CONSTRUÇÃO DE UM PROCESSO FORTE DE MARKOV

Dissertação apresentada em junho de 197[illegible] ao Instituto de Matemática e Estatística da U.S.P., para obtenção do Mestrado em Estatística Teórica e Probabilidade, pelo candidato Jefferson Antonio Galves..........

INTRODUÇÃO

Um processo forte de Markov é um processo markoviano que, adicionalmente, tem a propriedade forte de Markov. Duas questões, portanto: primeiramente constrùir um processo de Markov; depois caracterizá-lo como fortemente markoviano.

O primeiro problema, é tratado no §2; lá, utilizando a Teorema da Extensão de Daniell, nós construímos um processo de Markov, a partir de um semigrupo de operadores de transição.

Na primeira metade do §3, nós mostramos que uma condição suficiente para um processo de Markov ser forte é que êle tenha trajetórias contínuas à direita e seja realização de um semigrupo de Feller.

Na segunda metade do §3, nós apresentamos o Movimento Browniano, como exemplo de um processo cuja família de operadores de transição é de Feller. Em seguida, utilizando a teoria dos martingais, nós mostramos que o Movimento Browniano admite uma versão com trajetórias contínuas. Isso o caracteriza como processo forte de Markov.

O §1 é um resumo de análise, inspirado no capítulo IV, volume II, do texto clássico de Feller. Sendo algo mais longo do que o estritamente necessário para a sequência, êle é dedicado àqueles que, compulsòriamente, tem lido o Dixmier.

No final de cada parágrafo há notas e referências.

Para a redação dos parágrafos 2 e 3, consultamos e baseamo-nos nos seguintes textos:-

Blumenthal, R.M. - "An extended Markov property". Trans. Amer. M. Soc., 85, 1957, p57-72.

Hunt, G.A. - "Martingales et Processus de Markov" Monographies de la Société Mathématique de France.

Meyer, P.A. - "Probabilités et potentiel" Publications de l'Institut de Math. de l'Univ de Strasbourg

Meyer, P.A. - Processus de Markov". Lectura Notes im Math.Springer-Verlag.

Nosso roteiro geral foi fornecido pelas notas de aula e seminário, escritas pelo Dr. Hans Föllmer, durante a sua estada no I.M.E.

Para finalizar uma primeira errata: na 3a. linha da página 1, por êrro de datilografia, o símbolo $\leq$ foi substituído pelo símbolo $<$; aparentemente, essa substituição datilográfica ocorre com uma certa frequência ao longo do §1. Houve, também, um salto na numeração das páginas, de sorte que, a página 19 simplesmente não existe.

Í N D I C E

****

" Os martingais e as esperanças condicionais se situam, por sua generalidade, entre os instrumentos de base da Análise "

(Hunt)

§1 - Preliminares de Análise

A . Funções de Baire.

Seja Ω um conjunto qualquer; indicaremos por R^{Ω} o conjunto das funções de Ω em R. Há uma ordem natural para R^{Ω}, definida como se segue: $f \leq g$, se $f(\omega) \leq g(\omega)$, em ponto ω de Ω, sendo f e g dois elementos quaisquer de R^{Ω}. Por essa ordem, pares quaisquer de funções sempre possuam supremo e ínfimo (em R^{Ω}), assim definidos :

$$\sup (f,g) : \omega \in \Omega \to \sup (f(\omega), g(\omega)) \in R,$$

$$\inf (f, g) : \omega \in \Omega \to \inf (f(\omega), g(\omega)) \in R.$$

Há, também, uma adição e um produto por escalar, definidos de maneira natural sôbre R^{Ω}, como segue:

$$f + g : \omega \in \Omega \to f(\omega) + g(\omega) \in R$$

$$\lambda f : \omega \in \Omega \to \lambda(f(\omega)) \in R$$, sendo f e g elementos quaisquer de R^{Ω} e λ sendo um número real qualquer.

Seja, agora, S um subconjunto de R^{Ω} . Diremos que S é um reticulado vetorial, com respeito às operações de adição, produto por escalar, supremo e ínfimo, definidas acima sôbre R^{Ω} , se :(1.1) com relação às operações de adição e produto por escalar, S é um espaço vetorial sôbre R; (1.2) com relação às operações de sup e inf, definidas sôbre R^{Ω} , S é fechado.

Como $|\alpha| = \sup(\alpha,o) - \inf(\alpha,o)$, α sendo um real qualquer,(1.1) e (1.2) acarretam:

(1-3) Se f é um elemento de S , então a aplicação

$$|f|: \omega \in \Omega \to |f(\omega)| \in R$$

também é .

Seja, então, S um reticulado vetorial de funções numéricas, definidas sôbre Ω.

<u>Definição</u> : Um elemento de R^{Ω} será chamado de <u>função de Baire</u>, relativamente a S, se êle fôr o limite, ponto a ponto, de uma sequência de elementos de S .

É imediato que o conjunto das funções de Baire é também um reticulado vetorial.

Se A é um subconjunto de Ω , usaremos a notação I_A, para indicar a função indicadora de A, isto é, a função assumindo o valor 1, nos pontos de A, e o valor 0 nos pontos do complementar de A.

