$$= \int g(x) f(x) dx$$

Caso 4. Suponha-se que g é integrável:

$\int |g(x)| \mu(dx) < \infty$. Então $\int g^+(x) \mu(dx) < \infty$ e $\int g^-(x) \mu(dx) < \infty$

assim o caso 3 implica $\int g(x) \mu(dx) = \int g^+(x) \mu(dx) - \int g^-(x) \mu(dx) =$
$= \int g^+(x) f(x) dx - \int g^-(x) f(x) dx = \int g(x) f(x) dx$

1.6.11. Se $E|X|^k < \infty$ então para $0 < j < k$ $E|X|^j < \infty$ e além disso

$$E|X|^j \leq (E|X|^k)^{j/k}$$

Para $0 < j < k$, temos que

$$E|X|^j = \int |X|^j dP = \int_{\{|X| \leq 1\}} |X|^j dP + \int_{\{|X| > 1\}} |X|^j dP$$

$$\left.\begin{array}{l} \displaystyle\int_{\{|X|>1\}} |X|^j dP < \int_{\{|X|>1\}} |X|^k dP < \infty \\ \displaystyle\int_{\{|X|\leq 1\}} |X|^j dP \underset{|X|^j \leq 1}{<} \int_{\{|X|\leq 1\}} 1\, dP = 1 . \underbrace{P(|X| \leq 1)}_{\in [0,1]} < 1 \end{array}\right\} + \quad E|X|^j < \infty$$

Note que $E|X|^j \leq (E|X|^k)^{j/k} \Leftrightarrow (E|X|^j)^{k/j} \leq E|X|^k$ (1)

Obs. Uma função f é convexa se, e só se, $f'' > 0$

Seja $f(x) = |x|^{k/j}$ $\quad f'(x) = \frac{k}{j} |x|^{\frac{k}{j}-1}$

(Podemos usar Jensen, pois f e $|X|^j$ são integráveis, ou seja, $E(\varphi) < \infty$ e $E(|X|^j) < \infty$)

$$\varphi(E|X|^j) \leq E[\varphi(|X|^j)] \Rightarrow (\underbrace{|E|X|^j|}_{>0})^{k/j} \leq E\left[|\underbrace{|X|^j}_{>0}|^{k/j}\right] \Rightarrow$$

$$(E|X|^j)^{k/j} \leq E|X|^k \quad (2)$$

Por (1) e (2), segue o resultado.

1.6.12 Aplique a desigualdade de Jensen para $\varphi(x) = e^x$ e $P(X = \log y_m) = p(m)$ para concluir que se $\sum_{m=1}^{n} p_m = 1$ e $p_m > 0$, $y_m > 0$ então

$$\sum_{m=1}^{n} p(m)\, y_m \geq \prod_{m=1}^{n} y_m^{p(m)}$$

Qdo $p(m) = 1/n$ temos que a média aritmética é maior que a média geométrica.

Sendo $\sum_{m=1}^{n} p(m) = 1$ temos $EX = \sum_{m=1}^{n} \log y_m\, p(m)$

Como $\varphi(x) = e^x$ é convexa ($\varphi'' > 0$) segue pela des. de Jensen,

$$E(\varphi(X)) \geq \varphi(EX) \qquad (1)$$

$$\underbrace{E(\varphi(X))}_{\int \varphi(x)\,dP} = \sum_{m=1}^{n} \exp(\log y_m)\, p(m) = \sum_{m=1}^{n} y_m\, p(m)$$

$$\varphi(EX) = e^{EX} = e^{\sum_{m=1}^{n} \log y_m\, p(m)} = \prod_{m=1}^{n} e^{\log y_m\, p(m)} = \prod_{m=1}^{n} e^{\log y_m^{p_m}} = \prod_{m=1}^{n} y_m^{p(m)}$$

Por (1) segue que

$$\sum_{m=1}^{n} y_m\, p(m) \geq \prod_{m=1}^{n} y_m^{p(m)}$$

1.7.2. Seja $g \geq 0$ uma função mensurável em $(X, \mathcal{A}, \mu)$. Use o Teorema 1.7.2 para concluir que

$$\int_X g\, d\mu = (\mu \times \lambda)(\{x, y\} : 0 \leq y \leq g(x)\}) = \int_0^\infty \mu(\{x : g(x) \geq y\})\, dy$$

Teorema 1.7.2. Teo. de Fubini. Se $f \geq 0$ ou $\int |f|\, d\mu < \infty$, então

$$\int_X \int_Y f(x,y)\, \mu_2(dy)\, \mu_1(dx) = \int_{X \times Y} f\, d\mu = \int_Y \int_X f(x,y)\, \mu_1(dx)\, \mu_2(dy)$$

Basta mostrar que cada um dos termos da igualdades do enunciado corresponde a um dos termos das igualdades do Teorema. Dessa forma, ao tomarmos λ como a medida de Lebesgue, $Y := [0, \infty)$ e $\mathcal{B}$ como os borelianos, e fazendo

$$f(x,y) = \mathbb{1}_{\{(x,y) : 0 \leq y \leq g(x)\}}$$

observamos

$$\int_X \left[\int_Y f(x,y)\, dy\right] \mu(dx) = \int_X g(x)\, \mu(dx)$$

$$\int f\, d(\mu \times \lambda) = (\mu \times \lambda)(\{x, y\} : 0 \leq y \leq g(x)\})$$

$$\int_Y \left(\int_X f(x,y)\, \mu(dx)\right) \mu(dy) = \int_0^\infty \mu(g(x) \geq y)\, dy$$

$\int_Y$

2.1.1

i) Mostre que se X e Y forem independentes, então $\sigma(X)$ e $\sigma(Y)$ tbm serão.

ii) No sentido oposto, se $\mathcal{F}$ e $\mathcal{G}$ forem independentes, $X \in \mathcal{F}$ e $Y \in \mathcal{G}$, então X e Y são ind.

A primeira def. é, por sua vez, um caso especial da segunda.

i) Se $A \in \sigma(X)$, então segue da definição de $\sigma(X)$ que $A = \{X \in C\}$ para algum $C \in \mathcal{R}$. Da mesma forma, se $B \in \sigma(Y)$, então $B = \{Y \in D\}$ para algum $D \in \mathcal{R}$. Logo, usando esses fatos e a ind. de X e Y,

$$P(A \cap B) = P(X \in C, Y \in D) = P(X \in C) P(Y \in D) = P(A) P(B)$$

ii) Da def. de mensurabilidade, segue que $\{X \in C\} \in \mathcal{F}$, e $\{Y \in D\} \in \mathcal{G}$, para todos $C, D \in \mathcal{R}$. Uma vez que $\mathcal{F}$ e $\mathcal{G}$ são indep., segue que

$$P(\underbrace{X \in C}_{\substack{\text{evento} \\ \text{em} \\ \mathcal{F}}}, \underbrace{Y \in D}_{\substack{\text{evento} \\ \text{em} \\ \mathcal{G}}}) = P(X \in C) P(Y \in D) \quad \text{p/ todo } C, D \in \mathcal{R}$$

de modo que X e Y são independentes

2.1.2.

(i) Mostre que, se A e B forem independentes, então tbm A^c e B, A e B^c, e A^c e B^c serão.

(ii) Conclua que os eventos A e B são independentes se e somente se suas v.a.'s indicadoras $\mathbb{1}_A$ e $\mathbb{1}_B$ forem independentes.

i)

- $P(A^c\cap B) = P(B) - P(A\cap B) \overset{A \text{ e } B \text{ ind.}}{=} P(B) - P(A)P(B) = P(B)(1-P(A)) = P(A^c)P(B)$

∴ A^c e B são independentes

- Analogamente $P(B^c\cap A) = P(A) - P(A\cap B) = P(A) - P(A)P(B) = P(B^c)P(A)$.

A e B^c são ind.

- $P(A^c\cap B^c) = P(A^c) - P(A^c\cap B) \overset{A^c \text{ e } B \text{ são ind.}}{=} P(A^c) - P(A^c)P(B) = P(A^c)P(B^c)$.

∴ A^c e B^c são ind.

ii) Duas v.a.'s X e Y são ind. se, para todo $C, D \subseteq \mathbb{R}$,

$$P(X\in C, Y\in D) = P(X\in C, Y\in D)$$

i.e., os eventos $\{X\in C\}$ e $\{Y\in D\}$ são independentes.

No nosso caso, se $C, D \in \mathcal{B}$, então $\{\mathbb{1}_A \in C\} \in \{\emptyset, A, A^c, \Omega\}$ e $\{\mathbb{1}_B \in D\} \in \{\emptyset, B, B^c, \Omega\}$, de modo que há 16 coisas para verificar. Qdo algum dos conjuntos env. é $\emptyset$ ou Ω, a igualdade é válida.

$P(\emptyset\cap \_) = P(\emptyset)P(\_) = 0$ $\qquad$ $P(\Omega\cap \_) = P(\Omega)P(\_) = P(\_)$

$P(\_\cap\emptyset) = P(\_)P(\emptyset) = 0$ $\qquad$ $P(\_\cap\Omega) = P(\_)P(\Omega) = P(\_)$

$P(\Omega\cap\Omega) = P(\Omega)P(\Omega) = 1$

Por (i), temos

$\{1_A \in C\}$ \ $\{1_B \in D\}$	$\emptyset$	B	B^c	Ω
$\emptyset$	✓	✓	✓	✓
A	✓			✓
A^c	✓			✓
Ω	✓	✓	✓	✓

Restam quatro coisas para analisar, e estão todos abrangidos pelo item i:

$$P(A \cap B) = P(A)P(A)$$

$$\vdots$$

$$P(A^c \cap B^c) = P(A^c)\,P(B^c) \qquad \square$$

2.1.3 Sejam $A_1, A_2, \dots, A_n$ independentes. Mostre (i) que $A_1^c, A_2, \dots, A_n$ são ind.; (ii) $\mathbb{1}_{A_1}, \mathbb{1}_{A_2}, \dots, \mathbb{1}_{A_n}$ são independentes.

i) Sejam $B_1 = A_1^c$ e $B_i = A_i$, $2 \le i \le n$. Os conj. $B_1, \dots, B_n$ serão ind. se para todo $I \subset \{1, \dots, n\}$,

$$\mathbb{P}\Big(\bigcap_{i \in I} B_i\Big) = \prod_{i \in I} P(B_i) \qquad (*)$$

Como $A_1, A_2, \dots, A_n$ são ind., basta verificarmos $(*)$ para $I \subseteq \{1, 2, \dots, n\}$. Seja $J = I \setminus \{1\}$

$$P\Big(A_1^c \cap \bigcap_{i \in J} A_i\Big) = P\Big(\bigcap_{i \in J} A_i\Big) - P\Big(A_1 \cap \bigcap_{i \in J} A_i\Big) =$$

$A_1, \dots, A_n$ são ind.

$$= \prod_{i \in J} P(A_i) - \mathbb{P}(A_1) \prod_{i \in J} P(A_i) =$$

$$= [1 - P(A_1)] \prod_{i \in J} P(A_i) = P(A_1^c) \prod_{i \in J} P(A_i)$$

$\therefore$ $A_1^c, A_2, \dots, A_n$ são independentes.

(ii) $\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_n}$ serão independentes se, para todo $D_i \in \mathcal{R}$, $i = 1, \dots, n$, tivermos

$$P\Big(\bigcap_{i=1}^{n} \{\mathbb{1}_{A_i} \in D_i\}\Big) = \prod_{i=1}^{n} P(\mathbb{1}_{A_i} \in D_i) \qquad (*)$$

Mas para todo $D_i \in \mathcal{R}$, $i = 1, \dots, n$, $\underbrace{\{\mathbb{1}_{A_i} \in D_i\}}_{C_i} \in \{\emptyset, A_i, A_i^c, \Omega\}$. Se algum dos C_i for o conj. $\emptyset$, $(*)$ segue trivialmente. Caso contrário, usaremos o item (i) uma vez que, iterando o arg. daquele item, notaremos que, se $B_i \in \{A_i, A_i^c\}$, então $B_1, \dots, B_n$ são ind., de modo que (no item (ii)), se $C_i \in \{A_i, A_i^c, \Omega\}$, então $\mathbb{P}\Big(\bigcap_{i=1}^{n} C_i\Big) = \prod_{i=1}^{n} P(C_i)$

2.1.4 Suponha $(X_1, \dots, X_n)$ tem densidade $f(x_1, \dots, x_n)$, i.e.

$$P((X_1, \dots, X_n) \in A) = \int_A f(x)\, dx \quad \text{para } A \in \mathbb{R}^n$$

se $f(x)$ pode ser escrito como $g_1(x_1) \dots g_n(x_n)$ onde as $g_m \geq 0$ são mensuráveis, então $X_1, \dots, X_n$ são independentes. Note que as g_m não são assumidas densidades de probabilidade.

Temos que $f(x) = g_1(x_1) g_2(x_2) \dots g_n(x_n)$, com $g_m \geq 0$ mensuráveis, $1 \leq m \leq n$. Assim

$$P((X_1, \dots, X_n) \in A) = \int_A g_1(x_1) g_2(x_2) \dots g_n(x_n)\, dx_1 \dots dx_n$$

para cada conjunto de Borel $A \subset \mathbb{R}^n$. Vamos tomar $A = \mathbb{R}^n$. Precisamos fazer uso do Teorema de Fubini e como as funções $g_m(\cdot)$ são não-negativas, então

$$1 = P((X_1, \dots, X_n) \in \mathbb{R}^n) = \int_{\mathbb{R}^n} g_1(x) \dots g_n(x_n)\, dx_1 \dots dx_n =$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} g_1(x_1) g_2(x_2) \dots g_n(x_n)\, dx_1\, dx_2 \dots dx_n =$$

$$= \prod_{i=1}^{n} \int_{\mathbb{R}} g_i(x_i)\, dx_i = \prod_{i=1}^{n} C_i \quad , \text{ note que } C_i = \int_{\mathbb{R}} g_i(x_i)\, dx_i < \infty$$

pois $g_m(\cdot)$ é mensurável

Agora seja $B \subset \mathbb{R}$, logo podemos reescrever $A = \mathbb{R}^{i-1} B\, \mathbb{R}^{n-i}$, logo

Aplicando nov. o teorema de Fubini, temos

$$P\{(X_1,\dots,X_n)\in\mathbb{R}^n\} = \int_{\mathbb{R}^{i-1}\times B\times\mathbb{R}^{n-i}} g_1(x_1)\dots g_n(x_n)\,dx_1\dots dx_n$$

$$\overset{\text{Fubini}}{=} \int_{\mathbb{R}^{i-1}}\int_B\int_{\mathbb{R}^{n-i}} g_1(x_1)\dots g_n(x_n)\,dx_1\dots dx_n =$$

$$= \int_{\mathbb{R}^{i-1}}\int_{\mathbb{R}^{n-i}}\int_B g_1(x_1)\dots g_n(x_n)\,dx_1\dots dx_n =$$

$$= \int_{\mathbb{R}^{i-1}}\int_{\mathbb{R}^{n-i}} g_1(x_1)\dots g_{i-1}(x_{i-1})\,g_{i+1}(x_{i+1})\dots g(x_n)\,dx_1\dots dx_{i-1}\,dx_{i+1}\dots dx_n \int_B g_i(x_i)\,dx_i$$

$$= \left[\prod_{j\neq i}^{n}\int_{\mathbb{R}} g_j(x_j)\,dx_j\right]\int_B g_i(x_i)\,dx_i = \left[\prod_{j\neq i}^{n} C_j\right]\int_B g_i(x_i)\,dx_i$$

$$= \int_B \prod_{j\neq i}^{n} C_j\, g_i(x_i)\,dx_i = \int_B h_i(x_i)\,dx_i, \text{com } h_i(x_i) = \prod_{j\neq i}^{n} C_j\, g_i(x_i)$$

Note que $1 = \prod_{j=1}^{n} C_j \Leftrightarrow 1 = \prod_{j\neq i}^{n} C_j \times C_i \Leftrightarrow \prod_{j\neq i}^{n} C_j = C_i^{-1}$

Assim, $h_i(x_i) = C_i^{-1} g_i(x_i)$, mais

$\int_{\mathbb{R}} h_i(x_i)\,dx_i = 1$ pois $\int_{\mathbb{R}} h_i(x_i)\,dx_i = \int_{\mathbb{R}} \frac{g_i(x_i)\,dx_i}{C_i} = \frac{\int_{\mathbb{R}} g_i(x_i)\,dx_i}{\int_{\mathbb{R}} g_i(x_i)\,dx_i} = 1$

Como $h_i(\cdot)$ é mens. (pois $g_i(\cdot)$ é tbm mensurável então podemos dizer que h_i é a função de probabilidade de X_i então

2.1.4

$$P(X_1 \leq x_1, \dots, X_n \leq x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} g_1(y_1) \cdots g_n(y_n)\, dy_1 \cdots dy_n =$$

$$= \prod_{i=1}^{n} \int_{-\infty}^{x_i} g_i(y_i)\, dy_i = \underbrace{\left(\prod_{i=1}^{n} c_i\right)}_{=1} \prod_{i=1}^{n} \int_{-\infty}^{x_i} g_i(y_i)\, dy_i =$$

$$= \left(\prod_{i=1}^{n} c_i\right)^{-1} \prod_{i=1}^{n} \int_{-\infty}^{x_i} g_i(y_i)\, dy_i = \prod_{i=1}^{n} \int_{-\infty}^{x_i} c_i^{-1} g_i(y_i)\, dy_i =$$

$$= \prod_{i=1}^{n} \int_{-\infty}^{x_i} h_i(y_i)\, dy_i = \prod_{i=1}^{n} P(X_i \leq x_i)$$

Como o resultado obtido foi considerando $(X_1, \dots, X_n) \in \mathbb{R}^n$.

$\Omega = (0,1)$, $\mathcal{F} =$ σ-álg. de Borel, $P =$ med. de Lebesgue. X_n correl. mas não ind.

2.1.7. Seja $\Omega = (0,1)$, $\mathcal{F} =$ σ-álg. de Borel, $P =$ medida de Lebesgue.
$X_n(\omega) = \sin(2\pi n \omega)$, $n = 1, 2, \dots$, não correlacionadas, mas não independentes.

$$E X_n = \int_0^1 \sin(2\pi n x)\, dx = -(2\pi n)^{-1} \cos(2\pi n x)\Big|_0^1 = 0$$

integrando por parte duas vezes,

$$E X_m X_n = \int_0^1 \sin(2\pi n x)\, \mathrm{sen}(2\pi m x)\, dx =$$

$$= \frac{m}{n} \int_0^1 \cos(2\pi m x) \cos(2\pi n x)\, dx$$

$$= \frac{m^2}{n^2} \int_0^1 \mathrm{sen}(2\pi n x)\, \mathrm{sen}(2\pi m x)\, dx$$

$E X_m X_n$ $E(1-a) = 0 \Rightarrow E = 0,\ a \neq 1$

então se $m \neq n$, $E X_m X_n = 0$. Para ver que X_m e X_n não são independentes basta notar que $X_m(x) = 0$ quando $x = \frac{k}{2m}$,

$0 \leq k < 2m$ e neste conjunto $X_n(x)$ assume os valores $V_n = \{y_0, y_1, \dots, y_{2m-1}\}$.

Seja $[a,b] \subset [-1,1] - V_n$ com $a < b$. A continuidade da função seno implica que, $\epsilon > 0$ suficientemente pequeno, temos que

$$P(X_m \in [0,\epsilon],\ X_n \in [a,b]) = 0 < P(X_m \in [0,\epsilon])\, P(X_m \in [a,b])$$

pois quando $X_n \in [a,b]$, $X_n \notin V_n \Rightarrow X_m(x) \neq 0 \Rightarrow$ não ocorrência

Outra forma:

se $n \neq m$, então $E[X_n X_m] = \int_0^1 \sin(2\pi n \omega)\, \mathrm{sen}(2\pi m \omega)\, d\omega$ é igual

2.1.10. Seja $k \geq 3$ um número primo, e sejam X e Y variáveis aleatórias independentes uniformemente distribuídas em $\{0, 1, \dots, k-1\}$. Para $0 \leq n < k$, seja $Z_n = X + nY \bmod k$.

Mostre que $Z_0, Z_1, \dots, Z_{k-1}$ são ind. duas a duas, i.e., que cada par é ind.. Elas não são ind. porque, se conhecermos os valores de duas das variáveis, então saberemos os valores de todas.

<u>Obs</u>: a mod b = resto da divisão inteira de a por b

Notação $x + ny \bmod k = (x + ny) \bmod k$

O fato de k ser primo implica que, para todo $l > 0$,

$$\{lj \bmod k : 0 \leq j < k\} = \{0, 1, \dots, k-1\}$$

Por exemplo, para $k = 5$ e $l = 3$

$$3 \times 0 \bmod 5 = 0$$
$$3 \times 1 \bmod 5 = 3$$
$$3 \times 2 \bmod 5 = 1$$
$$3 \times 3 \bmod 5 = 4$$
$$3 \times 4 \bmod 5 = 2 \rightarrow \{0, 1, 2, 3, 4\}$$

Assim, mY tem dist. unif. no conjunto $\{0, 1, \dots, k\}$, da mesma forma que $X + nY^{\bmod k}$, i.e.,

$$P(X + nY = i) = 1/k, \text{ para } 0 \leq i < k$$

Se $m < n$ e $l = n - m$, o que acabamos de constatar implica que, se $0 \leq m < n < k$, então, para todo $0 \leq i, j < k$, há exatamente

um par $0 \leq x, y < k$ de tal modo que $x + my = i$ e $x + ny = j$.
Isso mostra que

$$P(X + mY = i, X + nY = j) = 1/k^2 =$$

$$= P(X + mY = i) P(X + nY = j)$$

de forma que as variáveis são ind. duas a duas

2.1.13 Mostre que se X e Y são independentes, variáveis aleatórias de valores inteiros, então

$$P(X+Y=n) = \sum_m P(X=m)P(Y=n-m)$$

X e Y v.a.'s valorizadas em $\mathbb{Z}$. Então

$$P(X=i, Y=j) = P(X=i)P(Y=j) \quad \forall\, i,j \in \mathbb{Z}$$

Observando que o evento $\{X+Y=n\}$ é igual a união disjunta dos eventos $\{X=m, Y=n-m\}$, em que m varia ao longo de $\mathbb{Z}$, temos que

$$P(X+Y=n) = \sum_{m\in\mathbb{Z}} P(X=m, Y=n-m) = \sum_{m\in\mathbb{Z}} P(X=m)P(Y=n-m)$$

2.1.14. No exemplo 1.6.4, introduzimos a distribuição de Poisson com parâmetro λ, que é dada por

$$P(Z=k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, \dots$$

Use o ex. anterior para mostrar que, se $X \sim \text{Poisson}(\lambda)$ e $Y \sim \text{Poisson}(\mu)$ forem ind., então $X+Y \sim \text{Poi}(\lambda+\mu)$.

Pelo ex. anterior, se X e Y são v.a. ind. e que assumem valores nos inteiros ($\mathbb{Z}$), então

$$P(X+Y=n) = \sum_m P(X=m)\,P(Y=n-m)$$

Sendo $X \sim \text{Poisson}(\lambda)$ e $Y \sim \text{Poisson}(\mu)$ independentes,

$$P(X+Y=n) = \sum_{m=0}^{n} P(X=m)\,P(Y=n-m) =$$

$$= \sum_{m=0}^{n} \frac{e^{-\lambda}\lambda^m}{m!} \frac{e^{-\mu}\mu^{n-m}}{(n-m)!} =$$

$$= \frac{e^{-(\lambda+\mu)}}{n!} \sum_{m=0}^{n} \binom{n}{m} \lambda^m \mu^{n-m} =$$

$$= \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^n}{n!}, \quad n = 0, 1, \dots$$

$\therefore X+Y \sim \text{Poisson}(\lambda+\mu)$

2.1.15. $X \sim Bin(n,p)$ se

$$P(X=m) = \binom{n}{m} p^m (1-p)^{n-m}$$

i) Mostre que se $X \sim Bin(n,p)$ e $Y \sim Bin(m,p)$ ind., então $X+Y \sim Bin(n+m,p)$

ii) Olhe o exemplo 1.6.3 e use indução para concluir que a soma de n v.a. de Bernoulli (p) ind. é uma binomial (n, p).

i)

$$P(X+Y=k) = \sum_{l=k-m \vee 0}^{n} P(X=l)\, P(Y=k-l) =$$

$$= \sum_{l=(k-m)\vee 0}^{n} \binom{n}{l} p^l (1-p)^{n-l} \binom{m}{k-l} p^{k-l} (1-p)^{m-(k-l)}$$

$$= \binom{m+n}{k} p^k (1-p)^{m+n-k} \underbrace{\sum_{l=(k-m)\vee 0}^{n} \frac{\binom{n}{l}\binom{m}{k-l}}{\binom{m+n}{k}}}_{=1}$$

(Hipergeométrica (m, n, k))

$\blacksquare$

ii) $X_1 \sim$ Bernoulli (p)

$$P(X_1=1) = \binom{1}{1} p^1 (1-p)^{1-1} = p$$

ou seja, $X_1 \sim Bin(1,p)$. Assim, $S_1 = X_1 \sim Bin(1,p)$.

Suponhamos, como hipótese de indução, que $S_n = \sum_{i=1}^{n} X_i \sim Bin(n,p)$ em que $X_1, \ldots X_n$ são v.a.'s i.i.d. com dist. de Ber (p). Suponhamos então $X_1, \ldots, X_{n+1}$ i.i.d. Ber (p) e most. que $S_n = \sum_{i=1}^{n+1} X_i \sim Bin(n+1,p)$

Pelo item (i), como $S_n \sim Bin(n,p)$, e $X_{n+1} \sim Bin(1,p)$ são ind., segue o resultado.

$\square$

Para cada $A \subset \mathbb{N}$, seja $d^-(A) = \liminf_{n \to \infty} \frac{1}{n} |A \cap \{1, n\}|$ e

$d^+(A) = \limsup_{n \to \infty} \frac{1}{n} |A \cap \{1, n\}|$.

Note que $d^-(A)$ e $d^+(A)$ sempre existem e que

A tem dens. natural $\Leftrightarrow d^-(A) = d^+(A) = d(A)$.

Agora, considere

$I_0 = \{1\}$

$I_1 = \{2, 3\}$

$I_2 = \{4, 5, 6, 7\}$

$I_3 = \{8, \dots, 15\}$

$I_4 = \{16, \dots, 31\}$ i.e. $I_n = \{2^n, 2^n+1, \dots, 2^{n+1}-1\}$

$O_n = I_n \cap O$ $\quad O = \{2n+1, n \in \mathbb{N}\}$

$E_n = I_n \cap E$ $\quad E = \{$even number$\}$

$\cdot\, C = \bigcup_n I_{2n}$

1.6.8

$g: (\mathbb{R}, \mathcal{R}) \to (\mathbb{R}, \mathcal{R})$.

$$(\Omega, \mathcal{F}, \mathbb{P}) \xrightarrow{g(X)} (\mathbb{R}, \mathcal{R})$$

$$\mu(A) = P(X \in A) = P(X^{-1}(A)) = \int_A f(x)\,dx$$

$(\mathbb{R}, \mathcal{R}, \mu)$, g

$\mu = P \circ X^{-1}$

Caso 1

$$g(y) = \mathbb{1}_A(y) = \begin{cases} 1, & \text{se } y \in A \\ 0, & \text{c.c} \end{cases}$$

$$E(g(X)) = \int_\Omega \mathbb{1}_A(x)\,d\mathbb{P} = P(X \in A) = \mu(A) = \int_A f(x)\,dx = \int \mathbb{1}_A(x) f(x)\,dx =$$

$$= \int g(x) f(x)\,dx$$

Caso 2

$$g(y) = \sum_{m=1}^{n} c_m \mathbb{1}_{A_m}(y)$$

$$E(g(x)) = \int_\Omega \sum_{m=1}^{n} c_m \mathbb{1}_{A_m}(y)\,d\mathbb{P} = \sum c_m \int_\Omega \mathbb{1}_{A_m}(x)\,d\mathbb{P} =$$

$$= \sum c_m \int \mathbb{1}_{A_m}(x) f(x)\,dx = \int \underbrace{\sum \mathbb{1}_{A_m}(x) c_m}_{g(x)} f(x)\,dx$$

Caso 3 $g \geq 0$

Nesse caso tentaremos escrever g como função simples

Seja $g_n(x) = \min\left(\frac{[2^n g(x)]}{2^n}, n\right) = \frac{[2^n g(x)]}{2^n} \wedge n$,

$[g(x)]$ é maior inteiro maior que $g(x)$

Nesse caso, g_n é função simples e $g_n \uparrow g$ ($g_n \geq 0$)

Então pelo TCM, segue que

$$E[g(X)] = \lim E(g_n(X)) = \lim \int g_n(x)\,d\mathbb{P} = \lim \int g_n(x) f(x)\,dx =$$

1.7.2.

$g \geq 0$ uma função mensurável em $(X, \mathcal{A}, \mu)$.

$$\int_X g \, d\mu = (\mu \times \lambda)(\{x,y\} : 0 \leq y \leq g(x)\}) = \int_0^{\infty} \mu(\{x : g(x) \geq y\}) dy$$

Teo. 1.7.2. Se $f \geq 0$ ou $\int |f| d\mu < \infty$, então

$$\int_X \int_Y f(x,y) \, \mu_2(dy) \, \mu_1(dx) = \int_{X \times Y} f \, d\mu = \int_Y \int_X f(x,y) \, \mu_1(dx) \, \mu_2(dy)$$

Basta mostrar que cada um dos termos do enunciado corresponde a um dos termos da ig. do teorema. Dessa forma, ao tomarmos λ como a medida de Lebesgue, $Y := [0, \infty)$ e $\mathcal{B}$ como os borelianos e fazendo

$$f(x,y) = \mathbb{1}_{\{(x,y) : 0 \leq y \leq g(x)\}}$$

Observamos

$$\int_X \Big[\underbrace{\int_Y \mathbb{1}(x,y) \, dy}_{g(x)} \Big] \mu(dx) = \int_X g(x) \, \mu(dx)$$

$$\int_Y \Big[\int_X f(x,y) \, \mu(dx) \Big] \mu(dy) = \int_0^{\infty} \mu(g(x) \geq y) \, dy$$

Se A tem densidade assintótica Θ então A^c tem densidade assintótica $1-\Theta$. No entanto, A não é fechada sob uniões.

Para provar que isto note que se A tem a propriedade de

$$|\{2k-1, 2k\} \cap A| = 1 \quad \forall\, k \in \mathbb{N} \quad \text{então}$$

$$\Theta(A) = \frac{1}{2}$$

$\{0, 1\}$

$k = 1$ $\quad \{1, 2\}$ $\qquad A = \{2n+1$

2 $\quad \{3, 4\}$

3 $\quad \{5, 6\}$

1.1.6.

$E = \{2n, n \geq 1\}$

$E \in A$ pois $\theta(A) = 1/2$.

Da mesma forma

$O = \{2n+1; n \geq 0\}$, $\theta(O) = 1/2$.

Para cada $m \geq 1$,

$$K_m = \begin{cases} E \cap \{3^m, \dots, 3^{m+1} - 1\} & \text{se } m \text{ é par} \\ E \cap \{3^m, \dots, 3^{m+1} - 1\} & \text{se } m \text{ é impar} \end{cases}$$

2.2 Lei Fraca dos Grandes Números

- $Y_n \to Y$ em probabilidade se para todo $\varepsilon > 0$, $P(|Y_n - Y| > \varepsilon) \to 0$ qdo $n \to \infty$.

2.2.1 Leis Fraca em L^2

- Uma família de v.a. $X_i, i \in I$ com $EX_i^2 < \infty$ é dita não-correlacionada se temos $E(X_i X_j) = EX_i EX_j$ sempre que $i \neq j$.

Teorema 2.2.1. Seja $X_1, \ldots, X_n$ com $E(X_i^2) < \infty$ e não-correlacionadas. Então

$$var(X_1 + X_2 + \ldots + X_n) = var(X_1) + \ldots + var(X_n)$$

Lema 2.2.2. Se $p > 0$ e $E|Z_n|^p \to 0$ então $Z_n \to 0$ em probabilidade.

Teorema 2.2.3. L^2 weak law. Se $X_1, X_2, \ldots$ são v.a. não correlacionadas com $EX_i = \mu$ e $var(X_i) \leq C < \infty$. Se $S_n = X_1 + \ldots + X_n$, então quando $n \to \infty$,

$$\frac{S_n}{n} \to \mu \text{ em } L^2 \text{ e em probabilidade.}$$

2.2.2 Arranjos Triangulares

Teorema 2.2.4. Seja $\mu_n = ES_n$, $\sigma_n^2 = Var(S_n)$. Se $\sigma_n^2 / b_n^2 \to 0$ então

$$\frac{S_n - \mu_n}{b_n} \to 0 \text{ em probabilidade.}$$

2.2.3 Truncation

Truncar uma v.a. X ao nível M significa considerar

$$\bar{X} = X \mathbb{1}_{(|X| \leq M)} = \begin{cases} X & \text{se } |X| \leq M \\ 0 & \text{se } |X| > M \end{cases}$$

Teorema 2.2.6 Lei Fraca Para Arranjos Triangulares.

Para cada n seja $X_{n,k}$, $1 \leq k \leq n$, independentes. Seja $b_n > 0$ com $b_n \to \infty$, e $\bar{X}_{n,k} = X_{n,k} \mathbb{1}_{(|X_{n,k}| \leq b_n)}$ Suponha que $n \to \infty$,

i) $\sum_{k=1}^{n} P(|X_{n,k}| > b_n) \to 0$ e,
ii) $b_n^{-2} \sum_{k=1}^{n} E\bar{X}_{n,k}^2 \to 0$

se fizermos $S_n = X_{n,1} + \dots + X_{n,n}$ e colocarmos $a_n = \sum_{k=1}^{n} E\bar{X}_{n,k}$, então

$$\frac{S_n - a_n}{b_n} \to 0 \text{ em probabilidade}$$

Teorema 2.2.7. Lei Fraca dos Grandes Números. Sejam $X_1, X_2, \dots$ i.i.d. com

$$x P(|X_1| > x) \to 0 \quad \text{qdo } x \to \infty$$

Seja $S_n = X_1 + \dots + X_n$ e $\mu_n = E(X_1 \mathbb{1}_{(|X_1| \leq n)})$, Então

$$\frac{S_n}{n} - \mu_n \to 0 \text{ em probabilidade}$$

Lema 2.2.8. Se $Y \geq 0$ e $p > 0$ então $E(Y^p) = \int_0^\infty p y^{p-1} P(Y > y)\, dy$

Teo. 2.2.9. Sejam $X_1, X_2, \dots$ i.i.d. com $E|X_i| < \infty$. Seja $S_n = X_1 + \dots + X_n$ e seja $\mu = EX_1$. Então $S_n/n \to \mu$ em probabilidade

2.3 Lemas de Borel-Cantelli

A_n é uma seq. de subconjuntos de Ω, fazemos

$$\limsup A_n = \lim_{m \to \infty} \bigcup_{n=m}^{\infty} A_n = \{\omega \text{ infinitas vezes}\}$$

$$\liminf A_n = \lim_{m \to \infty} \bigcap_{n=m}^{\infty} A_n = \{\omega \text{ inf. vezes, exceto num nº finito}\}$$

Teorema 2.3.1 Lema de Borel-Cantelli

Se $\sum_{n=1}^{\infty} P(A_n) < \infty$ então

$$P(A_n \text{ i.o.}) = 0$$

Teorema 2.3.2. $X_n \to X$ em prob. se e só se para toda subsequência $X_{n(m)}$ há uma outra subsequência $X_{n(m_k)}$ que converge q.c. à X.

Teorema 2.3.5. Sejam $X_1, X_2, \ldots$ i.i.d. com $EX_i = \mu$ e $EX_i^4 < \infty$. Se $S_n = X_1 + \ldots + X_n$, então $\frac{S_n}{n} \to \mu$ q.c.

Teorema 2.3.6. Segundo Lema de Borel-Cantelli. Se os eventos A_n são independentes, então $\sum P(A_n) = \infty$ implica $P(A_n \text{ i.o.}) = 1$.

2.4 Lei Forte dos Grandes Números

Teo. 2.4.1. Lei Forte dos Grandes Números. Sejam $X_1, X_2, \ldots$ ind. duas a duas e ident. distribuídas com $E|X_i| < \infty$. Seja $EX_i = \mu$ e $S_n = X_1 + \ldots + X_n$. Então

$$\frac{S_n}{n} \to \mu \text{ q.c. qdo } n \to \infty$$

CAPÍTULO 3

Exercício (4ª edição)	*Solutions M.* (2ª edição)	Página no *Solutions M.*	Exercício (4ª edição)	*Solutions M.* (2ª edição)	Página no *Solutions M.*	Exercício (4ª edição)	*Solutions M.* (2ª edição)	Página no *Solutions M.*
3.1.1	1.1	26	**3.3.12**	3.12	32	**3.6.3**	6.3	40
3.1.2	1.2	26	**3.3.13**	3.13	33	**3.6.4**	6.4	40
3.1.3	1.3	26	**3.3.14**	3.14	33	**3.6.5**	6.5	41
3.1.4	1.4	26	**3.3.15**	3.15	33	**3.6.6**	6.6	41
3.2.1	2.1	27	**3.3.16**	3.16	33	**3.6.7**	6.7	41
3.2.2	2.2	27	**3.3.17**	3.17	34	**3.6.8**	6.8	42
3.2.3	2.3	27	**3.3.18**	3.18	34	**3.6.9**	6.9	42
3.2.4	2.4	28	**3.3.19**	3.19	34	**3.6.10**	6.10	42
3.2.5	2.5	28	**3.3.20**	3.20	34	**3.6.11**	6.11	42
3.2.6	2.15	30	**3.3.21**	3.21	34	**3.6.12**	6.12	42
3.2.7	2.16	30	**3.3.22**	3.22	34	**3.6.13**	6.13	42
3.2.8	2.17	31	**3.3.23**	3.23	35	**3.7.1**	7.1	43
3.2.9	2.6	28	**3.3.24**	3.24	35	**3.7.2**	7.2	43
3.2.10	2.7	28	**3.3.25**	3.25	35	**3.7.3**	7.3	43
3.2.11	2.8	28	**3.3.26**	3.26	35	**3.7.4**	7.4	43
3.2.12	2.9	29	**3.3.27**	3.27	35	**3.7.5**	7.5	44
3.2.13	2.10	29	**3.3.28**	3.28	35	**3.7.6**	7.6	44
3.2.14	2.11	29	**3.4.1**	4.1	35	**3.7.7**	7.7	44
3.2.15	2.12	30	**3.4.2**	4.2	36	**3.7.8**	7.8	45
3.2.16	2.13	30	**3.4.3**	4.3	36	**3.8.1**	8.1	45
3.2.17	2.14	30	**3.4.4**	4.4	36	**3.8.2**	8.2	45
3.3.1	3.1	31	**3.4.5**	4.5	37	**3.8.3**	8.3	45
3.3.2	3.2	31	**3.4.6**	4.6	37	**3.8.4**	8.4	45
3.3.3	3.3	31	**3.4.7**	4.7	37	**3.9.1**	9.1	46
3.3.4	3.4	32	**3.4.8**	4.8	38	**3.9.2**	9.2	46
3.3.5	3.5	32	**3.4.9**	4.9	38	**3.9.3**	9.3	46
3.3.6	3.6	32	**3.4.10**	4.10	38	**3.9.4**	9.4	46
3.3.7	3.7	32	**3.4.11**	4.11	38	**3.9.5**	9.5	46
3.3.8	3.8	32	**3.4.12**	4.12	38	**3.9.6**	9.6	47
3.3.9	3.9	32	**3.4.13**	4.13	39	**3.9.7**	9.7	47
3.3.10	3.10	32	**3.6.1**	6.1	39	**3.9.8**	9.8	47
3.3.11	3.11	32	**3.6.2**	6.2	39			

$$\lim_{T\to\infty} (2\pi)^{-1}\int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it}\,\varphi(t)\,dt = \mu(a,b) + \frac{1}{2}\mu(\{a,b\})$$

3. Teorema Limite Central

3.2 Convergência Fraca

É dito que uma sequência de funções distribuições convergem fracamente à um limite F ($F_n \Rightarrow F$) se $F_n(y) \to F(y)$ para todo y que é ponto de continuidade de F.

