

Probabilidade Avançada, II

Aula 01

11/08/2015

Probabilidade e Esperança Condicionais

Sabemos

1) Dado $(1, 2, \dots, 6)$ com 1, 2, 3 brancos e 4, 5, 6 pretos

Lançar o dado e observar a cor preta. Qual a prob de ter caído na face ímpar?

Ex 2. (X, Y) têm função dens.

$$f_{X,Y}(x,y) = \begin{cases} \dots, 0.5 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, \text{ c.c.} \end{cases}$$

Qual a função densidade da dist. de X dado que $Y = y$ ($y \in [0,1]$)?

Soluções (tot. distintas)

$$1. P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2} = 1/3$$

$$2. f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \text{ onde } f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dx$$

Obs: é menor que momento.

A razão:

No ex., o condicionador tem probabilidade nula.

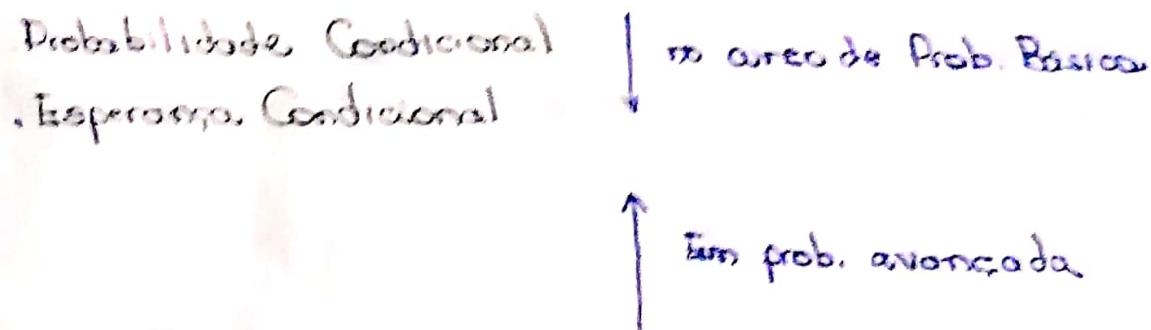
Apresentarão uma definição unificada da qual os casos abordados tot. dif. seguem como casos particulares.

Então tentaremos de definir (2) a partir de (1), fazendo
 $B \in \phi$ de maneira t.q. $(P(B) \neq 0)$

continuidade de $P(\text{no } \phi)$

$B \rightarrow \{Y = y\}$ (Barry James) "Prob. em sub-álgbras"

Exercício: Achar furos nos argumentos do Barry James



Definição: (Ω, \mathcal{F}, P) e \mathcal{G} é uma variável aleatória (\mathcal{G} é integrável).

\mathcal{G} sub- σ -álgebra de \mathcal{F} (\mathcal{G} é σ -álgebra e $\mathcal{G} \subset \mathcal{F}$).

Dai,

$E[\xi | \mathcal{G}]$ é a variável aleatória que sal.

① $E[\xi | \mathcal{G}]$ é \mathcal{G} -mensurável

② $\int_A E[\xi | \mathcal{G}] dP = \int_A \xi dP, \forall A \in \mathcal{G}^{\mathbb{F}}$??

Dúvida: Para que serve a esperança condicional?

Def. essa esp cond. daquela maneira, porque queremos ou porque somos obrigados?

σ -álgebras, variáveis aleatórias, mensurabilidade

\mathbb{R} - espaço de reais, equipado com a métrica tradicional (dist.)
que é separável (\exists subconjunto enumerável que está denso no \mathbb{R})

Por decreto (Por construção, decorre a separabilidade)

Conjuntos abertos (\leftarrow métrica) $\forall a \in A, \exists B(a, \delta) \subseteq A$.

$\mathcal{B}(\mathbb{R})$ é a notação pl. a menor σ -álgebra que contém todos os conjuntos abertos.

↓

\mathcal{B} (qualquer coleção de todos os conj. abertos de um espaço S)

$\equiv \mathcal{B}(S)$ chama-se σ -álgebra de Borel de S

Só $\mathcal{B}(\mathbb{R})$ denota-se por \mathcal{B} .

Definição. de álgebra e σ -álgebra

Uma coleção Σ de subconjuntos de Ω t.q.

1. $\emptyset, \Omega \in \Sigma$
2. Se $A \in \Sigma \Rightarrow A^c \in \Sigma$
3. Se $A, B \in \Sigma \Rightarrow A \cup B \in \Sigma$ ($\therefore A \cap B \in \Sigma$)
4. Se $A_1, A_2, \dots \in \Sigma \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \Sigma$ ($\therefore \bigcap_{n=1}^{\infty} A_n \in \Sigma$)
seq. enumerável

\uparrow 1, 2, 3 \rightarrow álgebra
1, 2, 4 \rightarrow σ -álgebra

Ω , \mathcal{A} é uma coleção de eventos de Ω .

$\sigma(A)$ é a notação para a menor σ -álgebra que contém todos os conjuntos de A .

Exemplo: $\Omega = \mathbb{R}$, $A = \{(-\infty, x], x \in \mathbb{R}\}$ é a coleção de todos os semi-intervalos semi-fechados.
 $\hat{=} \pi(\mathbb{R})$
 \downarrow aqui, a ponto veremos a razão para tal notação

Alego. $\sigma(\pi(\mathbb{R})) = \mathcal{B}(\mathbb{R}) \equiv \mathcal{B}$ (Vej o demonstração no livro de
Ver Prova (Exercício))

Willans, pg. 17)

David Willans, Prob. with Martingals

A construção fracionária de $\sigma(\mathcal{A})$

A visão da σ -álgebra da perspectiva da probabilidade.
Mé ou "conceber-se" $\sigma(\mathcal{A})$ dessa maneira.

Para sermos precisos, se $\mathcal{A} \subset \pi(\mathbb{R})$

Verificamos fato que $\pi(\mathbb{R})$ não é σ -álgebra.

$$\textcircled{1} \quad \bigcup_{n \in \mathbb{N}} (-\infty, x^{-1/n}] = (-\infty, x) \notin \pi(\mathbb{R})$$

Vamos acrescentar todos os intersecções e uniones, enunc. de conjuntos
de $\pi(\mathbb{R})$.

Chame esta coleção por $\pi_1(\mathbb{R})$

Ela é σ -álgebra? Sim - ótimo

Será: rep. o proc. Resultado $\pi_2(\mathbb{R})$

Continua. Pode ser que nunca pare.

Tome $\bigcap_{m=1}^{\infty} \pi_m(\mathbb{R})$ Este é $\sigma(\pi(\mathbb{R}))$

Comentário: $\pi^*(\mathbb{R}) = \{(-\infty, r), r \in \mathbb{Q}\}$

$\pi^*(\mathbb{R})$ é enumerável.

Por outro lado, tomado $r_0 \in \mathbb{R} \setminus \mathbb{Q}$, então

$$\bigcap_{m=1}^{\infty} (-\infty, r_0] = (-\infty, r_0]$$

Isto significa que se ocorresse de $\pi^*(\mathbb{R})$ o fizesse o complemento
só, então já no 2º passo estaria obrigado a acrescentar
um número não-enumerável de conjuntos.

Infelizmente, na maioria dos casos, o procedimento de complemento

são não chega à $\sigma(\mathcal{A})$

Por que? Não lembro (esqueci)

faltado prático, a dificuldade na demonstração da prop. de conjuntos de $\sigma(\mathcal{A})$

$\pi \rightarrow \lambda$ argumento (Williams, Dorret)

Shiryagin (aluno do Kolmogorov). Princípio de conjuntos aproximadores.

Definição. π -sistema. Uma coleção \mathcal{U} chama-se π -sistema se.

i. $A, B \in \mathcal{U}, A \cap B \in \mathcal{U}$

Exemplo: $\pi(\mathbb{R})$ é π -sistema

$$(-\infty, x] \cap (-\infty, y] = (-\infty, \min\{x, y\}] \in \pi(\mathbb{R})$$

Teorema. Se duas medidas coincidem nos conjuntos de um π -sistema, então as duas também coincidem no conjunto $\sigma(\pi\text{-sistema})$.

para entender o Teorema, era preciso introduzir a extensão de medida de \mathcal{A} para $\sigma(\mathcal{A})$. E antes disso, era necessário introduzir medida em geral.

Versão ...

Ex. Aula 01.

2. $f_{x \times y}(x) = \underline{f_{x,y}(x,y)}$, onde $f_y(y) = \int_{-\infty}^{\infty} f_{x,y}(x,y) dx$,
 $f_y(y)$

Não demonstrado!

2. $\sigma(\pi(\mathbb{R})) = \mathbb{B}$

Prova. Para cada $x \in \mathbb{R}$, $(-\infty, x] = \bigcap_{n \in \mathbb{N}} (-\infty, \infty + n^{-1})$, de modo que como união enumerável de intersecções de conj. abertos, o conjunto $(-\infty, x]$ está em \mathbb{B} .

Tudo o que resta é provar que todo subconjunto aberto G de \mathbb{R} está em $\sigma(\pi(\mathbb{R}))$. Mas todo tal G é uma união enumerável (esparso) de intervalos abertos, de modo que precisamos apenas mostrar que, para $a, b \in \mathbb{R}$ com $a < b$,

$$(a, b) \in \sigma(\pi(\mathbb{R}))$$

Mas, para qualquer w com $w > a$,

$$(a, w] = (-\infty, w] \cap (-\infty, a]^c \in \sigma(\pi(\mathbb{R})),$$

e desde que para $\epsilon = \frac{1}{2}(b-a)$,

$$(a, b) = \bigcup_n (a, b - \epsilon n^{-1}).$$

vermos que $(a, b) \in \sigma(\pi(\mathbb{R}))$, e a prova está completa, pois se $A \in \mathbb{B}$, então $A \in \sigma(\pi(\mathbb{R}))$ se $B \in \sigma(\pi(\mathbb{R})) \nsubseteq B \in \mathbb{B}$.

Lista Probabilidade Avançada II

Aula 01.

- Foco no Barry James;
- Construção dos números reais. Identificar se, da construção, decorre a separabilidade
- Rever definições: conjunto aberto, conjunto denso, espaço separável, espaço completo, espaço arquimediano, espaço ordenado.
- Demonstrar que $\sigma(\pi(\mathbb{R})) = \mathcal{B}$, onde $\pi(\mathbb{R}) = \{(-\infty, x], x \in \mathbb{R}\}$ (Pg 37 livro Probability with Martingales, por David Williams)
- Proc. de complementação de $\pi(\mathbb{R})$, p. ex., não chega à uma σ -álgebra! (afirmação a ser demonstrada)
- Relembre o argumento $\pi \rightarrow \lambda$ (Williams, D; Durret, R.)
De lo Shiryaev - princípio dos conjuntos aproximadores;
Extensão da Caratheodory;

Aula 02

- Demonstrar

Lema 1 (Cap. II, §4, Shiryaev) (Ω, \mathcal{F}, P) , $X: \Omega \rightarrow \mathbb{R}$.

Há col. \mathcal{F} de círculos de \mathbb{R} t.q. $\sigma(\mathcal{F}) = \mathcal{B}$ então, se

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

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

(Vejam como o autor fala com σ -alg. - "appropriate sets")

- Ler Cap. II, §4 de Shiryaev até Teo. 3. (Se \mathcal{F} é uma v.a. \mathcal{F} mens, então \mathcal{F} é boreliana t.q. $\mathcal{B} = \{f \circ \mathcal{F}\}$.)
- Recapitular Borel-Cantelli;
- $\mathcal{F}_g = \{g^{-1}(B), B \in \mathcal{B}\} \subseteq \mathcal{F}$ e é σ -álgebra

- Construa $\tilde{g}: \mathbb{R} \rightarrow \mathbb{R}$ t.q a coleção $\{\tilde{g}(B), B \in \mathcal{B}\}$ não é uma σ -álgebra

Seja $\tilde{g}: \Omega \rightarrow \mathbb{R}$
 $w \mapsto \tilde{g}(w) + 2$. Dai,

$$\begin{aligned} A \subset B &\quad \subset \\ A \cup C &\subset B \cup C \\ \text{Ex: } (\Omega, \mathcal{N}, \mathbb{P}) & \quad A \cap C \subset B \cap C \quad A \cap B \subset \subset \\ A \cap B &\subset \left\{ \begin{array}{l} B_{\text{inf}} \\ A_0 \in B_{\text{fin}}, |A| \leq |B| \end{array} \right. \\ \text{SP1.} & \quad A_i \in B_1 \quad (i) \\ \text{SP2.} & \quad A_i \in B_2 \\ \text{SP3.} & \quad A_i \in B_3 \end{aligned}$$

$\emptyset \subset A, A \in \mathcal{I} \quad (\emptyset \subset \Omega)$

SP1. A_1, A_2, \dots com $A_1 \subset A_{n+1}, B \in \mathcal{I}$,

$$B \subset A_n, \forall n, B \subset \bigcap_{m=1}^n A_m$$

$$B = \mathbb{N} \quad A_n = \mathbb{N} / \{0, \dots, n\}$$

$$B \subset A_n \quad M = \bigcup_{n=1}^{\infty} \{0, \dots, n\}$$

$$\bigcap_{m=1}^{\infty} A_m = \emptyset$$

$$\bigcap_{m=1}^{\infty} \emptyset = \bigcap_{m=1}^{\infty} A_m$$

$$B \subset \emptyset$$

$$\begin{cases} A \subset B \\ B^c \subset A^c \end{cases}$$

$A = \text{imparés}$

$$B = \mathbb{N}$$

$$B \subset A$$

$$A^c \subset B^c$$

$$P \subset \emptyset$$

$$\bigcap_{n=1}^{\infty} A_n / \{0, \dots, n\} = \emptyset$$

$$\text{Se } K \in \mathbb{N}, \text{ então } \in \bigcup_{n=1}^K A_n$$

$$0 \in$$

$$1 \in$$

$$\text{Supondo } K \in \mathbb{N}$$

Se $\exists n \in \mathbb{N}, n \in B$,

então $n \in A_n \quad \forall n \in \mathbb{N}$

Se $n \in \mathbb{N}$, então

n deve ser $\in A_{n+2}$ e que não ocorre

Probabilidade Avançada II

Aula 2

18/08/15

Variável Aleatória

$X: \Omega \rightarrow \mathbb{R}$ ($\forall \omega \in \Omega, X(\omega)$ é def. e é um n° real)

função \uparrow \uparrow

$\sigma\text{-alg. } P$ $\sigma\text{-alg.}$
 Borel Borel

P não importa em todos

A função X chama-se variável aleatória, se

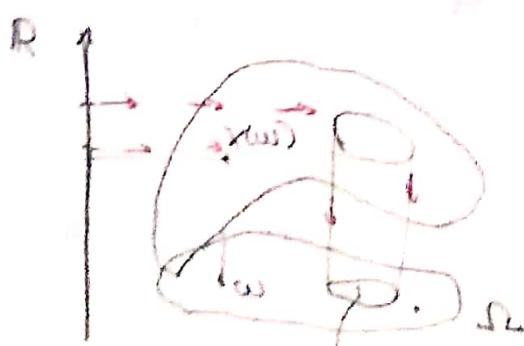
$$X^{-1}(B) \in \mathcal{F} \quad \forall B \in \mathcal{B} \quad (\text{presume-se que há } P \text{ em } \mathcal{F})$$

Caço Particular

$$\begin{array}{c} X: \Omega \rightarrow \mathbb{R} \\ \uparrow B \quad \uparrow B \\ \text{Se} \quad X^{-1}(B) \in \mathcal{B} \end{array}$$

então X chama-se função boreliano (Borel function)

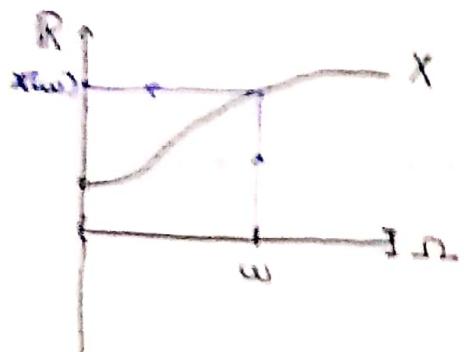
Vizualização



toda $\omega \in \Omega$ que $x(\omega) \in B$

\mathbb{B} é o $(0,1)$ e aberto.

Agora amplifiquemos o desenho, freq. lemniscata num $(0,1)$, $f = B((0,1))$ (σ -alg. de Borel no int. $(0,1)$), é um σ -alg. de todos os subconjuntos em $(0,1)$.



Como verificar que pluma dada função $X: \Omega \rightarrow \mathbb{R}$ é uma função f (σ -alg. de conj. de Ω), $X^{-1}(B) \in f$ para todo $B \in \mathbb{B}$?

Problema. \forall Boreliano?

Lema 1 (Cap II, S.4, Shiryayev)

Há $(\Omega, \mathcal{F}, \mathbb{P})$ e $X: \Omega \rightarrow \mathbb{R}$

Há uma coleção \mathbb{B} de conjuntos de \mathbb{R} t.q. $\sigma(\mathbb{B}) = \mathbb{B}(\mathbb{R}) = \mathbb{B}$.
Então, se

$$X^{-1}(E) \in \mathcal{F}, \forall E \in \mathbb{B}, \text{ então } (*)$$

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

Em pal. simples, para verif. que $X^{-1}(B) \in \mathcal{F} \forall B \in \mathbb{B}$, é suf. verif. que $X^{-1}(E) \in \mathcal{F} \forall E \in \mathbb{B}$ t.q. $\sigma(\mathbb{B}) = \mathbb{B}$.

A ideia é que a verif. de $(*)$ é mais fácil que a verif. de $(**)$.

Veja a demonstração do Lema.

Veja como o autor trabalha com o álgebra.

(lembre: o alg. não é construtivo, então, para que

$$X^*(B) \in \mathcal{V}BE(B)$$

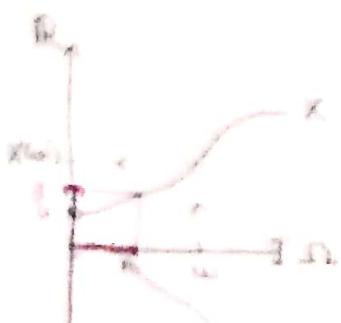
é preciso de um caminho indireto

Shiryayev usa o princípio de conjuntos "appropriate sets".

Exemplo. Sabemos que

$$\sigma(\pi(\mathbb{R})) = \mathcal{B}$$

$\pi(\mathbb{R}) = \{(-\infty, x], x \in \mathbb{R}\}$; a (afegan) menor alg. que contém todos os conj. dessa cl.



menor alg. que contém todos os conj. dessa cl.

$\mathcal{B} \neq \mathcal{A}$

Todo todos os intervalos fechados abertos e semi-abertos de $(0, 1)$ estão em \mathcal{B} .

$$X^*([-\infty, \log]) = \text{este intervalo } S \in$$

Tarefa. Ler Capítulo II, § 4 do Shiryayev até Teorema 3, a ser formulado a seguir.

Teorema 3 Se f é uma fn. d \mathbb{P} mensurável. Então existem funções f_1 e f_2 contínuas tal que $f = f_1 + f_2$.

1. Decidir se este Thm. pl. esp. construtivel.

2. Observar que obtemos imediatamente das bases da Teoria de Prob. o resultado que vimos no final.

Feliz sob E.O.

O significado de $\eta = f \circ \xi$.

Esq. de dizer que ha (Ω, \mathcal{B}, P) ; da $\xi: \Omega \rightarrow \mathbb{R}$ e ta $\eta: \Omega \rightarrow \mathbb{R}$ e ambas v.a.

f cuja exist. esta alegada aqui é vma $f: \mathbb{R} \rightarrow \mathbb{R}$

$f^{-1}(B) \in \mathcal{B}$ $\forall B \in \mathcal{B}$ (f é Boreliana)

$$\eta = f \circ \xi \Leftrightarrow \eta(\omega) = f(\xi(\omega)) \quad \forall \omega \in \Omega$$

Explicarei o significado de \mathcal{F}_ξ e " η é \mathcal{F}_ξ mensurável"

Considere a coleção de pré-imagens dos eventos da σ -álgebra de \mathcal{B} por funç. ξ .

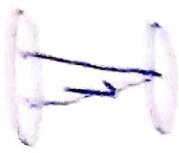
$$\{\xi^{-1}(B), B \in \mathcal{B}\}$$

①. Esta col. é σ -álgebra (Veja Ex 1-3)

②. Esta σ -álgebra de nota-se \mathcal{F}_ξ

Ex 1. $\xi: \Omega \rightarrow \mathbb{R}$ cons. $\{\xi^{-1}(B), B \in \mathcal{B}\}$ e prove que é σ -álgebra de \mathcal{B} (de eventos de Ω que pode ser mais pobre que \mathcal{F}).

Ex 2. Construa $\xi: \mathbb{R} \rightarrow \mathbb{R}$ t.q. a coleção $\{\xi(B), B \in \mathcal{B}\}$ não é uma σ -álgebra.



Ex. 3. Seja $f: \mathbb{R} \rightarrow \mathbb{R}$ e seja \mathcal{S} na $(f^{-1}(B))_{\mathcal{S}}, \forall B \in \mathcal{B}$

Consideraremos uma coleção \mathcal{A} de conj. de sub. tal que

$\mathcal{A} = \{f^{-1}((-\infty, x]), x \in \mathbb{R}\}$ em outras pal.

$$\mathcal{A} = f^{-1}(\pi(\mathcal{B})).$$

Prove que:

①. Se é π-sistema (elegendo verdade)

②. $\sigma(\mathcal{A}) = \mathcal{F}_f$

* import.

$$\begin{array}{ccc} \pi(\mathcal{B}) & \xrightarrow{\delta} & \mathcal{B} \\ \delta \downarrow & & \downarrow f^{-1} \\ \mathcal{A} & \xrightarrow{\sigma} & \mathcal{F}_f \end{array}$$

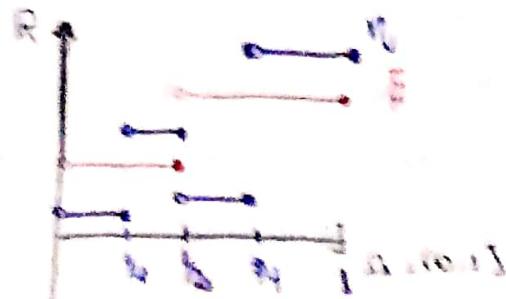
Notanda à 1.

Se chama-se \mathcal{F}_f mensurável se

$$(f^{-1}(B))_{\mathcal{F}_f} \forall B \in \mathcal{B}$$

Isso quer dizer \mathcal{F}_f é mais forte que f cumpre esta cond. por definição

Variabilizações



os átomos da partição

$D_\xi = \{ (0; \frac{1}{2}], (\frac{1}{2}, 1] \}$ - a partição gerada por ξ

$D_\eta = \{ (0; \frac{1}{4}), (\frac{1}{4}; \frac{1}{2}], (\frac{1}{2}, \frac{3}{4}], (\frac{3}{4}; 1] \}$ partição gerada por η

$\mathcal{F}_\xi = \{\emptyset, \Omega, (0; \frac{1}{2}], (\frac{1}{2}, 1]\}$

$\mathcal{F}_\eta = \{\emptyset, \Omega, (0; \frac{1}{4}], \dots, (\frac{3}{4}; 1], \{(0; \frac{1}{4}] \cup (\frac{1}{4}; \frac{1}{2}]\}, \dots\}$

↓
todas as reuniões dos
átomos da part. cão D_η .

Neste caso ξ e \mathcal{F}_η mensurável, mas η não é \mathcal{F}_ξ mensurável.

Cada átomo de D_ξ pode ser composto (reuniao) de alguns átomos de $D_\eta \Rightarrow$ Cap ξ shirayev.

De volta à def. assustadora da Esp. Condiciona!

$E(\xi | G)$ é G -mensurável

$$\int_A \xi dP = \int_A E(\xi | G) dP, \forall A \in G. \quad (*)$$

B_1	$\&$
B_2	
B_3	
B_n	

$$P(A) = \sum_{i=1}^n P(A|B_i) P(B_i) \quad \text{fórmula prob. tot. no caso de espaço \& discreto.}$$

Como espaço discreto:

1) Define-se a prob. cond. $P(A|B)$

2) Demonstra-se a fórmula (*), que é, uma das propriedades da prob. condicional.

No nosso curso, esta prop. torna-se como definição da esp. cond.
Quer dizer

$$\int_A g dP = \int_A E(g|G) dP, \forall A \in G \text{ é o "equiv." da fórmula}$$

As dif. da explicação:

- 1) No curso de prob. Básica, define-se $P(\cdot|B)$ e, com uso dela, define-se $E(X|B)$
Enq. no curso, def. $E[g|G] \in P(A|G) \triangleq E[1_A|G]$

Probabilidade Avançada II,

Aula 03

18/08/15

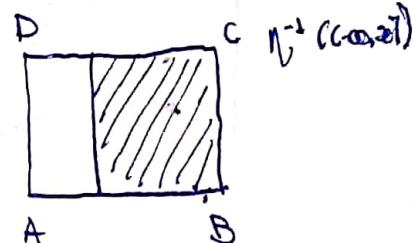
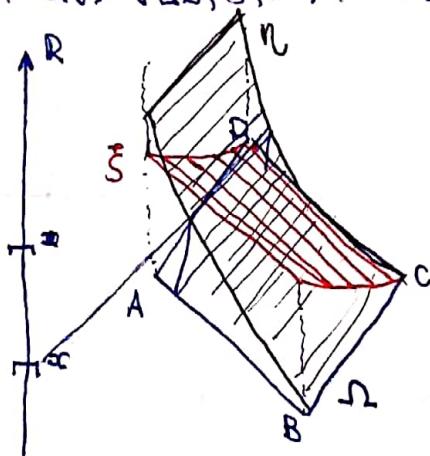
$\int_A \xi dP = \int_A E[\xi | \mathcal{G}] dP \quad \forall A \in \mathcal{G}$ define $E[\xi | \mathcal{G}]$, de maneira t.q. ela torna a ser fórmula da prob. total. (Esp por material)

Exemplo. $\xi \in \mathbb{I}$ duas v.v. em (Ω, \mathcal{F}, P) . Vou construir $E[\xi | \mathcal{G}_N]$.

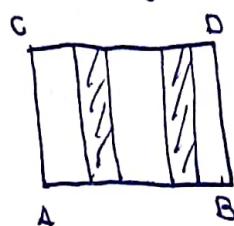
$$\Omega = [0,1] \times [0,1]$$

$$\mathcal{F} = \mathcal{B}([0,1] \times [0,1])$$

P. a medi. de Lebesgue



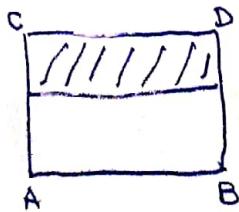
Todos os $f_N^{-1}((-\infty, \infty])$
tem esta cara \Rightarrow
 \mathcal{G}_N tem esta cara



$$\mathcal{G}_N = \{B \times [0,1], \forall B \in \mathcal{B}([0,1])\}$$

(σ-álgebra cilíndrica)

\mathcal{F}_ξ - σ-álg. em Ω gerada por ξ .



$$\mathcal{F}_\xi = \{[0,1] \times B, \forall B \in \mathcal{B}([0,1])\}$$

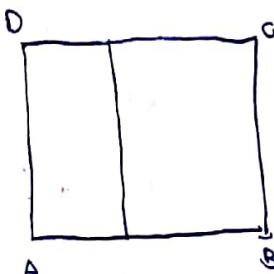
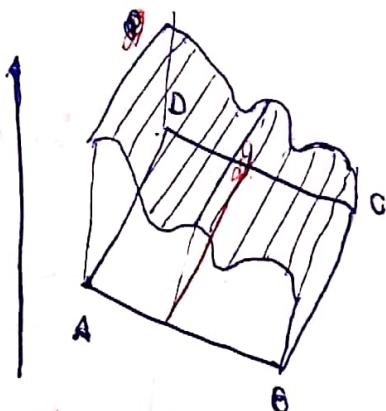
Quero visualizar

$$\mathbb{E}[\xi | \mathcal{G}_n](\omega)$$

1. É uma variável aleatória.

2. \mathcal{G}_n mensurável.

$\Rightarrow \mathbb{E}[\xi | \mathcal{G}_n]$ é uma "chapeadoada" no dir. de AB.

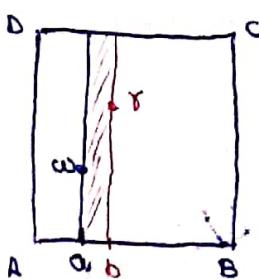


P (de Leb., no caso) dá pes. 0 ao $a \in [0,1] = Y \in \mathcal{G}_n$

$$\int_V \xi dP = \int_V \mathbb{E}[\xi | \mathcal{G}_n] dP$$

não define explicitamente o valor de $\mathbb{E}[\xi | \mathcal{G}_n]$ no V, pois este pode ser qqer um, já que sendo mul. por $P(V)$, dá zero.

Berry James



Vou calcular $\mathbb{E}[\xi | \mathcal{G}_n]$ no $[a,b] \times [0,1]$ e depois b $\rightarrow a$ e achar o limite.

$$\frac{\mathbb{E}[\xi | \mathcal{G}_n]}{(b-a)}$$

Problema.

$$\mathbb{E}[\xi | \mathcal{G}_n](\omega) \neq \mathbb{E}[\xi | \mathcal{G}_n](\gamma)$$

↑
pode ser

Acionamos nessa indução, que diz que, sendo que b está muito próximo de a , os valores de $\mathbb{E}[\xi | \mathcal{G}_n]$ são "basicamente" os mesmos em toda a faixa $[a, b] \times [0, 1]$

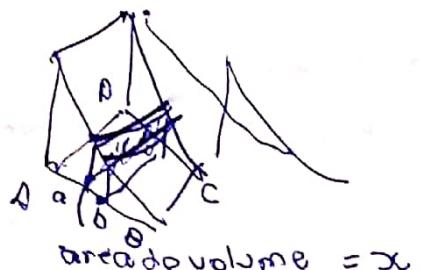
Chama este valor de x

$$\int_{[a, b] \times [0, 1]} \xi(\omega) dP = \int_{[a, b] \times [0, 1]} x dP = x \lambda([a, b] \times [0, 1]) = x(b-a)$$

Barry James

$$\lim_{b \rightarrow a} \frac{\int_A \xi^f dP}{b-a}, \quad A = [a, b] \times [0, 1]$$

$\therefore \mathbb{E}[\xi | \mathcal{G}_n](\omega), \omega \in \Omega \times [0, 1]$



Exercício. Vamos calcular

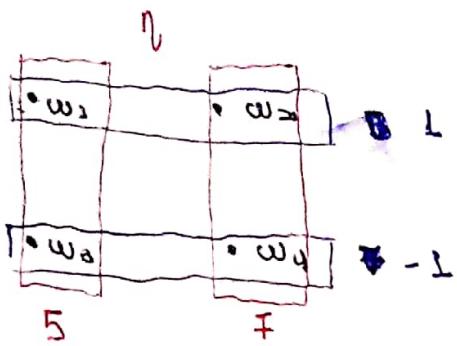
$\mathbb{E}[\xi | \mathcal{G}_n]$ no caso discreto usando a fórmula

$$\int_A \xi dP = \int_A \mathbb{E}[\xi | \mathcal{G}_n] dP.$$

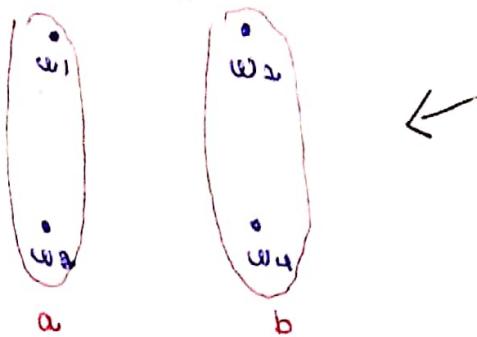
$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $\mathcal{F} = \sigma(\Omega)$, onde $\Omega = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}$

$$P(\omega_i) = \frac{i}{4}, \quad i=1, 2, 3, 4.$$

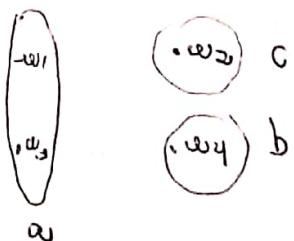
5



$$E[S | \mathcal{G}_n]$$



$E[S | \mathcal{G}_n]$ - esta não é \mathcal{G}_n mensurável



Tome (*) eache, com o uso deste, os val. de a e b.

Filtragem.

Filtragem \rightarrow tempo de parada \rightarrow T. Martingal e num martingal.
 \downarrow martingais \nearrow parada

Tomo X_1, X_2, Y_1 (e X_4 se precisar) talis que elas são

1) independentes

$$2) P[X_1=1] = P[X_1=-1] = \frac{1}{2}.$$

Comment: Y_1, Y_2, Y_3, \dots

Dizer que são ind. sign. todas as subcole

$\{Y_1, Y_2, \dots, Y_n\} \subset \{Y_1, \dots, Y_n\}$ são indep.
 $\cap = \emptyset$.

Comentário: Existe (Ω, \mathcal{F}, P) e $X_1, X_2, X_3 : \Omega \rightarrow \mathbb{R}$
 satisfaçõe que 1 e 2 valem?

Resposta: Sim. É o espaço produto!

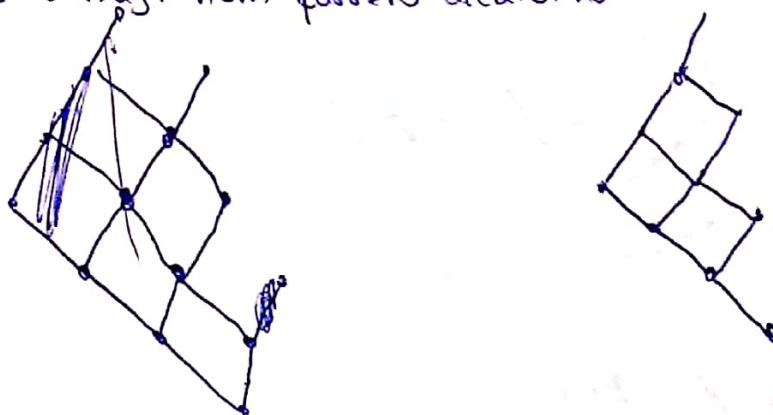
$$\Omega = \{HHH, \dots, TTT\}, \quad P[\omega] = 1/8, \quad \forall \omega \in \Omega.$$

$$X_i(\omega) = \begin{cases} 1 & \text{se a prim. destra do } \omega \text{ for } H \\ -1 & \text{se } T \end{cases} \quad X_2 \text{ e } X_3 \text{ ind.}$$

Defino S_0, S_1, S_2, S_3 da seg. forma:

$$\left. \begin{array}{l} S_0 \triangleq 0 \\ S_1 \triangleq S_0 + X_1 = X_1 \\ S_2 \triangleq S_1 + X_2 = X_1 + X_2 \\ S_3 \triangleq S_2 + X_3 = X_1 + X_2 + X_3 \end{array} \right\} \quad \begin{array}{l} \text{Constrói um processo aleatório não simples simétrico} \\ \text{de compri. 3.} \end{array}$$

Aqui são todas as 8 traj. num processo aleatório



Trajetória $(S_0(\omega), S_1(\omega), S_2(\omega), S_3(\omega))$

$$\Omega \ni \omega \mapsto$$

Nesta construção, Δ pode ser identificado com

$$\{\wedge, \wedge^*, \vee, \vee^*, \neg, \neg^*\}$$

Quero ver ver $D(S_1)$ - a partição gerada por S_1 ($\sigma(S_1) = \sigma(D(S_1))$)

Como: todos os valores α que S_1 pode assumir

$$f(S_1) = \{S_1^{-1}(\alpha), \text{ todos que } S_1 \text{ pode dar}\}$$

Quando S_1 assume valor 1:

$$\{\wedge, \wedge^*, \vee, \vee^*\}$$

Quando S_1 ass. valor -1

$$\{\vee, \vee^*, \wedge, \wedge^*\}$$

$D(S_2)$

$$2 \{\wedge, \wedge^*\}$$

$$0 \{\vee, \vee^*, \neg, \neg^*, \vee, \vee^*\}$$

$$-2 \{\wedge, \wedge^*\}$$

$D(S_1, S_2)$

$$1, 2 \{\wedge, \wedge^*\}$$

$$1, 0 \{\vee, \vee^*\}$$

são os átomos de $D(S_1, S_2)$

$$-1, 0 \{\vee, \vee^*\}$$

$$-1, -2 \{\wedge, \wedge^*\}$$

$D(S_0)$, $D(S_0, S_1)$, $D(S_0, S_1, S_2)$, $D(S_0, S_1, S_2, S_3) \rightarrow$ filtragem

$D(S_0)$, $D(S_1)$, $D(S_2)$, $D(S_3) \rightarrow$ não é filtragem!

"processo em paralelo"

Probabilidade Avançada II

Aula 04

20/08/15

- Martingais; Tempo de Parada; Martingais Paradas

Da aula passada:

$$S_0, S_1, S_2, S_3$$

fizeram aleatório simples simétrico.

Construimos

$\mathcal{D}(S_0), \mathcal{D}(S_0, S_1), \mathcal{D}(S_0, S_1, S_2), \mathcal{D}(S_0, S_1, S_2, S_3)$
sequência de partições geradas pelo processo S_0, \dots, S_3 .

Esta seq. de partições é um ex. daquilo que chamam-se "filtrações".

Definição de filtração. Temos (Ω, \mathcal{F}) . Uma sequência de sub- σ -álgebras de $\mathcal{F}, \mathcal{F}_0, \mathcal{F}_1, \dots$ chama-se filtração (do espaço Ω) caso

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots \subseteq \mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \dots$$

" $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ se $B \in \mathcal{F}_n \rightarrow B \in \mathcal{F}_{n+1}$, \mathcal{F}_n é mais grossa de \mathcal{F}_{n+1} ,
 \mathcal{F}_{n+1} é mais fina (refinada) que \mathcal{F}_n "

"As vezes começam a contar com 0, às vezes com 1 ($\mathcal{F}_1, \mathcal{F}_2, \dots$)"

"A seq. de sub- σ -álgebras pode ser finita"

Quando há um processo S_0, S_1, \dots qualquer, a seq. de σ -álgebras

"processo em pal. simples: coleção de v.a. defi no mesmo espaço, se qual será denotada por (Ω, \mathcal{F}) ".

$\sigma(S_0), \sigma(S_0, S_1), \dots, \sigma(S_0, S_1, \dots, S_n), \dots$ chama-se a filtragem (do ω) gerada pelo processo.



é a filtragem pois a condição $\sigma(S_0, \dots, S_n) \subseteq \sigma(S_0, \dots, S_n, S_{n+1})$ é naturalmente satisfeita.

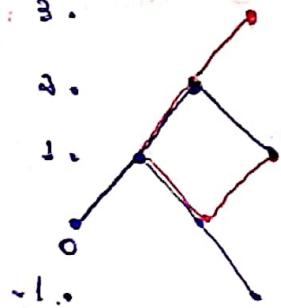
"Quando preciso pensar numa filtragem gerada pelo proc., tomo um proc. com n finito e penso $\mathcal{D}(S_0), \dots, \mathcal{D}(S_0, \dots, S_n)$.

"Na literatura, $\sigma(S_0, \dots, S_n)$ chama-se a inf. cont. no processo até o tempo n ".

No âmbito do exemplo acima, vamos entender o nome "inf. cont. no proc. até o tempo n " e logo vamos melhorá-lo.

Pergunta: Sabemos que S_1 assume valor 1, e que podemos dizer sobre a trajetória de todo o proc. S_0, S_1, S_2, S_3 ?

Resposta:



está num dos 2 átomos de $\mathcal{D}(S_0, S_1)$.

$\mathcal{D}(S_0, S_1)$ é tudo que pode ser revelado por observações dos valores de S_0 e S_1 , quer dizer, dos valores do proc. até tempo 1.

Conclusão: $\sigma(S_0, \dots, S_n)$, no caso geral, deve ser chamado por conjuntos de informações (sobre o and. do proc. todo), t.q. ao observar o proc. até n , você possa ident. un. eferm. destes conj.

Definição. (de tempo de parada) Ω -finito.

Tome $\mathcal{T} : \Omega \rightarrow \{0, 1, \dots, n\}$, randomas part.

Tome uma filtragem " D_0, D_1, \dots, D_n "

\mathcal{T} chama-se tempo de parada em relação desta filtragem, se

$$\{\omega \in \Omega : \mathcal{T}(\omega) \leq k\} \in D_k, \quad k = 0, 1, \dots, n$$

$$\{\omega \in \Omega : \mathcal{T}(\omega) = k\} \in D_k, \quad k = 0, 1, \dots, n$$

Caso Particular: Se há processo S_0, S_1, \dots, S_n t.q. $D_n = \mathcal{D}(S_0, \dots, S_n)$, então \mathcal{T} chama-se tempo de parada em relação ao proc. S_0, \dots, S_n .

Tome S_0, S_1, S_2, S_3 PASS. Define \mathcal{T} como o primeiro momento quando S_i é maior que o anterior. "Daqui a pouco, prec. compl. esta def."

Vamos construir \mathcal{T} explicitamente.

$\mathcal{T} : \text{Cto traj. PASS}_3 \rightarrow \{0, 1, 2, 3\}$

$\omega \in \Omega \quad \mathcal{T}(\omega)$

$\omega \in \Omega \quad \mathcal{T}(\omega)$



1



1



2



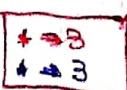
1



3



1



→ Tanto liberdade, limitada pelo sentido do \mathcal{T} e pelo emprego pret. para \mathcal{T} .

* Ex. das moedas

* Def. de tempo de parada.

8

Suponha que τ entende-se assim: apost. 1 em cada moeda e saia do jogo, logo que ganhar pela 1^a vez.

Suponha que lança-se 3 moedas só!

Suponha que quero que τ seja tempo de parada em relações à $\mathcal{D}(S_0), \mathcal{D}(S_0, S_1), \dots, \mathcal{D}(S_0, S_1, S_2, S_3)$. Nesse caso, $\tau = 3$ quando .

Observação Importante. Tome S_0, S_1, \dots PASS. Define τ da mesma maneira:

Tome ω , t.q. $S_0(\omega) \rightarrow S_1(\omega) \rightarrow S_2(\omega) \rightarrow \dots$ (há um único destino). Como completar $\tau(\omega)$? Estender

Definição genérica de tempo de parada: (Ω, \mathcal{F}) - quaisquer S_0, S_1, \dots numa filtragem $\mathcal{F}: \Omega \rightarrow \{\emptyset, 1, 2, \dots, \# \}$ as vezes, chama-se fbb, mas isto chama-se t. de parada em rel. f/filt. Poderá confr. com ex. Sc.

$$\{\omega \in \Omega : \tau(\omega) \leq k\} \in \mathcal{F}_k \quad \forall k \geq 1.$$

†

$$\{\omega \in \Omega : \tau(\omega) = k\} \in \mathcal{F}_k, k = 0, 1, \dots$$

Quero mostrar que τ do meu ex. é um tempo de parada com relações à filtragem " $\mathcal{D}(S_0), \mathcal{D}(S_0, S_1), \dots, \mathcal{D}(S_0, S_1, S_2, S_3)$ ".

Preciso verif. que

$$\{\omega : \tau(\omega) = 1\} \in \mathcal{D}(S_0, S_1)$$

$$\{\omega : \tau(\omega) = 2\} \in \mathcal{D}(S_0, S_1, S_2)$$

$$\{\omega : \tau(\omega) = 3\} \in \mathcal{D}(S_0, S_1, S_2, S_3)$$

Interprete $\tau(\omega) = k$ como "na traj. ω para no tempo k ". Então τ é tempo de parada em relação ao proc. S_0, S_1, \dots (que), se para saber se parar ou não no tempo k é suf. saber os valores do proc. até (inclusive) tempo k ."

Martingal. Tem-se Ω , finito, e filtragem " $\mathcal{D}_0, \dots, \mathcal{D}_n$ ". Um processo $\xi_0, \xi_1, \dots, \xi_n$ chama-se martingal em relação à filtragem se:

1. ξ_k é \mathcal{D}_k -mensurável;
2. $E[\xi_{k+1} | \mathcal{D}_k] = \xi_k, k=0, \dots, n-1$.

Se $\mathcal{D}_k = \mathcal{D}(\xi_0, \dots, \xi_k)$, então ξ_0, \dots, ξ_n chama-se martingal se

3. $E[\xi_{k+1} | \mathcal{D}_k] = \xi_k, k=0, \dots, n-1$
(pq, nat., ξ_k é $\mathcal{D}(\xi_0, \dots, \xi_k)$ -mensurável)

Entendendo:

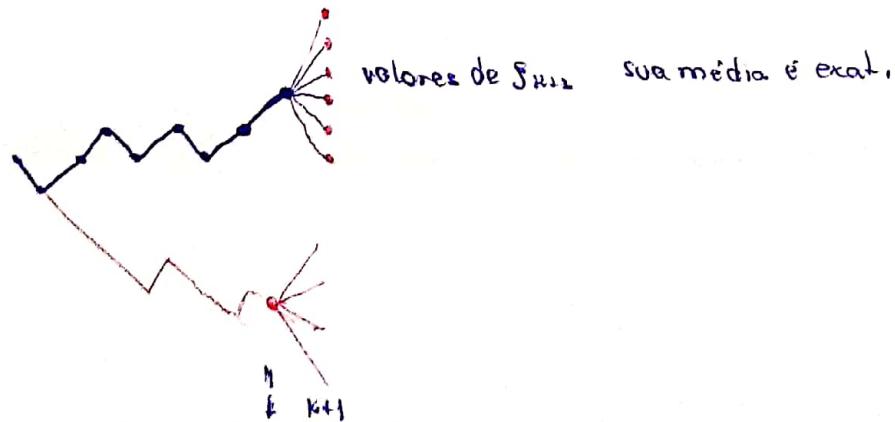
ξ_n é v.a., $E[\xi_{k+1} | \mathcal{D}_k]$ é v.a. $\xrightarrow{\text{(1)}} E[\xi_{k+1} | \mathcal{D}_k](\omega) = \xi_k(\omega) \forall \omega \in \Omega$.

ω : traj. do processo

$\xi_k(\omega)$: o valor do proc. na traj. ω no tempo k .

\mathcal{D}_k é uma partição de Ω . Em cada atomo de \mathcal{D}_k , sabemos os valores do proc. até (inclusive) o tempo k .

Tome um atomo de \mathcal{D}_k . Chame este de A .



Probabilidade Avançada II

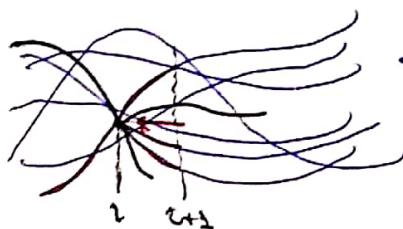
Aula 05

25/08

Martingal: ξ_1, \dots, ξ_n chama-se martingal em relação à "filtragem" $\mathcal{D}_1, \dots, \mathcal{D}_n$ de partição, se ξ_i é \mathcal{D}_i -mensurável e

$$E[\xi_{n+1} | \mathcal{D}_i] = \xi_i, i=1, \dots$$

Desenho da martingalidade

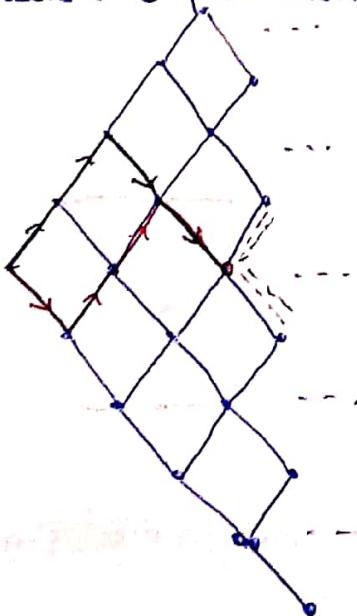


Traj. do proc.

em média, o processo fica no mesmo nível.

As cores indicam a possibilidade de dependência do passado

Ligado a *, existe o pior exemplo do martingal: é o PASS;



PASS é um Proc. de Markov
⇒ o fut. dep. só do presente.