Indicaremos o complementar de A, por A^c.

<u>Definição</u> : Se A é um subconjunto de Ω , tal que I_A é uma função de Baire, então A será chamado <u>conjunto de Baire</u>.

Usaremos a notação $Ba\ (S)$, para indicar a classe de todos os conjuntos de Baire.

Se Ω é um espaço topológico, então o reticulado vetorial empregado será geralmente $C\ (\Omega)$, o conjunto das funções de Ω em R. Nesse caso, escreveremos também $Ba\ (\Omega)$, ou simplesmente Ba, se não houver perigo de confusão.

(1.4) <u>Proposição</u> : $Ba\ (S)$ é estável sob operações de reunião e intersecção enumeráveis. Se I_{Ω} é uma função de Baire, então $Ba\ (S)$ é um σ - álgebra.

As afirmações são de verificação imediata, se nós lembramos de que

$$I_{\cup_n A_n} = \sup_n I_{A_n} = \lim_n \sup_{1<i<n} I_{A_i}$$

$$\text{e } I_{A^c} = 1 - I_A = I_{\Omega} - I_A.$$

Se f é uma função de Ω em R e se **B** é um subconjunto de R, usaremos a notação $\{f \in \mathbf{B}\}$, para indicar o subconjunto $f^{-1}(\mathbf{B})$ de Ω .

Em geral, escrevemos :

$\{ f < r \}$, para indicar o conjunto $\{ \omega \in \Omega \mid f(\omega) < r \}$

$\{ f < g \}$, para indicar o conjunto $\{ \omega \in \Omega \mid f(\omega) < g(\omega) \}$, etc.

(1.5) Proposição : Uma função f é de Baire, se e só se, para todo número real r, o conjunto $\{f \leq r\}$ fôr de Baire.

Demonstração : Se f é de Baire, então as funções $\sup(f, r + \frac{1}{n})$ e $\sup(f, r)$ também são de Baire, para todo n. Logo $\{f \leq r\}$ é um conjunto de Baire, pois

$$I_{\{f \leq r\}} = \lim_{\substack{n \to \infty \\ n \in N}} \frac{\sup(f, r + \frac{1}{n}) - \sup(f, r)}{\frac{1}{n}}$$

Reciprocamente, se $I_{\{f \leq r\}}$, para todo r, então as funções $I_{\{s<f \leq r\}} = I_{\{f \leq r\}} - I_{\{f \leq s\}}$ e

$$\phi_n := -nI_{\{f \leq -n\}} + \sum_{\kappa=1}^{2n-1} \left(-n + \kappa \frac{n}{2^n - 1}\right) I_{\{-n + \frac{(\kappa-1)n}{2^{n-1}} < f \leq -n + \frac{\kappa n}{2^{n-1}}\}} +$$

$$+ n I_{\{f > n\}}$$

também serão de Baire, qualquer que o natural u considerado. Donde f será de Baire, já que $f = \lim_{n \to \infty} \phi_n$ (o limite sendo tomado ponto a ponto).

Definição : Se Ω é um espaço topológico, a sua σ-álgebra de Borel, que nós indicaremos por $B(\Omega)$, é a menor σ-álgebra contendo todos os abertos do espaço.

Podemos, agora, enunciar uma consequência da proposição (1.5)

(1.6) Corolário : Sendo I_Ω uma função de Baire, então um elemento f de R^Ω é de Baire, se e só se, para todo subconjunto B, de Borel, de R, o conjunto $\{f \in B\}$ fôr de Baire,

Dos termos do corolário, nós tiramos a seguinte definição:

Definição : Seja Ω um conjunto qualquer, munido de σ-álgebra A; seja F um espaço topológico.

Uma aplicação

$$f : \Omega \to F$$

será chamada de mensurável - A se $\{f \in B\}$ pertencer a A, para todo conjunto B em $B(F)$

Então, o que o corolário(1.6) está dizendo é que as funções de Baire são mensuráveis - $\mathcal{B}_a$

Mais precisamente, $\mathcal{B}a$ é a menor σ-álgebra tornando tôdas as funções do reticulado original S, mensuráveis.

(1.7) <u>Proposição</u> : Se Ω é um espaço topológico, então

$$\mathcal{B}_a(C(\Omega)) \text{ está contido em } \mathcal{B}(\Omega)$$

<u>Demonstração</u>: Seja B um subconjunto aberto, qualquer de R. Se f é uma função contínua de Ω em R, então $\{ f \in B \}$ é, por definição, um subconjunto aberto de Ω e, portanto, pertencente a $\mathcal{B}(\Omega)$.

Ora, por definição, os abertos geram $\mathcal{B}$(R), donde a pertinência de $\{f \in B\}$ a $\mathcal{B}(\Omega)$ permanece verdadeira,se B é um boreliano qualquer de R(2).

Mas, então, tôda função contínua de Ω em R, é mensurável- $\mathcal{B}(\Omega)$.Como $\mathcal{B}_a(C(\Omega))$ é a menor σ-álgebra tornando mensuráveis todos os elementos do reticulado $C(\Omega)$, segue que $\mathcal{B}_a(C(\Omega))$ está contido em $\mathcal{B}(\Omega)$.