Diz-se que a seq. de variáveis aleatórias X_n é conv. frac. ou em dist. à um limite X_∞ ($X_n \Rightarrow X_\infty$) se suas f.d. $F_n(x) = P(X_n \leq x)$ conv. frac.

3.2.1 Exemplos

3.2.2 Teoria

Teorema 3.2.2. $F_n \Rightarrow F_\infty \Rightarrow \exists\, Y_n,\ 1 \leq n \leq \infty$, com $Y_n \sim F_n$ de modo que $Y_n \to Y_\infty$ q.c.

Teorema 3.2.7. Toda limite subsequencial é uma f.d. se e só se a seq. F_n é "tight", isto é, para todo $\varepsilon > 0$ há um M_ε de modo que

$$\limsup_{n \to \infty} 1 - [F_n(M_\varepsilon) - F_n(-M_\varepsilon)] \leq \varepsilon$$

3.3 Funções Características

3.3.1 Definição, Formula da Inversão

$$\varphi(t) = E e^{itX} = E \cos tX + i \operatorname{sen} tX$$

Teorema. 3.3.4. Teorema Formula da Inversão. Seja $\varphi(t) = \int e^{itx} \mu(dx)$ onde μ é uma med. de prob. Se $a < b$, então

$$\lim_{T \to \infty} (2\pi)^{-1} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)\, dt = \mu(a,b) + \frac{1}{2}\mu(\{a,b\})$$

3.3.2 Convergência Fraca

Teo. 3.3.6. Teo. da Continuidade. Seja μ_n, $1 \leq n \leq \infty$ as med. de probabilid. com f.c. φ_n.

i) Se $\mu_n \Rightarrow \mu_\infty$ então $\varphi_n(t) \to \varphi_\infty(t)$ $\forall t$

ii) Se $\varphi_n(t)$ converge ponto a ponto a um limite $\varphi(t)$ que é contínua em 0, então a sequência de distrib. μ_n associada é tight e converge fracamente a uma medida μ com f.c. φ.

3.4. Teorema Limite Central

3.4.1. Seq. i.i.d.

Teor. 3.4.1. Seja $X_1, X_2, \ldots$ i.i.d. com $EX_i = \mu$, $var(X_i) = \sigma^2 \in (0, \infty)$, se $S_n = X_1 + X_2 + \cdots + X_n$, então

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow \mathcal{X} \sim N(0,1)$$

3.4.2 Arr. Triangulares

Teo. 3.4.5. The Lindberg-Feller Theorem. Para cada n, seja $X_{n,m}$, $1 \leq m \leq n$, v.a. ind. com $EX_{n,m} = 0$. Suponha

(i) $\sum_{m=1}^{n} EX_{n,m}^2 \to 0$

(ii) $\forall \varepsilon > 0$, $\lim_{n\to\infty} \sum_{m=1}^{n} E(|X_{n,m}|^2; |X_{n,m}| > \varepsilon) = 0$

Então $S_n = X_{n,1} + \cdots + X_{n,n} \Rightarrow \sigma\mathcal{X}$ qdo $n \to \infty$.

Teorema Limite Central

1. Teorema de De Moivre-Laplace

$\{X_n\}_{n\geq 1}$ i.i.d. com $P(X_1=1)=P(X_1=-1)=1/2$. $S_n=\sum_{m=1}^{n} X_m$.
Se n e k são inteiros

$$P(S_{2n}=2k)=\binom{2n}{n+k}2^{-2n}$$

• Stirling's formula:

$$n! \sim n^n e^{-n}\sqrt{2\pi n} \quad \text{quando } n\to\infty$$

$$\binom{2n}{n+k} = \frac{(2n)!}{(n-k)!\,(n+k)!} \approx \frac{(2n)^{2n}}{(n+k)^{n+k}(n-k)^{n-k}}\cdot\frac{e^{2n}}{e^{n+k}e^{n-k}}\,\frac{\sqrt{2\pi\, 2n}}{\sqrt{2\pi(n+k)\cancel{2}\pi(n-k)}} = \frac{2^{2n}\,n^{2n}}{(n+k)^{n+k}(n-k)^{n-k}}\;\frac{n^{1/2}}{(n+k)^{1/2}(n-k)^{1/2}\pi^{1/2}}$$

Daí,

$$P(S_{2n}=2k)\approx \frac{1}{\left(\frac{n+k}{n}\right)^{n+k}}\cdot\frac{1}{\left(\frac{n-k}{n}\right)^{n-k}}\cdot\frac{1}{\left(\frac{n+k}{n}\right)^{1/2}\cdot\left(\frac{n-k}{n}\right)^{1/2}\cdot n^{1/2}\pi^{1/2}} =$$

$$= \underbrace{\left(1+\frac{k}{n}\right)^{-n-k}}_{a}\underbrace{\left(1-\frac{k}{n}\right)^{-n+k}}_{b}\left(1+\frac{k}{n}\right)^{-1/2}\left(1-\frac{k}{n}\right)^{-1/2}(n\pi)^{-1/2}$$

$$ab = \left(1+\frac{k}{n}\right)^{-n}\left(1-\frac{k}{n}\right)^{-n}\left(1+\frac{k}{n}\right)^{-k}\left(1-\frac{k}{n}\right)^{k} = \left(1-\frac{k^2}{n^2}\right)^{-n}\left(1+\frac{k}{n}\right)^{-k}\left(1-\frac{k}{n}\right)^{k}$$

Lema. Se $c_j\to 0$, $a_j\to\infty$ e $a_jc_j\to\lambda$ então $(1+c_j)^{a_j}\to e^{\lambda}$.

$$\lim_{x\to 0}\frac{\log(1+x)}{x}=1 \Rightarrow \lim_{c_j\to 0} a_jc_j\,\frac{\log(1+c_j)}{c_j}\to\lambda \Rightarrow \lim(1+c_j)^{a_j}=e^{\lambda}.$$

Exercício 3.1.1.

Usando o lema 3.1.1, vemos que se $2k = x\sqrt{2n}$, i.e., $k = x\sqrt{n/2}$,

$$\left(1-\frac{k^2}{n^2}\right)^{-n} = \left(1-\frac{x^2}{2n}\right)^{-n} \to e^{x^2/2},$$

$$c_j = -\frac{x^2}{2n} \to 0, \quad a_j = n \to \infty$$

$$c_j a_j = -\frac{x^2}{2} \ \forall j \to -\frac{x^2}{2}$$

$$\left(1+\frac{k}{n}\right)^{-k} = \left(1+\frac{x}{\sqrt{2n}}\right)^{-x\sqrt{n/2}} \to e^{-\frac{x^2}{2}}$$

$$c_j = \frac{x}{\sqrt{2n}} \to 0 \quad a_j = x\sqrt{\frac{n}{2}} \to \infty; \quad a_j c_j = \frac{x^2}{2}$$

$$\left(1-\frac{k}{n}\right)^{-k} = \left(1+\frac{x}{\sqrt{2n}}\right)^{x\sqrt{n/2}} \to e^{-x^2/2}$$

Para esta escolha de k, $k/n \to 0$, de modo que

$$\left(1+\frac{k}{n}\right)^{-1/2}\left(1-\frac{k}{n}\right)^{-1/2} \to 1$$

e colocando as coisas em cjto,

$$\frac{2k-0}{\sqrt{2n}} \to 0$$

Teorema 3.1.2. Se $2k/\sqrt{2n} \to x$, então

$$P(S_{2n} = 2k) \sim (\pi n)^{-1/2} e^{-x^2/2}.$$

Próximo passo.

$$P\left(a \le \frac{S_{2n}}{\sqrt{2n}} \le b\right) = P(a\sqrt{2n} \le S_{2n} \le b\sqrt{2n}) = \sum_{m \in [a\sqrt{2n},\, b\sqrt{2n}] \cap 2\mathbb{Z}} P(S_{2n} = m)$$

Fazendo $m = x\sqrt{2n}$,

$$\approx \sum_{x \in [a,b] \cap (2\mathbb{Z}/\sqrt{2n})} (2\pi)^{-1/2} e^{-x^2/2} (2/n)^{1/2} \xrightarrow[n\uparrow\infty]{} \int_a^b (2\pi)^{-1/2} e^{-x^2/2}\, dx$$

Teorema 3.1.3. O Teorema de De Moivre-Laplace.

Se $a < b$ e $m \to \infty$

$$P\left(a \leq \frac{S_m}{\sqrt{m}} \leq b\right) \to \int_a^b (2\pi)^{-1/2} e^{-x^2/2} \, dx$$

Exercícios: 3.1.2; 3.1.3; 3.1.4

3.2. Convergência Fraca

~~Uma seq. de~~ É dito que uma sequência de funções distribuições converge fracamente à um limite F ($F_n \Rightarrow F$) se $F_n(y) \to F(y)$ para todo y que é ponto de continuidade de F.

É dito que uma sequência de v.a. X_n converge fracamente ou converge em distribuição à um limite X_∞ ($X_n \Rightarrow X_\infty$) se suas funções distribuições $F_n(x) = P(X_n \leq x)$ convergem fracamente.

Obs: Converg. nos pto de cont. é suf. p/ identificar o limite, uma vez que F é contínua à direita e as descont. de F é, no máximo, enumerável.

Exemplos.

3.2.1. $\{X_n\}_{n \geq 1}$ i.i.d. com $P(X_i = 1) = P(X_i = -1) = 1/2$, $S_n = \sum_{m=1}^{n} X_m$. Teo. 3.1.3 implica

$$F_n(y) = P(S_n \leq y) \to \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx$$

3.2.2. $\{X_n\}_{n \geq 1}$ i.i.d. com f.d. F. O Teor. de Glivenko-Cantelli implica que p/ quase todo ω,

$$F_n(y) = n^{-1} \sum_{m=1}^{n} \mathbb{1}_{(X_m(\omega) \leq y)} \to F(y) \text{ para todo } y$$

Nos 2 últimos exemplos, a converg. ocorre p/ todo y, mesmo o segundo caso podendo ter pontos de descontinuidades.

Agora vamos, por meio de exemplos, mostrar porque a restrição da atenção aos pontos de continuidade.

Ex. 3.2.3. $X \sim F$. $Y_n = X + 1/n$ tem f.d.

$$F_n(x) = P(X + 1/n \leq x) = F(x - 1/n)$$

Quando $n \to \infty$, $F_n(x) \to F(x-) = \lim_{y \uparrow x} F(y)$, de modo que a convergência

só ocorre nos pontos de continuidade. Note que $n \to \infty$, Y_n fica muito próximo de X; ou seja, deveria ocorrer a conv., e ocorre, mas só nos ptos de continuidade de F.

$$X_n(\omega) = X(\omega) + \frac{1}{n} \xrightarrow{n \to \infty} X(\omega) \qquad Y_n \xrightarrow{q.c} X$$

Ex. 3.2.4. X_p = # tentativas nec. para obter um sucesso numa seq. de tentativas ind. com prob. p. Então

$$P(X_p \geq n) = (1-p)^{n-1} \quad , \text{para } n = 1, 2, 3, \ldots$$

$$P(X_p \geq n) = \sum_{m=n}^{\infty} P(X_p = m) = \sum_{m=n+1}^{\infty} (1-p)^{m-1} p = (1-p)^{n-1} \sum_{m=n}^{\infty} (1-p)^{m-n} p =$$

$$= (1-p)^{n-1} p \sum_{k=0}^{\infty} (1-p)^{k} = (1-p)^{n-1} p \frac{1}{1-(1-p)} = (1-p)^{n-1}.$$

Segue do Lema 3.1.1 que quando $p \to 0$

$$P(pX_p > x) \to e^{-x} \quad \text{para todo } x \geq 0$$

$x \to \frac{n}{p}$, $P(pX_p > x) = (1-p)^{px-1} - (1-p)^{px-1} p = (1-p)^{px-1}(1-p) = (1-p)^{px}$

$\to e^{-x}$, $\forall x \geq 0$. Em palavras,

pX_p converge frac. p/ um dist. exponencial.

Ex. 3.2.5. Problema do aniversário. $\{X_n\}_{n\geq 1}$ i.i.d. $\overset{unif.}{\text{distribuídos}}$ em $\{1,\dots,N\}$, e seja $T_N = \min\{n : X_n = X_m \text{ para algum } m < n\}$.

$$P(T_N > n) = \prod_{m=2}^{n}\left(1 - \frac{m-1}{N}\right) \qquad \underbrace{\frac{N}{N}\;\frac{N-1}{N}\;\frac{N-2}{N}\cdots\frac{N-n}{N}}_{n}$$

Qdo $N = 365$, esta é a prob. de 2 pessoas em um grupo de tamanho n não terem o aniversário no mesmo dia (assumindo que todos os nasc. são igualmente prováveis). Usando o Exercício 3.1.1, é fácil ver que

$$P\left(\frac{T_N}{N^{1/2}} > x\right) \xrightarrow{*} e^{-x^2/2} \quad \forall x \geq 0$$

*Nesse caso,

$$c_{m,n} = \frac{1-m}{N} \to 0 \text{ para } 1 \leq m \leq n \Rightarrow \max_{1\leq m\leq n} |c_{j,n}| \to 0.$$

Ademais, $\displaystyle\sum_{m=2}^{n} c_{m,n} = \sum_{m=2}^{xN^{1/2}}\left(\frac{1-j}{N}\right) = -\frac{1}{N}\sum_{m=2}^{xN^{1/2}}(j-1) = -\frac{1}{N}\,\frac{xN^{1/2}(xN^{1/2}+1)}{2} =$

$$= -\frac{x^2}{2} - \frac{x}{2N^{1/2}} \to -\frac{x^2}{2}$$

Daí,

$$P\left(\frac{T_N}{N^{1/2}} > x\right) \to e^{-x^2/2}, \; \forall x \geq 0$$

Podemos aproximar a prob $\{T_N > n\} = \left\{\frac{T_N}{N^{1/2}} > \underset{\substack{\downarrow \\ x}}{\frac{n}{N^{1/2}}}\right\}$ por $e^{-x^2/2}$.

Antes do próximo ex., precisamos de um resultado.

Teorema de Scheffé. Suponha termos as densidades f_n, $1 \leq n \leq \infty$, e $f_n \to f_\infty$ pontualmente quando $n \to \infty$. Então $\forall B \in \mathcal{R}$,

$$\left| \int_B f_n(x)\,dx - \int_B f_\infty(x)\,dx \right| \leqslant \int |f_n(x) - f_\infty(x)|\,dx =$$

$$= 2\int (f_\infty(x) - f_n(x))^+\,dx \to 0$$

pelo TGD, a igualdade segue do fato de que $f_n \geqslant 0$ e tem integral 1.

Sendo $g_n = f_\infty - f_n$, e seja g_n^+ a parte positiva de g_n e g_n^- a parte negativa de g_n. Note que $g_n^+ \leq f_\infty$ e $g_n^+ \to 0$ qdo $n \to \infty$ q.c. em $\mathbb{R}$. Uma vez que f é função densidade, ela é trivialmente integrável, então pelo TGD

$$\lim_{n\to\infty} \int_{\mathbb{R}} g_n^+\,dx = \int_{\mathbb{R}} 0\,dx = 0\text{, i.e., } \int_{\mathbb{R}} g_n^+\,dx \to 0 \text{ qdo } n \to \infty.$$

Mas

$$\underbrace{\int_{\mathbb{R}} g_n(x)\,d\mu = 0}_{1-1=0} \text{ de modo que } \int g_n^+(x)\,dx = \int g_n^-(x)\,dx$$

Portanto

$$\int_{\mathbb{R}} |g_n|\,dx = 2\int_S g^+\,d\mu \to 0 \text{ quando } n \to \infty$$

Escrevendo μ_n p/ as correspondentes medidas, mostramos que a ~~variação~~ norma da variação total

$$\|\mu_n - \mu_\infty\| \equiv \sup_B |\mu_n(B) - \mu_\infty(B)| \to 0$$

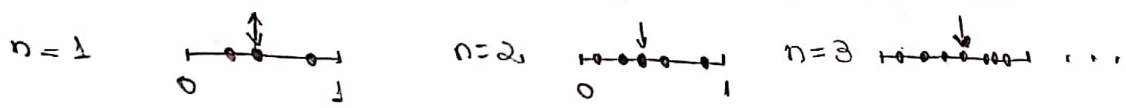
Uma conclusão mais forte que a convergência fraca. (Tome $B = (-\infty, x]$).
"O exemplo μ_n = n. ponto de massa em $1/n$ (com $1/\infty = 0$) mostra ~~em~~ que podemos ter $\mu_n \Rightarrow \mu_\infty$ com $\|\mu_n - \mu_\infty\| = 1\ \forall n$."

Exe. 3.2.1. Dê um exemplo de v.a. X_n com densid. f_n de modo que $X_n \Rightarrow$ uma distribuição uniforme em $(0,1)$ mas $f_n(x)$ não converge à 1 para qualquer $x \in [0,1]$.

Para tal, seja (?)

X_n : uma variável aleatória uniformemente distribuida em $\left[\frac{m}{2^n}, \frac{m}{2^n} + \frac{1}{2^n}\right]$, $0 < m < 2^n$ por

Exemplo 3.2.6. Central order statistic. Coloque $(2n+1)$ pontos ao aleatória em $(0,1)$, i.e., com localizações que são independentes e uniform. distribuídas. Seja V_{n+1} o $(n+1)$-ésimo maior ponto.



É fácil ver que.

ω_1, ω_2, ..., ω_{2n+1} Escolho o de interesse $\to$ $(2n+1)$ modos $\to$ colocos o n menores $\binom{2n}{n}$ $\to$ $\underbrace{x^n}_{<x}\underbrace{(1-x)^n}_{>x}$

Lema 3.2.1. V_{n+1} tem função densidade

$$f_{V_{n+1}}(x) = (2n+1)\binom{2n}{n} x^n (1-x)^n$$

Prova. Há $2n+1$ modos de escolher a observação que cai em x, daí temos que escolher n indices para as observações $< x$, que pode ser feito de $\binom{2n}{n}$ modos. Uma vez que decidimos sobre os indices que vão ficar $< x$ e $> x$, a prob. das corresp. v.a. fazerem o que queremos é $x^n(1-x)^n$, e a dens. de prob. de que o que resta caia em x é 1. Se não gostou da sentença prévia, compute a prob. $X_1 < x-\varepsilon, \dots, X_n < x-\varepsilon$, $x-\varepsilon < X_{n+1} < x+\varepsilon$, $X_{n+2} > x+\varepsilon, \dots, X_{2n+1} > x+\varepsilon$, então faça $\varepsilon \to 0$.

Exercício 3.2.2. Convergência do máximo.

Os limites do ex. acima são chamados distribuições de valores extremos. O último é chamado exp. dupla ou dist. de Gumbel.

Ex. 3.2.8. . . .

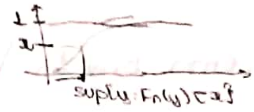
3.2.2 Teoria

Teorema. Se $F_n \Rightarrow F_\infty$ então há v.a. Y_n, $1 \leq n \leq \infty$, com distribuições de modo que $Y_n \to Y_\infty$ q.c.

Prova.

Seja $\Omega = (0,1)$, $\mathcal{F}$ = Cjtos de Borel, P = medida de Lebesgue e

$$Y_n(x) = \sup\{y : F_n(y) < x\}$$



Pelo Teo. 1.2.2, Y_n tem dist. F_n.

Vamos mostrar agora que $Y_n(x) \to Y_\infty(x)$ para todo x, com excessão de um cjto enumerável de x. Para fazer isso, é conveniente escrever $Y_n(x)$ como $F_n^{-1}(x)$ e retiramos o subscrito qdo $n=\infty$.

Começamos identificando o cjto excepcional. Seja $a_x = \sup\{y : F(y) < x\}$, $b_x = \inf\{y : F(y) > x\}$, e $\Omega_0 = \{x : (a_x, b_x) = \emptyset\}$ onde (a_x, b_x) é o intervalo aberto com os pontos de fim. $\Omega - \Omega_0$ é enumerável uma vez que (a_x, b_x) são disj. e cada int. não vazio contém um nº racional. Se

Se $x \in \Omega_0$ então $F(y) < x$ para $y < F^{-1}(x)$ e $F(z) > x$ para $z > F^{-1}(x)$.

Para mostrar que $F_n^{-1}(x) \to F^{-1}(x)$ para $x \in \Omega_0$, há duas coisas para mostrar.

a. $\liminf\limits_{n\to\infty} F_n^{-1}(x) \geq F^{-1}(x)$

Seja $y < F^{-1}(x)$ tal que F é contínua em y. Como $x \in \Omega_0$, $F(y) < x$ e se n é suf. grande $F_n(y) < x$, i.e., $F_n^{-1}(x) \geq y$. Uma vez que isto vale p/ todo y satisfazendo as restrições indicadas, o resultado segue.

b. $\limsup\limits_{n\to\infty} F_n^{-1}(x) \leq F^{-1}(x)$

Seja $y > F^{-1}(x)$ tal que F é contínua em y. Uma vez $x \in \Omega_0$, $F(y) > x$ e se n é suf. grande $F_n(y) > x$, i.e., $F_n^{-1}(x) \leq y$. Uma vez isto vale para todo y satisf. as restrições indicados, o res. segue e completamos a prova.

Ex. 3.2.4. Lema de Fatou. $g \geq 0$ cont. Se $X_n \Rightarrow X_\infty$,

$$\liminf Eg(X_n) \geq E(g(X_\infty))$$

Ex. 3.2.5. Integração no limite. g, h cont., $g(x) > 0$, e $|h(x)|/g(x) \to 0$ qdo $|x| \to \infty$. $F_n \Rightarrow F$ e $\int g(x)\,dF_n(x) \leq C < \infty$, $\int h(x)\,dF_n(x) \to \int h(x)\,dF(x)$.

O próximo ~~def~~ resultado demonstra a utilidade do Teorema 3.2.2 e dá uma def. equivalente de conv. fraca que faz sentido em qqer espaço topológico.

Teorema 3.2.3. $X_n \Rightarrow X_\infty$ se e somente se para toda função contínua limitada g, nós temos

$$Eg(X_n) \to Eg(X_\infty)$$

Prova. Seja Y_n com a mesma dist. de X_n e converge q.c. Desde que g é contínua, $g(Y_n) \to g(Y_\infty)$ q.c. e o teorema da convergencia limitada implica

$$Eg(X_n) = Eg(Y_n) \to Eg(Y_\infty) = Eg(X_\infty)$$

Para provar o inverso, seja

$$g_{x,\varepsilon} = \begin{cases} 1 & y \leq x \\ 0 & y \geq x+\varepsilon \\ \text{linear} & x \leq y \leq x+\varepsilon \end{cases}$$

Uma vez $g_{x,\varepsilon}(y) = 1$ para $y \leq x$, $g_{x,\varepsilon}$ é contínua, e $g_{x,\varepsilon}(y) = 0$ para $y > x+\varepsilon$,

$$\limsup_{n\to\infty} P(X_n \leq x) \leq \limsup_{n\to\infty} Eg_{x,\varepsilon}(X_n) = Eg_{x,\varepsilon}(X_\infty) \leq P(X_\infty \leq x+\varepsilon)$$

Fazendo $\varepsilon \to 0$ temos $\limsup_{n\to\infty} P(X_n \leq x) \leq P(X_\infty \leq x)$. A última conclusão é válida para qqer x. ... pg 101

Teorema 3.2.4. Continuous mapping theorem.

Seja g uma função mensurável e $D_g = \{x : g$ não é contínua em $x\}$. Se $X_n \Rightarrow X_\infty$ e $P(X_\infty \in D_g) = 0$, então

$$g(X_n) \Rightarrow g(X_\infty)$$

Se, em adição, g é limitada, então

$$E g(X_n) \to E g(X_\infty)$$

O próximo res. fornece um número de def. alternativas da convergência fraca.

- Teorema 3.2.5. As seguintes declarações são equivalentes.

(i) $X_n \Rightarrow X_\infty$

(ii) Para todo aberto G, $\liminf_{n\to\infty} P(X_n \in G) \geq P(X_\infty \in G)$

(iii) Para todo fechado K, $\limsup_{n\to\infty} P(X_n \in K) \leq P(X_\infty \in K)$

(iv) Para todo conjunto A com $P(X_\infty \in \partial A) = 0$, $\lim_{n\to\infty} P(X_n \in A) = P(X_\infty \in A)$

$\partial A \downarrow$ fronteira

- Teorema de seleção de Helly. 3.2.6. Para toda sequência F_n de funções distribuições, há uma subsequência $F_{n(k)}$ e uma função não-decrescente contínua à direita F de modo que

$$\lim_{k\to\infty} F_{n(k)}(y) = F(y)$$

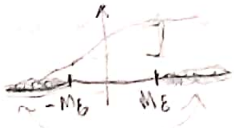
para todo ponto de continuidade y de F.

Remark: $a+b+c=1$, $\underbrace{F_n(x)}_{\text{ñ é f.d.}} = a\,\mathbb{1}_{\{x \geq n\}} + b\,\mathbb{1}_{\{x \geq -n\}} + c\,G(x)$, G um f.d.

$F_n(x) \to F(x) = b + cG(x)$. $\lim_{x\to\infty} F(x) = b+c = 1-a \neq 1$ que ñ é f.d. "Conv. Vaga"

Teorema 3.2.7. Todo limite subsequencial é a f.d. de probabilidade se e só se a sequencia F_n é tight, i.e., $\forall \varepsilon > 0$ há um M_ε de modo que

$$\limsup_{n \to \infty} 1 - [F_n(M_\varepsilon) - F_n(-M_\varepsilon)] \leq \varepsilon$$



massa das pontas não pode escapar

3.8. Funções Características

Longa seção → dividida em 5 partes. As 3 primeiras obrigatórias, as 2 últimas opcionais.

Na 1ª parte, mostramos que $\varphi(x)$ $\overset{E(e^{itX})}{=}$ determina $F(x) = P(X \leq x)$, e damos receitas p/ computar F a partir de φ. Na 2ª, relacionamos a converg. fraca de distrib. com o comport. dos φ's correspondentes. Na 3ª parte, relacionamos o comp. de φ em 0 com os momentos de X. Quarta parte: problema de Polya. Quinta: caracterização da distribuição por seus momentos.

3.8.1. Definição, Fórmula da Inversão

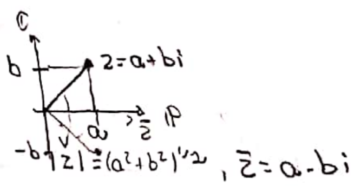
Se X é uma v.a.; definimos sua função característica por

$$\varphi(t) = Ee^{itX} = E\cos tX + iE\,\text{sen}\,tX$$

- Se Z é valorada nos complexos, definimos

$E[Z] = E(\text{Re}\,Z) + iE(\text{im}\,Z)$ onde $\text{Re}(a+bi) = a$ é a parte real e $\text{Im}(a+bi) = b$ é a pte imaginária

- $z = a+bi$, $|z| = |a+bi| = (a^2+b^2)^{1/2}$.
- $z = a+bi$, $\bar{z} = a-bi$ ↓ complexo conjugado
- $|z| = (z.\bar{z})^{1/2}$



Teorema 3.8.1. Todas as f. caract. tem as seguintes propriedades:

a) $\varphi(0) = 1$ $\qquad i.0 = 0$

$$\varphi(0) = E(e^{i0X}) = \cos(0) + i\,\text{sen}(0) = 1 + i.0 = 1$$

$$E(e^{i0X}) = E(e^0) = E(1) = 1.$$

b) $\varphi(-t) = \overline{\varphi(t)}$

$\varphi(-t) = E(e^{-itX}) = E(\cos(-tX) + i\,\mathrm{sen}(-tX)) = E(\cos(tX) - i\,\mathrm{sen}(tX)) = \overline{\varphi(t)}$

c) $|\varphi(t)| = |E\, e^{itX}| \leq E|e^{itX}| = 1.$

$\varphi(x,y) = (x^2+y^2)^{1/2}$ é convexa. Daí

d) $|\varphi(t+h) - \varphi(t)| \leq E|e^{ihX} - 1|$, assim $\varphi(t)$ é unif. cont. em $(-\infty,\infty)$

$$|\varphi(t+h) - \varphi(t)| = |E(e^{i(t+h)X} - e^{itX})|$$
$$\leq E|e^{i(t+h)X} - e^{itX}| = E|e^{ihX} - 1|$$

e) $E e^{it(aX+b)} = e^{itb}\varphi(at)$

Direto.

A principal razão p/ introduzir a f.c. é a seguinte.

Teo 3.8.2. X_1 e X_2 ind. com f.c. φ_1 e φ_2, respec., então X_1+X_2 tem f.c.
$$\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\,\varphi_{X_2}(t).$$

Prova. Direto.

Exemplo 3.3.1. Lançamentos de moedas. $P(X=1) = P(X=-1) = 1/2$, então
$$E(e^{itX}) = \frac{(e^{it}+e^{-it})}{2} = \cos t$$

$$E(e^{itX}) = \frac{1}{2}e^{it} + \frac{1}{2}e^{-it} = \frac{(e^{it}+e^{-it})}{2}$$

Ex. 3.8.2. Poisson.
$$\varphi(t) = E(e^{itX}) = \sum_{k=0}^{\infty} \frac{e^{itk}\lambda^k e^{-\lambda}}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(e^{it}\lambda)^k}{k!} = e^{-\lambda}e^{e^{it}\lambda} =$$
$$= e^{\lambda(e^{it}-1)}$$
//

Exemplo 3.3.3 Normal.

Densidade $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$$\varphi(t) = e^{-t^2/2}$$

Combinando este result. com (e) do Teo 3.3.1, vemos que a dist. normal com média μ e variância σ^2 tem $\varphi(t) = e^{i\mu t - \frac{\sigma^2 t^2}{2}}$.

$X \sim N(0,1)$

$\sigma^2 X + \mu \sim N(\mu, \sigma^2)$

$$\varphi_{\sigma^2 X + \mu}(t) = E(e^{it(\sigma^2 X + \mu)}) = e^{it\mu} E(e^{it\sigma^2 X}) = e^{it\mu} e^{-\frac{t\sigma^2}{2}} = e^{it\mu - \frac{t\sigma^2}{2}} \ \square$$

Prova física:

$$\int e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \overset{\text{compl. quadrado}}{=} e^{-t^2/2} \underbrace{\int \frac{1}{\sqrt{2\pi}} e^{-(x-it)^2/2} dx}_{= 1 \ N(it,1)}$$

Prova matemática:

$$\varphi(t) = \int e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \int \cos tx \, (2\pi)^{-1/2} e^{-x^2/2} dx$$

$\vdots$

Teorema 3.3.4. A Fórmula da Inversão.

$\varphi(t) = \int e^{itx} \mu(dx)$, μ uma medida de prob. Se $a < b$, então

$$\lim_{T \to \infty} (2\pi)^{-1} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt = \mu(a,b) + \frac{1}{2}\mu(\{a,b\})$$

Exercício 3.1.1.

Generalize a última prova para concluir que se $\max_{1\leq j\leq n} |c_{j,n}| \to 0$,

$\sum_{j=1}^{n} c_{j,n} \to \lambda$ e $\sup_n \sum_{j=1}^{n} |c_{j,n}| < \infty$ então

$$\prod_{j=1}^{n} (1 + c_{j,n}) \to e^{\lambda}$$

Uma vez que $\frac{\log(1+x)}{x} \to 1$ quando $x \to 0$, segue pela def. de limite que dado $\varepsilon > 0$ existe $\delta > 0$ tal que

$$|x - 0| < \delta \Rightarrow 1 - \varepsilon < \frac{\log(1+x)}{x} < 1 + \varepsilon \iff$$

$$|x| < \delta \Rightarrow (1-\varepsilon)x < \log(1+x) < (1+\varepsilon)x \quad (1)$$

Agora, como $\max_{1\leq j\leq n} |c_{j,n}| \to 0$, então $|c_{j,n}| \to 0$ para $1 \leq j \leq n$. Daí, segue também que $c_{j,n} \to 0$ para $1 \leq j \leq n$. Por (1), temos que para $\varepsilon > 0$ fixado existe um $\delta_{j,n} > 0$ tal que

$$|c_{j,n}| < \delta_{j,n} \Rightarrow (1-\varepsilon)c_{j,n} < \log(1+c_{j,n}) < (1+\varepsilon)c_{j,n}$$

Sendo agora $a_1(n) = (1-\varepsilon)c_{j,n}$, $a_2(n) = \log(1+c_{j,n})$, $a_3(n) = (1+\varepsilon)c_{j,n}$, e $s_1(n)$, $s_2(n)$, $s_3(n)$ as séries associadas a estas três sequências, respectivamente, temos que

$$s_1(n) < s_2(n) < s_3(n)$$

com $s_1(n) = \sum_{j=1}^{n} (1-\varepsilon)c_{j,n}$ e $s_3(n) = \sum_{j=1}^{n} (1+\varepsilon)c_{j,n}$.

Daí, segue que

finito pois $\sup_n \sum |c_{j,n}| < \infty$

$$(1-\varepsilon)\lambda \leq \lim_{n\to\infty} \sum_{j=1}^{n} \log(1+c_{j,n}) \leq (1+\varepsilon)\lambda$$

Como ε é arbitrário, temos

$$\lim_{n\to\infty} \log\left(\prod_{j=1}^{n}(1+c_{j,n})\right) = \lambda \Rightarrow$$ pela cont. do limite por f. contínuas

$$\lim_{n\to\infty} \prod_{j=1}^{n}(1+c_{j,n}) = e^{\lambda}$$

$\therefore$

$$\prod_{j=1}^{n}(1+c_{j,n}) \to e^{\lambda}$$

3.1.2.

Se X_i tem uma distribuição Poisson com média 1, então S_n tem uma distribuição Poisson com média n, i.e.,

$$P(S_n = k) = e^{-n} \frac{n^k}{k!}$$

Use a formula de Stirling p/ mostrar que se $\frac{(k-n)}{\sqrt{n}} \to x$ então

$$\sqrt{2\pi n}\, P(S_n = k) \to \exp(-x^2/2)$$

Como no caso do lançamento de moedas, segue que

$$P(a \leq (S_n - n)/\sqrt{n} \leq b) \to \int_a^b (2\pi)^{-1/2} e^{-x^2/2}\, dx$$

mas provar a última conclusão não é parte deste exercício.

CAPÍTULO 5

Exercício (4ª edição)	*Solutions M.* (2ª edição)	Página no *Solutions M.*	Exercício (4ª edição)	*Solutions M.* (2ª edição)	Página no *Solutions M.*	Exercício (4ª edição)	*Solutions M.* (2ª edição)	Página no *Solutions M.*
5.1.1	1.1	54	**5.2.10**	2.10	59	**5.4.8**	4.9	67
5.1.2	1.2	54	**5.2.11**	2.11	59	**5.4.9**	4.10	67
5.1.3	1.3	54	**5.2.12**	2.12	60	**5.4.10**	4.11	67
5.1.4	1.4	54	**5.2.13**	2.13	60	**5.5.1**	5.1	68
5.1.5	1.5	55	**5.2.14**	2.14	60	**5.5.2**	5.2	68
5.1.6	1.6	55	**5.3.1**	3.1	60	**5.5.3**	5.3	68
5.1.7	1.7	55	**5.3.2**	3.2	61	**5.5.4**	5.4	68
5.1.8	1.8	55	**5.3.3**	3.3	61	**5.5.5**	5.5	68
5.1.9	1.9	55	**5.3.4**	3.4	61	**5.5.6**	5.6	69
5.1.10	1.10	56	**5.3.5**	3.5	61	**5.5.7**	5.7	69
5.1.11	1.11	56	**5.3.6**	3.6	62	**5.5.8**	5.8	69
5.1.12	1.12	56	**5.3.7**	3.7	62	**5.6.1**	6.1	69
5.1.13	1.13	56	**5.3.8**	3.8	62	**5.6.2**	6.2	69
5.1.14	1.14	57	**5.3.9**	3.9	62	**5.6.3**	6.3	70
5.1.15	1.15	57	**5.3.10**	3.10	63	**5.6.4**	6.4	70
5.1.16	1.16	57	**5.3.11**	3.11	63	**5.6.5**	6.5	70
5.2.1	2.1	57	**5.3.12**	3.12	63	**5.7.1**	7.1	70
5.2.2	2.2	57	**5.3.13**	3.13	64	**5.7.2**	7.2	70
5.2.3	2.3	58	**5.4.1**	4.1	65	**5.7.3**	7.3	71
5.2.4	2.4	58	**5.4.2**	4.2	65	**5.7.4**	7.4	72
5.2.5	2.5	58	**5.4.3**	4.3	65	**5.7.5**	7.5	72
5.2.6	2.6	58	**5.4.4**	4.4	65	**5.7.6**	7.6	73
5.2.7	2.7	58	**5.4.5**	4.5	66	**5.7.7**	-	-
5.2.8	2.8	58	**5.4.6**	4.7	66	**5.7.8**	7.7	73
5.2.9	2.9	58	**5.4.7**	4.8	67	**5.7.9**	7.8	73

• Caso Discreto

$E \in \mathcal{F}, \quad \mathcal{S} = \mathcal{B}.$

• E um evento

$$P(E \mid \underbrace{X=x}_{\text{outro evento}}) = \frac{P(E, X=x)}{\underbrace{P(X=x)}_{>0 \text{ por c. discreto}}} = \frac{P(E, X=x)}{g(x)}, \quad x \in S$$

•

$$P(E, X \in A) = \sum_{x \in A} P(E, X=x) = \sum_{x \in A} g(x) P(E \mid X=x)$$

• $Q(x, E)$, def. para cada $x \in S$ e p/ eventos E, satisfaz

$$P(E, X \in A) = \sum_{x \in A} g(x) Q(x, E)$$

então $Q(x,E) = P(E \mid X=x)$ $\forall x \in S$ e evento E.

$\Omega = \{1, 2, 3, 4, \}$ $X(\omega) = \omega^2$

. . . .

$$P(\{3,4\}, X \in \{1,4\}) = \underbrace{P(\{3,4\}, X=1)}_{0} + \underbrace{P(\{3,4\}, X=4)}_{0} = 0$$

$$\sum_{\substack{x \in \{1,4\} \\ \downarrow \\ \in \mathcal{B}}} P(\{3,4\}, x) = \sum_{x \in \{1,2\}} p(x) P(\{3,4\} \mid X=x)$$

$$\frac{1}{2} \cdot 0 + \left(\frac{1}{2}\right) \cdot 0 = 0$$

• Caso contínuo $\quad (\Omega, \mathcal{F}, P) \xrightarrow{X} (S, \mathcal{A}, \mu)$

X tem dist. contínua em $S \subseteq \mathbb{R}^n$, com densidade g.

Assum $g(x) > 0$ $\forall x \in S$.

Infelizm não podemos simplesmente usar prob. condicional para definir de um evento E dado $\{X=x\}$, porque o evento condicionante tem prob. 0 p/ todo x.

No entanto, o conceito deveria fazer sentido. Se nós rodarmos um experimento, X tomará um certo valor x (embora a priori, este evento ocorre com prob. zero), e certamente a informação que $X=x$ deveria em geral alterar as probabilidades que nós atribuímos aos eventos.