Em geral, martingais não são abordados a serem processos de Markov.

~~~~~

Parece que Martingal só pode "crescer em cada ponto"



com aumento de  $t$ , o n.º de Val. de  $\xi_t$  cresce.

Vej. um exemplo de mart. que não "cresce" em n.º de seus valores

Exemplo:  $X_1, X_2, X_3$  i.i.d. t.q.  $P[X_1=1]=P[X_1=0]=\frac{1}{2}$

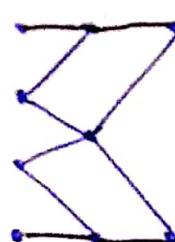
$$S_1 \triangleq X_1, S_2 \triangleq S_1 + X_2 = X_1 + X_2, S_3 \triangleq S_2 + X_3 = X_1 + X_2 + X_3$$

$\xi_1 = \frac{S_3}{3}, \xi_2 = \frac{S_2}{2}, \xi_3 = \frac{S_1}{1}$ . Não é martingal em relação à seq.  
 $D_1 = D(S_1), D_2 = D(S_1, S_2), D_3 = D(S_1, S_2, S_3)$ ,  
pois, p.e.,  $S_1$  não é  $D_1$ -monotônico.

No entanto, o é em relação à seq.  $D_1 = D(S_1), D_2 = D(S_3, S_2), D_3 = D(S_3, S_2, S_1)$ . (Basta verif.  $E[\xi_{...} | D_i] = \xi_i$ .)

Antes, vamos ver def. do mart. construído.

| $X_1$   | $X_2$ | $X_3$ | $S_3/3$ | $S_2/2$ | $S_1/1$ |   |
|---------|-------|-------|---------|---------|---------|---|
| 0, 0, 0 | 0/3   | 0/2   | 0       | 0       | 1       | + |
| 0, 0, 0 | 1/3   | 1/2   | 1       | 0       | 0       | + |
| 0, 1, 0 | 1/3   | 1/2   | 0       | 0       | 0       | 0 |
| 0, 0, 1 | 1/3   | 0/2   | 0       | 0       | 0       | 0 |
| 1, 1, 0 | 2/3   | 0/2   | 1       | 0       | 0       | 0 |
| 1, 0, 1 | 2/3   | 1/2   | 1       | 0       | 0       | 0 |
| 0, 1, 1 | 2/3   | 1/2   | 0       | 0       | 0       | 0 |
| 1, 1, 1 | 2/3   | 2/2   | 1       | 0       | 0       | 0 |



Ex. Verificar  $E[\xi_{...} | D_i] = \xi_i$ .

**Plano:** Usar este arg. e Teo. Mart. periodo para provar Ballot Theorem.

**Teorema:** Tomo uma 'filtração'  $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n$ . Seja  $S_1, \dots, S_n$  um martingal em relação à filtração e seja  $\tau$  um tempo de parada em relação à mesma filtração. Então

$$\mathbb{E}[\xi_2 | D_2] = \xi_2.$$

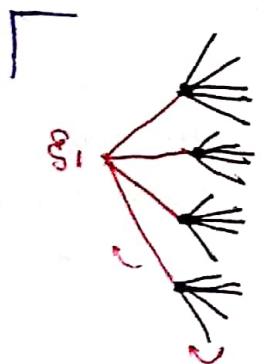
## Observaçõe:

$$\mathbb{E}[\xi_{\text{ad}} | \mathcal{D}_t] = \xi_t \quad (\text{prior } \{\xi_t\} \text{ & marginal})$$

$$\mathbb{E}[\xi_3 | \mathcal{D}_2] = \mathbb{E}[\mathbb{E}[\xi_3 | \mathcal{D}_2] | \mathcal{D}_1], \quad \mathbb{E}[\xi_2 | \mathcal{D}_1] = \xi_1.$$

→ (Indulgāns, 1.ter.?)

$$\begin{aligned} \mathbb{E}[\xi_n | \mathcal{D}_n] &= \mathbb{E}[\mathbb{E}[\xi_n | \mathcal{D}_{n-1}] | \mathcal{D}_n] = \mathbb{E}[\xi_{n-1} | \mathcal{D}_n] = \\ &= \mathbb{E}[\mathbb{E}[\xi_{n-1} | \mathcal{D}_{n-2}] | \mathcal{D}_n] = \dots = \mathbb{E}[\xi_2 | \mathcal{D}_n] = \xi_1. \end{aligned}$$



$$E[S_0] = E[E[S_0 | \mathcal{D}_0]] = E[S_0] = \dots$$

$$\xi_T(w) \triangleq \xi_{t(w)}(w) \quad \forall w \in \Sigma$$

↑  
Ultra noise.  
↓  
Proc.

$$\Gamma \vdash E(x/\beta)$$

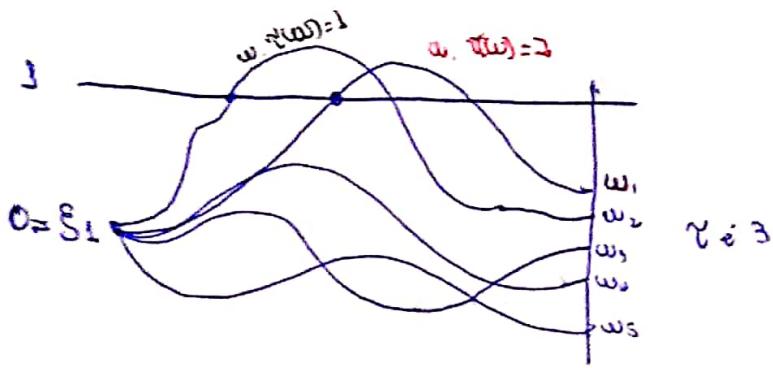
$$E(E(x/\beta)/\beta) =$$

$$E(x/\beta)$$

def

alpha conversion

P é a primeira vez que E<sub>g</sub> alcança o nível I (ou o que for no final)



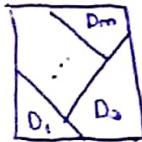
Suponha

$$\mathcal{D}_1 = \{D_1, D_2, \dots, D_m\}$$

Aqui partici<sup>o</sup><sup>n</sup>

$$E[\xi_\gamma | D_i] = g_i \Leftrightarrow E[\xi_\gamma | D] = \xi_\gamma(\omega) \quad \forall \omega \in D, \quad \forall D \in \mathcal{D}_1.$$

v.o.      v.o.      valor



$$\Gamma \vdash P(\cdot) \rightarrow E\xi$$

↓

$$P(\cdot | D) \rightarrow E[\xi | D] \quad \text{Na Avançado E, a partir de } E[\xi | D]$$

↓

↓

$$P[\cdot | D] \rightarrow E[\xi | D]$$

for \$\xi = \mathbf{1}\_A\$.

Tomo \$D\$ gerado de \$D\_1\$ e preciso provar que

$$E[\xi_\gamma | D] = \xi_\gamma(\omega) \quad \forall \omega \in D$$

$$\xi_\gamma = \sum_{i=1}^n \xi_i \mathbf{1}_{\{\tau_i = \gamma\}}$$

$$E[\xi_\gamma | D] = \frac{E[\xi_\gamma \mathbf{1}_A]}{P(D)} = \frac{1}{P(D)} \sum_{i=1}^n E[\xi_i \mathbf{1}_{\{\tau_i = \gamma\}} \mathbf{1}_D]$$

$$= \frac{1}{P(D)} \sum_{i=1}^n E[E[\xi_i | D] \mathbf{1}_{\{\tau_i = \gamma\}} \mathbf{1}_D]$$

$$\begin{aligned}
 &= \frac{1}{P(D)} \sum_{n=1}^{\infty} E[E[\xi_n \mathbb{1}_{\{Z=n\}} \mathbb{1}_D | \mathcal{G}_n]] \\
 &= \frac{1}{P(D)} \sum_{n=1}^{\infty} E[\xi_n \mathbb{1}_{\{Z=n\}} \mathbb{1}_D] \\
 &= \frac{1}{P(D)} E[\xi_n \mathbb{1}_D], \quad \text{markiert}
 \end{aligned}$$

Este é o voto para  $n = \infty$ . Quais são as alterações de fato que  $n = \infty$ ?

Conseq. unique application: "Ballot Theorem"

Dosso: Decaptação de carbonato

$$\xi_1, \dots, \xi_n \text{ i.i.d. com } P(\xi_i = \pm 1) = 1/2,$$

Som a e b para que  $a+b=0$  e  $a+b=n$

Isto garante que, no def. PASS

$$w_i > 0, \quad \delta_i = \delta_{0i} + \dots + \delta_{ii}, \quad i=1, \dots, n.$$

prontimmo que So pode assumir a-b

Queremos mostrar que

$$P(S_1 > 0, S_2 > 0, \dots, S_{n-1} > 0, S_n > 0 | S_n = \overset{\circ}{a} - b) = \frac{a-b}{a+b}$$

d.h. vek. ist nicht eingeschränkt

+ und - Zeichen

n

$\overset{\circ}{a} - b$

$$P(s_1 > 0, \dots, s_n > 0 | S_n = a - b) = P(s_1 < 0, \dots, s_n < 0 | S_n = -(a - b)) =$$

$$= P(S_1 + 1 \leq 2, S_2 + 2 \leq 2, \dots, S_{n-a} + n \leq n) = P(S_1 + \dots + S_{n-a} = n - (a-b)) =$$

$$\max \left( 0; 1 - \frac{b-a}{n} + \frac{(a-b)}{n} \right) = \frac{a-b}{n}$$

Sejam  $\eta_1, \eta_2, \dots, \eta_n$  i.i.d., assumindo valores não negativos (com dist. arb.)  
Defina

$$M_k = \eta_1 + \dots + \eta_k, \quad k=1, \dots, n$$

Então

$$P[M_k \leq k \text{ para todo } k=1, \dots, n | M_n] = \max \left( 0, 1 - \frac{M_n}{n} \right)$$

## Probabilidade Avançada II

Aula 06

24/08/15

Ler Capítulo 3, Williams

$(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0}^{\infty}, P)$ ,  $\{\mathcal{F}_n\}_{n=0}^{\infty}$  é uma filtragem

→ Não tem obrig de ser σ-algebra  $\Rightarrow$  à press. pendurar o σ nela

$\mathcal{F}_{\infty} := \sigma(\bigcup_{n=0}^{\infty} \mathcal{F}_n)$  - para que?

①  $\mathcal{F}_{\infty}$  não é σ-algebra candal!

$\mathcal{F}_{\text{candal}} := \bigcap_{n=0}^{\infty} \mathcal{F}_n$  é σ-algebra

② Preciso dela para definir corretamente o tempo de parada (a) e a v.a. limite (b) aper. correto.

2o. T é tempo de parada em relação  $\{\mathcal{F}_n\}_{n=0}^{\infty}$ , se T assume os valores  $\{0, 1, \dots\}$  e

$$\{\omega : T(\omega) = n\} \in \mathcal{F}_n \quad \forall n$$

Sabemos que em certas situações faz sentido tempos de parada que em alguns ω's não assumem nenhum valor inteiro (gê-gê vivo exemplo).

Suponha que desejamos considerar tais tempo de parada

$$\{1, 2, \dots, \# \}$$

Naturalmente

$$\{\omega : T(\omega) = \# \} = \Omega \setminus \{\omega : T(\omega) = 0\} \setminus \{\omega : T(\omega) = \infty\} \setminus \dots$$
$$= \Omega \setminus \left( \bigcup_{n=0}^{\infty} \{\omega : T(\omega) = n\} \right)$$

$$\{\omega : T(\omega) = \#\} \stackrel{?}{=} \bigcup J_n$$
$$\sigma(\bigcup J_n) = \emptyset$$

$T \in \{0, 1, 2, \dots, \infty\}$  é tempo de parada se

$$\{\omega : T(\omega) = n\} \in \mathcal{G}_n, n \in \{0, 1, \dots, \infty\}$$

" $\infty$ " é símbolo! Tanto é que se  $M_0, M_1, \dots$  for um processo, então  $M_T$  faz sentido só para  $\omega$ 's onde  $T(\omega) \neq \infty$ .

As vezes, " $\infty$ " torna-se uma notação muito comoda. Por exemplo, se for possível. Por exemplo, se for possível definir

$$M_\infty = \lim_{n \rightarrow \infty} M_n \text{ então podemos concordar}$$

que  $M_T$  significa  $M_\infty$  onde  $T = \infty$ .

ab) Termos processos  $M_0, M_1, M_2, \dots$  onde  $M_i$  é fi-mensurável.  
Em algumas situações, existirá

$$\lim_{n \rightarrow \infty} M_n(\omega) \text{ para quase todo } \omega \in \Omega.$$

Nestes casos, diremos  $M_\infty$  de acordo:

$$M_\infty(\omega) := \lim_{n \rightarrow \infty} M_n(\omega), \text{ a.s.}$$

(Acho que)  $M_\infty$  é  $\sigma$ -menorável.

Nas fics S.

Recorda começamos com  $(\Omega, \mathcal{S}, \mathcal{B}(\mathbb{R}), P)$ , logo não temos nada a f.

Processo previsível, estratégia chamada "martingal", martingal parada.

PASS servirá de exemplo.

$$X_1, X_2, X_3, \dots \quad \text{e} \quad P[X_1 = 1] = 1/2$$

$\downarrow$

$S_1 = S_0 + X_1, \dots$

Interpretação:  $X_i$  é o lançamento de uma moeda no inst.  $i$ . Pelo meu modelo (\*), não há lançamento nenhô 0, se ele não existir, moeda de ruim vai acontecer"

Outra historinha: Suponha "aposta C na i-ésima moeda"

"Estratégia": algo fixo (determinado) entre os lançamentos

Ex: A postar 1 em cada 1 dos 10 primeiros lanç. e não apostar nada a partir do 11º. Como calif. a estrat. do ex.?

$$1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, \dots$$

$$C_0, C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}, C_{11}, \dots$$

$$C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}, C_{11}, C_{12}, \dots$$

V.A.

Seq.  $\{C_n\}$  chama-se estratégia. São sentidos considerar estratégias exequíveis:

A quantidade da aposta no  $i$ -ésimo lançamento, pode dep. se dos resultados dos lançamentos  $1, 2, \dots, i-1$ .

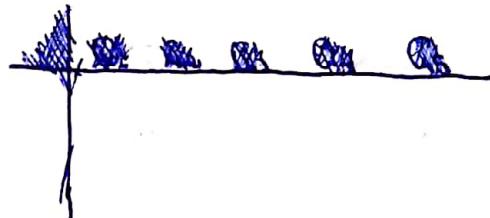
Estratégia chamada "martingal"

Apostar 1 no primeiro lançamento

Apostar ~~dobro~~ no lans.  $i$  o dobro da aposta no  $i-1$  se  $1^o, 2^o, \dots, (i-1)$  deram "-1".

"Sair do jogo" após a primeira "+1".

Apost. 0.



|                        | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $C_6$ |     |
|------------------------|-------|-------|-------|-------|-------|-------|-----|
| -2, 1, -1, ...         | 1     | 2     | 0     | 0     | 0     | 0     | ... |
| -2, -1, 1, -1, ...     | 1     | 2     | 2     | 4     | 0     | 0     | ... |
| -3, -2, -1, 1, -1, ... | 1     | 2     | 2     | 4     | 8     | 0     | ... |

Exequibilidade de estratégia  $\{C_n\}$  significa:

$C_n \in J_{n-1}, \quad 1, 2, \dots, n-1$ .

Podemos dizer que  $C_n$  significa a aposta no  $n$ -ésimo lançamento. Com esta escolha, a expressão da exequibilidade de estrat. torna-se clara:  $C_n \in J_{n-1}$

$\downarrow$   
 $C_n$  não existe

$\downarrow$   
motiva o Mo.

Uma estratégia exequível  $\{C_n\}_1^\infty$  em relação à  $\{\mathcal{F}_n\}_0^\infty$  chama-se "processo previsível" (predictable process):

$$C_n \in \mathcal{F}_{n-1}, n=1, 2, \dots$$

1. Tem-se  $S_0, S_1, S_2, S_3, \dots$

$$X_1 = S_1 - S_0, X_2 = S_2 - S_1, X_3 = S_3 - S_2, \dots$$

Os resultados dos lançamentos são os res. do proc. que avançam o tempo.

2. O ganho acumulado (até inclusive) o lançamento  $n$ , oriundo da estratégia  $\{C_n\}_0^\infty$  torna-se

$$C_0 X_0 + C_1 X_1 + \dots + C_n X_n = C_0 (S_1 - S_0) + \dots + C_n (S_{n+1} - S_n).$$

Definição:  $(\omega, \mathcal{F}, \{\mathcal{F}_n\}, P)$ .

$\{M_n\}_{n=0}^\infty$  é um processo adaptado à  $\{\mathcal{F}_n\}_{n=0}^\infty$  ( $M_n \in \mathcal{F}_n$  memoriável)

$\{C_n\}_{n=1}^\infty$  é um processo previsível em relação a  $\{\mathcal{F}_n\}_{n=0}^\infty$

Defina  $Y_n = C_0 (M_0 - M_0) + C_1 (M_1 - M_0) + \dots + C_n (M_n - M_{n-1}), n=1, 2, \dots$

Então  $\{Y_n\}_{n=1}^\infty$  chama-se a transformação de  $M$  por  $C$ .

Notação:  $Y = C \bullet M$ .

As vezes, quando comodo, podemos introduzir  $Y_0 = 0$ .

Teorema: Se  $\{M_n\}$  for martingal (+ condições), então  $\{Y_n\}$  também é martingal.

Se  $\{M_n\}$  for sub (sup)-martingal (+ condições), então  $\{Y_n\}$  também é sub (sup)-martingal.

Tempo de parada pode ser expresso em termos de proc. previsível.

$$\{M_n\}_{0}^\infty, T \Rightarrow \{M_{T \wedge n}\}_{0}^\infty \neq M_T$$

$\{M_n\}$  martingal  $\Rightarrow$  com é martingal  $\rightarrow E[M_{Tn}] = E[M_0]$

$$\Rightarrow E[M_T] = E[M_0]$$

# Probabilidade Avançada I

Aula 07

01/09/2015

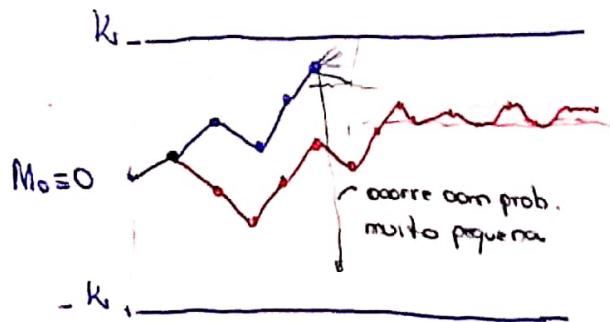
Teorema de Conv. pl Martingais (Cap. II Williams - emenda)

"Teorema" Se  $\{M_n\}_{n=0}^{\infty}$  for um martingal e se  $|M_n| \leq k \forall n$  então pl todos  $w \in \Omega$ , existe

$$\lim_{n \rightarrow \infty} M_n(w)$$

e este limite é finito. ( $\leq k$ ).

Porque ocorre



Precisamos provar, que pl cadas  $w$ ,  $\{M_n(w)\}_{n=0}^{\infty}$  é uma seq. convergente.

Ou conv. para a se  $\forall \epsilon > 0 \exists N(\epsilon)$  t.q.

$$\forall n > N(\epsilon), |a_n - a| < \epsilon.$$

$\forall \epsilon > 0$ , o número de  $n$ 's para os quais  $|a_n - a| > \epsilon$  é finito.

↓

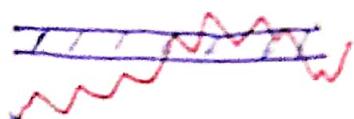
O conjunto de  $w$ 's nos quais não ocorre a converg. de  $M_n(w)$  para  $M_\infty(w)$  (o limite), é

$$\{w : |M_n(w) - M_\infty(w)| > \epsilon \text{ infinitas vezes}\}$$

O objetivo da dem. é provar que  $D$  deste conjunto é nula.

Neste problema da dem. há um prob: Moço não é cont.

$\mu_1$



$\{w : \exists \text{fase}(w) \text{ tq } (\text{Mo}(w))\}_{w=0}^{\infty}$  orça  
ato para inf. vetos

$\mu_2$

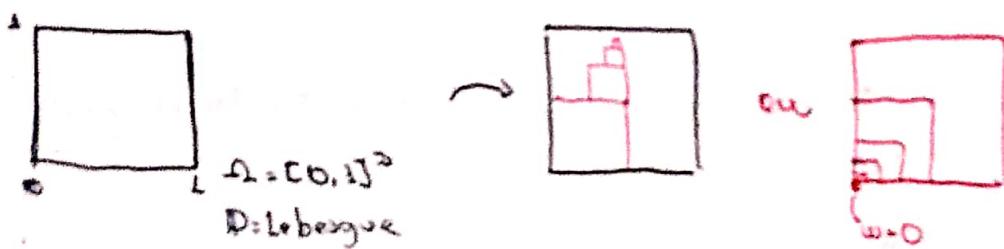
A idéia da dem. é provar que  $D$  deste conjunto é 0.

A conv. q.o. é o resultado do

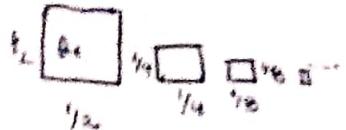
Teor. da Conv. de Mont. e do  
Lema de Borel-Cantelli.

Quem é mais forte? Talvez n<sup>o</sup> haja ligação entre os dois?

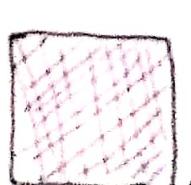
Lema de Borel-Cantelli.



1) Se  $\sum P(A_i) \leq c < \{A_i, i > 0\}$  tem Rnula

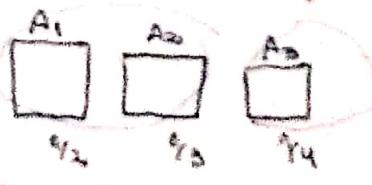


2)



3)

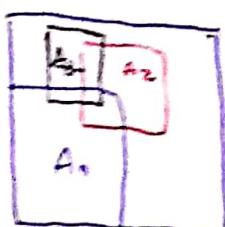
divergente



Queremos que o  
conjunto  $\{w : A(w) > 0\}$   
tenga menor medida.

Queremos que o csto  $P\{\omega : A_i(\omega) \text{ i.o.}\}$  seja menor possível.

E se force usar ind.?



prop. de P.

Ainda dependência obriga os eventos se posicionarem de uma forma que não é arbitrária.

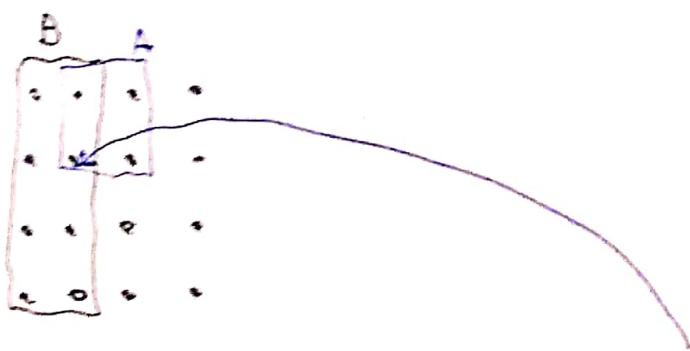
B+C. prova que esta forma garante  $P\{\omega : A_i \text{ i.o.}\} =$

Existem outras formas de dependência organizadas pela P que garantem

$P\{\omega : A_i \text{ i.o.}\} = \frac{1}{2}$  ?

Erdoes-Kac

Petrov



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

Se outra Q dí  $\frac{1}{2}, \dots$ , do pto de vista de Q, A e B são dep.

$$P(AB) = \frac{1}{16}$$

2. Se  $\sum P(A) = \infty$  e se  $A_1, A_2, \dots$  são ind. B.C. megarante que consigo cobrir  $\mathbb{R}$  inf. vezes. (a prob. X está env.).

$(\Omega, \mathcal{F}, P)$   $\xi_n : \Omega \rightarrow \mathbb{R}$ ,  $n=1, 2, \dots$ ,  $\xi : \Omega \rightarrow \mathbb{R}$ .

$$\xi_n(\omega) \rightarrow \xi(\omega), \forall \omega \in \Omega \rightarrow \text{B.env. } P.$$

$$\exists \dots \sum P\{\omega : |\xi_n(\omega) - \xi(\omega)| > \epsilon\} < \infty \quad \text{A.S.T.O. envolve Piamarro \& S.} \quad 26$$

tal forma que que ocorre a conv. em cada  $\omega$ .

Estamos à procura da rel. mais básica entre a convergência de  $\{\xi_n\}_{n=1}^{\infty}$  por uma medida e a propriedade " $\{\xi_n(\omega)\}_{n=1}^{\infty}$ " com A.U.

↑  
Shiryaev Cap. II, Seção 10 "Div. tipo de conv."

**Teorema I.** Para que  $(\xi_n \rightarrow \xi \text{ P.q.c}) \quad \xi_n(\omega) \rightarrow \xi(\omega) \text{ em cada } \omega \in \Omega$  é necessário e suficiente que

$$\forall \epsilon > 0, \quad P\{\omega : \sup_{k \geq n} |\xi_k(\omega) - \xi(\omega)| \geq \epsilon\} \xrightarrow{n \rightarrow \infty} 0.$$

A dem. do Teo. I contém as ideias da dem. do B.C. para o teorema de Conv. de Martingais.

$$A_n^\epsilon = \{\omega : |\xi_n(\omega) - \xi(\omega)| \geq \epsilon\}; \quad A^\epsilon = \overline{\lim}_n A_n^\epsilon = \limsup_n A_n^\epsilon = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m^\epsilon$$

$$\text{Então } \{\omega : \xi_n(\omega) \neq \xi(\omega)\} = \bigcup_{\epsilon > 0} A^\epsilon = \bigcup_{m=1}^{\infty} A'^{1/m}$$

$$\text{Mas, } P(A^\epsilon) = \lim_{n \rightarrow \infty} P\left(\bigcup_{m=n}^{\infty} A_m^\epsilon\right). \text{ Por isto,}$$

$$0 = P\{\omega : \xi_n(\omega) \neq \xi(\omega)\} = P\left(\bigcup_{\epsilon > 0} A^\epsilon\right) \Leftrightarrow P\left(\bigcup_{m=1}^{\infty} A'^{1/m}\right) = 0 \Leftrightarrow$$

$$P(A'^{1/m}) = 0, \forall m \in \mathbb{N} \Leftrightarrow P(A^\epsilon) = 0, \forall \epsilon > 0 \Leftrightarrow$$

$$\Leftrightarrow P\left(\bigcup_{n=1}^{\infty} A_n^\epsilon\right) = 0, \quad n \rightarrow \infty \Leftrightarrow P\{\omega : \sup_{k \geq n} |\xi_k(\omega) - \xi(\omega)| \geq \epsilon\} \rightarrow 0, \quad n \rightarrow \infty$$

$(\Omega, \mathcal{F}, \mathbb{P})$ ,  $\xi_n: \Omega \rightarrow \mathbb{R}$ ,  $n=1, \dots$

$\{\xi_n\}_1^\infty$  chama-se seq. de Cauchy de v.a. se  $\{\xi_n(\omega)\}_1^\infty$  é Cauchy  $\forall \omega$ .

↓  
I Cauchy que é conv. no esp., mas todo de Cauchy é conv. No esp completo - não é só  
vde que é de Cauchy e conv.

↓

Teorema 2. Para que  $(\xi_n \rightarrow \xi)$  P-q.c) é necessário e suficiente que

$$\forall \epsilon > 0 \quad \mathbb{P}\left\{\omega: \sup_{n \geq n_0} |\xi_n(\omega) - \xi(\omega)| \geq \epsilon\right\} \xrightarrow{n \rightarrow \infty} 0,$$

↓

$$\Leftrightarrow \forall \epsilon > 0, \quad \mathbb{P}\left\{\omega: \sup_{k \geq n} |\xi_k(\omega) - \xi(\omega)| \geq \epsilon\right\} \xrightarrow{n \rightarrow \infty} 0$$

↑

$$\Leftrightarrow \forall \epsilon > 0, \quad \mathbb{P}\left\{\omega: \sup_{k \geq 0} |\xi_{n+k}(\omega) - \xi_n(\omega)| \geq \epsilon\right\} \xrightarrow{n \rightarrow \infty} 0$$

# Probabilidade Avançada II

Aula 09

15/09/2011

Martingais Uniformemente Integráveis (Cap. 18 Williams, Durrel alternativo)

Teorema: Se  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ , ( $X \in L^1$  s.f.  $\mathbb{E}[|X|] < \infty$ )

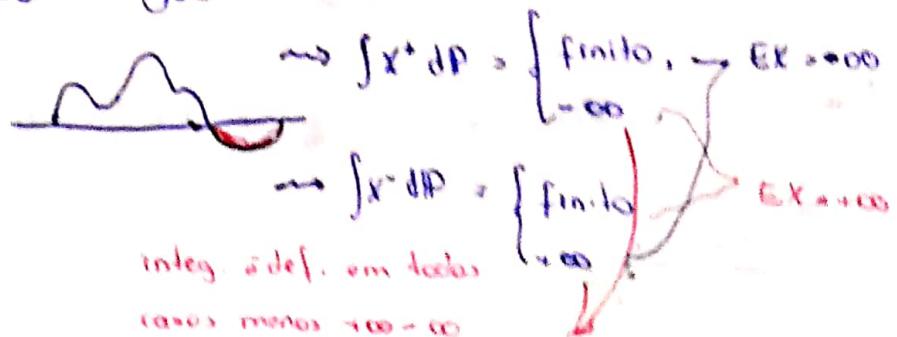
Então para cada  $\delta > 0$ ,  $\exists \delta_0 > 0$  t.q.  $\forall F \in \mathcal{F}$  com  $\mathbb{P}(F) \leq \delta$ , ocorre que

$$\mathbb{E}[|X| \mathbf{1}_F] \leq \delta.$$

**Comentário:** sobre a not.  $\mathbb{E}[X] = \infty$  e  $\mathbb{E}[|X|] = \infty$

Records a const. da int. de Lebesgue

$$X \begin{cases} X^+ = \max\{0, X\} \\ X^- = \min\{0, X\} \end{cases}$$

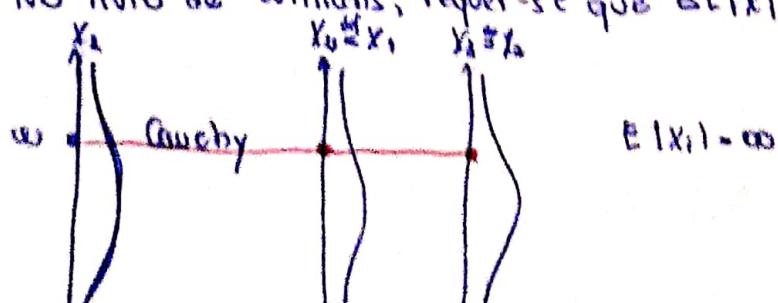


(Cont. com, Exemplo:  $f(x) = \max\{0, x\}$ )

Se  $\mathbb{E}(X) = \pm\infty$ , então  $\mathbb{E}(X/G)$  pode ser definida.

Se  $\mathbb{E}(X)$  não é definida, não se define  $\mathbb{E}(X/G)$ .

No livro de Williams, requer-se que  $\mathbb{E}[|X|] < \infty$  p/ definir  $\mathbb{E}(X/G)$ .



$X_2$  n Cauchy

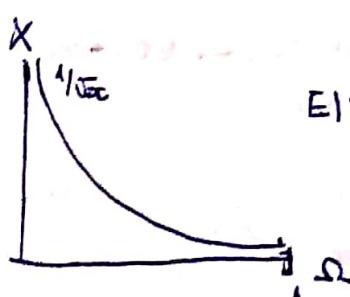
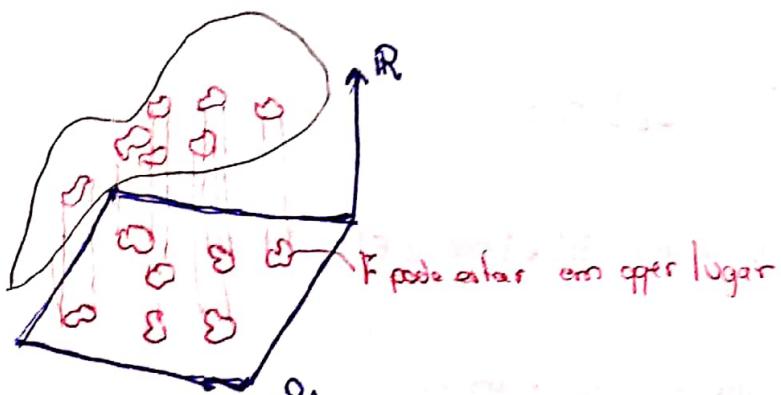
$$X_2/X_1 = \infty = \begin{cases} x+2, \text{ c.p. } 1/2 \\ \infty-2, \text{ com. p. } 1/2 \end{cases}$$

não é Cauchy

$X_1$  n nataura Cauchy  $\Rightarrow E[X_1] = \infty$

$$X_0/X_{n+1} = \infty = \begin{cases} x+1, \text{ c.p. } 1/2 \\ \infty-1, \text{ com. p. } 1/2 \end{cases}$$

Voltando ao Lema de Fatou



$$E[X] = 2x^{1/2} \Big|_0^1 = 2 \Rightarrow X \in \ell^1((0,1], \mathcal{B}, \text{Lebesgue})$$

Dem: Suponha o contrário. Então,  $\exists \epsilon_0 > 0$  s.t. q.  $\exists \{F_n\}_{n=1}^{\infty}$ , com  $R(F_n) = 1/2^n$  e  $E[|X|1_{F_n}] \geq \epsilon_0$ .

$H := \limsup_n F_n \equiv \{w: F_n(w) \text{ infinitas vezes}\} = B.C \quad P(H) = 0 \Rightarrow E[|X|1_H] = 0$

$\cap \cup F_n$ .

Pelo reverse Lema de Fatou,  $E[|X|1_H] \geq \epsilon_0$ .

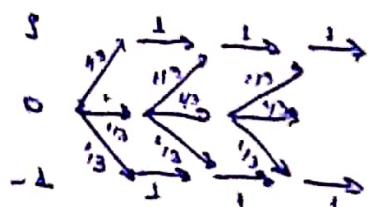
onde  $\mu$  é posit. e integrável

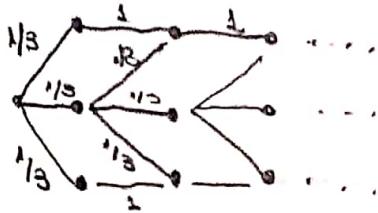
(Williams) Fato. Se  $\{f_n\}$  em  $(m\Sigma)^*$ , então

$$\mu(\liminf f_n) \leq \liminf \mu(f_n)$$

Fato Inverso: Se  $\{f_n\} \in (m\Sigma)^*$  e  $g \in (m\Sigma)^*$  t.q.  $f_n(x) \leq g(x) \forall x \in \Sigma$

$$\mu(\limsup f_n) \geq \limsup \mu(f_n).$$





$$X_0 = 0$$

$$X_1 = \begin{cases} 1, \text{ com prob. } 1/3 \\ 0, \text{ com prob. } 1/3 \\ -1, \text{ com prob. } 1/3 \end{cases}$$

$$X_2 = \begin{cases} 2, \text{ com prob. } 1/9 \\ 1, \text{ com prob. } 1/3 \\ 0, \text{ com prob. } 1/3 \\ -1, \text{ com prob. } 1/3 \\ -2, \text{ com prob. } 1/9 \end{cases}$$

$$X_n = \begin{cases} 1 & \text{se } X_1 = 1 \\ 1 - \text{com prob. } 1/3 & \text{se } X_1 = 0 \\ 0 & \text{se } X_1 = 0 \\ -1 & \text{se } X_1 = -1 \\ -1 - \text{com prob. } 1/3 & \text{se } X_1 = 0 \\ -2 & \text{se } X_1 = -2 \\ -2 - \text{com prob. } 1/9 & \text{se } X_1 = -1 \\ -1 & \text{se } X_1 = -1 \\ -1 - \text{com prob. } 1/3 & \text{se } X_1 = -2 \\ -2 & \text{se } X_1 = -2 \\ -2 - \text{com prob. } 1/9 & \text{se } X_1 = -1 \end{cases}$$

Exercício:  $(X_i)_{i \geq 0}$  é um martingal. Ademais

$$X_n \rightarrow X_\infty = \begin{cases} 1 \text{ c.p. } 1/2, & \text{quase certamente} \\ -1 \text{ c.p. } 1/2 \end{cases}$$

Solução

$(X_n)_{n \geq 0}$  será um martingal com relação à filtragem  $\{\mathcal{F}_n\}_{n \geq 0}$  se

$$1. X_n \in \mathcal{F}_n$$

$$2. E[X_{n+1} / \mathcal{F}_n] = X_n, \forall n \geq 0.$$

Tomando  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$  (a menor  $\sigma$ -álgebra que for  $(X_0, \dots, X_n)$  mensurável), segue a validade de (1) pois,  $\forall B \in \mathcal{B}(\mathbb{R})$ ,

$$\{X_n \in B\} = \{X_n \in B, X_k \in \mathbb{R}, k=0, \dots, n-1\} \in \mathcal{F}_n.$$

Basta averiguar (2). De fato,

$$E[X_{n+1} / \mathcal{F}_n] = \begin{cases} 1, \text{ com prob. } 1/3 + 1/9 + \dots + (1/3)^n \\ 0 = \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 1, \text{ com prob. } (1/3)^n \\ -1, \text{ com prob. } 1/3 + 1/9 + (1/3)^n \end{cases} := X_n$$

$$\mathcal{F}_1 = \left\{ \left\{ \begin{array}{c} D_1 \\ \diagdown \quad \diagup \end{array} \right\}, \left\{ \begin{array}{c} D_2 \\ \diagup \quad \diagdown \end{array} \right\}, \left\{ \begin{array}{c} D_3 \\ \diagup \quad \diagup \end{array} \right\} \right\}$$

$$\mathcal{F}_{20} = \left\{ \left\{ \begin{array}{c} 1,1 \\ \diagdown \quad \diagup \end{array} \right\}, \left\{ \begin{array}{c} 0,1 \\ \diagup \quad \diagdown \end{array} \right\}, \left\{ \begin{array}{c} 0,-1 \\ \diagup \quad \diagup \end{array} \right\}, \left\{ \begin{array}{c} 0,0 \\ \diagdown \quad \diagdown \end{array} \right\}, \left\{ \begin{array}{c} -1,-1 \\ \diagup \quad \diagdown \end{array} \right\} \right\}$$

$$E[X_{\infty} | \mathcal{F}_2] = \begin{cases} E[X_{\infty} | D_1] = 1 \cdot \text{em } D_1, \text{ prob} = 1/3 \\ E[Y_{\infty} | D_2] = + \text{ em } D_2, \dots 1/3 \\ E[X_{\infty}, D_3] = -1 \text{ em } D_3, \dots 1/3 \end{cases}$$

$$\begin{aligned} * &= 1 \cdot P(X_{\infty}=1) + 0 \cdot P(X_{\infty}=0) + (-1) P(X_{\infty}=-1) = \\ &= P(Y_{\infty}=1) - P(X_{\infty}=-1) = 0, \text{ pela simetria.} \end{aligned}$$

$$\begin{aligned} &\stackrel{\text{exp}}{=} \left( \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \dots \right) - \left( \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \dots \right) \\ &= \left( \frac{1}{3} \cdot \frac{1}{1-\frac{1}{3}} \right) + \left( \frac{1}{3} \cdot \frac{1}{1-\frac{1}{3}} \right) = 1/2 - 1/2 = 0 \end{aligned}$$

$$Y_{\infty} = \begin{cases} -1, \text{ com prob } 1/2 \\ 1, \dots 1/2 \end{cases}$$

$$\begin{aligned} P(X_{\infty}=1) &= P(X_n=1, \text{ para algum } n) \\ &= \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \dots = \frac{1}{2}. \end{aligned}$$

# Probabilidade Avançada II

Aula 10

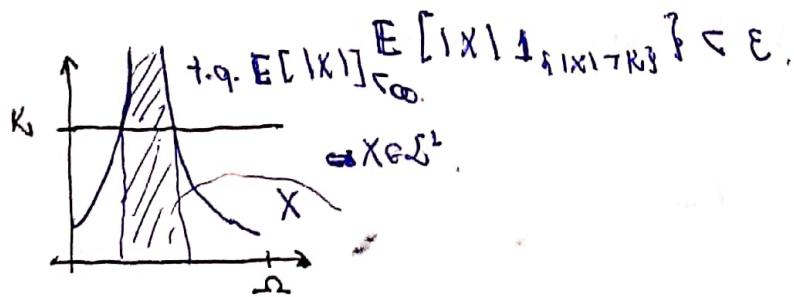
17/09/2015

Integrabilidade Uniforme (de uma família)

Uma só v.a.

Lema: Se  $\in L^1$  ( $E[|X|] < \infty$ ), então  $\forall \epsilon > 0$ ,  $\exists \delta > 0$  t.q.  $\forall x \in \mathbb{R}$  com  $P(|x| > \delta) < \epsilon$ , tem que  $E[|x| \mathbf{1}_{\{|x| > \delta\}}] < \epsilon$ .

Corolário: Se  $X \in L^1$ , então  $\forall \epsilon > 0$ ,  $\exists K \in [0, +\infty)$  t.q.



Uma família de v.a.

Def. Uma classe  $\mathcal{F}$  de variáveis aleatórias chama-se uniformemente integrável se  $\forall \epsilon > 0$   $\exists K \in [0, \infty)$  t.q.

$$E[|x| \mathbf{1}_{\{|x| > K\}}] < \epsilon, \forall X \in \mathcal{F}.$$

Sem nenhuma relação com martingais.

Pergunta:

Qual a relação das relações entre  $X_n \xrightarrow{P} X$  e  $X_n \xrightarrow{d} X$ .

$\forall \epsilon > 0$ ,  $P[|X_n - X| \geq \epsilon] \xrightarrow{n \rightarrow \infty} 0$   $\Rightarrow E[|X_n - X|] \xrightarrow{n \rightarrow \infty} 0$ .

Verifique: há ex. de  $\{X_n\}$  que conv. em  $P$  mas não em  $L^1$ .

Teorema. Seja  $\{X_n\}$  uma seq. t.q.  $X_n \in L^1 \forall n$ .

(nomenclatura:  $\{X_n\}$  é uma seq. em  $L^1 \Rightarrow \exists K \in \mathbb{R} \text{ s.t. } E[|X_n|] \leq K \forall n$ .  
 $\Rightarrow E[|X_n|] \leq K \forall n$ .

Afirmativa:

$$E[|X_n - X|] \xrightarrow{n \rightarrow \infty} 0 \Leftrightarrow \begin{cases} (i) X_n \xrightarrow{P} X \\ (ii) \{X_n\} \text{ é unif. integrável} \end{cases}$$

$\{X_n\}$  martingal + condições:  $X_n \xrightarrow{P} X_0$ .

Teorema (de Doob) de convergência de Martingais. Suponha que  $\{X_n\}_{n=0}^\infty$  é mart. limitado em  $L^1$ .

(nomenclatura:  $\{X_n\}_{n=0}^\infty$  é limitada em  $L^1 \Rightarrow \exists K \text{ t.q. } E[|X_n|] \leq K \forall n$ , ou, em outras pal.,  $\sup_n E[|X_n|] < \infty$ ).

Então  $\exists X_0$  t.q.  $X_n \xrightarrow{P} X_0$ .

Tomaremos  $\{X_n\}_{n=0}^\infty$  um martingal, que ao mesmo tempo, uma seq. unif. integrável.

$\hookrightarrow \{X_n\}_{n=0}^\infty$  limitado em  $L^1 \xrightarrow{\text{Doob}} X_n \xrightarrow{P} X \rightarrow X_n \xrightarrow{P} X_0 \xrightarrow{\text{unif. integr}} E[X_n - X_0] \rightarrow 0$

Tbm conv.  
em  $L^1$

O esq em lápis parece escrito levantando duas ppg. naturais.

- ① Qual a utilidade da conclusão  $\mathbb{E}[|X_n - x|] \xrightarrow{n \rightarrow \infty} 0$
- ② A condição  $\{X_n\}_{n=1}^m$  u.i. é natural?

Resp 2.

Tomo  $\mathbb{S}$  umma. ( $\text{com } \mathbb{E}[\mathbb{S}] < \infty$ )

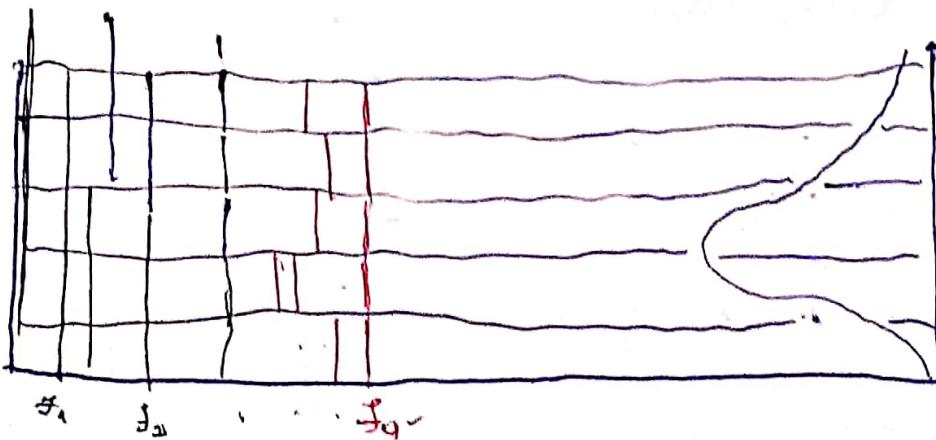
Tomo uma filtragão  $\{\mathcal{F}_n\}_{n=1}^{\infty}$ .

Faço

$$X_n := \mathbb{E}[\mathbb{S} | \mathcal{F}_n], n = 1, \dots$$

$\{X_n\}_{n=1}^{\infty}$  é martingal.  $\{X_n\}_{n=1}^{\infty}$  unsf. integrável

$$\mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n$$



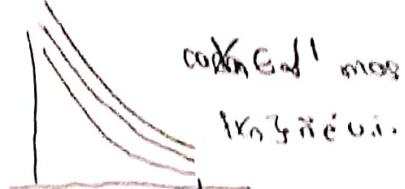
O desenho mostra que se " $f_n \rightarrow B$ " então  $X_n(\omega) \rightarrow f(\omega)$ .

Mas existe  $\{\mathcal{F}_n\}_{n=1}^{\infty}$ , filtragão " $f_n \rightarrow B$ " ( $f_n \rightarrow$  algo mais grosso que  $B$ )

Neste caso  $\exists X_{00}$  (pelo Teo. Conv.) Qual a relação entre  $X_{00}$  e  $\mathbb{S}$ ? 26

$$\{X_n\}_{n=1}^m$$

$$X_n = \frac{n}{100}, \text{ se } \Omega = [0, 1]$$



Sugestão:  $X_\infty = \mathbb{E}[\xi / \mathcal{F}_\infty]$ , a "limite" do  $\{\mathcal{F}_n\}$ ,  
 $\mathcal{F}_\infty = \sigma(\bigcup_{n=1}^{\infty} \mathcal{F}_n)$

$\mathcal{F}_\infty = \sigma(\bigcup_{n=1}^{\infty} \mathcal{F}_n)$ . Existe situação em que não necessita de  $\sigma$ ?

Tradição

$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \subseteq \dots \subseteq \mathcal{F}_n \subseteq \dots$  se  $\mathbb{E}[\xi_\infty] < \infty$  existe,

$\dots \subseteq \mathcal{F}_n \subseteq \dots \subseteq \mathcal{F}_3 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_1$  (\*)  
 'tríco de todos'

O + pobre, não existe.

Exemplos, int. acabado (B,  $\mathcal{F}$ ).

Tomo  $\xi$  uma v.a. ( $\text{com } \mathbb{E}(|\xi|) < \infty$ )

Defino

$$\dots X_n \dots X_3 X_2 X_1$$

$$\mathbb{E}[\xi / \mathcal{F}_n] \quad \mathbb{E}[\xi / \mathcal{F}_3] \quad \mathbb{E}[\xi / \mathcal{F}_2]$$

→ neste sentido é martingal.

$\{\mathbb{E}[\xi / \mathcal{F}_n]\}_{n=1}^{\infty}$  é martingal

Vocal

$X_1, \dots, X_n, \dots$  e  
 → martingal

Reverso

$\dots, X_2, X_1$   
 → martingal

→ convergencia? ← convergencia?

Exercício. Ver Demonstrações. Teo. Lévy

Mostrar que  $X_{-\infty} = \mathbb{E}[S | \mathcal{F}_{\infty}]$  e. o Teo. Lévy.

# Probabilidade Avançada II

Aula 10

Exercícios.

Teorema de Lévy. e mostrar que  $X_\infty = E[\xi | \mathcal{F}_\infty]$ .

Lévy's 'Upward' Theorem

Eja  $\xi \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ , e defina  $M_n := E[\xi | \mathcal{F}_n]$ , q.c. Então  $M$  é um UI martingal e

$$M_n \rightarrow \eta := E(\xi | \mathcal{F}_\infty)$$

quase cert. em  $\mathcal{L}'$ .

Prova. Sabemos que  $M$  é um martingal pela Tower Property. Sabemos do Teorema 13.4 que  $M$  é UI. Portanto  $M_\infty := \lim M_n$  existe q.c. e em  $\mathcal{L}'$ , e disto resta apenas provar que  $M_\infty = \eta$  q.c., onde  $\eta := E(\xi | \mathcal{F}_\infty)$

Sem perda de generalidade, podemos (e faremos) assumir que  $\xi \neq 0$ . Agora considere as medidas  $Q_1$  e  $Q_2$  em  $(\Omega, \mathcal{F}_\infty)$ , onde

$$Q_1(F) := E(\eta; F), \quad Q_2(F) := E(M_\infty; F), \quad F \in \mathcal{F}_\infty.$$

Se  $F \in \mathcal{F}_n$ , então desde que  $E(\eta | \mathcal{F}_n) = E(\xi | \mathcal{F}_n)$  pela Tower Property,

$$E(\eta; F) = E(M_n; F) = E(M_\infty; F),$$

a segunda inc. prov. na Seção 14.1. Deste modo,  $Q_1$  e  $Q_2$  concordam no  $\pi$ -sistema (algebra)  $\cup \mathcal{F}_n$ , e portanto concordam em  $\mathcal{F}_\infty$ .

Amboos  $\eta$  e  $M_\infty$  s̄o  $\mathcal{F}_\infty$  mensuráveis; mais estritamente,  $M_\infty$  pode ser tomada como  $\mathcal{F}_\infty$  mns. Definindo  $M_\infty := \limsup M_n$  para todos  $n$ . Deite modo,

$$F := \{\omega : \eta > M_\infty\} \in \mathcal{F}_\infty,$$

e desde que  $Q_1(F) = Q_2(F)$ ,

$$E(\eta - M_\infty; \eta > M_\infty) = 0.$$

Portanto  $P(\eta > M_\infty) = 0$ , e similarmente.  $P(M_\infty > \eta) = 0$ .

Argumentau igualdade nas esp. condic. para , a partirdai, mostrar que  $P(\eta > M_\infty) = P(M_\infty > \eta) = 0$ .

### • Lévy's 'Downward' Theorem

Suppenha ( $\Omega, \mathcal{F}, P$ ) é um esp prob e que  $\{G_n : n \in \mathbb{N}\}$  é uma coleção de sub- $\sigma$ -algebras de  $\mathcal{F}$  t.q.

$$G_\infty = \bigcap_k G_k \subseteq \dots \subseteq G_{-(n+1)} \subseteq G_n \subseteq \dots \subseteq G_1$$

Seja  $\gamma \in L^1(\Omega, \mathcal{F}, P)$  e defina

$$M_n := E(\gamma | G_n).$$

Então

$$M_\infty := \lim M_n \text{ existe q.c. e em } L^1$$

$$\text{e } M_\infty = E(\gamma | G_\infty) \text{ q.c.}$$

Prova. The Upcrossing Lemma aplicado ao martingal  $(M_k, G_k : -N \leq k \leq -1)$  pode ser usado p/ mostrar que  $\lim M_n$  existe q.c. O w.t. resultado Teo 13.4, mostra que  $\lim M_n$  ex. em  $L^1$ .

Então (\*) vale, ( $\text{se } g(\cdot) = \limsup M_n \in G_\infty$ ) segue por família racioc. PI GE  $g_\infty \in G_\infty$ ,

$$E(\gamma; G) = E(M_\infty; G) \neq \text{faca } r \neq 0.$$

# Probabilidade Avançada II

Aula 11

22/09

$X_1, X_2, \dots$  v.a. independentes.

Defino:  $T_n = \sigma(X_{n+1}, X_{n+2}, \dots)$  e  $T = \bigcap_{n=1}^{\infty} T_n$

34.3 Williams. Afirma se  $F \in \mathcal{T} \Rightarrow P(F) = 0 \text{ ou } 1$ .

Lei 0-1 de Kolmogorov.

Demost. com uso de martingais  
(não vai funcionar)

Comentário: onde foi definida  $P$ , que tem esta relação estranha com oper  $F \in \mathcal{T}$ ?

Resp: as funções  $X_1, X_2, \dots: \Omega \rightarrow \mathbb{R}$  são l.i.g. as v.a.s  $X_1, X_2, \dots$  são ind.

Pergunta:  $\sigma(X_{n+1})$  - conteúdo

$\sigma(X_{n+1}, \dots, X_{n+m})$  - conteúdo

$\sigma(X_{n+1}, X_{n+2}, \dots) = ?$

Dem. segundo Williams:

ao definir  $T_n$ , tem-se que

Defino  $\mathfrak{I}_n := \sigma(X_1, X_2, \dots, X_n)$  (!) Tome  $F \in \mathcal{T}$

Defino v.a.  $\eta = \mathbb{1}_F(1)$  Obs: em cada  $w \in \Omega$ ,  $\eta(w)$  é ou 0 ou 1. (!)

Quem é  $\mathfrak{I}_{\infty}$ ? Vamos ao Ex. do Cap. 10 de Williams, o qual indica que

$\mathfrak{I}_n := \sigma(X_1, \dots, X_n)$  e  $\mathfrak{I}_{\infty} := \sigma(\bigcup_{n=1}^{\infty} \mathfrak{I}_n)$ . Por que  $\eta$  é  $\mathfrak{I}_{\infty}$ -mensurável?

Ser que  $\eta$  é  $T$ -mensurável!

④ Se  $\eta$  for  $\mathbb{F}_n$ -mensurável, então, já que  $\mathbb{F}_\infty = \sigma(\cup \mathbb{F}_n)$ ,  