(1.8) <u>Proposição</u>: Se Ω é um espaço métrico, então : $\mathcal{B}_a(C(\Omega)) = \mathcal{B}(\Omega)$

<u>Demonstração</u> : Seja A um subconjunto aberto de Ω. Para todo natural n,definimos $A_n = \bigcup_{\omega \in A} B(\omega, \frac{1}{n})$, sendo $B(\omega, \frac{1}{n})$ a bola aberta de centro ω e raio $\frac{1}{n}$.

Pelo lema de Urysohn (3)existe uma função real contínua ϕ_n, definida sôbre Ω, assumindo valor 1 em A e valor 0 em A_n^c.

Ora, $I_A = \lim_{n\to\infty} \phi_n$, donde I_A é uma função de Baire.

Em outras palavras, $\mathcal{B}_a(C(\Omega))$ contém todos os subconjuntos abertos de Ω e, portanto, $\mathcal{B}_a(C(\Omega))$ **contém** $\mathcal{B}(\Omega)$, já que $\mathcal{B}(\Omega)$, por definição, é a menor σ-álgebra contendo todos os abertos de Ω.

B. <u>Integral de Daniell</u>

Outra vez, sejam Ω um conjunto qualquer e S um reticulado vetorial de funções numéricas definidas sôbre Ω.

Seja, agora, uma aplicação $\mu : S \to R$, satisfazendo as condições :

(1.9) μ é linear

(1.10) μ é positiva, isto é, se f é um elemento positivo de S, então $\mu(f) > 0$.

Daqui por diante, nós usaremos indiferentemente as notações $\mu(f)$ e $\int_\Omega f(\omega)\,\mu(d\omega)$ para indicar a imagem de f, pela aplicação μ, sendo f um elemento de S.

DEFINIÇÃO : Um subconjunto A, de Ω, será dito desprezível, se existir uma sequência $(f_n)_n$, monótona, de elementos de S, tendendo para ∞ nos pontos de A e tal que a sequência $(\mu(f_n))_n$ limitada.

Diremos que uma propriedade relativa a Ω vale "quase certamente", se ela fôr verdadeira exceto para um subconjunto desprezível de pontos de Ω. Assim, por exemplo, escreveremos que $\lim_{n\to\infty} f_n \overset{qc}{=} f$, se o conjunto $\{f_n \nrightarrow f\}$ fôr desprezível ("qc" será usado como abreviatura de "quase certamente").

A partir da aplicação μ, é possível definirmos uma semi-norma $\|\cdot\|_1$, sôbre S, da seguinte maneira:

se f é um elemento de S, $\|f\|_1 := \mu(|f|)$

É fácil verificar que $\|\cdot\|_1$ é, efetivamente, uma semi-norma.

Seja, agora, uma sequência $(f_n)_n$, de elementos de S. Diremos que $(f_n)_n$ é de Cauchy, se

$$\lim_{m,n\to\infty} \|f_n - f_m\|_1 = 0$$

Munidos dessas definições, é possível caracterizarmos dois subconjuntos fundamentais de R^Ω.

DEFINIÇÃO : Um elemento f de R^Ω, pertencerá a $L^1(S, \mu)$, se existir uma sequência $(f_n)_n$, de Cauchy, de elementos de S, convergindo quase que certamente para f.

Um elemento f de R^Ω, pertencerá a $L(S, \mu)$ se existir uma sequência $(f_n)_n$, qualquer, de elementos de S, convergindo quase certamente para f.

Chamaremos as funções de $L^1(S,\mu)$ de integráveis-μ; as de $L(S,\mu)$ serão chamadas de mensuráveis-μ.

A questão fundamental, em Teoria da Integração, é a extensão da integral μ, de S a L^1. Essa questão é respondida completamente pelo seguinte teorema fundamental:

(1.11) Teorema de Daniell : Se μ é uma aplicação linear e positiva, sôbre um reticulado vetorial S e, se além disso:

(1.12) $\mu(f_n) \to 0$, sempre que $(f_n)_n$ fôr uma sequência decrescente de elementos de S, convergindo quase certamente pafa 0;

então : i) μ admite uma extensão a L^1;

ii) em L^1 estão tôdas as funções que são limites quase certamente de sequências de Cauchy de elementos de L^1. (Em outras palavras, o processo de extensão que levou a L^1, encerra-se nele (4)).

Definição : Um par (S,μ) formado por um reticulado vetorial, mais uma aplicação linear positiva, satisfazendo (1.12), será por nós chamado de Sistema de Daniell.

Em particular se (S,μ) fôr um Sistema de Daniell satisfazendo ainda as condições:

(1.13) S contém as constantes

(1.14) $\mu(I_\Omega) = 1$,

nós o chamaremos de "esbôço de espaço de probabilidades". Nesse caso, nós usaremos a letra E, para indicar a integral.