Uma aproximação natural é usar os resultados obtidos no caso discreto como uma definição no caso contínuo. Deste modo, baseado na car. acima, definimos a prob. condicional

$$P(E \mid X=x), \quad x \in S$$

pelo requerimento que p/ qualquer (mensurável) A de S,

$$P(E, X \in A) = \int_A g(x) P(E \mid X=x)\, dx$$

Martingais

5.1. Esp. Condicional

Dado um esp. $(\Omega, \mathcal{F}_0, \mathbb{P})$, uma σ-álgebra $\mathcal{F} \subset \mathcal{F}_0$ e uma v.a. $X \in \mathcal{F}_0$ com $E|X| < \infty$, definimos a esperança condicional de X dado $\mathcal{F}$, $E(X/\mathcal{F})$, como qqr v.a. Y que tem

i) $Y \in \mathcal{F}$, i.e., é $\mathcal{F}$ mensurável, e

ii) $\forall A \in \mathcal{F}$, $\int_A X \, dP = \int_A Y \, dP$

Integrável, Existe, Única

ν é abs. cont. com resp. a μ ($\nu \ll \mu$) se $\mu(A) = 0 \Rightarrow \nu(A) = 0$.

R.-N. Teorema. Sejam μ e ν medidas σ-finitas em $(\Omega, \mathcal{F})$. Se $\nu \ll \mu$, há uma função $f \in \mathcal{F}$ de modo que $\forall A \in \mathcal{F}$,

$$\int_A f \, d\mu = \nu(A)$$

f é usual. denotada $d\nu/d\mu$ e é chamada deriv. de Radon-Nikodym.

$\Omega_1, \Omega_2, \ldots$ parti Ω e $\mathcal{F} = \sigma(\Omega_1, \Omega_2, \ldots)$

$$E(X/\mathcal{F}) = \frac{E(X; \Omega_i)}{P(\Omega_i)} \text{ em } \Omega_i$$

$P(A/\mathcal{G}) = E(\mathbb{1}_A / \mathcal{G})$ =

$E(X/Y) = E(X/\sigma(Y))$

5.2 Prop

a $E(aX + Y/\mathcal{F}) = aE(X/\mathcal{F}) + E(Y/\mathcal{F})$

b $X \leq Y$ então $E(X/\mathcal{F}) \leq E(Y/\mathcal{F})$

c $X_n \geq 0$ e $X_n \uparrow X$ com $EX < \infty$, então $E(X_n/\mathcal{F}) \uparrow E(X/\mathcal{F})$.

Se φ é conv. e $E|X|, E|\varphi(X)| < \infty$,

$$\varphi(E(X/\mathcal{F})) \leq E(\varphi(X)/\mathcal{F})$$

Te. se $\mathcal{F} \subset \mathcal{G}$ e $E(X/\mathcal{G}) \in \mathcal{F}$, então $E(X/\mathcal{F}) = E(X/\mathcal{G})$.

Te: se $\mathcal{F}_1 \subset \mathcal{F}_2$, então

i) $E(E(X/\mathcal{F}_1)/\mathcal{F}_2) = E(X/\mathcal{F}_1)$

ii) $E(E(X/\mathcal{F}_2)/\mathcal{F}_1) = E(X/\mathcal{F}_1)$

o menor wins

Teor. Se $X \in \mathcal{F}$ e $E|Y|, E|XY| < \infty$, então

$$E(XY/\mathcal{F}) = X\,E(Y/\mathcal{F})$$

Martingais

Seja $\mathcal{F}_n$ uma filtragem, i.e., uma seq. crescente de σ-álgebras.

Uma seq. X_n é dita adaptada à $\mathcal{F}_n$ se $X_n \in \mathcal{F}_n \ \forall n$.

Se X_n é seq. com

i) $E|X_n| < \infty$

ii) X_n é adap. à $\mathcal{F}_n$,

iii) $E(X_{n+1}/\mathcal{F}_n) = X \ \forall n$

então X é dita um martingal c.r.a $\mathcal{F}_n$.

iii $\leq$ super

$\geq$ sub

Te. X_n é um sup. então p/ $n \geq m$, $E(X_n/\mathcal{F}_m) \leq X_m$.

Teor 5.2.3. Se X_n é um mart. c.r.a. $\mathcal{F}_n$ e φ é uma f. convexa com $|E\varphi(X_n)| < \infty$ p/ todo n, então $\varphi(X_n)$ é um submartingal c.r.a $\mathcal{F}_n$. Consequent., se $p \geq 1$ e $E|X_n|^p < \infty$ p/ todo n, $|X_n|^p$ é um submarting c.r.a $\mathcal{F}_n$.

5.2.4. Se X_n é um sub c.r.a. $\mathcal{F}_n$ e φ é uma f. convexa crescente com $E|\varphi(X_n)| < \infty$ p/ todo n, então $\varphi(X_n)$ é um sub. c.r.a. $\mathcal{F}_n$

Conseq.

i) Se X_n é sub, então $(X_n - a)^+$ é sub

ii) Se X_n é super, então $X_n \wedge a$ é super

Seja $\mathcal{F}_n, n \geq 0$.

$H_n, n \geq 1$ é dita uma seq preditivel se $H_n \in \mathcal{F}_{n-1}$ p/ todo $n \geq 1$.

Em p., o valor de H_n pode ser predito (com certeza) a partir do valor avaliado em $n-1$.

H_n é qtde dinheiro apost. em n.

$$(H \cdot X)_n = \sum_{m=1}^{n} H_m (X_m - X_{m-1}) \text{ ganho no tempo } n$$

$X_m - X_{m-1} = \pm 1$.

Teor. 5.2.5. Seja $X_n, n \geq 0$ um super martingal. Se $H_n \geq 0$ é predizivel e cada H_n é limitada $(H \cdot X_n)_n$ é um super.

N é um t. parada $\Leftrightarrow \{N = n\} \in \mathcal{F}_n$ p/ todo $n < \infty$

teor. 5.2.10 Dec de Doob. Qquer sub. $X_n, n \geq 0$, pode ser escrito em um único modo como $X_n = M_n + A_n$, onde M_n é um marti e A_n é uma sequen pred, crescente com $A_0 = 0$.

$$\binom{n}{m} \frac{m!\,(n-m)!}{(n+1)!} = \frac{1}{n+1}$$

Teo 5.2.9 Se $X_n \geq 0$ é um sup, então qdo $n \to \infty$, $X_n \to X$ q.c., e

$EX \leq EX_0$.

T. 5.2.8. Teor. Conv. Mart. Se X_n é um sub. com $\sup E X_n^+ < \infty$, então qdo $n \to \infty$, X_n converge q.c. a um limite X com $E|X| < \infty$.

Exercício 5.1.1.

Generalize o último argumento para mostrar que se $X_1 = X_2$ em $B \in \mathcal{F}$, então $E(X_1/\mathcal{F}) = E(X_2/\mathcal{F})$ q.c. em B.

Sejam $Y_1 = E[X_1/\mathcal{F}]$ e $Y_2 = E[X_2/\mathcal{F}]$ versões das esperanças condicionais de X_1 e X_2, dado $\mathcal{F}$, respectivamente. Por definição, sabemos que, para $A \in \mathcal{F}$

$$\int_A Y_1 \, dP = \int_A X_1 \, dP \quad \text{e}$$

$$\int_A Y_2 \, dP = \int_A X_2 \, dP$$

Para $B \in \mathcal{F}$, em especial, temos que $X_1 = X_2$ q.c. $\Rightarrow \int_B X_1 dP = \int_B X_2 dP$. Isto, para $B \in \mathcal{F}$

$$\int_B Y_1 \, dP = \int_B Y_2 \, dP$$

Ademais, seja $C = \{Y_1 - Y_2 \geq \varepsilon > 0\} \cap B$. Temos que

$$0 = \int_C X_1 - X_2 \, dP = \int_C Y_1 - Y_2 \, dP \geq \varepsilon \mu(C) \Rightarrow \mu(C) = 0$$

Daí, em B, $E[X_1/\mathcal{F}] = E[X_2/\mathcal{F}]$ q.c.

Exercício 5.1.2. Fórmula de Bayes.

Seja $G \in \mathcal{G}$ e mostre que

$$P(G|A) = \int_G P(A|\mathcal{G})\,dP \Big/ \int_\Omega P(A|\mathcal{G})\,dP$$

Basta notar que

$$P(G/A) = \frac{P(G \cap A)}{P(A)},$$

Mas, uma vez que $G \in \mathcal{G}$,

$$P(G \cap A) = \int_G \mathbb{1}_A \, dP \overset{*}{=} \int_G P(A|\mathcal{G})\,d\mathbb{P}$$

* pelo fato de $P(A|\mathcal{G})$ ser uma versão da esperança condicional de $X = \mathbb{1}_A$ dado $\mathcal{G}$.

e $\quad \mathbb{P}(A) = P(A \cap \Omega) \overset{*}{=} \int_\Omega P(A|\mathcal{G})\,d\mathbb{P}$

* Basta tomar $B = \Omega \in \mathcal{G}$, uma vez que ela é σ-álgebra

Exercício 5.1.3. Prove a Desigualdade de Chebyshev. Se $a > 0$, então

$$P(|X| \geq a | \mathcal{F}) \leq a^{-2} E(X^2 / \mathcal{F})$$

Para tal, sejam

$$W = a^2 \mathbb{1}_{\{|X| \geq a\}} \quad e \quad Z = X^2$$

Temos que $W \leq Z$, e portanto, para $A \in \mathcal{F}$

$$\int_A E(W|\mathcal{F})\, dP = \int_A a^2 \mathbb{1}_{\{|X| \geq a\}}\, dP \leq \int_A X^2\, dP = \int_A E(Z/\mathcal{F})\, dP$$

Ao considerar $A = \{E(W|\mathcal{F}) - E(Z/\mathcal{F}) \geq \varepsilon > 0\}$, temos

$$\varepsilon \mu(A) \leq \int_A E(W/\mathcal{F}) - E(Z/\mathcal{F})\, dP = \underbrace{\int_A a^2 \mathbb{1}_{\{|X| \geq a\}} - X^2\, dP} \leq 0 \Rightarrow \mu(A) \leq 0 \Rightarrow$$

$$\mu(A) = 0.$$

Ou seja,

$E(W|\mathcal{F}) \leq E(Z/\mathcal{F})$ i.e. pela linearidade de $E(\cdot/\mathcal{F})$

$$a^2 E(\mathbb{1}_{\{|X| \geq a\}} / \mathcal{F}) \leq E(X^2/\mathcal{F}) \Rightarrow \text{por } E(\mathbb{1}_A | \mathcal{F}) = P(A/\mathcal{F})$$

$$P(|X| \geq a / \mathcal{F}) \leq \frac{E(X^2/\mathcal{F})}{a^2}$$

Exercício 5.1.4. Suponha $X \geq 0$ e $EX = \infty$. (Não há a se provar qdo $EX < \infty$) Mostre que existe uma única Y $\mathcal{F}$-mensurável com $0 \leq Y \leq \infty$ de modo que

$$\int_A X\, dP = \int_A Y\, dP \quad \text{para todo } A \in \mathcal{F}$$

Sugestão: Seja $X_M = X \wedge M$, $Y_M = E(X_M/\mathcal{F})$, e faça $M \to \infty$

Seja

$$X_m = X \wedge m \quad e \quad Y_m = E[X_m \mid \mathcal{F}]$$

Temos que, para $A \in \mathcal{F}$,

$$\int_A Y_m \, d\mathbb{P} = \int_A X \wedge m \, d\mathbb{P} = \int_{A \cap \{X > m\}} m \, d\mathbb{P} + \int_{A \cap \{X \leq m\}} X \, d\mathbb{P} \leq$$

$$= \int_{A \cap \{X > m\}} X \, d\mathbb{P} + \int_{A \cap \{X \leq m\}} X \, d\mathbb{P} =$$

$$= \int_A X \, d\mathbb{P} = \int_A Y \, d\mathbb{P}.$$

Na verdade com $X_{m_1} \leq X_{m_2}$, para $m_1 \leq m_2$, segue que $Y_m \uparrow$ e é limitada por Y.

$$X \wedge m \uparrow X \Rightarrow \int X \, d\mathbb{P} = \lim_{m \to \infty} \int X \wedge m \, d\mathbb{P} = \lim_{m \to \infty} \int Y_m \, d\mathbb{P} = \int Y \, d\mathbb{P}$$

Logo Y é uma versão da esp. cond de X dado $\mathcal{F}$.

Teorema 5.1.3. Se φ é convexa e $E|X|, E|\varphi(X)| < \infty$, então

$$\varphi(E(X/\mathcal{F})) \leq E(\varphi(X)/\mathcal{F}) \qquad (5.1.4)$$

Prova. Se φ é linear, o resultado é trivial:

$$\varphi(x) = ax + b \quad , \forall A \in \mathcal{F}$$

$$\int_A \varphi(X) dP = a\int_A X dP + b = \varphi(E(X/\mathcal{F})) \text{, i.e. } E(\varphi(X)/\mathcal{F}) = \varphi(E(X/\mathcal{F}))$$

Vamos então supor que φ é não-linear.
Nos fazemos isso de modo que se $S = \{(a,b) : a, b \in \mathbb{Q}, ax+b \leq \varphi(x)\ \forall x\}$, então $\varphi(x) = \sup\{ax+b : (a,b) \in S\}$.

$$E(\varphi(X)/\mathcal{F}) \geq a E(X/\mathcal{F}) + b \text{ q.c.}$$

Tomando o supremo sobre $(a,b) \in S$

$$E(\varphi(X)/\mathcal{F}) \geq \varphi(E(X/\mathcal{F})) \text{ q.c}$$

que prova o res.

Exercício 5.1.5. Imite a prova no remark após o Teo. 1.5.2 para provar a desigualdade condicional de Cauchy-Schwarz

$$E(XY/\mathcal{G})^2 \leq E(X^2/\mathcal{G})E(Y^2/\mathcal{G})$$

Note que $(X+\theta Y)^2/\mathcal{G} \geq 0$ q.o. Portanto

$$0 \leq E((X+\theta Y)^2/\mathcal{G}) = E(X^2 + 2\theta XY + \theta^2 Y^2/\mathcal{G}) \stackrel{\text{linearidade}}{=}$$

$$= E(X^2/\mathcal{G}) + 2\theta E(XY/\mathcal{G}) + \theta^2 E(Y^2/\mathcal{G}),$$

uma forma quadrática em θ^2. Pelo fato da mesma ser positiva, devemos ter

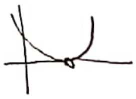
$$[\cancel{2} E(XY|\mathcal{G})]^2 \leq \cancel{4}.E(X^2|\mathcal{G}).E(Y^2|\mathcal{G}) \quad , \text{i.e.}$$

$$(E(XY|\mathcal{G}))^2 \leq E(X^2|\mathcal{G})E(Y^2|\mathcal{G})$$

//

* $ax^2+bx+c \geq 0 \Leftrightarrow$

$ax^2+bx+c>0$  não existe raiz $\therefore \Delta<0$

$ax^2+bx+c=0$  $\exists$ apenas uma raiz $\therefore \Delta=0$

i.e. $\Leftrightarrow b^2-4ac \leq 0$.

Teo. 5.1.4. Esperança condicional é uma contração em L^p, $p \geq 1$.

Prova.

(5.1.4) $\Rightarrow |E(X/\mathcal{F})|^p \leq E(|X|^p/\mathcal{F})$. Tomando o valor esperado

$$E(|E(X/\mathcal{F})|^p) \leq E(E(|X|^p/\mathcal{F})) = E|X|^p$$

⋮

Exercício. 5.1.6 Dê um exemplo em $\Omega=\{a,b,c\}$ em que

$$E(E(X/\mathcal{F}_1)|\mathcal{F}_2) \neq E(E(X/\mathcal{F}_2)|\mathcal{F}_1)$$

Seja $X:(\Omega,\mathcal{P}(\Omega),P)$ em $(\mathbb{R},\mathcal{R},\mu)$ t.q.

$X(a)=1$ e $P\{a\}=1/3$

$X(b)=2$ $\quad P\{b\}=1/3$

$X(c)=3$ $\quad P\{c\}=1/3$

Ao considerarmos

$\mathcal{F}_1 = \{\Omega, \{a\}, \{b,c\}, \emptyset\}$ e

$\mathcal{F}_2 = \{\Omega, \{b\}, \{a,c\}, \emptyset\}$

Temos que $\mathcal{F}_1 \not\subset \mathcal{F}_2$ e $\mathcal{F}_2 \not\subset \mathcal{F}_1$.

Sejam agora $Y_1 = E(X/\mathcal{F}_1)$ e $Y_2 = E(X/\mathcal{F}_2)$

Temos que $Y_1 \in \mathcal{F}_1$ e $Y_2 \in \mathcal{F}_2$ e

		a	b	c
$E(X/\mathcal{F}_1)$	Y_1	1	2,5	2,5
$E(X/\mathcal{F}_2)$	Y_2	2	2	2

Note agora que uma candidata para $E(E(X/\mathcal{F}_1)/\mathcal{F}_2) = W_1$ seria tal que

	a	b	c
W_1	1,75	2,5	1,75
W_2	2	2	2

$\Rightarrow W_1 \neq W_2$ q.c

Do livro:

$\mathcal{F}_1 = \{\Omega, \{a\}, \{b,c\}, \emptyset\}$, $\mathcal{F}_2 = \{\Omega, \{c\}, \{a,b\}, \emptyset\}$, $X(b) = 1$

$X(a) = X(c) = 0$

Neste caso

	a	b	c
$E(X/\mathcal{F}_1)$	0	1/2	1/2
$E(E(X/\mathcal{F}_1)/\mathcal{F}_2)$	1/4	1/4	1/2

Por outro lado

$E(X/\mathcal{F}_2)$	1/2	1/2	0
$E(E(X/\mathcal{F}_2)/\mathcal{F}_1)$	1/2	1/4	1/4

Na vdd, basta notar que

$E(E(X/\mathcal{F}_1)/\mathcal{F}_2) \notin \mathcal{F}_1$

$B = (-\infty, 1/4)$, $W_1^{-1}(B) = \{a,b\} \notin \mathcal{F}_1$

Teorema 5.1.7. Se $X \in \mathcal{F}$ e $E|Y|$, $E|XY| < \infty$, então

$$E(XY/\mathcal{F}) = X E(Y/\mathcal{F})$$

Exercício 5.1.7. Mostre que quando $E|X|$, $E|Y|$, e $E|XY|$ são finitas, cada declaração implica a próxima, e dê exemplos com $X, Y \in \{-1, 0, 1\}$ q.c. que mostram que a implicam reversa é falsa:

(i) X e Y são independentes;

(ii) $E(Y/X) = EY$,

(iii) $E(XY) = EX\,EY$

a) (i) $\Rightarrow$ (ii)

Sabemos que $E(Y/X) = E(Y/\sigma(X))$. Mas Y e ind. de X, e portanto, $\sigma(X)$ não trás nenhuma informação sobre Y. Logo $E(Y/X) = E(Y/\sigma(X)) = EY$

O reverso falha pelo ex. 4.2 do Cap. 1.

(ii) $\Rightarrow$ (iii)

$$E(XY) = E\,E(\underbrace{XY}_{XY \in \sigma(X)}/X) = E(X E(Y/X)) = E(X EY) = EX EY$$

Para ver que o inverso é falso

$X \backslash Y$	1	-1
1	1/4	0
0	0	1/2
-1	1/4	0

$EX = 0 = EY = 0 = EXY$ mas

$$E(Y/X) = -1 + 2X^2 = \begin{cases} 1 & \text{se } X \in \{-1,1\} \\ -1 & \text{se } X = 0 \end{cases}$$

Ex. 5.1.8. Mostre que se $\mathcal{G} \subset \mathcal{F}$ e $EX^2 < \infty$, então

$$E(\{X - E(X/\mathcal{F})\}^2) + E(\{E(X/\mathcal{F}) - E(X/\mathcal{G})\}^2) = E(\{X - E(X/\mathcal{G})\}^2)$$

Seja $Z = E(X/\mathcal{F}) - E(X/\mathcal{G}) \in \mathcal{F}$ e usando uma eq. da prova

($\mathcal{G}$-mensurável $\Rightarrow$ $\mathcal{F}$-mensurável)

de (1.4)

$$E\{X - E(X/\mathcal{F}) + Z\}^2 = E\{X - E(X/\mathcal{F})\}^2 + EZ^2$$

Subs Z

$$E\{X - E(X/\mathcal{F}) + E(X/\mathcal{F}) - E(X/\mathcal{G})\}^2 = E(X - E(X/\mathcal{F}))^2 + E\{E(X/\mathcal{F}) - E(X/\mathcal{G})\}^2)$$

$$E(X - E(X/\mathcal{F}))^2 + E\{E(X/\mathcal{F}) - E(X/\mathcal{G})\}^2 = E\{X - E(X/\mathcal{G})\}^2$$

Tirando o segundo termo do membro esquerdo da equação, temos uma desigualdade que diz geometricamente, maior o subespaço, mais próxima a projeção $(E(X - E(X/\mathcal{F}))^2 \leq E(X - E(X/\mathcal{G}))^2)$, ou estatisticamente, mais inf. significa um erro quadrado médio menor. Um caso importante ocorre quando $\mathcal{G} = \{\emptyset, \Omega\}$.

$$E(\{X - E(X/\mathcal{F})\}^2) \leq E(X - EX)^2 = Var(X) \quad //$$

5.1.9.

Seja $var(X/\mathcal{G}) = E(X^2/\mathcal{G}) - (E(X/\mathcal{G}))^2$. Mostre que

$$var(X) = E(var(X/\mathcal{G})) + var(E(X/\mathcal{G}))$$

$var(X/\mathcal{G}) = E(X^2/\mathcal{G}) - E(X/\mathcal{G})^2$ e $E(E(X^2/\mathcal{G})) = EX^2$, temos

$$E(var(X/\mathcal{G})) = EX^2 - E(E(X/\mathcal{G})^2) \qquad (1)$$

Uma vez $E(E(X/\mathcal{G})) = EX$, temos

$$var(\underbrace{E(X/\mathcal{G})}_{Y}) = E(\underbrace{E(X/\mathcal{G})^2}_{Y^2}) - \underbrace{(EX)^2}_{(E(Y))^2} \qquad (2)$$

Somando (1) e (2) temos

$$E(var(X/\mathcal{G})) + var(E(X/\mathcal{G})) = EX^2 - (EX)^2 = var(X) \;//$$

5.1.10

Sejam $Y_1, Y_2, \ldots$ i.i.d com média μ e variância σ^2, N uma v.a. indep. valorada nos int. positivos com $EN^2 < \infty$ e $X = Y_1 + \cdots + Y_N$. Mostre que $var(X) = \sigma^2 E(N) + \mu^2 var(N)$. Para entender e ajudar a entender o problema, pense sobre dois casos especiais em que N ou Y é constante.

Seja $\mathcal{F} = \sigma(N)$.

Nosso 1º passo é provar

$$E(X/N) = \mu N$$

Para tal, basta verificar que

i) $\mu N \in \mathcal{F}$

Com $N \in \mathcal{F}$, segue $\mu N \in \mathcal{F}$

ii) $\forall A \in \mathcal{F}$, $\int_A X \, dP = \int_A \mu N \, dP$

Para tal é suf. considerar $A = \{N = n\}$. Mas neste caso

$$\int_{\{N=n\}} X \, dP = E[(Y_1 + Y_2 + \cdots + Y_n) 1_{(N=n)}] \overset{ind.}{=}$$

$$= n\mu P(N=n) = \int_{\{N=n\}} N \mu \, dP$$

De modo similar, segue que

$$E(X^2/N) = \sigma^2 N + (\mu N)^2$$

Daí,

$$var(X/N) = \sigma^2 N + (\mu N)^2 - (\mu N)^2 = \sigma^2 N$$

Usando o ex. ant. $E(var(X/N)) + var(E(X/N))$

$$var(X) = \sigma^2 EN + \mu^2 var(N)$$

Ex. 5.1.11. X, Y, v.a. com $E(Y|\mathcal{G}) = X$ e $EY^2 = EX^2$, então $X = Y$ q.c.

Ex. 1.8. com $\mathcal{G} = \{\emptyset, \Omega\} \Rightarrow$

$$E(Y-X)^2 + E(X-EY)^2 = E(Y-EY)^2$$

como $\overset{E(E(Y|\mathcal{G}))}{EY} = EX$, e $EX^2 = EY^2$, $E(X-EX)^2 = E(Y-EY)^2$. Daí,

~~tendo~~ $E(X-EY)^2 = E(Y-EY)^2 \Rightarrow$

$$E(Y-X)^2 = 0 \qquad (*)$$

como $(Y-X)^2 \geq 0$, segue ~~que~~ que $Y-X=0$ q.c, i.e,

$$Y = X \text{ q.c.}$$

5.1.12. O result. no último ex. implica que se $EY^2 < \infty$ e $E(Y|\mathcal{G})$ tem a mesma distri. que Y, então $E(Y|\mathcal{G}) = Y$ q.c. Prove isto sob a suposição $E|Y| < \infty$.

Exercicio 5.2.1. Suponha X_n é um martingal c.r.a $\mathcal{G}_n$ e seja $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Então $\mathcal{G}_n \supset \mathcal{F}_n$ e X_n é um martingal c.r.a $\mathcal{F}_n$.

$\mathcal{F}_n$ é a menor σ-álgebra que faz $(X_1, \ldots, X_n)$ mensurável. Mas $X_1, \ldots, X_n \in \mathcal{G}_n$, de modo que $\mathcal{G}_n \supset \mathcal{F}_n$, uma vez que $\mathcal{F}_n$ é a menor.

Vemos então que $\mathcal{G}_n \supset \mathcal{F}_n$. Daí, pelos result. de esp. condicional,

$$E(X_{n+1} | \mathcal{F}_n) = E(E(X_{n+1} | \mathcal{G}_n) | \mathcal{F}_n) = E[X_n | \mathcal{F}_n] = X_n$$

\# $\mathcal{F}_n$ é a menor σ-álgebra que faz $X_1, X_2, \ldots, X_n$ mensuráveis. I.e. $\mathcal{F}_n$ é a menor σ-álgebra t.q. $X_i \in \mathcal{F}_n$, $1 \leq i \leq n$.

Como X_n é um martingal com respeito à filtragem $\{\mathcal{G}_n\}_{n \geq 1}$, segue que $X_1 \in \mathcal{G}_n, \ldots, X_n \in \mathcal{G}_n$, i.e., $\mathcal{G}_n$ é uma σ-álgebra que faz $X_1, X_2, \ldots, X_n$ mensuráveis. Uma vez que $\mathcal{F}_n$ é a menor com tal prop., segue que $\mathcal{G}_n \supset \mathcal{F}_n$. Daí,

$$E[X_{n+1} / \mathcal{F}_n] = E[E[X_{n+1} | \mathcal{G}_n] / \mathcal{F}_n] = E[X_n / \mathcal{F}_n] = X_n \,,,$$

Ex. 5.2.3. Dê um exemplo de um submartingal X_n de modo que X_n^2 é um supermartingal.

Dica: X_n não precisa ser aleatória.

Dizemos que $\{X_n\}_{n \geq 1}$ é um submartingal c.r.a $\{\mathcal{F}_n\}$ se

- $E|X_n| < \infty$
- X_n é adaptada à $\mathcal{F}_n$
- $E(X_{n+1} | \mathcal{F}_n) \geq X_n$

Seja $X_n = \frac{1}{n} - 1$,

$$E(X_{n+1} | \mathcal{F}_n) = E\left(\frac{1}{n+1} - 1 \,\middle|\, \mathcal{F}_n\right) = \frac{1}{n+1} - 1 > \frac{1}{n} - 1 = X_n$$

mas $X_n^2 = \left(\frac{1-n}{n}\right)^2$

$$E(X_{n+1}^2 | \mathcal{F}_n) = E\left(\left(\frac{1-n}{n}\right)^2 \,\middle|\, \mathcal{F}_n\right) = \left(\frac{1-n}{n}\right)^2 \leqslant \left(\frac{1n}{n-1}\right)^2$$

MAE 5811 - **Probabilidade Avançada I- Prova I**

Justifique suas respostas.

A nota máxima desta prova é 10

1. (Valor máximo 2,5)
   (a) (1.0) Sejam $A_1, A_2, \ldots$ subconjuntos de Ω. Mostre que $\limsup \mathbf{1}_{A_n} = \mathbf{1}_{\limsup A_n}$.
   (b) (1.0) Defina A_n, $n \geq 1$, sub-conjuntos da reta real, da seguinte forma: $A_n = (-1/n, 1]$ se n for impar e $A_n = (-1, 1/n]$ se n for par. Determine $\limsup A_n$ e $\liminf An$.
   (c) (1.0) Se $\mathcal{F}_i$ é uma σ-álgebra para cada $i \in I$, onde I é um conjunto não vazio de índices, mostre que $\cap_{i \in I} \mathcal{F}_i$ é uma σ-álgebra.
2. (Valor máximo 2,5) Mostre que Convergência em L^2 implica em
   (a) (1,5) convergência em L^1.
   (b) (1,5) convergência em Probabilidade.
3. (Extra 1.0) Seja $X : (\Omega, \mathcal{F}) \to (\mathbf{R}, \mathcal{B})$, com $\Omega = \{1, 2, 3\}$ e $\mathcal{F} = \{\Omega, \emptyset, \{1, 2\}, \{3\}\}$, $\mathbf{R}$ = números reais e $\mathcal{B}$ a σ-álgebra de Borel, definida por $X(1) = 1$, $X(2) = 2$, $X(3) = 3$.
   (a) (0.5) $(\Omega, \mathcal{F})$ é um espaço mensurável ?
   (b) (0,5) X é uma variável aleatória?
4. (2,0) Dadas duas classes de eventos $\mathcal{C}_1$ e $\mathcal{C}_2$ independentes entre si discutimos em aula um teorema que estabelece uma **condição** extra sob a qual podemos concluir que $\sigma(\mathcal{C}_1)$ e $\sigma(\mathcal{C}_2)$ são independentes.
   Seja $\Omega = \{1, 2, 3, 4\}$, $\mathcal{F}$ = classe de todos os sub-conjuntos de Ω e P tal que $P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\})$.
   (a) (1,0) Mostre que as duas classes de eventos $\mathcal{C}_1 = \{\{1, 2\}\}$ e $\mathcal{C}_2 = \{\{2, 3\}, \{2, 4\}\}$ são independentes mas que $\sigma(\mathcal{C}_1)$ e $\sigma(\mathcal{C}_2)$ não são independentes.
   (b) (1,0) Identifique a **condição** extra que não é verificada neste exemplo de $\mathcal{C}_1$ e $\mathcal{C}_2$.
5. (3,0) Seja, para cada $n \geq 1$, $X_n = \{X_{n,1}, X_{n,2}, \ldots, X_{n,n}\}$ uma coleção de variáveis aleatórias tal que $X_{n,j}$ tem distribuição Bernoulli(p), com $p = \frac{1}{n}$ (ou seja, $P(X_{n,j} = 1) = 1 - P(X_{n,j} = 0) = \frac{1}{n}$) para $1 \leq j \leq n$. Considere S_n o total de sucessos nesses n experimentos Bernoulli: $S_n = \sum_{j=1}^{n} X_{n,j}$.
   (a) (1,0) Se $X_{n,1}, X_{n,2}, \ldots, X_{n,n}$ são independentes, mostre que $P(S_n = k) \to e^{-\lambda}\frac{\lambda^k}{k!}$, quando $n \to \infty$ (distribuição de Poisson com média 1).
   (b) Construa agora $X_n = \{X_{n,1}, X_{n,2}, \ldots, X_{n,n}\}$ da seguinte forma:
   Para cada $n \geq 1$ Considere o espaço de probabilidade $(\Omega_n, \mathcal{F}_n, P_n)$ com $\Omega_n = \{1, 2, \ldots, n\}^n$, $\mathcal{F}_n$ = classe de todos os sub-conjuntos de Ω_n. Para definir $P_n(\omega)$, probabilidade de selecionar $\omega = (\omega_1, \omega_2, \ldots, \omega_n) \in \Omega$ ($\omega_j \in \{1, 2, \ldots, n\}$, $1 \leq j \leq n$) temos o seguinte procedimento: com probabilidade $\frac{1}{n}$ seleciono aleatoriamente em $\{(1, 1, \ldots, 1), (2, 2, \ldots, 2), \ldots, (n, n, \ldots, n)\} \subset \Omega$; com probabilidade $\frac{n-1}{n}$ seleciono aleatoriamente uma das $n!$ permutaçoes de $(1, 2, \ldots, n)$.
   Construa agora $X_n = \{X_{n,1}, X_{n,2}, \ldots, X_{n,n}\}$, tal que $X_{n,j}(\omega) = 1$ se $\omega_j = 1$ e $X_{n,j}(\omega) = 0$ caso contário.
      i. (0,5) Mostre que $X_{n,j}$ tem distribuição Bernoulli(p), com $p = \frac{1}{n}$.
      ii. (0,5) Mostre que $X_{n,j}$ e $X_{n,i}$ são independentes, para $i \neq j$.
      iii. (0,5) Mostre que a distribuição de $S_n = \sum_{j=1}^{n} X_{n,j}$ não converge para Poisson.
      iv. (0,5) Porque $S_n = \sum_{j=1}^{n} X_{n,j}$ não converge para Poisson mas isto era verdade no caso anterior (item a)?
6. (Extra 2.0) Entregue a resolução desta prova até a próxima segunda (28/04) as 14:00h. Esta resolução deve ser individual mas voce pode consultar livros e anotações pessoais.

Maicon Aparecido Pinheiro - 8637401

Questão 1.

a. Sejam $A_1, A_2, \dots$ subconjuntos de Ω. Mostre que $\limsup \mathbb{1}_{A_n} = \mathbb{1}_{\limsup A_n}$.

Por definição, temos que

$$\limsup A_n = \bigcap_{n \geq 1} \bigcup_{m \geq n} A_m$$

ou, de forma equivalente,

$$\limsup A_n = \bigcap_{n \geq 1} B_n, \quad \text{com } B_n = \bigcup_{m \geq n} A_m \qquad (1)$$

Como $B_1, B_2, \dots$ formam uma sequência de eventos não-crescente, pois se $\omega \in B_{k+1}$, então $\omega \in \bigcup_{m \geq k+1} A_m \subset \bigcup_{m \geq k} A_m = B_k$; isto é, $B_k \supset B_{k+1}$ para todo $k \geq 1$ natural. Daí, segue que $B_n \downarrow \bigcap_{n=1}^{\infty} B_n$, ou seja,

$$\lim_{n \to \infty} B_n = \bigcap_{n=1}^{\infty} B_n = \bigcap_{n=1}^{\infty} \bigcup_{m \geq n} A_m = \limsup A_m \qquad (2)$$

Agora, seja $\mathbb{1}_{B_n}$ denotando a variável indicadora de B_n para $n \geq 1$, e $\mathbb{1}_{\bigcap_{n=1}^{\infty} B_n} = \mathbb{1}_{\limsup A_n}$. Temos que.

$$\lim_{n \to \infty} \mathbb{1}_{B_n} = \mathbb{1}_{\limsup A_n} \qquad (3)$$

Vamos verificar que vale a identidade (3). Uma maneira rápida seria notar $\mathbb{1}_{B_n}$, $n \geq 1$ formam uma sequência de funções não-crescentes (Se A e C são conjuntos, $A \subset C$, então $\mathbb{1}_A \leq \mathbb{1}_C$) e limitada, de modo que para todo $\omega \in \Omega$, $\mathbb{1}_{B_n}(\omega)$ converge.

Note agora que se $\omega \in \bigcap_{n=1}^{\infty} B_n$, então $\omega \in B_n$ para cada $n \geq 1$. Neste caso, $\mathbb{1}_{B_n}(\omega) = 1$ para cada $n \geq 1$ e $\mathbb{1}_{\bigcap_{n=1}^{\infty} B_n} = \mathbb{1}_{\limsup A_n} = 1$. Se $\omega \notin \bigcap_{n=1}^{\infty} B_n$ então $\omega \notin B_k$ para algum k. Como os eventos são não-crescentes, $\omega \notin B_n$ para todo $n \geq k$. Neste caso, $\mathbb{1}_{B_n}(\omega) = 0$ para $n \geq k$ e $\mathbb{1}_{\bigcap_{n=1}^{\infty} B_n} = \mathbb{1}_{\limsup A_n} = 0$. Em ambos os casos,

$$\lim_{n\to\infty} \mathbb{1}_{B_n} = \mathbb{1}_{\limsup A_n}$$

Mas $\lim_{n\to\infty} \mathbb{1}_{B_n} = \lim_{n\to\infty} \mathbb{1}_{\bigcup_{i=n}^{\infty} A_n} \overset{*}{=} \lim_{n\to\infty} \max\{\mathbb{1}_{A_i} : i \geq n\} \overset{**}{=} \limsup \mathbb{1}_{A_n}$.

** segue pela definição de limsup

* Seja $A_1, A_2, \ldots$ conjuntos: $\mathbb{1}_{\bigcup_{m\geq n} A_m}(\omega) = 1 \Leftrightarrow \omega \in A_i$ para algum $i \geq n \Leftrightarrow \mathbb{1}_{A_i}(\omega) = 1$ para algum $i \geq n \Leftrightarrow \max\{\mathbb{1}_{A_i}(\omega) : i \geq n\} = 1$. Por outro lado, $\mathbb{1}_{\bigcup_{m\geq n} A_n}(\omega) = 0 \Leftrightarrow \omega \notin A_i$ para todo $i \geq n \Leftrightarrow \mathbb{1}_{A_i}(\omega) = 0\ \forall i \geq n \Leftrightarrow \max\{\mathbb{1}_{A_i}(\omega) : i \geq n\} = 0$.

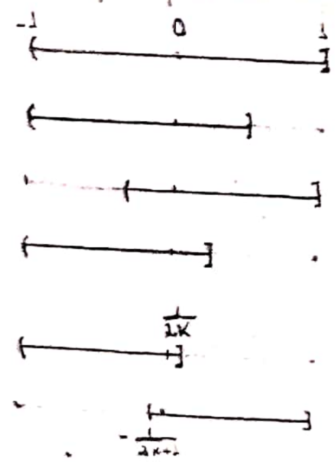
b) Defina A_n, $n \geq 1$, sub-conjuntos da reta real, da seguinte forma: $A_n = (-1/n, 1]$ se n for ímpar e $A_n = (-1, 1/n]$ se n for par. Determine $\limsup A_n$ e $\liminf A_n$.

Para tal, sejam $B_n = \bigcup_{m\geq n} A_m$ e $C_n = \bigcap_{m \geq n} A_m$. Nestas condições, temos que

i) $\limsup A_n = \bigcap_{n=1}^{\infty} B_n$;

ii) $\liminf A_n = \bigcup_{n=1}^{\infty} C_n$.

Vejamos agora que, para k natural,



$n=1$ A_1: ; $B_1 = \bigcup_{m\geq 1} A_m = (-1, 1]$; $C_1 = \bigcap_{m\geq 1} A_m = [0] = \{0\}$

$n=2$ A_2: ; $B_2 = \bigcup_{m\geq 2} A_m = (-1, 1]$; $C_2 = \bigcap_{m\geq 2} A_m = \{0\}$

$n=3$ A_3: ; $B_3 = \bigcup_{m\geq 3} A_m = (-1, 1]$; $C_3 = \{0\}$

$n=4$ A_4: ; $B_4 = \bigcup_{m\geq 4} A_m = (-1, 1]$; $C_4 = \{0\}$

$n=2k$ A_{2k}: ; $B_{2k} = \bigcup_{m\geq 2k} A_m = (-1, 1]$; $C_{2k} = \{0\}$

$n=2k+1$ A_{2k+1}: ; $B_{2k+1} = \bigcup_{m\geq 2k+1} A_m = (-1, 1)$; $C_{2k+1} = \{0\}$

Maicon Aparecido Pinheiro - 8637401

Com isso, segue que

i) $\limsup A_n = \bigcap_{n\geq 1} \bigcup_{m\geq n} A_m = \bigcap_{n\geq 1} B_n = (-1,1]$ e;

ii) $\liminf A_n = \bigcup_{m\geq 1} \bigcap_{n\geq m} A_m = \bigcup_{m\geq 1} C_n = \{0\}$

c) Se $\mathcal{F}_i$ é uma σ-álgebra para cada $i \in I$, onde I é um conjunto não-vazio de índices, mostre que $\bigcap_{i\in I} \mathcal{F}_i$ é uma σ-álgebra.