 a série  $\{\eta_n\}$  é  $\mathbb{F}_\infty$ -mensurável

Pois se  $\{\eta_n\}$  é  $\mathbb{F}_n$ -mensurável em relação a  $\sigma(X_1, \dots, X_n)$  p/ algum  $n$ ?

⑤ Como se controla  $\sigma(X_{n+1}, X_{n+2}, \dots)$ ?

Portanto  $q \in b\mathbb{F}_\infty$ ,  $q \in \Gamma$  (que dizer  $E[|q|] < \infty$ )!!

Port.  $\eta$  satisfaz as condições do Lévy Upward Theorem (Sessão 14.2) e, def.  
 $M_n = E[\eta | \mathbb{F}_n]$  tem-se que  $M_n \rightarrow E[\eta | \mathbb{F}_\infty]$

!

$$E[|\eta| | \mathbb{F}_\infty] = \lim_{n \rightarrow \infty} E[|\eta| | \mathbb{F}_n]$$

pois  $\{\eta_n\} \in \mathbb{F}_n = \sigma(X_1, X_2, \dots)$  - mensurável, e  $\mathbb{F}_n = \sigma(X_1, \dots, X_n)$ . então  $\eta$  não depende de  $\mathbb{F}_n$  e portanto  $E[\eta | \mathbb{F}_n] = E[\eta]$

A seq.  $\{E[\eta | \mathbb{F}_n]_{n=1}^\infty\}_{n=1}^\infty$  é a menor seq. p/ cada w.

Lem 0.3 Kolmogorov (Sem Martingais) (Shiryaev, Cap 4, §3)

Se  $A \notin \mathfrak{B}$ , então  $P(A) \leq 4 \cdot 2^{-n} \cdot 0$ , onde  $\mathfrak{B}$  é a álgebra contabil:  $\mathfrak{B} = \bigcap_{n=1}^\infty \mathbb{F}_n$   
 $\mathbb{F}_n^* = \sigma(\xi_1, \xi_2, \dots)$

Se  $A \in \mathfrak{B}$  então  $A \in \mathbb{F}_n^* = \sigma(\xi_1, \xi_2, \dots) = \underbrace{\sigma(\bigcup_{i=1}^n \mathbb{F}_i)}$

é a respectiva da massa dividida ①

e é possível escolher  $A_n \in \mathbb{F}_n^*$ ,  $n=1, 2, \dots$ , tal que

$$P(A \Delta A_n) \rightarrow 0, \quad n \rightarrow \infty \quad \text{tal que } P(A \cap A_n) = P(A)P(A_n) \Rightarrow P(A) = 0 \text{ ou } 1$$

Prob. 8 do Cap II, Seção 2.

Seja  $(\Omega, \mathcal{I}, P)$  um espaço probabilístico e seja  $\mathcal{A}$  uma álgebra de subconjuntos de  $\Omega$ .

Suponha  $\sigma(\mathcal{A}) = \mathcal{F}$ .

Prove que para qualquer  $\epsilon > 0$ , se existir  $B \in \mathcal{F}$ , podemos achar  $A \in \mathcal{A}$  t.q.

$$P(B \Delta A) \leq \epsilon \quad (\text{usa prop. conj. "suitable"})$$

O modo de raciocinar: pode ser que não para cada  $B \in \mathcal{F}$ , poderemos achar  $A \in \mathcal{A}$  t.q.  $P[A \Delta B] \leq \epsilon$  para  $\epsilon$  dado anteriormente.

Mas para alguns  $B$ , isto talvez seja possível. Vamos juntar tais  $B$ 's num conjunto:

$$M := \{B \in \mathcal{F} : \forall \epsilon > 0 \exists A \in \mathcal{A} \text{ t.q. } P(B \Delta A_\epsilon) \leq \epsilon\}$$

Se mostrar (por alg. caminho) que  $M = \mathcal{F}$ , então resolvemos o problema.

Caminho: provar que  $M$  é classe monótona  
 $M$  é álgebra  $\left( \begin{array}{l} D_n \uparrow D \text{ se } D_n \in M \forall n \\ \text{então } D \in M \end{array} \right)$

Isto ajuda porque:

$$\mathcal{F} \subset M \Rightarrow \sigma(\mathcal{F}) \subset \sigma(M)$$

mas  $M$  é álgebra e classe monótona  $\Rightarrow M$  é  $\sigma$ -álgebra  $\Rightarrow \sigma(M) = M$ . Mas  $M \subset \mathcal{F} \Rightarrow M = \mathcal{F}$ .

Shiryayev - Cap 2, Seção 3 (Subseção Espaço Mensurável  $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ )

# Probabilidade Avançada II

Aula 13

29/09

Problema (causado pela Lei 0-1)

$$\xi_1, \xi_2, \dots$$
$$\sigma(\xi_1, \dots, \xi_n)$$

$$\sigma\left(\bigcup_{n=1}^{\infty} \sigma(\xi_1, \dots, \xi_n)\right)$$

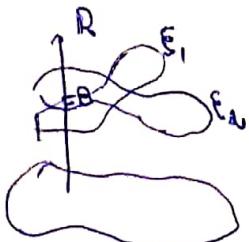
$A \in \sigma(\xi_1, \xi_2, \dots)$ ,  $P(A)?$

$\forall \varepsilon > 0$ ,  $\exists n \in \mathbb{N}$   $\exists A_n \in \sigma(\xi_1, \dots, \xi_n)$  t.q.  $P(A \Delta A_n) \leq \varepsilon$ .

Seja  $A$  uma álgebra. Prove que para cada  $B \in \sigma(A)$ , existe  $A = A(\xi, B) \in A$  t.q.  $P(B \Delta A) \leq \varepsilon$ .

Onde procurar tal resposta?

Observ. da pergunta: O que é  $\sigma(\xi_1, \dots, \xi_n)$ ? É uma o-álg. de subconj. do que.

①  $\xi_i : \Omega \rightarrow \mathbb{R}$ ,  $B(\mathbb{R})$ -meas.  
 $\sigma\{\xi_1^{-1}(B) \cap \dots \cap \xi_n^{-1}(B), B \in B(\mathbb{R})\}$

② Shiryaev. Espaço  $(\mathbb{R}^k, B(\mathbb{R}^k)) \subset (\mathbb{R}^\infty, B(\mathbb{R}^\infty))$   
 $\sigma(\xi_1, \dots, \xi_n) \subseteq B(\mathbb{R}^\infty)$

• Se admitir que o Lema está adocado no ambiente ①, então tem-se

$\Omega, \mathcal{A}, \sigma(\mathcal{A})$  num  $\mathbb{A}$  e sua extensão pl  $\sigma(\mathcal{A})$ .  $\rightarrow$  Halmos, Measure Theory

• Se admitir ②, a mais tem-se  $B(\mathbb{R}^n)$  e  $B(\mathbb{R}^\infty)$ . Estes foram construídos usando conj. abertos  $\Rightarrow$  topologia ou métrica.  $\rightarrow$  livro Kolmogorov e Fomin (\*). 30

\* "Elements of the Theory of Functions and Functional Analysis", Kolmogorov e Fomin

A-álgebra  $\sigma(A)$  -  $\sigma$ -álgebra

Halmos

R-anel  $\sigma(\mathbb{R})$  -  $\sigma$ -anel

A-álgebra é anel que contém  $X (= \Omega)$

Qual é o interesse (necessidade) de tirar  $\Omega$  das considerações?

$$\Omega = \{-\infty, +\infty\}$$

A classe de todas uniões fin. de intervalos finitos fechados à esq. e abr. à esquerda.

Uma classe não vazia  $\mathcal{R}$  chama-se anel se

$$E \in \mathcal{R} \text{ e } F \in \mathcal{R} \Rightarrow E \cup F \in \mathcal{R}$$
$$E \setminus F \in \mathcal{R}$$

Com  $\emptyset \in \mathcal{R} \neq \emptyset$ ,  $\exists E \in \mathcal{R}, \Rightarrow E \setminus E = \emptyset \in \mathcal{R}$ , p/  $F = E$ .

Se que  $E \Delta F = (E - F) \cup (F - E) \Rightarrow E \Delta F \in \mathcal{R}$ .

Se que  $E \cap F = (E \cup F) - (E \Delta F) \Rightarrow E \cap F \in \mathcal{R}$ .

$\Rightarrow \bigcup_{i=1}^n E_i \in \mathcal{R}, \bigcap_{i=1}^n E_i \in \mathcal{R}, \forall i, \forall E_1, \dots, E_n \in \mathcal{R}$ .

- Medida em um anel.

extended real function

↓  
pode ser  $+\infty$ .

(Caratheodory) Seja  $\mu$  uma medida no anel  $\mathcal{R}$ . Então  $\exists$  é única extensão  $\tilde{\mu}$  de  $\mu$  para  $\sigma(\mathcal{R})$  s.t.  $\tilde{\mu}|_{\mathcal{R}} = \mu$  (quer dizer  $\forall E \in \sigma(\mathcal{R}), \tilde{\mu}(E) = \mu(E)$ ).

Tome  $\mu$  no  $\mathbb{R}$ . Tome  $\{E_i\}_{i=1}^{\infty}$  de  $\mathbb{R}$ , d.s.g. entresi.

Sab. que  $\mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$ .

(Obs.  $\mu(E \Delta E_0) \leq \epsilon, E_0 \in \mathcal{R}, E \in \sigma(\mathcal{R})$ )

$E \in \sigma(\mathcal{R})$

$$\mu(E) = \inf \left\{ \sum_{i=1}^{\infty} \mu(E_i) : E \subseteq \bigcup_{i=1}^{\infty} E_i, E_i \in \mathcal{V}_i \right\}$$

$\mathcal{V}(\mathcal{R}) \Rightarrow \exists$  seq  $\{E_i\}$  de conj. em  $\mathcal{R}$  t.q.

$$E \subseteq \bigcup_{i=1}^{\infty} E_i \rightarrow \mu(\bigcup_{i=1}^{\infty} E_i) \leq \bar{\mu}(E) + \frac{\epsilon}{2}$$

## Prob. Avançada II

01/10

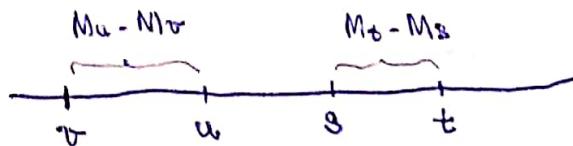
Aula 14

Anulação da Correlação linear de incrementos de um martingal

Lembrete: Def. de martingal.

$$\mathbb{E}[M_{n+1} | \mathcal{F}_n] = M_n$$

$$+ M_0 = 0 \Rightarrow \mathbb{E}[M_n] = 0$$



$$\mathbb{E}[\{M_u - M_v\} \cdot \{M_t - M_s\}] \times \{(M_t - M_s) - \mathbb{E}(M_t - M_s)\} = 0 \Leftrightarrow$$

$$\mathbb{E}\{(M_u - M_v)(M_t - M_s)\} = 0 \quad \text{pois } \mathbb{E}(M_u - M_v) = \mathbb{E}(M_t - M_s) = 0. \Leftrightarrow$$

Passo 10

$$\mathbb{E}[\mathbb{E}[\{M_u - M_v\} \{M_t - M_s\} | \mathcal{F}_n]] = \mathbb{E}[\{M_u - M_v\} \mathbb{E}[\{M_t - M_s\} | \mathcal{F}_u]]$$

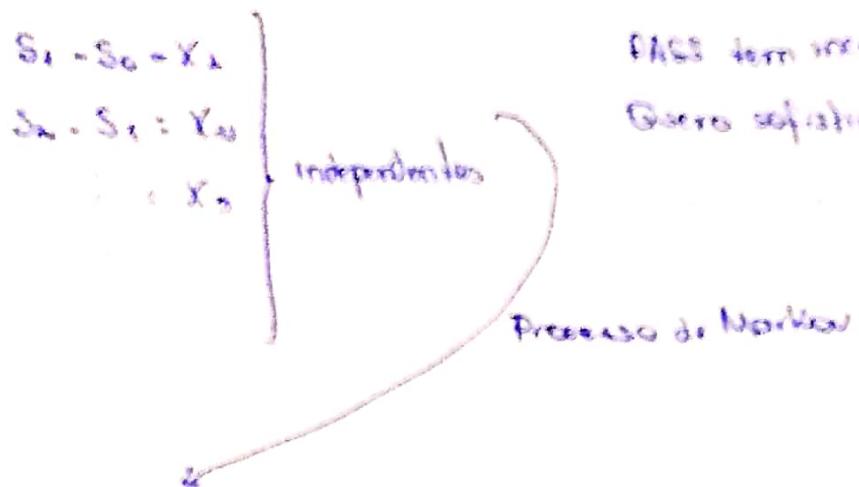
$$\begin{aligned} &= \mathbb{E}[\{M_u - M_v\} \mathbb{E}[\{M_t - M_s\} | \mathcal{F}_u]] \\ &= \mathbb{E}[\{M_u - M_v\} \{M_u - M_u\}] \\ &= 0 \end{aligned}$$

Sugestão: Se tomar  $M_0 \equiv 0$  e  $M_1 - M_0, M_2 - M_1, \dots$  não lin. correl.  
então

$\{M_n\}_{n=0}^{\infty}$  será martingal!!!  $M \neq 0$

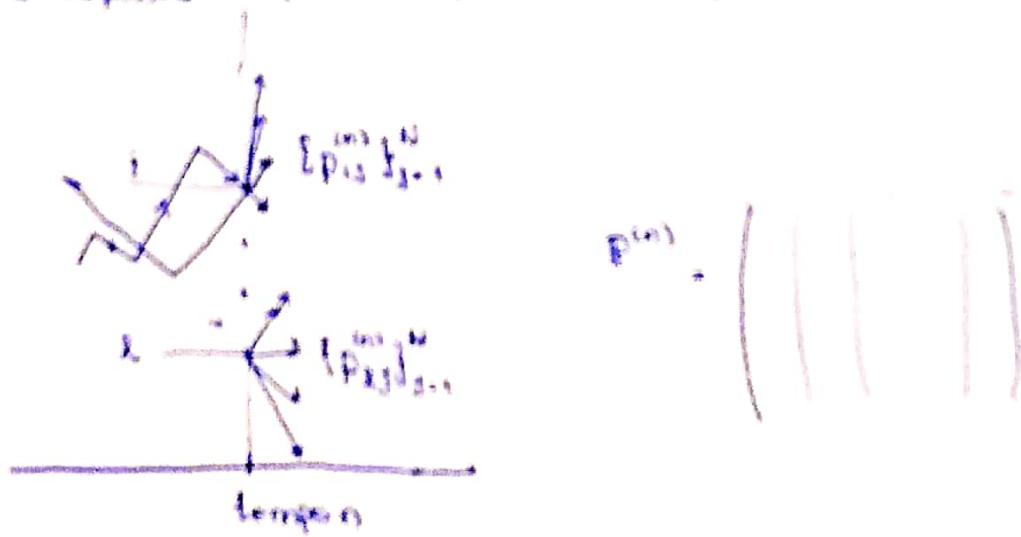
## Comentário

DASS  $S_0 = 0, S_1, S_2, \dots$



FASES tem incrementos independentes.  
 Quero refletir sobre processos.

O futuro depende do presente, mas não do passado



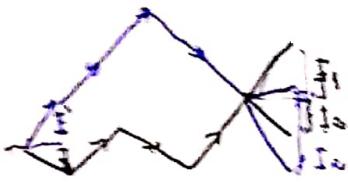
- \*  $E[\text{incremento}] = 0$
- \*  $E[\text{incremento} \neq 0 \text{ given passado}] = 0$
- \*  $E[M_{n+1} - M_n | S_n] = 0$

¶

$E[M_{n+1} | S_n] = M_n \Rightarrow$  anulação da  
 correlação linear  
 de incrementos

$$0 = E[(M_u - M_v)(M_t - M_s)]$$

$$= E[(M_u - M_v) \underbrace{E[(M_t - M_s)|\mathcal{F}]}_0]$$



$$\begin{aligned} & (I_1 \times I_2) P(I_1, I_2) \\ & + I_1 \times I_3 P(I_1, I_3) \\ & + I_2 \times I_3 P(I_2, I_3) \\ & + I_3 \times I_3 P(I_3, I_3) \\ & = I_1 (I_1 P(I_1, I_1) + I_2 P(I_1, I_2)) + \\ & I_2 (I_1 P(I_2, I_1) + I_2 P(I_2, I_2)) \\ & = \end{aligned}$$

Exercício 4. Seção "Martingais" do Cap. I do Shirayev.

Motivação:  $\xi_1, \dots, \xi_n$ .

martingal em relação à "filtragem"

$D_1 \subseteq D_2 \subseteq \dots \subseteq D_n$ .

$T$  - momento de parada em relação a

Esfôco:

$$E[\xi_T] = E[\xi_t].$$

Aprove o contrário.

Se  $\xi_i$  é  $D_i$ -mensurável ( $i=1, 2, \dots, n$ )

tempo finito deve ser importante

para ser subst. por qqr tempo de parada que assume dñs valores.

Se para qqr tempo de parada  $T$ , acontece.

$$E[\xi_T] = E[\xi_t]$$

então  $\{\xi_i\}_{i=1}^n$  é um martingal em relação a  $\{D_i\}_{i=1}^n$ .

Se o tempo é finito,  $\xi_1 = E[\xi_n / D_1]$ . Se  $n \rightarrow \infty$ , " $\xi_\infty$ " pode não existir.

The Ballot Theorem:

Durret:  $\{\xi_j : 1 \leq j \leq n\}$  i.i.d.

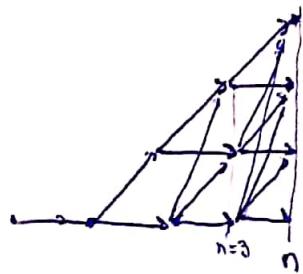
nonnegative, integer r.v.'s

$$S_n = \xi_1 + \dots + \xi_n.$$

$$P[S_j \leq j \text{ para cada } j=1, \dots, n \mid S_n] \geq (1 - s_n/n)^+ = \max(0, 1 - s_n/n)$$

with equality if  $P[\xi_j \leq j] = 1$ .

Shiryaev  $\Rightarrow$  " $=$ " sem a ressalva.



$$\xi_1 = \begin{cases} 0 & p_0 \\ 1 & p_1 \\ 2 & p \\ 3 & q \\ 4 & p_3 \end{cases}$$

A Matemática do Koltchegov.

## Probabilidade Avergada II

Aula 15.

06/10/15

$\eta_1, \eta_2, \dots, \eta_n$  independentes i.d. assumindo valores inteiros não negativos.

$$\delta_1 = \eta_1, S_2 = \eta_1 + \eta_2, \dots, S_n = \eta_1 + \dots + \eta_n.$$

$$P[S_1 \leq 1, S_2 \leq 2, \dots, S_{n-1} \leq n-2, S_n \leq n \mid S_n = l] = \max\left(0, 1 - \frac{S_n}{n}\right)$$

Exemplo: Seja  $\eta$  t.q. seu conjunto de valores é maior que  $\{0, 1, 2\}$ .

Soponha que  $\eta$  pode assumir qualquer valor em  $\mathbb{N}$ .

$$n=5, l=4 \Leftrightarrow S_5=4. \text{ De um lado, } (0; 1 - \frac{S_5}{5}) = (0; 1 - 0) = 1/5.$$

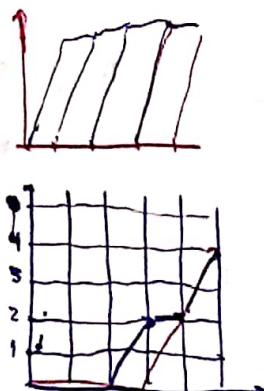
Quero calcular:

$$P[S_1 \leq 1, S_2 \leq 2, S_3 \leq 3, S_4 = 4, S_5 \leq 5 \mid S_5 = 4] = 1/5$$

Tamanho do passo

|   | Caso 1<br>Quantidade |
|---|----------------------|
| 0 | 4                    |
| 1 | 0                    |
| 2 | 0                    |
| 3 | 0                    |
| 4 | 1                    |
| 5 | 5                    |

Caso 1  
Quantidade



$$P[S_1 \leq 1, \dots, S_5 \leq 5 \mid S_5 = 4, \text{ todos } \eta_1, \dots, \eta_4 = 0 \text{ e } \eta_5 = 4] = 1/5$$

$$P[S_1 \leq 1, \dots, S_4 \leq 4 \mid S_4 = 4, \text{ } \exists \eta_5 = 0 \text{ e } \eta_5 = 2] =$$

$$\rightarrow \frac{2}{C_5^3}$$

| Tamanho do passo | Caso 2 | Caso 3 |
|------------------|--------|--------|
|                  | Oldo   | Oldo   |
| 0                | 3      | 3      |
| 1                | 0      | 3      |
| 2                | 2      | 0      |
| 3                | 0      | 1      |
| 4                | 0      | 0      |
| ...              | {      | }      |

$$\rightarrow \frac{4}{C_2^1 C_2^2} = \frac{4}{20}$$



①  $\eta_1, \eta_2, \dots, \eta_n$  i.i.d.  $S_n = \eta_1 + \dots + \eta_n$

$$E[\eta_n | S_n]^{(\omega)}, \frac{S_n^{(\omega)}}{n}$$

Porque

$$E[\eta_n | S_n](\omega) = E[\eta_n | S_n](\omega)$$

$$E[\eta_n | S_n]^{(\omega)} = \frac{S_n^{(\omega)}}{n}$$

②

$$\text{Define } \xi_1 = \frac{S_n}{n}, \xi_2 = \frac{S_{n-1}}{n-1}, \dots, \xi_k = \frac{S_{n-k+1}}{n-k+1}, \dots, \xi_{n-1} = \frac{S_2}{2}, \xi_n = \frac{S_1}{1}$$

$$\text{e } D_K = D(\xi_1, \dots, \xi_K), K=2, 2, \dots, n$$

$(\xi_n, D_n)_{n \in \mathbb{N}}$  é martingal.

③ Defina  $\tau$ , tempo de parada em relação à  $D_1, D_2, \dots, D_n$ .

$$\tau = \min \{ 1 \leq k \leq n, \xi_k \geq 0 \}$$

$$\tau = n \text{ se } \xi_1 \leq 0, \xi_2 \leq 0, \dots, \xi_{n-1} \leq 0.$$

④  $\Leftarrow$  ② ③

$$\mathbb{E}[\xi_n | \mathcal{D}_n] = \mathbb{E}[\xi_n | \xi_1] = \mathbb{E}[\xi_n | S_n] = \frac{S_n}{n}$$

$\xi_1$

$$\mathbb{E}[\xi_n | S_n] = \frac{S_n}{n}$$

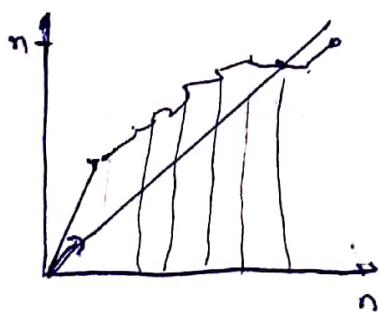
Direção da confirmação errada.

$$\{S_1 \in 1, S_2 \in 2, \dots, S_n \in n\} = \{\xi_1 \in 1, \dots, \xi_n \in 1\}$$

$$= \{T=n\} \setminus \{\xi_1 \in 1, \xi_2 \in 2, \dots, \xi_n \in 1\}$$

$$= \{T=n\} \setminus \{\xi_1 \in 1, \xi_2 \in 2, \dots, \xi_n \in n\}$$

Analisar esse evento num  $\omega$  t.q.  $S_n(\omega) = n$ .



$$\{T=n\} \cap \{S_1 \in 1, S_2 \in 2, \dots, S_n \in n\} \subseteq \{S_1 = 1\}$$

Continuação Certa:

$$(*) \quad \mathbb{E}[\xi_n | S_n](\omega) = \mathbb{P}[S_1 \in 1 \text{ ou } \xi_2 \in 2 \text{ ou } \dots \text{ ou } \xi_n \in n | S_n](\omega) \neq \omega \text{ t.q. } S(\omega) \in n$$

$$\Rightarrow \frac{S_n}{n} = 1 - \mathbb{P}[S_1 \in 1, \dots, \xi_n \in n | S_n](\omega) \quad \omega \text{ t.q. } S(\omega) \in n$$

$$\textcircled{4} \quad 1 - \frac{S_n}{n} = \mathbb{P}[S_1 \in 1, S_2 \in 2, \dots, S_n \in n | S_n](\omega) \quad \omega \text{ t.q. } S(\omega) \in n$$

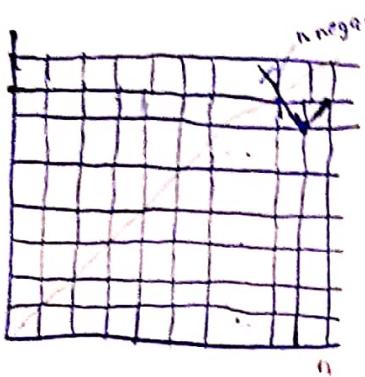
Para w.t.q.  $s_n(w) \geq n$ ,  $P[s_1 < 2, s_2 < 2, \dots, s_n < n | s_n](w) = 0$

conf 110.

Dai, segue

$$P[s_1 < 1, s_2 < 2, \dots, s_n < n | s_n] = \max \left\{ 0, 1 - \frac{6n}{n} \right\}$$

Falta averiguar a igualdade (\*).



w.t.q.  $s_n \leq n$

$s_n \geq 2$

# Probabilidade Avançada II

Aula 36

08/10/15

Movimento Browniano.

Def: Um processo  $\{B_t\}_{t \in [0,1]}$  ( $B_t, t \in [0,1]$ )

chama-se M.B. (no intervalo de tempo  $[0,1]$ ), se

①  $B_0 = 0$

② Os incrementos são independentes

$$\forall n, \forall 0 = t_0 < t_1 < \dots < t_{n-1} < t_n = 1$$

$B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$  são independentes

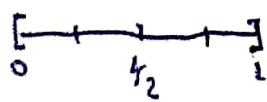
③  $\forall t, s \in [0,1]$  com  $s < t$ ,

$$(B_t - B_s) \sim N(0, t-s)$$

④ Para cada  $\omega$ ,  $B_t(\omega), t \in [0,1]$ , é função contínua

Comentário: Sobre 3. Tiro ③ e quero  $B_0=0$  construir o processo que satisfaça ③ e ④

Será que no lugar de ③, posso calcular a dist. on-f.  $B_t - B_s \sim U[0, t-s]$ ?



Da sugestão:

$$(B_{t_1} - B_0) \sim U[0, 1/2], B_1 - B_0 \sim U[0, 1];$$

$$(B_1 - B_{t_2}) \sim U[0, 1/2] \quad \text{é errado pois } B_1 - B_{t_2} \sim U[0, 1/2]$$



## Fato Trivial

$$B_1 = (B_{01} - B_{112}) + (B_{112} - B_0)$$

$$B_2 = (B_1 - B_{214}) + (B_{214} - B_{210}) + (B_{210} - B_0)$$

$$B_3 = (B_2 - B_{310}) + \dots + (B_{310} - B_0)$$

A condição ③ faz com que  $B_3$  tenha distribuição infinitamente divisível para todos  $t \in [0,1]$

↳ Normal, ou Poisson, ou mixtuas

## Infinitely Divisible Distribution

V.a.  $X$  com certa distribuição  $\mathbb{D}$

Suponha que  $X$  possua a seguinte propriedade:  $\forall n$

$$X = Y_1 + \dots + Y_n,$$

onde  $\{Y_i\}_i$  são i.i.d.

lucada

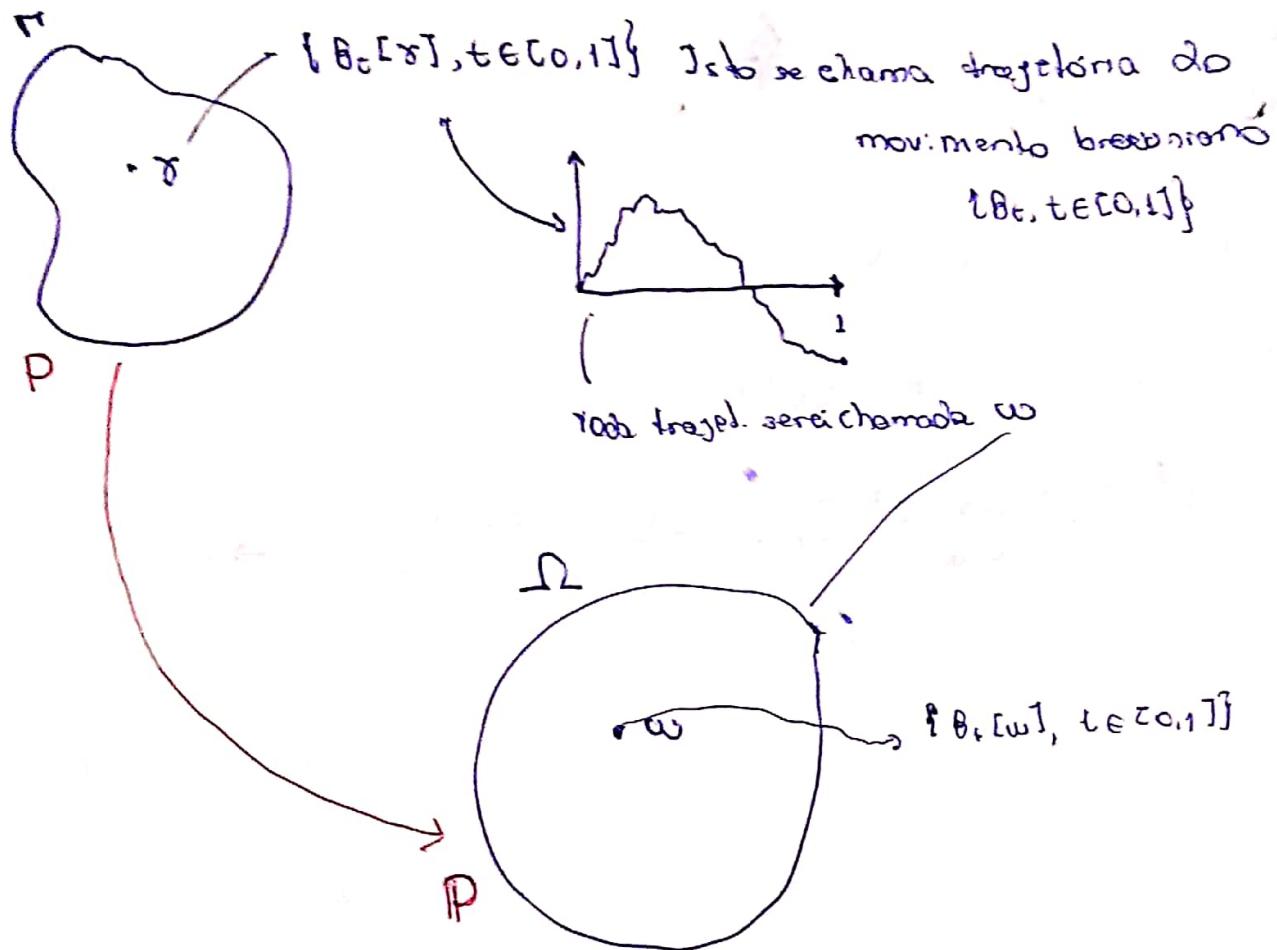
Se desejar construir um proc. com incrementos independentes, ou  
primeira pergunta é se você deseja que  $B_t - B_s$  dep. somente  
de  $t-s$ . Se sim  $\Rightarrow$  a opção é uma e única.

O que significa "processo  $\{B_t, t \in [0,1]\}$ "?

Uma edeção indexada por  $t \in [0,1]$  de v.a.  $\{B_t\}$  def. num mesmo espaço amostral  $(\Omega, \mathcal{F}, \mathbb{P})$ .

Qual deve ser  $\mathcal{F}_t$ ?  $\Omega$  pode ser enumerável?  
 $\Omega$  pode ser frnito?

Suponha que  $\exists (\Omega, \mathcal{P}, \mathcal{G})$  no qual a coleção  $\{\theta_t, t \in [0,1]\}$  foi def. da maneira t.q. ②, ⑤, ③, ④ sejam satisfeitas

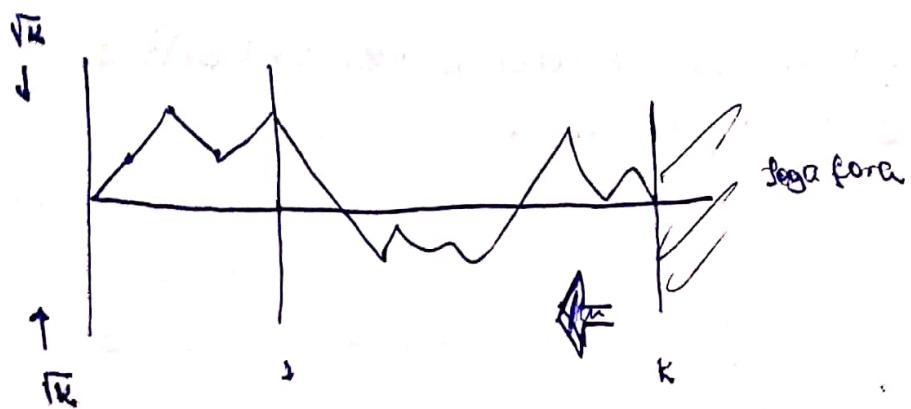


Eu fiz o seguinte: o M.B. este def. é o conjunto de todas as traj. de próprio M.B.

Com isto  $\theta_t[\omega] = \omega_t$ .

$$\mathcal{F}_s := \sigma(\theta_t, t \leq s)$$

J. Michael Steele  
Stochastic Calculus  
and financial Applications



Billingsley: "Conv. of Prob. Measures"

$$\delta_t^{\omega}, t \in [0,1] \rightarrow \mathbb{B}_t, t \in [0,1]$$

$$\mathbb{E}^{S^{\omega}}[f] \xrightarrow{t \nearrow \infty} \mathbb{E}^{\mathbb{B}}[f], \forall f \quad (\text{espace de trajectoires}) \rightarrow \mathbb{R}.$$

Lebesgue measure  $\lambda$  after Billingsley and the weak convergence of probability measures in weak topology

weak topology  $\sigma(\mathcal{B}, \mathcal{B}^*)$

weak topology  $\sigma(\mathcal{B}, \mathcal{B}^*)$

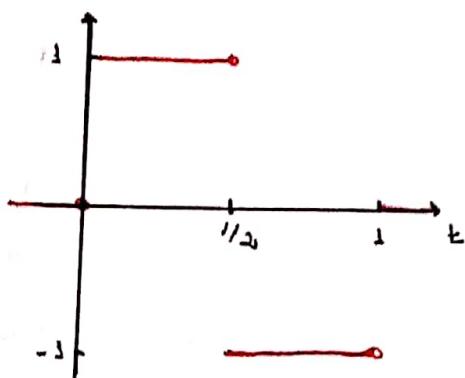
$$(\mathcal{B}, \mathcal{B}^*) \rightarrow (\mathcal{B}, \mathcal{B}^*)$$

# Probabilidade Avançada II

Aula 17

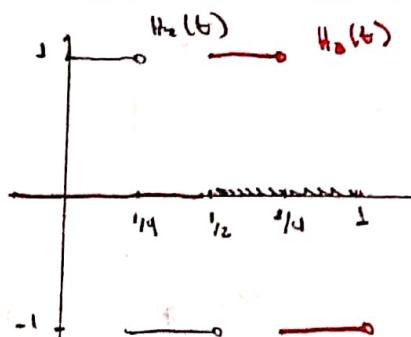
Construção do Mov. Browniano

18/10



$$H(t) = \begin{cases} 1, & \text{se } t \in [0, 1/2] \\ -1, & \text{se } t \in [1/2, 1] \\ 0, & \text{caso contrário} \end{cases}$$

A partir de  $H(t)$ , vamos construir  $\{H_n(t)\}_{n=1}^{\infty}$



Objetivo: construir base ortonormal no espaço  $L^2([0, 1])$ , funções quadrado integráveis.

A indexação de funções

|       |       |   |       |   |   |   |       |   |    |    |    |    |    |    |       |         |
|-------|-------|---|-------|---|---|---|-------|---|----|----|----|----|----|----|-------|---------|
| 1     | 2     | 3 | 4     | 5 | 6 | 7 | 8     | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16    | $\dots$ |
| $2^0$ | $2^1$ |   | $2^2$ |   |   |   | $2^3$ |   |    |    |    |    |    |    | $2^4$ | $\dots$ |

$\underbrace{H_2}_{[0, 1]} \quad \underbrace{H_3}_{[0, 1]}$

$[0, 1]$  for

div. em  $2^1$

pedaços

este bloco

tem ind.

$s=1$

$H_4$

$H_5$

$H_6$

$H_7$

$[0, 1]$  for div.

em  $2^2$  pedaços

este bloco tem

ind.

$s=2$

A transformação das funções

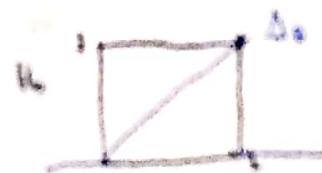
$$\rightarrow h_n(t) = 2^{\frac{n}{2}} \cdot H\left(\frac{t}{2} - k\right), \quad k = n - 2^{\frac{n}{2}}$$

expansão de Fourier

As  $n$  e  $k$  são determinadas via tabela 1 anterior

$$\text{exemplo} \rightarrow \int_0^1 h_n(t) dt = 1, \quad \forall n \rightarrow \sum_{n=1}^{\infty} \lambda_n \Delta_n = 1$$

Outras formas



$$\Delta_0(t) = 2 \cdot 2^{\frac{n}{2}} \int_0^t h_n(t) dt \quad \forall n, \quad f = g(n) \text{ ver tab}$$

Logo temos uma seq. de v.a.  $\lambda_n \Delta_n(t)$ .

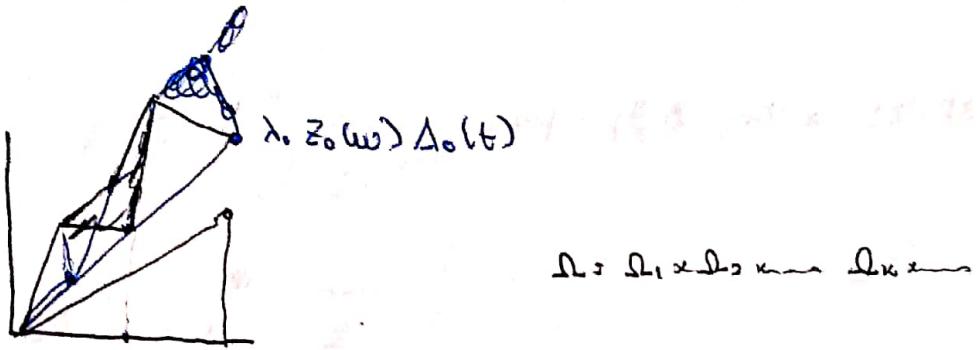
$$f_0(t) = \lambda_0 \Delta_0(t)$$

$$f_1(t) = \lambda_1 \Delta_1(t) + \lambda_0 \Delta_0(t)$$

$$f_2(t) = \lambda_2 \Delta_2(t) + \lambda_1 \Delta_1(t) + \lambda_0 \Delta_0(t)$$

$$f_n(t)$$

$$\lambda_n = \frac{1}{2 \cdot 2^{\frac{n}{2}}} \text{ onde } g = g(n)$$



Verificação / Motivação:

Por que desse tipo?

$$f_n(t) \rightarrow B_n(t), t \in [0,1]$$

M.B.

Observe: no  $t=1/2$ ,  $f_n(t) = f_3(t) = f_4(t) = \dots \Rightarrow f_2(1/2) \sim N(0, 1/2)$   
 $\lambda_{20} \frac{1}{2} + \lambda_1 z_1 \sim N(0, 1/2)$  pois  
 $\Rightarrow \lambda_1 = 1/2$        $B(1/2) = B(1/2) \cdot B(0)$

Pelo mesmo razão,  $f_0(1) \sim N(0, 1) \Rightarrow \lambda_0 = 1 \sim N(0, 1/2)$

Queremos provar que para toda  $w \in \Omega$ ,

$$f_1(t)[w], \dots, f_s(t)[w]$$

converge uniformemente para em  $t \in [0,1]$ .

↓  
a função limite é cont. em  $t \in [0,1]$ .

Depois, percorremos por todos  $w \in \Omega$  e obter  $f(t)[w]$  num sentido comum.  
Este caso é M.B. → será provado.

① Base ortonormal de funções

② Normalidade das  $z^i$ 's.

Lema: No  $(\Omega, \mathcal{F}, P)$  (onde vivem  $z^i$ 's) podemos def.  $C: \mathbb{R} \rightarrow \mathbb{R}^+$  com  $P(w: (w), \omega)^n$   
 f.g.  $\forall \omega \in \Omega, |z_n(\omega)| \leq C(\omega) \sqrt{\log n}$ ,  $n \geq 0$ ,  $1.L.P.C$

$$Z_n \sim N(0,1) \Rightarrow H_0, \quad P[|Z_n| > x] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$$

$$\leq \sqrt{\frac{2}{\pi}} \int_{-\infty}^x u e^{-u^2/2} du = e^{-x^2/2} \sqrt{\frac{2}{\pi}}.$$

$\lambda$  um par. livre cujo valor será def. depois.

$$P(A_n) := P[|Z_n| > \sqrt{2x \log n}] = \sqrt{2/\pi} e^{-x^2/2} \cdot \sqrt{2/\pi} n^{-x^2}$$

$\infty < \text{for s.t. q. } \sum_{n=2}^{\infty} n^{-x} < \infty \Rightarrow \sum_{n=2}^{\infty} P(A_n) < \infty \Rightarrow$

$P\{A_n \text{ i.u.}\} = 0 \Leftrightarrow P\{|Z_n| > \sqrt{2x \log n} \text{ finitas vezes}\} = 1.$

$$|z_0(\omega)|, |z_1(\omega)|, |z_2(\omega)|, |z_3(\omega)|, \dots, |z_{m-1}(\omega)|, |z_m(\omega)|, |z_{m+1}(\omega)|, \dots$$

$\sqrt{2x \log 2}, \quad \sqrt{2x \log 3}, \quad \sqrt{2x \log 4}, \quad \dots$ 
  
 $\sqrt{2x \log m}, \quad \sqrt{2x \log (m+1)}$

$$C^*(\omega) = \max_{\omega \in \Omega} \left\{ \max_{2 \leq k \leq m+1} \frac{|z_k(\omega)|}{\sqrt{2x \log k}} \geq 1 \right\} \Rightarrow$$

$$\forall n \quad |z_n(\omega)| \leq C^*(\omega) \sqrt{2x \log n}$$

$$\leq C(\omega) \sqrt{\log n}.$$

\* Usamos que  $Z_n \sim N(0,1)$ , mas usamos para este resultado distribuição,

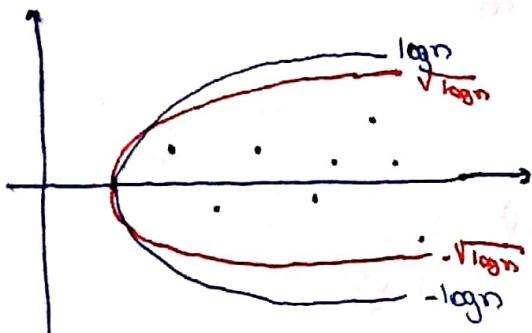
# Probabilidade Avançada II

Aula 18

B.M.

15/10/15

$\forall \omega, \exists C(\omega) < \infty : |Z_n(\omega)| \leq C(\omega) \sqrt{\log n}, \forall n$



$\omega \in \Omega \rightarrow Z_1(\omega), Z_2(\omega), \dots$

z.i.i.d.  $Z_i \sim N(0, 1)$

Queremos provar que  $\forall \omega$ ,

$\{f_0(t)[\omega], f_1(t)[\omega], \dots, t \in [0,1]\}$  convergem uniformemente

em  $t \in [0,1]$ , pois assim, como cada  $f_n(t)$  é contínua, temos que que a função limite tbm o será. (preparação)

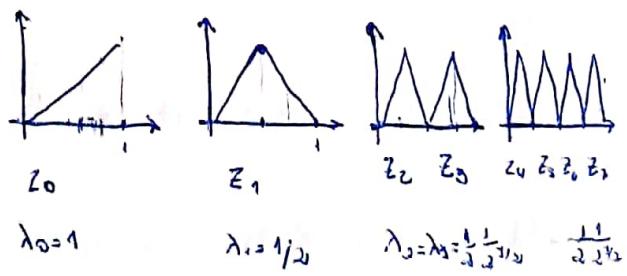
$$f_0(t) = \lambda_0 \Delta_0(t) Z_0$$

$$f_1(t) = f_0(t) + \lambda_1 \Delta_1(t) Z_1$$

$$f_2(t) = f_1(t) + \lambda_2 \Delta_2(t) Z_2 + \lambda_3 \Delta_3(t) Z_3$$

⋮

$f_n(t)$  conv. unif. em t



$\sum_{n=0}^{\infty} \lambda_n \Delta_n |Z_n(\omega)|$  absolutamente convergente em  $t \in [0,1]$

$$\sum_{n=0}^{\infty} \lambda_n |\Delta_n(t) f_2(\omega)| \leq C(\omega) \sum_{n=0}^{\infty} \lambda_n \sqrt{\log n} |\Delta_n(t)| \stackrel{\text{cuidado}}{=}$$

$$\leq C(\omega) \sum_{j=1}^{\infty} \sum_{k=0}^{2^j-1} \frac{1}{2^j} \frac{1}{2^j} \sqrt{(j+1) \log 2} |\Delta_{2^j+k}(t)|$$

desigualdade

$$\leq C(\omega) \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{1}{2^j} \sqrt{j+1} \sqrt{\log 2} \rightarrow \infty.$$

$$+ \text{Se } n \text{ no bloco } j \Rightarrow 2^j \leq n \leq 2^{j+1}-1$$

$$\log 2^j \leq \log n \leq \log(2^j \cdot 2)$$

$$\log 2 \leq \log n \leq (j+1) \log 2$$

Logo, a partir de  $\forall \epsilon > 0$

Logo, a partir de  $(\Omega, \mathcal{F}, P)$ , definimos  $\{f_i\}_{i=0}^{\infty}$  tal que

Para  $\omega \in \Omega$ , temos

$$\{f_n(t)[\omega], t \in [0,1]\}_{n=0}^{\infty}$$

de funções contínuas, que convergem uniformemente em t. Isto permite

$$\omega \mapsto \text{uma função contínua } f(t)[\omega]$$

Desejo provar que  $f_t, t \in [0,1]$  é M.B.

①  $f_0 \equiv 0$  é da construção

④ f. é contínua. Acabamos de provar

② ?

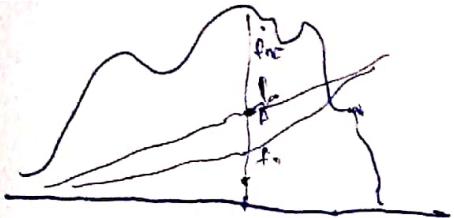
⑤ ?

Ex. Qual a dist. de

$$f_{3/4} - f_{1/2} \quad (\text{deveria ser } N(0, 3/4 - 1/2) = N(0, 1/4))$$

$$f'_{1/2} = \frac{1}{2} Z_0 + \frac{1}{2} Z_1 \Rightarrow E(f') = 0$$

$$f_{3/4} = \frac{3}{4} Z_0 + \frac{1}{2} Z_1 \lambda_1 + \lambda_3 Z_3 \quad \text{Var}(-) = 1/4$$



$f_n(t)$  é normal  $\Rightarrow$   
normal.

"Falta". Como  $\{f_n(t)\}_{n=0}^{\infty}$  convergem a  $f(t)$ ? Em prob.? Ou q.c.?

Qual seja o tipo da conv. (que existe!)

$$\{f_n(t)\}_{n=0}^{\infty} \xrightarrow{d} f(t) \quad ("d" - \text{em distribuição})$$

Usarei funções características pt provar que se  $\{X_n\}_{n=0}^{\infty}$  normais conv. em d, então a dist. do limite é normal.

$$F.T. (x)(t) = e^{imt - \frac{1}{2}\sigma^2 t^2}$$

$$X \sim N(m, \sigma^2)$$

$$X \sim N(m_i, \sigma^2)$$

$$m_i \rightarrow m$$

$$\sigma_i^2 \rightarrow \sigma^2 \text{ (que pode ser 0)}$$



$\Rightarrow$  F.T. conv. pt  $e^{imt - \frac{1}{2}\sigma^2 t^2}$  e esta é cont.  
 $\text{no } t=0$ .

$\Rightarrow$  a.v.a. limite existe e sua dist.

Todas as dificuldade a surgir na demonstração de ③ + ④ estão escondidas no livro do Steele na.

Construção do processo Gaussiano

e a demonstração (fácil) que  $f_t$ ,  $t \in [0,1]$ , é Gaussiano com

$$\text{Cov}(f_t, f_s) = \min(t, s).$$

## Probabilidade Avergada II

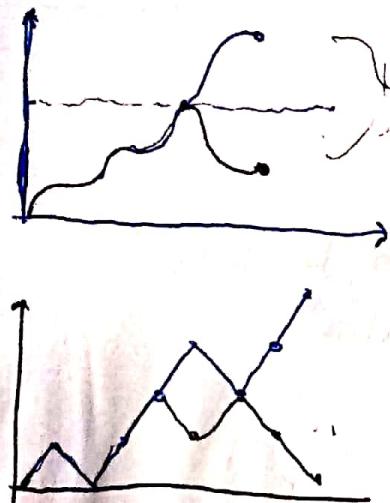
Aula 19

20/10

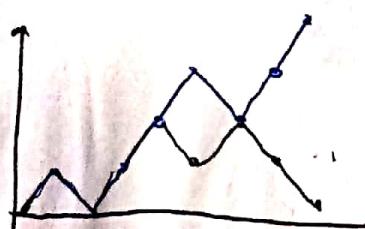
PASS (pass. al. simples simétrico) atende princ. da reflexão.

MB padrão (Mov. Brown. sem drift e de volatilidade 1)

$B_t - B_s \sim N(0, \sqrt{t-s})$  atende princ. das Reflexões.



A prob. do MB seguir gger



PASS (seguir gger branca).  $P_{\text{PASS}}$  (seguir tr. ~~amarelo~~ verde.)

Mais filosofia: por outro lado, sabemos que PASS converge (fracamente) para o MB.

O obj. que mal. a lista: Tornar o fato sobre o MB que é ddm. d. PQ, e tent derivar PR p/ PASS e a conv. da mesm ao MB.

Comentário sobre a conv. fraca do PASS ao MB.

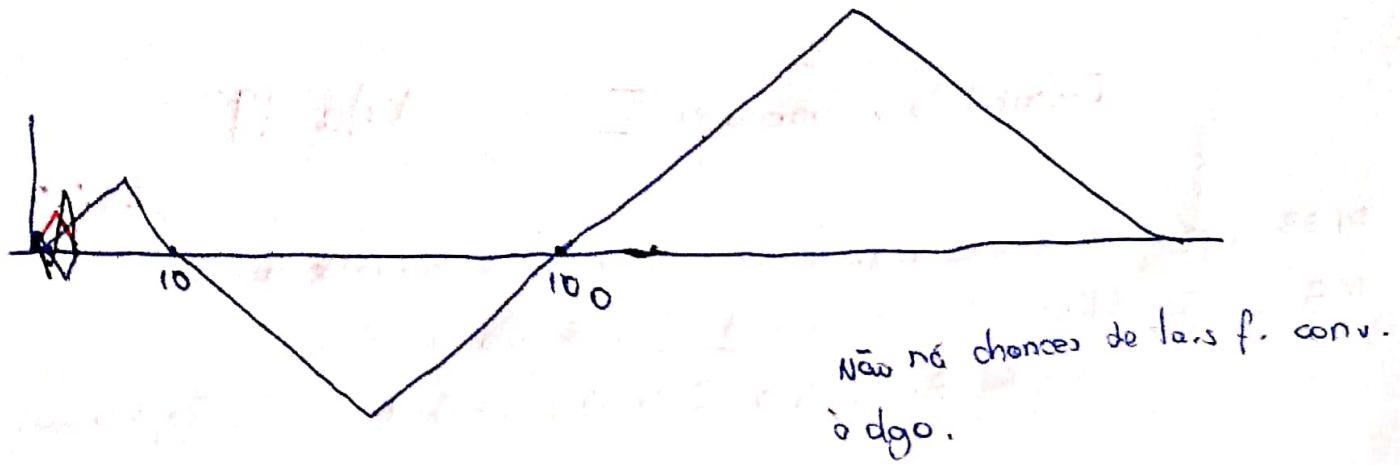
O que causa a ausência de conv. q.c?

$(\Omega, \mathcal{F}, P)$  onde  $X_1, X_2, \dots$  i.i.d. com  $P(X_i = \pm 1) = 1/2$

neste espaço vive  $\{W_k\}_{k=0}^{\infty}$  PASS

$\{S_{kL}\}_{k=0}^{\infty} \Rightarrow$  compressir em vezes  $\downarrow$  e em vezes  $\leftarrow \{S_k^{(n)}\}_{k=0}^{\infty}$

$\Rightarrow$  deixa só a parte no  $[0,1] \Rightarrow \{S_t^{(n)}\}_{t \in [0,1]}$  42



Capítulos do Steele.

Trabalho para casa: Ler Cap. 3 com seus exercícios de Steele  
Ler Cap. 5: "Richness of paths"

Comentários:

Cap. 3: Exercício 3.1 ("Brownian Bridge")

Definição: (Tradicional da Brownian Bridge) O processo  $w_t, t \in [0,1]$  def. como  $B_t, t \in [0,1]$  condicionado ao evento  $B_0 = 0$ , chama-se Brownian Bridge.

Ex 3.1. Recorde:  $B_t = \sum_{n=0}^{\infty} \lambda_n Z_n \Delta_n(t), \lambda_0 = 1, \Delta_n(t) = z_n N(0,1)$

Defino  $V_t = \sum_{n=0}^{\infty} \lambda_n \Delta_n(t) Z_n$



Observe que  $V_t(w) = 0 \forall w$ .

É possível provar que  $V_t = U_0$ .

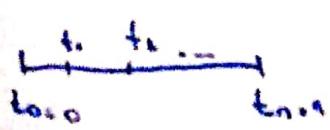
"Cálculo de Mallman"

Ex. 5.1.

$$W_t = B_t - t B_1$$

Teorema, o qual não acho no Capítulo 5 do Steele's.

$B_t, b \in [0,1]$ , a variação quadrática  $B_t(w)$ ,  $t \in [0,1]$  é 1



$$\left| \begin{array}{l} 1 \\ 2 \\ \vdots \\ n \\ \vdots \\ n-t_0, t_1, \dots, t_n \end{array} \right| \Delta(t_0, \dots, t_n) := \max(t_i - t_{i-1})$$

agora que seja a seq. de partições, se

$$\Delta(t_0, \dots, t_n) \xrightarrow{n \rightarrow \infty} 0$$

$$\text{então } \lim_{n \rightarrow \infty} [\dots] = 1.$$

$$(B_{t_2} - B_{t_1})^2 + \dots + (B_{t_n} - B_{t_{n-1}})^2]$$

2 \* Estat. Avançada I

2 \* Teoria da Decisão

3 \* Probabilidade Avançada \*

\*

\* Lema de Itô, Teorema de Itô, teorema de Itô-Tanaka

\*

3. Seja  $\{X_n\}_{n \geq 1}$  uma seq. de var. aleatória tal que

$$\{X_2 - X_1, X_3 - X_2, X_4 - X_3, \dots\} \text{ são i.i.d.}$$

$\{X_n - X_{n-1}\}_{n \geq 2}$  são não correlacionadas

Então  $\{X_n\}_{n \geq 1}$  é necessariamente um martingal com r. à filtragem  $\{\mathcal{F}_n\}_{n \geq 1}$ , com  $S_n = \sigma(X_1, \dots, X_n)$ .

$$\mathbb{E}[(X_m - \mathbb{E}X_m)(Y_n - \mathbb{E}Y_n)]$$

$$= \mathbb{E}X_m X_n - \mathbb{E}X_m \mathbb{E}X_n - ?$$

$$= \mathbb{E}X_m \mathbb{E}X_n - \mathbb{E}X_m \mathbb{E}X_n$$

$$= 0, m \neq n$$

$$\mathbb{E}[(X_m - \mathbb{E}X_m) - \mathbb{E}(X_m - \mathbb{E}X_m)]$$

$$\cdot ((X_k - \mathbb{E}X_k) - \mathbb{E}(X_k - \mathbb{E}X_k))$$

$$= \mathbb{E}(X_m X_n) - \mathbb{E}X_m \mathbb{E}X_n$$

$$- \mathbb{E}(X_m X_k) + \mathbb{E}X_m \mathbb{E}X_k$$

$$- \mathbb{E}(X_n X_k) + \mathbb{E}X_n \mathbb{E}X_k$$

$$+ \mathbb{E}(X_n X_k) - \mathbb{E}X_n \mathbb{E}X_k$$

$$= \text{Cov}(X_m, X_k) - \text{Cov}(X_m, X_k)$$

$$- [\text{Cov}(X_n, X_k) - \text{Cov}(X_n, X_k)]$$

## Probabilidade Avançada II

08/13

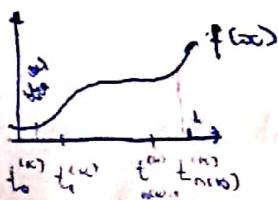
Aula 2.0

Elementary Stochastic Calculus ; Thomas Mikosch

Integral de Itô , fórmula de Ito

Integral de R-S , fórmula de N-P

Lembrete sobre integral de A.S.



Para cada  $k \in \mathbb{N}$ , tem-se uma partição

$t_0^{(w)} = t_0^{(k)}, \dots, t_{n(k)}^{(w)} = t_n^{(k)}$  de  $[0,1]$   
em  $n(w)$  pedaços

$$\lim_{n(w) \rightarrow \infty} \sum_{i=0}^{n(w)-1} f(u_i^{(w)}) [g(t_{i+1}^{(w)}) - g(t_i^{(w)})]$$

existe conforme  $(t_0^{(w)}, \dots, t_{n(w)}^{(w)}) \rightarrow \infty$  em não dep. da escolha das partições  $\{t_i^{(w)}\}$  e dos pontos  $\{u_i^{(w)}\}$ . Então este limite denota-se por

$$\int_0^1 f(x) dg(x)$$

Teorema (Cap. 2 do Mikosch)

$$\text{Int. de R-S : } \int_0^1 f(x) dg(x)$$

Teorema (Cap. 2 Mikosch) (Mulher com sobrenome Sargent)

Integral de R-S  $\int_0^1 f(x) dg(x)$  existe (no sentido de  $(\infty)$ ) se

(a)  $f \circ g$  não tem descont. nos mesmos pontos de  $[0,1]$

(b)  $\frac{1}{p} + \frac{1}{q} > 1$  onde  $p$  é a ordem de var. de  $f$  e  $q$  é a ordem de var. de  $g$ .

**Teorema:** A variação quadrática de M.B. é 1, e a variação de ordens  $\leq 2$  não existe p/ mov. B.

Onde quer-se chegar:  $\int_0^t B_s dB_s$  existe no sent. (a)

p de  $B_t$  é 2, q de  $B_t$  é 2  $\Rightarrow$  no caso, temos  $1/2 + 1/2 \neq 1$

Então Teo. do livro do Mikosch não garante a exist. da int. de R-S, pois (a) e (b) são cond. suficientes e (b) não é satisfeita.

(\*)

(II) Ainda pode ser que  $\int_0^t B_s [w] dB_s [w]$  existe no sent. R-S, pois o Teo. de Mikosch é só condição suficiente.

Vamos confirmar que, de fato, a abordagem R-S não funciona para o int.

$$\int_0^t B_s dB_s.$$

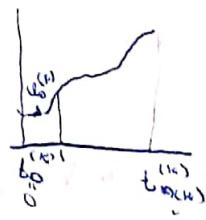
Sig. de partição:  $t_0^{(n)}, \dots, t_{n(k)}^{(n)} = 0, 1/k, 2/k, \dots, n/k$ .

Três escolhas do ponto  $w_i^{(n)}$ :

(1)  $w_i^{(n)} < t_i^{(n)}$  (ponto à esquerda)

(2)  $w_i^{(n)} = t_{i+1}^{(n)}$  (ponto à direita)

(3)  $w_i^{(n)} = \frac{t_{i+1}^{(n)} + t_i^{(n)}}{2}$  (ponto médio)



$$\lim_{k \rightarrow \infty} \sum_{i=0}^{k-1} B_{t_i^{(n)}} [B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}]$$

$$\lim_{k \rightarrow \infty} \sum_{n=0}^{k-1} B_{t_n^{(w)}} [B_{t_{n+1}^{(w)}} - B_{t_n^{(w)}}]$$

$$\lim_{k \rightarrow \infty} \sum_{n=0}^{k-1} B_{\frac{t_{n+1}^{(w)} + t_n^{(w)}}{2}} [B_{t_{n+1}^{(w)}} - B_{t_n^{(w)}}]$$

Todas são calculáveis (em cada  $w$ ) e têm val. diferentes.

**1º Caminho:** cons. só  $\int_0^1 u_t dB_t$  onde  $u_t$  tem variações l.q. (p. 1/97)

$$\int_0^1 B_t dB_t \text{ e/ou } \int_0^1 f(B_t) dB_t \text{ e}$$

**2º Caminho:** fazer construção que seja aplicável a  $\int_0^1 B_t dB_t$ .

O exemplo mostra que a posição do "ponto"  $\{u_i^{(w)}\}$  é importante.

Vamos tentar fixando que em cada int. de cada partição, o ponto  $u_i^{(w)}$  se o extremo esq. do intervalo.

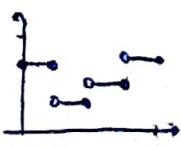
Se seg. o caminho 2. quais funções serão integráveis?

(o  $\lim_{k \rightarrow \infty}$  existe e não dep. da escolha da sep. do part.)

• Se  $f(w, t) = \text{const. } \forall w \forall t$ ,



• Se  $f(w, t)$  é step-function



• Se  $f(w, t)$  é step-function p/ cada  $w$ .

Fixado  $w$ ,  $f(w, t)$  é função escada mas  $\forall w$  a mesma função



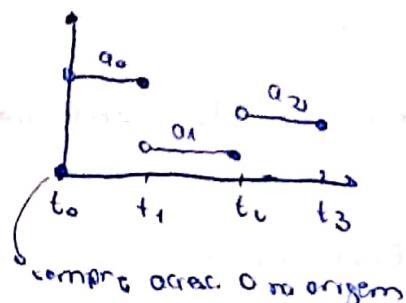
Se para algum  $\omega \in \Omega$  e alguma variável aleatória "non-unintegrating"

$$a_0, a_1, \dots, a_{n-1}, \omega \in \Omega \text{ s.t. } E[a_i^2] < \infty \quad \forall i,$$

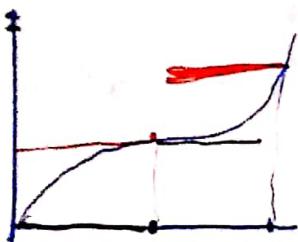
e alguma partição  $0 = t_0, t_1, \dots, t_n = 1$ ,

$$f(\omega, t) = \sum_{i=0}^{n-1} a_i(\omega) \mathbb{1}_{(t_i, t_{i+1}]}(t)$$

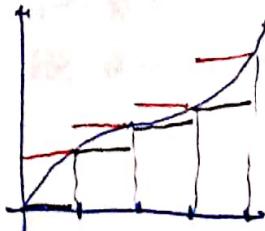
então existe



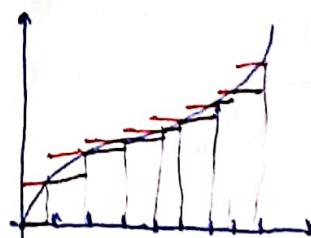
$\mathbb{E}[f(\omega, \cdot)]$



$$f_1(\omega, s)$$



$$f_2(\omega, s)$$