Nesse último caso, chamaremos $L^1(S,E)$ de espaço de probabilidades, chamaremos os elementos de $L(S,E)$ de variáveis aleatórias, e a integral de um elemento f, de L^1, será chamada a "esperança da variável aleatória f". Finalmente, o conjunto Ω, domínio das variáveis aleatórias, será chamado de "espaço amostral".

Vamos agora, estabelecer ligação entre a classe das funções de Baire e a classe das funções integráveis e mensuráveis-μ.

As seguintes afirmações são de verificação imediata.

(1.15) Tôda função mensurável -μ é igual, quase certamente, a uma função de Baire.

Definição : $\mathcal{A}_\mu = \{ A \in \Omega | I_A$ é mensurável-μ }

(1.16) $\mathcal{A}_\mu$ é uma σ - álgebra..

Definição: Seja F uma σ-álgebra de partes de Ω. Nós diremos que F é completa, com respeito a μ, se contém todos os subconjuntos de Ω, desprezíveis com respeito a μ. Se não é completa, nós indicaremos por $\overline{F}^\mu$ a σ- álgebra obtida,"completando-se" F,isto é, a menor σ-álgebra contendo F e os conjuntos desprezíveis.

(1.17) $\mathcal{A}_\mu$ é completa;$\mathcal{A}_\mu \supset \mathcal{B}_a(S)$;$\mathcal{A}_\mu = \overline{\mathcal{B}_a(S)}^\mu$

(1.18) Tôda função mensurável -μ é mensurável-$\mathcal{A}_\mu$ e recìprocamente.

Chamaremos os elementos de $\mathcal{A}_\mu$ de conjunto mensuráveis-μ. No caso de um espaço de probabilidades, êles serão chamados simplesmente de eventos. Noto que o espaço Ω é necessàriamente um evento, em virtude de (1.13); é o evento certo.

Se f é uma função mensurável-μ,positiva, mas não integrável, nós diremos, por convenção, que o valor de μ em f é $+\infty$. Com essa convenção, é possível, a partir da integral, definirmos uma "medida" sôbre os elementos de $\mathcal{A}_\mu$; da seguinte maneira:

$$A \in \mathcal{A}_\mu \longmapsto \int_\Omega I_A(\omega)\mu(d\omega)$$

Como é habitual, nós indicaremos a integral e a medida derivada dela, pela mesma letra; escreveremos, portanto, $\mu(A) = \int I_A(\omega)\ \mu\ (d\omega)$. No caso de um espaço de probabilidades, porém, por razões históricas,a medida associada à integral será chamada de probabilidade e indicada pela letra P.

C. Esperança Condicional

Definição : Um subconjunto H de um espaço de probabilidades $L^1(S,E)$ será chamado de subespaço de probabilidade se êle fôr um sub-reticulado vetorial de L^1,completo com respeito à semi-norma $\| \ \|_1$, contendo as constantes.

Se G é uma σ-álgebra contida em $\mathcal{A}_E$, nós usaremos a notação $L^1(\Omega,G,P)$ para indicar o conjunto dos elementos de $L^1(S,E)$ que são mensuráveis-G.É fácil ver que $L^1(\Omega,G,P)$ é um subespaço de probabilidade de $L^1(S,E)$.

Note que $L^1(\Omega, A_E, P) = L^1(S, E)$

Se $(f_j)_{j\varepsilon J}$ é uma família de elementos de $L^1(S, E)$, indicamos com $r(f_j)_{j\varepsilon J}$ o menor reticulado vetorial contendo as funções f_j e as constantes. É claro, então, que $(r(f_j)_{j\varepsilon J}, E)$ é um esboço de espaço de probabilidades e o espaço construido a partir dêle é um subespaço de probabilidade de $L^1(S, E)$.

Note que $L^1(r(f_j)_j, E) = L^1(\Omega, Ba(r(f_j)_{j\varepsilon J}), P)$.

Um subespaço de probabilidade é um espaço de probabilidade. Possui, então, suas próprias funções mensuráveis. Nós as chamaremos de <u>variáveis aleatórias associadas</u> ao subespaço, para identificá-las no meio de tôdas as variáveis aleatórias do espaço maior de probabilidades considerado. Assim, por exemplo, se G é uma sub σ-álgebra de A_E, as variáveis aleatórias associadas ao subespaço $L^1(\Omega, G, P)$ são os elementos de $L(S, E)$ que são mensuráveis-G.

Um evento é dito associado a um subespaço de probabilidades, se a sua função indicadora fôr associada ao subespaço.

<u>Definição</u>: Se f é um elemento de L^1 e H é um sub-espaço de probabilidade de L^1, nós chamaremos de <u>esperança condicional</u> de f, dado H, a qualquer elemento de g de H tal que

(1.19) $E(hf) = E(hg)$, para tôda variável aleatória limitada h, associada a H.

As diversas versões da esperança condicional de f, dado H são iguais quase certamente (isso é uma decorrência da definição). Qualquer das versões será indicada por $E(f/H)$.

Se $H = L^1(\Omega, G, P)$, sendo G uma sub σ-álgebra de A_E, nós escreveremos, simplesmente, $E(f/G)$.