Por definição, uma coleção $\mathcal{A}$ não-vazia de subconjuntos de Ω que satisfaz

(1) $\Omega \in \mathcal{A}$, $\emptyset \in \mathcal{A}$;

(2) Se $A \in \mathcal{A}$ então $A^c \in \mathcal{A}$;

(3) Se $A_1, A_2, \ldots$ estão em $\mathcal{A}$, então $\bigcup_j A_j \in \mathcal{A}$.

*: Sendo $\mathcal{F}_i$ uma σ-álgebra para cada $i \in I$, segue que $\mathcal{F}_i$ satisfaz as condições (1), (2) e (3) para cada $i \in I$.

Verificar se $\bigcap_{i\in I} \mathcal{F}_i$ é uma σ-álgebra é equivalente a verificar se $\bigcap_{i\in I} \mathcal{F}_i$ satisfaz as condições (1), (2) e (3). Vamos verificá-las:

(1) $\Omega \in \mathcal{F}_i \ \forall i \in I \Rightarrow \Omega \in \bigcap_{i\in I} \mathcal{F}_i$. Logo, $\Omega \in \bigcap_{i\in I} \mathcal{F}_i$ e $\emptyset \in \bigcap_{i\in I} \mathcal{F}_i$.

$\emptyset \in \mathcal{F}_i \ \forall i \in I \Rightarrow \emptyset \in \bigcap_{i\in I} \mathcal{F}_i$

(2) $A \in \bigcap_{i\in I} \mathcal{F}_i \Rightarrow A \in \mathcal{F}_i \ \forall i \in I \overset{*}{\Rightarrow} A^c \in \mathcal{F}_i \ \forall i \in I \Rightarrow A^c \in \bigcap_{i\in I} \mathcal{F}_i$. Isto é, se $A \in \bigcap_{i\in I} \mathcal{F}_i$ então $A^c \in \bigcap_{i\in I} \mathcal{F}_i$ (vale condição 2).

(3) $A_1, A_2, \ldots$ em $\bigcap_{i\in I} \mathcal{F}_i \Rightarrow A_1, A_2, \ldots$ em $\mathcal{F}_i \ \forall i \in I \overset{*}{\Rightarrow} \bigcup_j A_j \in \mathcal{F}_i \ \forall i \in I \Rightarrow$ $\bigcup_j A_j \in \bigcap_{i\in I} \mathcal{F}_i$, i.e., se $A_1, A_2, \ldots$ estão em $\bigcap_{i\in I} \mathcal{F}_i$ então $\bigcup_j A_j \in \bigcap_{i\in I} \mathcal{F}_i$.

Como valem as três condições, segue que se $\mathcal{F}_i$ é uma σ-álgebra para cada

$i \in I$, então $\bigcap_{i \in I} \mathcal{F}_i$ é uma σ-álgebra.

Questão 2. Mostre que convergência em L^2 implica em

a) Convergência em L^1

b) Convergência em Probabilidade

Por definição, se X_n converge à X e L^2, com $E|X_n|^2 < \infty$ e $E|X|^2 < \infty$, então

$$E|X_n - X|^2 = E(X_n - X)^2 \to 0 \text{ quando } n \to \infty.$$

Obs: $EX = \int_\Omega X \, dP$

item a)

Dizemos que a sequência $\{X_n\}_{n \geq 1}$, com $E|X_n| < \infty$ e $E|X| < \infty$, converge em L^1 se

$$E|X_n - X| \to 0 \text{ quando } n \to \infty$$

Uma vez que $X_n \to X$ em L^2, temos que

$$E|X_n|^2 = \int_\Omega |X_n|^2 \, dP < \infty \text{ . Daí,}$$

$$E|X_n| = \int_\Omega |X_n| \, dP = \int_{\{|X_n| \leq 1\}} |X_n| \, dP + \int_{\{|X| > 1\}} |X_n| \, dP \leq \int_{\{|X| \leq 1\}} 1 \, dP + \int_{\{|X| > 1\}} |X_n|^2 dP =$$

$$= \underbrace{P(|X| \leq 1)}_{\in [0,1]} + E|X_n|^2 \leq 1 + \underbrace{E|X_n|^2}_{< \infty} < \infty$$

Maicon Aparecido Pinheiro - 8637401

De modo análogo, vemos que $EX^2 < \infty \Rightarrow E|X| < \infty$. Ou seja, as condições para a convergência em L^1 ocorrem.

Consideremos agora as funções

$$f = |X_n - X| \quad e \quad g = \mathbb{1}_{\Omega}$$

Pela Desigualdade de Cauchy-Schwarz (caso particular da Desigualdade de Hölder quando $p=q=2$), temos que

$$\int_{\Omega} |X_n - X| \mathbb{1}_{\Omega} \, d\mathbb{P} \leq \left(\int_{\Omega} |X_n - X|^2 d\mathbb{P} \right)^{1/2} \left(\int_{\Omega} |\mathbb{1}_{\Omega}|^2 \, d\mathbb{P} \right)^{1/2}, \text{ i.e.,}$$

$$\int_{\Omega} |X_n - X| d\mathbb{P} \leq \left(\int_{\Omega} |X_n - X|^2 d\mathbb{P} \right)^{1/2} \left(\underbrace{\int_{\Omega} d\mathbb{P}}_{P(\Omega)=1} \right)^{1/2} =$$

$$= \left(E(X_n - X)^2 \right)^{1/2}.$$

para cada n. Note agora que $|X_n - X| \geq 0$, e portanto, $\int_{\Omega} |X_n - X| d\mathbb{P} \geq 0$. Daí,

$$0 \leq \int_{\Omega} |X_n - X| d\mathbb{P} \leq \left(E(X_n - X)^2 \right)^{1/2} \quad \text{para cada } n. \quad (1)$$

Como $|X_n - X|^2 \geq 0$, segue também que $\int |X_n - X|^2 d\mathbb{P} \geq 0$, para todo n. Ademais, como $f(y) = \sqrt{y}$ é uma função contínua de $\mathbb{R}_+$ em $\mathbb{R}$, segue que

$$\lim_{n \to \infty} \sqrt{\underbrace{E(X_n - X)^2}_{\in \mathbb{R}_+}} = \sqrt{\lim_{n \to \infty} E(X_n - X)^2} = \sqrt{0} = 0.$$

Ou seja, uma vez que $E(X_n - X)^2 \to 0$ qdo $n \to \infty$, $(E|X_n - X|^2)^{1/2}$ também

Assim, quando $n \to \infty$, temos pela equação (1) que

$$0 \leq \int_{\Omega} |X_n - X| \, dP = E|X_n - X| \leq 0 \quad , \text{i.e.}$$

$$E|X_n - X| \to 0 \text{ quando } n \to \infty.$$

Logo, provamos que se $X_n \to X$ em L^2, então $X_n \to X$ em L^1.

item b)

É dito que a sequência $\{X_n\}_{n \geq 1}$ converge à X em probabilidade se $\forall \varepsilon > 0$, temos

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

Aqui, basta notar pela Desigualdade de Chebyshev que, $\forall \varepsilon > 0$ fixado,

$$0 \leq P(|X_n - X| > \varepsilon) = P(|X_n - X|^2 > \varepsilon^2) \leq \frac{E(|X_n - X|^2)}{\varepsilon^2} \to 0$$

quando $n \to \infty$, pois decorre da convergência em L^2 que $E|X_n - X|^2 \to 0$ quando $n \to \infty$.

Daí, provamos que a convergência em L^2 implica a convergência em probabilidade.

Questão 3. Se $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$, com $\Omega = \{1, 2, 3\}$ e $\mathcal{F} = \{\Omega, \emptyset, \{1, 2\}, \{3\}\}$, $\mathbb{R}$ = números reais e $\mathcal{B}$ a σ-álgebra de Borel, definida por $X(1) = 1$, $X(2) = 2$, $X(3) = 3$.

Maicon Aparecido Pinheiro - 8637401

a) $(\Omega, \mathcal{F})$ é um espaço mensurável?

Por definição, o par $(\Omega, \mathcal{F})$ será dito um espaço mensurável se $\mathcal{F}$ for uma σ-álgebra de subconjuntos de Ω, i.e., uma coleção não-vazia de subconjuntos de Ω satisfazendo as seguintes condições:

1. $\Omega \in \mathcal{F}$ e $\emptyset \in \mathcal{F}$

Nesse caso, $\mathcal{F} = \{\Omega, \emptyset, \{1,2\}, 3\}$, de modo que $\Omega \in \mathcal{F}$ e $\emptyset \in \mathcal{F}$.

Logo, vale a condição 1.

2. Se $A \in \mathcal{F}$ então $A^c \in \mathcal{F}$.

No caso, se $A \in \mathcal{F}$, então

$$A = \Omega \Rightarrow A^c = \emptyset \in \mathcal{F} \text{ ou}$$
$$A = \emptyset \Rightarrow A^c = \Omega \in \mathcal{F} \text{ ou}$$
$$A = \{3\} \Rightarrow A^c = \{1,2\} \in \mathcal{F} \text{ ou}$$
$$A = \{1,2\} \Rightarrow A^c = \{3\} \in \mathcal{F} .$$

Logo, vale a condição 2.

3. Se $A_1, A_2, \ldots$ estão em $\mathcal{F}$, então $\bigcup_i A_i \in \mathcal{F}$.

Aqui, basta notar que para este caso,

- $\bigcup_i A_i = \emptyset \in \mathcal{F}$, quando $A_i = \emptyset$ para $i \geq 1$; ou
- $\bigcup_i A_i = \Omega \in \mathcal{F}$, quando pelo menos um dos A_i's é Ω; ou quando há ao menos dois A_i's - digamos A_{i_1} e A_{i_2} - com um deles igual à $\{1,2\}$ e o outro à $\{3\}$; ou
- $\bigcup_i A_i = \{3\} \in \mathcal{F}$, quando ao menos um dos A_i's é igual à $\{3\}$, e o restante todos iguais ao $\emptyset$; ou
- $\bigcup_i A_i = \{1,2\} \in \mathcal{F}$, quando ao menos um dos A_i's é $\{1,2\}$ e os demais são todos iguais à $\emptyset$.

Ou seja, vale a condição 3. ∴ $\mathcal{F}$ é uma σ-álgebra de subconjuntos de Ω.

Logo, $(\Omega, \mathcal{F})$ é um espaço mensurável

b) X é uma variável aleatória?

Por definição, uma variável aleatória definida em $(\Omega, \mathcal{F})$ é uma função mensurável de $(\Omega, \mathcal{F})$ à $(\mathbb{R}, \mathcal{B})$.

Logo, verificar se X dada é uma variável aleatória é equivalente a verificar se para todo $B \in \mathcal{B}$, temos

$$X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F} ;$$

ou seja, se X é mensurável de $(\Omega, \mathcal{F})$ à $(\mathbb{R}, \mathcal{B})$.

Como visto no enunciado, temos que $X(\omega) = \omega \ \forall \omega \in \Omega = \{1,2,3\}$. Daí, ao considerarmos $B = (0, 3/2) \in \mathcal{B}$, teremos que

$$X^{-1}(B) = \{\omega : X(\omega) \in (0, 3/2)\} = \{1\} \notin \{\Omega, \emptyset, \{1,2\}, \{3\}\} = \mathcal{F}$$

Logo, como conseguimos exibir um $B \in \mathcal{B}$ tal que $X^{-1}(B)$ não está em nossa $\mathcal{F} = \{\Omega, \emptyset, \{1,2\}, \{3\}\}$, concluímos que X não é mensurável de $(\Omega, \mathcal{F})$ à $(\mathbb{R}, \mathcal{B})$ e, portanto, não é uma variável aleatória.

Questão 4. Dadas duas classes de eventos $\mathcal{C}_1$ e $\mathcal{C}_2$ independentes entre si discutimos em aula um teorema que estabelece uma condição extra sob a qual podemos concluir que $\sigma(\mathcal{C}_1)$ e $\sigma(\mathcal{C}_2)$ são independentes.

Seja $\Omega = \{1,2,3,4\}$, $\mathcal{F}$ = classe de todos os subconjuntos de Ω e P tal que $P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\})$.

a) Mostre que as duas classes de eventos $\mathcal{C}_1 = \{\{1,2\}\}$ e $\mathcal{C}_2 = \{\{2,3\}, \{2,4\}\}$ são independentes mas que $\sigma(\mathcal{C}_1)$ e $\sigma(\mathcal{C}_2)$ não são independentes

Maicon Aparecido Pinheiro - 8637401

Para falar da independência entre as classes C_1 e C_2 e da não-independências entre $\sigma(C_1)$ e $\sigma(C_2)$, precisamos definir as probabilidades dos eventos que estarão envolvidos nas constatações.

Do enunciado, sabemos que

$$\Omega = \{1,2,3,4\} = \bigcup_{i=1}^{4} \{i\} \text{ . e que} \qquad (1)$$

P - a medida de probabilidade definida em $\mathcal{F} = \mathcal{P}(\Omega)$ - é tal que

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = a \qquad (2)$$

Como (1) nos mostra uma representação de Ω como uma união disjunta de evento, segue pela aditividade de P (decorrente de sua σ-aditividade) e de (2) que

$$P(\Omega) = \sum_{i=1}^{4} P(\{i\}) = \sum_{i=1}^{4} a = 4a$$

Como $P(\Omega) = 1$, temos que $a = P(\{i\}) = 1/4$, para $i = 1, 2, 3, 4$.

Daí, temos que se

- $A \in \mathcal{F}$ é tal que $A = \{i, j\}$, $i \neq j$, $i, j \in \{1,2,3,4\}$, $P(A) = 1/4 + 1/4 = 1/2$; e
- $A \in \mathcal{F}$ é tal que $A = \{i, j, k\}$, $i \neq j \neq k$, $i, j, k \in \{1,2,3,4\}$, $P(A) = 1/4 + 1/4 + 1/4 = 3/4$

Com isso, podemos seguir para a resolução do item a.

Por definição C_1 e C_2 serão independentes se

$$P(A_1 \cap A_2) = P(A_1) P(A_2) \qquad (3)$$

para todo $A_1 \in C_1$ e $A_2 \in C_2$.

Como $C_1 = \{\{1,2\}\}$ e $C_2 = \{\{2,3\},\{2,4\}\}$, temos apenas duas situa-ções para verificar. Quando

- $A_1 = \{1,2\}$, $A_2 = \{2,3\}$;

$$P(A_1 \cap A_2) = P(\{2\}) = 1/4 \quad \therefore \quad P(A_1 \cap A_2) = P(A_1)P(A_2)$$
$$P(A_1) \cdot P(A_2) = \frac{1}{2} \cdot \frac{1}{2} = 1/4$$

- $A_1 = \{1,2\}$, $A_2 = \{2,4\}$;

$$P(A_1 \cap A_2) = P(\{2\}) = 1/4 \quad \therefore \quad P(A_1 \cap A_2) = P(A_1)P(A_2)$$
$$P(A_1)P(A_2) = 1/2 \cdot 1/2 = 1/4$$

Logo, como vale (3) para todo $A_1 \in C_1$ e $A_2 \in C_2$, segue que as classes de eventos C_1 e C_2 são independentes.

Agora sejam $\sigma(C_1)$ e $\sigma(C_2)$ as menores σ-álgebras contendo C_1 e C_2, respectivamente. Então

$$\sigma(C_1) = \{\emptyset, \Omega, \{1,2\}, \{3,4\}\} \text{ e}$$

$$\sigma(C_2) = \{\emptyset, \Omega, \{2,3\}, \{1,4\}, \{2,4\}, \{1,3\}, \{1,2,3\}, \{4\}, \{1,3,4\}, \{2\}, \{2,3,4\}, \{1\}, \{1,2,4\}, \{3\}, \{1,2\}, \{3,4\}\} = \mathcal{P}(\Omega)$$

Teremos que $\sigma(C_1)$ e $\sigma(C_2)$ seriam independentes se

$$P(A_1 \cap A_2) = P(A_1)P(A_2) \qquad (*)$$

para todo $A \in \sigma(C_1)$ e $A_2 \in \sigma(C_2)$. Mas note que a tomarmos

$$A_1 = \{1,2\} \quad \text{e} \quad A_2 = \{2\},$$
$$P(A_1 \cap A_2) = P(\{2\}) = 1/4 \quad \text{, i.e. } P(A_1 \cap A_2) \neq P(A_1)P(A_2)$$
$$P(A_1)P(A_2) = 1/2 \cdot 1/4 = 1/8$$

Como $(*)$ não vale para todos, segue que $\sigma(C_1)$ e $\sigma(C_2)$ não são independentes.

Maicon Aparecido Pinheiro - 8637401

b) Identifique a condição extra que não é verificada neste exemplo de C_1 e C_2.

O Teorema visto em aula era o seguinte:

Se $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ é uma coleção de classe de conjuntos independentes e cada $\mathcal{A}_i$ é um sistema-π*, então

$$\sigma(\mathcal{A}_1), \dots, \sigma(\mathcal{A}_n) \text{ é independente}$$

* Uma classe de conjuntos $\mathcal{P}$ é dita um sistema-π se for fechada por intersecções.

A condição extra é a de que C_1 e C_2 deveriam ser sistemas-π, i.e., classes de conjuntos fechadas por intersecções. Mas isso não ocorre com C_2 pois dados $\{2,3\}$ e $\{2,4\}$ em C_2, $\{2\} = \{2,3\} \cap \{2,4\}$ não está em C_2, o que mostra que tal classe não é um sistema π, o que implica na não validade da condição extra: de que ambas sejam sistemas π.

Questão 5.

Seja, para cada $n \geq 1$, $X_n = \{X_{n,1}, X_{n,2}, \dots, X_{n,n}\}$ uma coleção de variáveis aleatórias tal que $X_{n,j}$ tem distribuição Bernoulli(p), com $p = 1/n$ (ou seja, $P(X_{n,j} = 1) = 1 - P(X_{n,j} = 0) = 1/n$) para $1 \leq j \leq n$. Considere S_n o total de sucessos nesses n experimentos Bernoulli: $S_n = \sum_{j=1}^{n} X_{n,j}$.

a) Se $X_{n,1}, X_{n,2}, \dots, X_{n,n}$ são independentes, mostre que

$$P(S_n = k) \to \frac{e^{-1}}{k!} \quad \text{(distribuição de Poisson com média 1)}$$

Como $X_{n,1}, X_{n,2}, \dots, X_{n,n}$ são independentes e identicamentes distribuídas com distribuição Bernoulli ($1/n$), segue que

$$S_n = \sum_{j=1}^{n} X_{n,j} \sim \text{Binomial}(n, 1/n) \qquad (1)$$

isto é,

$$P(S_n = k) = \binom{n}{k}\left(\frac{1}{n}\right)^{k}\left(1-\frac{1}{n}\right)^{n-k}, \quad k = 0, 1, \dots, n$$

Como (1) vale para cada n, segue que quando $n \to \infty$,

$$P(S_n = k) = \lim_{n\to\infty} \binom{n}{k}\left(\frac{1}{n}\right)^{k}\left(1-\frac{1}{n}\right)^{n-k} =$$

$$= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \frac{1}{n^k} \left(\frac{n}{n-1}\right)^{k} \left(1-\frac{1}{n}\right)^{n} =$$

$$= \frac{1}{k!}\left[\lim_{n\to\infty} \underbrace{\frac{n}{n}\cdot\frac{(n-1)}{n}\cdots\frac{(n-k+1)}{n}}_{k \text{ termos}}\frac{\cancel{(n-k)!}}{(n-k)!} \cdot \lim_{n\to\infty}\left(\frac{n}{n-1}\right)^{k} \cdot \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{n}\right] =$$

$$= \frac{1}{k!}\left[1 \cdot 1 \cdot e^{-1}\right] =$$

$$= \frac{e^{-1}}{k!}, \text{ pois}$$

$$\lim_{n\to\infty} \frac{n}{n}\frac{(n-1)}{n}\cdots\frac{(n-k+1)}{n} = \lim_{n\to\infty}\frac{(n-1)}{n} \cdots \lim_{n\to\infty}\frac{(n-k+1)}{n} =$$

$$= \lim_{n\to\infty}(1 - 1/n) \cdots \lim_{n\to\infty}\left(1-\frac{(k-1)}{n}\right) = 1 \cdots 1 = 1$$

$$\lim_{n\to\infty}\left(\frac{n}{n-1}\right)^{k} = \underbrace{1 \cdots 1}_{k \text{ vezes}} = 1^k = 1$$

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{n} = e^{-1} \quad \text{(limite fundamental do cálculo)}$$

Maicon Aparecido Pinheiro - 8637401

Ou seja

$$P(S_n = k) \to \frac{e^{-1}}{k!}, \text{ quando } n \to \infty$$

PS: Demonstração da validade de (1).

Sejam $X_1, X_2, \dots$ variáveis aleatórias independentes e identicamente distribuídas com distribuição Bernoulli (p). Então $Y_n = \sum_{k=1}^{n} X_k$ é uma binomial com parâmetros n e p.

Vamos averiguar isso por indução. Note que para $n=1$, $Y_n = Y_1 = X_1$ e,

$$P(Y_1 = k) = P(X_1 = k) = p^k (1-p)^{1-k} = \binom{1}{k} p^k (1-p)^{1-k}, k=0,1.$$

ou seja, $Y_1 \sim$ Binomial$(1,p)$. (Vale p/ $n=1$)

Suponhamos agora que vale $Y_n \sim$ Binomial(n,p). Vamos averiguar para $k=n+1$. Como $Y_{n+1} = Y_n + X_{n+1}$, temos que Y_{n+1} é a soma de duas binomiais independentes - $Y_n \sim$ Binomial(n,p) e $X_{n+1} \sim$ Binomial$(1,p)$. Daí, segue que $Y_{n+1} \sim$ Binomial$(n+1,p)$. Com isso provamos que a soma de bernoullis i.i.d. é uma binomial.

O único problema é que usamos o fato de que se $X \sim$ Binomial(n,p) e $Y \sim$ Binomial(m,p) são independentes então $X+Y \sim$ Binomial$(n+m,p)$ na passagem "se vale para $k=n$ então vale para $k=n+1$" no parágrafo anterior. Para resolvê-lo, vamos mostrar a validade de tal resultado. Basta notar que

$$\{X+Y=z\} = \bigcup_{k=0}^{n \wedge m} \{X=k, Y=z-k\} \quad \text{(união de eventos disjuntos)}$$

onde $n \wedge m = \min(n,m)$. Daí,

$$P(X+Y=z) = \sum_{k=0}^{n\wedge m} P(X=k, Y=z-k) \overset{\text{ind. de } X \text{ e } Y}{=} \sum_{k=0}^{n \wedge m} P(X=k)P(Y=z-k) =$$

$$= \sum_{k=0}^{n\wedge m} \binom{n}{k} p^{k} (1-p)^{n-k} \binom{m}{z-k} p^{z-k} (1-p)^{m-z+k} =$$

$$= p^{z} (1-p)^{n+m-z} \sum_{k=0}^{n\wedge m} \binom{n}{k} \binom{m}{z-k}$$

Mas dos resultados de combinatória, temos que

$$\sum_{k=0}^{n} \binom{n}{k}\binom{m}{z-k} = \sum_{k=0}^{m} \binom{n}{k}\binom{m}{z-k} = \binom{n+m}{z}$$

Daí,

$$P(X+Y=z) = \binom{n+m}{z} p^{z} (1-p)^{n+m-z}, \quad z = 0, 1, 2, \ldots, n+m.$$

ou seja, $X+Y \sim$ Binomial $(n+m, p)$.

b) Construa agora $X_n = \{X_{n,1}, X_{n,2}, \ldots, X_{n,n}\}$ da seguinte forma:

Para cada $n \geq 1$ conside o espaço de probabilidade $(\Omega_n, \mathcal{F}_n, P_n)$ com $\Omega_n = \{1, 2, \ldots, n\}^n$, $\mathcal{F}_n$ = classe de todos os subconjuntos de Ω_n. Para definir $P_n(\omega)$, probabilidade de selecionar $\omega = (\omega_1, \omega_2, \ldots, \omega_n) \in \Omega_n$ $(\omega_j \in \{1, 2, \ldots, n\}$ $1 \leq j \leq n)$ temos o seguinte procedimento: com probabilidade $1/n$ seleciono aleatoriamente em $\{(1,1,\ldots,1), (2,2,\ldots,2), \ldots, (n,n,\ldots,n)\} \subset \Omega_n$; com probabilidade $\frac{n-1}{n}$ seleciono aleatoriamente uma das $n!$ permutações de $(1, 2, \ldots, n)$.

Construa agora $X_n = \{X_{n,1}, X_{n,2}, \ldots, X_{n,n}\}$, tal que

$$X_{n,j}(\omega) = \begin{cases} 1 & \text{se } \omega_j = 1 \\ 0 & \text{c.c.} \end{cases}$$

Maicon Aparecido Pinheiro _ 8637401

i) Mostre que $X_{n,j}$ tem distribuição Bernoulli (p), com $p = 1/n$

Como visto no enunciado, para cada n, Ω_n é o espaço contendo todos os vetores de dimensão n, com cada uma de suas componentes tomando valores em $\{1, 2, 3, \dots, n\}$, i.e., $\Omega_n = \{1, 2, \dots, n\}^n$; $\mathcal{F}_n$ é a classe de subconjuntos de Ω_n e P_n é tal que

$$P(I_n) = \frac{1}{n}, \quad P(D_n) = \frac{n-1}{n}, \quad P(R_n) = 0, \text{ onde}$$

$I_n = \{(1, 1, \dots, 1), (2, 2, \dots, 2), \dots, (n, \dots, n)\}$ = conjunto dos vetores que possuem valores idênticos nas componentes

D_n = conjunto com todos o $n!$ vetores que tem todas as componentes assumindo valores distintos um dos outros: $(1, 2, \dots, n)$; $(2, 1, \dots, n)$, ..

$R_n = \Omega_n - (I_n \cup D_n)$ = o restante dos vetores.

Daí, dado esse espaço de probabilidade $(\Omega_n, \mathcal{F}_n, P_n)$, a variável aleatória

$$X_{n,j}(\omega) = \begin{cases} 1, & \text{se } \omega_j = 1 \\ 0, & \text{c.c.} \end{cases}$$

é, em palavras, uma variável aleatoria que indica o fato da j-ésima componente do vetor ω ser 1 ou diferente do mesmo. Daí, segue que

$$P(X_{n,j} = 1) = P(\{\omega : X_{n,j}(\omega) = 1\}) =$$
$$= P[(\{\omega : X_{n,j}(\omega) = 1\} \cap I_n) \cup (\{\omega : X_{n,j}(\omega) = 1\} \cap D_n) \cup (\{\omega : X_{n,j}(\omega) = 1\} \cap R_n)$$

são eventos disjuntos e vetores em D_n não são selecionados

$$= P(\{\omega : X_{n,j}(\omega) = 1\} \cap I_n) + P(\{\omega : X_{n,j}(\omega) = 1\} \cap D_n) =$$

seleção dentro de I_n e D_n é feita de forma aleatória

$$= \frac{1}{n} \cdot \frac{1}{n} + \frac{(n-1)!}{n!} \frac{(n-1)}{n} = \frac{1}{n^2} + \frac{n-1}{n^2} = \frac{n}{n^2} = \frac{1}{n}$$

Como $P(X_{n,j} = 0) = 1 - P(X_{n,j} = 1)$, segue $P(X_{n,j} = 0) = 1 - 1/n$.
Ou seja, em termos gerais

$$P(X_{n,j} = k) = \left(\frac{1}{n}\right)^{k} \left(1 - \frac{1}{n}\right)^{1-k}, \quad k = 0, 1.$$

isto é, $X_{n,j} \sim \text{Bernoulli}(1/n)$ para cada n.

ii) Mostre que $X_{n,j}$ e $X_{n,i}$ são independentes, para $i \neq j$
Pelo item anterior, podemos constatar que

$$X_{n,j} \sim \text{Bernoulli}(1/n) \quad e \quad X_{n,i} \sim \text{Bernoulli}(1/n), \quad i \neq j.$$

Nesse caso, mostrar que $E[X_{n,j} X_{n,i}] = E[X_{n,j}]\, E[X_{n,i}]$, $i \neq j$, implica que $X_{n,j}$ e $X_{n,i}$ são independentes.
Do lado esquerdo, temos que

$$E[X_{n,j}]\, E[X_{n,i}] = \frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n^2}$$

Agora, por outro lado, devemos notar que

$$X_{n,j} X_{n,i} = \begin{cases} 1, & \text{se } \omega_j = \omega_i = 1 \\ 0, & \text{c.c.} \end{cases}$$

Daí,

[illegible]

$$P(X_{n,j} X_{n,i} = 1) = P(\{\omega : X_{n,j} X_{n,i}(\omega) = 1\} \cap I_n) + \overbrace{P(\{\omega : X_{n,j} X_{n,i}(\omega) = 1\} \cap D_n)}^{0}$$

$$= \frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n^2}$$

já que em D_n não ocorrem valores tomando um mesmo valor

Maicon Aparecido Pinheiro - 8637401

em mais de uma componente. Logo,

$$X_{n,j} X_{n,i} \sim \text{Bernoulli}(1/n^2), \quad i \neq j$$

Daí, $E[X_{n,j} X_{n,i}] = 1/n^2 = E[X_{n,j}] E[X_{n,i}]$; ou seja, $X_{n,j}$ e $X_{n,i}$, para $i \neq j$, são independentes.

Nota:

Seja $X \sim \text{Bernoulli}(p)$ e $Y \sim \text{Bernoulli}(p)$. Note que

$$P(X=1, Y=1) = P(X=1)P(Y=1) \Leftrightarrow$$
$$E(XY) = E(X).E(Y)$$

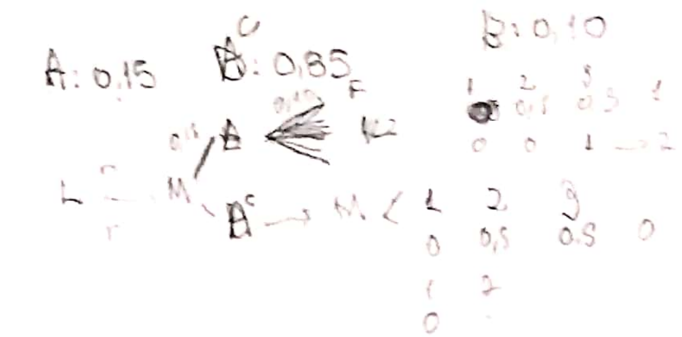
Daí, com $P(X=1, Y=1) = P(X=1)P(Y=1)$, é fácil obter por complementação que todas as demais probabilidades conjuntas podem ser fatoradas no produto das marginais envolvidas.

iii) Mostre que a distribuição de $S_n = \sum_{j=1}^{n} X_{n,j}$ não converge para a Poisson.

Para cada $\omega \in \Omega_n$, podemos entender $S_n(\omega)$ como o número de componentes de ω que assumem o valor 1. Como só há probabilidade positiva de se selecionar vetores em I_n ou em D_n, segue que $S_n(\omega)$ assumirá o valor n se $\omega \in I_n$ e $\omega = (1, 1, \ldots, 1)$; o valor 0 se $\omega \in I_n$ e $\omega \neq (1, 1, \ldots, 1)$; e o valor 1 se $\omega \in D_n$, uma vez que só uma das componentes de ω em D_n é necessariamente igual à 1. Isto é,

$$S_n(\omega) = \begin{cases} 0, & \text{se } \omega \in I_n \text{ e } \omega \neq (1, \ldots, 1); \\ 1, & \text{se } \omega \in D_n; \\ n, & \text{se } \omega \in I_n \text{ e } \omega = (1, \ldots, 1). \end{cases}$$

$A \quad A^c$



Daí, segue que

$$P(S_n=0) = \frac{1}{n}\cdot\frac{(n-1)}{n} = \frac{(n-1)}{n^2} \to 0 \text{ quando } n\uparrow\infty$$

$$P(S_n=1) = \frac{(n-1)}{n} \to 1 \text{ quando } n\uparrow\infty$$

$$P(S_n=n) = \frac{1}{n}\cdot\frac{1}{n} = \frac{1}{n^2} \to 0 \text{ quando } n\uparrow\infty$$

Isto é, S_n converge para uma variável aleatória degenerada, que toma o valor 1 com probabilidade 1, o que não é o caso da Poisson.

iv) Porque $S_n = \sum_{j=1}^{n} X_{n,j}$ não converge para a Poisson mais isto era verdade no caso anterior (item a)?

No item a), para que S_n convergisse para a Poisson, era suposto que $X_{n,1}, \dots, X_{n,n}$ fossem independentes. No caso do item b, não temos isso. No item b), elas são independentes duas a duas, mas não temos a independência entre todas as variáveis envolvidas na soma que compõe S_n.

Note que para $i \neq j \neq k$,

$$P(X_{n,i}=1, X_{n,j}=1, X_{n,k}=1) = \frac{1}{n^2} \quad (\text{só ocorre se } w=(1,1,\dots,1))$$

enquanto que $P(X_{n,i}=1)P(X_{n,j}=1)P(X_{n,k}=1) = 1/n^3$. $X_{n,1}, X_{n,2}, \dots, X_{n,n}$ não são independentes! Essa é a razão!

MAE 5811 - Probabilidade Avançada I- Prova II
Justifique suas respostas.

1. (a) (1,0) Considere a coleção de variáveis aleatórias $\{X_n\}_{n\geq 0}$ com $P(X_n = 1 - \frac{1}{n}) = P(X_n = 1 + \frac{1}{n}) = \frac{1}{2}$. X_n converge em distribuição para alguma v.a.? Qual?

   (b) (1,5) Sejam $X_1, X_2, \ldots$ variáveis aleatórias i.i.d. com $E(X_i) = 0$ e $var(X_i) \in (0, \infty)$. Se $S_n = X_1 + X_2 + \ldots + X_n$, determine $P(\limsup \frac{S_n}{\sqrt{n}} < \infty)$. Justifique (claro).

   (c) (2,0) Seja $\{\mathcal{F}_n\}_{n\geq 1}$ uma filtragem e $X_n = \sum_{m\leq n} 1_{B_m}$, com $B_n \in \mathcal{F}_n$, $n \geq 1$. Mostre que $\{X_n, \mathcal{F}_n\}$ é um submartingal e determine a *Decomposição de Doob* correspondente.

2. (2,5) Seja $\{Z_n\}_{n\geq 1}$ uma coleção de variáveis aleatórias independentes com $P(Z_n = 0) = 1 - \frac{1}{n}$ e $P(Z_n = 1) = P(Z_n = -1) = \frac{1}{2n}$. Defina agora $\{X_n\}_{n\geq 1}$ tal que $X_1 = Z_1$ e, para $n \geq 2$,

$$X_n = \begin{cases} Z_n \text{ se } X_{n-1} = 0 \\ nX_{n-1}|Z_n| \text{ se } X_{n-1} \neq 0 \end{cases}$$

   e defina $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$.

   (a) (1,0) Mostre que $\{X_n, \mathcal{F}_n\}$ é um martingal.

   (b) (0,5) Valem as condições do teorema de convergência de martingais vista em aula?

   (c) (0,5) X_n converge em probabilidade para alguma v.a.? Qual?

   (d) (0,5) X_n converge quase certamente para alguma v.a.? Qual?

3. (1,5) Se $\xi_1, \xi_2, \ldots$ são variáveis aleatórias i.i.d. com $P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}$ defina $S_0 = 1$, $S_n = S_0 + \sum_{k=1}^{n} \xi_k$, $n \geq 1$, e $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$. Se $N = \inf\{n : S_n = 0\}$ e $X_n = S_{n\wedge N}$, X_n converge em L^1? X_n converge quase certamente?

4. (1,5) Sejam $X_1, \ldots X_n$ variáveis aleatórias com distribuições de probabilidade $F_1, \ldots, F_n$ e funções caracteristicas $\psi_1, \ldots, \psi_n$, respectivamente. Seja tambem Z uma variável aleatória assumindo valores em $\{1, \ldots, n\}$ com $P(Z = i) = \lambda_i$. Assuma independência entre todas estas variáveis aleatórias.

   (a) (0,5) Usando $\{X_i\}_{i=1}^n$ e Z, construa uma variável aleatória Y com função distribuição $\sum_{i=1}^n \lambda_i F_i$.

   (b) (1,0) Determine a função característica de uma variável aleatória com distribuição $\sum_{i=1}^n \lambda_i F_i$ em função de $\{\psi_i\}$.

07 de abril

Teoria da Medida , Durret

$(\Omega, \mathcal{F}, P)$: espaço de probabilidade

Ω: conjunto dos resultados

$\mathcal{F}$: conjunto dos eventos

$P: \mathcal{F} \to [0,1]$ uma função que associa probabilidade aos eventos

Assumimos que $\mathcal{F}$ é uma σ-álgebra, i.e., uma coleção não-vazia de subconjuntos de Ω que satisfaz

(i) Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$, e

(ii) Se $A_i \in \mathcal{F}$ é uma sequência enumerável de conjuntos, então $\bigcup_i A_i \in \mathcal{F}$.

* Como $\bigcap_i A_i = (\bigcup_i A_i)^c$, segue que uma σ-álgebra é fechada por intersecções enumeráveis*

$(\Omega, \mathcal{F})$ é chamado espaço mensurável, i.e., um espaço onde podemos colocar uma medida. Uma medida é uma função set não-negativa σ-aditiva, i.e., uma função $\mu: \mathcal{F} \to \mathbb{R}$ com

(i) $\mu(A) \geq \mu(\emptyset) = 0$ para todo $A \in \mathcal{F}$, e

(ii) Se $A_i \in \mathcal{F}$ é uma seq. enumerável de conj. disjuntos, então

$$\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i)$$

Se $\mu(\Omega) = 1$, μ é uma medida de probabilidade, que denotaremos por P.

Algumas consequências do def. de medida.

Teorema 1.1.1. Seja μ um medida em $(\Omega, \mathcal{F})$

1. Monotocidade. Se $A \subset B$ então $\mu(A) \leq \mu(B)$

Seja $B - A = B \cap A^c$ a diferença dos dois conjuntos. Usando + para denotar união disjunta, $B = A + (B-A)$. Assim

$$\mu(B) \overset{(ii)}{=} \mu(A) + \mu(B-A) \overset{(i)}{\geq} \mu(A)$$

ii. Subaditividade. Se $A \subset \bigcup_{m=1}^{\infty} A_m$ então $\mu(A) \leq \sum_{m=1}^{\infty} \mu(A_m)$

Seja $A_n' = A_n \cap A$, $B_1 = A_1'$ e p/ $n > 1$, $B_n = A_n' - \left(\bigcup_{m=1}^{n-1} (A_m')\right)^c$. Como os B_ns são disjuntos e tem união A, temos, usando (i) da def. de medida, $B_m \subset A_m$, e (i) deste teorema

$$\mu(A) = \sum_{m=1}^{\infty} \mu(B_m) \leq \sum_{m=1}^{\infty} \mu(A_m)$$

iii. Continuidade por baixo. Se $A_i \uparrow A$ (i.e., $A_1 \supset A_2 \supset \dots$ e $\cap_i A_i = A$), com $\mu(A_1) < \infty$ então $\mu(A_i) \downarrow \mu(A)$.

$B_n = A_n - A_{n-1}$. B_n são disjuntos e $\bigcup_{m=1}^{\infty} B_m = A$, $\bigcup_{m=1}^{n} B_m = A_n$.

Logo.

$$\mu(A) = \sum_{m=1}^{\infty} \mu(B_m) = \lim_{n \to \infty} \sum_{m=1}^{n} \mu(B_m) = \lim_{n \to \infty} \mu(A_n)$$

iv. Continuidade por cima. Se $A_i \downarrow A$ (i.e., $A_1 \supset A_2 \supset \dots$ e $\cap_i A_i = A$), com $\mu(A_1) < \infty$ então $\mu(A_i) \downarrow \mu(A)$

$A_1 - A_n \uparrow A_1 - A$, Daí, (iii) implica $\mu(A_1 - A_n) \uparrow \mu(A_1 - A)$. Como $A_1 \supset B$, $\mu(A_1 - B) = \mu(A_1) - \mu(B)$ e segue que $\mu(A_n) \downarrow \mu(A)$.