$$f_3(\omega, s)$$

$$s \in [0, 1]$$

$$\int_0^1 f_s(\omega, t) dB_t(\omega) = B_{1/2}[\omega] (B_{1/2}[\omega] - B_0[\omega]) + B_1[\omega] (B_1[\omega] - B_{1/2}[\omega])$$

$$\rightarrow B_0[\omega] (B_{1/2}[\omega] - B_0[\omega]) + B_{1/2}[\omega] (B_1[\omega] - B_{1/2}[\omega])$$

$\{f_t\}_{t=0}^1$  gera. M.B.

$$h(\cdot, \cdot) : \Omega \times [0, 1] \rightarrow \mathbb{R}$$

$$h(\cdot, \cdot) \in \mathcal{F} \times \mathcal{G}$$

$$\forall t, h(\cdot | t) \in \mathcal{F}_t$$

$$\left. \begin{array}{l} \int_0^1 g(\omega, t) dB_t(\omega) \\ g : \mathbb{R} \rightarrow \mathbb{R}, h(\omega, t) = g(\omega_t) \end{array} \right\}$$

$$g : \mathbb{R} \rightarrow \mathbb{R}, h(\omega, t) = g(\omega_t)$$

## Probabilidade Avançada II

Aula 21

06/11/15

A integral de Itô -  $\int f \, d\omega$  - é de uma  $f(\cdot, \cdot) \in \mathcal{H}^2 \subseteq L^2(dP, dt)$  definida como  $L^2$  limite de  $\{\int (f_n(\cdot, \cdot))\}_{n=1}^\infty$  no espaço  $L^2(dP)$ , onde  $f_n$  é uma seq. de funções contínuas ( $f_n(\cdot, \cdot) \in \mathcal{H}_0 \subseteq \mathcal{H}^2$ ) que aproxima  $f(\cdot, \cdot)$  no sentido de norma em  $L^2(dP, dt)$ .

Esta construção e outras de complexidades são demonstradas envolvendo a integral de Itô.

med. de Lebesgue

$[0, T]$ : intervalo de tempo

$B_t, t \in [0, T]$  o M.B. padrão em  $[0, T]$

↓  
def. em  $(\Omega, \mathcal{F}, \mathbb{P})$

Em conjunto,  $\Omega \times [0, T]$  com a med.  $dP \times dt$ .

$f(\cdot, \cdot) : \Omega \times [0, T] \rightarrow \mathbb{R}$

$\mathcal{H}^2 = \{f : \Omega \times [0, 1] \rightarrow \mathbb{R}, \|f\|_2 < \infty\}, \|f\|_2 \text{ equiv.}$

$$\iint_{\Omega \times [0, T]} f^2(\omega, s) \, dP \, ds < \infty$$

↑ adaptado a filtragem do M.B.

$$\int_{\Omega} \left( \int_{[0, T]} f^2(\omega, s) \, ds \right) \, dP$$

$$\mathbb{E} \left( \int_{[0, T]} f^2(\omega, s) \, ds \right), \quad 46$$

Ex:  $h(\omega, s) = s \cdot \max_{t \in [0, T]} (\omega_t)$  não é função na qual a int. de Itô será construída

Para construir Integral de Itô de  $f \in H_0^2$  (notação  $I(f)$  ou  $\int_0^T f(\omega, s) dB_s$ ), se faz o seguinte.

*funções contínuas*

a) Constrói-se  $\{f_n(\omega, s)\}_{n=1}^\infty$  t.q.  $f_n \in H_0^2$  e  $f_n \rightarrow f$  em  $L^2(dP, dt)$ .

b) O fato  $f_n \in H_0^2$  permite definir  $I(f_n)$  sem ambiguidades.

Toma-se  $L^2(dP)$ - $\lim_{n \rightarrow \infty}$  de  $\{I(f_n)\}_{n=1}^\infty$  e este limite declara-se  $I(f)$ .

A const. de  $I(f_n)$  faz com  $I(f_n)$  é uma função  $\Omega \rightarrow \mathbb{R}$

(A-S  $\int_0^T g(\omega, s) ds = Q(\omega)$  nos par. numéricos)

D) Programa (a)-(b) funciona por causa da Isometria de Itô de  $H_0^2$ .

$$\|f_n\| = \|I(f_n)\|$$

$\in L^2(dP)$

$H_0^2$

$\sup_{t \in [0, T]} \|\cdot\|_{L^2(dt)}$

A isometria de Itô garante que  $L^2$ - $\lim I(f_n)$  existe e seu valor não depende da escolha da sequência  $\{f_n\}_{n=1}^\infty$ .

$$H_0^2 \quad f(\omega, t) = \sum_{i=0}^{n-1} \alpha_i(\omega) \mathbb{1}_{[t_i, t_{i+1}]}^{(t)}$$

$$I(f)(\omega) = \sum_{i=0}^{n-1} \alpha_i(\omega) [B_{t_{i+1}}(\omega) - B_{t_i}(\omega)]$$

mas em relação a filtragem  
gerada pelas M.B. em  $[0, T]$

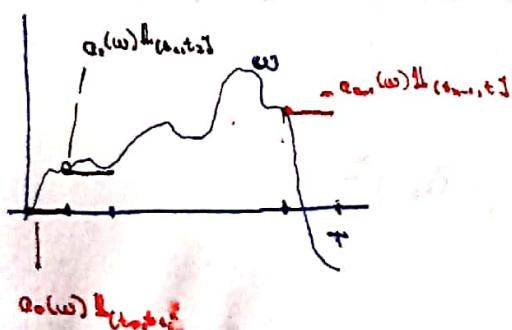
Executam-se os passos (a, b) para achar  $\int_0^T B_s dB_s$ .

$$B_s = f(\omega, s) = B_s[\omega]$$

é uma f. de duas variáveis

Exercício (a): sugerir uma seq. de f. escada que aproxima  $f(\omega, s) \cdot B_s[\omega]$  em  $L^2(dP, dt)$ .

Para cada  $n$ , divido  $[0, T]$  em  $n$  pedaços iguais e chamo  $t_0, t_1, \dots, t_n$  os pontos de separação.



(fig. \*)

$$a_0^{(\omega)} = B_0[\omega] \leq \omega_0$$

$$a_1(\omega) = B_{t_1}[\omega] = \omega_{t_1}$$

$$a_{n+1}(\omega) = B_{t_{n+1}}[\omega] = \omega_{t_{n+1}}$$

No ex. de construção  $\int_0^T B_s dB_s$ , as funções escadas aproximadoras de  $f(\omega, s) \cdot B_s[\omega]$  são construídas tomando o valor de  $B_s$  no extremo esquerdo de cada subintervalo  $(t_i, t_{i+1}]$ .

Já se induz a pensar que a mesma coisa ( $\dots f(\dots) \dots$ ) procede para a aprox. de  $f \in L^2$ . Acontece que não!

É a parte difícil de toda a construção, provar que para cada  $f \in L^2$  há  $\{f_n\}_{n=0}^{\infty}$  em  $L^2$  t.q.  $\|f_n - f\|_{L^2(dP, dt)} \rightarrow 0$ .

No Ottendal esta parte é incompreensível.

Steel usa prop. da MARTINGAL p/ provar este fato.

$$f_n(\omega, t) = \sum_{i=0}^{n-1} B_{t_i}[\omega] \mathbb{1}_{(t_i, t_{i+1}]}(t)$$

$$I(f_n) = \sum_{i=0}^{n-1} B_{t_i} (B_{t_{i+1}} - B_{t_i})$$

Quero provar "Bo" que  $\|f_n - f\|_{L^2(dB, dt)}$

Quero mostrar  $L^2$  - lim de  $I(f_n)$

- Será que as convergências neste exemplo não são w-a-w?

↓

Como  $\omega: [0, T] \rightarrow \mathbb{R}$  é contínua e, portanto, unif. cont., então  $f_n \rightarrow f$  conv. em cada  $\omega$ .

$$B_{t_i} (B_{t_{i+1}} - B_{t_i}) = \frac{1}{2} (B_{t_{i+1}}^2 - B_{t_i}^2) - \frac{1}{2} (B_{t_{i+1}} - B_{t_i})^2$$

⇒

$$\begin{aligned} I(f_n) &= \sum_{i=0}^{n-1} B_{t_i} (B_{t_{i+1}} - B_{t_i}) = \frac{1}{2} B_{t_n}^2 - \frac{1}{2} \sum_{i=0}^{n-1} (B_{t_{i+1}} - B_{t_i})^2 \\ &= \frac{1}{2} B_T^2 - \frac{1}{2} \sum_{i=0}^{n-1} (B_{t_{i+1}} - B_{t_i})^2 = \frac{1}{2} B_T^2 - \frac{1}{2} T \end{aligned}$$

De novo, o ex.  $\int B_s dB_s$  não dá ex. de int. de Itô incalculável w-a-w.

Acabei o ex.  $\left( \int_0^T B_s dB_s = \frac{1}{2} B_T^2 - \frac{1}{2} T \right)$  e quero mostrar que a compreensão w-a-w consegue a gente.

Como cada  $\int_0^t f dB_s$  foi definido como  $L^2$ -limite, então para cada  $t$ , pode existir conjunto  $A_t \subseteq \mathbb{R}$ , onde a integral é ruim (com  $P(A_t) = 0$ )

Então o processo  $\{\int_0^t f dB_s, t \in [0, T]\}$  é ruim no conjunto  $\sum_{t \in A_T} A_t = \Omega$

Teorema 6.2. Existe um processo  $\{X_t^f, t \in [0, T]\}$  no ru, com traj. q.c. contínuas tal que

$$X_t^f = \int_0^t f dB_s, \quad \forall t \in [0, T] \quad (X_t, t \in [0, T] \text{ é martingal}).$$

A dem. do Teo. 6.2.

Tomo  $\{f_n\}$ . Considero  $\{X_n(t) := \int_0^t f_n dB_s, t \in [0, T]\}$ .

(F6.2)  $X_n(t), t \in [0, T]$  é mart. com traj. (ou traj. contínuas)

Considero  $\lim X_n(\cdot)(w)$ .

Teorema. Este limite é um martingal com traj. contínuas

Chamo ele de  $\{X(t), t \in [0, T]\}$  e provo que  $X(t) := \int_0^t f dB_s, \forall t \in [0, T]$ .

## Probabilidade Avançada II

Aula 2.2

00/11/2015

Lembrete:  $f(\cdot, \cdot) : \Omega \times [0, T] \rightarrow \mathbb{R}$

$f$  adaptada à  $\mathcal{F}$ . Horação natural de  $B_t$ ,  $t \in [0, T]$

$$\int_{\Omega \times [0, T]} f^2(\omega, t) dP dt = E \left[ \int_0^T f^2(\omega, t) dt \right] < \infty$$

Tais funções formam o espaço completo denotado por  $H^2$ .

$$\{f_n\}_{n=1}^{\infty} \quad (f_n \in H^2) \quad \|f_n - f\|_L^2 \rightarrow 0$$

$$f_n \rightarrow I(f_n) \quad I(f) : \lim_{n \rightarrow \infty} \|I(f_n) - I(f)\|_{L^2(dP dt)} = 0$$

Integração de Riemann-Stieltjes  $\rightarrow$  Fórmula Newton-Leibniz  $F(b) - F(a) = \int_a^b f'(x) dx$

Integral de Itô  $\rightarrow$  Fórmula (Lema) de Itô

$$g(\omega, t) = b \max \{ \omega_s : 0 \leq s \leq t \} \leftarrow \text{Fórmula de Itô}$$

$f(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$  duas vezes diferenciável

$$f(B_T) - f(B_0) = \int_0^T f'(B_s) dB_s$$

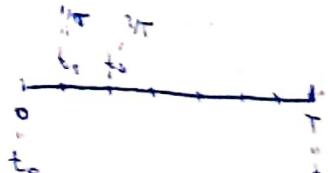
$$g(\omega, t) = f(\omega_t) \leftarrow \text{caso particular da fórmula de Itô}$$

$$+ \frac{1}{2} \int_0^T f''(B_s) ds$$

$g(\cdot, t)$

$$g(\cdot, t) = \int_0^t a ds + \int_0^t b dB_s \quad \text{Fórmula de Itô pl caso mais geral}$$

$$f(B_T) - f(B_0) = \sum_{i=0}^{n-1} f(B_{t_{i+1}}) - f(B_{t_i})$$



$$f(y) = f(x) + f'(x)(y-x) + \frac{f''(x)(y-x)^2}{2} + r_{xy}(y-x)^2$$

$$\sum_{i=0}^{n-1} f(B_{t_{i+1}}) - f(B_{t_i}) = \sum_{i=0}^{n-1} f'(B_i)(B_{t_{i+1}} - B_i) + \sum_{i=0}^{n-1} \frac{f''(B_i)}{2} (B_{t_{i+1}} - B_i)^2 + \sum_{i=0}^{n-1} 2(B_{t_{i+1}} - B_i)^2$$

$\uparrow \int_0^T f'(B_s) dB_s$       - converg. em prob.  
 $\downarrow \int_0^T \frac{1}{2} f''(B_s) ds$       converg. em prob.  
 $\downarrow 0$

Queremos integrar  $f'$ . Precisamos construir  $\{f_n\}$  que  $\xrightarrow{L^2} f'$ . Mas a construção genérica não pega  $f'$  no extremo esquerdo de cada subintervalo da partição.

## Capítulo 7 - Localisation and Itô's Integral

Extensão da  $I(t)$  para  $f \in L^2_{loc}[0,T] \equiv L^2_{loc}$ .

$$L^2_{loc} = \left\{ f: \Omega \times [0,T] \cdot \mathbb{P} \left[ \int_0^T f^2(\omega, s) ds < \infty \right] = 1 \right\}$$

$\downarrow$   
adaptação filtração rel. do M.B.

Em nenhum  $\omega$ ,  $\int_0^T f^2(\omega, s) ds$  é igual a  $\infty$   
(experi.)

#

$$\mathbb{E} \left[ \int_0^T f^2(\omega, s) ds \right] < \infty$$

Definição integral de Itô para  $f \in L^2_{loc}$ .

Definição: Uma seq. crescente de tempo de parada  $\{\tau_n\}$  chama-se "Itô's localizing for function  $f$ ", se  $f_n(t)(\omega, t) := f(\omega, t) \mathbf{1}_{t \leq \tau_n} \in H^2[0, T]$  e se

$$\mathbb{P} \left( \bigcup_{n=1}^{\infty} \{\omega : \tau_n(\omega) = T\} \right) = 1.$$

(Ω, ℙ, P)

$\mathcal{F}_t = \sigma(\beta_0, s \cdot t)$

$\{\mathcal{F}_t\}_{t=0}^T$  a filtragem nat. do mB

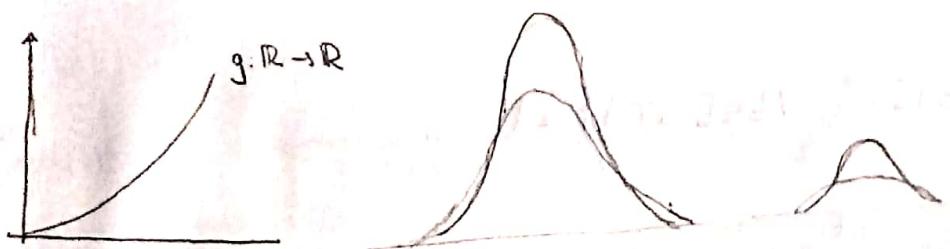
• em cada  $w$ , se vc percorrer por números

$v_1(w), v_2(w), \dots, v_n(w), \dots$  então um deles é o T.

Proposição F.J. Para cada  $f \in L^2_{loc}[0, T]$ , existe "fl<sup>2</sup>. localizing sequence", e na verdade, uma de tais é dada via

$$\chi_n(w) := \inf \{s \in [0, T] : (\int_0^s f^2(w, t) dt)^{1/2} \leq n \}$$

se  $\int_0^s f^2(w, t) dt \leq n$   
então  $\chi_n(w) = T$ .



Como seria agora, o int. de I<sup>2</sup> para  $f \in L^2_{loc}[0, T]$ ?

Tomo uma seq. localizadora  $\{v_n\}^\infty$ .

Considero  $f_n(w, t) = f(w, t) \mathbb{1}_{\{t \leq v_n(w)\}} \}_{n=1}^\infty$

estas não são funções-excedos

mas são funções da fl<sup>2</sup>

→ I(f<sub>n</sub>(·)) existe

Daí, defino

I(f) como P-lim de I(f<sub>n</sub>)

Recorda que para  $f: \Omega \times [0,T]$  em  $L^2$  posso calcular

$$I(f) = \int_0^T f(\omega, t) dB_t$$

Mas posso também, calcular  $\int_0^t f(\omega, s) dB_s \quad \forall t \in [0, T]$  de forma l.q.

$\{X_t = \int_0^t f(\omega, s) dB_s\}_{t \in [0, T]}$  seja um martingal com trajetórias contínuas ( $X_t(\omega), t \in [0, T]$  é contínua)

$$f_n = f \mathbb{1}_{\{t \leq n\}} \longrightarrow X_n(t), t \in [0, T] \quad (X_n(t) = \int_0^t f_n(\omega, s) dB_s)$$

Fato: Existe um martingal  $\overset{X(t), t \in [0, T]}{\curvearrowright}$  com traj. contínuas l.q.

$$\mathbb{P}\{ \omega : X_n(t)[\omega] \xrightarrow{n \rightarrow \infty} X(t)[\omega] \quad \forall t \in [0, T] \} = 1$$

$X(t), t \in [0, T]$  dedoro-se ser  $\int_0^t f(\omega, s) dB_s$  para  $f \in L^2_{loc}[0, T]$

Teorema F.1. Seja  $f: \mathbb{R} \rightarrow \mathbb{R}$  contínua (pode ser localmente crescente de modo l.q.  $E[\int_0^T f^2(\omega, t) dt] = \infty$ )

Tomo a partição  $t_0, t_1, \dots, t_n$  de  $[0, T]$  em  $n$  pedaços iguais, então

$$\int_0^T f(B_s) ds = \mathbb{P}\lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} f(B_{t_i}) [B_{t_{i+1}} - B_{t_i}]$$

Int. da ITB  $f$

$f \in L^2_{loc}[0, T]$ .

Sumas parciais de Riemann com  $f$  no est. esquerdo de cada int. de partição

## Probabilidade Avançada II

Aula 23

12/11/15

Lq. Dif. Estocásticas

Teorema de Girsanov

O proc.  $X_t, t \in [0, T]$  (ou  $t \in [0, \infty)$ ) apresenta no intervalo infinitesimal de tempo o drift proporcional ao seu valor e a variância prop. ao seu valor.

Portanto

$$dX_t = \mu X_t dt + \sigma X_t dB_t$$

outra

$$X_{t+\Delta t} - X_t \approx \mu X_t \Delta t + \sigma X_t (B_{t+\Delta t} - B_t) \quad \text{onde } B \in \text{o.m.B.}$$

O que é a versão intuitiva das equações

$$X_t = X_0 + \int_0^t \mu X_s ds + \int_0^t \sigma X_s dB_s \quad (*)$$

O tamanho da pop no tempo  $t$ ,  $f(t)$ , cresce no int. infinitesimal de, prop ao seu tamanho corrente

$$\frac{f(t+\Delta t) - f(t)}{\Delta t} = \lambda f(t)$$

$$f'(t) = \lambda f(t) \Rightarrow f(t) = f(0) e^{\lambda t}.$$

A explicação & construção

Tendo isso feito

$$\text{Sup. que } P(\text{ter } 1 \text{ no } [t, t+\Delta t] / \mathcal{F}_t) = \lambda \Delta t + o(\Delta t)$$

$$P(>1 \dots I \geq t) = e^{-\lambda t}$$

$$P(0 \dots I \geq t) = 1 - e^{-\lambda t} + o(\Delta t)$$

$$\Rightarrow P_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, n \in \mathbb{N}$$

Ques.

Tenho um processo estocástico real. Consigo as medições com certa precisão ate uma certa magnitude de grandeza de  $\Delta t$ . Repetindo este processo muitas vezes (pois preciso estimar as distribuições), e medindo, vejo que

$$(1) X_{t+\Delta t} - X_t \stackrel{D}{=} \lambda X_t \Delta t + \mu X_t N(0, \Delta t) + o(\Delta t)$$

$$\Rightarrow X_t = X_0 + \int_0^t \lambda X_s ds + \int_0^t \mu X_s dB_s$$

$pN(\Delta t), (1-p)N(0, \Delta t)$

# Probabilidade Avançada II

Aula 24

17/11/2015

## Teorema de Girsanov

- Entender (e saber aplicar) a form. do Teo. em sua versão mais geral.
- Entender toda a simplicidade e amplitude de aplicações.

Problema (a ser res. como ex. do T. de Girsanov)

Seja  $B_t$ ,  $t \in [0, T]$  o m.B. com volatilidade 1 e drift 0.

Tomo  $b \in \mathbb{R}_+$ . Para cada  $g \in \mathbb{R}$ , quero achar

$$\Pr \left[ \sup_{0 \leq t \leq T} B_t \geq b, B_T \leq g \right]$$

Distribuição do retorno de um ativo no tempo T, sabendo que seu retorno ultra passou a barreira b em  $[0, T]$

(visto pelo medida  $\mathbb{P}^*$ )

Teorema de Girsanov:  $W_t$ ,  $t \in [0, T]$ . m.B. padrão (drift 0, volatilidade  $\sigma$ )

$X_t$ ,  $t \in [0, T]$  - gpar processo adaptado e bom pl que os integrais sejam definidas.

$$\xi(X)_t := \exp \left\{ \int_0^t X_s dW_s - \frac{1}{2} \int_0^t X_s^2 ds \right\}, \quad t \in [0, T]$$

$$\tilde{W}_t := W_t - \int_0^t X_s ds, \quad t \in [0, T]$$

Significa  $\mathbb{Q}$  é nova  
medida em  
 $(\Omega, \mathcal{F})$

$$\mathbb{E}^{\mathbb{Q}}[\mathbb{1}_A] = \mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_A - \xi(X)_T] \quad \forall A, (\text{de onde?}) \in \mathcal{F}$$

# Probabilidade Avançada II

Aula 24

17/11/2015

## Teorema de Girsanov

1. Entender (e saber aplicar) a form. do Teo. em sua versão mais geral.
2. Entender toda a simplicidade e amplitude de aplicações.

Problema (a ser res. com aux. do T. de Girsanov)

Seja  $B_t$ ,  $t \in [0, T]$  o m.B. com volatilidade 1 e drift  $\alpha$ .

Tomo  $b \in \mathbb{R}_+$ . Para cada  $g \in \mathbb{R}$ , quero achar

$$\Pr \left[ \sup_{0 \leq t \leq T} B_t \geq b, B_T \leq g \right]$$

Distribuição do retorno de um ativo no tempo  $T$ , sabendo que seu retorno ultra-passe a barreira  $b$  em  $[0, T]$ "

Teorema de Girsanov:  $W_t$ ,  $t \in [0, T]$ . m.B. padrão (drift 0, volatilidade 1)

$X_t$ ,  $t \in [0, T]$  - é um processo adaptado e bom pl que os integrais sejam definidos.

$$^{\circ} \xi(X)_t := \exp \left\{ \int_0^t X_s dW_s - \frac{1}{2} \int_0^t X_s^2 ds \right\}, \quad t \in [0, T]$$

$$\tilde{W}_t := W_t - \int_0^t X_s ds, \quad t \in [0, T]$$

$$E^Q[\mathbb{1}_A] = Q(A) = E^P[\mathbb{1}_A - \xi(X)_T] \quad \forall A, \quad (\text{de onde?}) \in \mathcal{F}$$

Significa Q é revisão  
de medida em  
 $(\Omega, \mathcal{F})$

Então  $\tilde{W}_t$ ,  $t \in [0, T]$  é M.B. padrão sendo visto pela  $\mathbb{Q}$ .

Nesta relação,  $\mathbb{Q} \perp \mathbb{P}$ .

Seja  $\Omega$  o coto de todos os traj. do M.B. padrão  $\Rightarrow$  seja  $\mathbb{P}$  a prob. em  $\Omega$ , com relação à qual, este M.B. padrão é M.B. padrão.

$(\Omega, \mathbb{P})$  tal que um proc. segue os traj. do acordo com  $\mathbb{P}$ , então sua distribuição das trajetórias medida nas trajetórias é a do M.B. padrão.

$\downarrow$   $\{f_t\}$  filtrado ger. por  $\{\omega_s\}$

Há  $(\Omega, \mathbb{P})$  e há coto do v.a.  $\{\omega_t\}_{t \in [0, T]}$  ( $\tilde{W}_t$ ,  $\mathbb{Q} \perp \mathbb{P}$ ) de maneira t.q., visto pelo  $\mathbb{P}$ ,  $\tilde{W}_t$ ,  $t \in [0, T]$  é M.B. padrão.

$X_t: \Omega \rightarrow \mathbb{R}$ ,  $t \in [0, T]$

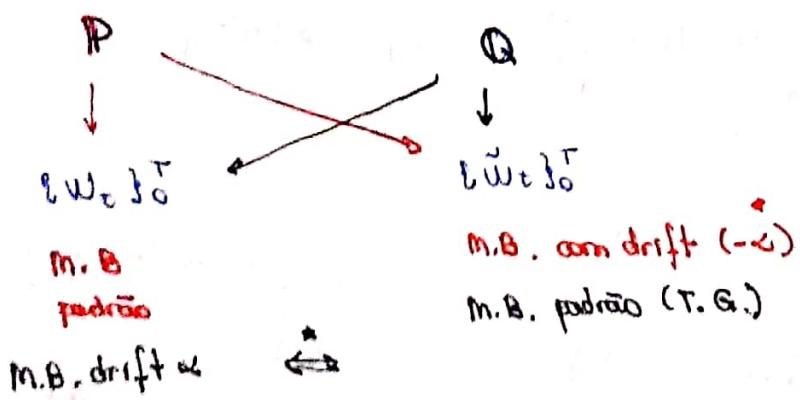
$\mathbb{E}(X_t)_t: \Omega \rightarrow \mathbb{R}$ ,  $t \in [0, T]$

A seq. de v.a.  $\{\tilde{W}_t\}_{t \in [0, T]}$  ( $\tilde{W}_t: \Omega \rightarrow \mathbb{R}$ ) forma um proc., cujas distribuições, vistas pela medi.  $\mathbb{Q}$  é o de M.B. padrão (i.e., com volatilidades e drift 0).

Escolhe  $V_t(\omega) = \alpha \mathbb{H}\omega \in \Omega$ ,  $\forall s \in [0, T]$ . Então  $\int_0^t X_s ds = \int_0^t \omega ds = \omega t$ .

Com isto, a const. genérica da T. do G. torna a ser:

$$\tilde{W}_t = W_t - \alpha t \quad (\Leftrightarrow W_t = \tilde{W}_t + \alpha t)$$



$$P\{ \sup_{0 \leq t \leq T} B_t > b, B_T \leq 3 \} = O(A), \text{ onde } A \approx \ln(3) \cdot \sup_{0 \leq t \leq T} W_t[w] > b$$

$\star O[\sup_{0 \leq t \leq T} W_t > b, W_T \leq 3] \quad e \quad W_T[w] \leq 3 \}$

$$Q[A] = E^0 [\mathbb{1}_A \xi(x)_T] = E^0 [\mathbb{1}_{\{\sup_{0 \leq t \leq T} W_t > b, N_T \leq 3\}} e^{cW_T - \frac{1}{2} c^2 T}]$$

$$\xi(x)_T = \exp \left\{ \int_0^T x \, dw_s - \frac{1}{2} \int_0^T dx_s^2 \right\}$$

$$= \exp \left\{ cW_T - \frac{1}{2} c^2 T \right\}$$

$$\Rightarrow \xi(x)_T = cW_T - \frac{1}{2} c^2 T$$



Steelt Imp. Sampling.

$$p_{00} = 1/2$$

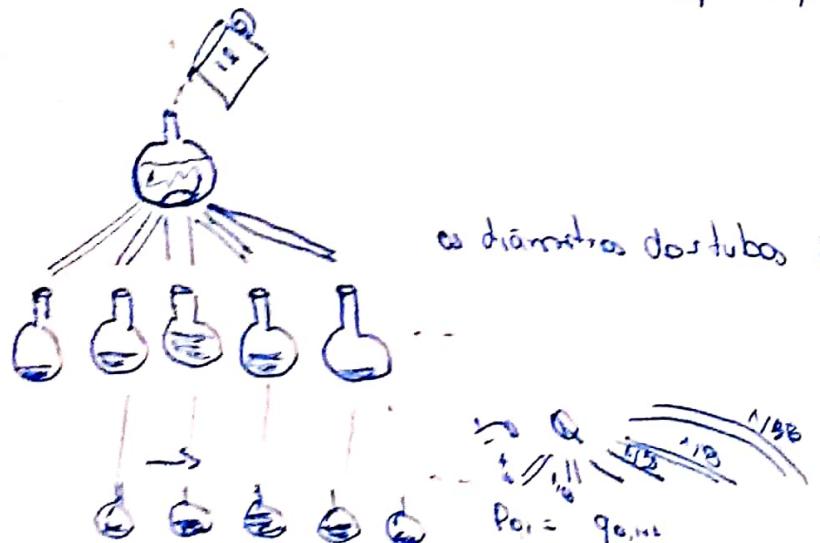
$$p_{01} = 1/8$$

$$p_{02} = 1/16$$

$$p_{03} = 1/32$$

$$\dots = 1/32$$

$$\dots = 1/32$$



$\omega$  diâmetros dos tubos, chama o c. ID

A gente que vê os seguintes pratos, vê  $X_2 + d/l$

## Teorema Ergódico - Prob. Av. II

Aula

26/11/18

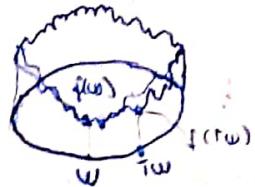
$$T: \Omega \rightarrow \Omega \cdot T = \text{rotogão}$$



círculo de pun.  $\omega$ . ( $\Omega$ )

$B$ -σ-alg. de Borel  $\mathcal{B}$ .

$P$ -med. de Lebesgue ( $\Omega, \mathcal{B}, P$ )



$f, g \rightarrow \mathbb{R}$  (finitas, B-mensuráveis)

Toss D

$$P(T^n A) = P(A) \forall A \in \mathcal{B}$$

$T$  pres. medida  $P$

Pergunta:  $\lim_{n \rightarrow \infty} \frac{(f(\omega) + f(T\omega) + \dots + f(T^{n-1}\omega))}{n}$

Intuição: Suponha que o ângulo de rotação é irracional, que não seja  $\alpha = q\pi$ ,  $q \in \mathbb{Q}$ .  
~~Então~~ Então  $\omega, T\omega, T^2\omega, \dots$  varre todo o círculo por igual (a dens. das pontas  $\omega, T\omega, T^2\omega, \dots, T^{n-1}\omega$  é +/- uniforme em  $\Omega$ ), e portanto

$$\frac{1}{n} (f(\omega) + f(T\omega) + \dots + f(T^{n-1}\omega)) \rightarrow \int_{\Omega} f(\omega) dP(\omega).$$

Teorema (Cap. 1 Billingsley "Ergodic Theory & Information")

(Acresc. pelo prof.: Seja  $(\Omega, \mathcal{B}, P)$  um espaço de probabilidades e seja  $T: \Omega \rightarrow \Omega$  a transformação que preserva a medida  $P$ ). Se  $f$  for ( $\mathcal{B}$ -mens) e integrável, então  $\exists$  uma func.  $\hat{f}: \Omega \rightarrow \Omega$  que é ( $\mathcal{B}$ -mens) e integrável que satisfaça

$$E^P(f) = E^P(\hat{f}),$$

\* Invariante com relações a  $T$  ( $\hat{f}(\omega) = \hat{f}(T\omega) \forall \omega \in \Omega$ ) e tal que

$$\lim \frac{1}{n} \sum f(T^n w) = \bar{f}(w) \quad \forall w \in \Omega.$$

Se ainda  $T$  for ERGODICO, então  $\bar{f}(w) = E^P(f) \quad \forall w \in \Omega$ , e portanto,  
 $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k w) = E^P(f) \quad \forall w \in \Omega.$

Té é ergódica com relação à medida  $P$  se para ger  $A$  com  $T^*(A) = A$ , ocorre  
que  $P(A) = 1$  ou 0.

Maximum Ergodic Lemma  $\xrightarrow{\text{Dem. Gorcia (1960): conta simp. e anti-intuitiva}}$   
 $\xrightarrow{\text{Dem. do Reisz (1945) do livro Billingsley}}$

Para cada  $\lambda \in \mathbb{R}$ , defino  $\Lambda_\lambda$  como  $\{w: \sup_{n \geq 1} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k w) > \lambda\}$ . Então

$$\int_{\Lambda_\lambda} f dP \geq \lambda P(\Lambda_\lambda).$$

$$f \rightarrow f - \lambda$$

$G := \{w: \sup_{n \geq 1} \sum_{k=0}^{n-1} f(T^k w) \geq 0\}$  e queremos provar que  $\int_G f dP \geq 0$

$\forall K$  defino  $G_K := \{w: \max_{1 \leq k \leq K} \sum_{k=0}^{n-1} f(T^k w) \geq 0\}$ , óbvio,  $G_K \xrightarrow{K \rightarrow \infty} G$ .

garante que  $\int_{G_K} f dP \xrightarrow{K \rightarrow \infty} \int_G f dP$

Fato: se  $a_n \rightarrow a$ , então  $\frac{1}{n}(a_1 + \dots + a_n) \rightarrow a$ .

Se provarmos que  $\sum_{k=1}^n \int_{G_K} f dP \geq 0 \quad \forall n = 1, 2, \dots$ , teríamos que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \int_{G_K} f dP \text{ pode ser } \geq 0 \iff$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left( \int_{G_1} f dP + \int_{G_2} f dP + \dots + \int_{G_n} f dP \right) \text{ só pode ser } \geq 0$$

$$\Rightarrow \int f dP \text{ só pode ser } \geq 0$$

Para provar "obj." p/ n=2 queremos provar que  $\int_{G_1} f dP + \int_{G_2} f dP \geq 0$  ( $\sum_{i=1}^2 \int_{G_i} f dP \geq 0$ )

$$G_1 = \{\omega : f(\omega) \geq 0\}, G_2 = \{\omega : \text{ou } f(\omega) \geq 0 \text{ ou } f(T\omega) \geq 0\}$$

$$G'_1 = \{\omega : f(T\omega) \geq 0\} \Rightarrow T^{-1}(G_1) = G'_1.$$

$$\int_{G'_1} f(T\omega) dP(d\omega) + \int_{G_2} f(\omega) dP(d\omega) \geq 0$$

$$\int_{G'_1 \cap G_2} (f(T\omega) + f(\omega)) dP + \int_{G'_1 \setminus G_2} f(T\omega) dP + \int_{G_2 \setminus G'_1} f(\omega) dP \geq 0$$

Em  $G'_1 \cap G_2$ ,  $f(T\omega) \geq 0$  devido à def. de  $G'_1$ . Port., se  $f(T\omega) + f(\omega) \leq 0$ ,  $f(\omega) \leq 0$ .

No caso contrário,  $f(\omega)$  seria obrigado a ser  $\leq -f(T\omega)$ , mas neste caso,  $\omega \notin G_2$ .

(

Portanto, só pode ser que  $f(\omega) + f(T\omega) \geq 0 \Rightarrow \int_{G'_1 \cap G_2} (f(T\omega) + f(\omega)) dP \geq 0$

$$\text{Em } G'_1 \setminus G_2, f(T\omega) \geq 0 \text{ (devido à def. de } G'_1) \Rightarrow \int_{G'_1 \setminus G_2} f(T\omega) dP \geq 0.$$

$$\text{Em } G_2 \setminus G'_1, f(T\omega) \leq 0 \text{ (devido à def. } G'_1\text{). Mas, se } \omega \in G_2 \setminus G'_1, \text{ ou } f(\omega) \geq 0 \text{ ou } f(\omega) + f(T\omega) \geq 0 \Rightarrow f(\omega) \text{ deve ser } \geq 0 \Rightarrow \int_{G_2 \setminus G'_1} f(\omega) dP \geq 0.$$

$$G_1 = \{\omega : f(\omega) > 0\}$$

$$G_2 = \{\omega : f(\omega) > 0 \text{ ou } f(\omega) + f(T\omega) > 0\}$$

$$G_3 = \{\omega : f(\omega) > 0 \text{ ou } f(\omega) + f(T\omega) > 0 \text{ ou } f(\omega) + f(T^2\omega) > 0\}$$

$$G_{n-2} = \{\omega : \text{ou } f(\omega) > 0 \text{ ou } f(\omega) + f(T\omega) > 0 \text{ ou } \dots \text{ ou } f(\omega) + f(T\omega) + \dots + f(T^{n-3}\omega) > 0\}$$

$$G_{n-1} = \{\omega : \text{ou } f(\omega) > 0 \text{ ou } \dots \text{ ou } f(\omega) + \dots + f(T^{n-3}\omega) > 0 \text{ ou } f(\omega) + \dots + f(T^{n-2}\omega) > 0\}$$

$$G_n = \{\omega : \text{ou } f(\omega) > 0 \text{ ou } \dots \text{ ou } f(\omega) + \dots + f(T^{n-1}\omega) > 0\}$$

Obs.  $\int_{G_1} f(\omega) dP + \dots + \int_{G_n} f(\omega) dP > 0.$

$$H_0 = G_n = \{\omega : f(\omega) > 0 \text{ ou } f(\omega) + f(T\omega) > 0 \text{ ou } \dots \text{ ou } f(\omega) + f(T\omega) + \dots + f(T^{n-1}\omega) > 0\}$$

$$H_1 = T^{-1}(G_{n-1}) = \{\omega : f(T\omega) > 0 \text{ ou } f(T\omega) + f(T^2\omega) > 0 \text{ ou } \dots \text{ ou } f(T\omega) + f(T^2\omega) + \dots + f(T^{n-1}\omega) > 0\}$$

$$H_2 = T^{-2}(G_{n-2}) = \{\omega : f(T^2\omega) > 0 \text{ ou } f(T^2\omega) + f(T^3\omega) > 0 \text{ ou } \dots \text{ ou } f(T^2\omega) + f(T^3\omega) + \dots + f(T^{n-1}\omega) > 0\}$$

$$H_{n-3} = T^{-(n-3)}(G_3) = \{\omega : f(T^{n-3}\omega) > 0 \text{ ou } f(T^{n-3}\omega) + f(T^{n-2}\omega) > 0 \text{ ou } f(T^{n-3}\omega) + f(T^{n-2}\omega) + f(T^{n-1}\omega) > 0\}$$

$$H_{n-2} = T^{-(n-2)}(G_2) = \{\omega : f(T^{n-2}\omega) > 0 \text{ ou } f(T^{n-2}\omega) + f(T^{n-1}\omega) > 0\}$$

$$H_{n-1} = T^{-(n-1)}(G_1) = \{\omega : f(T^{n-1}\omega) > 0\}$$

Como  $T$  preserva  $P$ , queremos provar

$$\int_{H_0} f(\omega) dP(d\omega) + \int_{H_1} f(T\omega) dP(d\omega) + \dots + \int_{H_{n-3}} f(T^{n-3}\omega) dP(d\omega) + \int_{H_{n-2}} f(T^{n-2}\omega) dP(d\omega) \\ + \int_{H_{n-1}} f(T^{n-1}\omega) dP(d\omega) > 0$$

soma de laderes

$$\forall \omega \in \bigcup_{n=0}^{\infty} I_n(\omega), f(\omega) I_{n_0}(\omega) + f(T\omega) I_{n_1}(\omega) + \dots + f(T^{n-2}\omega) I_{n_{n-2}}(\omega) + f(T^{n-1}\omega) I_{n_{n-1}}(\omega) > 0.$$

$> 0$

Lema: Em  $c_1, \dots, c_n$  (finita), chama  $c_i$  de "líder" se

$$\max(c_i, c_i + c_{i+1}, \dots, c_i + c_{i+n-1} + \dots + c_n) > 0$$

Então a soma dos líderes é positiva.

# Probabilidade Avançada

Aula

01/12/2015

## Relação entre T. Ergódicos e D. Est.

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

$T: \Omega \rightarrow \Omega$  que preserva medida ( $P(T^{-1}A) = P(A) \forall A \in \mathcal{F}$ ).

$f: \Omega \rightarrow \mathbb{R}$  (v.a.)

Resultado:

$\lim_{n \rightarrow \infty} \frac{1}{n} (f(\omega) + f(T\omega) + \dots + f(T^{n-1}\omega))$  existe e temos a expressão.

1. Relação assunto aula passada com proe. estacionários?

Definição: Um processo  $\{\xi_i\}_{i=0}^{\infty}$  chama-se estacionário se para cada  $n$  e cada  $m \in \mathbb{N}$ ,

$$(\xi_0, \xi_1, \dots, \xi_n) \in (\xi_m, \xi_{m+1}, \dots, \xi_{m+n})$$

têm a mesma distribuição.

Teorema: Para qualquer processo estacionário  $\{\xi_i\}_{i=1}^{\infty}$ , existem

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

2.  $T$  que preserva  $P$  ( $T: \Omega \rightarrow \Omega$ )

3. v.a.  $X: \Omega \rightarrow \mathbb{R}$

tais que

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

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

$$\{\xi_i\}_{i=0}^{\infty} \stackrel{D}{=} \{X \circ T^i\}_{i=0}^{\infty}$$

## Prova do Teorema

Tomo  $(\Omega, \mathcal{F}, P)$ ,  $T: \Omega \rightarrow \Omega$  que preserva  $P$ , e  $X: \Omega \rightarrow \mathbb{R}^n$ , e quero verificar que  $X, X_0 T, X_0 T^2, \dots$  é estacionário.

Tomo  $B \in \mathcal{B}(\mathbb{R}^3)$  e quero verificar que

$$P(\{\omega : (X(\omega), X_0 T(\omega), X_0 T^2(\omega)) \in B\}) =$$

$$P(\{\omega : (X_0 T(\omega), X_0 T^2(\omega), X_0 T^3(\omega)) \in B\})$$

$$A = \{\omega : (X(\omega), X_0 T(\omega), X_0 T^2(\omega)) \in B\}$$

$$A_1 = \{\omega : (X_0 T(\omega), X_0 T^2(\omega), X_0 T^3(\omega)) \in B\}$$

Tomo  $\omega \in A_1 \Rightarrow X(\omega), X_0 T(\omega), X_0 T^2(\omega) \in B \Rightarrow \omega \in A$

Tomo  $\omega \in A_1 \Rightarrow (X_0 T(\omega), X_0 T^2(\omega), X_0 T^3(\omega)) \in B \Rightarrow$

$$(X_0(T\omega), X_0 T^2(T\omega), X_0 T^3(T\omega)) \in B \Rightarrow T\omega \in A \text{ e } T=T\omega$$

$$\vec{\xi} = \{\xi_i\}_{i=1}^\infty$$

$P_{\vec{\xi}}$  é a medida induzida pelo processo em  $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$

$(\Omega, \mathcal{F}, P)$   $\vec{\xi} = \{\xi_i\}_{i=1}^\infty$  gera  $P_{\vec{\xi}}$  def. em  $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$

Seja  $\vec{X} = \{X_i\}_{i=1}^\infty$  essa vol. em  $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty)) \rightarrow (\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$

$$\stackrel{\uparrow}{P_{\vec{\xi}}}$$

$$\tau = (\tau_1, \tau_2, \dots)$$

$$\vec{X}(\tau) = \vec{X}(\tau_1, \tau_2, \tau_3, \dots) \stackrel{\text{definição}}{=} (\tau_1, \tau_2, \dots)$$

leto  $\epsilon$ ,  $X_\epsilon = \text{Id}: \mathbb{R} \rightarrow \mathbb{R}$

Defino  $T: \mathbb{R}^\infty \rightarrow \mathbb{R}^\infty$

$$T(\tau_1, \tau_2, \tau_3, \dots) \stackrel{\text{def}}{=} (\tau_1, \tau_2, \dots) \text{ T apaga a primeira estrela e shift o que restou de esquerda}$$

2. Quais os eventos invariantes para  $T$ ?

Definição inv. se  $T^{-1}(B) = B$ , se  $B$  é inv. p/  $T$ , onde  $B$  não pode dep. do prim. coordenada para

$$\{\sigma \in \mathbb{R}^\omega : \sum_{i=1}^{\infty} \sigma_i > 0\} \text{ não é contável}$$

$$\{\sigma \in \mathbb{R}^\omega : \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \sigma_i > 0\} \text{ é contável}$$

$$T^{-1}(\sigma_1, \sigma_2, \dots) = (\sigma_1, \sigma_2, \sigma_3, \dots)$$

A  $\sigma$ -álgebra de ev. invariantes para  $T$  é a  $\sigma$ -álg. contável em  $(\mathbb{R}^\omega, \mathcal{B}(\mathbb{R}^\omega))$ .

Teorema Ergódico. Seja  $T$  uma transf. que preserva medida em  $(\Omega, \mathcal{F}, P)$  e seja  $\xi : \Omega \rightarrow \mathbb{R}$  com  $E^P[\xi] < \infty$ . Então,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \xi \circ T^i(\omega) = E^P[\xi | \mathcal{F}] \text{ P-q.o.}$$

onde  $\mathcal{F}$  é a  $\sigma$ -álgebra de ev. invariantes de  $T$ .

Se  $T$  é ergódico em relação com  $P$  (cada  $B \in \mathcal{F}$  t.q.  $P(B) \neq 0$  ou 0) então  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \xi \circ T^i(\omega) = E^P[\xi] \text{ P-q.o.}$

## SETS AND CLASSES

[Sec. 3]

As the final and frequently very important operation on sets we introduce the symmetric difference of two sets  $E$  and  $F$ , denoted by

and defined by

$$E \Delta F,$$

The formation of limits, complements, and differences of sets requires a bit of practice for ease in manipulation. The reader is accordingly advised to carry through the proofs of the most important properties of these processes, listed in the exercises that follow.

(1) Another heuristic motivation of the convention

is the desire to have the identity

$$\bigcap_{\gamma \in \Gamma} E_\gamma = X$$

which is valid for all non empty index sets  $\Gamma$ , remain valid for  $\Gamma = 0$ .

(2) If  $E_* = \liminf_n E_n$  and  $E^* = \limsup_n E_n$ , then

$$E_* = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} E_m \subset \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m = E^*.$$

(3) The superior limit, inferior limit, and limit (if it exists) of a sequence of sets are unaltered if a finite number of the terms of the sequence are changed.

(4) If  $E_n = A$  or  $B$  according as  $n$  is even or odd, then

$\liminf_n E_n = A \cap B$  and  $\limsup_n E_n = A \cup B$ .

(5) If  $\{E_n\}$  is a disjoint sequence, then

$$\lim_n E_n = 0.$$

(6) If  $E_* = \liminf_n E_n$  and  $E^* = \limsup_n E_n$ , then

$$(E_*)' = \limsup_n E_n' \text{ and } (E^*)' = \liminf_n E_n'.$$

More generally,

$$F - E_* = \limsup_n (F - E_n) \text{ and } F - E^* = \liminf_n (F - E_n).$$

$$(7) E - F = E - (E \cap F) = (E \cup F) - F,$$

$$E \cap (F - G) = (E \cap F) - (E \cap G), \quad (E \cup F) - G = (E - G) \cup (F - G).$$

$$(8) (E - G) \cap (F - G) = (E \cap F) - G,$$

$$(E - F) - G = E - (F \cup G), \quad E - (F - G) = (E - F) \cup (E \cap G),$$

$$(E - F) \cap (G - H) = (E \cap G) - (F \cup H).$$

[Sec. 13]

## EXTENSION OF MEASURES

The results of §§ 10, 11, and 12, together with Theorem

... .

## MEASURES AND OUTER MEASURES

... on a hereditary

[Sec. 4]

## SETS AND CLASSES

$$(9) E \Delta F = F \Delta E, \quad E \Delta (F \Delta G) = (E \Delta F) \Delta G,$$

$$E \cap (F \Delta G) = (E \cap F) \Delta (E \cap G),$$

$$E \Delta 0 = E, \quad E \Delta X = E,$$

$$E \Delta E = 0, \quad E \Delta E' = X,$$

$$E \Delta F = (E \cup F) - (E \cap F).$$

(10) Does the class of all subsets of  $X$  form a group with respect to the operation  $\Delta$ ?

(11) If  $E_* = \liminf_n E_n$  and  $E^* = \limsup_n E_n$ , then

$$\chi_{E_*}(x) = \liminf_n \chi_{E_n}(x) \text{ and } \chi_{E^*}(x) = \limsup_n \chi_{E_n}(x).$$

(The expressions on the right sides of these equations refer, of course, to the usual numerical concepts of superior limit and inferior limit.)

$$(12) \chi_{E^*} = 1 - \chi_E, \quad \chi_{E-P} = \chi_E(1 - \chi_P),$$