Se $H = L^1(r(f_j)_{j\varepsilon J}, E)$, sendo $(f_j)_{j\varepsilon J}$ uma família de elementos de $L^1(S, E)$, nós escreveremos simplesmente $E(f/(f_j)_{j\varepsilon J})$.

(1.20) Teorema . Dados f e H, existe $E(f|H)$.

Demonstração: Se $f \in L^2(S,E)$, tomamos $E(f|H)$ igual, quase certamente, à projeção ortogonal de f no subespaço fechado $L^2 \cap H$ (existência da projeção ortogonal estando assegurada pelo Teorema da Projeção Ortogonal em espaços de Hilbert(5)).

Noto que se h é um elemento limitado de H, então h é um elemento, na verdade, de $H \cap L^2$, de sorte que a projeção ortogonal de f em $H \cap L^2$ satisfaz completamente a condição (1.19).

Seja, agora, f um elemento positivo qualquer de L^1. Para todo número natural n, definimos $f_n = \inf(f,n)$. Assim, definido, f_n é um elemento de L^2, estando assegurado a existência de $E(f_n|H)$.

Nossa intenção é mostrar que a sequência $(E(f_n|H))_n$ converge quase certamente, seu limite sendo, precisamente, $E(f|H)$.

Ora, pela forma como definimos a esperança condicional, mais o fato de E ser positivo, faz com que a monotonocidade crescente de $(f_n)_n$, acarrete a de $(E(E(f_n|H)))_n$.

Além disso, $E(E(f_n|H)) \overset{\text{def}}{=} E(f_n) < E(f)$

Isto é, $(E(E(f_n|H)))_n$ é uma sequência monótona, crescente, limitada de números reais, logo, de Cauchy, Então, supondo $n > m$:

$$|| E(f_n|H) - E(f_m|H)||_1 := E||E(f_n|H) - E(f_m|H)| =$$

$$= E(E(f_n|H) - E(f_m|H)) = E(E(f_n|H)) - E(E(f_m|H)) =$$

$$= |E(E(f_m|H)) - E(E(f_n| H))| \to 0,$$ quando n e m tendem a infinito.

Sendo, como acabamos de ver, uma sequência de Cauchy de elementos de L^1, $(E(f_n|H))_n$ contém uma subsequência convergindo quase certamente (6).

Sendo monótona , $(E(f_n|H))_n$ só pode admitir uma subsequência convergindo quase certamente, se ela própria convergir quase certamente.

Seja então, g o limite quase certo de $(E(f_n|H))$.

Pela segunda parte do Teorema de Daniell, não há dúvida de que g é um elemento de H.

Para testar(1.19), tomamos em elemento h, qualquer, limitado de H.

Pelo Teorema da Convergência Monótona(4), nós temos que:

$$E(hg) = E(h \lim_{n\to\infty} E(f_n|\mathcal{H})) = \lim_{n\to\infty} E(h\, E(f_n|\mathcal{H}))$$

Mas, pela definição de esperança condicional:

$$E(h\, E(f_n|\mathcal{H})) = E(h\, f_n).$$

Finalmente, o Teorema da Convergência Dominada(5), assegura

$$\lim_{n\to\infty} E(h\, f_n) = E(h \lim_{n\to\infty} f_n) = E(hf)$$

Encadeando-se tôdas as igualdades, obtemos realizada a condição (1.19) : $E(hg) = E(hf)$, sendo h um elemento qualquer, limitado, de $\mathcal{H}$.

O caso em que f é um elemento, de sinal qualquer, de L^1 é reduzido ao anterior, pela decomposição habitual de f nas suas partes positiva e negativa.

D. Processos Estocásticos.

a) Definição : Um processo estocástico é um objeto formado por:

i) um espaço de probabilidades $L^1(S, E)$

ii) uma família $(X_j)_{j\in J}$ de funções, com domínio em Ω, espaço amostral do espaço de probabilidades, assumindo valores num espaço topológico F, cada uma delas sendo mensuráveis- $\mathcal{A}_E$.

Nós sempre tomaremos $J \subset [0, +\infty)$

O espaço topológico F será chamado de espaço de fase do processo.

Chamaremos de trajetória associada ao ponto ω do espaço amostral, à função:

$$traj(\omega) = j \in J \rightarrow X_j(\omega) \in F$$

Os pontos de J serão chamados de instantes de tempo, e para cada ω, o ponto $X_j(\omega)$ de F será chamado de estado do processo no instante j, correspondendo a ω.

Frequentemente, nós identificaremos ω com a trajetória que lhe é associada, tomando, portanto, o espaço amostral Ω como subconjunto de $F^{[0,\infty)}$. Nesse caso, Ω é um espaço topológico(munido da topologia produto), nós tomaremos como reticulado básico do espaço de probabilidade a classe $C(\Omega)$ e definiremos X_j como sendo a projeção na j-ésima coordenada de Ω .

Assim definida, X_j será um elemento de $C(\Omega,F)$, sendo portanto, mensurável-A_E.

Finalmente, se f é um elemento de $C(F)$, a composta $f(X_j)$ será um elemento do reticulado básico $C(\Omega)$, havendo, portanto, sentido em considerarmos a sua esperança $E(f(X_j))$.