Exemplo 1.1.1. Espaços de probabilidade discretos. Ω um conjunto enumerável. $\mathcal{F}$ o conjunto de todos os subconjuntos de Ω. Seja

$$P(A) = \sum_{\omega \in A} p(\omega) \text{ onde } p(\omega) \geq 0 \text{ e } \sum_{\omega \in \Omega} p(\omega) = 1$$

É a medida mais geral neste espaço. Em muitos casos quando Ω é finito, temos $p(\omega) = 1/|\Omega|$ onde $|\Omega|$ é o nº de pontos em Ω.

Exemplo Concreto. Astrogali, dado feito de osso do tornozelo de uma ovelha usado no antigo Egipto. Poderia sair 4, 3, 6 ou 1 pontos. Logo, precisaríamos de p_1, p_3, p_4 e p_6 para caracterizar P.

Exercício 1.1.1. Intersecção de σ-álgebras é σ-álgebra. $\sigma(\mathcal{A})$ é a menor σ-álgebra contendo $\mathcal{A}$ (ou σ-álgebra gerada por $\mathcal{A}$), onde $\mathcal{A}$ é coleção de subconjuntos de Ω.

Cons. $\mathbb{R}^d$ e $\mathcal{R}^d$ os conjuntos de Borel, a menor σ-álgebra contendo os conjuntos abertos. Qdo $d=1$, omitimos o sobreescrito.

Exemplo 1.1.2. Medida na reta real. Medidas em $(\mathbb{R}, \mathcal{R})$ são def. by giving probability a Stieltjes measure function com as seg. propriedades:

(i) F é não-decrescente.

(ii) F é contínua à direita, i.e., $\lim_{y \downarrow x} F(y) = F(x)$.

Qdo $F(x) = x$ a medida resultante é chamada medida de Lebesgue.

Teorema 1.1.2 Associado com cada função medida de Stieltjes F há uma única medida μ em $(\mathbb{R}, \mathcal{R})$ com $\mu((a,b]) = F(b) - F(a)$

$$\mu((a,b]) = F(b) - F(a) \qquad (1.1.1)$$

A escolha do int. "fechado à direita" em $(a,b]$ é ditada pelo fato de que se $b_n \downarrow b$, então

$$\cap_n (a, b_n] = (a,b]$$

A próx. def. explicará a escolha do "aberto a esquerda".

Uma coleção $\mathcal{S}$ de conjuntos é dita uma semi-álgebra se
(i) é fechada sob intersecção, i.e., $S, T \in \mathcal{S} \Rightarrow S \cap T \in \mathcal{S}$, e
(ii) se $S \in \mathcal{S}$, então S^c é uma união finita de conjuntos disjuntos em $\mathcal{S}$.

Um importante exemplo de semi-álgebra é:

$\mathcal{S}_d$ = cjto vazio mais todos os conjuntos da forma

$(a_1, b_1] \times \dots \times (a_d, b_d] \subset \mathbb{R}^d$ onde $-\infty \leq a_i < b_i \leq \infty$

A def. em (1.1.1) dá os valores de μ na semiálgebra $\mathcal{S}_1$. Pra ir de uma semi-álgebra à uma σ-álgebra, usamos um passo intermediário. Uma coleção $\mathcal{A}$ de subconjuntos de Ω é dita uma álgebra se $A, B \in \mathcal{A} \Rightarrow A^c$ e $A \cup B$ estão em $\mathcal{A}$. Como $A \cap B = (A^c \cup B^c)^c$, segue que $A \cap B \in \mathcal{A}$. É óbvio que uma σ-álgebra é uma álgebra. Um exemplo de que o contrário é falso.

$\Omega = \mathbb{Z}$. $\mathcal{A}$ = coleção dos $A \subset \mathbb{Z}$ de modo que A ou A^c é finito é uma álgebra.

- $A \in \mathcal{A}$ então $\begin{cases} A \text{ é finito} \Rightarrow (A^c)^c = A \text{ é finito} \Rightarrow A^c \in \mathcal{A} \\ A^c \text{ é finito} \Rightarrow A^c \in \mathcal{A} \end{cases}$

- Se $A, B \in \mathcal{A}$, então

  $A \cup B$ é finita, A e B finitos $\Rightarrow A \cup B \in \mathcal{A}$

  A^c ou B^c é finito $\Rightarrow (A^c \cap B^c)$ é finito $\Rightarrow (A \cup B)^c$ é finito $\Rightarrow A \cup B \in \mathcal{A}$

  A e B não podem ser infinitos, a não ser que sejam idênticos.

1.2 Distribuições

Espaços de probabilidade se tornam ainda mais interessantes quando definimos uma variável aleatória nos mesmos.

Uma função real X definida em Ω é dita uma variável aleatória se para todo $B \subset \mathbb{R}$, temos

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$$

Quando for necessário enfatizar a σ-álgebra, dizemos que X é $\mathcal{F}$-mensurável ou escrevemos $X \in \mathcal{F}$.

- Um exemplo trivial, mas útil, de v.a. é a função indicadora de um conjunto $A \in \mathcal{F}$:

$$\mathbb{1}_A = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

Se X é uma variável aleatória, então X induz uma medida em $\mathbb{R}$ chamada ~~de~~ sua distribuição ao declarar $\mu(A) = P(X \in A)$ para conjuntos de Borel A. Usando a notação introduzida acima, o lado direito pode se escrito como $P(X^{-1}(A))$. Em palavras, we pull $A \in \mathcal{R}$ back to $X^{-1}(A) \in \mathcal{F}$ and then take P of that set.

Para verificar que μ é uma medida de probabilidade observamos que se os A_i são disjuntos, então usando a def. de μ; o fato de que X está na união se e só se está em um dos A_i; o fato de que se os $A_i \in \mathcal{R}$ são disjuntos então os eventos $\{X \in A_i\}$ são disjuntos; e a def. de μ outra vez, temos:

$$\mu(\cup_i A_i) \overset{1}{=} P(X \in \cup_i A_i) \overset{2}{=} P(\cup_i \{X \in A_i\}) \overset{3}{=} \sum_i P(X \in A_i) \overset{4}{=} \sum_i \mu(A_i)$$

1. Definição de $\mu(\cdot)$

2. Note que

$\omega \in X^{-1}(\cup_i A_i) \Leftrightarrow X(\omega) \in \bigcup_i A_i \Leftrightarrow X(\omega) \in A_i$ para algum i $\Leftrightarrow \omega \in X^{-1}(A_i)$ para algum $i \Leftrightarrow \omega \in \bigcup_i X^{-1}(A_i) \Leftrightarrow$ $\omega \in \cup_i \{X \in A_i\}$. Ou seja

$$X^{-1}(\bigcup_i A_i) = \bigcup_i X^{-1}(A_i)$$

Logo,

$$P(X^{-1}(\bigcup_i A_i)) = P(X \in \bigcup_i A_i) = P(\bigcup_i X^{-1}(A_i)) = P(\bigcup_i \{X \in A_i\})$$

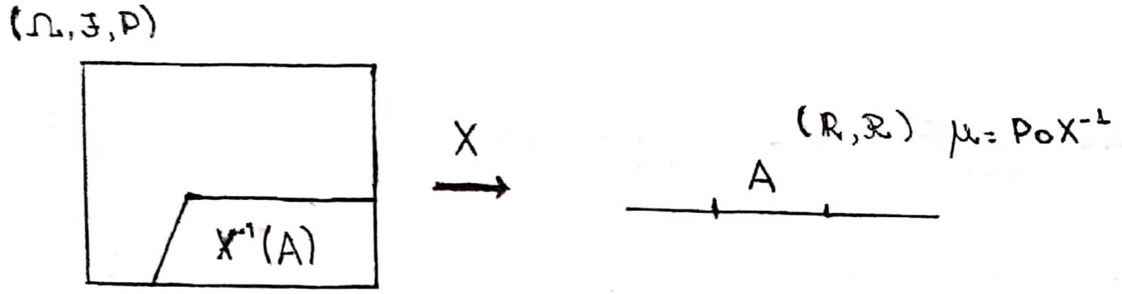
3. Dados A_{i_1} e A_{i_2}, em $\mathcal{R}$, disjuntos,

$$\{X \in A_{i_1}\} = \{\omega \in \Omega : X(\omega) \in A_{i_1}\} \text{ e } \{X \in A_{i_2}\} = \{\omega \in \Omega : X(\omega) \in A_{i_2}\}$$

Se $\{X \in A_{i_1}\} \cap \{X \in A_{i_2}\} \neq \emptyset$, então existiria ω^* pertencendo a ambos os conjuntos, fazendo com que $X(\omega^*)$ toma-se valores em A_{i_1} e A_{i_2}. Como $A_{i_1} \cap A_{i_2} = \emptyset$, isto é equivalente a $X(\omega^*)$ tomar dois valores distintos, o que é um absurdo, pois $X(\cdot)$ é uma função. Logo, pela propriedade aditiva de P,

$$P(\bigcup_i \{X \in A_i\}) = \sum_i P(X \in A_i)$$

4. Def. de μ. Para $A \in \mathcal{R}$, $\mu(A) = P(X \in A)$



$P : \mathcal{F} \to [0,1]$ Assim

$X^{-1} : \mathcal{R} \to \mathcal{F}$

$$\begin{tikzcd} \mathcal{R} \arrow[r, "X^{-1}"] \arrow[rr, bend right, "\mu = P \circ X^{-1}"'] & \mathcal{F} \arrow[r, "P"] & {[0,1]} \end{tikzcd}$$

A distribuição de uma v.a. X é usualmente descrita dando-se sua função distribuição; $F(x) = P(X \leq x) = P(X \in (-\infty, x])$

Teorema 1.2.1. Qualquer função distribuição tem as seguintes propriedades:

i) F é não-decrescente

De fato, para $x \leq y$, x, y reais,

$$\{X \leq x\} \subset \{X \leq y\} \qquad (1)$$

pois se $\omega \in \{X \leq x\}$, então $X(\omega) \leq x \Rightarrow X(\omega) \leq y \Rightarrow$ $\omega \in \{X \leq y\}$. Logo, vale (1).

Daí, pela monotonicidade da medida P, segue que

$$P(X \leq x) \leq P(X \leq y), \quad x \leq y.$$

ii) $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$

• $\lim_{x \to \infty} F(x) = 1$.

Se $x \uparrow \infty$, então $\{X \leq x\} \uparrow \Omega$. Como sabemos

$$\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\}.$$

Qdo $x \uparrow \infty$, $\forall \omega \in \Omega$, temos $X(\omega) \leq \infty$, e como $\{X \leq x\}$ é uma seq. crescente de eventos quando x cresce, segue

$$\{X \leq x\} \uparrow \Omega$$

Analogamente, segue que $\lim_{x \to -\infty} F(x) = P(\emptyset) \overset{P1.}{=} 0$.

iii) F é contínua à direita, isto é,

$$\lim_{y \downarrow x} F(y) = F(x).$$

Para ver que quando $y \downarrow x$, $\{X \leq y\} \downarrow \{X \leq x\}$.

Daí, pela cont. da prob. por seq. decrescentes, segue que

$$\lim_{y \downarrow x} P(\{X \leq y\}) = P(\lim_{y \downarrow x} \{X \leq y\}) \Leftrightarrow$$

$$\lim_{y \downarrow x} F(y) = P(X \leq x)$$

(iv) Se $F(x-) = \lim_{y\uparrow x} F(y)$ então $F(x-) = P(X < x)$

Quando $y \uparrow x$, temos

$$\{X \le y\} \uparrow \{X < x\}$$

Pois $X(\omega) \le \lim_{y\uparrow x} y \Leftrightarrow X(\omega) < x$

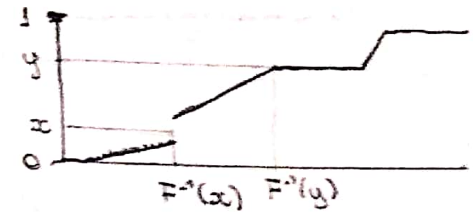
(v) $P(X=x) = F(x) - F(x-)$

Note que $P(X=x) = P(X\le x) - P(X<x)$.

Teorema 1.2.2. Se F satisfaz (i), (ii) e (iii) no Teorema 1.2.1, então esta é a função distribuição de alguma variável aleatória.

Prova. Seja $\Omega = (0,1)$, $\mathcal{F}$ = os conjuntos de Borel, e P a medida de Lebesgue. Se $\omega \in (0,1)$, seja

$$X(\omega) = \sup\{y : F(y) < \omega\}$$



Uma vez que mostrarmos que

$$\{\omega : X(\omega) \le x\} = \{\omega : \omega \le F(x)\} \qquad (*)$$

o resultado desejado segue imediatamente uma vez que

$P(\omega : \omega \le F(x)) = F(x)$ (P é a med de Lebesgue)

Para checar (*), observamos que* $\omega \le F(x)$ então $X(\omega) \le x$, já que

$x \notin \{y : F(y) < \omega\}$. Por outro lado, se $\omega > F(x)$, então uma vez que F é cont. à direita, há um $\varepsilon > 0$ de modo que $F(x+\varepsilon) < \omega$ e $X(\omega) \geq x+\varepsilon > x$.

$\blacksquare$

Repetindo. Basta mostrar que $\{\omega : X(\omega) \leq x\} = \{\omega : \omega \leq F(x)\}$

Se $\omega \leq F(x)$ então $X(\omega) \leq x$ pois $x \notin \sup\{y : F(y) < \omega\}$.
Se $\omega > F(x)$; por F ser contínua à direita; $\exists\, \varepsilon > 0$ t.q.

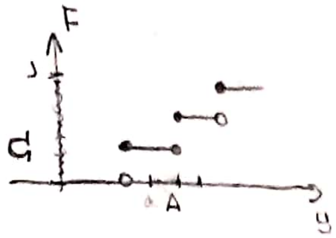
$$F(x+\varepsilon) < \omega \quad \text{e} \quad X(\omega) \geq x+\varepsilon > x.$$

Isto é, $\exists$ um cara maior que x e próximo o suficiente com a F nele sendo a F em $x \Rightarrow F(x+\varepsilon) = F(x) < \omega$ Mas nessas condições $X(\omega) \geq x+\varepsilon > x$

Logo, se $\omega \leq F(x)$, $X(\omega) \leq x$. $\Rightarrow \{\omega\ X(\omega) \leq x\} = \{\omega . \omega \leq F(x)\}$,
se $\omega > F(x)$, $X(\omega) > x$

Mesmo que F possa não ser um-a-um

Visão geométrica.



F satisfaz i, ii e iii.

Tomando $\Omega = (0,1)$, $\mathcal{F}$: conjuntos de Borel

P: medida de Leb.

$X(\omega) = \sup\{y : F(y) < \omega\}$, $y \in supp(F)$
é v.a.

$X^{-1}(A) = \{\omega \in [0,1] : X(\omega) \in A\} = (\omega_1, 1)$ tal que $\omega_1 > \max\limits_{X(\omega) \in A} (F(X(\omega)))$

Embora F possa não ser um-a-um e sobrejetora, chamaremos X de a inversa de F e denotamos-a por F^{-1}. O esquema na prova do Teorema 1.2.2 é útil em gerar variáveis aleatórias em um computador.

Se X e Y induzem a mesma distribuição μ em $(\mathbb{R}, \mathcal{R})$, dizemos que X e Y são iguais em distribuição. Em vista do Teorema 1.1.2, isto vale se e somente se X e Y tem a mesma função distribuição, isto é, $P(X \leq x) = P(Y \leq x)$ para todo x. Quando X e Y tem a mesma distribuição, escrevemos

$$X \stackrel{d}{=} Y$$

Quando a função distribuição $F(x) = P(X \leq x)$ tem a forma

$$F(x) = \int_{-\infty}^{x} f(y)\, dy \qquad (1.2.1)$$

dizemos que X tem função densidade f. É frequentemente útil pensar em $f(x)$ como $P(X = x)$ embora

$$P(X = x) = \lim_{\varepsilon \to 0} \int_{x-\varepsilon}^{x+\varepsilon} f(y)\, dy = 0$$

- Dada f, determinamos F
- Alguns exemplos úteis: uniforme, exponencial, normal padrão
- Distribuição Uniforme no conjunto de Cantor
- Uma medida de prob. P é dita discreta se há um conjunto enumerável S tal que $P(S^c) = 0$. (ou sua f.d. associada)

- Exemplo de f.d. discreta.
  - Degenerada em 0
  - Descontinuidades densas Se $q_1, q_2, \dots$ uma enumeração dos racionais.

    Seja $\alpha_i > 0$ tendo $\sum_{i=1}^{\infty} \alpha_i = 1$ e seja

$$F(x) = \sum_{i=1}^{\infty} \alpha_i \mathbb{1}_{[q_i, \infty)}$$

Legal: A F é como uma escada que ganha um degrau de altura α_i em q_i

1.3. Variáveis Aleatórias

- Desenvolver alguns resultados que nos ajudarão depois a provar que qtdes que definimos são v.a., i.e., são mensuráveis.

First we need a definition. Uma função $X : \Omega \to S$ é dita uma função mensurável de $(\Omega, \mathcal{F})$ à $(S, \mathcal{S})$ se

$$X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F} \quad \text{para todo } B \in \mathcal{S}$$

- Se $(S, \mathcal{S}) = (\mathbb{R}^d, \mathcal{R}^d)$ e $d > 1$, então X é chamado **vetor aleatório**. Claro, se $d = 1$, X é dita uma **variável aleatória**, ou v.a. como sigla.

O próximo resultado é útil para provar que 'maps' são mensuráveis.

Teorema 1.3.1. Se $\{\omega : X(\omega) \in A\} \in \mathcal{F}$ para todo $A \in \mathcal{A}$ e $\mathcal{A}$ gera $\mathcal{S}$ (i.e., $\mathcal{S}$ é a menor σ-álgebra que contém $\mathcal{A}$), então X é mensurável.

Prova. Escrevendo $\{X \in B\}$ para $\{\omega : X(\omega) \in B\}$, temos

$$\{X \in \cup_i B_i\} = \cup_i \{X \in B_i\}$$
$$\{X \in B^c\} = \{X \in B\}^c$$

de modo que a classe $\mathcal{B} = \{B : \{X \in B\} \in \mathcal{F}\}$ é uma σ-álgebra. Uma vez que $\mathcal{B} \supset \mathcal{A}$ e $\mathcal{A}$ gera $\mathcal{S}$, $\mathcal{B} \supset \mathcal{S}$.

Idéia 1. Pego uma geral, de interesse. $\mathcal{B} = \{B : \{X \in B\} \in \mathcal{F}\}$
2. Mostro que $\mathcal{B}$ é σ-álgebra
3. Mostro que $\mathcal{A} \subset \mathcal{B} \Rightarrow \mathcal{S} = \sigma(\mathcal{A}) \subset \mathcal{B}$

Segue das duas eq. dispostas anteriormente na prova que se $\mathcal{A}$ é uma σ-álgebra, então $\{\{X \in B\} : B \in \mathcal{A}\}$ é uma σ-álgebra. Esta é a menor σ-álgebra em Ω que faz X uma função mensurável. É chamada a σ-álgebra gerada por X e é denotada por $\sigma(X)$. Para referências futuras, notamos que

$$\sigma(X) = \{\{X \in B\} : B \in \mathcal{A}\}$$

Exemplo 1.3.1. Se $(S, \mathcal{A}) = (\mathbb{R}, \mathcal{R})$, então as possíveis escolhas de $\mathcal{A}$ no Teo. 1.3.1 são $\{(-\infty, x] : x \in \mathbb{R}\}$ ou $\{(-\infty, x), x \in \mathbb{Q}\}$.

Exemplo 1.3.2. Se $(S, \mathcal{A}) = (\mathbb{R}^d, \mathcal{R}^d)$, uma escolha útil de $\mathcal{A}$ é

$$\{(a_1, b_1) \times \dots \times (a_d, b_d) : -\infty < a_i < b_i < \infty\}$$

ou ocasionalmente a maior coleção de conjuntos abertos.

Teorema 1.3.2. Se $X : (\Omega, \mathcal{F}) \to (S, \mathcal{A})$ e $f : (S, \mathcal{A}) \to (T, \mathcal{J})$, são mapas mensuráveis, então $f(X)$ é um mapa mensurável de $(\Omega, \mathcal{F})$ à $(T, \mathcal{J})$.

Prova.

Seja $B \in \mathcal{J}$. $\{\omega : f(X(\omega)) \in B\} = \{\omega : X(\omega) \in f^{-1}(B)\} \in \mathcal{F}$, uma vez que por suposição $f^{-1}(B) \in \mathcal{A}$

- Do Teo. 1.3.2 segue imediat. que se X é uma v.a. então cX $\forall c \in \mathbb{R}$, X^2, $\sin(X)$, tbm o são.

Teorema 1.3.3. Se $X_1, \dots, X_n$ são v.a. e $f : (\mathbb{R}^n, \mathcal{R}^n) \to (\mathbb{R}, \mathcal{R})$ é mensurável, então $f(X_1, \dots, X_n)$ é uma variável aleatória.

Prova. Em vista do Teorema 1.3.2, é suficiente mostrar que $(X_1, \dots, X_n)$ é um vetor aleatório. Para isto, observamos que se $A_1, \dots, A_n$ são conjuntos de Borel então

$$\{(X_1, \dots, X_n) \in A_1 \times \dots \times A_n\} = \cap_i \{X_i \in A_i\} \in \mathcal{F}$$

Uma vez que os conjuntos $A_1 \times \dots \times A_n$ geram $\mathcal{R}^n$, o resultado desejado segue do Teorema 1.3.1.

Teorema 1.3.4. Se $X_1, \dots, X_n$ são variáveis aleatórias então $X_1 + X_2 + \dots + X_n$ é uma variável aleatória.

Prova. Pelo Teo. 1.3.3, é suf. mostrar que $f(x_1, \dots, x_n) = x_1 + \dots + x_n$ é mensurável. Para tal, usamos o exemplo 1.3.1 e notamos que $\{x : x_1 + \dots + x_n < a\}$ é um conjunto aberto e portanto está em $\mathcal{R}^n$.

Teorema 1.3.5. Se $X_1, X_2, \dots$ são v.a. então também o são

$$\inf_n X_n \qquad \sup_n X_n \qquad \limsup_n X_n \qquad \liminf_n X_n$$

Prova. Uma vez que o ínfimo de uma sequência é menor que a se e somente se alguns termos são menores que a (se todos são maiores que a, então tbm será o ínfimo), temos

$$\{\inf_n X_n < a\} = \cup_n \{X_n < a\} \in \mathcal{F}$$

Um argumento similar mostra que $\{\sup_n X_n > a\} = \cup_n \{X_n > a\} \in \mathcal{F}$.

Para os dois últimos, observamos

$$\liminf_{n\to\infty} X_n = \sup_n \left(\inf_{m\geq n} X_m\right)$$

$$\limsup_{n\to\infty} X_n = \inf_n \left(\sup_{m\geq n} X_m\right)$$

Para completar a prova no primeira, note que $Y_n = \inf_{m\geq n} X_m$

é uma variável aleatória para cada n, assim $\sup_{n} Y_n$ tbm o é.

Do Teorema 1.3.5, vemos que

$$\Omega_0 \equiv \{\omega : \lim_{n\to\infty} X_n \text{ existe}\} = \{\omega : \limsup_{n\to\infty} X_n - \liminf_{n\to\infty} X_n = 0\}$$

é um conjunto mensurável. (Aqui $\equiv$ indica que a primeira igualdade é uma definição.) Se $P(\Omega_0) = 1$, dizemos que X_n converge quase certamente, ou a.s (q.c). Este tipo de conv. é chamada quase todo lugar em teoria da medida. Para ter um limite def. no espaço inteiro, é conveniente deixar

$$X_\infty = \limsup_{n\to\infty} X_n$$

mas esta variável aleatória pode tomar valores $+\infty$ ou $-\infty$. Para acomodar esta e outras dores de cabeça, generelazimos a def. de v.a.

Uma função cujo domínio é um conjunto $D \in \mathcal{F}$ e cuja imagem é $\mathbb{R}^* \equiv [-\infty, \infty]$ é dita uma variável aleatória se para todo $B \in \mathcal{R}^*$ temos $X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}$. Aqui, $\mathcal{R}^*$ = os conj. de Borel de $\mathbb{R}^*$ com $\mathbb{R}^*$ dada a topologia usual, isto é, aquela gerada por int. da forma $[-\infty, a)$, (a, b) e $(b, \infty]$ onde $a, b \in \mathbb{R}$. Deve-se notar que a reta real extendida $(\mathbb{R}^*, \mathcal{R}^*)$ é um espaço mensurável, de modo que todos os resultados anteriores generalizam imediatamente.

1.4 Integração

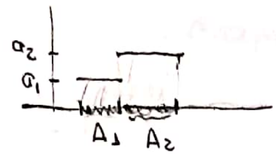
- μ uma medida σ-finita em $(\Omega, \mathcal{F})$
- Primeiro interesse: qdo μ é uma med. de prob., mas alg. vezes será necessário integrar com resp. à uma medida infinita, e não é difícil desenvolver os res. em geral.

Nesta seção vamos definir $\int f \, d\mu$ para uma classe de funções mensuráveis. Isto é um procedimento de quatro passos:

1. Funções Simples
2. Funções Limitadas
3. Funções Não-negativas
4. Funções gerais

Passo 1. φ é dita uma função simples se $\varphi(\omega) = \sum_{i=1}^{m} a_i \mathbb{1}_{A_i}$ e A_i são cjtos disjuntos com $\mu(A_i) < \infty$. Se φ é uma função simples, temos

$$\int \varphi \, d\mu = \sum_{i=1}^{n} a_i \, \mu(A_i)$$



$\mu(A_1).a_1 + \mu(A_2) \, a_2 + \cdots$

Largura

- A rep. de φ não é única uma vez que não supomos a_i's diferentes. No entanto, é fácil ver que a última def. não contradiz a si mesma.

- $\varphi \geq \psi$ μ-quase em todo lugar (ou $\varphi \geq \psi$ μ-a.e.) significa que $\mu(\{\omega : \varphi(\omega) < \psi(\omega)\}) = 0$. Qdo não há dúvida sobre a medida, retiramos μ.

Lema 1.4.1. Seja φ e ψ funções simples.

i) Se $\varphi \geq 0$ a.e. então $\int \varphi \, d\mu \geq 0$.

Pela def. temos

$$\int \varphi \, d\mu = \sum_{i=1}^{n} a_i \, \mu(A_i) = \sum_{i \in I_1} a_i \, \mu(A_i) + \sum_{i \in I_2} a_i \, \underbrace{\mu(A_i)}_{0} =$$

$$= \sum_{i \in I_1} a_i \, \mu(A_i) \geq 0$$

ii) Para qualquer $a \in \mathbb{R}$, $\int a\varphi \, d\mu = a \int \varphi \, d\mu$.

$$\int a\varphi \, d\mu = \sum_{i=1}^{n} a \, a_i \, \mu(A_i) = a \sum_{i=1}^{n} a_i \, \mu(A_i) = a \int \varphi \, d\mu.$$

iii) $\int \varphi + \psi \, d\mu = \int \varphi \, d\mu + \int \psi \, d\mu$.

Suponha

$$\varphi = \sum_{i=1}^{m} a_i \, \mathbb{1}_{A_i} \quad e \quad \psi = \sum_{j=1}^{n} b_j \, \mathbb{1}_{B_j}$$

Para tornar o suporte das duas funções o mesmo, tomamos $A_0 = \cup_i B_i - \cup_i A_i$, $B_0 = \cup_i A_i - \cup_i B_i$, e $a_0 = b_0 = 0$. Agora

$$\varphi + \psi = \sum_{i=0}^{m} \sum_{j=0}^{n} (a_i + b_j) \, \mathbb{1}_{(A_i \cap B_j)}$$

e a $A_i \cap B_j$ são disjuntos dois a dois, assim

$$\int (\varphi + \psi) \, d\mu = \sum_{i=0}^{m} \sum_{j=0}^{n} (a_i + b_j) \, \mu(A_i \cap B_j) =$$

$$= \sum_{i=0}^{m}\sum_{j=0}^{n} a_i\, \mu(A_i \cap B_j) + \sum_{i=0}^{m}\sum_{j=0}^{n} b_j\, \mu(A_i \cap B_j) =$$

$$= \sum_{i=0}^{m} a_i\, \mu(A_i) + \sum_{j=0}^{n} b_j\, \mu(B_j) = \int \varphi\, d\mu + \int \psi\, d\mu$$

pois $A_i = +_j (A_i \cap B_j)$ e $B_j = +_i (A_i \cap B_j)$.

• Provaremos (i) – (iii) mais três vezes conf. generalizamos nossa integral. Como cons. de (i) - (iii), temos mais três prop. úteis.

Lema: Se (i) e (iii) valem então temos:

iv) Se $\varphi \leq \psi$ a.e. então $\int \varphi\, d\mu \leq \int \psi\, d\mu$

v) Se $\varphi = \psi$ a.e. então $\int \varphi\, d\mu = \int \psi\, d\mu$

Se, em adição, (ii) vale $a = -1$ temos

vi) $\left|\int \phi\, d\mu\right| \leq \int |\phi|\, d\mu$

Prova.

Por (iii), $\int \psi\, d\mu = \int \phi\, d\mu + \int (\psi - \phi)\, d\mu$ e a segunda integral ≥ 0 por (i), de modo que (iv) vale. $\varphi = \psi$ a.e. implica $\varphi \leq \psi$ e $\psi \leq \varphi$ a.e., assim (v) segue de duas aplicações de (iv). Para provar (vi), note que $\phi \leq |\phi|$, assim (iv) implica $\int \phi\, d\mu \leq \int |\phi|\, d\mu$ $-\phi \leq |\phi|$, assim (iv) e (ii) implicam $-\int \phi\, d\mu \leq \int |\phi|\, d\mu$. Uma vez que $|y| = \max(y, -y)$, o resultado segue.

D. 2 

Passo 2. E um conjunto com $\mu(E) < \infty$ e seja f uma função limitada que se anula em E^c. Para definir a integral de f, observamos que se φ, ψ são simples tal que $\varphi \leq f \leq \psi$, então queremos ter

$$\int \varphi \, d\mu \leq \int f \, d\mu \leq \int \psi \, d\mu$$

assim temos

$$\int f \, d\mu = \sup_{\phi \leq f} \int \varphi \, d\mu = \inf_{\psi \geq f} \int \psi \, d\mu \qquad (1.4.1)$$

Aqui e no resto do Passo 2, assumimos que φ e ψ anulam-se em E^c. Para justificar a definição, temos que provar que o sup e inf são iguais. Isto segue de (iv) no Lema 1.4.2. que

$$\sup_{\phi \leq f} \int \varphi \, d\mu \leq \inf_{\psi \geq f} \int \psi \, d\mu$$

Para provar a outra desigualdade, suponha $|f| \leq M$ e seja

$$E_k = \left\{ x \in E : \frac{kM}{n} \geq f(x) > \frac{(k-1)M}{n} \right\} \quad \text{para } -n \leq k \leq n$$

$$\psi_n(x) = \sum_{k=-n}^{n} \frac{kM}{n} \mathbb{1}_{E_k} \qquad \varphi_n(x) = \sum_{k=-n}^{n} \frac{(k-1)M}{n} \mathbb{1}_{E_k}$$

Por definição, $\psi_n(x) - \varphi_n(x) = (M/n) \mathbb{1}_E$, assim

$$\int \psi_n(x) - \varphi_n(x) \, d\mu = \frac{M}{n} \mu(E)$$

Uma vez que $\varphi_n(x) \leq f(x) \leq \psi_n(x)$, segue de (iii) no Lema 1.4.1. que

$$\sup_{\phi \leq f} \int \varphi \, d\mu \geq \int \varphi_n \, d\mu = -\frac{M}{n} \mu(E) + \int \psi_n \, d\mu \geq -\frac{M}{n} \mu(E) + \inf_{\psi \geq f} \int \psi \, d\mu$$

A última des. vale para todo n, assim a prova está completa.

Lema 1.4.3. Seja E um conjunto com $\mu(E) < \infty$. Se f e g são funções limitadas que se anulam em E^c então

i) Se $f \geq 0$ a.e. então $\int f d\mu \geq 0$
ii) Para qualquer $a \in \mathbb{R}$, $\int af d\mu = a\int f d\mu$.
iii) $\int f+g \, d\mu = \int f d\mu + \int g d\mu$
iv) Se $g \leq f$ a.e. então $\int g d\mu \leq \int f d\mu$
v) Se $g = f$ a.e. então $\int g d\mu = \int f d\mu$
vi) $\left|\int f d\mu\right| \leq \int |f| d\mu$.

Prova. Uma vez que podemos tomar $\phi \equiv 0$, (i) é clara da definição.

Para provar (ii) observamos que se $a > 0$, então $a\varphi \leq af$ se e só se $\varphi \leq f$, assim

$$\int af d\mu = \sup_{\phi \leq f} \int a\varphi d\mu = \sup_{\phi \leq f} a\int \varphi d\mu = a \sup_{\phi \leq f} \int \varphi d\mu = a\int f d\mu$$

Para $a < 0$, observe que $a\varphi \leq af \Leftrightarrow \varphi \geq f$, assim

$$\int af d\mu = \sup_{\phi \geq f} \int a\varphi d\mu = \sup_{\phi \geq f} a\int \varphi d\mu = a \inf_{\phi \geq f} \int \varphi d\mu = a\int f d\mu$$

Para provar (iii), observe que se $\psi_1 \geq f$ e $\psi_2 \geq g$, então $\psi_1 + \psi_2 \geq f+g$, assim

$$\inf_{\psi \geq f+g} \int \psi d\mu \leq \inf_{\psi_1 \geq f, \psi_2 \geq g} \int \psi_1 + \psi_2 \, d\mu$$

usando a linearidade para f. simples, segue que

$$\int f+g \, d\mu = \inf_{\psi \geq f+g} \int \psi \, d\mu \leq \inf_{\psi_1 \geq f, \psi_2 \geq g} \int \psi_1 d\mu + \int \psi_2 d\mu =$$

$$= \int f d\mu + \int g \, d\mu$$

Para provar a outra des., observe que a última conclusão aplicada à $-f$ e $-g$ e (ii) implicam

$$-\int f+g \, d\mu \leq -\int f d\mu - \int g \, d\mu$$

(iv)-(vi) seguem de (i)-(iii) pelo Lema 1.4.2.

Notação.

Usando a linearidade para funções simples, segue que

$$\int f+g \, d\mu = \inf_{\psi \geq f+g} \int \psi \, d\mu$$

$$\leq \inf_{\psi_1 \geq f, \psi_2 \geq g} \int \psi_1 \, d\mu + \int \psi_2 \, d\mu = \int f \, d\mu + \int g \, d\mu$$

Para provar a outra desigualdade, observe que a última conclusão aplicadas a $-f$ e $-g$ e (ii) implicam que

$$-\int f+g \, d\mu \leq -\int f \, d\mu - \int g \, d\mu$$

(iv)-(vi) seguem de (i)-(iii) pelo Lema 1.4.2.

Notação. Definimos a integral de f sobre um conjunto E.

$$\int_E f \, d\mu \equiv \int f \cdot \mathbb{1}_E \, d\mu$$

Passo 3. Se $f \geq 0$, então seja

$$\int f \, d\mu = \sup \left\{ \int h \, d\mu : 0 \leq h \leq f, \; h \text{ limitada e } \mu\{x : h(x) > 0\} < \infty \right\}$$

A última def. é boa uma vez que é claro que esta é bem definida.
O próximo resultado nos ajudará a calcular os val. desta integral.

Lema 1.4.4. Seja $E_n \uparrow \Omega$ com $\mu(E_n) < \infty$ e seja $a \wedge b = \min(a,b)$. Então

$$\int_{E_n} f \wedge n \, d\mu \uparrow \int f \, d\mu \quad \text{qdo } n \uparrow \infty$$

Prova.

É claro, de (iv) no Lema 1.4.3, que o lado esquerdo cresce conforme n cresce. Uma vez que

$$h = (f \wedge n) \mathbb{1}_{E_n}$$ (h é o mínimo entre f e n restrita à E_n)

é uma possibilidade no sup, cada termo é menor que a integral à direita. Para provar que o limite é $\int f \, d\mu$, observe que se $0 \leq h \leq f$, $h \leq M$, e $\mu(\{x : h(x) > 0\}) < \infty$, então para $n \geq M$ usando $h \leq M$, (iv), e (iii)

"$n \geq M$, $h = (f \wedge n) \mathbb{1}_{E_n} \leq M \leq f \wedge M$"

$$\int_{E_n} f \wedge n \, d\mu \geq \int_{E_n} h \, d\mu = \int h \, d\mu - \int_{E_n^c} h \, d\mu$$

$a \leq a - \epsilon_n$

$\liminf \geq a$

Agora, $0 \leq \int_{E_n^c} h \, d\mu \leq M \mu(E_n^c \cap \{x : h(x) > 0\}) \to 0$ qdo $n \to \infty$,

$E_n \uparrow \Omega \Rightarrow E_n^c \downarrow \emptyset \overset{c.p}{\Rightarrow} \mu(E_n^c) \downarrow 0$

assim

$$\liminf_{n \to \infty} \int_{E_n} f \wedge n \, d\mu \geq \int h \, d\mu \cdot \int f \, d\mu$$

e

$$\limsup_{n \to \infty} \int_{E_n} f \wedge n \, d\mu \leq \int f \, d\mu \quad \text{por def.}$$

o que prova o resultado desejado uma vez que h é um membro arbitrário da classe que define a integral de f.

Lema 1.4.5. Suponha $f, g \geq 0$.

i) $\int f d\mu \geq 0$

ii) Se $a > 0$ então $\int af \, d\mu = a \int f \, d\mu$.

iii) $\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu$

iv) Se $0 \leq g \leq f$ q.c. então $\int g \, d\mu \leq \int f \, d\mu$

v) Se $0 \leq g = f$ q.c. então $\int g \, d\mu = \int f \, d\mu$

Aqui, usamos vi pois é trivial para $f \geq 0$. $\left|\int f \, d\mu\right| \leq \int |f| \, d\mu$

i) Por definição

$$\int f \, d\mu = \sup \left\{ \int h \, d\mu \ : \ 0 \leq \underset{\text{limitada}}{h} \leq f \ \text{ com } \mu(\{x : h(x) > 0\}) < \infty \right\}$$

Ou seja h é limitada, em $E = \{x : h(x) > 0\}$ e é nula em E^c, com $\mu(E) < \infty$. Como $h \geq 0$, segue do Lema 1.4.3 que $\int h \, d\mu \geq 0$. Daí, $\int f \, d\mu \geq 0$.

ii)

Passo 4. Dizemos que f é integrável se

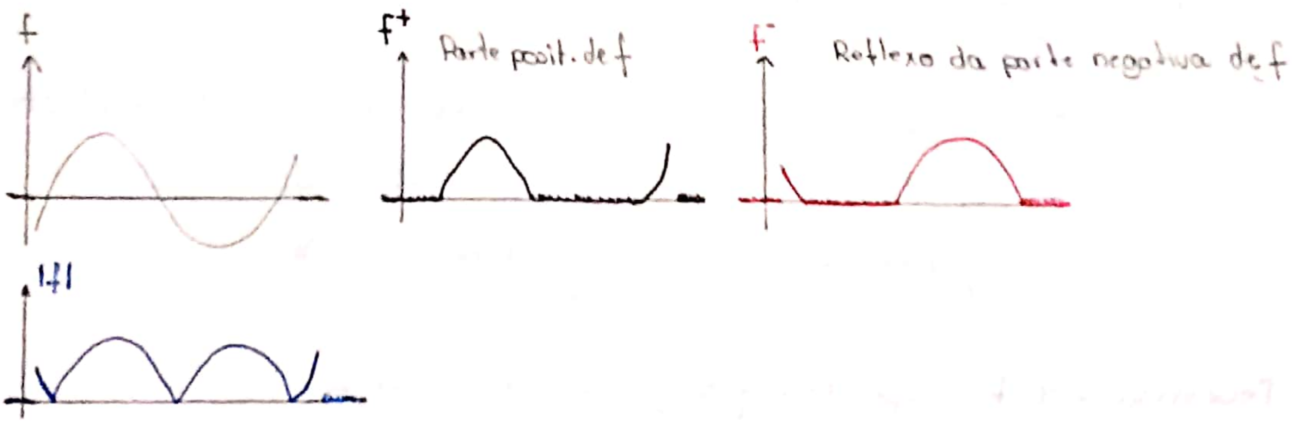
$$\int |f| \, d\mu < \infty.$$

Seja

$$f^+(x) = f(x) \vee 0 \quad e \quad f^-(x) = (-f(x)) \vee 0$$

onde $a \vee b = \max(a,b)$. Claramente,

$$f(x) = f^+(x) - f^-(x) \quad e \quad |f(x)| = f^+(x) + f^-(x)$$



Definimos a integral de f por

$$\int f \, d\mu = \int f^+ d\mu - \int f^- d\mu$$

o lado direito está bem definido já que $f^+, f^- \leq |f|$, f^+ e f^- não-neg.
(iv) Lema 1.4 $\Rightarrow$ $\int f^+ d\mu \leq \int |f| \, d\mu < \infty$ e $\int f^- d\mu \leq \int |f| \, d\mu < \infty$.