$$\chi_{E \Delta P} = |\chi_E - \chi_P| = \chi_E + \chi_P \pmod{2}.$$

(13) If  $\{E_n\}$  is a sequence of sets, write

$$D_1 = E_1, \quad D_2 = D_1 \Delta E_2, \quad D_3 = D_2 \Delta E_3,$$

and, in general,

$$D_{n+1} = D_n \Delta E_{n+1} \text{ for } n = 1, 2, \dots$$

The limit of the sequence  $\{D_n\}$  exists if and only if  $\lim_n E_n = 0$ . If the operation  $\Delta$  is thought of as addition (cf. (12)), then this result has the following verbal phrasing: an infinite series of sets converges if and only if its terms approach zero.

## § 4. RINGS AND ALGEBRAS

A ring (or Boolean ring) of sets is a non empty class  $R$  of sets such that if

$$E \in R \text{ and } F \in R,$$

$$E \cup F \in R \text{ and } E - F \in R.$$

In other words a ring is a non empty class of sets which is closed under the formation of unions and differences.

The empty set belongs to every ring  $R$ , for if

$$E \in R,$$

then

$$0 = E - E \in R.$$

### Aula 3. Martingais; Conceitos Básicos

1. Nesta aula e em série de aulas subsequentes, eu vou definir o que significa um processo ser martingal, e vou mostrar algumas propriedades de martingais. Entre estas, há três que são as mais importantes no sentido de que se você me perguntar sobre a utilidade de martingais para fins teóricos e práticos e se analisar minhas respostas então você identificará facilmente três propriedades que são as mais usadas. São estas que rotulei por mais importantes. Eis são elas:

- (1) os incrementos de um martingal possuem correlação linear nula;
- (2) um martingal parado é martingale, ou, sendo formulado em outros termos que

$$\mathbb{E}[M] = \mathbb{E}[M_0] \quad \{ M_{\tau \wedge n} \}_{n=0}^{\infty} \text{ é martingal}$$

onde  $M_0, M_1, \dots$  é um martingal e  $\tau$  é um tempo de parada;

- (3) um martingale limitado converge quase certamente.

Então, minhas aulas sobre conceitos básicos de martingais foram divididas em três módulos. O conteúdo de cada módulo é o que há de mais atrelado com a correspondente propriedade.

2. Só martingais de tempo discreto serão tratados neste texto. Martingais de tempo contínuo aparecerão lá muito pela frente quando falarmos do Movimento Browniano e Cálculo Estocástico. O tempo discreto é o que proporciona maior comodidade para compreensão intuitiva do assunto.

3. Também, com o intuito de se livrar dos detalhes que mais chateiam que ensinam, vamos trabalhar só com o caso quando

$$\rightarrow M_0 \equiv 0, \text{ isto é: no tempo } 0, \text{ o valor do martingal é zero.} \quad (1)$$

Os casos opostos podem ser separados em duas classes: na primeira classe entram aqueles nos quais  $M_0$  é uma constante diferente de zero, e no segundo classe entram aqueles nos quais  $M_0$  é uma variável aleatória. Os casos da primeira classe pode ser trivialmente levados ao caso (1). A segunda classe não é mais distante que a primeira do caso (1) só quando  $\mathbb{E}[|M_0|] < \infty$ <sup>1</sup>. Já quando  $M_0$  é tal que  $\mathbb{E}[M_0]$  ou não existe, ou é  $-\infty$  ou  $+\infty$ , então temos um problema. A raiz do problema é que cada  $\mathbb{E}[M_k]$  repete o comportamento de  $\mathbb{E}[M_0]$ , mas o que incomoda são as brotinhos crescidos a partir desta raiz: a variância de incrementos do processo definida como  $\text{Var}(M_{k+1} - M_k)$  não existe. Isto nos tristeia pois muitas aplicações de martingais baséiam-se na finitude da variância de seus incrementos. A saída aliviadora é muito simples: definir um processo  $X_1, X_2, \dots$  de variáveis com variâncias finitas e definir que dado que  $M_0 = x$ , a distribuição de  $M_k$  é  $x + X_1 + \dots + X_k$ ; aqui  $x$  pode ser qualquer número real e a distribuição das  $X$ 's pode até depender de  $x$ .

4. Junto com o presente texto você recebe a Seção 11 do Cap I e Seção 1 do Cap. VII do livro de Shiriaev, e Capítulos 10 e 11 do livro de Williams. É essencial que sejam feitos por você todos os exercícios sugeridos pelo Shiriaev na sua Seção 11 do Cap I. Alguns destes são idênticos ou parecidos com exercícios formulados abaixo no meu texto. Não faz mal. Ah, tenho que comentar que não entendo mais que 90% do conteúdo do Cap. VII do livro de Shiriaev.

<sup>1</sup>Certifique-se que você entende o significado de " $\mathbb{E}[|M_0|] < \infty$ "; para um auto-teste, responda se este é a mesma coisa que dizer "X é quadrado integrável".

### Aula 3. Martingais; Conceitos Básicos. A não correlação linear de incrementos.

#### Definição de Martingal

5. Eis definição de martingal: tendo um processo estocástico  $\{M_n\}_{n=0}^{\infty}$  e uma filtração  $\{\mathcal{F}_n\}_{n=0}^{\infty}$ , o processo chama-se martingale em relação à filtração caso cada  $M_n$  for  $\mathcal{F}_n$  mensurável e

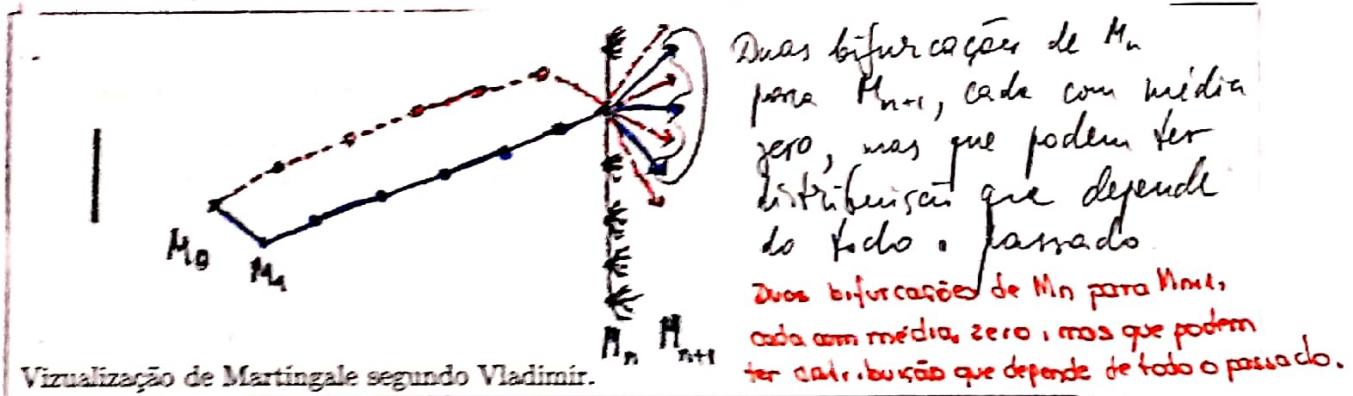
$$E[M_{n+1} | \mathcal{F}_n] = M_n, \text{ valha para todo } n \geq 0. \quad (2)$$

Por certo, devemos ter dito em (2) que a igualdade ocorre quase certamente, mas omitimos esta ressalva aqui assim como omitiremos-a em todos os outros lugares onde esta não requer uma atenção especial (durante a aula foi explicada a necessidade do acrescimo de "quase certamente" à relação (2)).

O nome martingale em relação à filtração é substituído pelo simples *martingale* quando a filtração for a "natural gerada pelo próprio processo", isto é, quando

$$\mathcal{F}_n = \sigma(M_0, M_1, \dots, M_n) \text{ para cada } n \geq 0.$$

A definição de martingale está interpretada graficamente no desenho abaixo. É a imagem que eu tenho na cabeça caso alguém me dizer: "Imagine um martingale..."



6. A propriedade básica de martingais, empregada muito em demonstrações de propriedades mais refinadas e menos genéricas, alega: "os incrementos de um martingale têm correlação linear nula". Em termos mais formais, isto significa o seguinte: para quaisquer quatro instantes de tempo  $s, t, u, v$  satisfazendo  $0 \leq s < t \leq u < v < \infty$  (observe:  $t$  e  $u$  podem coincidir) vale que

$$E[(M_t - M_s - E[M_t - M_s])(M_v - M_u - E[M_v - M_u])] = 0. \quad (3)$$

Exercício 1. Prove que  $E[E(M_t - M_s) | \mathcal{F}_s] = E[M_t - E(M_t) | \mathcal{F}_s] = 0$ .

$$E[M_t - M_s] = 0 \text{ para quaisquer } t \text{ e } s. \quad (4)$$

Observe que o fato provado justifica que (3) pode ser re-escrito assim:

$$E[(M_t - M_s)(M_v - M_u)] = 0. \quad (5)$$

A expressão (5) é a mais usada que a (3).

Gracias à relação (4) podemos provar (5) em vez de (3). Eis a conta que prova a relação:

$$\begin{aligned} E[(M_t - M_s)(M_v - M_u)] &= E[E\{(M_t - M_s)(M_v - M_u) | \mathcal{F}_t\}] \\ &= E[(M_t - M_s)E\{(M_v - M_u) | \mathcal{F}_t\}] \\ &= E[(M_t - M_s) \times 0] = 0. \end{aligned} \quad (6)$$

A ser acrescida.

(cf. (6))

Aqui desenhamos aquilo que a conta acima fez para gente.

**Exercício 2.** Há pessoas que confundem incrementos de processo com valores de processo. Preste atenção: a expressão  $\mathbb{E}[M_s M_t] = 0$  não corresponde ao não correlacionamento linear entre os incrementos de martingal  $\{M_n\}_{n=0}^{\infty}$ . Para que este fato seja bem claro, sugiro que faça o seguinte: expresse  $\mathbb{E}[M_s M_t]$  em termos de esperanças e variâncias das variáveis aleatórias do martingal. Se o caso genérico causou problemas insuperáveis, considere um caso particular no qual seu martingal é o Passo Aleatório Simples Simétrico (veja Ex. 5). Dê exemplo de martingale para qual  $\mathbb{E}[M_s M_t]$  cresce em  $t$  (estou querendo que o valor desta expressão cresça quando  $t$  aumenta mas  $s$  fica congelado). Você conseguiria construir martingale para qual  $\mathbb{E}[M_s M_t]$  decresce em  $t$ ?

7. O uso da propriedade de não correlacionamento linear entre incrementos de martinlal é tal frequente que induz a gente a pensar que essa é equivalente à propriedade principal de martingalidade, isto é, à (2). Em outras palavras, muitas pessoas – eu entre tais – se equivocam ao pensar que

Equívoco: se construir um processo  $\{M_n\}$  da maneira tal que  $\mathbb{E}[M_i] = 0$  para todo  $i$ , e tal que seus incrementos tenham correlação linear nula, então o processo construído será obrigatoriamente um martingal.

Faça então o próximo exercício para eliminar toda e qualquer dúvida a respeito deste pensamento.

**Exercício 3.** Construa um processo  $\{M_n\}$  que não seja martingale, mas que  $\mathbb{E}[M_i] = 0$  para todo  $i$ , e tal que seus incrementos tenham a correlação linear nula.

A tarefa talvez seja pouco difícil para o estado presente de seu conhecimento da estrutura de martingais. Pode adiar a solução deste exercício para depois do Ex. 10, mas faça o exercício seguinte. Ele não é tal complicado como o anterior, mas mostra que o não correlacionamento linear entre incrementos de martingal significa somente a anulação de esperança de produtos cruzados mas não a ausência da dependência. Isto sugere que é possível construir um processo  $(\{M_n\}_{n=0}^{\infty})$  de variáveis com médio nula ( $\mathbb{E}[M_n] = 0, \forall n$ ) cujos incrementos tenham a correlação linear nula ( $\mathbb{E}[(M_t - M_s)(M_v - M_u)] = 0$ ), mas que as variáveis do processo estejam amarrados pela dependência específica tal que  $\mathbb{E}[M_{n+1} | \mathcal{F}_n] \neq M_n$ . Se você conseguir isto, entenderá então que a estrutura de dependência entre variáveis de um martingale não é qualquer uma dentro das infinitas possíveis que proporciona a não correlacionamento linear.

**Exercício 4.** Escolhe arbitrariamente  $t$  e  $s$ , dois instantes de tempo com  $t > s$ . Tome uma função  $f : \mathbb{R}^{s+1} \rightarrow \mathbb{R}$ . Prove que

$$\mathbb{E}[(M_t - M_s) \times f(M_s, M_{s-1}, \dots, M_0)] = 0; \quad (7)$$

Observe que a demonstração de (7) não exigiu nenhuma idéia nova em comparação com que precisava para provar (5). É esta observação que deve indicar para você que a propriedade (2), que define martingais, proporciona muito mais de que simples não correlacionamento linear de incrementos.

Um resumo ao equívoco discutido na presente seção, cabe aqui a seguinte observação. Eu acho que este equívoco visto de modo simplista sobre a estrutura de martingais, e esta, por sua vez, foi causada pelo exemplo que é sempre o primeiro a aparecer em todos os livros texto sobre martingais: é o processo chamado Processo Aleatório Simples Simétrico. Sobre este é o próximo exercício.

**Exercício 5.** Sejam  $X_1, X_2, \dots, X_n$  independentes e identicamente distribuídos com  $P[X_i = \pm 1] = 1/2$ . Vamos dizer como incrementos do processo  $S_0, S_1, \dots, S_n$ , onde  $S_0 = 0$ . O processo  $\{S_i\}$  chama-se Processo Aleatório Simples Simétrico (PASS). Mostre que ele é martingal. Verifique que ele é Cadêa de Markov. Verifique que seus incrementos são independentes. Estas duas propriedades não são típicas para martingais. E por isso que não gosto de PASS como exemplo de martingal.

**8.** A seção anterior enfatizou o fato de que a ausência da correlação linear entre incrementos é uma das muitas consequências da definição de martingal. Isto desencadeia a pergunta: "Qual forte é a dependência entre variáveis que possuem correlação linear nula entre si?" Neste sentido, é bom lembrar que a correlação linear é uma medida muito fraca e frágil de independência entre variáveis aleatórias. Para colocar este fato no lugar adequado de sua mente, lhe convido a fazer os exercícios da presente seção.

**Exercício 6.** A tabela de desenho abaixo apresenta a distribuição conjunta de variáveis  $X$  e  $Y$ . Escolhe as probabilidade  $p_1, p_2, p_3$  para que  $E[X] = 0$ ,  $E[Y] = 0$  e  $\text{cov}(X, Y) = E[XY] = 0$ . Observe que há três equações lineares para três incógnitas ( $p_1, p_2, p_3$ ). Será que a solução existe? É única? Esta solução (ou soluções, caso existam mais que uma) implica na independência entre  $X$  e  $Y$ ?

| $X \setminus Y$ | -1    | 0                     | 1     |                                                                                                                                                                                               |
|-----------------|-------|-----------------------|-------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| -1              | $p_1$ | $p_2$                 |       | $p_1 + p_2 = 1/4$                                                                                                                                                                             |
| 0               |       |                       | $p_3$ | $1 - p_1 - p_2 = p_1 + p_3 = 0$                                                                                                                                                               |
| 1               | $p_3$ | $1 - p_1 - p_2 - p_3$ |       | $1 - 2p_1 - 2p_3 = 0 \quad p_1 = p_3 = 0$<br>$1 - 2p_2 - 2p_3 = 0 \quad p_2 = p_3 = 0$<br>$1 - 2p_1 - 2p_2 = 0 \quad p_1 = p_2 = 0$<br>$1 - 2p_1 - 2p_2 - 2p_3 = 0 \quad p_1 = p_2 = p_3 = 0$ |

Distribuição bivariada do Ex. 6

Diagrama de dispersão do par  $(X, Y)$  do Ex. 6

**Exercício 7.** Sejam  $X$  e  $Y$  quaisquer com médias nulas e variâncias finitas. Suponha que  $\text{cov}(X, Y) = E[XY] \neq 0$ . Mostre que é possível achar  $\alpha \in \mathbb{R}$  tal que  $\text{cov}(X + \alpha Y, Y) = 0$ . Observe que  $X + \alpha Y$  não é obrigado a ser independente de  $Y$ . Tal observação deve reforçar a compreensão de que a anulação da covariância entre duas variáveis aleatórias não implica no estabelecimento de independência entre elas.

**Exercício 8.** O diagrama de dispersão do desenho acima aparece na maioria de livros-texto sobre conceitos básicos da Estatística com o intuito de mostrar, da maneira visual e inegável, que duas variáveis aleatórias podem ser fortemente dependentes e mesmo assim possuir o coeficiente de correlação linear igual a zero. Os pontos do desenho marcam os pares de valores nas quais a distribuição de par  $(X, Y)$  está concentrada. Suponha que todos os pontos têm o mesmo peso e confirme que  $\text{cov}(X, Y) = 0$ . Qual é o argumento formal que prova que  $X$  e  $Y$  não são independentes?

**9.** Por causa da ausência de covariância entre incrementos de um martingale, este possui a seguinte propriedade

$$\text{Var}[M_n] = \text{Var}[M_1 - M_0] + \text{Var}[M_2 - M_1] + \cdots + \text{Var}[M_n - M_{n-1}]. \quad (8)$$

Esta é muito útil quando combinada com o Teorema de Convergência de Martingais. Os resultados desta combinação estão apresentados no Capítulo 12 do livro de Williams.

Exercício 9. Verifique que (8) é de fato, válido.

A propriedade (8) é a porta para o próximo exercício, porém o exercício vai nos proporcionar muito mais de que uma simples aplicação de (8).

Exercício 10. Nas minhas aulas da disciplina “Estatística Básica” ministrada para a faculdade de Zootecnia e Medicina Veterinária da USP, eu ofereci para alunos o seguinte problema:

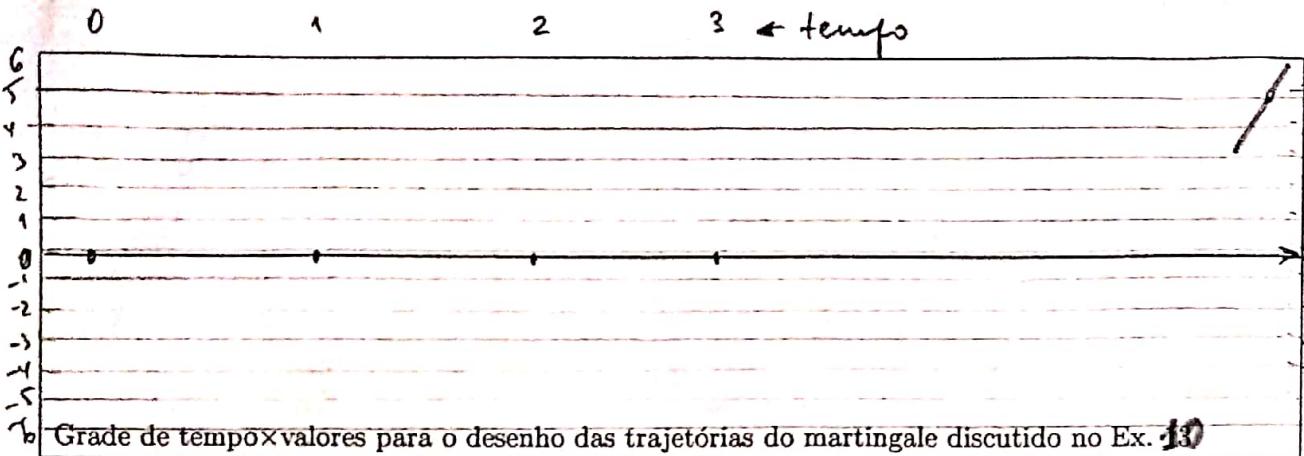
*Pedro e José resolveram jogas o seguinte jogo. O jogo constitui-se de dois passos. No primeiro passo lança-se uma moeda honesta (digamos que esta é a moeda de R\$1) e Pedro ganha R\$1 de José caso der “cara”, e paga R\$1 a José caso der “coroa”. Caso o primeiro lançamento acabar em “cara”, então no segundo passo do jogo será lançada uma moeda honesta (digamos que esta é a moeda de 50 centavos), e Pedro ganhará R\$1 de José na “cara”, pagará R\$1 a José na “coroa”. Já caso o primeiro lançamento der “coroa”, então no segundo passo será lançado um tetraedro equilibrado, cujas faces são marcadas com os números -2, -1, 1, 2, e Pedro ganhará/perderá o valor igual ao resultado do lançamento deste dado (imagine que este tetraédro é azul – a cor que lhe permitirá distingui-lo de outros tetráedros que aparecerão no item (iii) abaixo).*

*Designamos por  $X$ ,  $Y$  respectivamente o ganho/perda de Pedro no primeiro e no segundo passo do jogo, e designamos por  $Z$  o ganho/perda total de Pedro neste jogo.*

- Construa as distribuições das variáveis aleatórias  $X$ ,  $Y$  e  $Z$ .*
- Verifique se  $X$  e  $Y$  não são independentes.*
- Calcule  $E[X]$ ,  $E[Y]$ ,  $E[Z]$ ,  $\text{Var}[X]$ ,  $\text{Var}[Y]$ ,  $\text{Var}[Z]$ .*
- Note que as nossas definição garantem que  $Z = X + Y$ . Esta igualdade garante que  $E[Z] = E[X] + E[Y]$ ? Verifique se esta igualdade é válida usando os resultados do item (c).*
- A igualdade  $Z = X + Y$  garante que  $\text{Var}[Z] = \text{Var}[X] + \text{Var}[Y]$ ? Verifique se esta igualdade é válida usando os resultados do item (c).*

Este exercício foi composto para chamar a atenção de alunos ao fato de que  $\text{Var}[X + Y]$  não sempre é igual a  $\text{Var}[X] + \text{Var}[Y]$ . Minha intenção era construir  $X$  e  $Y$  da maneira tal que elas sejam dependentes, e solicitar que os alunos calculem  $\text{Var}[X + Y]$  e  $\text{Var}[X] + \text{Var}[Y]$  e comparem os valores calculados confirmando a desigualdade entre estes. É óbvio que poderia construir variáveis aleatórias dependentes da mesma maneira que usei no Ex. 6 acima. Mas tal construção é abstrata demais para alunos de Veterinária, enquanto que lançar moedas, dados, tetráedros é o caminho mais intuitivo para eles. Isto me motivou para a inventar toda esta história sobre Pedro e José. Minha expectativa era que ao fazer as contas, os alunos obteriam valores diferentes para  $\text{Var}[X + Y]$  e para  $\text{Var}[X] + \text{Var}[Y]$ . Porem, para minha total surpresa, os dois valores coincidiram! Achei que foi uma coincidência infeliz. Só muito tempo depois descubri que a coincidência é uma regra que está garantida pela propriedade genérica. Isto é o conteúdo do item (i) da série de minhas perguntas a você:

- Prove que (8) garante que no meu exemplo ocorre a igualdade  $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$ .*
- Na solução do item anterior você usou minhas  $X$  e  $Y$  para construir seu martingale  $M_0, M_1, M_2$ . Desenhe todas as trajetórias do processo  $M_0, M_1, M_2$  na grade fornecida no desenho abaixo.*



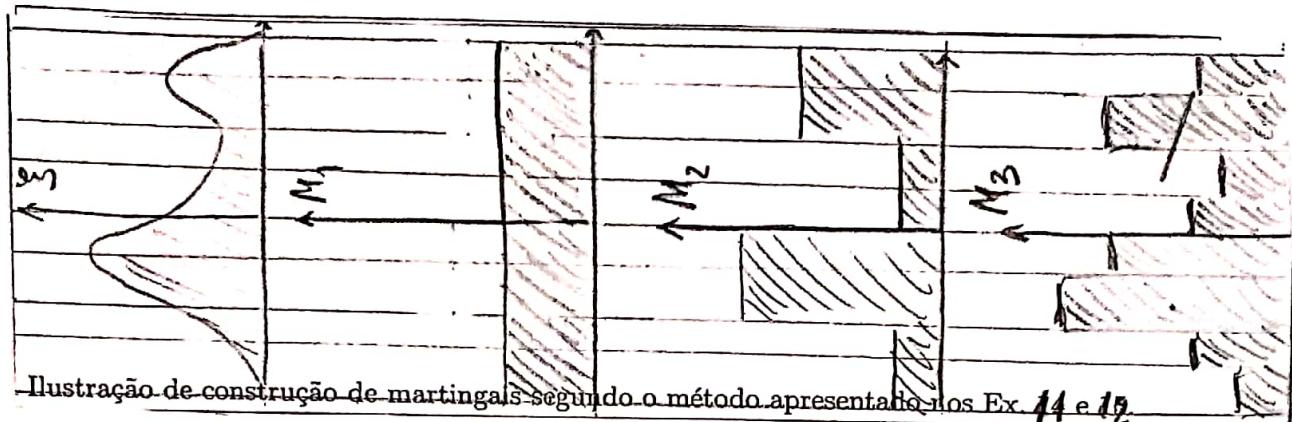
- (iii) Olhando no seu desenho (o que foi solicitado no item anterior), você se convence facilmente que seu processo  $M_0, M_1, M_2$  é Cadeia de Markov. A razão para tal é simples: o passado pode ter um tempo só, e assim não dá para construir a dependência da passado pretérito que faça com que o processo tenha memória longa e consequentemente deixa de ser uma Cadéia de Markov. Em virtude desta observação, vamos acrescentar mais um tempo para quebrar a markovidade: defina  $M_3$  da maneira tal que  $M_0, M_1, M_2, M_3$  não seja Cadéia de Markov. Dou-lhe moedas honestas de 25, 10 e 5 centavos, e tetráedros equilibrados de cores distintas - pega tanto quanto for necessário; use todo este material para definir  $M_3$ .
- (iv) Aquí a tarefa é igual a do item anterior, só que agora eu obrigo você usar a moeda de R\$1 (aquela lançada no começo) para construir uma das bifurcações que leva de  $M_2$  a  $M_3$ . Mostre que assim você não vai conseguir atender o requisito (2) de martingalidade. (Que seja claro: você não pode lançar a moeda de R\$1 novamente; você tem que usar o resultado de seu lançamento que determinou o valor de  $X$ .)
- (v) Eu insisto: use a moeda de R\$1, mas de maneira mais sofisticada: multiplique valores de  $M_3$  por 2 caso a moeda deu cara e por 3, caso deu coroa. Mas multiplique valores só de algumas bifurcações de  $M_3$ . Você deve conseguir fazer com que seu  $M_0, M_1, M_2, M_3$  seja martingale.

10. Suponha agora que eu quero contruir um martingale  $M_0, M_1, \dots, M_n$  e desejo que ele possua estrutura complicada de dependência entre seus valores. Um dos caminhos da construção é definir uma sequencia  $X_1, \dots, X_n$  e usá-la como os incrementos do martingal, isto é, usar  $M_i + X_{i+1}$  para definir  $M_{i+1}$ . Imagine que seguimos este caminho. Neste caso, o desejo de ter dependência complexa entre  $M_0, M_1, \dots, M_n$  implica no que os incrementos devem ser dependentes. Quanto a isto, o Ex. 10 mostra que montar uma dependência complexa entre incrementos não é uma tarefa simples. O que é simples, conforme o exercício mostrou, é fazer com que a distribuição do incremento  $X_{i+1}$  dependa dos valores de  $M_0, \dots, M_i$ : para tal, você simplesmente pega uma família de distribuições, todas de média zero, e escolhe uma delas para cada realização de  $M_0, \dots, M_i$ . Mas se você quiser construir uma dependência mais complicada, ai não tudo é fácil e possível; isto foi revelado no item (v) do exercício.

Então, qual é a estrutura de dependência entre  $M$ 's de um martingal  $\{M_i\}$ ? Qual é a sua complexidade? Parece que tais perguntas não permitem uma resposta clara e completa. Assim sentimos apôs ter lido o argumento do paraágrafo acima. Pois bem, a realidade nos supreende neste quesito: a resposta existe. Ela será clara para você depois de ter feito os dois próximos exercícios. Só que tome cuidado: a resposta vale só para martngias de tempo finito, isto é, para o caso quando  $M_0, \dots, M_n$  é uma sequencia finita.

Exercício 11. Seja  $\xi$  uma variável aleatória qualquer mas com média nula. No espaço  $\Omega$ , onde está variável "vive", escolheremos uma sequencia de  $\sigma$ -álgebras encaixadas  $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n$  (o encaixe denota-se aqui por  $\subseteq$  e significa que cada próxima é mais refinada que sua anterior), sendo que  $\mathcal{F}_0$

é trivial (isto é:  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ ). Definiremos  $M_i := E[\xi | \mathcal{F}_i]$ . Prove que, assim construído, o processo  $M_0, M_1, \dots, M_n$  é um martingale.



Exercício 12. Seja  $M_0, \dots, M_n$  um martingale. Mostre que  $M_i = E[M_n | \mathcal{F}_i]$ , para todo  $i = 0, 1, \dots, n$ .

**Comentário final, meio filosófico, sobre a não correlação linear dos incrementos de um martingal:** Quando deseja-se construir processos estocásticos, a primeira e mais natural idéia é empregar a independência. Isto dá em processos com incrementos independentes. Os mais famosos destas são o Passeio Aleatório, o Movimento Browniano e os Processos de Lévy. Após obter avanços significativos no estudo dos processos com incrementos independentes, começa-se pensar em relaxar a independência e ver até onde poderemos ir no sentido de obtenção de resultados genéricos. O problema que a independência é a relação única, enquanto que a dependência – não; existem infinitas formas de dependência. Então é a questão de gosto pessoal a definir a dependência que seja a mais próxima da independência. Para alguns, esta é quando o passado e o futuro são independentes, condicionados por conhecimento do presente. Quem acredita nisso, toma os processos de Markov (para os incrementos) como a classe de processos mais próxima à classe dos processos com incrementos independentes. Para outros, a dependência mais próxima da independência é aquela que exige que a correlação linear seja zero. Para os crentes nesta dogma, os martingais formam a classe de processos mais próxima à classe dos processos com incrementos independentes.

$\{X_n\}_{n \geq 0}$  uma seq. de r.a. iid. com

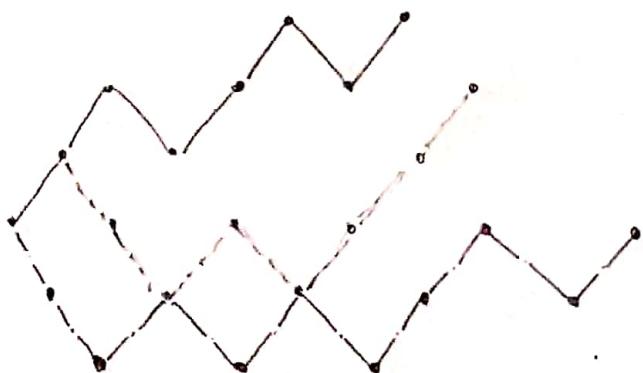
$$X_1 = \begin{cases} +1 & \text{com prob. } p \\ -1 & \text{c.c.} \end{cases}$$

Definir  $S_0 = 0$  e

$$S_n = \sum_{k=1}^n X_k \dots$$

Note que

$$P(S_n = l) = \binom{n}{n+l/2} p^{\frac{n+l}{2}} \cdot (1-p)^{\frac{n-l}{2}}, \quad l \in \{-n, -n+2, \dots, n-2, n\}$$



Traj. de um processo aleatório

$$\begin{aligned} E[M_1 M_{21}] &= E[X_1 (X_1 + X_2)] = E(X_1^2 + X_1 X_2) \\ &\stackrel{p=1/2}{=} 1 + E(X_1 X_2) \stackrel{\text{ind}}{=} 1 + EX_1 EX_2 = 1 + 0 \cdot 0 = 1 \end{aligned}$$

$$E[M_1 M_3] = E[X_1 (X_1 + X_2 + X_3)] = 1 + 0 \cdot 0 \cdot 0 = 1$$

## Probabilidade Avançada II - Tarefas

Capítulo 3, Steele.

### • O Movimento Browniano

O M.B. é o processo estocástico mais importante. Como uma ferramenta prática, este teve profundo impacto em quase todo ramo da ciência física, bem como das ciências sociais. Como uma criação da mat. pura, é uma entidade de beleza incomum. Reflete uma perfeição que parece mais próximo à lei da natureza do que à invenção humana.

A primeira importante aplicação do M.B. foi feita por L. Bachelier e A. Einstein. Bachelier é o reconhecido pai dos métodos quant. em fin. e Einstein é... bem... Einstein. O obj. de Bachelier era fornecer um modelo para mercado financeiro, e Einstein queria modelar o comportamento de uma partícula suspensa em um líquido.

Bachelier : def. uma base sólida para precificação de opções;

Einstein : fornecer um meio de medir o nº de Avogadro, nº de mol. em um ml d'água

Wiener direcionou sua atenção a mat. de mov. Browniano. Dentre as muitas cont., está a primeira prova de que M.B. existe como um objeto matemático definido rigorosamente. As cont. foram tantas, que por vezes o M.B. é chamado de Processo de Wiener.

Definição 3.1. Um proc. estocástico cont. no tempo  $\{B_t : 0 \leq t \leq T\}$  é dito um **Movimento Browniano Padrão em  $[0, T]$**  se satisfizer as seguintes quatro propriedades:

(i)  $B_0 = 0$ .

(ii) Os incrementos de  $B_t$  são independentes; i.e., para qualquer conjunto finito de tempos  $0 \leq t_1 < t_2 < \dots < t_n \leq T$  as variáveis aleatórias

$$B_{t_0} - B_{t_1}, B_{t_1} - B_{t_2}, \dots, B_{t_n} - B_{t_{n-1}}$$

são independentes;

- (iii) Para qualquer  $0 \leq s < t \leq T$ , os incrementos  $B_t - B_s$  tem distribuição normal com média 0 e variância  $t-s$ .
- (iv) Para todo  $w$  em um cto de prob. 1,  $B_t(w)$  é uma f. contínua de t.

O princ. objetivo desse capítulo é mostrar que o MB pode ser repres. como uma soma aleatória de integrais de funções ortogonais. Esta rep. satisfaz a nec. do teórico p/ provar a ex. de um proc com as quatro prop. def. de mov. Browniano, mas tbm serve p/ dem. mais concretas. Em particular, a represent. por série pode ser usada p/ derivar a maior parte das prop. import. de MB. Esta rep. fornece um pod. método numérico para gerar traj. do MB. que são exigidas em sim. de computadores.

### 3.1 Introdução às Covariâncias e Funções Características

A fim de confirmar que ~~o~~ cons. do MB é honesta, precisaremos de alguns fatos básicos sobre a normal multivariada. Em part., vamos precisar de ferramentas que nos ajudem a checar que os inc. de nosso processo são normais e ~~que~~ indep.

#### Normal Multivariada

V um vetor aleatório d-dimensional. Def. o vetor média de

$$V = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_d \end{bmatrix} \quad \text{como} \quad \mu = E[V] = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_d \end{bmatrix} \quad \text{e def. a matriz de cov. de } V$$

com

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1d} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \dots & \sigma_{dd} \end{bmatrix}, \text{ onde } \sigma_{ij} = E[(V_i - \mu_i)(V_j - \mu_j)],$$

Enfatizou o fato mais imp. sobre a Normal Multiv. seja de que sua f.d. é dada pelo vet de médias e pela sua matr. de covariância.

Def. 3.2. Um vetor aleatório d-dimensional é dito ter uma dist. normal multivariada com média  $\mu$  e covariância  $\Sigma$  se a densidade de  $V$  é dada por

$$(4\pi)^{-d/2} (|\Sigma|)^{-1/2} \exp\left(-\frac{1}{2} (x-\mu)^T \Sigma^{-1} (x-\mu)\right) \quad \forall x \in \mathbb{R}^d$$

**Gaussian Miracle:** As coordenadas  $\{V_i : 1 \leq i \leq d\}$  de um vetor com dist. normal d-varioda são ind. se, e somente se,  $\Sigma$  é uma matriz diagonal.

### Funções Características

Se  $X$  uma v.a., a f. caract. de  $X$  é definida como

$$\begin{aligned} \phi : \mathbb{R} &\rightarrow \mathbb{C} \\ t &\rightarrow \phi(t) = E(e^{itX}) \end{aligned}$$

De modo similar, se  $V$  é um vetor d-dimensional de v.a. e  $\theta = (\theta_1, \dots, \theta_d)$  um vetor real, então a f. característica multivariada é def. por

$$\phi(\theta) = E[e^{i\theta^T V}]$$

Há fórmulas de inv. pl. ambo os casos que permitem det. ou f.d. subjacentes a partir da função característica, mas pl. nossas prop. é suf. notar que a f.c. é det. unicamente pela f.d. da v.a. em estudo.

Se  $X \sim N(\mu, \sigma^2)$

$$E(e^{itx}) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-(x-\mu)^2/2\sigma^2} dx = e^{it\mu} e^{-t^2\sigma^2/2}$$

Se  $V \sim N_d(\mu, \Sigma)$ ,

$$E(e^{i\theta^T V}) = e^{i\theta^T \mu - \frac{1}{2}\theta^T \Sigma \theta}$$

**Proposição 3.1 (Caracterização Gaussiana)** O vetor aleatório d-dimensional  $V$  tem dist. normal d-var. se, e só se, toda comb. linear  $\theta_1 V_1 + \dots + \theta_d V_d$  é normal univariada.

Demonstração: pg 21. Steele.

### Funções de Covariância e o Processo Gaussiano

Se um proc. estocástico  $\{X_t, 0 \leq t \leq \omega\}$  tem a prop. de que o vet.  $(X_{t_1}, \dots, X_{t_n})$  tem dist. normal n-var. pl. qqr seq. finita  $0 \leq t_1 < t_2 < \dots < t_n$ , então

$\{X_t\}$  é dito Processo Gaussiano

O Proc. Gaussiano é mais fácil pl. cálculos exatos em relação à quase qualq. outro proc. pois sua f.d. conjuntas de um Proc. Gauss.. são completamente determinadas pela função média

$$\mu(t) = E[X_t]$$

e pela função de covariância

$$f(s,t) = \text{Cov}(X_s, X_t).$$

Poderemos cal. facilmente a função de cov. p/ o mov. Browniano apenas notando que p/ qqr s < t, temos

$$\text{Cov}(B_s, B_t) = E[(B_t - B_s + B_s) B_s] = E[B_s^2] = s, \text{ de modo}$$

que em geral

$$\text{Cov}(B_s, B_t) = \min(s, t), \quad 0 \leq s, t \leq \infty.$$

Qqr Proc Gaussiano com f. cov. dada por é dever ter inc. independentes

Lema 3.1. Se um proc. Gaussiano  $\{X_t : 0 \leq t \leq T\}$  tem  $E(X_t) = 0 \quad \forall 0 \leq t \leq T$  e se

$$\text{Cov}(X_s, X_t) = \min(s, t) \quad \forall 0 \leq s, t \leq T,$$

então o proc.  $\{X_t\}$  tem incr. indep. Ademais, se o proc. tem traj. cont. e  $X_0 = 0$ , então este é um M.B. padrão em  $[0, T]$ .

Dem: pg. 83 Steele.

### 3.2 Visão de umas Aprox. Por Série

### 3.2 Visão de Uma Aprox. por Série

Vamos est. a existência do m.b. forn. visto explicitas expansões em série. Os calc. são quase básicos, mas ainda é preciso desvendar alguns fatos sobre espaço de funções. Primeiro, relembrarmos que pl qqr pr.  $L^p[0,1]$  denota o conjunto de funções  $f: [0,1] \rightarrow \mathbb{R}$  tais que

$$\int_0^1 |f(x)|^p dx < \infty$$

e  $L^p[0,1]$  é visto como um esp. linear normado com norma

$$\|f\|_p = \left( \int_0^1 |f(x)|^p dx \right)^{1/p}.$$

No caso especial em que  $p=1/2$ , também podemos ver  $L^2[0,1]$  como um produto espaço com prod. interno com prod. def. por

$$\langle f, g \rangle = \int_0^1 f(x)g(x) dx < \infty.$$

Finalmente, se as funções no conjunto  $\{\phi_n \in L^2[0,1]: 0 \leq n \leq \infty\}$  satisfazem  $\langle \phi_n, \phi_n \rangle = 1 \forall n \geq 0$  e  $\langle \phi_n, \phi_m \rangle = 0 \forall n \neq m$ , então  $\{\phi_n\}$  é chamada **sequência ortonormal**. Se os comb. lineares finitos de  $\{\phi_n\}$  também formarem um conjunto denso em  $L^2[0,1]$ , então  $\{\phi_n\}$  é uma **sequência ortonormal completa**.

Há 2 fatos que dão um seq. ord. comp. importantes aqui. O 1º é que  $\forall f \in L^2[0,1]$ ,

$$f = \sum_{n=0}^{\infty} \langle f, \phi_n \rangle \phi_n$$

onde o significado preciso da conv. da soma é que

$$\left\| f - \sum_{n=0}^N \langle f, \phi_n \rangle \phi_n \right\|_2 \rightarrow 0 \text{ qdo } N \rightarrow \infty.$$