Um processo dêsse tipo será chamado de "canônico". No §2, nós encontraremos um processo dêsse tipo.

Nos esforçaremos, doravante, em notar as aplicações de Ω em R com as letras gregas ϕ, ψ, etc.; as aplicações de Ω em F com as letras latinas maiúsculas X,Y,Z,etc e as aplicações de F em R, com as letras minúsculas latinas f,g,h,...

Seja $(L^1(S,E),(X_j)_{j\varepsilon J})$ um processo. Uma família crescente $(H_j)_{j\varepsilon J}$ de subespaços de probabilidade de L^1, será chamada de família de passados do processo, se, para tôda função f, de F em R, sendo mensurável -$B(F)$, e para todo instante j, a função composta $f(X_j)$ fôr uma variável aleatória associada a H_j.

b) Definição : Uma aplicação $T : \Omega \to J$, sendo Ω o espaço amostral de um espaço de probabilidades, será chamada tempo aleatório, com respeito à família$(H_j)_{j\,\varepsilon J}$ se, para todo instante j:

(1.21) o conjunto $\{T<j\}$, fôr um evento associado ao passado H_j.

exemplo. Seja $(L^1(S,E),(X_t)_{t\geq 0})$ um processo estocástico, tendo F como espaço de fase; seja $(H_t)_{t\geq 0}$ uma família de passados do processo.

Se A é um aberto de F, então a aplicação

$$T_A: \omega \in \Omega \to \inf\{s\geq 0 : X_s(\omega)\varepsilon A\}\}$$

é um tempo aleatório, com respeito a $(H_t)_{t\geq 0}$ (7),

De fato, para cada instante t fixado, nós temos:

$$\{T_A<T\} = \bigcup_{\substack{s<t\\ s\varepsilon Q}}\{X_s\varepsilon A\} = \bigcup_{\substack{s<t\\ s\varepsilon Q}}\{I_A(X_s) = 1\}$$

Mas, para cada instante s, o conjunto $\{I_A(X_s) = 1\}$é um evento associado a H_s e, portanto, associado a H_t, para t>s, já que a família$(H_j)_j$ é crescente.

Como a família dos eventos associados a H_t forma uma σ-álgebra, concluimos que a reunião enumerável $\bigcup_{\substack{s<t \\ s\in Q}}\{I_A(X_s) = 1\}$ é, também, um evento associado a H_t, o que encerra nossa verificação.

(1.23) <u>Proposição</u> : A aplicação T de Ω em $[0,+\infty)$ é um tempo aleatório, com respeito à família $(H_t)_{t\geq 0}$, se e sòmente se, a aplicação inf(T, t) fôr, para cada t, uma variável aleatória associada a H_t.
A demonstração é feita com técnica semelhante à que empregamos para demonstrar (1.5)

<u>Definição</u>: Seja T um tempo aleatório, com respeito à família $(H_t)_{t\geq 0}$.
Chamaremos de "passado até o instante aleatório T" e indicaremos por H_T, a classe das funções ϕ, integráveis, tais que, qualquer que seja o instante t considerado, a função $\phi I_{\{T< t\}}$ é um elemento de H_t.
É simples vèrificar que H_T é um subespaço de probabilidades.

Indicaremos por A_T a classe dos eventos associados ao subespaço H_T.

(1.24) <u>Proposição</u> : Um subconjunto A, de Ω, pertence a A_T se e só se, para todo instante $t \geq 0$, o conjunto $A \cap \{T< t\}$ fôr um evento associado a H_t.
Aqui não há o que demonstrar.
Não há dificuldade em se verificar que:
(1.25) $H_T = L^1(\Omega, A_T, P)$

c) Tempos aleatórios aparecem em conexão com processos estocásticos. Seja, então, um processo $(L^1(S,E)\ ,\ (X_t)_{t\geq 0})$, tendo o espaço topológico F como espaço de fase e munido de uma família de passados $(H_t)_{t\geq 0}$. Seja T um tempo aleatório com respeito à família $(H_t)_{t\geq 0}$.

Definimos a aplicação X_T como :

$$X_T : \omega \in \Omega \to X_{T(\omega)}(\omega) \in F$$

Para que a aplicação X_T funcione ela tem que ser uma variável aleatória associada a H_T.

A proposição seguinte, cuja demonstração se encontra no Meyer("Probabilités et...", IV - T47 e T49), nos dá uma condição suficiente para que isso aconteça:

(1.26) Proposição : Condição suficiente para que X_T seja uma variável aleatória associada a H_T é que as trajetórias do processo sejam contínuas à direita.

D. MARTINGAIS

Definição : Um processo $(L^1(S,E),(X_j)_{j\varepsilon J})$, tendo R como espaço de fase, será dito um martingal , com respeito à família de passados$(H_j)_{j\varepsilon J}$, se para todo par de instantes i e j, sendo i < j, nós tivermos:

(1.24) X_j é um elemento de H_j,

(1.25) $E(X_j|H_i)= X_i$

exemplo: Seja $L^1(S,E)$ um espaço de probabilidades; seja $(H_j)_{j\varepsilon J}$ uma família crescente qualquer de subespaços de L^1, finalmente seja X um elemento qualquer, fixado, de L^1.