Lema. Se $f = f_1 - f_2$ onde $f_1, f_2 \geq 0$ e $\int f_i \, d\mu < \infty$ então

$$\int f \, d\mu = \int f_1 \, d\mu - \int f_2 \, d\mu$$

Sabemos que

$$f = f^+ - f^-$$

e, por outro lado,

$$f = f_1 - f_2$$

Logo,

$$f^+ - f^- = f_1 - f_2 \Leftrightarrow f^+ + f_2 = f^- + f_1$$

Todas são funções não-negativas. Logo, por (iii) do Lema 1.4.5,

$$\int f^+ d\mu + \int f_2 d\mu = \int f^- d\mu + \int f_1 d\mu \Leftrightarrow$$

$$\int f^+ d\mu - \int f^- d\mu = \int f_1 d\mu - \int f_2 d\mu \Leftrightarrow$$

$$\int f d\mu = \int f_1 d\mu - \int f_2 d\mu \qquad \blacksquare$$

Teorema 1.4.7. Suponha f e g são integráveis.

i. Se $f \geq 0$ q.c. então $\int f d\mu \geq 0$.

ii. Para todo $a \in \mathbb{R}$, $\int af d\mu = a \int f d\mu$

iii. $\int f + g \, d\mu = \int f d\mu + \int g d\mu$

iv. Se $g \leq f$ q.c. então $\int g d\mu \leq \int f d\mu$

v. Se $g = f$ q.c. então $\int g d\mu = \int f d\mu$

vi. $\left| \int f d\mu \right| \leq \int |f| d\mu$

i. trivial Se $f \geq 0$, então f é positiva e por i do Lema 1.4.5, segue.

ii. Se $a \geq 0$, $(af^+) = a(f^+)$ e $(af^-) = a(f^-) \Rightarrow \int af d\mu = a \int f d\mu$.

Se $a < 0$, $(af^+) = a(f^-)$ e $(af^-) = a(f^+) \Rightarrow \int af d\mu = a \int f d\mu$

iii. Observe que $f + g = (f^+ + g^+) - (f^- + g^-)$, assim pela Lemas 1.4.6 e 1.4.5,

$$\int f + g \, d\mu = \int f d\mu + \int g d\mu$$

(iv) - (vi) seguem de (i) - (iii) e Lema 1.4.2.

Notação para casos especiais.

a) Quando $(\Omega, \mathcal{F}, \mu) = (\mathbb{R}^d, \mathcal{R}^d, \lambda)$, escrevemos

$$\int f(x)\,dx \text{ para } \int f\,d\lambda$$

b) Qdo $(\Omega, \mathcal{F}, \mu) = (\mathbb{R}, \mathcal{R}, \lambda)$ e $E = [a,b]$, escrevemos

$$\int_a^b f(x)\,dx \text{ para } \int_E f\,d\lambda$$

c) Qdo $(\Omega, \mathcal{F}, \mu) = (\mathbb{R}, \mathcal{R}, \mu)$ com $\mu((a,b]) = G(b) - G(a)$, escrevemos $\int f(x)\,dG(x)$ para $\int f\,d\mu$

d) Qdo Ω é enumerável, $\mathcal{F}$: todos os subconjuntos de Ω, e μ uma medida de contagem, escrevemos

$$\sum_{i \in \Omega} f(i) \text{ para } \int f\,d\mu$$

1.5 Propriedades da Integral

Teorema 1.5.1. Desigualdade de Jensen. Suponha φ convexa, isto é,

$$\lambda\varphi(x) + (1-\lambda)\varphi(y) \geq \varphi(\lambda x + (1-\lambda)y)$$

para todo $\lambda \in (0,1)$ e $x, y \in \mathbb{R}$. Se μ é uma medida de probabilidade, e f e $\varphi(f)$ são integráveis então

$$\varphi\left(\int f\, d\mu\right) \leq \int \varphi(f)\, d\mu$$

Prova. Seja $c = \int f\, d\mu$ e $\ell(x) = ax + b$ uma função linear que tem $\ell(c) = \varphi(c)$ e $\varphi(x) \geq \ell(x)$. Para ver que tal função existe, relembre que a convexidade implica

$$\lim_{h \downarrow 0} \frac{\varphi(c) - \varphi(c-h)}{h} \leq \lim_{h \downarrow 0} \frac{\varphi(c+h) - \varphi(c)}{h}$$

(O limite existe uma vez que as sequências são monotonas). Se a como um número qualquer entre os dois limites e fazer

Ver apêndice

$$\ell(x) = a(x-c) + \varphi(c),$$

então ℓ tem as propriedades desejadas. Com a existência do ℓ estabelecida, o resto é fácil. (iv) no Teorema 1.4.7 implica

$$\int \varphi(f)\, d\mu \geq \int af + b\, d\mu = a\int f\, d\mu + b = \varphi\left(\int f\, d\mu\right)$$

$c = \int f d\mu$, $\ell(c) = \varphi(c)$.

Seja $\|f\|_p = \left(\int |f|^p \, d\mu\right)^{1/p}$ para $1 \leq p < \infty$, e note que $\|cf\|_p = |c| \|f\|_p$ para qualquer c real.

De fato,

$$\|cf\|_p = \left(\int |cf|^p \, d\mu\right)^{1/p} = \left(\int |c|^p |f|^p d\mu\right)^{1/p} = |c| \left(\int |f|^p \, d\mu\right)^{1/p} = |c| \|f\|_p$$

Teorema. Desigualdade de Holder. Se $p, q \in (1, \infty)$ com $\frac{1}{p} + \frac{1}{q} = 1$, então

$$\int |fg| \, d\mu \leq \|f\|_p \|g\|_q$$

Prova.

Se $\|f\|_p$ ou $\|g\|_q = 0$, então $|fg| = 0$ q.c.

$$\|f\|_p = \left(\int |f|^p d\mu\right)^{1/p} = 0 \Leftrightarrow \int |f|^p d\mu = 0 \Leftrightarrow |f|^p = 0 \text{ q.c.} \Leftrightarrow |f| = 0 \text{ q.c.} \Leftrightarrow f = 0 \text{ q.c.}$$

Da. $fg = 0$ q.c. $\Leftrightarrow |fg| = 0$ q.c.

assim, basta provar o resultado quando $\|f\|_p$ e $\|g\|_q > 0$ ou dividindo ambos os ...

Fixe y e seja

$$\varphi(x) = \frac{x^p}{p} + \frac{y^q}{q} - xy \quad \text{para } x \geq 0$$

$$\varphi'(x) = x^{p-1} - y \quad \text{e} \quad \varphi''(x) = (p-1)x^{p-2} > 0$$

assim φ tem mínimo em $x_0 = y^{\frac{1}{p-1}}$. $q = p/(p-1)$ e $x_0^p = y^q$, assim

$$\varphi(x_0) = y^q(1/p + 1/q) - y^{1/p-1} y = 0$$

Uma vez que x_0 é o mínimo, segue que a diferença entre $\frac{x^p}{p} + \frac{y^q}{q}$ e xy

é sempre positiva, isto é

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

Fazendo $x = |\tilde{f}|$ e $y = |\tilde{g}|$ e integrando temos

$$\int |fg| \, d\mu \leq \|f\|_p \|g\|_q$$

com $\tilde{f} = \frac{|f|}{\|f\|_p}$, $\tilde{g} = \frac{|g|}{\|g\|_q}$, , $x = \tilde{f}$ é maior que zero

Veja que

$$\|\tilde{f}\|_p = \left(\int \frac{|f|^p}{((\int |f|^p d\mu)^{1/p})^p} \, d\mu \right)^{1/p} = 1^{1/p} = 1$$

Daí,

$$|\tilde{f}\tilde{g}| \leq \frac{1}{p} |\tilde{f}|^p + \frac{1}{q} |\tilde{g}|^q \Rightarrow$$

$$\underbrace{\int |\tilde{f}\tilde{g}| \, d\mu}_{\|} \leq \frac{1}{p} \underbrace{\int |\tilde{f}|^p d\mu}_{1} + \frac{1}{q} \underbrace{\int |\tilde{g}|^q \, d\mu}_{1} = 1 \Rightarrow$$

$$\int |fg| \, d\mu \leq \|f\|_p \|g\|_q$$

Remark. O caso especial $p = q = 2$ é chama desigualdade de Cauchy-Schwarz. Pode-se ter uma prova direta observando que

$$0 \leq \int (f + \theta g)^2 \, d\mu = \int f^2 d\mu + \theta \left(2 \int fg \, d\mu\right) + \theta^2 \left(\int g^2 d\mu\right)$$

Nosso próximo objetivo é dar as condições para garantir que

$$\lim_{n \to \infty} \int f_n \, d\mu = \int \left(\lim_{n \to \infty} f_n\right) d\mu$$

Primeiro, precisamos de uma definição.

Dizemos que

$$f_n \to f \quad \text{em medida}$$

i.e., para todo $\varepsilon > 0$, $\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) \to 0$ qdo $n \to \infty$. Em um espaço de medida finita, esta é uma suposição mais fraca que $f_n \to f$ q.c., mas o próximo resultado é mais fácil de provar na maior generalidade.

Teorema 1.5.3. Teorema da Convergência Limitada. Seja E um conjunto com $\mu(E) < \infty$. Suponha que f_n se anula em E^c, $|f_n(x)| \leq M$, e $f_n \to f$ em medida. Então

$$\int f \, d\mu = \lim_{n \to \infty} \int f_n \, d\mu$$

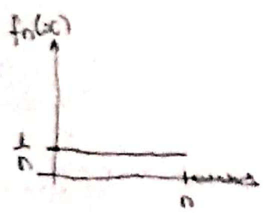
Exemplo 1.5.1. Considere a reta real $\mathbb{R}$ equipada com os conjuntos de Borel $\mathcal{R}$ e a medida de Lebesgue λ. As funções $f_n(x) = 1/n$ em $[0,n]$ e 0 c.c. mostram que a conclusão do Teorema 1.5.3. não vale qdo $\mu(E) = \infty$.

De fato,

limitada

$$f_n(x) = \begin{cases} 1/n & \text{se } x \in [0,n] \\ 0 & \text{, c.c} \end{cases} \qquad f_1 = \begin{cases} 1, & \text{se } x \in [0,1] \\ 0, & \text{c.c.} \end{cases}$$

$$f_2 = \begin{cases} 1/2, & \text{se } x \in [0,2] \\ 0, & \text{c.c.} \end{cases}$$



$$f_n \to 0 \Rightarrow \int \lim_{n \to \infty} f_n \, d\mu = 0$$

mas $\int f_n \, d\mu = 0 \, \mu(E^c) + \frac{1}{n} \mu([0,n]) = 1 \Rightarrow \lim_{n \to \infty} \int f_n \, d\mu = 1$

Prova. Seja $\varepsilon > 0$, $G_n = \{x : |f_n(x) - f(x)| \leq \varepsilon\}$ e $B_n = E - G_n$. Usando (iii) e (vi) do Teorema 1.4.7,

$$\left| \int f \, d\mu - \int f_n \, d\mu \right| = \left| \int f - f_n \, d\mu \right| \leq \int |f - f_n| \, d\mu$$

$|a+b| \leq |a| + |b|$

$$= \int_{G_n} |f - f_n| \, d\mu + \int_{B_n} |f - f_n| \, d\mu$$

* $|f_n| \leq M \Rightarrow |f| \leq M$

$|f_n - f| \leq |f_n| + |f| \leq 2M$

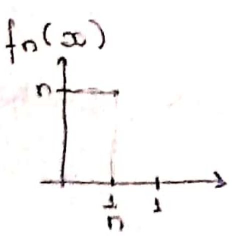
$$\leq \varepsilon \mu(G_n) + 2M \mu(B_n)$$

$f_n \to f$ em medida $\Rightarrow \mu(B_n) \to 0$. $\varepsilon > 0$ é arbitrário e $\mu(E) < \infty$, de modo que a prova está completa.

Teorema 1.5.4. Lema de Fatou. Se $f_n \geq 0$ então

$$\liminf_{n \to \infty} \int f_n \, d\mu \geq \int (\liminf_{n \to \infty} f_n) \, d\mu$$

Exemplo 1.5.2. Exemplo 1.5.1 mostra que podemos ter a desigualdade estrita ($>$) no Teorema 1.5.4. As funções $f_n = n \mathbb{1}_{[0,1/n]}(x)$ em $(0,1)$ equipada com os conjuntos de Borel e a medida de Lebesgue mostram que isso pode acontecer em um espaço de medida finita.



$$f_n(x) \to f(x) = \begin{cases} +\infty & , \text{ se } x = 0 \\ 0 & , \text{ c.c.} \end{cases}$$

$$\int f_n \, d\mu = 1 \Rightarrow \liminf_{n \to \infty} \int f_n \, d\mu = 1$$

$$\int \liminf_{n \to \infty} f_n \, d\mu = \int f \, d\mu = +\infty \, \mu(\{0\}) + 0 \, \mu((0,1]) = 0$$

Prova.

Seja

$$g_n(x) = \inf_{m \geq n} f_m(x).$$

$f_n(x) \geq g_n(x)$ e conforme $n \uparrow \infty$,

$$g_n(x) \uparrow g(x) = \liminf_{n \to \infty} f_n(x) = \sup_{n \geq 0} \inf_{m \geq n} f_m(x) = \lim_{n \to \infty} \left(\inf_{m \geq n} f_m(x) \right)$$

Uma vez que $\int f_n \, d\mu \geq \int g_n \, d\mu$, basta provar que

$$\liminf_{n \to \infty} \int g_n \, d\mu \geq \int g \, d\mu$$

Seja $E_m \uparrow \Omega$ conjuntos de medida finita. Desde que $g_n \geq 0$ e para m fixado

$$(g_n \wedge m) . \mathbb{1}_{E_m} \to (g \wedge m) \mathbb{1}_{E_m} \text{ q.c.}$$

o teorema da Convergência Limitada, 1.5.3, implica

$$\liminf_{n \to \infty} \int g_n \, d\mu \geq \int_{E_m} g_n \wedge m \, d\mu \to \int_{E_m} g \wedge m \, d\mu$$

Tomando o sup sobre m e usando Teorema 1.4.4, temos o resultado desejado.

1.5.6.

Teorema da Convergência Dominada ~~[illegible]~~. Se $f_n \to f$ q.c., $|f_n| \leq g$ para todo n, e g é integrável, então $\int f_n \, d\mu \to \int f \, d\mu$

Prova. $f_n + g \geq 0$, assim, pelo Lema de Fatou,

$$\liminf_{n \to \infty} \int f_n + g \, d\mu \geq \int f + g \, d\mu$$

subtraindo $\int g\,d\mu$ de ambos os lados na eq. anterior,

$$\liminf_{n\to\infty} \int f_n \, d\mu \geq \int f \, d\mu$$

Aplicando o último resultado à $-f_n$ $(-f_n + g \geq 0)$, temos

$$\limsup_{n\to\infty} \int f_n \, d\mu \leq \int f \, d\mu$$

e a prova está completa.

1.5.5 Teorema da Convergência Monótona. Se $f_n \geq 0$ e $f_n \uparrow f$ então

$$\int f_n \, d\mu \uparrow \int f \, d\mu$$

Prova. Lema de Fatou, Teorema 1.5.4, implicam que $\liminf \int f_n \, d\mu \geq \int f \, d\mu$.
Por outro lado, $f_n \leq f \Rightarrow \limsup \int f_n \, d\mu \leq \int f \, d\mu$.

1.6 Valor Esperado

Agora especializamos a integração com resp. a medida de probabi-lidade P. Se $X \geq 0$ é uma v.a. em $(\Omega, \mathcal{F}, P)$ então definimos o seu valor esperado como

$$EX = \int_\Omega X\, dP$$

que sempre faz sentido, mas pode ser ∞. Para reduzir o caso geral ao caso não-negativo, seja $x^+ = \max\{x, 0\}$ a parte positiva e $x^- = \max\{-x, 0\}$ a parte negativa de x. Declaramos que EX existe e def. $EX = EX^+ - EX^-$ sempre que a subtração faz sentido, isto é, $EX^+ < \infty$ ou $EX^- < \infty$.

EX é freq. chamada a média de X e é denotada por μ. EX é definida integrando X, assim ela tem todas as propriedade da integração. Dos Teoremas 1.4.5 e 1.4.7 e da obs. trivial que $E(b) = b$ para qqer b real, temos o seguinte:

Teorema 1.6.1 Suponha $X, Y \geq 0$ ou $E|X|, E|Y| < \infty$.

a) $E(X+Y) = E(X) + E(Y)$

b) $E(aX+b) = aE(X) + b$ para qualquer num. reais a, b.

c) Se $X \geq Y$ então $EX \geq EY$.

1.6.1 Desigualdades

Para medidas de prob., Teorema 1.5.1 se torna:

Teorema 1.6.2. Desigualdade de Jensen. Suponha que φ é convexa, i.e.,

$$\lambda \varphi(x) + (1-\lambda)\varphi(y) \geq \varphi(\lambda x + (1-\lambda) y)$$

para todo $\lambda \in (0,1)$ e $x, y \in \mathbb{R}$. Então

$$E(\varphi(X)) \geq \varphi(EX)$$

provided que ambas as esperanças existem, i.e., $E|X|$ e $E|\varphi(X)| < \infty$.

Dois casos especiais. $|EX| \leq E|X|$

$$(EX)^2 \leq E(X^2)$$

Teorema 1.6.3. Desigualdade de Holder. Se $p, q \in [1, \infty]$ com $\frac{1}{p} + \frac{1}{q} = 1$, então

$$E|XY| \leq \|X\|_p \|Y\|_q$$

Aqui, $\|X\|_r = (E|X|^r)^{1/r}$ para $r \in [1, \infty)$; $\|X\|_\infty = \inf\{M : P(|X| > M) = 0\}$.

$$E(X; A) = \int_A X \, dP$$

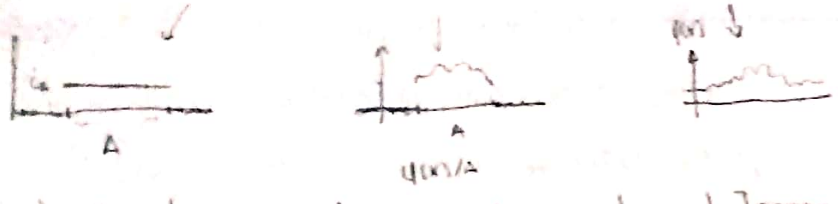
Teorema 1.6.4. Desigualdade de Chebyshev. Suponha $\varphi : \mathbb{R} \to \mathbb{R}$ com $\varphi \geq 0$, seja $A \in \mathcal{R}$ e $i_A = \inf\{\varphi(y) ; y \in A\}$.

$$i_A P(X \in A) \leq E(\varphi(X); X \in A) \leq E\varphi(X).$$

Prova.

A def. de i_A e o fato de que $\varphi \geq 0$ implicam que

$$i_A \mathbb{1}_{(X \in A)} \leq \varphi(X) \mathbb{1}_{(X \in A)} \leq \varphi(X)$$



Daí, tomando o valor esperado e usando a parte c do Teorema 1.6.1, temos o resultado.

Remark. Alguns autores chamam este res. de Desigualdade de Markov e usam o nome Des. de Chebyshev para o caso especial em que $\phi(x) = x^2$ e $A = \{x : |x| \geq a\}$.

$$a^2 P(X \geq a) \leq EX^2 \Rightarrow P(X \geq a) \leq \frac{EX^2}{a^2}$$

1.6.2 Integração Para o Limite

Teorema 1.6.5. Lema de Fatou. Se $X_n \geq 0$ então

$$\liminf_{n\to\infty} EX_n \geq E(\liminf_{n\to\infty} X_n)$$

Teorema 1.6.6. Teorema da Convergência Monótona. ~~[illegible]~~ Se $0 \leq X_n \uparrow X$ então $EX_n \uparrow EX$.

~~[illegible]~~

Teorema 1.6.7. Teorema da Convergência Dominada. Se $X_n \to X$ q.c., $|X_n| \leq Y$ para todo n, e $EY < \infty$, então $EX_n \to EX$.

O caso especial de Teorema 1.6.7 em que Y é constante é chamado teorema da convergência limitada.

Teorema 1.6.8. Suponha $X_n \to X$ q.c.. Seja g, h funções contínuas com

i) $g \geq 0$ e $g(x) \to \infty$ quando $|x| \to \infty$

ii) $|h(x)|/g(x) \to 0$ quando $|x| \to \infty$, e

(iii) $Eg(X_n) \leq K < \infty \ \forall n$.

Então $Eh(X_n) \to Eh(X)$

Prova.

Subtraindo uma cte de h, podemos supor sem perda de generalidade $h(0)=0$. Escolha M grande t.q. $P(|X|=M)=0$ e $g(x)>0$ qdo $|x|\geq M$.

Dada a v.a. Y, seja $\bar{Y}=Y\mathbb{1}_{(|Y|\leq M)}$. Uma vez $P(|X|=M)=0$, $\bar{X}_n\to\bar{X}$ q.c.. Uma vez $h(\bar{X}_n)$ é limitada e contínua, segue do teorema da conv. limitada que

$$E\,h(\bar{X}_n)\to E\,h(\bar{X}) \qquad (a)$$

⋮

1.7 Medidas Produto, Teorema de Fubini

Sejam $(X, \mathcal{A}, \mu_1)$ e $(Y, \mathcal{B}, \mu_2)$ dois espaços de medida σ-finita. Seja

$$\Omega = X \times Y = \{(x,y) : x \in X, y \in Y\}$$

$$\mathcal{S} = \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$$

Conjuntos em $\mathcal{S}$ são chamados retângulos. É fácil ver que $\mathcal{S}$ é uma semialgebra:

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$
$$(A \times B)^c = (A^c \times B) \cup (A \times B^c) \cup (A^c \times B^c)$$

Seja $\mathcal{F} = \mathcal{A} \times \mathcal{B}$ a σ-álgebra gerada por $\mathcal{S}$.

Teorema 1.7.1. Há uma única medida μ em $\mathcal{F}$ com

$$\mu(A \times B) = \mu_1(A)\,\mu_2(B)$$

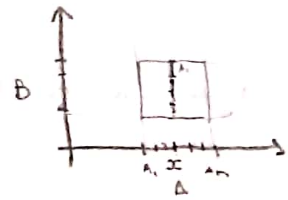
Notação. μ é frequentemente denotada por $\mu_1 \times \mu_2$.

Prova. Pelo Teorema 1.1.4 é suficiente mostrar que se $A \times B = +_i (A_i \times B_i)$ é uma união disjunta finita ou contável, então

$$\mu(A \times B) = \sum_i \mu(A_i \times B_i)$$

Para cada $x \in A$, seja $I(x) = \{i : x \in A_i\}$. $B = +_{i \in I(x)} B_i$, assim

$$\mathbb{1}_A(x)\, \mu_2(B) = \sum_i \mathbb{1}_{A_i}(x)\, \mu_2(B_i)$$



Integrando com respeito a μ_1 e usando o Ex. 1.5.6, temos

$$\mu_1(A)\, \mu_2(B) = \sum_i \mu_1(A_i)\, \mu_2(B_i)$$

o que prova o resultado.

Usando 1.7.1 e indução, segue que se $(\Omega_i, \mathcal{F}_i, \mu_i)$, $i = 1, \dots, n$, são ~~[illegible]~~ espaços de medidas σ-finitas e $\Omega = \Omega_1 \times \dots \times \Omega_n$, há uma única medida μ na σ-álgebra $\mathcal{F}$ gerada por conjuntos da forma $A_1 \times \dots \times A_n$, $A_i \in \mathcal{F}_i$, que tem

$$\mu(A_1 \times \dots \times A_n) = \prod_{m=1}^{n} \mu_m(A_m)$$

$(\Omega, \mathcal{F}, \mu)$ é o prod. de dois esp. mensuráveis $(X, \mathcal{A}, \mu_1)$ e $(Y, \mathcal{B}, \nu)$

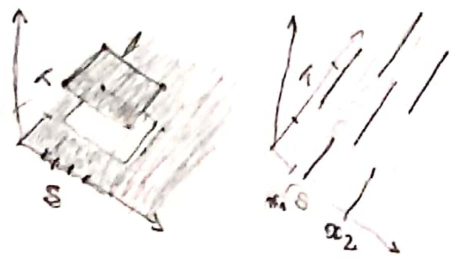
Teorema 1.7.2. Teorema de Fubini. Se $f \geq 0$ ou $\int |f|\, d\mu < \infty$, então

$$(*) \quad \int_X \int_Y f(x,y)\, \mu_2(dy)\, \mu_1(dx) = \int_{X \times Y} f\, d\mu = \int_Y \int_X f(x,y)\, \mu_1(dx)\, \mu_2(dy)$$

Prova. Vamos provar apenas a 1ª igualdade, uma vez que a outra segue por simetria.

Seja $S\times T = \bigcup_i S_i\times T_i$, $S, S_i, i\geq 1 \in \mathcal{S}_1$ e $T, T_i, i\geq 1 \in \mathcal{S}_2$.
Temos que a função

$$\mathbb{1}_{S\times T} = \sum_{i\geq 1} \mathbb{1}_{S_i\times T_i} : \Omega_1\times\Omega_2 \to [0,+\infty] \quad ,$$



mas $(x,y)\in S\times T \Leftrightarrow x\in S$ e $y\in T$, i.e.

$$\mathbb{1}_{S\times T}(x,y) = \mathbb{1}_S(x)\,\mathbb{1}_T(y)$$

Daí,

$$\mathbb{1}_S(x)\,\mathbb{1}_T(y) = \sum_{i\geq 1} \mathbb{1}_{S_i}(x)\,\mathbb{1}_{T_i}(y) \qquad (2)$$

Se fixarmos x e deixar y livre em (2) temos uma função de Ω_2 em $[0,+\infty]$, isto é,

$$\mathbb{1}_S(x)\,\mathbb{1}_T = \sum_{i\geq 1} \mathbb{1}_{S_i}(x)\,\mathbb{1}_{T_i} : \Omega_2 \to [0,+\infty]$$

Integrando em y

$$\int \mathbb{1}_S(x)\,\mathbb{1}_T\, d\mu_2(y) = \sum_{i\geq 1} \int \mathbb{1}_{S_i}(x)\,\mathbb{1}_{T_i}\, d\mu_2(y) \Leftrightarrow$$

$$\mathbb{1}_S(x)\,\mu_2(T) = \sum_{i\geq 1} \mathbb{1}_{S_i}(x)\,\mu_2(T_i)$$

Int. em x,

$$\mu_1(S)\,\mu_2(T) = \sum_{i\geq 1} \mu_1(S_i)\,\mu_2(T_i) \qquad ▨$$

Teorema de Fubini.

Se $f \geq 0$ ou $\int |f| < \infty$ para $f : (\Omega, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B})$, $\Omega = \Omega_1 \times \Omega_2$

$$\underbrace{\int_{\Omega_1} \int_{\Omega_2} f(x,y)\, \mu_2(dy)}_{h : (\Omega_1, \mathcal{F}_1) \to (\mathbb{R}, \mathcal{B})} \mu_1(dx) = \int_{\Omega} f\, d\mu = \int_{\Omega_2} \int_{\Omega_1} f(x,y)\, \mu_1(dx)\, \mu_2(dy)$$

→ $h \in \mathcal{F}_1$

→ Para cada $E \in \mathcal{F}$ e $x \in \Omega_1$, $E_x = \{ y \in \Omega_2 : (x,y) \in E \} \in \mathcal{F}_2$.

fixa $x \in \Omega_1$ e declara

$$\mathcal{G}_x = \{ E \in \mathcal{F} : E_x \in \mathcal{F}_2 \}$$

$\mathcal{G}_x$ contém os retângulos pois $E_x = (a,b) \in \mathcal{F}_2$.

. Se $E \in \mathcal{G}_x$, $E^c \in \mathcal{G}_x$

$\Downarrow$

$E_x \in \mathcal{F}_2 \Rightarrow (E_x)^c \in \mathcal{F}_2 \Rightarrow (E^c)_x \in \mathcal{F}_2 \Rightarrow E^c \in \mathcal{G}_x$

. Se $E_1, E_2, \ldots$ em $\mathcal{G}_x$ então $\cup E_i \in \mathcal{G}_x$

$$(\cup E_i)_x = \cup (E_i)_x$$

$\mathcal{G}_x$ é uma álg. que contém os ret. $\Rightarrow \mathcal{G}_x \supset \mathcal{F}$. Logo, $E \in \mathcal{F} \Rightarrow E \in \mathcal{G}_x$

$E_x \in \mathcal{F}_2$.

→ $\forall E \in \mathcal{F}$ seja

$$g(x) = \mu_2(E_x)$$

Lema

é mensurável em $\mathcal{F}_1$, e

$$\int_{\Omega_1} g(x)\, \mu_1(dx) = \mu(E).$$

• $\mathcal{G}$ = cjto dos E t.q. (*) é verdadeiro

• $\mathcal{G}$ é um sist. λ

• $\mathcal{G}$ cont. um sist. π, pois $\mathcal{G}$ contem os ret. • $(E = A \times B$ [illegible] $g(x) = \mathbb{1}_A \mu_2(B) \in \mathcal{F}_1)$

Se $A \in \mathcal{G} \Rightarrow A^c \in \mathcal{G}$

Teorema de Fubini

$f: (\Omega_1, \mathcal{F}, \mu_1) \to (\mathbb{R}, \mathcal{B})$

$f: (\Omega_1 \times \Omega_2, \mathcal{F}, \mu)$

Se $f \geq 0$ ou $\int |f| < \infty$

$$\int_{\Omega_1} \int_{\Omega_2} f(x,y)\, \mu_2(dy)\, \mu_1(dx) = \int_{\Omega} f\, d\mu = \int_{\Omega_2} \int_{\Omega_1} f(x,y)\, \mu_1(dx)\, \mu_2(dy)$$

Um lado

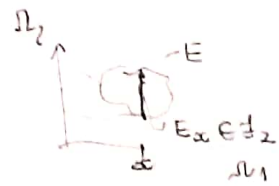
- $h(x) = \int_{\Omega_2} f(x,y)\, \mu_2(dy) \in \mathcal{F}_1$?

$h: (\Omega_1, \mathcal{F}_1) \to (\mathbb{R}, \mathcal{B})$ $h \in \mathcal{F}_1 \Leftrightarrow \forall B \in \mathcal{B}, h^{-1}(B) = \{w : h(w) \in B\} \in \mathcal{F}_1$

Lema: "fatias de $E \in \mathcal{F}$ estão nas σ-álgebras marginais

Lema. Se $E \in \mathcal{F}$,

$$E_x = \{y \in \Omega_2 : (x,y) \in E\}$$



então $E_x \in \mathcal{F}_2\ \forall x \in \Omega_1$

Dem

Estudo do Livro Probability and Measure, by Patrick Billingsley

Capítulo 1. Probabilidade

04/03/14

1.1 Borel's Normal Number Theorem

Apesar de suficiente para muitos tópicos, a teoria de espaços de probabilidade discretos (Feller, Vol. I) não fornece um tratamento rigoroso à dois tipos de problemas: os que envolvem uma operação repetida infinitamente, como uma sequência infinita de lançamentos de moedas e aqueles envolvendo uma "infinitely fine operation", como a escolha aleatória de um ponto de um segmento.

O Intervalo Unitário

O objetivo é construir, ao mesmo tempo, um modelo para os dois problemas citados acima.

Seja $\Omega = [0,1]$; pra ser definido, tome int. abertos à esquerda e fechado à direita. Seja ω um ponto genérico de Ω. Denote o comprimento de um intervalo $I = (a,b]$ por $|I|$:

$$|I| = |(a,b]| = b - a \qquad (1.1)$$

Se

$$A = \bigcup_{i=1}^{n} I_i = \bigcup_{i=1}^{n} (a_i, b_i], \qquad (1.2)$$

onde os intervalos $I_i = (a_i, b_i]$ são disjuntos e estão contidos em Ω, atribuímos a A a probabilidade

$$P(A) = \sum_{i=1}^{n} |I_i| = \sum_{i=1}^{n} (b_i - a_i) \qquad (1.3)$$

É importante entender que nesta seção $P(A)$ é definida apenas se A é uma união finita de subintervalos disjuntos de $(0,1]$ - nunca para conjuntos A de qualquer outro tipo.

Se A e B são da forma acima, e A e B são disjuntos, então $A \cup B$ é uma união finita de intervalos disjuntos e

$$P(A \cup B) = P(A) + P(B) \qquad (1.4)$$

Esta relação, que é certamente óbvia intuitivamente, é consequência da aditividade da integral de Riemann:

$$\int_0^1 (f(\omega) + g(\omega))\, d\omega = \int_0^1 f(\omega)\, d\omega + \int_0^1 g(\omega)\, d\omega \qquad (1.5)$$

onde

$$f(\omega) = I_A(\omega) \quad \text{e}$$
$$g(\omega) = I_B(\omega)$$

. Se $f(\omega)$ é uma função step tomando valor c_j no intervalo $(x_{j-1}, x_j]$, onde $0 = x_0 < x_1 < \ldots < x_k = 1$, então sua integral de Riemann tem valor:

$$\int_0^1 f(\omega)\, d\omega = \sum_{j=1}^{k} c_j (x_j - x_{j-1}) \qquad (1.6)$$

Como dito, sendo $f(\omega) = I_A(\omega)$ e $g(\omega) = I_B(\omega)$, (1.4) segue de (1.5) e (1.6), uma vez que A e B são disjuntos. Isto mostra tbm que a def. (1.3) não é ambígua - note que A terá muitas representações da forma (1.2) pois $(a,b] \cup (b,c] = (a,c]$. Mais tarde, estes fatos serão derivados outra vez a partir da teoria geral da integração de Lebesgue.

De acordo com os modelos usuais, se uma substância radiativa emitiu uma única partícula-α durante um intervalo unitário de tempo, ou se uma única chamada de telefone chegou a uma estação durante um intervalo unitário de tempo, então o instante no qual a emissão ou a chegada ocorreu é aleatório no sentido que cai em (1.2) com probabilidade (1.3). Assim, (1.3) é o ponto inicial para a descrição da seleção de um ponto ao aleatório do um intervalo unitário: Ω é considerado como um espaço amostral, e o conjunto (1.2) é identificado com o evento que o ponto aleatório cai nele.

A def. (1.3) é também o ponto inicial para a representação matemática de uma sequência infinita de lançamentos de uma moeda. A cada ω associamos sua "nonterminating dyodic expansion

$$\omega = \sum_{n=1}^{\infty} \frac{d_n(\omega)}{2^n} = .d_1(\omega)\, d_2(\omega) \ldots, \tag{1.7}$$

cada $d_n(\omega)$ sendo 0 ou 1 [A31: parágrafo 31 do apêndice do livro]. Assim,

$$(d_1(\omega), d_2(\omega), \ldots) \tag{1.8}$$

é a sequência de dígitos binários na expansão de ω. "For definiteness", um ponto como $\frac{1}{2} = .1000\ldots = .0111\ldots$, que tem duas expansões: 1 toma a expansão $.111\ldots$.

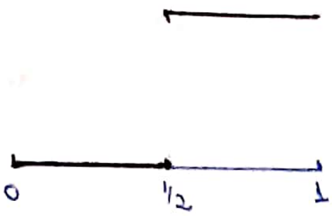


Gráfico de $d_1(\omega)$

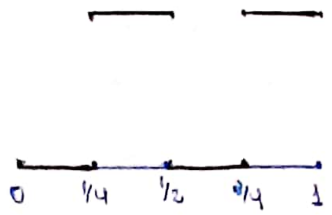


Gráfico de $d_2(\omega)$

Imagine agora uma moeda com faces nomeadas 1 e 0 ao invés da usual cara e coroa. Se ω é escolhido aleatoriamente, então (1.8) se comporta como se fosse o resultado de uma sequência infinita de lançamentos de uma moeda. Para ver isto, considere primeiro o conjunto de ω para o qual $d_i(\omega)=u_i$ para $i=1,\dots,n$, onde $u_1,\dots,u_n$ é a seq. de 0's e 1's. Tais ω satisfazem

$$\sum_{i=1}^{n} \frac{u_i}{2^i} < \omega \le \sum_{i=1}^{n} \frac{u_i}{2^i} + \sum_{i=n+1}^{\infty} \frac{1}{2^i},$$

onde os valores extremos de ω correspondem ao caso $d_i(\omega)=0$ para $i>n$ e ao caso $d_i(\omega)=1$ para $i>n$. O segundo caso pode ser alcançado, mas desde que a expansão binária representada pelos $d_i(\omega)$ é nonterminating – não termina em 0's – o primeiro não pode, e ω deve, na verdade, exceder $\sum_{i=1}^{n} u_i/2^i$. Deste modo

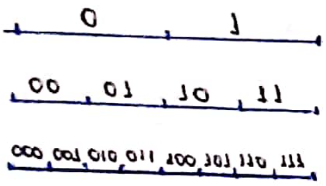
$$[\omega : d_i(\omega)=u_i,\ i=1,\dots,n] = \left(\sum_{i=1}^{n} \frac{u_i}{2^i},\ \sum_{i=1}^{n} \frac{u_i}{2^i} + \underbrace{\frac{1}{2^n}}_{\frac{1}{1-1/2} - \frac{(1-(1/2)^{n+1})}{1-1/2}} \right] \qquad (1.9)$$

O intervalo aqui é aberto à esquerda e fechado à direita precisamente por causa da exp. (1.7) ser nonterminating one. No modelo

No modelo para o lançamento de uma moeda, o conjunto (1.9) representa o evento em que os primeiros n lançamentos deram as saídas $u_1, \dots, u_n$ em sequência. Por (1.3) e (1.9),

$$P[\omega: d_i(\omega) = u_i, i = 1, \dots, n] = \frac{1}{2^n} \qquad (1.10)$$

que é o que a intuição probabilística requer.



Decomposição por intervalos diádicos.