A seq. é que podemos comp. o prod. interno de dois modos distintos por meio da **Identidade de Parseval**:

$$\int_0^1 f(x) g(x) dx = \sum_{n=0}^{\infty} \langle f, \phi_n \rangle \langle g, \phi_n \rangle.$$

Se vislumbrarmos no horizonte, podemos ver que Identidade de Parseval pode ser útil no est. do M.B. A conexão vem esp.  $f$  e  $g$  funç. indicadoras nos int.  $[0,s]$  e  $[0,t]$ , i.e.,

$$f(x) \cdot \mathbb{1}_{[0,s]}(x) = g(x) \cdot \mathbb{1}_{[0,t]}(x).$$

Neste caso, a id. de Parseval simplifica um pouco e basta dizer que

$$\min(s, t) = \sum_{n=0}^{\infty} \int_0^s \phi_n(x) dx \int_0^t \phi_n(x) dx \quad (3.8)$$

$\int_0^s \phi_n(x) dx$

Apesar de a prim. vista longa, a fórm. acima é mto útil.

### A Geração $\text{Min}(s, t)$

O Benefício da exp de Parseval (3.8) adverem do fato de que o M.B. tem cov.  $\text{Cov}(B_s, B_t) = \min(s, t)$ , de modo que a identidade (3.8) eria uma conexão entre o M.B. e integrais de uma seq. completa orthonormal. Há vários caminhos para explorar tal vínculo, mas talvez a ideia mais natural é consid. o prod. est. definido por

$$X_t = \sum_{n=0}^{\infty} Z_n \int_0^t \phi_n(x) dx$$

onde  $\{Z_n : 0 \leq n < \infty\}$  é uma seq. de v.a. l.i.d. com a m.m.d. de  $Z \sim N(0,1)$ .

### 3.3 Two Wavelets

Visto em aula: introdução  $\lambda_n$  e  $\Delta_n$ .

### 3.4 Wavelet Prepr. do M.B.

**Teorema 3.1.** Se  $\{Z_n : 0 \leq n < \infty\}$  é uma seq. de normais ind. com média 0 e variâncas  $1$ , então a série def. por

$$X_t = \sum_{n=0}^{\infty} \lambda_n Z_n \Delta_n(t)$$

converge uniformemente em  $[0,1]$  com prob. 1. Ademais, o proc.  $\{X_t\}$  definido pelo limite é um M.B. padrão para  $0 \leq t \leq 1$ .

Introduz alguns resultados sobre via. norm. ind. e depois vai à prova.

Lema da aula:  $|Z_n| \leq \sqrt{\log n} \quad \forall n \geq 2 \quad \Rightarrow \mathbb{P}(|Z_n| \geq 1) = 1$

Mais, o que é Proc. Gaussiano. Depois, que é M.B.

### 3.5. Escalonando e Invertendo M.B.

A partir do B.M. em  $[0,1]$ , M.B. em  $[0,\infty)$ : vários modos.

Juntar no ~~desmembrável~~ de M.B. com o sucessor começando onde o antec. acaba



Para cada  $1 \leq n < \infty$ , formamos um ind. M.B. padrão  $B_t^{(n)}$  em  $[0,1]$ , e para qqr  $t < \infty$  def.  $B_t$  pela soma

$$B_t = \sum_{n=1}^{\infty} B_t^{(n)} + B_{t-n}^{(n+1)}, \text{ sempre que } t \in [n, n+1],$$

Assim, por ex.

$$B_{0,5} = B_i^{(1)} + B_i^{(2)} + B_i^{(3)} + B_{0,5}^{(4)}$$

• Mais natural  $[0, \infty)$  como dom. para M.B.

•

**Proposição 3.2. (Existe e Inversão)** Para  $a > 0$ , o proc. escalado def. por

$$X_t = \frac{1}{\sqrt{a}} B_{at} \quad \text{para } t \geq 0$$

e proc. invertido def por

$$Y_0 = 0 \quad Y_t = t B_{1/t} \quad \text{para } t \geq 0$$

são ambos M.B. podendo em  $[0, \infty)$ .

**SOLUTION FOR PROBLEM 2.10.** To calculate  $E(M_{n+1}|\mathcal{F}_n)$ , first note that  $M_{n+1}$  is equal to  $(A_n + 1)/(A_n + B_n + 1)$  with probability  $M_n = A_n/(A_n + B_n)$  and  $M_{n+1}$  equals  $A_n/(A_n + B_n + 1)$  with probability  $1 - M_n$ . Thus,

$$\begin{aligned} E(M_{n+1}|\mathcal{F}_n) &= \frac{A_n + 1}{(A_n + B_n + 1)} \frac{A_n}{A_n + B_n} + \frac{A_n}{(A_n + B_n + 1)} \frac{B_n}{A_n + B_n} \\ &= \frac{1}{(A_n + B_n + 1)} \left\{ \frac{A_n(A_n + B_n + 1)}{A_n + B_n} \right\} \\ &= \frac{A_n}{A_n + B_n} = M_n. \end{aligned}$$

Since  $\{M_n\}$  is bounded between 0 and 1, the martingale convergence theorem tells us that with probability one  $\{M_n\}$  converges to some random variable  $Y$ . Incidentally, one can further show that  $Y$  has the uniform distribution on  $[0, 1]$ , and this model is equivalent to the simplest case of the famous Pólya urn model.

**SOLUTION FOR PROBLEM 2.11.** We have  $E(|M_n|) \leq E(|M_n|^p)^{1/p} \leq B^{1/p} < \infty$  by Jensen's inequality, so  $\{M_n : n \geq 0\}$  is also an  $L^1$ -bounded martingale. The  $L^1$  convergence theorem tells us that there is an  $M_\infty \in L^1$  such that  $M_n$  converges almost surely to  $M_\infty$ , confirming the first assertion of (2.36).

To prove the second assertion, we first note by Fatou's Lemma that

$$E(|M_\infty|^p) = E(\liminf_{n \rightarrow \infty} |M_n|^p) \leq \liminf_{n \rightarrow \infty} E(|M_n|^p) \leq B < \infty,$$

so  $M_\infty \in L^p$ . If we set  $D = \sup_n |M_n|$ , then  $D \in L^p$  by Doob's maximal inequality (2.23), so by the elementary Jensen bound (2.32), we have

$$|M_n - M_\infty|^p \leq 2^{p-1} \{ |M_n|^p + |M_\infty|^p \} \leq 2^{p-1} \{ D^p + |M_\infty|^p \}.$$

Since  $D^p + |M_\infty|^p \in L^1$ , we can therefore apply the dominated convergence theorem to the sequence  $|M_n - M_\infty|^p$  to obtain

$$\lim_{n \rightarrow \infty} E(|M_n - M_\infty|^p) = E(\lim_{n \rightarrow \infty} |M_n - M_\infty|^p) = 0;$$

In other words,  $\lim_{n \rightarrow \infty} \|M_n - M_\infty\|_p = 0$ , just as we hoped.

### Chapter 3

**SOLUTION FOR PROBLEM 3.1.** For part (a), complete the square in the exponent to get

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2/2} dx = e^{t^2/2} \left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-t)^2/2} dx \right\}$$

and observe that the braced integral is equal to one since  $e^{-(x-t)^2/2}/\sqrt{2\pi}$  is the density of a Gaussian random variable with mean  $\mu = t$  and variance 1.

## Capítulo 5 . Riqueza de Caminhos

Pode-se passar uma vida toda explorando as delicadas - e fascinantes - propriedades das trajetórias do movimento Browniano. A maioria não pode arcar com tal investimento - escolhas dev ser feitas. Ainda assim, sem quer dúvida, há duas questões fundamentais que devem ser consideradas

- Quão suave - ou rugosa - é uma traj. Browniana?
- Qual a relação entre traj. do M.B. e do P.A.S.?

Vamos ver que a wavelet construction do M.B. fornece uma ferramenta bem afiada para responder sobre a suavidade na primeira parte. A resposta para a rugosidade requer um desenvolvimento independente, mas os resultados anteriormente mencionados nos auxiliam. Vamos ver que com prob. 1, as traj. do M.B. não são diferenciáveis em qqer ponto. Esta é a realidade geométrica que força o cálculo estocástico a divergir do cálc. elem. de f. suaves.

Nossa exploração da seg. ~~pergunta~~ questão tem duas partes. Na prim., examinaremos alguns dos resultados em que o PAS pode nos inf. sobre o M.B. Um dos links mais persistentes é simplesmente analogia, e exploraremos este link juntamente com desenv. paralelo para o princípio da reflexão para PAS e M.B. A partir daí, daremos uma breve introd. às pontas mais formais entre PAS e M.B, esp. o Princípio da Invariância e o Teorema de Donsker.

A última parte fornece a conexão mais decisiva. Veremos que todo 'unbiased' processo aleatório com passos com variância finita podem ser imerso em um M.B. de uma forma que é tão explícita que mtoas perg. pl o PAS se tornam corolários das res. corresp. pl M.B.

### 6.6. Approximation in $\mathcal{H}^2$

Our definition of the Itô integral on  $\mathcal{H}^2$  used the fact that  $\mathcal{H}_0^2$  is a dense subset of  $\mathcal{H}^2$ , and the main goal of this section is to prove an explicit approximation theorem that provides the desired density result as an immediate corollary. The proof of this approximation theorem is interesting for several reasons, not the least of which is that it illustrates the important fact that one can sometimes approximate a "nonmartingale" with a martingale. Still, this proof does require a detour from our main path, and the reader who is pressed for time should know that the techniques of this section will not be used in subsequent chapters.

#### APPROXIMATION OPERATOR

Given any integer  $n \geq 1$ , the approximation operator  $A_n$  is defined to be the mapping from  $\mathcal{H}^2$  into  $\mathcal{H}_0^2$  that is given by taking

$$(6.30) \quad A_n(f) = \sum_{i=1}^{2^n-1} \left\{ \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} f(\omega, u) du \right\} 1(t_i < t \leq t_{i+1}),$$

where  $t_i = iT/2^n$  for  $0 \leq i \leq 2^n$ . Our interest in this operator is explained by the following theorem and its immediate corollary —  $\mathcal{H}_0^2$  is a dense subset of  $\mathcal{H}^2$ .

**THEOREM 6.5 (Approximation Theorem).** *The approximation operator  $A_n$  defines a bounded linear mapping from  $\mathcal{H}^2$  into  $\mathcal{H}_0^2$  such that*

$$(6.31) \quad \|A_n(f)\|_\infty \leq \|f\|_\infty,$$

$$(6.32) \quad \|A_n(f)\|_{L^2(dP \times dt)} \leq \|f\|_{L^2(dP \times dt)},$$

and

$$(6.33) \quad \lim_{n \rightarrow \infty} \|A_n(f) - f\|_{L^2(dP \times dt)} = 0 \quad \text{for all } f \in \mathcal{H}^2.$$

#### FIRST CHECKS

The operator  $A_n$  is clearly linear. To verify that  $A_n(f) \in \mathcal{H}_0^2$ , we first consider measurability. First, note that for all  $0 < i < 2^n$ , the coefficients

$$(6.34) \quad a_i(\omega) = \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} f(\omega, u) du$$

are in  $\mathcal{F}_{t_i}$ . For  $t \in [0, T2^{-n}]$  we have  $A_n(f)(t) = 0$ , while for  $1 \leq i < 2^n$  and  $t_i < t \leq t_{i+1}$  we have  $A_n(f)(t) = a_i(\omega) \in \mathcal{F}_{t_i} \subset \mathcal{F}_t$ . Therefore, for all  $0 \leq t \leq T$ , we have  $A_n(f)(t) \in \mathcal{F}_t$ . Next, to show that  $A_n(f) \in \mathcal{H}_0^2$ , we only need to prove that  $E[(A_n(f))^2] < \infty$ , and this follows immediately if we show  $E(a_i^2) < \infty$ . By Jensen's inequality, we have

$$(6.35) \quad a_i^2(\omega) \leq \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} f(\omega, u)^2 du,$$

and since the right-hand side is bounded by  $(t_i - t_{i-1})^{-1} \|f\|_{L^2(dP \times dt)}^2 < \infty$ , we find  $E(a_i^2) < \infty$ . This bound tells us that  $A_n(f) \in \mathcal{H}_0^2$ , exactly as required.

The proof of the  $L^\infty$  bound (6.31) is trivial, and the proof of the  $L^2$  bound (6.32) is not much harder. When we square  $A_n(f)$  and integrate over  $[0, T]$ , we find

that inequality (6.35) gives us

$$A_n^2(f) = \sum_{i=1}^{2^n-1} 1(t_i < t \leq t_{i+1}) a_i^2(\omega),$$

so, when we take expectations and apply the bound (6.35), we find

$$E\left[\int_0^T A_n^2(f) dt\right] = E\left[\sum_{i=1}^{2^n-1} (t_{i+1} - t_i) a_i^2(\omega)\right] \leq E\left[\int_0^T f(\omega, t)^2 dt\right],$$

precisely the required  $L^2$  bound (6.32).

### A MARTINGALE COUSIN

We now come to the crux of Theorem 6.5, the proof of the limit (6.33). Here, it is fruitful to introduce a new operator that looks a lot like  $A_n$  but which has an interpretation as a martingale. This new operator  $B_n$  is defined for  $f \in \mathcal{H}^2$  by taking  $t_i = iT/2^n$  for  $0 \leq i \leq 2^n$  and setting

$$(6.36) \quad B_n(f) = \sum_{i=1}^{2^n} \left\{ \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} f(\omega, u) du \right\} 1(t_{i-1} < t \leq t_i).$$

The operator  $B_n$  is closely related to  $A_n$ , although one should note that  $B_n$  no longer maps  $\mathcal{H}^2$  into  $\mathcal{H}_0^2$ ; the coefficients have been shifted in such a way that the coefficient of the indicator of the interval  $(t_{i-1}, t_i]$  is no longer measurable with respect to  $\mathcal{F}_{t_{i-1}}$ . This means that  $B_n(f)$  no longer has the measurability that is required of a function in  $\mathcal{H}_0^2$ . In compensation, we have for all  $f \in L^2(dP \times dt)$  and all  $\omega \in \Omega$  that the process defined by

$$\{M_n(\cdot) = B_n(f)(\omega, \cdot) : n \geq 0\}$$

is a martingale, or at least it is a martingale once we spell out the filtration and the probability space where the martingale property holds.

In fact, our new martingale is conceptually identical to the one introduced in Exercise 4.7, but in this case we have some extra baggage since  $B_n(\omega, t)$  is a function of two variables. Our task is thus to show that when we fix  $\omega$  we obtain a martingale in the  $t$  variable that exactly replicates the martingale of Exercise 4.7. This is a little tedious to write out, but it is easy to think through.

To punch out the formalities, we first fix  $\omega \in \Omega$  and then define a new probability space  $(\Omega', \mathcal{F}', Q)$  by taking the base space  $\Omega' = \omega \times [0, T]$  and a new  $\sigma$ -field  $\mathcal{F}' = \{(\omega, A) : A \in \mathcal{B}\}$ . We then define a probability measure  $Q$  on  $\mathcal{F}'$  by taking  $Q(\omega, A) = P(\Omega \times A)/T$  for all  $A \in \mathcal{B}$ . Finally, the filtration for our martingale is taken to be  $\{\mathcal{G}_n\}$ , where  $\mathcal{G}_n$  is the smallest  $\sigma$ -field contained in  $\mathcal{F}'$  such that all of the functions of the form

$$(6.37) \quad \sum_{i=1}^{2^n} c_i 1(t_{i-1} < t \leq t_i) \text{ where } c_i \in \mathbb{R}$$

are  $\mathcal{G}_n$ -measurable.

The benefit of this choice of  $\mathcal{G}_n$  is that it gives us a representation of  $B_n(f)(\omega, \cdot)$  as a conditional expectation

$$(6.38) \quad E_Q(f(\omega, t) | \mathcal{G}_n) = B_n(f)(\omega, t),$$

so the martingale property for  $M_n$  comes to us as a consequence of the tower property of conditional expectations. We are now ready to prove a lemma that tells us  $B_n(f)$  is a good approximation to  $f$  provide that  $f \in \mathcal{H}^2 \subset L^2(dP \times dt)$  is bounded.

**LEMMA 6.3.** *If  $f \in \mathcal{H}^2$  and  $|f(\omega, t)| \leq B$  for all  $(\omega, t)$ , then*

$$(6.39) \quad \|B_n(f) - f\|_{L^2(dP \times dt)} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

**PROOF.** If we fix  $\omega$ , our interpretation of  $B_n(f)(\omega, \cdot)$  as a martingale lets us apply the  $L^2$ -bounded martingale convergence theorem to conclude that  $B_n(f)(\omega, t)$  converges for all  $t \in [0, T]$  except possibly a set of measure zero. We denote this limit by  $B_\infty(f)(\omega, t)$  and note by the bound on  $f$  this limit is bounded by  $B$ . We can then apply the dominated convergence theorem to find

$$(6.40) \quad \lim_{n \rightarrow \infty} \int_A B_n(f)(\omega, t) dt = \int_A B_\infty(f)(\omega, t) dt \text{ for all } A \in \mathcal{B}.$$

We also know directly from the definition of the sequence  $\{B_n\}$  that

$$(6.41) \quad \int_A B_m(f)(\omega, t) dt = \int_A f(\omega, t) dt$$

for all  $m \geq n$  and all  $A$  such that  $(\omega, A) \in \mathcal{G}_n$ . From equations (6.40) and (6.41), we can conclude (e.g., by Exercise 6.5) that for all  $\omega \in \Omega$  we have

$$(6.42) \quad B_\infty(f)(\omega, t) = f(\omega, t)$$

for all  $t \in [0, T]$  except a set of measure zero. By a second application of the DCT, we therefore have for all  $\omega \in \Omega$  that

$$(6.43) \quad \lim_{n \rightarrow \infty} \int_0^T |B_n(f)(\omega, t) - f(\omega, t)|^2 dt = 0.$$

Finally, we can take expectations in equation (6.43) and apply the DCT one last time to complete the proof of the limit (6.39) and the lemma.  $\square$

### THE MARTINGALE MEETS THE APPROXIMATION

The key connection between  $A_n$  and  $B_n$  is given by the following lemma.

**LEMMA 6.4.** *For any  $f \in \mathcal{H}^2$  and any fixed integer  $m$ ,*

$$\lim_{n \rightarrow \infty} \|(A_n(B_m(f)) - B_m(f))\|_{L^2(dP \times dt)} = 0.$$

**PROOF.** Two basic observations yield the lemma. First, we note that by direct consideration of the definition of  $B_m$  we have that  $B_m(f)(\omega, t)$  is constant on each of the intervals  $(i2^{-m}T, (i+1)2^{-m}T]$  for  $i = 0, 1, \dots, 2^m - 1$ , so from the definition of  $A_n$  we find for  $n \geq m$  and all  $\omega$  that

$$A_n(B_m(f))(\omega, t) = B_m(f)(\omega, t) \text{ for all } t \notin \cup_{1 \leq i < 2^m} [i2^{-m}T, i2^{-m}T + 2^{-n}T].$$

In view of this identity, we have

$$\lim_{n \rightarrow \infty} A_n(B_m(f))(\omega, t) = B_m(f)(\omega, t)$$

for all  $(\omega, t)$  such that  $t \neq i2^{-m}T$ .

Second, we note that the converging terms are dominated. In particular, for all  $n \geq m$ , we have

$$|A_n(B_m(f))(\omega, t)| \leq |B_m(f)(\omega, t)| + |B_m(f)(\omega, t - 2^{-n}T)|,$$

and  $|B_m(f)(\omega, t)| \in L^2(dP \times dt)$ , so for all  $n \geq 1$  the difference  $|A_n(B_m(f))(\omega, t) - B_m(f)(\omega, t)|^2$  is bounded by a fixed integrable function. The DCT then completes the proof of the lemma.

### PROOF OF THE APPROXIMATION THEOREM: LAST STEP

We only need prove the limit result (6.33). First, we consider a bounded  $f \in \mathcal{H}^2$ . By the triangle inequality and the  $L^2$  bound (6.32), we have

$$\begin{aligned}\|A_n(f) - f\|_{L^2(dP \times dt)} &\leq \|A_n(f - B_m(f))\|_{L^2(dP \times dt)} + \|A_n(B_m(f)) - f\|_{L^2(dP \times dt)} \\ &\leq \|(f - B_m(f))\|_{L^2(dP \times dt)} + \|A_n(B_m(f)) - f\|_{L^2(dP \times dt)},\end{aligned}$$

and by Lemma 6.4 we have  $\|A_n(B_m(f)) - B_m(f)\|_{L^2(dP \times dt)} \rightarrow 0$ , so for all  $m$  we have

$$\limsup_{n \rightarrow \infty} \|A_n(f) - f\|_{L^2(dP \times dt)} \leq 2\|(f - B_m(f))\|_{L^2(dP \times dt)}.$$

Since  $m$  is arbitrary, this inequality and Lemma 6.3 tell us that

$$\lim_{n \rightarrow \infty} \|A_n(f) - f\|_{L^2(dP \times dt)} = 0$$

for all bounded  $f \in \mathcal{H}^2$ .

Now, at last, we consider a possibly unbounded  $f \in \mathcal{H}^2$ . For any  $\epsilon > 0$ , we have a bounded  $f_0 \in \mathcal{H}^2$  such that  $\|f - f_0\|_{L^2(dP \times dt)} \leq \epsilon$ , and now we can check that the contraction property of  $A_n$  will complete the proof. Specifically, by the triangle inequality and the  $L^2$  bound (6.32), we have

$$\begin{aligned}\|A_n f - f\|_{L^2(dP \times dt)} &= \|A_n f_0 - f_0 + A(f - f_0) - (f - f_0)\|_{L^2(dP \times dt)} \\ &\leq \|A_n f_0 - f_0\|_{L^2(dP \times dt)} + \|A(f - f_0)\|_{L^2(dP \times dt)} + \|(f - f_0)\|_{L^2(dP \times dt)} \\ &\leq \|A_n f_0 - f_0\|_{L^2(dP \times dt)} + 2\|(f - f_0)\|_{L^2(dP \times dt)}.\end{aligned}$$

If we now let  $n \rightarrow \infty$ , the boundedness of  $f_0$  and the first half of our argument tell us  $\|A_n f_0 - f_0\|_{L^2(dP \times dt)}$  goes to zero. Finally, the arbitrariness of  $\epsilon$  completes the proof of the fundamental limit result (6.33) and the proof of Theorem 6.5.  $\square$

### 6.7. Exercises

The first three exercises offer “must do” practice with the Itô isometry and the distinction between ordinary and stochastic integrals. The last two exercises are more technical. In particular, Exercise 6.5 sketches the proof of an intuitive result that we needed in our proof of the approximation theorem, and Exercise 6.4 suggests an alternative way to find a dense subset of  $\mathcal{H}^2$  in  $\mathcal{H}_0^2$ .

**EXERCISE 6.1.** Use the Itô isometry to calculate the variances of

$$\int_0^t |B_s|^{\frac{1}{2}} dB_s \text{ and } \int_0^t (B_s + s)^2 dB_s.$$

**EXERCISE 6.2.** The integrals

$$I_1 = \int_0^t B_s ds \text{ and } I_2 = \int_0^t B_s^2 ds$$

are *not* stochastic integrals, although they are random variables. For each  $\omega$  the integrands are nice continuous functions of  $s$  and the  $ds$  integration is just the traditional calculus integration. Find the mean and variance of the random variables  $I_1$  and  $I_2$ .

Ex. 1 Ache a distribuição da variável aleatória  $\tau$  definida assim (é um caso particular da definição genérica que você acha na primeira página do Capítulo 1):

$$\tau := \min\{n \geq 0 : S_n = 2 \text{ ou } S_n = -1\}$$

onde  $S_0, S_1, S_2, \dots$  é PASS (Paséio aleatório simples simétrico). Dica: Faça o desenho do PASS para primeiros tempos. Você vai achar um padrão de comportamento que facilita a resolução da tarefa. Calcule também a distribuição de  $S_\tau$ .

Ex. 2 Apresente a conta que deriva a fórmula (1.1) a partir da definição de  $f(k)$  (esta definição está 5 linhas acima de (1.1)).

Ex. 3. Com a Seção 1.2 dedica-se à demonstração da finitude da variável aleatória  $\tau$ , sugere naturalmente a pergunta: qual seria o problema se  $\tau$  assumisse *infty* com uma probabilidade não nula?

Ex. 4 Apresente a conta que deriva a fórmula (1.7) a partir da definição de  $g(k)$  (esta definição está 5 linhas acima de (1.7)).

Ex. 5 Faça o Exercício 1.1 da Seção 1.6. Obs.: Não tentei fazê-lo e não sei o nível de sua dificuldade.

Ex. 6 Faça o Exercício 1.3 da Seção 1.6. Obs.: Não tentei fazê-lo e não sei o nível de sua dificuldade; talvez, sua resolução baseia-se parcialmente, nos resultados do Exercício 1.2 da Seção 1.6. Este foi substituído pela série de exercícios entitulados "Princípio de Reflexão e sua Aplicação na Derivação da Lei de Arco Seno", que está abaixo.

Ex. 7 O texto abaixo é uma série de definições, explicações e exercícios elaborada por mim para que você aprenda algumas propriedades do PASS. Leia o texto e faça os exercícios.

## Princípio de Reflexão e sua Aplicação na Derivação da Lei de Arco Seno

**Definição 1** (de caminho). Considere um passeio aleatório simples em tempo discreto, quer dizer, a sequencia de variáveis aleatórias  $S_0, S_1, \dots$  definidas pela seguinte maneira:

$$\begin{aligned} S_0 &\equiv 0 \\ S_1 &:= S_0 + X_1 \\ S_2 &:= S_1 + X_2 \\ &\dots \\ S_n &:= S_{n-1} + X_n \\ &\dots \end{aligned}$$

onde  $X_1, X_2, \dots$  são variáveis aleatórias i.i.d. que obedecem à condição

$$\mathbb{P}[X_i = 1] = p, \mathbb{P}[X_i = -1] = 1 - p, \text{ para algum } p \in [0, 1].$$

Seja  $s_0, s_1, s_2, \dots$  a notação genérica para realização do processo  $S_0, S_1, S_2, \dots$ . Para qualquer uma de tais realizações, desenhe no plano  $\mathbb{R}^2$  os pontos  $(0, s_0), (1, s_1), (2, s_2), \dots$  e junte por linhas retas os pontos consecutivos. Você então construiu um exemplo daquilo que chamamos por *caminho*. Os caminhos correspondentes a PAS sempre começam no ponto  $(0, 0)$ , mas precisaremos as vezes, relaxar esta limitação. Por isto, formularemos a definição de *caminho* de maneira pouco diferente daquela te guiou agora para construir seu exemplo; nossa definição soa assim: *uma linha poligonal em  $\mathbb{R}^2$  com vértices*

$$(n, s_n), (n+1, s_{n+1}), \dots, (n+k, s_{n+k}), \text{ com } n, s_n \in \mathbb{Z} \text{ e } s_{j+1} - s_j = \pm 1$$

chama-se caminho;  $k$  interpreta-se como seu comprimento.

**Lema 1.** O número de caminhos de comprimento  $n$  que partem de  $(0, 0)$  e chegam a  $(n, y)$  é igual a  $C_n^{(n-y)/2}$  caso  $(n-y)/2$  for inteiro, e 0 caso contrário.

**Demonstração.** Use o seguinte fato: para um caminho chegar ao nível  $y$  no tempo  $n$ , é necessário que

$$\text{o número de passos para cima} - \text{o número de passos para baixo} = y.$$

Mas estes dois números dão  $n$  em soma. Daí o número de passos para cima é  $(n-y)/2$ . Estes passos podem ser distribuídos por  $n$  “tempos” de maneira arbitrária. A quantidade de maneiras diferentes desta distribuição é  $C_n^{(n-y)/2}$ . Verifique os detalhes deste argumento e termine a demonstração.  $\square$

**Lema 2** (princípio de reflexão). Sejam  $A = (x, n)$  e  $B = (y, k)$  dois pontos em  $\mathbb{R}^2$  tais que

$$\begin{aligned} x, y, n, k \in \mathbb{Z} &\quad \text{as coordenadas são inteiros} \\ x, y > 0 &\quad \text{os pontos ficam no semiplano superior} \\ n < k &\quad A \text{ fica à “esquerda” de } B \end{aligned}$$

Considere os caminhos que ligam a a  $B$  e que ou tocam no eixo de abscissas ou cruzam-o. Então, o número desses é igual ao número de todos os caminhos que ligam o ponto  $A' = (-x, n)$  ao ponto  $B$ .

**Demonstração.** Usando a Figura A, é fácil sugerir a bijeção entre o conjunto de caminhos ligando  $A$  e  $B$  e o conjunto de caminhos ligando  $A'$  e  $B$ . Por trás desta bijeção está a reflexão da parte de caminho entre  $A$  e o primeiro ponto do encontro do caminho com eixo de abscissas. Faça esta bijeção com cuidado: verifica que os encontros posteriores do caminho com o eixo – caso existam – não estragam a bijeção.  $\square$

**Teorema 1** (sobre eleições). Fixaremos dois inteiros  $n$  e  $y$ , mas com a condição adicional  $y > 0$ . Consideraremos os caminhos

$$(0, 0), (1, s_1), (2, s_2), \dots, (n, s_n)$$

tais que

$$\begin{aligned} s_n &= y && (\text{os caminhos chegam ao ponto } (n, y)) \\ s_1, s_2, \dots, s_{n-1} & && \text{são todos positivos} \end{aligned}$$

Então, o número de tais caminhos é igual a  $(y/n) C_n^{(n-y)/2}$ .

**Dicas para a demonstração.** Os caminhos, que satisfazem às condições do teorema, são os que ligam  $(1, 1)$  com  $(n, y)$  e não tocam nem cruzam o eixo de abscissas. O número de tais caminhos

pode ser expresso como a diferença entre o número de

(i) todos os caminhos que ligam  $(1, 1)$  e  $(n, y)$

e o número dos

(ii) caminhos que ligam  $(1, 1)$  e  $(n, y)$  e tocam ou cruzam o eixo de abcissas.

Verifique, usando as lemas já formulados, que a quantidade dos caminhos em (i) é  $C_{n-1}^{((n-1)-(y-1))/2}$ , enquanto que a dos (ii) é  $C_{n-1}^{((n-1)-(y+1))/2}$  e faça o cálculo que deduz que a diferença entre essa quantidades dá o resultado formulado no teorema.  $\square$

**Aviso 1.** O Princípio de Reflexão será usado para analizar o passeio aleatório simples simétrico (PASS). Tal análise ficará mais direta, caso considerarmos o passeio aleatório simples de comprimento pár (quer dizer,  $S_0, S_1, \dots, S_k$  com  $k = 2n$ ,  $n \in \mathbb{N}$ ). É isto que faremos daqui para frente.

**Teorema 2.** Para  $k = 1, 2, \dots$ , definimos  $L_{2k} := \frac{1}{k+1} C_{2k}^k$ . Então:

(a) há exatamente  $L_{2n-2}$  caminhos de comprimento  $2n$  que ligam  $(0, 0)$  e  $(2n, 0)$  e são tais que  $s_1 > 0, s_2 > 0, \dots, s_{2n-1} > 0$ ;

(b) há exatamente  $L_{2n}$  caminhos de comprimento  $2n$  que ligam  $(0, 0)$  a  $(2n, 0)$  e são tais que  $s_1 \geq 0, s_2 \geq 0, \dots, s_{2n-1} \geq 0$ .

Dicas para a demonstração. Os caminhos que nos interessam no item (a) são iguais em número aos que ligam  $(0, 0)$  com  $(2n-1, 1)$  e não tocam e não cruzam o eixo de abcissas. Estes últimos foram contados no Teorema 1. Dai chega-se ao resultado.

Tome qualquer caminho do item (b). Levante este por 1 e depois desloque à direita por 1. O resultado é um caminho que liga  $(1, 1)$  com  $(2n+1, 1)$ . Acrecente a este o elo  $(0, 0) \rightarrow (1, 1)$  e o elo  $(2n+1, 1) \rightarrow (2n+2, 0)$ . Você então construiu um caminho de  $(0, 0)$  a  $(2n+2, 0)$  cujo comprimento é  $2n+2$  e tal que  $s_1 > 0, s_2 > 0, \dots, s_{2n+1} > 0$ . Use agora Teorema 2(a) para concluir a demonstração.  $\square$

**Aviso 2.** Até o momento, nosso foco era a contagem de caminhos com certas propriedades. Os resultados obtidos apliam-se então para calcular probabilidades de passeio aleatório simples, que pode ser tanto simétrico ( $p = 1/2$ ) quanto assimétrico ( $p \neq 1/2$ ). A partir de agora, nossas definições e resultados são exclusivamente sobre o passeio aleatório simples simétrico (PASS).

**Definição 2.** Considere o PASS  $S_0, S_1, S_2, \dots$

Para  $k = 1, 2, 3, \dots$ , definimos

$$u_{2k} := \mathbb{P}[S_{2k} = 0]$$

e acrescemos  $u_0 := 1$ . Note a interpretação:  $u_{2k}$  é a probabilidade do PASS voltar ao origem no tempo  $2k$ .

Para  $k = 1, 2, 3, \dots$ , definimos

$$f_{2k} := \mathbb{P}[S_1 \neq 0, S_2 \neq 0, \dots, S_{2k-1} \neq 0, S_{2k} = 0]$$

e acrescemos  $f_0 = 0$ . Note a interpretação:  $f_{2k}$  é a probabilidade que a primeira volta ao origem do PASS ocorra no tempo  $2k$ .

**Theorema 3.**

$$(a) \quad u_{2n} = 2^{-2n} C_{2n}^n, \text{ para } n = 0, 1, 2, \dots$$

$$(b) \quad f_{2n} = \frac{1}{2n} u_{2n-2}, \text{ para } n = 1, 2, 3, \dots$$

**Dicas para a demonstração.** O fator  $2^{-2n}$  apareceu por que esta é a probabilidade de PASS seguir um dos seus possíveis caminhos de comprimento  $2n$ . Para contar os caminhos de comprimento  $2n$  que adequam-se a  $u_{2n}$ , usa-se Lema 1. Isto nos leva ao resultado (a). Para provar (b) use Teorema 1, mas não se esqueça de que há duas maneiras para um caminho chegar ao 0 pelo primeira vez no tempo  $2n$  saindo do 0 no tempo 0: uma é passar por valores positivos, outra – por valores negativos.  $\square$

**Comentário.** Existem relações entre  $u_{2k}$  e  $f_{2k}$  que são úteis, mas não para o objetivo do presente documento. Eis algumas delas:

$$\begin{aligned}f_{2n} &= u_{2n-2} - u_{2n}, \quad n = 1, 2, \dots \\u_{2n} &= \mathbb{P}[S_1 \neq 0, S_2 \neq 0, \dots, S_{2n} \neq 0], \quad n = 1, 2, \dots \\u_{2n} &= \mathbb{P}[S_1 \geq 0, S_2 \geq 0, \dots, S_{2n} \geq 0], \quad n = 1, 2, \dots \\f_{2n} &= \mathbb{P}[S_1 \geq 0, S_2 \geq 0, \dots, S_{2n-2} \geq 0, S_{2n-1} < 0], \quad n = 1, 2, \dots\end{aligned}$$

**Lema 3.** Para qualquer  $n \geq 1$ , ocorre que

$$u_{2n} = \sum_{r=1}^n f_{2r} u_{2n-2r}$$

**Dicas para a demonstração.** Use a interpretação probabilística de  $f$ 's e  $u$ 's.

**Definição 3.** Denotamos por  $p_{2k,2n}$  a probabilidade de que o PASS de comprimento  $2n$  passe  $2k$  unidades de tempo acima do eixo de abcissas (veja Figura B); aqui  $n \geq 1$  e  $k = 0, 1, 2, \dots, n$ .

**Teorema 4.**  $p_{2k,2n} = u_{2k} \cdot u_{2n-2k}$ .

**Dicas para a demonstração.** A demonstração é por indução matemática.

A base da indução (quer dizer, a afirmação do teorema para  $n = 1$ ) justifica-se sem problemas.

A suposição da indução é: a afirmação vale para todo  $m \leq n - 1$ .

Vamos provar que então a afirmação vale para  $n$ . Fixaremos  $k$ , e consideraremos todos os caminhos (de comprimento  $2n$ ) que ficam  $2k$  unidades de tempo no semiplano positivo. Dividimos o conjunto deles em “sacos” rotulados  $2, 4, \dots, 2n$ . Ao saco 2 irão todos os caminhos cujo primeiro retorno ao origem ocorre no tempo 2. Ao saco 4 irão todos os caminhos cujo primeiro retorno ao origem ocorre no tempo 4. E assim por diante, até o saco  $2n$ . (Alguns sacos podem ficar vazios; por exemplo, se  $k \neq n$ , então o saco  $2n$  não contém nenhum elemento.)

Considere o saco  $2r$ . Tome um caminho que está nele. O caminho tomado volta para a origem pela primeira vez no instante  $2r$ , e isto pode acontecer um das maneiras: a primeira é quando todo o tempo até a volta, o caminho ficou acima do eixo de abcissas, e a segunda é quando ficou abaixo. Mas o caminho tomado deve permanecer acima do zero por tempo  $2k$  durante o período de tempo  $[0, 2n]$ . Isto implica no que no intervalo de tempo  $[2r, 2n]$  o caminho permaneceu acima do eixo por tempo  $2k - 2r$  no primeiro dos casos descritos acima, e por tempo  $2k$  no segundo dos casos. Da extensão cuidadosa deste raciocínio sai a seguinte relação:

$$p_{2k,2n} = \frac{1}{2} \sum_{r=1}^k f_{2r} p_{2k-2r,2n-2r} + \frac{1}{2} \sum_{r=1}^{n-k} f_{2r} p_{2k,2n-2r}$$

Use agora a suposição da indução e propriedades de  $f$ 's e  $u$ 's para deduzir desta relação que  $p_{2k,2n} = u_{2k} \cdot u_{2n-2k}$ .  $\square$

**Lema 4.**

$$p_{2k,2n} \sim \frac{1}{\pi \sqrt{k(n-k)}}$$

sendo que o sinal “~” significa que o quociente entre os lados da relação converge a 1 quando  $n \rightarrow \infty$ .

Dicas para a demonstração. Tudo apoia-se fortemente na Fórmula de Stirling:

$$n! \sim (2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n}$$

Desta e da expressão  $u_{2k} = C_{2k}^k 2^{-2k}$  você consegue deduzir que  $u_{2k} \sim (\pi k)^{-\frac{1}{2}}$ . Analogamente, segue-se que  $u_{2n-2k} \sim (\pi(n-k))^{-\frac{1}{2}}$ . Juntando os dois, conclui-se a afirmação do lema.  $\square$