Definimos $X_j := E(X|H_j)$, para todo j em J.

Decorre, então, da definição de esperança condicional, que o processo $(L^1(S, E), (X_j)_{j\varepsilon J})$ é um martingal.

Se em vez de (1.25), nós tivermos a condição mais fraca

(1.26) $E(X_j|H_i)\leq X_i$,

Nós diremos que o processo é um supermartingal

Martingais e supermartingais são as nossas ferramentas básicas no ataque ao problema de Dirichlet. A razão disso está no bom comportamento dos martingais; esse bom comportamento está resumido nos teoremas seguintes, cujas demonstrações podem ser encontradas no Meyer("Probabilités et..." Cap. VI).

Antes, porém, duas últimas definições:

Definição : Uma família de passados$(H_t)_{t\geq 0}$ é dita contínua se, para todo instante t:

$$H_t = \bigcap_{s>t} H_s$$

Definição : Dados dois processos $(L^1(S,E)$, $(X_t)_{t\geq 0})$ e $(L^1(S,E),(X'_t)_{t\geq 0})$, sôbre um mesmo espaço de probabilidades $L^1(S,E)$, nós diremos que um dêles é uma modificação do outro, se $X'_t = X_t$ q.c., para todo $t\geq 0$.

Seja, então, $(L^1(S,E),(X_t)_{t\geq 0})$ um supermartingal, com respeito à família $(H_t)_{t\geq 0}$.

(1.27) Teorema : Se $(H_t)_{t\geq 0}$ é contínua, então existe uma modificação $(L^1(S,E),(X'_t)_{t\geq 0})$ do processo original, com trajetórias contínuas e que é um supermartingal com respeito à família $(H_t)_{t\geq 0}$, se e só se a função

$$t \in |0, \infty) \rightarrow E(X_t) \in \mathbb{R}$$

é contínua à direita.

(1.28) Teorema da convergência :Se as trajetórias do supermartingal são contínuas à direita, e se $\sup E(X_t^-) < \infty$, então as variáveis aleatórias X_t convergem quase certamente para uma variável aleatória integrável X_∞, quando $t \rightarrow \infty$.

(1.29) Teorema da parada : Sejam S e T dois tempos de parada, com respeito à família $(H_t)_{t\geq 0}$ e tais que $S \leq T$. Se existir uma variável aleatória integrável Y , tal que,

$$E(Y \mid H_t) \leq X_t$$, para todo $t\geq 0$, então X_T e X_S são variáveis aleatórias integráveis e satisfazendo a desigualdade dos supermartingais

$$E(X_T \mid H_s) > X_s \quad \text{q.c.}$$

(1.30) Teorema : Seja $(L^1(S,E),(X^n_t)_{t\geq 0})$ $n\in N$ uuma sequência de martingais com respeito à família de passados $(H_t)_{t\geq 0}$, todos êles tendo trajetórias contínuas à direita e tal que, para cada t, a sequência $(X^n_t)n\in N$ é monótona crescente.

Para cada $t > 0$ e para cada ponto ω do espaço amostral, nós definimos

$$X_t(\omega) = \sup_n X^n_t(\omega)$$

Então as trajetórias de $(L^1(S,E),(X_t)_{t\geq 0})$ são quase certamente contínuas à direita e desprovidas de descontinuidades oscilatórias.

Notas e referências

O material dos ítens A e B foi baseado principalmente no capítulo IV, volume II, do livro de Feller, "Introduction to Probability Theory and its applications". O material sôbre a Integral de Daniell pode ser encontrado no texto clássico de Riesz e Nagy, "Leçons d'Analyse Fonctionelle". Em particular, nossas definições são baseadas numa variante do caminho de Daniell, sugerido por Bochner. Essa variante é muito bem exposta, no caso da integral de Lebesgue, no R^n, por Dixímier ("Integral de Lebesgue", por Diximier - na coleção de Notas de Matemática, do Impa).

(1)-(2) -Em ambos os casos, empregamos, sem mencionar, a seguinte proposição:

Seja M um conjunto, munido da σ-álgebra $\mathcal{M}$, seja f uma aplicação de M em N, seja $\mathcal{N}_0$ uma classe qualquer de subconjunto de N. Então, condição necessária e suficiente para que $\{f \in A\} \in \mathcal{M}$, para todo elemento A de $\mathcal{N}$, sendo $\mathcal{N}$ a σ-álgebra gerada por $\mathcal{N}_0$, é que $\{f \in A\} \in \mathcal{M}$, para todo conjunto A, pertencente a $\mathcal{N}_0$.

A demonstração decorre imediatamente do fato que

$$f^{-1}(\bigcup_n A_n) = \bigcup_n f^{-1}(A_n) \quad \text{e} \quad f^{-1}(A^c) = (f^{-1}(A))^c,$$

qualquer que seja a função f.