Os intervalos (1.9) são chamados intervalos diádicos, os pontos finais sendo racionais diádicos adjacentes $k/2^n$ e $k+1/2^n$ com o mesmo denominador, e n é o rank ou a ordem do intervalo. Para cada n os 2^n intervalos diádicos de ordem n decompõem ou particionam o intervalo unitário. Na passagem da partição para n para aquela para n+1, cada intervalo (1.9) é dividido em duas partes de igual comprimento, uma metade esquerda em que $d_{n+1}(\omega)$ é 0 e uma metade à direita em que $d_{n+1}(\omega) = 1$. Para $u = 0$ e $u = 1$, o conjunto $[\omega: d_{n+1}(\omega) = u]$ é deste modo a união disjunta de 2^n intervalos de comp. $1/2^{n+1}$ e portanto tem probabilidade $1/2$: $P[\omega: d_n(\omega) = u] = 1/2$ para todo n.

Note que $d_i(\omega)$ é constante sobre cada intervalo diádico de rank i e que para $n > i$ cada intervalo diádico de rank n é inteiramente contido em um único intervalo diádico de rank i. Portanto, $d_i(\omega)$ é constante sobre cada intervalo diádico de rank n se $i \leq n$.

As prob. de vários eventos familiares pode ser escritos abaixo imediatamente. A soma $\sum_{i=1}^{n} d_i(\omega)$ é o número de 1's entre $d_1(\omega), \ldots, d_n(\omega)$, a ser pensado como o número de caras em n lançamentos de uma moeda justa. A formula binomial usual é

$$P\left[\omega : \sum_{i=1}^{n} d_i(\omega) = k\right] = \binom{n}{k} \frac{1}{2^n}, \quad 0 \leq k \leq n, \quad (1.11)$$

Isto segue diretamente da definição. O conjunto à esequerda em (1.11) é união daqueles intervalos (1.9) correspondente as sequências $\omega_1, \ldots, \omega_n$ contendo k 1's e n-k's 0; cada um destes intervalos tem comprimento $1/2^n$ por (1.10) e há $\binom{n}{k}$ deles, e assim (1.11) segue de (1.3).

As funções $d_n(\omega)$ podem ser vistos de dois modos. Fixando n e deixando ω variar temos uma função real $d_n = d_n(\cdot)$ no intervalo unitário. Fixando ω e variando n, temos a sequência (1.8) de 0's e 1's. As prob. (1.10) e (1.11) envolvem apenas "infinitely many" das componentes $d_i(\omega)$. O interesse aqui, no entanto, será indicar principalmente as propriedades de sequências inteiras (1.8). Será visto que as propriedades matemáticas desta sequência espelham as propriedades esperadas de um processo da lançamentos de uma moeda que continua para sempre.

Como a expansão (1.7) é a "nonterminating one", há o defeito de que para nenhum ω é (1.8), por exemplo, a sequência (1,0,0,...). Parece claro que a chance de uma moeda levar a cara no primeiro lançamento e coroas para sempre depois disto é igual à 0, de modo que a ausência de (1,0,0,0,...) - ou de qualquer outra sequência única (isolada) - não deveria importar.

A Lei Fraca dos Grandes Números

No estudo da conexão com o lançamento de uma moeda é instrutivo começar com um resultado que pode, de fato, ser tratado dentro do quadro de prob. discreta, isto é, a lei fraca dos grandes números

Teorema 1.1. Para cada ϵ,

$$\lim_{n\to\infty} P\left[\omega : \left| \frac{1}{n}\sum_{i=1}^{n} d_i(\omega) - \frac{1}{2}\right| \geq \varepsilon \right] = 0 \qquad (1.12)$$

Interpretado probabilisticamente, (1.12) diz que se n é grande, então há prob. pequena de que a fração ou freq. relativa de caras em n lançamentos se desvie muito de $1/2$, uma idéia que encontra-se na base da concepção frequentista de probabilidade. Como uma declaração sobre a estrutura dos números reais, (1.12) é também interessante aritmeticamente.

Desde que $d_i(\omega)$ é cte sobre cada intervalo diádico de rank n se $i \leq n$, a soma $\sum_{i=1}^{n} d_i(\omega)$ é também constante sobre cada intervalo diádico de rank n. O conjunto em (1.12) é portanto a união de determinados intervalos (1.9), e assim sua probabilidade está bem definida por (1.3).

Com a int. de Riemann no papel de valor esperado, a aplicação usual da desigualdade de Chebyshev leva à prova de (1.12). O argumento se torna mais simples se o $d_n(\omega)$ são substituídos pelas funções de Rademacher,

$$r_n(\omega) = 2d_n(\omega) - 1 = \begin{cases} +1 & \text{se } d_n(\omega) = 1, \\ -1 & \text{se } d_n(\omega) = 0. \end{cases} \qquad (1.13)$$

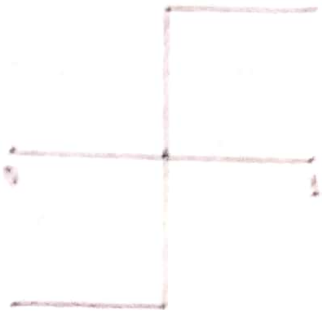


Gráfico de $r_1(\omega)$

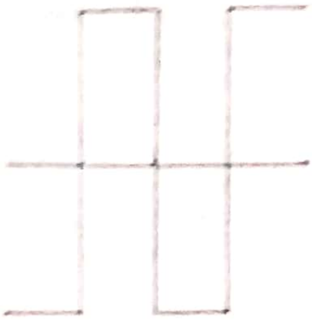


Gráfico de $r_2(\omega)$

Considere a soma parcial

$$S_n(\omega) = \sum_{i=1}^{n} r_i(\omega) \qquad (1.14)$$

Desde que $\sum_{i=1}^{n} d_i(\omega) = (S_n(\omega) + n)/2$, [$\sum d_i(\omega) = \sum(\ldots)$ ⇒ $2\sum d_i(\omega) = S_n(\omega) + n$ [illegible]]

(1.12) com $\varepsilon/2$ no lugar de ε é a mesma coisa que

$$\lim_{n\to\infty} P\left[\omega : \left|\frac{1}{n} S_n(\omega)\right| \geq \varepsilon\right] = 0. \qquad (1.15)$$

Esta é a forma em que o teorema será provado.

As funções de Rademacher tem elas próprias o significado probabilístico direto. Se uma moeda é lançada sucessivamente, e se uma partícula começando da origem performa um caminho aleatório na reta real movendo uma unidade na direção positiva ou negativa conforme a moeda cai cara ou coroa, então $r_i(\omega)$ representa a distância que ele percorre no i-ésimo passo e $S_n(\omega)$ representa sua posição depois de n passos. Há tbm uma interpretação de jogo: se um jogador aposta 1 dólar e em cada lança:

mento da moeda, $r_i(\omega)$ representa seu ganho ou perda na i-ésima jogada e $s_n(\omega)$ representa seu ganho ou perda em n rodadas.

Cada intervalo diádico de rank $i-1$ divide em dois intervalos diádicos de rank i; $r_i(\omega)$ tem o valor -1 em um destes e $+1$ no outro. Deste modo $r_i(\omega)$ é -1 em um conjunto de intervalos de comprimento total $1/2$ e $+1$ em um cjto de comprimento total $1/2$. Portanto $\int_0^1 r_i(\omega)\,d\omega = 0$ por (1.6), e

$$\int_0^1 s_n(\omega)\,d\omega = 0 \qquad (1.16)$$

por (1.5). Se a integral é vista como um valor esperado, então (1.16) diz que a posição média depois de n passos de um caminho aleatório é 0.

Suponha que $i < j$. Em um intervalo diádico de rank $j-1$, $r_i(\omega)$ é constante e $r_j(\omega)$ tem valor -1 na metade à esquerda e $+1$ na outra metade à direita. O produto $r_i(\omega)\,r_j(\omega)$ portanto integra 0 sobre cada dos int. diádicos de rank $j-1$, e assim

$$\int_0^1 r_i(\omega)\,r_j(\omega)\,d\omega = 0 \, , \quad i \neq j \qquad (1.17)$$

Isto corresponde ao fato de que variáveis aleatórias independentes são não-correlacionadas. Desde que $r_i^2(\omega) = 1$, expandindo o quadrado da soma (1.14), temos

$$\int_0^1 s_n^2(\omega)\,d\omega = n. \qquad (1.18)$$

Isto corresponde ao fato de que a variâncias de v.a. independentes são somadas. Claro (1.16), (1.17), e (1.18) permanecem por si só, de modo algum dependo de alguma interpretação probabilística

Aplicando a Desigualdade de Chebyshev em um modo formal para a probabilidade em (1.15) temos

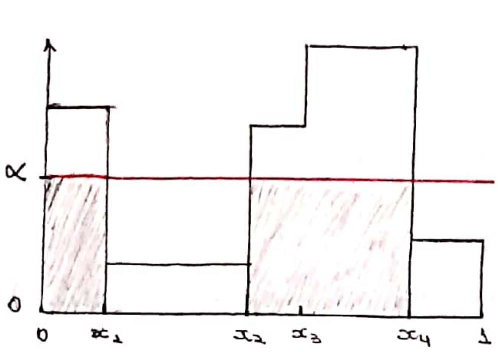
$$P[\omega : |s_n(\omega)| \geq n\epsilon] \leq \frac{1}{n^2\epsilon^2}\int_0^1 s_n^2(\omega)\,d\omega = \frac{1}{n\epsilon^2} \qquad (1.19)$$

O seguinte lema justifica a desigualdade.

Seja f uma função step como em (1.6): $f(\omega) = c_j$ para $\omega \in (x_{j-1}, x_j]$, onde $0 \leq x_0 < x_1 < \ldots < x_k = 1$.

Lema. Se f é uma função step não-negativa, então $[\omega : f(\omega) \geq \alpha]$ é para $\alpha > 0$ uma união finita de intervalos e

$$P[\omega : f(\omega) \geq \alpha] \leq \frac{1}{\alpha}\int_0^1 f(\omega)\,d\omega \qquad (1.20)$$



A região sombreada tem área $\alpha P[\omega : f(\omega) \geq \alpha] = \alpha[x_1 + (x_4 - x_2)]$

Prova. O conjunto em questão é a união de intervalos $(x_{j-1}, x_j]$ para o qual $c_j \geq \alpha$. Se $\sum'$ denota a soma sobre tais j satisfazendo $c_j \geq \alpha$, então $P[\omega : f(\omega) \geq \alpha] = \sum'(x_j - x_{j-1})$ pela definição (1.3). Por outro lado, como c_j são todos não-negativos por hipótese, (1.6) nos dá

$$\int_0^1 f(\omega)\,d\omega = \sum_{j=1}^{k} c_j(x_j - x_{j-1}) \geq \sum{}' c_j(x_j - x_{j-1}) \geq \sum{}' \alpha(x_j - x_{j-1}).$$

Daí, segue (1.20).

Tomando $\alpha = n^2 \epsilon^2$ e $f(\omega) = s_n^2(\omega)$ em (1.20) temos (1.19). Claramente, (1.19) implica (1.15), e como já observado, isto volta a implicar (1.12).

A Lei Forte dos Grandes Números

É possível com um mínimo de aparato técnico provar um resultado mais forte que pode até mesmo não ser formulado na teoria de probabilidade discreta. Considere o conjunto

$$N = \left\{ \omega : \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} d_i(\omega) = \frac{1}{2} \right\} \qquad (1.21)$$

consistindo daqueles ω para os quais a frequência relativa assintótica de 1's na sequência (1.8) é $1/2$.

Axioma da Escolha

Seja A_s, $s \in I$, uma família de conjuntos não-vazios, onde I é um conjunto arbitrário (não-vazio) de índices. Então podemos construir um conjunto A tomando ("escolhendo") um elemento a_s de cada conjunto A_s. Em termos mais técnicos, o axioma diz que há funções $F: I \to \bigcup_{s \in I} A_s$ tais que $F(s) \subset A_s$ para todo $s \in I$.

Como discutiremos, o Axioma da Escolha afirma que o produto cartesiano $\underset{s \in I}{X} A_s$ é não-vazio.

A primeira vista desse axioma parece ser uma obviedade. Sucede, porém, que, sobretudo pelo fato de o cjto I de índices ser arbitrário (podendo até ser um cjto infinito e não-contável), a afirmativa que o mesmo contém não pode ser derivada de princípios mais básicos. O axioma faz uma afirmação de existência (de uma função como a F, ou de um conjunto como A formado por elementos escolhidos de A_s) que, geralmente, não pode ser demonstrada construtivamente, ou seja, por exibição explícita de uma tal função ou de um conjunto A.

Relações de Equivalência

Uma relação $E \subset A \times A$ é dita ser uma relação de equivalência em um conjunto não-vazio A se os seguintes quesitos forem satisfeitos:

1. Reflexividade: $(a,a) \in E$ para todo $a \in A$
2. Simetria: $(a,b) \in E \Rightarrow (b,a) \in E$
3. Transitividade: se $(a,b) \in E$ e $(b,c) \in E$, então $(a,c) \in E$

$a \overset{E}{\sim} b$: a e b são equivalentes segundo E, isto é, $(a,b) \in E$.

Ex: $W = \{(x,y) \in \mathbb{R} \times \mathbb{R}$ tal que $x - y \in \mathbb{Q}\}$ é uma relação de equivalência

Relações

João Borata

Seja A e B conjuntos e seja o produto cartesiano $A \times B$. Um subconjunto de $A \times B$ é dito ser uma relação binária, ou simplesmente relação entre A e B.

Dada uma relação $G \subset A \times B$ há duas noções importantes associadas: a de domínio da relação e a de imagem da relação. Define-se por domínio de G o conjunto

$$Dom(G) := \{a \in A \text{ tal que } (a,b) \in G \text{ para algum } b \in B\}$$

Define-se por imagem de G o conjunto

$$Im(G) := \{b \in B \text{ tal que } (a,b) \in G \text{ para algum } a \in A\}$$

Note que $Dom(G) \subset A$ e $Im(G) \subset B$.

Funções

Este é talvez o mais importante exemplo de relação. Sejam A e B conjuntos e F uma relação entre A e B. Então a relação F é dita ser uma função de A em B se $Dom(F) = A$ e se $(a,b) \in F$ e $(a,b') \in F$ só for possível caso $b = b'$. Em outras, a cada elemento a de A a função associa um e apenas um elemento b de B que faz o papel de segundo elemento do par ordenado (a,b). Este segundo elemento associado pela função F ao elemento a, é mais conveniente denotá-lo por $F(a)$. Assim, uma função é o conjunto de pares $\{(a, F(a)) \in A \times B, a \in A\}$*. Freq. denotamos uma função F de A em B por $F: A \to B$.

* Definição por meio do gráfico.

*

Rick Durret, Apêndice A, Measure Theory Details

A.1. Caratheodory's Extension Theorem

Teorema A.1.1. Seja $\mathcal{S}$ uma semialgebra, e seja μ definida em $\mathcal{S}$ com $\mu(\emptyset)=0$. Suponha

(i) se $S \in \mathcal{S}$ é uma união disjunta de $S_i \in \mathcal{S}$, então $\mu(S) = \sum_i \mu(S_i)$ e,

(ii) se $S_i, S \in \mathcal{S}$ com $S = \sum_{i \geq 1} S_i$ então $\mu(S) \leq \sum_i \mu(S_i)$.

Então μ tem uma única extensão $\bar{\mu}$ que é uma medida em $\bar{\mathcal{S}}$, a álgebra gerada por S. Se a extensão é σ-finita, então há uma única extensão ν que uma medida em $\sigma(\mathcal{S})$.

- Lema 1.1.3. Se $\mathcal{S}$ é uma semiálgebra, então $\bar{\mathcal{S}} = \{$uniões disjuntas finitas de conjuntos em $\mathcal{S}\}$ é umálgebra, chamada a àlgebra gerada por $\mathcal{S}$.

- Uma coleção $\mathcal{S}$ de conjuntos é dita uma semi-álgebra se
  (i) É fechada pela intersecção, i.e., se $S, T \in \mathcal{S}$, então $S \cap T \in \mathcal{S}$, e
  (ii) se $S \in \mathcal{S}$, então S^c é união disjunta finita de conjuntos de $\mathcal{S}$.

- Um importante exemplo de semialgebra é:
  Exemplo 1.1.3. $\mathcal{S}_d$ = o conjunto vazio mais todos os conjuntos da forma
  $(a_1, b_1] \times \dots \times (a_d, b_d] \subset \mathbb{R}^d$ onde $-\infty \leq a_i \leq b_i \leq \infty$

- A definição em 1.1.1 dá os valores de μ na semialgebra $\mathcal{S}_1$.

. Measure on the real line. Medidas em $(\mathbb{R}, \mathcal{B})$ são definida by giving probability a Stieltjes measure function com as seguintes propriedades

i) F é não-decrescente

ii) F é contínua à direita, i.e., $\lim_{y \downarrow x} F(y) = F(x)$

. Teorema 1.1.2. Associado com cada Stieltjes measure function F há uma única medida em $(\mathbb{R}, \mathcal{B})$ com $\mu((a,b]) = F(b) - F(a)$

$$\mu((a,b]) = F(b) - F(a) \qquad (1.1.1)$$

Quando $F(x) = x$, a medida resultante é chamada <u>medida de Lebesgue</u>.

com $\mu(A_1) < \infty$ então $\mu(A_n) \downarrow \mu(A)$.

- $A_1 - A_n \uparrow A_1 - A$, assim por (iii)

$$\mu(A_1 - A_n) \uparrow \mu(A_1 - A).$$

Uma vez que $A_1 \supset B$, $\mu(A_1 - B) = \mu(A_1) - \mu(B)$, i.e.

$$\mu(A_1) - \mu(A_n) \uparrow \mu(A_1) - \mu(A) \Leftrightarrow$$
$$\mu(A_n) \downarrow \mu(A)$$

Exemplo 1.1.1. **Espaços de probabilidade discretos.** Se Ω um cjto enumerável, i.e., finito ou infinito enumerável. Seja $\mathcal{F}$ o conjunto das partes de Ω. Seja

$$P(A) = \sum_{\omega \in A} p(\omega) \quad \text{onde } p(\omega) \geq 0 \text{ e } \sum_{\omega \in \Omega} p(\omega) = 1$$

é a mais geral medida de prob. neste espaço. Em muitos casos quando Ω é um cjto finito, temos $p(\omega) = 1/|\Omega|$.

Um exemplo: Astragali, dado usado no antigo Egito feito de osso do tornozelo de ovelha.

- Se $\mathcal{F}_i$, $i \in I$ são σ-álgebras, então $\bigcap_{i \in I} \mathcal{F}_i$ também é. Aqui, $I \neq \emptyset$
- Dado $\mathcal{A}$ uma coleção de subconjuntos de Ω,

$$\sigma(\mathcal{A}) = \bigcap_{\mathcal{F} \in F_{\mathcal{A}}} \mathcal{F}$$, com $F_{\mathcal{A}}$ o conjunto de todas a σ-álgebras que contém $\mathcal{A}$

Em todos os casos a seguir, assumimos que os conjuntos estão em $\mathcal{F}$.

Teorema 1.1.1. Seja μ uma medida em $(\Omega, \mathcal{F})$

(i) Monotonicidade. Se $A \subset B$, então $\mu(A) \leq \mu(B)$

• Seja $B - A = B \cap A^c$ a diferença dos dois cjtos. Usando $+$ para denotar a união disjunta, $B = A + (B - A)$ de modo que

$$\mu(B) = \mu(A) + \underbrace{\mu(B-A)}_{\geq 0} \geq \mu(A)$$

(ii) Subaditividade. Se $A \subset \bigcup_{m=1}^{\infty} A_m$, então $\mu(A) \leq \sum_{m=1}^{\infty} \mu(A_m)$.

• Seja $A_n' = A_n \cap A$, $B_1 = A_1'$ e para $n > 1$, $B_n = A_n' - \bigcup_{m=1}^{n-1} (A_m')$. Desde que os B_n são disjuntos e tem união A, temos por (i) da def. de medida, $B_m \subset A_m$, e (i) deste teorema,

$$\mu(A) = \sum_{m=1}^{\infty} \mu(B_m) \leq \sum_{m=1}^{\infty} \mu(A_m)$$

(iii) Continuidade por ~~baixo~~ cima. Se $A_i \uparrow A$ (i.e., $A_1 \subset A_2 \subset A_3 \subset \ldots$ e $\bigcup_i A_i = A$), então $\mu(A_i) \uparrow \mu(A)$.

• Seja $B_n = A_n - A_{n-1}$. Então os B_n são disjuntos e tem $\bigcup_{m=1}^{\infty} B_m = A$, $\bigcup_{m=1}^{n} B_m = A_n$, assim

$$\mu(A) = \sum_{m=1}^{\infty} \mu(B_m) = \lim_{n\to\infty} \sum_{m=1}^{n} \mu(B_m) = \lim_{n\to\infty} \mu(A_n)$$

(iv) Continuidade por baixo. Se $A_i \downarrow A$ (i.e., $A_1 \supset A_2 \supset A_3 \supset \ldots$ e $\bigcap A_i = A$), Rick Durret Probability: Theory and Examples

1. Measure Theory

1.1 Espaços de Probabilidade

Um espaço de probabilidade é uma tripla $(\Omega, \mathcal{F}, P)$, onde Ω é o conjunto dos "resultados possíveis", $\mathcal{F}$ é um conjunto de "eventos" e $P: \mathcal{F} \to [0,1]$ é uma função que associa probabilidade à eventos. Assumimos de $\mathcal{F}$ é uma σ-álgebra, que é uma coleção não-vazia de subconjuntos de Ω que satisfaz

(i) Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$, e
(ii) Se $A_i \in \mathcal{F}$ é uma sequência enumerável de conjuntos, então $\bigcup_i A_i \in \mathcal{F}$

Aqui e adiante, contável significa finito ou infinito enumerável. Uma vez que $\bigcap A_i = (\bigcup A_i^c)^c$, segue que uma σ-álgebra é fechada por intersecções enumeráveis.

Sem P, $(\Omega, \mathcal{F})$ é chamado um espaço mensurável, i.e., é um espaço onde podemos colocar uma medida. Uma medida é função st não-negativa contávelmente aditiva; i.e. uma função $\mu: \mathcal{F} \to \mathbb{R}$ com

i) $\mu(A) \geq \mu(\emptyset) = 0$ para todo $A \in \mathcal{F}$ e
ii) Se $A_i \in \mathcal{F}$ é uma seq. enumerável de conj. disjuntos, então
$$\mu(\bigcup_i A_i) = \sum_i \mu(A_i)$$

Se $\mu(\Omega) = 1$, chamamos μ de uma medida de probabilidade, denotadas aqui por P.

Royden

3. Uniões, Intersecções e Complementos

. X, $\mathcal{P}(X)$ é um conjunto com todos os subconjuntos de X

. $A, B \subset X$,

$$A \cap B = \{x : x \in A \text{ e } x \in B\}$$

→ $A \cap B = B \cap A$

→ $A \cap B \subset A$

→ $A \cap B = A \Leftrightarrow A \subset B$

→ $(A \cap B) \cap C = A \cap (B \cap C) = A \cap B \cap C$

$$A \cup B = \{x : x \in A \text{ ou } x \in B\}$$

. $A \cup B = B \cup A$

. $A \cup (B \cup C) = (A \cup B) \cup C = A \cup B \cup C$

. $A \subset A \cup B$

. $A = A \cup B \Leftrightarrow B \subset A$

. Leis Distributivas

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

. $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$

$A \cup X = X$, $A \cap X = A$

$$A^c = \tilde{A} = \{x \in X : x \notin A\}$$

. $\emptyset^c = X$, $X^c = \emptyset$

. $(A^c)^c = A$, $A \cup A^c = X$, $A \cap A^c = \emptyset$

. $A \subset B \Leftrightarrow B^c \subset A^c$.

. Leis De Morgan

$$(A \cup B)^c = A^c \cap B^c$$

$$(A \cap B)^c = A^c \cup B^c$$

. Diferença dos conjuntos B e A.

$$B - A = \{x : x \in B \text{ e } x \notin A\} = B \cap A^c$$

. Diferença simétrica de A e B

$$A \Delta B = (A - B) \cup (B - A) =$$

$$= \{x : x \in B, x \in A, x \notin A \cap B\}$$

. Se $A \cap B = \emptyset$, A e B são disjuntos.

. Uma coleção $\mathcal{C}$ de cjtos é dita uma coleção disjunta de cjtos ou uma coleção de conjuntos disjuntos dois a dois se quaisquer dois cjtos em $\mathcal{C}$ são disjuntos.

c

- $f: X \to Y$, $A \subset X$, $g: A \to Y$, $g(x) = f(x)\ \forall x \in A$

g é chamada restrição de f à A $(f|A)$.

g e f tem imagens distintas, e as imagens inversas por g são distintas das por f.

- Pares ordenados podem ser definidos em termos da noção de uma função. Um par ordenado é uma função cujo domínio cujo domínio é $\{1, 2\}$. De modo similar, uma sequência finita, ou n-upla, é uma função cujo domínio é os primeiros n números naturais, i.e., o cjto $\{i \in \mathbb{N} : i \leq n\}$. Ainda usando o mesmo raciocínio, uma sequência infinita é uma função cujo domínio é $\mathbb{N}$. Usamos o termo "sequência" para uma seq. finita ou infinita.

- Se a imagem da sequência está em um cjto X, falamos ~~de uma~~ seq. de (ou em) X ou de uma seq. de elementos de X. O valor da função em i é x_i, o i-ésimo elemento da sequência. Freq. denotaremos n-uplas ordenados por $\langle x_i \rangle_{i=1}^{n}$ e seq. infinitas por $\langle x_i \rangle_{i=1}^{\infty}$. A imagem de $\langle x_i \rangle_{i=1}^{n}$ será denotada por $\{x_i\}_{i=1}^{n}$.

$X \times X \times \dots \times X = X^n$

A é enumerável se é imagem de alguma sequência e finita se é imagem de alg. sequência finita. Um cjto que não é finito é infinito. Usamos o termo "countably infinite" para conjuntos infinitos enumeráveis.

Princípio da Definição Recursiva: Seja f uma função de um cjto X à si mesmo, e seja a um elemento de X. Então há uma única seq. infinita $\langle x_i \rangle$ de X de modo que $x_1 = a$ e $x_{i+1} = f(x_i)$ p/ cada i.

A existência dessa seq. é intuitivamente clara: Definimos $x_1 = a$, $x_2 = f(a)$, $x_3 = f(f(a))$, e assim por diante. Uma prova mais formal pode ser dada como segue. Primeiro provamos por indução em n que para cada nº natural n há uma única seq. finita

$$x_1^{(n)}, x_2^{(n)}, \dots, x_n^{(n)}$$

de modo que

$$x_1^{(n)} = a \quad \text{e} \quad x_{i+1}^{(n)} = f(x_i^{(n)}) \quad \text{para } 1 \leq i < n.$$

- É conveniente considerar um cjto que não tem membros. Uma vez que um cjto é determinado por seus elementos, há apenas um cjto, que chamamos de conjunto vazio e denotamos por $\emptyset$.
- Se A um cjto qualquer, então cada membro de $\emptyset$ (não há nenhum) é um membro de A, e assim $\emptyset \subset A$. Logo, $\emptyset$ é subconjunto de todo conjunto.
- O conjunto $\{x\}$ (só tem o elemento x) é chamado **unit set**, ou o **singleton** de x. Devemos distinguir cuidadosamente x e $\{x\}$. Por exemplo, sempre temos $x \in \{x\}$, enquanto que raramente temos $x \in x$.
- $\{x,y\} = \{y,x\}$ é um par não-ordenado. Freq., é útil considerar o par ordenado $\langle x,y\rangle = \langle y,x\rangle \Leftrightarrow x=y$. Podemos estender a ideia: $\langle x,y,z\rangle$, $\langle x,y,z,w\rangle$, ...
- Produto Cartesiano (Produto direto) entre dois conjuntos X e Y é o conjunto $\{\langle x,y\rangle\}$ de todos os pares ordenados cuja primeira componente pertence à X e a segunda à Y. Similarmente, $X \times Y \times Z$ é o conjunto $\{\langle x,y,z\rangle\}$ de todas as triplas ordenadas, com $x \in X$, $y \in Y$ e $z \in Z$.

$X = \mathbb{R}$, $X \times X = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ pontos no plano

$X \times X \ldots \times X = X^n$

Problemas

1. $\{x : x \neq x\} = \emptyset$

Suponha que $y \in \{x : x \neq x\}$. Daí, teríamos que $y \neq y$, o que é um absurdo.

Logo $\{x : x \neq x\}$ não possui elementos, $\{x : x \neq x\} = \emptyset$.

2. Funções

Por uma função f de um conjunto X em um conjunto Y entendermos uma regra que associa a cada x em X um único elemento em Y. A coleção G de pares da forma $\langle x, f(x)\rangle$ em $X \times Y$ é chamada gráfico da função f. Um subconjunto G de $X \times Y$ é o gráfico de uma função de X se, e só se, para cada $x \in X$ há um único par em G cujo primeiro elemento é x. Uma vez que uma função é determinada por seu gráfico, muitos gostam de definir uma função como seu gráfico.

A palavra "mapping" é freq. usada como um sinônimo para "função". Expressamos o fato de que f é uma função de X em Y escrevendo

$$f : X \to Y$$

Roydon

- X: domínio (ou domínio de definição) de f
- Y: contradomínio de f
- $Im(f) = \{y \in Y : y = f(x), x \in X\} \subset Y$: imagem de f
- Se $Im(f) = Y$, dizemos que f é sobre Y. ↳ sobrejetiva

- $A \subset X$, denotamos por imagem de A pela função f ao conjunto $f[A]$ formado pelos valores $f(x)$ que f assume nos pontos $x \in A$.

$$f[A] = \{f(x) ; x \in A\} = \{y \in Y : y = f(x), x \in A\}$$

Logo, $Im(f) = f[X]$. Daí, f é sobre Y se $Y = f[X]$.
↳ sobrejetiva

Mais importante que a noção de imagem de um conjunto por f é a noção de imagem inversa. Se B é um subconjunto de Y, definimos a imagem inversa $(f^{-1}(B))$ de B pela função f como o conjunto daqueles $x \in X$ para os quais $f(x) \in B$; i.e.

$$f^{-1}[B] = \{x \in X : f(x) \in B\}$$

Deve ser notado que f é sobre Y se, e só se, a imagem inversa de cada subconjunto não-vazio de Y é não-vazio.

- $f : X \to Y$ é chamada um-a-um (ou univalente, ou injetiva) se $f(x_1) = f(x_2)$ apenas quando $x_1 = x_2$

Funções que são um-a-um de X sobre Y são chamadas correspondências um-a-um de X e Y. (São chamadas tbm de bijetivas). Neste caso, há uma função $g : Y \to X$ de modo que para todo x e y temos $g(f(x)) = x$ e $f(g(y)) = y$. A função g é chamada a inversa de f e é algumas vezes denotada por f^{-1}.

- Deve ser notado que se denotarmos g por f^{-1}, então $f^{-1}[E]$ pode ser considerado a imagem inversa de E por f ou a imagem de E por f^{-1}. Afortunadamente para nossa notação, são os mesmos conjunto.

- $f : X \to Y$ e $g : Y \to Z$, definimos $h : X \to Z$ declarando $h(x) = g(f(x))$. A função h é chamada a composição de g com f e é denotada por $g \circ f$.

- É conveniente considerar um cjto que não tem membros. Uma vez que um cjto é determinado por seus elementos, há apenas um cjto, que chamamos de conjunto vazio e denotamos por $\emptyset$.
- Se A um cjto qualquer, então cada membro de $\emptyset$ (não há nenhum) é um membro de A, e assim $\emptyset \subset A$. Logo, $\emptyset$ é subconjunto de todo conjunto.

- O conjunto $\{x\}$ (só tem o elemento x) é chamado **unit set**, ou o **singleton** de x. Devemos distinguir cuidadosamente x e $\{x\}$. Por exemplo, sempre temos $x \in \{x\}$, enquanto que raramente temos $x \in x$.

- $\{x,y\} = \{y,x\}$ é um par não-ordenado. Freq., é útil considerar o par ordenado $\langle x,y\rangle = \langle y,x\rangle \Leftrightarrow x=y$. Podemos estender à ideia: $\langle x,y,z\rangle$, $\langle x,y,z,w\rangle, \dots$

- Produto Cartesiano (Produto direto) entre dois conjuntos X e Y é o conjunto $\{\langle x,y\rangle\}$ de todos os pares ordenados cuja primeira componente pertence à X e a segunda à Y. Similarmente, $X \times Y \times Z$ é o conjunto $\{\langle x,y,z\rangle\}$ de todos as triplas ordenados, com $x \in X$, $y \in Y$ e $z \in Z$.

$X = \mathbb{R}$, $X \times X = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ pontos no plano

$X \times X \times \dots \times X = X^n$

Problemas

1. $\{x : x \neq x\} = \emptyset$

Suponha que $y \in \{x : x \neq x\}$. Daí, teríamos que $y \neq y$, o que é um absurdo

Logo $\{x : x \neq x\}$ não possui elementos, $\{x : x \neq x\} = \emptyset$.

2. Funções

Por uma função f de um conjunto X em um conjunto Y entendemos uma regra que associa a cada x em X um único elemento em Y. A coleção G de pares da forma $\langle x, f(x)\rangle$ em $X \times Y$ é chamada gráfico da função f. Um subconjunto G de $X \times Y$ é o gráfico de uma função de X se, e só se, para cada $x \in X$ há um único par em G cujo primeiro elemento é x. Uma vez que uma função é determinada por seu gráfico, muitos gostam de definir uma função como seu gráfico.

A palavra "mapping" é freq. usada como um sinônimo para "função". Expressamos o fato de que f é uma função de X em Y escrevendo

$$f : X \to Y$$

Ex 2. Mostre que a correspondência $x \mapsto \alpha_n(x)$ definida em (1.1) é uma aplicação de $]0,1]$ em $\{0,1\}$

· Fixado n, α_n é de fato uma função de x, pois,

· Para $x \in]0,1]$, $\exists\, y \in \{0,1\}$ tal que $y = \alpha_n(x)$, que segue de (1.1), pois $\alpha_n(x) = \begin{cases} 0, \dots \\ 1, \dots \end{cases}$ $\forall x \in]0,1]$.

· Seja $y_1 = \alpha_n(x)$ e $y_2 = \alpha_n(x)$.

Se $y_1 \neq y_2$ então, por um lado, $x \in \left]\frac{2p}{2^n}, \frac{2p+1}{2^n}\right]$, $p = 0,1,\dots n$

e por outro lado, $x \in \left]\frac{2p+1}{2^n}, \frac{2p+2}{2^n}\right]$, $p = 0,1,\dots,n$

o que é suficiente para dizer que $x \neq x \Rightarrow\Leftarrow$. Logo $y_1 = y_2$.

A questão que podemos colocar agora é a de saber até que ponto é que a sucessão $(\alpha_n)_{n\in\mathbb{N}^*}$, em que $\alpha_n \in \{0,1\}$, permite localizar sem qualquer ambiguidade o ponto x. Uma primeira ideia natural é a que decorre da observação que o ponto x pertence a uma sucessão de intervalos encaixados, $(I_n(x))_{n\in\mathbb{N}^*}$ que foram sendo construídos pelas etapas descritas acima e que tem as seguintes propriedades.

1. $\forall n \in \mathbb{N}^*$, $x \in I_n(x)$
2. $\forall n \in \mathbb{N}^*$, $\exists\, m \in \{0,\dots,2^n-1\}$, $I_n(x) = \left]\frac{m}{2^n}, \frac{m+1}{2^n}\right]$
3. $\forall n \in \mathbb{N}^*$, $I_n(x) = \left]\frac{m}{2^n}, \frac{m+1}{2^n}\right] \Rightarrow I_{n+1}(x) = \left]\frac{2m}{2^{n+1}}, \frac{2m+1}{2^{n+1}}\right] \vee$

$$I_{n+1}(x) = \left]\frac{2m+1}{2^{n+1}}, \frac{2m+2}{2^{n+1}}\right].$$

Na segunda possibilidade, consideramos a partição:

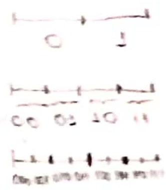
$$]1/2, 1] =]2/2^2, 3/2^2] \cup [3/2^2, 1]$$

e claro está, o código correspondente para esta segunda possibilidade:

$$\alpha_2 = \alpha_2(x) := \begin{cases} 0, \text{ se } x \in]2/2^2; 3/2^2] \\ 1, \text{ se } x \in]3/2^2, 1] \end{cases}$$

Ou seja, em resumo para as duas possibilidades da segunda etapa

$$\alpha_2 = \alpha_2(x) = \begin{cases} 0, & \text{se } x \in]0, 1/2^2] \cup]2/2^2, 3/2^2] \\ 1, & \text{se } x \in]1/2^2, 2/2^2] \cup]3/2^2, 1] \end{cases}$$



$$\vdots$$

$$\alpha_n = \alpha_n(x) = \begin{cases} 0, \text{ se } & x \in \bigcup_{p=0}^{2^{n-1}} \left]\frac{2p}{2^n}, \frac{2p+1}{2^n}\right] \\ 1, \text{ se } & x \in \bigcup_{p=0}^{2^{n-1}} \left]\frac{2p+1}{2^n}, \frac{2p+2}{2^n}\right] \end{cases}$$

Podemos aceitar agora que é válido o seguinte resultado.

Proposição 1. Dado $x \in [0,1]$ existe uma sucessão $(\alpha_n)_{n\in\mathbb{N}^*}$ tal que

$$\alpha_n = \alpha_n(x) := \begin{cases} 0 & \text{se } x \in \left]\frac{2p}{2^n}, \frac{2p+1}{2^n}\right], \; p = 0, \ldots, 2^{n-1}-1 \\ 1 & \text{se } x \in \left]\frac{2p+1}{2^n}, \frac{2p+2}{2^n}\right], \; p = 0, \ldots, 2^{n-1}-1 \end{cases} \quad (1.1)$$

Dem. Leitor, pelo metodo da indução, p.e.

Obs. Note que para cada $n \in \mathbb{N}^*$ fica assim definida uma correspondência que a cada $x \in [0,1]$ associa $\alpha_n(x) \in \{0,1\}$.

26/02/2014

1. A Teoria da Probabilidade, segundo Borel, para sucessões de Bernoulli

1.1 A Representação Binária dos Números de $[0,1]$

Consideremos o seguinte problema: como localizar um dado número x do intervalo $]0,1]$? Uma ideia inspirada no algoritmo da caça ao leão no deserto do Saara é a seguinte. Na primeira etapa, consideramos uma partição do intervalo $]0,1]$ em dois int. de igual comprimento:

$$]0,1] =]0,1/2] \cup]1/2,1]$$

Seja α_1 o código para o resultado da primeira etapa, definido da seguinte forma:

$$\alpha_1 = \alpha_1(x) := \begin{cases} 0 & \text{se } x \in]0;1/2] \\ 1 & \text{se } x \in]1/2;1] \end{cases}$$

Na segunda etapa consideramos duas possibilidades e a ambas aplicamos o processo aplicado na primeira etapa. Na primeira possibilidade supomos que $x \in]0,1/2]$ e consideramos uma nova partição

$$]0,1/2] =]0,1/2^2] \cup]1/2^2, 1/2],$$

dando origem a um código α_2 para o resultado da primeira possibilidade da segunda etapa definida por:

$$\alpha_2(x) = \begin{cases} 0, & \text{se } x \in]0,1/2^2] \\ 1, & \text{se } x \in]1/2^2, 1/2] \end{cases}$$

✦ O sistema binário ou de base 2 é um sistema de numeração posicional em que todas as quantidades se representam com base em dois números ou seja, zero e um.

$11001 = 2^4 + 2^3 + 0.2^2 . 0.2^1 + 1.2^0 = 25$;

25 | 2
1 12 | 2
0 6 | 2
0 3 | 2
1 1 | 2
1 0

$0,101 = 0.2^0 + 1.2^{-1} + 0 2^{-2} + 1.2^{-3} = \frac{1}{2} + \frac{1}{8} = \frac{5}{8}$

**

Classe de Equivalência

Seja A um cjto e $E \subset A \times A$ uma relação de equivalência em A. Para cada $a \in A$ podemos definir o conjunto

$$E(a) := \{a' \in A \text{ tal que } (a,a') \in E\}.$$

Esse conjunto é chamado classe de equivalência de a (pela relação de equivalência E). Na literatura, encontra-se também frequentemente a notação $[a]$ para denotar a classe de equivalência do elemento a.