**Teorema 5** (a primeira lei de arco seno). Fixe arbitrariamente  $\alpha \in (0, 1)$ . Considere PASS de comprimento  $2n$ ,  $n \in \mathbb{N}$ , e considere a probabilidade deste ser positivo em proporção de tempo inferior a  $\alpha$ . Então, o valor da probabilidade tende ao

$$2\pi^{-1} \arcsin(\alpha^{1/2}) \quad (1)$$

conforme  $n \rightarrow \infty$ .

Dica para a demonstração. Por razão da comodidade na exposição dos cálculos, tomo  $\alpha \in (1/2, 1)$  e procuro pela probabilidade da proporção do tempo passado pelo PASS acima do zero estar entre  $1/2$  e  $\alpha$ . De acordo com Theorema 4, a probabilidade em interesse expressa-se por

$$\sum_{k: \frac{1}{2} \leq \frac{2k}{2n} \leq \alpha} p_{2k,2n}$$

Lema 4 garante que esta expressão comporta-se assintoticamente como

$$\sum_{k: \frac{1}{2} \leq \frac{2k}{2n} \leq \alpha} \frac{1}{\pi \sqrt{k(n-k)}}$$

Curioso é, que esta somatória é calculável. Para o cálculo, expressaremos esta como

$$\frac{1}{\pi n} \sum_{k: \frac{1}{2} \leq \frac{2k}{2n} \leq \alpha} \frac{1}{\sqrt{\frac{k}{n}(1 - \frac{k}{n})}}$$

Em seguida, ao introduzir  $k^*$  como o maior inteiro tal que  $k^*/n \leq \alpha$ , re-escrevemos a última expressão assim:

$$\frac{1}{\pi n} \sum_{k=0,1,\dots,k^*} \frac{1}{\sqrt{(\frac{1}{2} + \frac{k}{n})(1 - (\frac{1}{2} + \frac{k}{n}))}}$$

A Figura C convence você que esta é a soma de Riemann correspondente à integral

$$\pi^{-1} \int_{1/2}^{\alpha} \frac{dx}{\sqrt{x(1-x)}}$$

Com a ajuda de nossos amigos Newton e Leibnitz, achamos o valor da integral:

$$\frac{2}{\pi} \arcsin(\alpha^{1/2}) - 1/2$$

Daqui até o término da demonstração não sobra muito trabalho. Fica por sua conta.  $\square$

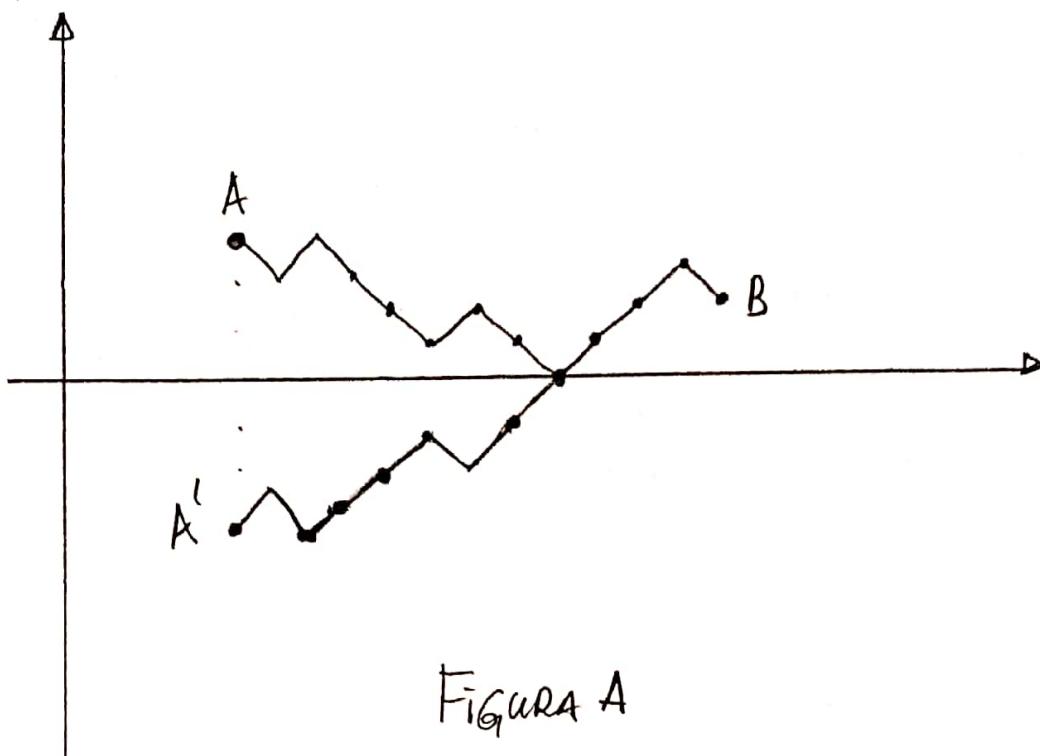


Figura A

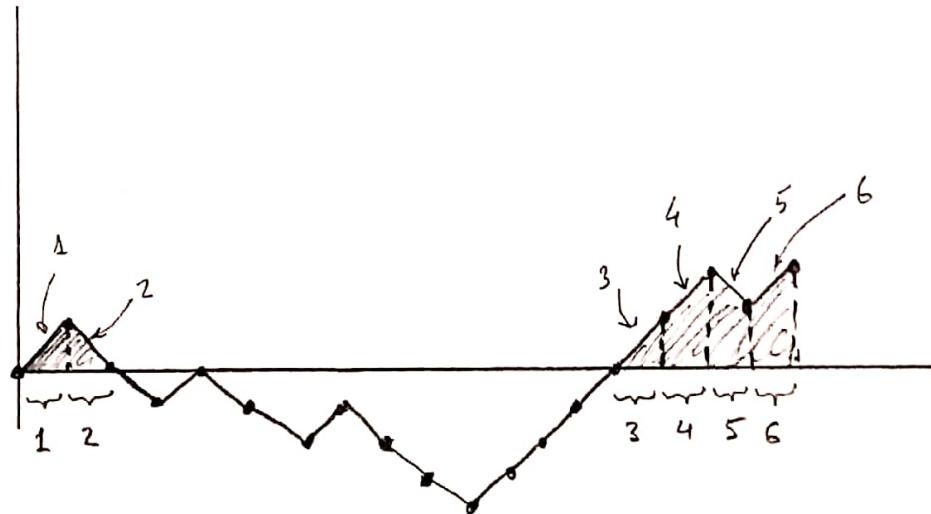


FIGURA B: este caminho passa  
tempo 6 acima do eixo de  
abscissas, ou, em outras palavras,  
no semi-plano positivo.

Curioso é que esta somatória é calculável. Para o cálculo, expressaremos esta como

$$\frac{1}{\pi n} \sum_{k: \frac{1}{2} \leq \frac{2k}{n} \leq \alpha} \frac{1}{\sqrt{\frac{k}{n}(1 - \frac{k}{n})}}$$

Em seguida, ao introduzir  $k^*$  como o maior inteiro tal que  $k^*/n \leq \alpha$ , re-escrevemos a última expressão assim:

$$\frac{1}{\pi n} \sum_{k=0,1,\dots,k^*} \frac{1}{\sqrt{(\frac{1}{2} + \frac{k}{n})(1 - (\frac{1}{2} + \frac{k}{n}))}}$$

A Figura C convence você que esta é a soma de Riemann correspondente à integral

$$\pi^{-1} \int_{1/2}^{\alpha} \frac{dx}{\sqrt{x(1-x)}}$$

Com a ajuda de nossos amigos Newton e Leibnitz, achamos o valor da integral:

$$\frac{2}{\pi} \arcsin(\alpha^{1/2}) - 1/2$$

Daqui até o término da demonstração não sobra muito trabalho. Fica por sua conta.  $\square$

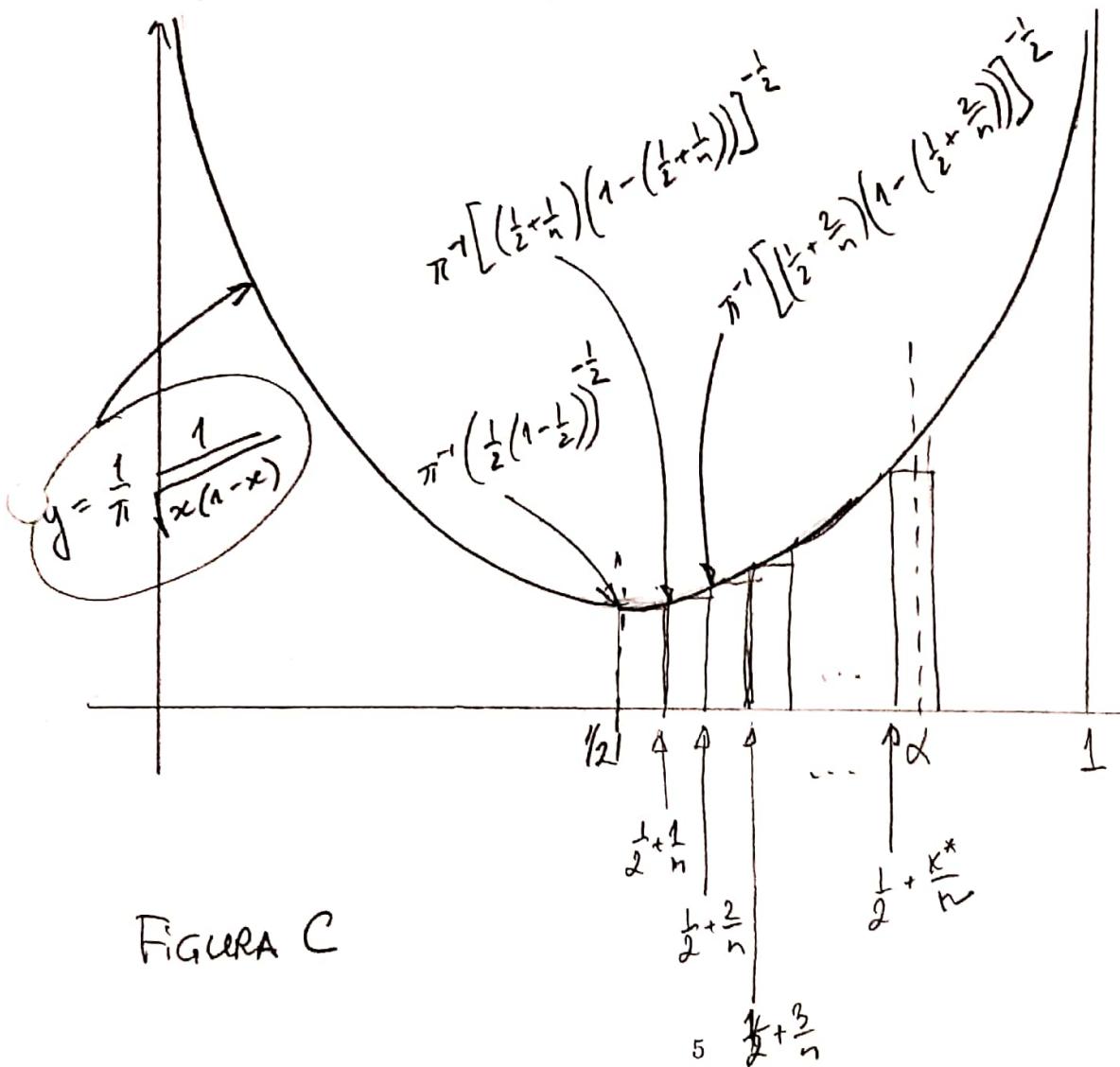


FIGURA C

## 1-a Aula

### Passeio aleatório simples

Fixamos um número inteiro positivo finito<sup>1</sup>  $n$  e um número real  $p$  do intervalo  $[0; 1]$  e consideraremos i.i.d. (quer dizer independentes e identicamente distribuidas) variáveis aleatórias  $X_1, X_2, \dots, X_n$  com a seguinte distribuição:

$$P[X_i = +1] = p, \quad P[X_i = -1] = 1 - p \quad (1)$$

Definimos variáveis aleatórias  $S_0, S_1, \dots, S_n$  pela seguinte maneira<sup>2</sup>

$$\begin{aligned} S_0 &\equiv 0, \\ S_1 &:= S_0 + X_1, \\ S_2 &:= S_1 + X_2 = X_1 + X_2, \\ \dots \\ S_n &= S_{n-1} + X_n = X_1 + \dots + X_n \end{aligned} \quad (2)$$

A sequencia<sup>3</sup> das variáveis aleatórias  $\{S_i\}$  agora construída se chama *passeio aleatório simples de comprimento n e de parâmetro p*. Quando  $p = 1/2$ , ele se chama *passeio aleatório simples simétrico*. As referências explícitas para o comprimento e o valor da parâmetro geralmente são omitidas, o que obriga leitor pescar esses valores no texto.

*Interpretação do passeio aleatório usando partícula que se move nos números inteiros.* Uma partícula coloca-se na origem de  $\mathbb{Z}$  e move-se nos instantes de tempos  $1, 2, \dots, n$  pela seguinte regra: em cada  $t$  do conjunto de instantes, a partícula pula para o número inteiro à direita da sua posição no instante  $t - 1$  caso  $X_t = +1$  e à esquerda caso  $X_t = -1$ . Claro que a posição da partícula no instante  $t$  corresponde a  $S_t$ .

A posição da partícula no instante  $t$  pode ser interpretada como o capital de um jogador que começa com capital 0, ganha 1R\$ cada vez que  $X$  assume valor 1 e perde 1R\$ cada vez que  $X$  assume valor -1. Esse joga "cara-coroa" tem, do ponto de vista desse jogador, a probabilidade de sucesso  $p$  em cada rodada; é por isso que o parâmetro  $p$  do passeio aleatório se chama as vezes *probabilidade de sucesso*.

*Interpretação do passeio aleatório usando caminhos em  $\mathbb{R}^2$ .* Cada realização  $0, s_1, s_2, \dots, s_n$  das variáveis aleatórias  $S_0, S_1, \dots, S_n$  pode ser apresentada por uma linha poligonal em  $\mathbb{R}^2$  cujos vértices têm coordenadas  $(0; 0), (1; s_1), \dots, (n; s_n)$ . Note que a associação entre caminho e realização  $0, s_1, s_2, \dots, s_n$  pode ser dada em termos da realização  $x_1, \dots, x_n$  que "gerou" a realização  $0, s_1, s_2, \dots, s_n$ : o  $i$ -ésimo elo do caminho sobe ou desce conforme  $x_i$ .

<sup>1</sup> Alias, números inteiros só podem ser finitos; os símbolos  $-\infty$  e  $+\infty$  são nada mais que símbolos. Eles adquiriram interpretações que são cómodas para formalização de argumentos e resultados. Quando for usar  $-\infty$  e  $+\infty$ , darei maiores detalhes.

<sup>2</sup>No meu texto " $:=$ " significa "igual por definição", enquanto " $=$ " indica o resultado de uma conta, por exemplo, no caso de (2) ele é o resultado da substituição de  $S_i$  por  $X_1 + \dots + X_i$ . O símbolo " $\equiv$ " se lê "sempre igual"; ele ressalta o fato de que  $S_0$  é sempre 0 embora foi introduzido como uma variável aleatória.

<sup>3</sup>Existem diversas notações para "sequencia"; por exemplo:  $\{S_i\}_{i=1}^n$  e  $\{S_i, i = 0, \dots, n\}$ , sendo que no lugar de ",," da última podem aparecer ";" ou ":". Uso no texto a abreviação  $\{S_i\}$ . Outros autores gostam outras abreviações, por exemplo,  $\{S\}$  ou simplesmente  $S$ .



for +1 ou -1. Note também, que essa associação implica imediatamente que existem  $2^n$  caminhos de comprimento  $n$ .

Cada uma das interpretações tem suas vantagens e desvantagens, e ambas serão utilizadas em nossos futuros argumentos (as vezes de maneira bem misturada).

**Lema 1.** O número de caminhos de comprimento  $n$  que partem de  $(0; 0)$  e chegam ao ponto  $(n, y)$  é igual  $C_n^{(n-y)/2}$  (onde  $C_n^{(n-y)/2}$  denota o número de combinações de  $n$  em  $(n-y)/2$  caso  $(n-y)/2$  é inteiro, e deve ser entendido como 0 caso  $(n-y)/2$  não é inteiro).

*Demonstração do Lema 1.* É óbvio que um caminho sai do  $(0, 0)$  e chega ao  $(n, y)$  se o somente se o número de seus elos que vão para cima menos o número de seus elos que vão para baixo é igual a  $y$ . Denotando esses números por  $n_{\text{up}}$  e  $n_{\text{down}}$  respectivamente, temos então, que um caminho liga  $(0, 0)$  com  $(n, y)$  se e somente se  $n_{\text{up}} - n_{\text{down}} = y$  e  $n_{\text{up}} + n_{\text{down}} = n$  (a última equação expressa o fato que o número total de elos é  $n$ ). Resolvendo esse sistema de duas equações lineares com dois incognitos (que são  $n_{\text{up}}$  e  $n_{\text{down}}$ ), obtivemos que  $n_{\text{up}} = (n-y)/2$ . Então o número de caminhos que ligam  $(0, 0)$  com  $(n, y)$  é igual a número de possibilidades de distribuir  $(n-y)/2$  "etiquetas" "up" para  $n$  elos (pois os que não receberão "up" serão obrigatoriamente "down"). Quem estudou análise combinatória (que é o primeiro assunto de qualquer curso básico de probabilidade) sabe que esse número é  $C_n^{(n-y)/2}$ .  $\square$

Abaixo no texto o conceito "caminho" será entendido de maneira mais ampla de que foi definido acima: agora um caminho pode partir de qualquer enteiro do  $\mathbb{R}^2$ .

**Lema 2.** (Princípio de reflexão) Sejam  $A$  e  $B$  dois pontos inteiros em  $\mathbb{R}^2$  tais que  $y$ -coordenada de  $A$  é positiva e  $x$ -coordenada de  $B$  é maior de que  $x$ -coordenada de  $A$ . Então o número de caminhos de  $A$  até  $B$  que tocam ou cruzam o eixo de abscissas é igual ao número de caminhos que vão de  $A'$  até  $B$ , onde  $A'$  é o ponto do plano obtido pela reflexão do ponto  $A$  em torno do eixo de abscissas.

*Demonstração do Lema 2* está na Figura 1.  $\square$

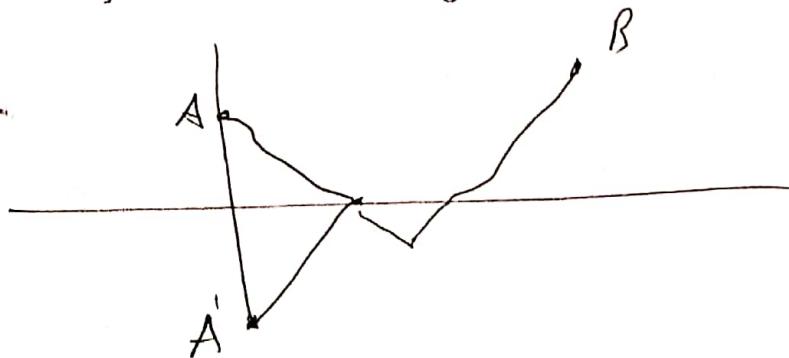
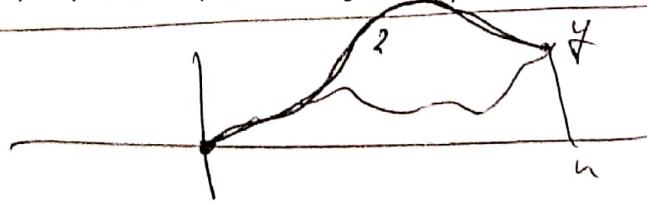


Figura 1. É claro que o argumento não muda de caso quando caminho crusa o eixo para o caso quando ele só toca no eixo.

**Teorema 1.** (Teorema sobre eleições.) Fixamos  $n$  e  $y > 0$ . Então o número de caminhos com comprimento  $n$  que ligam ponto  $(0, 0)$  com ponto  $(n, y)$  e satisfazem a condição que  $s_1 > 0, s_2 > 0, \dots, s_{n-1} > 0, s_n > 0$  é igual a  $y/n \cdot C_n^{(n-y)/2}$ .



*Demonstração do Teorema 1.* Os caminhos que satisfazem as condições do teorema, devem obrigatoriamente passar pelo ponto  $(1, 1)$ . Portanto, o número deles é igual ao número de caminhos que ligam  $(1, 1)$  com  $(n, y)$  e não tocam nem cruzam o eixo de abscissas. Esse número é igual à diferença entre todos os caminhos de  $(1, 1)$  até  $(n, y)$  e o número de caminhos de  $(1, 1)$  até  $(n, y)$  que ou tocam ou cruzam o eixo. Usando Lema 2, temos que o último número é igual ao número de caminhos de  $(1; -1)$  até  $(n, y)$ . Usando o resultado do Lema 1, podemos escrever a diferença:

$$\begin{aligned} C_{n-1}^{((n-1)-(y-1))/2} - C_{n-1}^{((n-1)-(y+1))/2} &= \frac{(n-1)!}{\left(\frac{n-y}{2}\right)!\left(\frac{n+y}{2}-1\right)!} - \frac{(n-1)!}{\left(\frac{n-y}{2}-1\right)!\left(\frac{n+y}{2}\right)!} \\ &= \frac{\left\{\frac{n+y}{2} - \frac{n-y}{2}\right\}(n-1)!}{\left(\frac{n-y}{2}\right)!\left(\frac{n+y}{2}\right)!} = \frac{y}{n} \cdot \frac{n!}{\left(\frac{n-y}{2}\right)!\left(\frac{n+y}{2}\right)!} = \frac{y}{n} \cdot C_n^{(n-y)/2} \end{aligned}$$

□

*Observação.* Infelizmente, dado  $n$ , não qualquer ponto  $(n, y)$  é "alcançável" pelos caminhos do passeio aleatórios; por exemplo, nenhum dos caminhos de comprimento 3 que parte de  $(0; 0)$  pode chegar ao ponto  $(3; 0)$ . Essa é uma incomodidade, pois cada vez que quisermos afirmar algo sobre propriedades dos caminhos teríamos que ressaltar que a afirmação extende-se somente aos pontos alcançáveis. Infelizmente temos que conviver com essa incomodidade pois não há maneira de eliminá-la completamente. Existe porém, uma maneira que é pouco mais eficiente que as outras no sentido de diminuição da incomodidade. Seguindo essa maneira, consideram-se somente passeios aleatórios com comprimento par. Isso implica que os caminhos desses passeios podem passar somente por inteiros pares nos instantes pares e somente nos inteiros ímpares nos instantes ímpares. Esse fato considera-se consumado, e assume-se implicitamente. Claro que passeios aleatórios com comprimento ímpar também têm direito a existir. Mas caso algem querer saber alguma coisa sobre esses "recusados" passeios, é só usar os métodos de análise desenvolvidos para passeios com comprimento par. Em fim, a partir de agora usaremos  $2n$  no lugar de  $n$  para os comprimentos de passeios aleatórios e seus caminhos.

**Teorema 2.** Definimos  $L_{2k} := \frac{1}{k+1} C_{2k}^k$ ,  $k = 1, 2, \dots$ . Então

- (a) há exatamente  $L_{2n-2}$  caminhos de comprimento  $2n$  que ligam ponto  $(0, 0)$  com  $(2n; 0)$  e satisfazem a condição de que  $s_1 > 0, s_2 > 0, \dots, s_{2n-1} > 0$  (e, óbvio, que  $s_0 = 0$  e  $s_{2n} = 0$ );
- (b) há exatamente  $L_{2n}$  caminhos de comprimento  $2n$  que ligam ponto  $(0, 0)$  com  $(2n; 0)$  e satisfazem a condição de que  $s_1 \geq 0, s_2 \geq 0, \dots, s_{2n-1} \geq 0$  (e, óbvio, que  $s_0 = 0$  e  $s_{2n} = 0$ ).

*Demonstração do Teorema 2 (a).* Os caminhos que nos interessam ligam o ponto  $(0; 0)$  com o ponto  $(2n-1; 1)$  e não tocam e não cruzam o eixo de abscissas. Pelo Teorema 1, o número de tais caminhos é

$$\frac{1}{2n-1} C_{2n-1}^{n-1} = \frac{1}{n} C_{2n-2}^{n-1} = L_{2n-2} \quad (3)$$

$\binom{2k}{k}$

□

*Comentário sobre Teorema 2.* O teorema junto com a demonstração dada parece ser uma enrolação pois, após uma análise meticulosa, descubra-se que o teorema introduziu um novo símbolo  $L$  e mostrou que ele expressa o número de caminhos que já foi estudado antes desse teorema. Na verdade, Teorema 2 oferece um resultado auxiliar para nossos futuros cálculos. Esses cálculos estudam os retornos de passeio aleatório para a origem. Já que os retornos podem acontecer somente nos instantes pares, então seria cômodo ter todas as expressões em termos dos instantes pares. É isso que Teorema 2 faz para nos.

*Comentário sobre a continuação da aula.* Até agora contavamos números de caminhos que satisfazem certas condições. A partir de agora nos interessaremos pelas probabilidade do passeio aleatório simples simétrico satisfazer certas condições. Em maioria dos casos, para obter essas probabilidades, contaremos o número de caminhos correspondente e o multiplicaremos por  $2^{-2n}$ , já que esta é a probabilidade do passeio aleatório simples simétrico seguir qualquer um dos caminhos. Aparentemente a mudança do objeto do nosso interesse prejudica o estudo de passeios não simétricos, mas na verdade tal estudo é muito mais complicado de que o de passeios simétricos, pois no caso de assimetria as probabilidades de seguir caminhos diferentes são diferentes; esse fato complica e, às vezes, até impossibilita a obtenção de resultados exatos sobre passeios aleatórios não simétricos.

**Teorema 3.** Para um passeio aleatório simples simétrico  $\{S\}$  de comprimento  $n$  definimos

$$u_{2k} := P[S_{2k} = 0], \quad f_{2k} := P[S_1 \neq 0, \dots, S_{2k-1} \neq 0, S_{2k} = 0], \quad \text{para } 1 \leq k \leq n \quad (4)$$

enquanto para  $k = 0$  definimos  $u_0 := 1, \quad f_0 := 0$

Então,



$$u_{2n} = C_{2n}^n 2^{-2n} \text{ para } n = 0, 1, \dots, \quad \text{e } f_{2n} = \frac{1}{2n} u_{2n-2} \text{ para } n = 1, 2, \dots \quad (5)$$

*Comentário sobre a interpretação de  $u$ 's e  $f$ 's e sobre a maneira como eles forma introduzidos.*  $u_{2k}$  é a probabilidade de que passeio aleatório simples simétrico de comprimento  $2n$  volta ao origem no instante  $2k$ , enquanto  $f_{2k}$  é a probabilidade de no instante  $2k$  ocorrer a primeira volta desse passeio a origem (a primeira volta sem contar que o passeio esteve na origem no tempo inicial; é por isso que colocamos  $f_0 := 0$ ; agora, quanto às razões da definição  $u_0 := 1$ , notarei que elas são puramente técnicas).

Nota-se que  $u_{2k}$  e  $f_{2k}$  foram definidas para passeio aleatório de comprimento  $2n \geq 2k$ . É fácil ver que  $u_{2k}$  e  $f_{2k}$  assumem os mesmos valores para todos os passeios de comprimento  $\geq 2k$ . Portanto, seria possível introduzir  $u$ 's e  $f$ 's da forma que não depende do comprimento do passeio envolvido na sua definição. A saber: poderíamos introduzir  $u$ 's e  $f$ 's para passeio aleatório de comprimento não limitado, quer dizer, para uma sequência infinita  $S_0, S_1, S_2, \dots$ . Isso é perfeitamente possível desde que se sabe que tal sequência existe. Você prova essa existência no Exercício ??.

*Demostração do Teorema 3.* Já sabemos que há exatamente  $C_{2n}^n$  caminhos que ligar  $(0; 0)$  com  $(2n; 0)$ . A probabilidade do passeio aleatório simples simétrico seguir qualquer um desses caminhos é  $2^{-2n}$ . Portanto,  $u_{2n} = 2^{-2n} C_{2n}^n$ . Seguindo o resultado do Teorema 2, há  $L_{2n-2}$  caminhos que voltam para origem no tempo  $2n$  pela primeira vez passando por

cima do eixo de abcissas. Tomando em conta que é possível também passar por baixo do eixo e que a probabilidade de seguir qualquer caminho é  $2^{-2n}$ , temos portanto que  $f_{2n} = 2L_{2n-2} \cdot 2^{-2n} = u_{2n-2}$ .  $\square$

Serão deixadas como exercícios as demonstrações das seguintes relações que serão usados a seguir.

$$f_{2n} = u_{2n-2} - u_{2n}, n = 1, 2, \dots \quad (6)$$

$$u_{2n} = P[S_1 \neq 0, S_2 \neq 0, \dots, S_{2n} \neq 0], n = 1, 2, \dots \quad (7)$$

$$u_{2n} = P[S_1 \geq 0, S_2 \geq 0, \dots, S_{2n} \geq 0], n = 1, 2, \dots \quad (8)$$

$$f_{2n} = P[S_1 \geq 0, S_2 \geq 0, \dots, S_{2n-2} \geq 0, S_{2n-1} < 0], n = 1, 2, \dots \quad (9)$$

**Lema 3.** Para qualquer  $n \geq 1$ , ocorre que

$$u_{2n} = \sum_{r=1}^n f_{2r} u_{2n-2r} \quad (10)$$

*Demonstração do Lema 3.* Escolhemos o instante  $2r$  maior que 0 mas menor que  $2n$ . Então a probabilidade de que o passeio aleatório volta a origem no instante  $2r$  pela primeira vez e esteja na origem também no instante  $2n$  é dada por  $f_{2r} u_{2n-2r}$ . A soma dessas probabilidades por todos  $2r$  entre 2 e  $2n$  dá, conforme a fórmula da probabilidade total (veja Exercício ??), a probabilidade de estar na origem no instante  $2n$ . Esta última é expressa por  $u_{2n}$ , da onde segue-se a identidade 10.  $\square$

**Teorema 4.** Designamos por  $p_{2k,2n}$  a probabilidade de que durante intervalo de tempo  $[0; 2n]$  o passeio aleatório passa  $2k$  unidades de tempo acima do eixo de abscissas (e, consequentemente,  $2n - 2k$  unidades de tempo abaixo do eixo de abscissas). Então

$$p_{2k,2n} = u_{2k} \cdot u_{2n-2k} \quad (11)$$

*Coemntário.* Note que os caminhos 0, 1, 0, 1, 2 e 0, 1, 2, 1, 2 passa acima do eixo de abscissas durante o mesmo tempo, que é 4, embora o primeiro deles tocou no eixo no instante 2, enquanto o segundo nunca tocou no eixo. Com essa observação quero chamar a atenção do leitor que "instante" não tem "duração" na nossa nomenclatura.

*Demostraçao do Teorema 4* será pelo método de indução matemática. Só que na fórmula (11) temos dois parâmetros que são  $n$  e  $k$ . Portanto a indução funcuinara da seguinte forma: assumimos que a fórmula (11) é verdadeira para todos os pares  $(k', n')$  onde  $n'$  é menor que um fixo  $n$  e demostramos a fórmula para qualquer par  $(k, n)$ .

Fixaremos então  $n$  e  $k$ , tal que  $1 \leq k < n$ . Fixaremos também  $r$ . Entre os caminhos de comprimento  $2n$  que passam tempo  $2k$  acima do eixo, consideraremos somente aqueles que voltam para a origem pela primeira vez no instante  $2r$ . Separaremos esses em dois sacos: para primeiro saco irão todos os caminhos que passaram por cima do eixo antes voltar a origem no instante  $2r$ , enquanto para o segundo saco irão aqueles que passaram por baixo do eixo. A quantidade de caminhos no primeiro saco se calcula pelo seguinte argumento: em primeiro lugar, há  $2^{2r} f_{2r}/2$  caminhos de comprimento  $2r$  que voltam para origem no instante  $2r$  pela primeira vez e estejam acima do eixo; em segundo lugar

há  $2^{2n-2r} p_{2k-2r, 2n-2r}$  caminhos de comprimento  $2n - 2r$  que saem de  $(2r; 0)$  e passam tempo  $2k - 2r$  acima do eixo. É claro que todo caminho do primeiro saco pode ser obtido pela colagem desses dois tipos de caminhos, portanto, no primeiro saco temos  $2^{2n-1} f_{2r} p_{2k-2r, 2n-2r}$  caminhos.

Calcularemos agora o número de caminhos no segundo saco. De modo análogo à argumentação acima, temos que qualquer um dos caminhos desse saco pode ser composto juntando a parte até instante  $2r$  com parte depois do instante  $2r$ . Ao total, temos  $2^{2r} f_{2r}/2$  caminhos que chegam pela primeira vez a origem no instante  $2r$  e passam por baixo do eixo, e temos  $2^{2n-2r} p_{2k, 2n-2k}$  caminhos de comprimento  $2n - 2r$  que partem de  $(2r; 0)$  e ficam tempo  $2k$  acima do eixo. Portanto, o segundo saco contém  $2^{2n-1} f_{2r} p_{2k, 2n-2r}$  caminhos.

Colocando todos os sacos junto, concluimos que, para  $1 \leq k < n$ ,

$$p_{2k, 2n} = \frac{1}{2} \sum_{r=1}^k f_{2r} p_{2k-2r, 2n-2r} + \frac{1}{2} \sum_{r=1}^{n-k} f_{2r} p_{2k, 2n-2r} \quad (12)$$

$$\text{usando suposição da indução} = \frac{1}{2} u_{2n-2k} \sum_{r=1}^k f_{2r} u_{2k-2r} + \frac{1}{2} u_{2k} \sum_{r=1}^{n-k} f_{2r} u_{2n-2k-2r} \quad (13)$$

Usando (10), concluimos que a primeira soma em (13) é  $u_{2k}$  e a segunda é  $u_{2n-2k}$ , e portanto (11) está provado para  $1 \leq k < n$ . Os casos  $k = 0$  e  $k = n$  seguem-se diretamente.  $\square$

Precisaremos de aproximação para

$$p_{2k, 2n} = u_{2k} \cancel{u_{2n-2k}} = C_{2k}^k 2^{-2k} \cdot C_{2n-2k}^{n-k} 2^{\cancel{k}-k} \quad (14)$$

que obtiveremos abaixo usando a fórmula de Stirling

$$n! \sim (2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n} \quad (15)$$

onde o sinal " $\sim$ " significa que o quociente de um lado pelo outro converge ao 1 quando  $n \rightarrow \infty$  (de fato, a diferença absoluta entre os dois lados em (15) tende a crescer conforme  $n$  cresce, porém, a diferença relativa diminui muito rápido: para  $n = 10$ , o erro relativo da aproximação é de 0,8%, enquanto para  $n = 100$  esse erro é só de 0,08%). Usando (15), temos que

$$\begin{aligned} u_{2k} &= C_{2k}^k 2^{-2k} = \frac{(2k)!}{k!(2k-k)!} \cdot 2^{-2k} \sim \\ &\sim \frac{(2\pi)^{\frac{1}{2}} (2k)^{2k+\frac{1}{2}} e^{-2k}}{(2\pi)^{\frac{1}{2}} (k)^{k+\frac{1}{2}} e^{-k} \cdot (2\pi)^{\frac{1}{2}} (k)^{k+\frac{1}{2}} e^{-k}} \cdot 2^{-2k} = \frac{2^{2k+\frac{1}{2}} k^{2k+\frac{1}{2}}}{(2\pi)^{\frac{1}{2}} k^{2k+1}} \cdot 2^{-2k} = \\ &= (\pi k)^{-\frac{1}{2}} \end{aligned}$$

Da maneira análoga, obtemos que  $u_{2n-2k} \sim (\pi(n-k))^{-\frac{1}{2}}$ , e juntando os dois temos portanto que

$$p_{2k, 2n} \sim \frac{1}{\pi \sqrt{k(n-k)}} \quad (16)$$

sendo que a rigor, devemos interpretar " $\sim$ " da mesma maneira como em (15) sem esquecer de acrescentar "conforme  $k \rightarrow \infty$  e  $n - k \rightarrow \infty$ ".

Usamos (16) para provar o seguinte resultado:

**Teorema 5.** (A primeira lei de arco seno.) *Para qualquer  $\alpha$  entre 0 e 1, a probabilidade que a proporção do tempo que passeio aleatório simples simétrico passa acima do eixo de abscissas esteja menor que  $\alpha$  tende ao valor*

$$\pi^{-1} \int_0^\alpha \frac{dx}{\{x(1-x)\}^{1/2}} = 2\pi^{-1} \arcsin(\alpha^{1/2}) \quad (17)$$

conforme  $n \rightarrow \infty$  (quer dizer, conforme a duração do passeio crescer).

*Demonstração do Teorema 5.* Fixaremos  $\alpha$  do intervalo  $(1/2; 1)$ . Por razões que serão claras para você quando fizer Exercício ? da 1-a Lista de Exercícios, em vez de procurar por probabilidade de que passeio aleatório esteja acima do eixo de abscissas a proporção  $\leq \alpha$  do tempo da sua vida, procurarei pela probabilidade dessa proporção estar entre  $1/2$  e  $\alpha$ . Devido à relação (16), a proporção em questão tem a seguinte expressão:

$$\sum_{k: \frac{1}{2} \leq \frac{2k}{2n} \leq \alpha} p_{2k, 2n} \approx \pi^{-1} \left\{ \frac{1}{n} \sum_{k: \frac{1}{2} \leq \frac{k}{n} \leq \alpha} \frac{1}{[\frac{k}{n}(1 - \frac{k}{n})]^{1/2}} \right\} \quad (18)$$

O sinal " $\approx$ " foi colocado para avisar que ainda não foi estabelecido que os dois lados de (18) são assintoticamente iguais no sentido do sinal " $\sim$ ". De fato, pelo que (16) nos garante, cada membro da soma no lado esquerdo em (18) está  $\sim$ -próximo ao correspondente membro da soma no lado direito. Porém, conforme  $n$  cresce, a quantidade de membros em cada soma cresce também e nada garante que os erros da  $\sim$ -aproximação não possam se acumular resultando no comportamento assintótico distinto dos dos lados de (18). Entretanto, presupussemos que essa divergência não ocorre e que os dois lados de (18) são iguais (sim, iguais!) e continuaremos nosso argumento sob essa suposição.

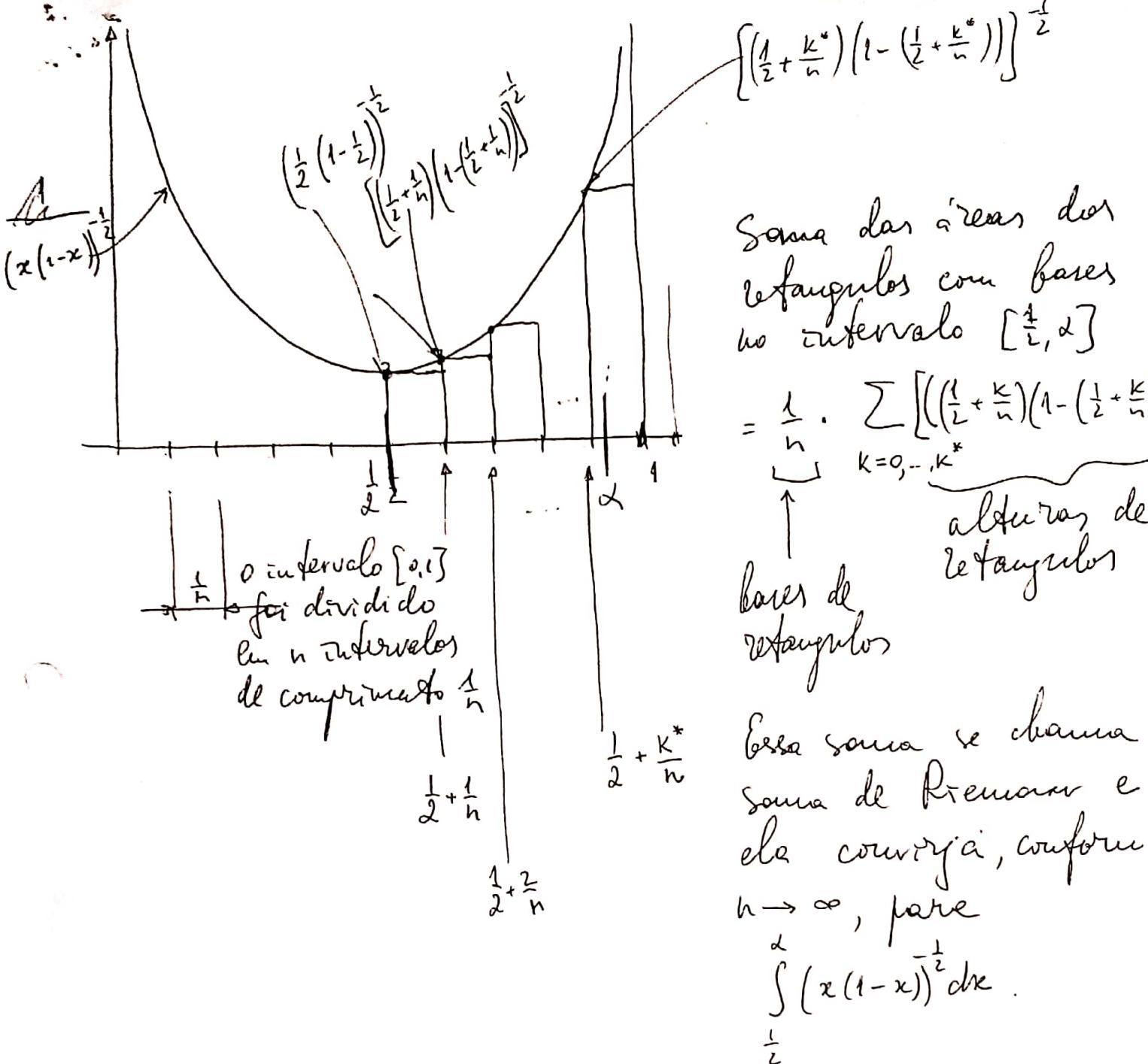
Considere Figura 2. Seja  $k^*$  o valor inteiro máximo para o qual  $k^*/n < \alpha$ . Então a soma das áreas dos retângulos cujos bases estão no intervalo  $[1/2; \alpha]$  é igual

$$\frac{1}{n} \cdot \sum_{k=0,1,\dots,k^*} \frac{1}{[(\frac{1}{2} + \frac{k}{n})(1 - (\frac{1}{2} + \frac{k}{n}))]^{1/2}} \quad (19)$$

pois  $1/n$  é o comprimento da base de cada retângulo e a soma contempla as alturas deles. A expressão (19) converge para

$$\int_{1/2}^\alpha (x(1-x))^{-1/2} dx \quad (20)$$

devido à própria definição de integral de Riemann. Mas a fórmula de Newton-Leibnitz nos permite calcular o valor de (20) explicitamente; ele é  $2 \arcsin(\alpha^{1/2}) - \pi/2$ . Combinando esse resultado com (18) concluímos que o limite da proporção procurada é  $2\pi^{-1} \arcsin(\alpha^{1/2}) - 1/2$ . Lembramos agora que a diferença entre a proporção em questão e a proporção em interesse do teorema é a probabilidade de que passeio aleatório esteja acima do eixo de abscissas a proporção do seu tempo de vida  $\leq 1/2$ . Por simetria, esta probabilidade deve ser  $1/2$ .  $\square$



De outro lado,

$$y = \arcsin x$$

$$y' = \frac{1}{\sqrt{1-x^2}}$$

$$y = \arcsin(\sqrt{x})$$

$$y' = \frac{1}{2} \frac{1}{\sqrt{1-(\sqrt{x})^2} \cdot \sqrt{x}} = \frac{1}{2} \frac{1}{\sqrt{(1-x)x}}$$

## AULA: Martingais limitados em $\mathcal{L}^2$ , decomposição de Doob, colchetes angulares, aplicação para Urna de Polya e Urna de Friedman

A junção dos temas que compõem o presente capítulo foi motivada pela seguinte argumento. Há uma classe de processos chamados “urna de Polya” que deve fazer parte da cultura de pós-graduandos em Probabilidade e Estatística. Não somente o resultado final é o que interessa nestes processos; seu tratamento (que apresento abaixo seguindo o livro de Durrett) envolve a Extensão de Lévy para o Lema de Borel-Cantelli, a qual, por sua vez, envolve a Relação entre a convergência de martingale e a convergência do processo de seus colchetes. Então, decidi apresentar os processos e demonstrar sua convergência na aula, e deixar para trabalho de casa a demonstração dos resultados clássicos sobre martingais que ampararam a apresentação da aula.

**Urna de Bernard Friedman.** Há uma urna com bolas verdes e vermelhas. Pense na evolução da proporção da bolas verdes que surge se formos repetir recursivamente nos tempos  $1, 2, \dots$ , o seguinte procedimento:

(\*) uma bola é retirada da urna, devolvida à urna, e ainda

$$a \text{ bolas da mesma cor e } b \text{ bolas da outra cor} \quad (1)$$

são acrescidas à urna.

Suponha, para apimentar a discussão, que você apostou na proporção de bolas verdes e ganhará de mim se ao repetir o passo (\*), essa proporção convergir a algum valor maior que  $1/2$ .

Suponha ainda, que te permito escolher os valores de  $a$  e de  $b$ , restringindo a escolha pela condição  $b > 0$ , e te permito escolher as quantidades de bolas no início deste processo todo (as quais denotarei no que se segue por  $G_0$  e  $R_0$ ; “G” de green e “R” de red).

Sua intuição te assopra o seguinte. Peça  $G_0$  muito grande e peça que  $a$  seja muito maior que  $b$ . No primeiro passo então, a probabilidade de retirar bola verde é quase que 1, e ao acontecer tal evento, você adicionará à urna mais bolas verdes que vermelhas, causando com isto a dominância de  $G_1$  acima de  $R_1$  que será ainda maior que era a de  $G_0$  acima de  $R_0$ . Isto te garante que no passo seguinte a situação se repetirá com grande probabilidade, e assim, no passar do tempo,  $G_t$  vai voar ao infinito deixando  $R_t$  pra trás. Como o resultado,  $g_t$ , a proporção de bolas verdes, vai tender a 1, e você ganhará do professor.

Curioso é, mas a conclusão sustentada pelo sua intuição, está errada. A verdade é que para quaisquer  $G_0, R_0, a$  e  $b$  (desde que  $b > 0$ ), acontece que

$$g_t \rightarrow 1/2 \text{ quase certamente, conforme } t \rightarrow \infty \quad (2)$$

A demonstração desta convergência é o conteúdo da presente seção. Eu não sei te dizer onde nossa intuição se enrosca e erra acerca do comportamento assintótico de  $g_t$ . Talvez, a fonte da confusão está no que a intuição sente e pensa em termos das quantidades absolutas das bolas, enquanto que a proporção vive, aparentemente, sendo guiada por leis diferentes das que guiam as frequências absolutas. Eis um exemplo. Suponha que

$G_0 = 100, R_0 = 1, a = 100, b = 1$ , e suponha que aconteceu o mais natural que pode: nos primeiros 10 instantes, foram retiradas bolas verdes. Então, as primeiras 11 proporções são

$$\frac{100}{101}, \frac{200}{202}, \frac{300}{303}, \dots, \frac{1100}{1111}$$

e todas são iguais! Eu não pretendo mais explorar os equívocos de nossa intuição. Vou diretamente às contas e torço pelo que essas revelem para você – se você for mais esperto que eu – onde a intuição se enroscou na derivação da conclusão de que  $g_t$  cresceria para 1.

Vamos fazer uma conta auxiliar que será utilizada em seguida para convencer nossa intuição da plausibilidade do resultado  $g_t \rightarrow 1/2$ . Antecipando-me, digo que o resultado da conta é nada intuitivo, e é por isto que sem o amparo desta a intuição pensava que  $g_t \rightarrow 1$ .

Assusmo que num momento  $t$  qualquer, a urna tem

$$\begin{array}{c} j \text{ bolas verdes} \\ i \text{ bolas vermelhas} \end{array} \quad (3)$$

e me atarefo a calcular a diferença entre

$$g_t = \frac{j}{j+i} \quad (\text{recorde, } g_t \text{ denota a proporção de bolas verdes no tempo } t)$$

e

$$\mathbb{E}[g_{t+1} \mid \text{no tempo } t \text{ há } j \text{ bolas verdes e } i \text{ vermelhas}]$$

Observe que o evento condicionador na última esperança condicional é mais abrangente que  $g_t$ , isto é, a esperança condicional acima não é (ou melhor, pode não ser)  $\mathbb{E}[g_{t+1} \mid g_t]$ . A diferença entre as duas será lhe clara no futuro quando você ver o resultado final da presente tarefa. Recapitulando e concluindo, quero calcular

$$\mathbb{E}[g_{t+1} \mid \text{no tempo } t \text{ há } j \text{ bolas verdes e } i \text{ vermelhas}] - g_t \quad (4)$$

cuja interpretação é a média do drift que age entre os tempos  $t$  e  $t+1$  dada a condição (3) no tempo  $t$ .

É óbvio que a cor da bola a ser vista na retirada no tempo  $t+1$  é

$$\begin{array}{ll} \text{verde} & \text{com a probabilidade } \frac{j}{j+i} \\ \text{vermelha} & \text{com a probabilidade } \frac{i}{j+i} \end{array}$$

e, portanto, em respectivos casos,

$b$  bolas verdes e  $a$  bolas vermelhas serão acrescidas à urna levando assim  $g_{t+1}$  a assumir o valor  $(j+a)/(j+i+a+b)$   
ou

$a$  bolas verdes e  $b$  bolas vermelhas serão acrescidas à urna levando assim  $g_{t+1}$  a assumir o valor  $(j+b)/(j+i+a+b)$

Portanto, (4) é igual a

$$\begin{aligned}
 & \left\{ \frac{j+a}{j+i+a+b} \cdot \frac{j}{j+i} + \frac{j+b}{j+i+a+b} \cdot \frac{i}{j+i} \right\} - \frac{j}{j+i} \\
 &= \left\{ \frac{j+a}{j+i+a+b} \cdot \frac{j}{j+i} + \frac{j+b}{j+i+a+b} \cdot \frac{i}{j+i} \right\} - \left\{ \frac{j}{j+i} + \frac{i}{j+i} \right\} \frac{j}{j+i} \\
 &= \left\{ \frac{j+a}{j+i+a+b} - \frac{j}{j+i} \right\} \frac{j}{j+i} + \left\{ \frac{j+b}{j+i+a+b} - \frac{i}{j+i} \right\} \cdot \frac{i}{j+i} \\
 &= \frac{ai-bj}{(j+i)(j+i+a+b)} \cdot \frac{j}{j+i} + \frac{bi-aj}{(j+i)(j+i+a+b)} \cdot \frac{i}{j+i} \\
 &= \frac{b(i^2-j^2)}{(j+i)^2(j+i+a+b)} = b \left( \frac{i}{j+i} - \frac{j}{j+i} \right) \cdot \frac{1}{j+i+a+b}
 \end{aligned} \tag{5}$$

O resultado desencadeia as seguintes observações