E.19. Exercício Importante Prove que se A é um cjto e $E \subset A \times A$ é uma relação de equivalência em A, então A é a união disjunta de classes de equivalência de seus elementos.

Sendo $A = \{a_1, a_2, \ldots, a_n\}$, então

$$E \subset \{(a,b);\ a, b \in A\} \text{ de modo que}$$

- $(a,a) \in E \Rightarrow \{(a_i, a_i), i = 1, \ldots, n\} \subset E$
- $(a,b) \in E \Rightarrow (b,a) \in E$
- $(a,b) \in E$, $(b,c) \in E$, então $(a,c) \in E$

. Relações de Equivalência Minimal e Maximal

$$E_{min} := \{(x,x),\ x \in X\} \quad , \quad E_{max} := \{(x,y),\ x,y \in X\} = X \times X$$

Um cjto A é dito ter n elementos (para um número natural n) se for equivalente ao conjunto $\{1, \ldots, n\}$.

. O conj. $\mathbb{Q}_+$ dos n. racionais pos. é enumerável. $\mathbb{Q}$ e $\mathbb{Z}$ o são!

**

· O conjunto dos números reais não é contável

· Mostrar que $U \subset \mathbb{R}$ não o é;

· $U = \{u \in [0,1)$, tais que apenas os nº 0 e 1 aparecem em sua representação decimal$\}$, $U \subset [0,1)$

$$u = \sum_{n=1}^{\infty} \frac{d_n(u)}{10^n} \quad , \quad u_i = 0 \text{ ou } 1$$

Se U é contável, então há $f: \mathbb{N} \to U$ cuja imagem é U. Seja

$$a = \sum_{n=1}^{\infty} \frac{1 - d_n(f(n))}{10^n}$$

- complementar de todos

u_1 $d_1(u_1)$ $1-d_1(u_1)$
u_2 $1-d_1(u_1)$ $d_1(u_2)$
u_3 $d_1(u_1)$ $1-d_1$
⋮
a é diferente de todos

Como $1 - d_n(f(n))$ é 0 ou 1, segue que $a \in U$.

Mas $a \notin Im(f)$. Se $a \in Im(f)$, $\exists\, m$ t.q. $f(m) = a$. Mas isso significaria então que o m-ésimo dígito de a seria $d_m(a) = d_m(f(m))$. Mas pela def. de a, seu m-ésimo dígito é $1 - d_m(f(m))$. Assim, teríamos $d_m(f(m)) = 1 - d_m(f(m))$, o que não é possível

. Prec. def. P em conjuntos não enumeráveis, em infinitos ou não contáveis ctos. com $P(\Omega)$. Duro def. numa col. menor e estab. condições de consistência para estender a medida.

π-sistema Dado um conjunto Ω

σ-álgebra é π-sistema - π

σ-álgebra é sistema - λ mas o inverso não precisa acontecer.

Uma classe que tanto sist. π quanto sist. λ é σ-álgebra.

Teor. (Dinkin) Se $\mathcal{P}$ é um sist. π e $\mathcal{L}$ é um sist. λ, então $\mathcal{P} \subset \mathcal{L} \Rightarrow \sigma(\mathcal{P}) \subset \mathcal{L}$

1. Começar com uma semi-álgebra $\mathcal{S}$
2. mostrar extensão única para uma álgebra $\mathcal{A}(\mathcal{S})$ gerada por $\mathcal{S}$
3. " " " " " " σ- " " $\sigma(\mathcal{A}(\mathcal{S})) = \sigma(\mathcal{S})$

. $\mathcal{S}$ é semiálgebra de Ω e $P: \mathcal{S} \to [0,1]$ é σ-aditiva em $\mathcal{S}$ e $P(\Omega) = 1$. Há uma única extensão P' de P à $\mathcal{A}(\mathcal{S})$ def por

$$P'(S_i) = \sum_i P(S_i)$$

. P em $\mathcal{A}$ tem uma única extensão P' em $\sigma(\mathcal{A})$

. Definir a extensão

$$P^*(A) = \inf_{A \subset \cup A_j} \sum P(A_j)$$

. P^* é cont. subaditiva e $\forall A \in \mathcal{A}$, $P^*(A) = P(A)$

. Checando mensurabilidade da extensão (conjuntos mensuráveis)

E mens. se $P^*(A) \geq P^*(A \cap E) + P^*(A \cap E^c)$ $\mathcal{M}$ é σ-álgebra

. $\mathcal{A} \subset \mathcal{M} \Rightarrow \sigma(\mathcal{A}) \subset \mathcal{M}$ e P^* estende de $\mathcal{A}$ à $\sigma(\mathcal{A})$

. Unicidade.

Espaços de Probabilidade

- $(\Omega, \mathcal{F}, \mathbb{P})$: espaço de probabilidade
  - Ω: conjunto dos resultados
  - $\mathcal{F}$: conjunto dos eventos (σ-álgebra)
  - $\mathbb{P}: \mathcal{F} \to [0,1]$ uma função que associa prob. aos eventos (medida)

- σ-álgebra: coleção não-vazia de subconjuntos de Ω: i) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$, ii) $A_1, \dots, A_n \in \mathcal{F}$
  $\bigcup_i A_i \in \mathcal{F}$

$(\Omega, \mathcal{F})$: espaço mensurável

medida: Uma função de conjuntos não-negativa σ-aditiva, i.e., uma função $\mu: \mathcal{F} \to \mathbb{R}$

i) $\mu(A) \geq \mu(\emptyset) = 0 \quad \forall A \in \mathcal{F}_e$,

ii) se $A_i \in \mathcal{F}$ é uma seq. enumerável de cjtos disjuntos, $\mu(\bigcup_i A_i) = \sum_i \mu(A_i)$

iii) $\mu(\Omega) = 1$, $(\Omega, \mathcal{F}, \mu)$ é um esp. probabilidade.

Propriedades de μ:
1. Monotonicidade: $A \subseteq B$, $\mu(A) \leq \mu(B)$
2. Subaditividade: $A \subseteq \bigcup_{m=1}^{\infty} A_m \Rightarrow \mu(A) \leq \sum_{m=1}^{\infty} \mu(A_m)$
3. Cont. por baixo. $A_1 \supset A_2 \supset$ $A_n \uparrow A$, então, com $\mu(A_1) < \infty$. $\mu(A_i) \downarrow \mu(A)$
4. Cont. por cima. $A_n \downarrow A$ ($A_1 \subset A_2 \subset \dots$) com $\mu(A_1) < \infty$, $\mu(A_i) \downarrow \mu(A)$.

- $\bigcap_{\mathcal{F} \in \mathbb{F}} \mathcal{F} = \sigma(\mathcal{A})$, $\mathbb{F} = \{\mathcal{F}, \text{t.q. } \mathcal{A} \subset \mathcal{F}\}$

$\mathcal{S}$ é dita uma semi-algebra (i) fechada por interserção, i.e., $S, T \in \mathcal{S}$, $S \cap T \in \mathcal{S}$
(ii) $S \in \mathcal{S} \Rightarrow S^c$ é uma união finita de cjtos disjuntos de $\mathcal{S}$

Resumo de Estudo, 07/04/2014

- $(\Omega, \mathcal{F}, P)$

- O que é uma σ-álgebra?

- O que é uma medida?

- Propriedades da medida: Monotocidade, Subaditividade, Continuidade por baixo; Continuidade por cima.

- Espaço de probabilidade discreto. (Exemplo do Astragali.)

- Se $\mathcal{F}_i$, $i \in I$ são σ-álgebras, então $\bigcap_i \mathcal{F}_i$ tbm é. Aqui, $I \neq \emptyset$ é um conjunto de índices arbitrário (i.e., possivelmente não enumerável). Dado uma coleção $\mathcal{A}$ de subconjuntos de Ω, existe uma menor σ-álgebra contendo $\mathcal{A}$: $\sigma(\mathcal{A})$.

- $\mathcal{R}^d$ os conjuntos de Borel: a menor σ-álgebra contendo os cjtos abertos

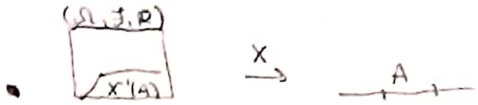
- Medida na reta: Stieltjes measure function F. A cada F há uma única μ com $\mu((a,b]) = F(b) - F(a)$. Se $F(x) = x$, temos a medida de Lebesgue.

- O que é uma semi-algebra.

Distribuições

- $(\Omega, \mathcal{F}, P)$ + interessante qdo def. uma v.a. no mesmo
- Uma f. X def. em Ω é uma v.a. se $\forall B \in \mathcal{B}$, temos
  $X^{-1}(B) = \{\omega \in B : X(\omega) \in B\} \in \mathcal{F}$. X é $\mathcal{F}$-mensurável $(X \in \mathcal{F})$
- Exemplo trivial $X = \mathbb{1}_A$, $A \in \mathcal{F}$

- X induz medida em $\mathbb{R}$, chamada sua distribuição $\mu(A) = P(X \in A) = P(X^{-1}(A))$.
  μ é prob. : . $\mu(A) = P(X \in A) = P(\underbrace{X^{-1}(A)}_{\in \mathcal{F}}) \geq P(\emptyset) = 0$
  . $\mu(\bigsqcup_i A_i) \overset{def}{=} P(X \in \bigsqcup_i A_i) = P(\bigsqcup_i \{X \in A_i\}) = \sum_{i=1}^{\infty} P(X \in A_i) \overset{def}{=} \sum_i \mu(A_i)$

- 

$(\mathbb{R}, \mathcal{B})$ $\mu = P \circ X^{-1}$

$P : \mathcal{F} \to [0,1]$ Assim $P \circ X^{-1} : \mathcal{B} \to [0,1]$
$X^{-1} : \mathcal{B} \to \mathcal{F}$

$$\mathcal{B} \xrightarrow{X^{-1}} \mathcal{F} \xrightarrow{P} [0,1]$$
$\mu = P \circ X^{-1}$

- A dist. de X é descrita dando sua f.d. $F(x) = P(X \leq x)$
- Qqer f.d. tem as seg. prop.: i) F é não-decrescente
  ii) $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$
  iii) F é cont. à direita $\lim_{y \downarrow x} F(y) = F(x)$
  iv) Se $F(x-) = \lim_{y \uparrow x} F(y) \Rightarrow F(x-) = P(X < x)$
  v) $P(X = x) = F(x) - F(x-)$

- Se F satisfaz i-iii, esta é f.d. de alg. v.a.: $X(\omega) = \sup\{y : F(y) < \omega\}$ X

- $X \overset{d}{=} Y \Leftrightarrow P(X \leq x) = P(Y \leq x)$ $\forall x$
- Se $F(x) = \int_{-\infty}^{x} f(y)\,dy$, f é a f.d. de X. Dada f $\Rightarrow$ determ. F

Variáveis Aleatórias

- $X: \Omega \to S$ é dita uma função mensurável de $(\Omega, \mathcal{F})$ à $(S, \mathcal{S})$ se

$$X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F} \quad \forall B \in \mathcal{S}$$

- Se $\{\omega : X(\omega) \in A\} \in \mathcal{F} \ \forall A \in \mathcal{A}$ e $\mathcal{A}$ gera $\mathcal{S}$ $[\sigma(\mathcal{A}) = \mathcal{S}]$, então X é mensurável.

$$\begin{array}{ccc} \mathcal{A} & \longrightarrow & \mathcal{S} \\ \downarrow & & \downarrow \\ X^{-1}(A) \in \mathcal{F} & \longrightarrow & X^{-1}(B) \in \mathcal{F} \end{array}$$

1. Pego uma geral $\mathcal{B} = \{B : \{X \in B\} \in \mathcal{F}\}$
2. Mostro que $\mathcal{B}$ é σ-álgebra e contém $\mathcal{A}$
3. $\mathcal{B} \subset \sigma(\mathcal{A}) = \mathcal{S}$

- Se $\mathcal{S}$ é uma σ-álgebra, $\{\{X \in B\} : B \in \mathcal{S}\}$ é uma σ-álgebra: σ-álgebra gerada por X: $\sigma(X)$

$\sigma(X) = \{\{X \in B\} : B \in \mathcal{S}\}$

- $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ e $f : (S, \mathcal{S}) \to (T, \mathcal{T})$, são f mens. então $f(X)$ é mensurável de $(\Omega, \mathcal{F})$ à $(T, \mathcal{T})$.

- $X_1, \dots, X_n$ v.a. e $f : (\mathbb{R}^n, \mathcal{R}^n) \to (\mathbb{R}, \mathcal{R})$ é mensurável, $f(X_1, \dots, X_n)$ é v.a.

- $X_1, \dots, X_n$ v.a. $\Rightarrow X_1 + X_2 + \dots + X_n$ é v.a.

- $\inf_n X_n$, $\sup_n X_n$, $\limsup_n X_n$, $\liminf_n X_n$ são v.a.

- $\Omega_0 = \{\omega : \lim_{n\to\infty} X_n \text{ existe}\} = \{\omega : \limsup_{n\to\infty} X_n - \liminf_{n\to\infty} X_n = 0\}$, $P(\Omega_0) = 1 \Rightarrow X_n \xrightarrow{q.c.} X_\infty$

Integração

1. Funções Simples
2. Funções Limitadas
3. Funções Não-Negativas
4. Funções Integráveis

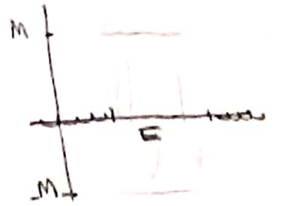
Propriedades

1. Se $f \geq 0$ q.c. $\int f \, d\mu \geq 0$
2. $\forall a \in \mathbb{R}$, $\int a f \, d\mu = a \int f \, d\mu$
3. $\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu$
4. Se $g \leq f$ q.c. $\Rightarrow \int g \, d\mu \leq \int f \, d\mu$
5. Se $g = f$ q.c. $\Rightarrow \int g \, d\mu = \int f \, d\mu$
6. $|\int f \, d\mu| \leq \int |f| \, d\mu$

1. φ é simples: $\varphi(\omega) = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$ e A_i disj. com $\mu(A_i) < \infty$: $\boxed{\int \varphi \, d\mu = \sum_{i=1}^{n} a_i \mu(A_i)}$

- $\varphi \geq \psi$ "q.c." significa $\mu(\{\omega : f(\omega) \leq \psi(\omega)\}) = 0$

2. E um cjto com $\mu(E) < \infty$ e seja f uma função limitada que se anula em E^c. Para definir a integral de f, observamos que φ, ψ são f. simples tal que $\varphi \leq f \leq \psi$, então queremos ter



$$\int \varphi \, d\mu \leq \int f \, d\mu \leq \int \psi \, d\mu$$

Daí,

$$\boxed{\int f \, d\mu = \underbrace{\sup_{\varphi \leq f} \int \varphi \, d\mu}_{A} = \underbrace{\inf_{\psi \geq f} \int \psi \, d\mu}_{B}}$$

- Mostrar que $A = B$; $A \leq B$: $\varphi \leq \psi \Rightarrow \int \varphi \, d\mu \leq \int \psi \, d\mu \Rightarrow \sup_{\varphi \leq f} \int \varphi \, d\mu \leq \inf_{\psi \geq f} \int \psi \, d\mu$
- $B \leq A$. $|f| \leq M$ e seja $E_k = \{x \in E : \frac{kM}{n} \geq f(x) > (k-1)\frac{M}{n}\}$, $-n \leq k \leq n$

$$\psi_n(x) = \sum_{k=-n}^{n} \frac{kM}{n} \mathbb{1}_{E_k}, \quad \varphi_n(x) = \sum_{k=-n}^{n} \frac{(k-1)M}{n} \mathbb{1}_{E_k} \Rightarrow \psi_n(x) - \varphi_n(x) = \frac{M}{n} \mathbb{1}_E$$

$$\sup_{\varphi \leq f} \int \varphi \, d\mu \geq \int \varphi_n \, d\mu = -\frac{M}{n}\mu(E) + \int \psi_n \, d\mu \geq -\frac{M}{n}\mu(E) + \inf_{\psi \geq f} \int \psi \, d\mu$$

$$\int \frac{\prime}{} d\mu = \frac{M}{n}\mu(E)$$

- $\int_E f \, d\mu = \int f \mathbb{1}_E \, d\mu$

3. $f \geq 0$

$$\int f\,d\mu = \sup\left\{\int h\,d\mu : 0 \leq h \leq f,\ h \text{ limitada e } \mu\{x : h(x) > 0\} < \infty\right\}$$

- $E_n \uparrow \Omega$ com $\mu(E_n) < \infty$

$$\int_{E_n} f \wedge n\,d\mu \uparrow \int f\,d\mu \text{ qdo } n \uparrow \infty$$

4. f é integrável se $\int |f|\,d\mu < \infty$

$$\int f\,d\mu = \int f^+\,d\mu - \int f^-\,d\mu \quad , \quad f^+ = f \vee 0$$

$$f^- = -f \vee 0$$

Propriedades da Integral

- Desi. de Jensen: Suponha φ convexa, f e μ é uma med. de prob. e f e $\varphi(f)$ são integráveis então

$$\varphi\left(\int f\,d\mu\right) \leq \int \varphi(f)\,d\mu$$

- $\|f\|_p = \left(\int |f|^p\,d\mu\right)^{1/p}$, $1 \leq p < \infty$

- Des. de Holder: $p, q \in (1, \infty)$ com $\frac{1}{p} + \frac{1}{q} = 1$, então

$$\int |fg|\,d\mu \leq \|f\|_p + \|g\|_q$$

  - $p = q = 2 \Rightarrow$ Des. de Cauchy

- $f_n \to f$ em medida, $\forall \varepsilon > 0$, $\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) \to 0$, qdo $n \to \infty$

- Converg. Limitada. E tq $\mu(E) < \infty$. $f_n = 0$ em E^c, $|f_n(x)| \leq M$ e $f_n \to f$ em medida. Então

$$\int f\,d\mu = \lim_{n\to\infty} \int f_n\,d\mu$$

- Lema de Fatou Se $f_n \geq 0$ então

$$\liminf_{n\to\infty} \int f_n\,d\mu \geq \int \left(\liminf_{n\to\infty} f_n\right) d\mu$$

- Teo. Convergência Dominada. $f_n \to f$ q.c., $|f_n| \leq g$ $\forall n$, e g e integrável, então $\int f_n\,d\mu \to \int f\,d\mu$.

- Teo. Conv. Monótona. $f_n \geq 0$ e $f_n \uparrow f$ então

$$\int f_n\,d\mu \uparrow \int f\,d\mu$$

Valor Esperado

- $EX = \int_\Omega X \, dP$

- Teor. $X, Y \geq 0$, $E|X|, E|Y| < \infty$

  a) $E(X+Y) = EX + EY$

  b) $E(aX+b) = aEX + b \quad \forall a, b \in \mathbb{R}$

  c) $X \geq Y \Rightarrow EX \geq EY$

- Desigualdades:
  - Jensen: $E(\varphi(X)) \geq \varphi(EX)$, $E|X|$ e $E|\varphi(X)| < \infty$
  - Hölder. $E|XY| \leq \|X\|_p \|Y\|_q$, $\|X\|_r = (E|X|^r)^{1/r}$ $r \in [1, \infty)$
  - Chebyshev. $\varphi: \mathbb{R} \to \mathbb{R}$, $\varphi \geq 0$

    $\underset{\inf\{\varphi(y);\, y \in A\}}{i_A} P(X \in A) \leq E(\varphi(X); X \in A) \leq E\varphi(X)$

- Integração p/ o limite:
  - Fatou. $X_n \geq 0 \Rightarrow \liminf\limits_{n\to\infty} EX_n \geq E(\liminf\limits_{n\to\infty} X_n)$
  - Conv. monótona: $0 \leq X_n \uparrow X \Rightarrow EX_n \uparrow EX$
  - " Dominada: $X_n \to X$ q.c., $|X_n| \leq Y \ \forall n$ e $EY < \infty \Rightarrow EX_n \to EX$

Medidas Produto, Teorema de Fubini

$(X, \mathcal{A}, \mu_1)$, $(Y, \mathcal{B}, \mu_2)$ de med. σ-finitas

$\Omega = X \times Y = \{(x,y) : x \in X, y \in Y\}$

$\mathcal{D} = \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$

- $\mathcal{D}$ conjto dos "retângulos". $\mathcal{D}$ é semi-algebra $\mathcal{F} = \sigma(\mathcal{D}) \stackrel{=}{} \text{"}\mathcal{A} \times \mathcal{B}\text{"}$

- Teor. Há uma única med. μ em $\mathcal{F}$ com

$$\mu(A \times B) = \mu_1(A)\mu_2(B) \quad \cdot \ \mu \equiv \mu_1 \times \mu_2$$

- teor. Fubini: $f \geq 0$ ou $\int |f| d\mu < \infty$, então

$$\int_X \int_Y f(x,y)\, \mu_2(dy)\, \mu_1(dx) = \int_{X \times Y} f\, d\mu = \int_Y \int_X f(x,y)\, \mu_1(dx)\, \mu_2(dy)$$

Independência

Def. $\mathcal{F}_1, \dots, \mathcal{F}_n$ σ-álg. Dizemos que esta coleção é ind se $P(A_1 \cap \dots \cap A_n) = \prod_{i=1}^{n} P(A_i)$

$\forall A_1 \in \mathcal{F}_1, \dots, A_n \in \mathcal{F}_n$

Def. $X_1, \dots, X_n$ são v.a.. Esta col. é dita ind. se $P(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^{n} P(X_i \in B_i)$

$\forall B_1 \in \mathcal{B}, B_2 \in \mathcal{B}, \dots, B_n \in \mathcal{B}_n$

Def. Uma coleção de eventos $A_1, \dots, A_n$ é dita independentes se

$$P\Big(\bigcap_{i \in J} A_i\Big) = \prod_{i \in J} P(A_i) \quad \forall J \subset \{1, 2, \dots, n\} = \text{cjto índices}$$

Def: Uma coleção $\mathcal{A}_1, \dots, \mathcal{A}_n$ de classes de cjtos é dita indep. se $P\Big(\bigcap_{i \in J} A_i\Big) = \prod_{i \in J} P(A_i)$

$\forall A_1 \in \mathcal{A}_1, \dots, A_n \in \mathcal{A}_n$

Def Alt. $\mathcal{A}_1, \dots, \mathcal{A}_n$ é ind. se

$$P\Big(\bigcap_{i=1}^{n} A_i\Big) = \prod_{i=1}^{n} P(A_i), \quad A_i \in \bar{\mathcal{A}}_i = \mathcal{A}_i \cup \{\Omega\}$$

Teorema: se $\mathcal{A}_1, \dots, \mathcal{A}_n$ é uma col. de classes de cjtos independentes e cada $\mathcal{A}_i$ é um sistema-π, então

$$\sigma(\mathcal{A}_1), \dots, \sigma(\mathcal{A}_n) \text{ é independente.}$$

Corolário. $X_1, \dots, X_n$ v.a. e, p/ toda coleção $x_1, x_2, \dots, x_n \in \mathbb{R} \cup \{\infty\}$, vale

$$P(X_1 \le x_1, \dots, X_n \le x_n) = \prod_{i=1}^{n} P(X_i \le x_i)$$

2.1.1. Condições Suficientes para Independência

O principal resultado é o Teorema 2.1.3.

Coleções de conjuntos $\mathcal{A}_1, \dots, \mathcal{A}_n \in \mathcal{F}$ são independentes se sempre que $A_i \in \mathcal{A}_i$ e $I \subset \{1, \dots, n\}$ temos

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

Se cada coleção é singular ($\mathcal{A}_i = \{A_i\}$) a def. se reduz àquela para conjuntos.

Lema 2.1.1. Sem perda de generalidade podemos supor cada $\mathcal{A}_i$ contendo Ω. Neste caso a condição é equivalente à

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i) \quad \text{sempre que } A_i \in \mathcal{A}_i$$

uma vez que podemos fixar $A_i = \Omega$ para $i \notin I$.

Prova. Se $\mathcal{A}_1, \dots, \mathcal{A}_n$ são indep. e $\bar{\mathcal{A}}_i = \mathcal{A}_i \cup \{\Omega\}$ então $\bar{\mathcal{A}}_1, \dots, \bar{\mathcal{A}}_n$ são independentes, uma vez que se $A_i \in \bar{\mathcal{A}}_i$ e $I = \{j : A_j = \Omega\}$

Teorema 2.1.2. π-λ Theorem. Se $\mathcal{P}$ é um sistema π e $\mathcal{L}$ é um sistema λ que contém $\mathcal{P}$, então $\sigma(\mathcal{P}) \subset \mathcal{L}$

Teorema 2.1.3. Suponha $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ são independentes e cada $\mathcal{A}_i$ é um sistema π. Então $\sigma(\mathcal{A}_1), \sigma(\mathcal{A}_2), \dots, \sigma(\mathcal{A}_n)$ são independentes.

Prova. $A_2, \dots, A_n$ cj tos com $A_i \in \mathcal{A}_i$, seja $F = A_2 \cap \dots \cap A_n$ e seja $\mathcal{L} = \{A : P(A \cap F) = P(A)P(F)\}$.

- $\Omega \in \mathcal{L}$: $P(\Omega \cap F) = P(\Omega)P(F)$.
- $A, B \in \mathcal{L}$, $A \subset B$, $(B - A) \cap F = (B \cap F) - (A \cap F) \Rightarrow P((B-A) \cap F) = P(B)P(F) - P(A)P(F) = P(B-A)P(F)$
  $B - A \in \mathcal{L}$
- $B_k \in \mathcal{L}$, $B_k \uparrow B$. Note $(B_k \cap F) \uparrow (B \cap F)$ $P(B \cap F) = \lim_k P(B_k \cap F) = \lim_k P(B_k)P(F) = P(B)P(F)$

Aplicando o teor λ-π temos $\mathcal{L} \supset \sigma(\mathcal{A}_1)$. Segue que se $A_1 \in \sigma(\mathcal{A}_1)$ e $A_i \in \mathcal{A}_i$ p/ $2 \leq i \leq n$, então

$$P(\cap_{i=1}^{n} A_i) = P(A_1) P(\bigcap_{i=2}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$$

Daí, se $\mathcal{A}_1, \dots, \mathcal{A}_n$ são ind., $\sigma(\mathcal{A}_1), \mathcal{A}_2, \dots, \mathcal{A}_n$ tbm são.

Itero o argumento!

Teorema 2.1.4. Em ordem para $X_1, \dots, X_n$ ser independentes, é suficiente que p/ todo $x_1, \dots, x_n \in (-\infty, \infty]$

$$P(X_1 \leq x_1, \dots, X_n \leq x_n) = \prod_{i=1}^{n} P(X_i \leq x_i)$$

Prova. Seja $\mathcal{A}_i$ = os cjtos da forma $\{X_i \leq x_i\}$. Uma vez que

$$\{X_i \leq x_i\} \cap \{X_i \leq y\} = \{X_i \leq x \wedge y\},$$

$\mathcal{A}_i$ é um sistema π. Uma vez que permitimos $x_i = \infty$, $\Omega \in \mathcal{A}_i$.
Ex. 1.3.1 implica $\sigma(\mathcal{A}_i) = \sigma(X_i)$, de modo que o resultado segue do Teorema 2.1.3.

Ex. 1.3.1 Se $\mathcal{A}$ gera $\mathcal{S}$ então $X^{-1}(\mathcal{A}) = \{\{X \in A\}, A \in \mathcal{A}\}$ gera $\sigma(X) = \{\{X \in B\}, B \in \mathcal{S}\}$
Para mostrar tal, temos que mostrar que
- $\sigma(X)$ é σ-álgebra $\quad \Omega \in \mathcal{R}$, pois $B \in \mathcal{S}$
- $\sigma(X) \supset X^{-1}(\mathcal{A})$
- $\sigma(X)$ é a menor que contém $X^{-1}(\mathcal{A})$

• Coleções $\mathcal{A}_1, \dots, \mathcal{A}_n \subset \mathcal{F}$ são independentes se sempre que $A_i \in \mathcal{A}_i$ e $I \subset \{1, \dots, n\}$, então

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i).$$

• Se $\mathcal{A}_1, \dots, \mathcal{A}_n$ são independentes, então $\sigma(\mathcal{A}_1), \dots, \sigma(\mathcal{A}_n)$ tbm o são.

Prova. Se $A_2, \dots, A_n$ com $A_i \in \mathcal{A}_i$. Seja $F = \bigcap_{i=2}^{n} A_i$ e

$$\mathcal{L} = \{B : P(B \cap F) = P(B) P(F)\}$$

Note que $\mathcal{L} \subset \mathcal{A}_1$.

Ademais: • $\Omega \in \mathcal{L}$

• Se $A, B \in \mathcal{L}$, $A \subset B$, $B - A \in \mathcal{L}$

• Se $B_n \uparrow B$, $B_n \in \mathcal{L}$, $B \in \mathcal{L}$

Logo, $\mathcal{L} \subset \sigma(\mathcal{A}_1)$.

Itero o argumento!

▲ Para que $X_1, \dots, X_n$ sejam independentes, é suficiente que para todo $x_1, \dots, x_n \in (-\infty, \infty]$,

$$P(X_1 \leq x_1, \dots, X_n \leq x_n) = \prod_{i=1}^{n} P(X_i \leq x_i)$$

Prova. Tomo $\mathcal{A}_i =$ conjuntos da forma $\{X_i \leq x_i\}$.

$\mathcal{A}_i$ é um sistema π

Se $\mathcal{A}_i$ indep. $\overset{\text{Teo Ant.}}{\Rightarrow}$ $\sigma(A_1), \dots, \sigma(A_n)$ indep.

Mas $\sigma(\mathcal{A}_i) = \sigma(X_i) = \{\{X \in B\}, B \in \mathcal{B}\}$

▲ Suponha $\mathcal{F}_{i,j}$, $1 \leq i \leq n$, $1 \leq j \leq m(i)$ independentes e seja $\mathcal{G}_i = \sigma(\cup_j \mathcal{F}_{i,j})$. Então $\mathcal{G}_1, \dots, \mathcal{G}_n$ são independentes.

Prova. Seja $\mathcal{A}_i$ uma coleção de cjtos da forma $\cap_j A_{i,j}$ onde $A_{i,j} \in \mathcal{F}_{i,j}$. $\mathcal{A}_i$ é um sistema π que contém Ω e $\cup_j \mathcal{F}_{i,j}$, assim Teo. 21.3 implica que $\sigma(\mathcal{A}_i) = \mathcal{G}_i$ são independentes.

Independência, Distribuição e Esperança

Teorema 2.1.7. $X_1, \dots, X_n$ v.a.is independentes e X_i tem dist. μ_i.
Então $(X_1, \dots, X_n)$ tem distribuição $\mu_1 \times \dots \times \mu_n$

$$P((X_1, \dots, X_n) \in A_1 \times A_2 \times \dots \times A_n) = P(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) =$$

$$P(X_1 \in A_1) \dots P(X_n \in A_n) = \mu_1(A_1) \dots \mu_n(A_n) = \mu_1 \times \dots \mu_n (A_1 \times \dots \times A_n)$$

A última formula mostra que a distribuição de $(X_1, \dots, X_n)$ e a medida $\mu_1 \times \dots \times \mu_n$ concordam em conjuntos da forma $A_1 \times \dots \times A_n$, um sistema π que gera $\mathbb{R}^n$. Daí, pelo Teorema 2.1.2 segue que devem concordar.

Teorema 1.6.9. Formula da Mudança de Variáveis.

Seja X um elemento aleatório de $(S, \mathcal{S})$ com distribuição μ, i.e.,

$$P(X \in A) = \mu(A)$$

Se f é uma função mensurável de $(S, \mathcal{S})$ à $(\mathbb{R}, \mathcal{R})$ de modo que $f \geq 0$ ou $E|f(X)| < \infty$, então

$$E f(X) = \int_S f(y) \mu(dy)$$

$$* \int_\Omega f(\underbrace{h(\omega)}_{X}) dP = \int_S f(y) \, d(P \circ h^{-1})$$

Prova.

Caso 1: Funções Indicadoras

Se $B \in \mathcal{S}$ e $f = \mathbb{1}_B$, então

$$E \mathbb{1}_B(X) = P(X \in B) = \mu(B) = \int_S \mathbb{1}_B(y) \mu(dy)$$

Probabilidade Avançada I, Aula 1. Tarefas 26/02

Exercicio 1. Seja $\mathcal{F}_\theta$, $\theta \in \Theta$, Θ um espaço de índices, σ-álgebras. então

$$\mathcal{F} = \bigcap_{\theta \in \Theta} \mathcal{F}_\theta \text{ é σ-álgebra.}$$

Sabemos que $\mathcal{F}$ é σ-álgebra em Ω se

1. $\Omega \in \mathcal{F}$, $\emptyset \in \mathcal{F}$

   Como $\Omega, \emptyset \in \mathcal{F}_\theta \ \forall \theta \in \Theta$, segue que $\Omega, \emptyset \in \bigcap_{\theta \in \Theta} \mathcal{F}_\theta$

2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

   $A \in \mathcal{F} \Rightarrow A \in \mathcal{F}_\theta \ \forall \theta \in \Theta \overset{\mathcal{F}_\theta \text{ é σ-alg}}{\Rightarrow} A^c \in \mathcal{F}_\theta \ \forall \theta \in \Theta \Rightarrow A^c \in \bigcap_{\theta \in \Theta} \mathcal{F}_\theta = \mathcal{F} \Leftrightarrow$
   $A^c \in \mathcal{F}$.

3. Se $A_1, A_2, \ldots$ estão em $\mathcal{F}$, então $\bigcup_i A_i \in \mathcal{F}$

   $A_1, A_2, \ldots$ estão em $\mathcal{F} \Rightarrow A_1, A_2, \ldots$ estão em $\mathcal{F}_\theta$, $\forall \theta \in \Theta \Rightarrow$
   $\bigcup_i A_i \in \mathcal{F}_\theta \ \forall \theta \in \Theta \Rightarrow \bigcup_i A_i \in \bigcap_{\theta \in \Theta} \mathcal{F}_\theta \Rightarrow \bigcup_i A_i \in \mathcal{F}$.

Exercicio. Se $A, B \in \mathcal{F} \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Pela definição de $P(\cdot)$, sabemos que se $C_1, C_2 \in \mathcal{F}$ são tais que $C_1 \cap C_2 = \emptyset$, então $P(C_1 \cup C_2) = P(C_1) + P(C_2)$.

A ideia aqui é rescrever $A \cup B$, $A, B \in \mathcal{F}$, como união de conjuntos disjuntos. Basta notar que

$$A \cup B = A \cup (B \cap A^c) \Rightarrow$$
$$P(A \cup B) = P(A) + P(B \cap A^c)$$

Mas $B = (B \cap A^c) \cup (B \cap A) \Rightarrow P(B) = P(B \cap A^c) + P(B \cap A) \Rightarrow$
$$P(B \cap A^c) = P(B) - P(A \cap B).$$

$\therefore$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Generalização. Se $A_1, A_2, \dots, A_n$ estão em $\mathcal{F}$,

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) +$$

$$+ \dots + (-1)^{n+1}\, \mathbb{P}(A_1 \cap \dots \cap A_n)$$

Demonstração. Vimos do anterior que a fórmula acima vale para $n=2$. $(\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2))$. Vamos supor que vale para $n=k$, vejamos se vale para $n=k+1$.

$$\mathbb{P}\left(\bigcup_{i=1}^{k+1} A_i\right) = \mathbb{P}\Big[\underbrace{\Big(\bigcup_{i=1}^{k} A_i\Big)}_{B} \cup A_{k+1}\Big] = \mathbb{P}\left(\bigcup_{i=1}^{k} A_i\right) + \mathbb{P}(A_{k+1}) - \mathbb{P}\left(\bigcup_{i=1}^{k} A_i \cap A_{k+1}\right) =$$

$$= \sum_{i=1}^{n} \mathbb{P}(A_i) + \mathbb{P}(A_{k+1}) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) +$$

$$\dots + (-1)^k\, \mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_k) - \mathbb{P}\left(\bigcup_{i=1}^{k} (A_i \cap A_{k+1})\right) =$$

$$= \sum_{i=n}^{k+1} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \dots + (-1)^k\, \mathbb{P}(A_1 \cap \dots \cap A_k) -$$

$$\Big[\sum_{l=1}^{k} \mathbb{P}(A_i \cap A_{k+1}) - \sum_{i<j} \mathbb{P}(A_i \cap A_j \cap A_{k+1}) + \dots + (-1)^k\, \mathbb{P}(A_1 \cap \dots \cap A_{k+1})\Big] =$$

$$= \sum_{i=1}^{k+1} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) + \dots + (-1)^{k+1}\, \mathbb{P}(A_1 \cap \dots \cap A_{k+1})$$

$i = 1, \dots, k+1$

Exercício. "Baralho embaralhado". N cartas, $N \geq 1$.

$(\Omega, \mathcal{F}, \mathbb{P})$ $\quad \Omega =$ cjto das perm das N cartas $= \{x = (a_1, \dots, a_N), a_i \neq a_j \text{ se } i \neq j, a_i \in \{1, \dots, N\}\}$

A_N : "nenhuma carta está na posição inicial."

$\mathbb{P}(A_N) = 1 - \mathbb{P}(A_N^c)$, A_n^c = "pelo menos uma carta está na posição inicial".

2. Seja $\{Z_n\}_{n\geq 1}$ uma coleção de variáveis aleatórias independentes com $P(Z_n=0)=1-1/n$ e $P(Z_n=1)=P(Z_n=-1)=1/2n$. Defina agora $\{X_n\}_{n\geq 1}$ tal que $X_1=Z_1$ e, para $n\geq 2$

$$X_n=\begin{cases} Z_n & \text{se } X_{n-1}=0, \\ n X_{n-1}|Z_n| & \text{se } X_{n-1}\neq 0 \end{cases}$$

e defina $\mathcal{F}_n=\sigma(X_1,\dots,X_n)$.

a) Mostre que $\{X_n,\mathcal{F}_n\}$ é um martingal

b) Valem as condições do teorema de convergência de martingais vista em aula?

c) X_n converge em probabilidade para alguma v.a.? Qual?

d) X_n converge quase certamente para alguma v.a? Qual?

a) Seja $\mathbb{1}_{k} = \mathbb{1}_{\{X_k=0\}}$. Temos

$$E(X_n|\mathcal{F}_{n-1}) = E\left(Z_n\mathbb{1}_{n-1}+nX_{n-1}|Z_n|(1-\mathbb{1}_{n-1})\,|\,\mathcal{F}_{n-1}\right)=$$

$$=\mathbb{1}_{n-1}\underbrace{E(Z_n)}_{0}+n\,X_{n-1}(1-\mathbb{1}_{n-1})\underbrace{E|Z_n|}_{\frac{1}{n}}=X_{n-1}$$

pois $E(Z_n)=0$ e $E|Z_n|=\frac{1}{n}$. Também $E|X_n|\leq E\{|Z_n|(1+n|X_{n-1}|)\}$ e $E|X_1|<\infty$, de onde $E|X_n|<\infty$. Portanto, $\{X_n,\mathcal{F}_n\}$ é um martingal.

b) c) d)

Agora, $X_n=0 \Leftarrow Z_n=0$. Portanto, $P(X_n=0)=P(Z_n=0)=1-1/n\to 1$ quando $n\to\infty$, implicando que $X_n\xrightarrow{P}0$. Por outro lado, $\sum_n P(Z_n\neq 0)=\infty$, e portanto $P(X_n\neq 0 \text{ i.v.})=1$ pelo seg. Lema B-C. No entanto, X_n toma apenas valores inteiros e portanto X_n não converge à 0 q.c.

O teorema da convergência de martingais é inaplicável uma vez que $\sup_n E|Y_n| = \infty$.