**Observação 1** oriunda da conta (4). Se  $b = 0$ , quer dizer, se bolas da cor oposta à retirada, não são acrescidas a urna, então a proporção de bolas verdes comporta-se como um martingal. Isto sugere que a teoria de martingais pode ajudar em alguma coisa. No que extatamente? Você acha a resposta fazendo o exercício formulado abaixo.

**Exc. 1.** Considere o caso especial da urna de Friedman caracterizado por seguintes valores de seus parâmetros:

- $G_0 = 1$  no início, o número de bolas verdes é 1
  - $R_0 = 1$  no início, o número de bolas vermelhas é 1
  - $a = 1$  o número de bolas da mesma cor que a da bola escolhida, acrescidas à urna é 1
  - $b = 0$  o número de bolas da mesma cor que a da bola escolhida, acrescidas à urna é 1
- (6)

Prove que  $g_t$ , proporção de bolas verdes no tempo  $t$ , converge quase certamente a variável aleatória distribuída uniformemente no  $[0, 1]$ .

**Comentário ao Exc. 1.** A conta (5) nos mostra que no presente caso, a sequência  $\{g_t\}_0^\infty$  é um martingal. Como cada  $g_t$  é uma proporção, então este martingal é limitado por cima e por baixo pelo 1 e 0 respectivamente. Portanto,  $g_t$  converge quase certamente. Denotando por  $g_\infty$  a variável limite, temos então que  $g_t \rightarrow g_\infty$  q.c. conforme  $t \rightarrow \infty$ , e isto é tudo que podemos exprimir do teorema sobre a convergência de martingais. Já se nós nos interessar pela distribuição da  $g_\infty$ , aí precisamos fazer contas específicas relacionadas à natureza do processo  $\{g_t\}$ . Você acha as contas no livro de Durrett, na seção intitulada pelo "Polya's urn scheme", pois este é o nome para o caso específico da urna de Friedman aqui considerado. É curioso observar que para  $G_0, R_0, a$  arbitrários (quer dizer, não "todos iguais a 1" como requeremos em (6)), a distribuição de  $g_\infty$  é Beta com parâmetros  $G_0/a$  e  $R_0/a$  (é o resultado citado mas não provado pelo Durrett na seção supracitada de seu livro). A distribuição Beta colocaria mais massa probabilística nos valores próximos ao 1 se  $G_0$  fosse maior que  $R_0$ . Se for pensar, este comportamento limite do  $g_t$  pode ser concebido por argumentos usados pela nossa intuição no nosso raciocínio (\*\*), mas eu não quero entrar em detalhes a respeito. Só queria mesmo indicar que o comportamento da urna de Polya não é totalmente contraintuitivo.

**Observação 2** oriunda da conta (4). Onde que estiver  $g_t$ , o drift médio que a leva para o tempo  $t+1$ , age para garantir o equilíbrio entre as proporções das cores na urna: se

$g_t > 1/2$ , então, em média,  $g_{t+1}$  será menor que  $g_t$ , enquanto que se  $g_t < 1/2$ , então em média,  $g_{t+1}$  será maior que  $g_t$ . Este fato faz acreditar, no nível intuitivo, que  $g_t$  converge para  $1/2$ ? Parece que sim, mas espere, pois há mais uma observação a vir que distruirá a crença. É a observação seguinte.

Observação 3 oriunda da conta (4). Deixamos de ver uma propriedade de  $\{g_t\}$  embutida em (4) por que não refletimos que as frequências absolutas crescem com o crescer do tempo. Vamos recuperar esta propriedade agora. Para tal, introduzimos  $G_t$  e  $R_t$  para denotar as frequências absolutas de bolas verdes e vermelhas no tempo  $t$ , e usando estas notações junto com já introduzidas notações  $g_t$  e  $r_t$  para representar as frequências relativas, reescrevemos o resultado final da conta (4) assim:

$$\frac{b(r_t - g_t)}{G_{t+1} + R_{t+1}} = \frac{b(1 - 2g_t)}{G_{t+1} + R_{t+1}} \quad (7)$$

Agora, recordamos que em cada instante de tempo, acrecemos  $a + b$  bolas, ao total, independentemente da cor da bola escolhida. Portanto,  $G_{t+1} + R_{t+1} = G_0 + R_0 + (t+1)(a+b)$ , expressão que, ao ser incerrida no (7), dá o seguinte resultado:

$$\frac{b(1 - 2g_t)}{G_0 + R_0 + (t+1)(a+b)} \quad (8)$$

Olhando no (8), podemos observar a seguinte propriedade: O valor absoluto do drift médio ao qual  $g_t$  está submetido no intervalo de tempo  $[t, t+1]$ , diminui com o crescimento de  $t$ . Portanto, para os grandes valores de  $t$ , o drift pode ser insuficiente para poder “devolver”  $g_t$  ao valor  $1/2$ , embora tenda a cuasar tal devolução, de acordo com Observação 2 feita acima. Concluindo: não tempos razões fortes o suficiente para acreditar – no nível de nossa intuição – que  $g_t$  é obrigado a convergir a  $1/2$  quase certamente.

Vamos então para a **demonstração rigorosa** da convergência quase certa de  $g_t$  ao  $1/2$ . Denotamos:

por  $B_n$  o evento que a  $n$ -ésima retirada resultou em bola verde (9)

por  $D_n$  o número de bolas verdes adicionadas nos  $n$  primeiros passos (10)

O resultado que ocupa a posição central na nossa demonstração é que

$$\frac{D_n}{\sum_{m=1}^n g_{m-1}} \rightarrow 1 \text{ quase certamente conforme } n \rightarrow \infty \quad (11)$$

Antecipando-me, aviso que este resultado dervar-se-á com uso de resultados fortes e genéricos sobre comportamento de martingais. Recordo: todo o exemplo tratado agora, veio à tona por causa desta sua relação com martingais.

A convergência alegada em (11) ajuda na derivação de resultado desejado via a seguinte corrente de argumentos. Primeiramente, definimos

$$x := \limsup_t g_t \quad (12)$$

Observe que  $g_t$  está sempre entre 0 e  $t$ , e isto garante que  $x$  é um número finito não negativo.

Em segundo lugar, com o uso da relação central (11), deduzimos que

$$\limsup_n \frac{D_n}{n} \leq x \text{ (onde } x \text{ foi definido acima como } \limsup_n g_n) \quad (13)$$

A derivação de (13) a partir de (12) é simples, mas, para a completude de contas, será apresentada após fechar o corrente raciocínio.

Em terceiro lugar, escrevemos o óbvio e um pouquinho a mais:

$$g_{n+1} = \frac{G_0 + aD_n + b(n - D_n)}{G_0 + R_0 + n(a + b)} = \frac{G_0 + bn + D_n(a - b)}{G_0 + R_0 + n(a - b)} \quad (14)$$

(A primeira igualdade foi chamada de "obvia", pois: (i) o numerador é a quantidade de bolas verdes no início ( $G_0$ ), acrescida a quantidade de bolas verdes colocadas na urna quando a retirada resultava na cor verde ( $aD_n$ ), e acrescida ainda a quantidade de bolas verdes quando a retirada resultava na cor vermelha ( $b(n - D_n)$ ); portanto, o numerador apresenta a quantidade total de bolas verdes na urna após completar o passo  $n$ ; (ii) o denominador apresenta a quantidade total de bolas na urna após o término do  $n$ -ésimo passo.) E assumindo, por um instante, que

$$a > b \quad (15)$$

chegamos a concluir a partir de (14), com a ajuda de (12) e (13), que

$$x \leq \frac{b + x(a - b)}{a + b} \quad (16)$$

desigualdade que equivale à  $x \leq 1/2$ , quer dizer,  $\limsup_n g_n \leq 1/2$ . Aqui paramos e voltaremos ao início, repetindo todo o argumento, mas com  $r_t$  no lugar de  $g_t$ . Tal repetição nos dá que  $\limsup_n r_n \leq 1/2$ . As duas últimas desigualdades garantem que  $\limsup_n g_n = \liminf_n g_n = 1/2$ , o que termina a demonstração da tal desejada propriedade  $\lim_n g_n = 1/2$ .

Antes de decretar a vitória, é preciso recordar que no caminho, tivemos que assumir que  $a > b$ . Precisamos pensar nos casos complementares. O caso  $a = b$  é simples, pois sob este,

$$g_{n+1} = \frac{G_0 + an}{G_0 + R_0 + 2an} \rightarrow \frac{1}{2}$$

O caso  $a < b$  é pouco mais exigente, mas felizmente, tem saída. A idéia é considerar  $g_t$  e  $r_t$  ao mesmo tempo, em vez de em separado, como foi feito acima. Então, deixando  $x$  de acordo com a já introduzida definição (12), acrescentamos ao argumento o limite de  $r_t$  definido assim:

$$y := \limsup_n r_n \quad (17)$$

Este  $y$  surge com a tomada de  $\limsup$  dos ambos os lados da igualdade abaixo, cuja derivação segue as mesmas idéias que usamos acima na derivação de (14):

$$r_{n+1} = \frac{R_0 + bD_n + a(n - D_n)}{G_0 + R_0 + n(a + b)} \quad (18)$$

Então temos

$$y \leq \frac{a+x(b-a)}{a+b} \text{ e } x \leq \frac{a+y(b-a)}{a+b} \quad (19)$$

Estas duas desigualdades implicam no que  $x \leq 1/2$  e  $y \leq 1/2$  por via do argumento recursivo da seguinte natureza: Suponhamos que  $x \leq 1$  e  $y \leq 1$  e colocando 1's nas desigualdades (19), obtemos que  $y \leq b/(a+b)$  e  $x \leq b/(a+b)$ . Repetimos o argumento, mas com o limite  $b/(a+b)$  no lugar de 1. Como resultado, obtivemos um novo limite para  $y$  e  $x$  que é ainda menor de  $b/(a+b)$ . Continuando o processo, chegaremos ao resultado  $x \leq 1/2$  e  $y \leq 1/2$  (pois  $1/2$  é o ponto fixo da função  $u = \frac{a+(b-a)t}{a+b}$ , já que estamos tratando o caso  $a < b$ ). Este resultado garante que  $x = y = 1/2$ , o que fecha a demonstração da convergência  $g_t \rightarrow 1/2$ . (Eu só não verifiquei se o resultado  $x = y = 1/2$  seja a consequência imediata e direta das desigualdades  $x \leq 1/2$  e  $y \leq 1/2$ , ou se seja preciso repetir o argumento, trocando  $\limsup$  pelo  $\liminf$ , ou algo do gênero semelhante a troca que fizemos para o caso  $a > b$ .)

Vamos agora tamar buracos que deixamos no argumento acima. O primeiro é a demonstração da propriedade  $\limsup_n (D_n/n) \leq x$  a partir da definição  $x := \limsup_n g_n$  e da propriedade  $D_n/(\sum g_{m-1}) \rightarrow 1$ . É um exercício de cálculo puro, mas eu apresentarci-o neste texto pois não quero que você gaste seu tempo com isto.

$$\begin{aligned} x := \limsup_n g_n &\Rightarrow \forall \delta > 0 \exists N(\delta) \text{ tal que } g_n \leq x + \delta \forall n \geq N(\delta) \\ &\Rightarrow \forall \delta > 0 \exists M(\delta) \text{ tal que } \frac{\sum_{m=1}^n g_{m-1}}{n} \leq x + 2\delta \forall n \geq M(\delta) \\ &\Rightarrow \forall n \geq M(\delta) \text{ tem se: } \frac{D_n}{\sum_{m=1}^n g_{m-1}} = \frac{(D_n/n)}{\sum_{m=1}^n g_{m-1}/n} \geq \frac{(D_n/n)}{x + 2\delta} \\ &\Rightarrow \text{se fosse } \limsup_n \frac{D_n}{n} > x \text{ teria } \limsup_n \frac{D_n}{\sum_{m=1}^n g_{m-1}} > 1 \\ &\quad \text{contradizendo } D_n/(\sum g_{m-1}) \rightarrow 1 \end{aligned}$$

O segundo buraco deixado no argumento é a demonstração da propriedade  $D_n/(\sum g_{m-1}) \rightarrow 1$ . Tal demonstração é o cerne da presente capítulo, conforme já avisado. Mas antes de partir para o formalismo, vou apresentar a demonstração intuitiva. Começo, recordando que  $B_i$  (a ser substituído abaixo pela notação  $B(i)$ ) denota o evento "a  $i$ -ésima retirada resultou na cor verde". Deste jeito,

$$D_n = \mathbb{I}_{B(1)} + \mathbb{I}_{B(2)} + \cdots + \mathbb{I}_{B(n)}$$

e, consequentemente,

$$\frac{D_n}{\sum g_{m-1}} = \frac{\mathbb{I}_{B(1)} + \mathbb{I}_{B(2)} + \cdots + \mathbb{I}_{B(n)}}{g_0 + g_1 + \cdots + g_{n-1}} \quad (20)$$

Suponha que o processo (de retiradas de bolas) esteja no seu estado estacionário desde o início. Isto significa que todas as proporções  $\{g_i\}$  são mais ou menos iguais; denotamos por  $g$  o "valor comum" delas. Com esta notação, podemos substituir o denominador de (20)

por  $ng$ . Agora vamos lembrar que  $g_0$  é a probabilidade de acontecer  $B_1$ , e que  $g_1$  é a probabilidade de acontecer  $B_2$ , e assim por diante. Portanto, em  $n$  ensaios, aproximadamente  $ng$  dos  $B$ 's aconteceram e  $n(1 - g)$  não. Isto significa que o valor do numerador dc (20) é aproximadamente  $ng$ , o que é também o valor de denominador, conforme descoberto acima. Logo, a fração toda é aproximadamente 1, conclusão que bate com o resultado do argumento rigoroso a ser apresentado em seguida.

Mas antes de partir para o argumento rigoroso, gostaria de revelar uma propriedade do processo estudado, quer dizer, da Urna de Friedman. A propriedade esta é a resposta na pergunta "Por que conseguimos o resultado tal forte sobre o comportamento assintótico de  $g_t$  na Urna de Friedman?"

Ao meu ver, a Urna de Friedman possui uma propriedade intrínseca que "amarra" os concitos e variáveis permitindo assim a derivação do resultado  $g_t \rightarrow 1/2$ . A propriedade é que  $g_t$  tem papel duplo no tudo que acontece: por um lado, é a proporção de esferas verdes no tempo  $t$ , mas por outro lado é a probabilidade que guia a evolução de  $g_{t+1}$  a partir de  $g_t$ . Esta duplidade pode ser vista em, no mínimo, dois lugares. O primeiro é a conta (5). Observe,  $\frac{j}{j+i}$  no final da primeira linha desta conta; neste lugar, esta fração denota  $g_t$ . Agora observe a mesma fração dentro das chaves na primeira linha; lá, ela representa a probabilidade. Mas para frente, fica pior: observe as primeiras duas apariências de  $\frac{j}{j+i}$  na terceira linha; veja, uma delas é  $g_t$  e outra é a probabilidade (de assunção do valor  $(j+a)/(j+i+a+b)$  pela variável aleatória  $g_{t+1}$ ). O segundo dos dois lugares supramencionados, é a fração (20). Nela, cada  $g_i$  do denominador é a probabilidade de acontecer  $B_{i+1}$ , enquanto que o numerador possui relação rígida com  $g_n$ : eis a relação

$$\frac{aD_n + b(n - D_n)}{n} \approx g_{n+1}, \text{ sendo que o erro da aproximação é } \frac{G_0}{n} \text{ que desaparece com } n \uparrow \infty$$

Enfim, esta seria minha resposta intuitiva na pergunta sobre o por quê conseguiu-se o resultado tão preciso e poderoso no caso da Urna de Friedman.

**Decomposição de Doob.** Suponha que temos um processo  $\{X_n\}_{n=0}^{\infty}$ . O termo "temos" significa que sabemos a lei probabilística que rege a distribuição conjunta do conjunto  $\{X_n\}$  de variáveis aleatórias. Vamos chamar por  $(\Omega, \mathcal{F}, \mathbb{P})$  o triplo probabilístico no qual este conjunto está definido; com isto, aquilo que foi chamado acima por "lei" adquire a notação  $\mathbb{P}$ . Ainda, para a necessidade, que surgirá daqui a pouco, vamos introduzir a filtragem  $\{\mathcal{F}_n\}_{n=0}^{\infty}$ ; nosso acordo é que  $X_n$  é  $\mathcal{F}_n$  mensurável para cada  $n$ , ou, em outras palavras, que o processo  $\{X_n\}$  é adaptado à filtragem  $\{\mathcal{F}_n\}$ .

Suponha agora que você observou seu processo até (inclusive) o instante  $n$ , quer dizer, você sabe as realizações das variáveis  $X_0, \dots, X_n$ . Lhe pergunto: "O que você pode falar sobre  $X_{n+1}$ ?" Sua resposta é:

- Só sei a distribuição do valor; esta é

$$\mathbb{E}[X_{n+1} | x_0, x_1, \dots, x_n] \text{ (aqui, } m_i \text{ é o valor observado de } X_i\text{)}. \quad (21)$$

Em palavras, o valor no instante  $n + 1$  está incerto, mas a caráter probabilístico da incerteza está totalmente conhecido.

Agora vamos calcular a média da distribuição de  $X_{n+1}$  condicionada ao conhecimento dos valores de  $X_0, \dots, X_n$ ; eis a expressão formal daquilo que sugiro calcular:

$$\mathbb{E}\left\{\mathbb{E}[X_{n+1} | x_0, x_1, \dots, x_n]\right\} =: A_{n+1}$$

Eu ainda introduzi a notação para esta média:  $A_{n+1}$ . Observe, como concordamos que toda a distribuição do processo  $\{X_n\}$  está nos conhecida, então o valor de  $A_{n+1}$  é conhecido também. Isto te permite responder à minha pergunta “O que você pode falar sobre  $X_{n+1}$ ?” de uma maneira diferente daquela usada acima. Você responde: “Em média, vou estar na posição  $A_{n+1}$ , mas o valor exato só é conhecido em forma de distribuição probabilística, e esta é a distribuição de

$$\mathbb{E}[X_{n+1} | x_0, x_1, \dots, x_n] = A_{n+1} \quad (22)$$

Suas duas respostas dão a mesma informação, só que em maneiras diferentes; a primeira delas dá a distribuição da posição futura, e a segunda dá o valor médio e a distribuição em torno deste valor. Observe dois fatos:

- (i) Se  $x_0, x_1, \dots, x_n$  mudarem, o valor de  $A_{n+1}$  a princípio, pode mudar também. Portanto,  $A_{n+1}$  é uma variável aleatória, e consequentemente,  $\{A_{n+1}\}_{n=0}^{\infty}$  é um processo estocástico. Pale nossa construção,  $A_{n+1}$  é  $\mathcal{F}_n$  mensurável (de fato,  $A_{n+1}$  depende só do histórico das  $X_0, X_1, \dots, X_n$ , o que é um evento em  $\mathcal{F}_n$ ). Devido minha escolha de indexação, o processo  $\{A_n\}_{n=1}^{\infty}$  é previsível em relação à filtração  $\{\mathcal{F}_n\}_{n=0}^{\infty}$ .
- (ii) Devido à própria construção, a flutuação ao redor de  $A_{n+1}$  possui média nula.

O argumento acima apresentado, acrescido de um pouquinho de formalismo, nos dá

**Decomposição de Doob.** Se  $\{X_n\}_{n=0}^{\infty}$  for um processo adaptado à uma filtração  $\{\mathcal{F}_n\}_{n=0}^{\infty}$ , e se  $X_n \in \mathcal{L}^1$  para cada  $n$  (isto é,  $\mathbb{E}[|X_n|]$  existe, a exigência necessária para garantir a existência das esperanças condicionais envolvidas na construção), então existem o processo previsível  $\{A_n\}_{n=1}^{\infty}$  e um martingale  $\{M_n\}_{n=0}^{\infty}$  com  $M_0 \equiv 0$  (ambos os processos são relacionados à filtração supracitada  $\{\mathcal{F}_n\}_{n=0}^{\infty}$ ) tais que

$$X_n = X_0 + A_n + M_n, \forall n \quad (\text{esta é o que chama-se decomposição de Doob}) \quad (23)$$

Os detalhes da Decomposição de Doob você pega no livro de Williams, na seção 12.11.

**O processo de colchetes angulares.** A decomposição de Doob introduzida acima, pode ser aplicada a qualquer processo  $X$ . Um caso particular é quando

$$X = M^2 \text{ onde } M \text{ é uma martingal satisfazendo } M_0 \equiv 0 \text{ e } \mathbb{E}[M_n^2] < \infty, \forall n \quad (24)$$

Neste caso, na decomposição de Doob

$$M_n^2 = N_n + A_n \text{ onde } \{N_n\} \text{ é martingal com } N_0 \equiv 0 \text{ e } \{A_n\}_{n=1}^{\infty} \text{ é previsível.} \quad (25)$$

o segundo, quer dizer, o  $\{A_n\}$ , chama-se *the angle-brackets process* de  $M$  e denota-se por  $\langle M \rangle$ .

**Exercício 2.** Considere o passeio aleatório simples simétrico  $S_0, S_1, S_2, S_3$ . Construa os processos  $A$  e  $N$  da decomposição de Doob do processo  $S_0^2, S_1^2, S_2^2, S_3^2$ . (Observação: este é um simples exercício que é bom para você entender o que acontece na decomposição.)

**Exercício 3.** Observe que se  $\{M\}$  for um martingal, então  $\{M^2\}$  é submartingal. Deduza desta propriedade o seguinte fato:

$$\text{para cada } \omega \in \Omega, \{A_n(\omega)\}_{n=1}^{\infty} \text{ é uma sequência não decrescente} \quad (26)$$

Observe que o fato (26) garante que a definição

$$A_{\infty} : \lim_{n \rightarrow \infty} A_n \quad (27)$$

define corretamente uma variável aleatória extendida  $A_{\infty} : \Omega \rightarrow \mathbb{R} \cup \{\infty\}$ . (Observação: Aqui e em tudo, “extendida” significa que a variável aleatória pode assumir “valor”  $\infty$ . Este é atribuído por decreto: se  $a_n \uparrow \infty$  então  $a_{\infty} := \infty$ . Mas é preciso tomar cuidado com as operações envolvendo variáveis aleatórias extendidas; por exemplo, se  $A_{\infty}(\omega) = \infty$  e  $B_{\infty}(\omega) = \infty$ , então,  $A_{\infty}(\omega) + B_{\infty}(\omega)$  é naturalmente  $\infty$ , mas  $A_{\infty}(\omega) - B_{\infty}(\omega)$  não é definido.)

**Exercício 4.** Prove que martingal  $M$  é limitado em  $L^2$  se e somente se  $\mathbb{E}[A_{\infty}] < \infty$ . (Observação 1: Se a medida do conjunto onde  $A_{\infty}$  assume  $\infty$  for não nulo, então, por decreto,  $\mathbb{E}[A_{\infty}] = \infty$ . Observe que esta última igualdade pode valer mesmo quando  $A_{\infty}$  não assume valor  $\infty$  em nenhum  $\omega$ . Então, a condição  $\mathbb{E}[A_{\infty}] < \infty$  quer dizer que nenhum destes dois casos ocorre. Observação 2: Para desconfundir sua cabeça, caso ela de fato ficou confusa, esclareço que “martingal em  $L^2$ ” e “martingal limitado em  $L^2$ ” são conceitos diferentes. A saber:

“martingal em  $L^2$ ” significa:  $\mathbb{E}[M_n^2] < \infty$  para cada  $n$

“martingal limitado em  $L^2$ ” significa: existe  $L < \infty$  tal que  $\mathbb{E}[M_n^2] \leq L$  para cada  $n$

Observação 3: Talvez a tarefa deste exercício seja a consequência direta da propriedade (28) do exercício que segue.

**Exercício 5 (a).** Seja  $M_0 \equiv 0, M_1, M_2, M_3$  o martingale cujos acréscimos  $X_1, X_2, X_3$  são variáveis aleatórias independentes, cada uma com média 0 e variância finita  $\sigma_i^2$  (isto é:  $\text{Var}(X_i) = \sigma_i^2 < \infty$ ). Considere os processos de colchetes angulares do  $M$ . Qual é a relação dele com a sequência  $\{\sigma_i^2\}_{i=1}^{\infty}$ ?

(b) Prove que em geral, para qualquer martingale  $M$ , ocorre que

$$A_n - A_{n-1} = \mathbb{E} [M_n^2 - M_{n-1}^2 | \mathcal{F}_{n-1}] = \mathbb{E} [(M_n - M_{n-1})^2 | \mathcal{F}_{n-1}] \quad (28)$$

**Exercício 6.** Leia a demonstração do Teorema da Seção 12.13 e verifique que você entende todas as suas passagens. Para a completude, estou apresentando abaixo a formulação do teorema.

**Teorema** que relaciona a convergência de martingal à finitude do limite de seu processo de colchetes angulares.

Seja  $M$  um martingal em  $\mathcal{L}^2$  e tal que  $M_0 \equiv 0$ . Seja  $A$  seu processo de colchetes angulares.

- (a)  $\lim_n \rightarrow \infty M_n(\omega)$  existe para quase todo  $\omega$  que satisfaz a relação  $A_\infty(\omega) < \infty$ .
- (b) Suponha que os incrementos do  $M$  são limitados uniformemente, quer dizer, suponha que existe um número  $K$  finito tal que

$$|M_n(\omega) - M_{n-1}(\omega)| \leq K, \text{ for all } n, \forall \omega$$

Então  $A_\infty(\omega) < \infty$  em quase todo  $\omega$  para o qual  $\lim_n \rightarrow \infty M_n(\omega)$  existe.

**Exercício 7.** O passeio aleatório simples simétrico  $S_0, S_1, S_2, \dots$  é um martingal que não converge (pois sabemos dos estudos de passeios aleatórios que este passeio particular não converge). Você consegue deduzir esta não convergência a partir do Teorema formulado no Exc. 6?

**Uma pequena dúvida.** Aliás, como você demonstraria rigorosamente que  $S$  citado no Exc. 7 não converge sem o uso do Teorema formulado no Exc. 6?

**Exc. 8.** Leia a formulação do Teorema da Seção 12.15 do livro de Williams. Para a completude, estou apresentando abaixo a formulação do teorema.

**Teorema** de Lévy sobre a extensão do Lema de Borel-Cantelli.

Seja  $\{\mathcal{F}_n\}_{n=0}^\infty$  uma filtração num espaço de probabilidades  $(\Omega, \mathbb{P}, \mathcal{F})$ . Considere uma sequência de eventos  $\{E_n\}_{n=1}^\infty$  tais que  $E_n \in \mathcal{F}_n$  para cada  $n = 1, 2, \dots$ , e define

$$Z_n := \sum_{1 \leq k \leq n} \mathbb{I}_{E_k} = \text{o número de } E_k (k \leq n) \text{ que ocorreram} \quad (29)$$

$$\xi_k := \mathbb{P}[E_k | \mathcal{F}_{k-1}] \quad (30)$$

$$Y_n := \sum_{1 \leq k \leq n} \xi_k \quad (31)$$

Então valem as duas seguintes implicações (com a reserva “quase certamente”):

- (a)  $(Y_\infty < \infty) \Rightarrow (Z_\infty < \infty)$ ,
- (b)  $(Y_\infty = \infty) \Rightarrow (Z_n/Y_n \rightarrow 1)$ .

(a) Sertifique-se que você entende por que a indexação da filtração começa com 0 enquanto que a da sequência  $E$  começa com 1, e que você entende também o significado das notações  $Y_\infty$  e  $Z_\infty$ .

(b) Já que o teorema supraformulado foi chamado de “extensão do Lema de Borel-Cantelli”, então espera-se que o lema pode ser derivado como o caso particular do teorema. Faça isto!

(c) Leia a demonstração do teorema, mas saiba que no texto abaixo pretendo demonstrar um resultado semelhante, mas de acordo com o argumento do livro de Durrett. Curioso, mas a demonstração seguindo Durrett é mais comprida que a que vimos no livro de Williams. Talvez, por causa desta curiosidade já é valioso ler a demonstração do Williams.

**Exercício 9.** A conclusão do teorema formulado no Exc. 8 motiva a pergunta do presente exercício: Construa um processo  $(S_n)$  tal que o processo  $(A_n)$  da decomposição de Doob  $S_n = S_0 + A_n + N_n$  possua a seguinte propriedade: existe  $\omega \in \Omega$  tal que  $A_n(\omega) < \infty$  assim como existe  $\omega \in \Omega$  tal que  $A_n(\omega) = \infty$ .

**Demonstração do Teorema de Lévy que generaliza o Lema de Borel-Cantelli.** Recorde que largamos o tratamento da Urna de Friedman deixando um buraco na demonstração do resultado sobre seu comportamento. Este resultado baseia-se nos resultados a serem apresentados agora.

**Teorema** que se encontra na Seção 4 do capítulo do livro de Durrett sobre martingais.

Seja  $\{X\}$  um martingal em  $L^2$  (recordo, isto significa que  $E[X_n^2] < \infty$  para cada  $n$ ), e tal que  $X_0 \equiv 0$ . Seja  $f : \mathbb{R} \rightarrow \mathbb{R}$ , função real crescente para  $\infty$ , maior que 1, e tal que

$$\int_0^\infty \frac{dt}{f^2(t)} < \infty \quad (32)$$

Então, no conjunto  $\{\omega; A_\infty(\omega) = \infty\}$  ocorre, quase certamente, que  $\frac{X_n(\omega)}{f(A_n(\omega))} \rightarrow 0$  conforme  $n \rightarrow \infty$ .

**Demonstração.** Recordo que  $A$  significa o processo de colchetes angulares para o martingal  $X$ . Este  $A$  é previsível e

$$A_{n+1} - A_n = E[(X_{n+1} - X_n)^2 | \mathcal{F}_n] \quad (33)$$

Como  $A$  é previsível, assim será também o processo  $H$  definido via  $H_m := 1/(f(A_m))$ . Deviso a esta previsibilidade, a transformação do martingal  $X$  pelo processo  $H$  é um martingal. Denotamos este por  $Y$ , isto é:

$$Y_n := (H \circ X)_n = \sum_{m=1}^n \frac{X_m - X_{m-1}}{f(A_m)} \quad (34)$$

(a primeira igualdade é a definição e a segunda pode ser verificada diretamente). Denotamos por  $B$  o processo de colchetes angulares para o martingal  $Y$ . Para  $B$ , temos que

$$B_{n+1} - B_n = E[(Y_{n+1} - Y_n)^2 | \mathcal{F}_n] = E\left[\frac{(X_{n+1} - X_n)^2}{f^2(A_{n+1})} | \mathcal{F}_n\right] = \frac{A_{n+1} - A_n}{f^2(A_{n+1})}$$

(a primeira igualdade é a propriedade genérica (cf. (33)), e a segunda segue de (34)).

Agora, como  $\{A_n(\omega)\}_{n=1}^\infty$  é sequência positiva crescente então (observe, para a validade da conta abaixo, é suficiente que a sequência seja não decrescente, e também, a conta usa  $A_0(\omega)$  que é 0 por decreto, e este 0 acrescido não quebra a monotonicidade de  $\{A_n(\omega)\}$ )

$$\sum_{n=1}^\infty \frac{A_{n+1} - A_n}{f^2(A_{n+1})} \leq \sum_{n=0}^\infty \int_{(A_n, A_{n+1})} \frac{dt}{f^2(t)} < \infty \quad (35)$$

Observe que a estimativa " $< \infty$ " veio da suposição (32) e é válida tanto no caso  $A_n(\omega) \rightarrow \infty$  quanto no caso  $A_n \rightarrow \text{algo finito}$ . A estimativa (35) significa que  $\{\omega : B_\infty < \infty\} = \Omega$ , e portanto o martingal  $Y$  converge em todo  $\Omega$  para uma variável aleatória finita:

$$Y_n(\omega) = \frac{X_1(\omega) - X_0(\omega)}{f(A_1(\omega))} + \frac{X_2(\omega) - X_1(\omega)}{f(A_2(\omega))} + \cdots + \frac{X_n(\omega) - X_{n-1}(\omega)}{f(A_n(\omega))} \rightarrow Y_\infty(\omega) \quad (36)$$

onde  $Y_\infty(\omega)$  é algo finito. Agora tomo  $\omega$  tal que  $A_n(\omega) \uparrow \infty$  (cuidado:  $A_n$  é sempre  $\uparrow$ ; eu tomo  $\omega$  onde ela  $\uparrow \infty$ ). Por causa da suposição imposta na  $f$  (que  $f \uparrow \infty$ ) tem-se que  $f(A_n(\omega)) \uparrow \infty$ . Então, o Lema de Kronecker, sendo aplicado ao (36), conclui que

$$\frac{(X_1(\omega) - X_0(\omega)) + (X_2(\omega) - X_1(\omega)) + \cdots + (X_n(\omega) - X_{n-1}(\omega))}{f(A_n(\omega))} = \frac{X_n(\omega)}{f(A_n(\omega))} \rightarrow 0 \text{ com } n \rightarrow \infty$$

o que fecha a demonstração. Só, para a completude, vale citar o lema usado:

$$\begin{aligned} \text{se} \quad & a_n \uparrow \infty \text{ e } \left( \frac{x_1}{a_1} + \frac{x_2}{a_2} + \cdots + \frac{x_n}{a_n} \right) \text{ converge} \\ \text{então} \quad & \frac{x_1 + \cdots + x_n}{a_n} \rightarrow 0 \text{ as } n \rightarrow \infty \end{aligned}$$

**Fechando o tratamento da Urna de Friedman com uso de resultados derivados para martingais.** Como bem se lembra, nos faltou a demostrar que

$$\frac{D_n}{\sum_{m=1}^n g_{m-1}} \rightarrow 1 \text{ quase certamente conforme } n \rightarrow \infty \quad (37)$$

onde

por  $B_n$  o evento que a  $n$ -ésima retirada resultou em bola verde  
por  $D_n$  o número de bolas verdes adicionadas nos  $n$  primeiros passos

notações que nos permitem reescrever a tarefa da seguinte maneira:

$$\text{provar que } \frac{\mathbb{I}_{B(1)} + \mathbb{I}_{B(2)} + \cdots + \mathbb{I}_{B(n)}}{g_0 + g_1 + \cdots + g_{n-1}} \rightarrow 1 \quad (38)$$

onde cada evento  $B_n$  é  $\mathcal{F}_n$ -mensurável,  $(39)$

$$\text{e } g_{k-1} = \mathbb{P}[B_k | \mathcal{F}_{k-1}] \quad (40)$$

$$\text{e } \sum_{k=1}^{\infty} g_k = \infty \text{ quase certamente} \quad (41)$$

Acima,  $\{\mathcal{F}_n\}$  é a filtracão gerada pelo processo, é, naturalmente,  $B_k \in \mathcal{F}_k$ . Também, segue-se da definição da Urna de Friedman, que a proporção de bolas verdes na urna após  $k-1$ -ésimo passo ser completado é a probabilidade de retirar bola verde no passo  $k$  (recorda, um passo é duas ações: a retirada de uma bola com a devolução da mesma e o acrescimo de bolas verdes e vermelhas de acordo com a cor da bola retirada). Por fim,

recordando que o número total de bolas acrescidas em cada passo é sempre  $a+b$ , conclui-se que  $g_n \geq (R_0 + G_0 + n(a+b))^{-1}$  (recorda,  $R_0$  e  $G_0$  são os números de bolas vermelhas e verdes, respectivamente, no início na urna. Esta estimativa garante a propriedade (41).

**Afirmção** que fecha os argumentos: (39), (40), (41)  $\Rightarrow$  (38).

**Demonstração.** Definimos o martingal  $X$  via:  $X_0 \equiv 0$ , e  $X_n - X_{n-1} := \mathbb{I}_{B_n} - P[B_n | \mathcal{F}_{n-1}]$ , para  $n \geq 1$ . Com esta definição, tem-se:

$$\frac{\mathbb{I}_{B(1)} + \mathbb{I}_{B(2)} + \cdots + \mathbb{I}_{B(n)}}{g_0 + g_1 + \cdots + g_{n-1}} - 1 = \frac{X_n}{g_0 + g_1 + \cdots + g_{n-1}} \quad (42)$$

igualdade que explica o por que nossa tarefa torna a ser a demonstração da convergência de  $X_n/(g_0 + g_1 + \cdots + g_{n-1})$  ao 0. Estabelecer esta convergência nos ajudará o processo dos colchetes angulares do martingal  $\{X\}$ . Tal processo de colchetes pode apresentar dois comportamentos: ou  $A_\infty < \infty$  ou  $A_\infty = \infty$ . Daí, são dois tratamentos distintos. Primeiramente, consideramos o que acontece no conjunto  $\{\omega : A_\infty(\omega) < \infty\}$ . Aqui a situação é simples, pois neste conjunto o martingal converge a algo finito (isto é,  $X_\infty < \infty$ ), enquanto que  $\sum_{i=1}^n g_{i-1} \rightarrow \infty$  de acordo com o pressuposto (41). Assim, conclui-se que a fração  $X_n/(g_0 + g_1 + \cdots + g_{n-1})$  vai ao 0. No segundo caso, consideramos o que acontece no conjunto  $\{\omega : A_\infty(\omega) = \infty\}$ . Aplicando o “Teorema que se encontra na Seção 4 do capítulo do livro de Durrett sobre martingais”, com  $f(x) = \max\{1, x\}$  (aqui 1 mata os pequenos valores de  $x$ , que são insignificantes para nossa necessidade mas atrapalham a finitude da integral  $\int_0^\infty x^{-2} dx$ ), concluimos que no conjunto de  $\omega$ 's por hora considerado acontece o seguinte:

$$\frac{X_n}{A_n} \rightarrow 0 \quad (43)$$

Esta conclusão ainda não é o que precisamos, pois, infelizmente, o denominador ainda não é  $g_0 + g_1 + \cdots + g_{n-1}$ . Mas vai ser já-já. Para tal, só precisa lembrar da fórmula principal que vincula os incrementos do processo de colchetes angulares com os do martingal a partir do qual os colchetes surgiram:

$$\begin{aligned} A_n - A_{n-1} &= \mathbb{E}[(X_n - X_{n-1})^2 | \mathcal{F}_{n-1}] \\ &= \mathbb{E}[(\mathbb{I}_{B_n} - g_{n-1})^2 | \mathcal{F}_{n-1}] = g_{n-1} - g_{n-1}^2 \leq g_{n-1} \end{aligned}$$

o que garante que  $A_n \leq g_0 + g_1 + \cdots + g_{n-1}$ , e, junto com (43) nos dá a desejada convergência  $X_n/(g_0 + g_1 + \cdots + g_{n-1}) \rightarrow 0$ .