Em (1) nos valemos do fato da classe

$\{(-\infty, x) : x \in R\}$ gerar $\mathcal{B}$ (R)

(3) lema de Urysohn : Seja M um espaço topológico normal, e sejam A e B dois subespaços disjuntos e fechados de M.

Então existe uma função contínua, definida em M, tendo imagem contida em $[0,1]$ e tal que $f(A) = 0$ e $f(B) = 1$.

(4) É precisamente êsse fato que garante, a priori, a viabilidade dos teoremas limites, razão de ser da Teoria da Integração, desde Lebesgue.

Vale a pena, aliás, enunciá-los aqui e de uma vez por tôdas:

Teorema de Beppo- Levi (ou da Convergência Monótona): Se $(f_n)_n$ é uma sequência monótona de funções integráveis cujas integrais formam um conjunto limitado, então $(f_n)_n$ converge quase certamente para alguma função integrável f e

$$\int f = \lim_{n\to\infty} \int f_n$$

Lema de Fatou : Se $(f_n)_n$ é uma sequência de funções integráveis e positivas, então

$$\int \underline{\lim}\, f_n < \underline{\lim} \int f_n$$

TE
Teorema de Lebesgue (da convergência dominada) : Se $(f_n)_n$ é uma sequência de funções integráveis, convergindo quase certamente para f, e tal que $|f_n| \overset{qc}{\leq} g$, sendo g uma função integrável, então f é integrável e $\lim_{n\to\infty} \int f_n = \int f$.

(5) Teorema da Projeção Ortogonal : Seja H um espaço de Hilbert e seja F um subespaço fechado de H. Então a todo elemento x de H, corresponde um e um só elemento x_F de F, tal que, $(x - x_F | y) = 0$, qualquer que seja o elemento y de F.

O elemento x_F é chamado projeção ortogonal de x em F.

(6) Tôda sequência de Cauchy de elementos de L^1, admite uma subsequência convergindo quase certamente.

(7) Para que o exemplo funcione é preciso que o processo tenha trajetórias contínuas e que estas encontrem quase certamente o aberto A.

"No sentido amplo, a propriedade de Markov significa independência do futuro X_{s+u} do processo, do passado X_{s-u}, quando o presente X_s é conhecido. Sequências aleatórias $X_0, X_1, X_2 \ldots$ tendo essa propriedade fôram pesquisadas, pela primeira vez, por Markov em 1907"(Dynkin)

§2- Processos de Markov

A- Definição : Vamos, sem que isso nos traga qualquer limitação, supor que Ω é um espaço topológico.

Um processo estocástico $(L^1(C(\Omega),E),(X_t)_{t\geq 0})$, tendo F, como espaço de fase, será dito de Markov com respeito à família de passados $(H_t)_{t\geq 0}$ se:

(2.1) $E(f(X_t)|H_s) = E(f(X_t)|(g(X_s))_{g\varepsilon C(F)})$ para todo elemento f de $C(F)$ e para todo par de instantes s e t, com $s\leq t$.

De agora em diante, escreveremos a esperança condicional que aparece no segundo membro da expressão (2.1) simplesmente como $E(f(X_t)|\ X_s)$.

Isso não acarretará nenhuma ambiguidade, já que X_s sòzinho é o que distingue o subespaço $L^1((g(X_s))_{g\varepsilon C(F)},E)$

Como entender essa definição? Os subespaços H_s e $L^1((g(X_s)_{g\varepsilon C(F)},E)$ representam acervos de conhecimento: H_s é o conhecimento de tôda a história do processo, até o instante s, isto é, o instante presente; $L^1((g(X_s))_{g\varepsilon C(F)},E$ é o conhecimento da situação presente do processo. Uma esperança condicional é uma predição do futuro, feita a partir do conhecimento de resultados já occoridos no sistema. O que a igualdade (2.1) está dizendo, então, é que processos de Markov são aquêles nos quais uma predição sôbre o futuro do sistema é igualmente boa, quer se baseie no conhecimento de tôda a história do processo, quer se baseie no simples conhecimento da situação presente do processo.

Assim, a noção de processo de Markov generaliza a noção de processo determinístico.

B - Construção de um processo de Markov

Nosso ponto de partida é um espaço de fase F(isto é, um espaço topológico que, por razões técnicas, será também compacto e com base enumerável de abertos) e um semigrupo de operadores de transição $(P_t)_{t\geq 0}$.

Um operador de transição é um endoformismo de $C(F)$ (o espaço $C(F)$ estando munido de sua topologia uniforme habitual) tal que, se f é um elemento de $C(F)$ e t, um instante qualquer, então:

(2.2) $f>0$ acarreta $0<P_t f<||f||_\infty$

Suporemos ainda que

(2.3) $P_t \; I_\Omega = 1$

Os operadores de transição se articulam numa estrutura de semi-grupo, expressa pela relação:

(2.4) $P_t P_s = P_{t+s}$, t e s sendo instantes quaisquer.

Veremos, em seguida, que o caráter Markoviano do processo a ser construido está contido, em germe no caráter de semi-grupo da família inicial de operadores de transição.

Fixando-se um elemento x, ao acaso, em F, obtemos uma aplicação / $f\varepsilon C(F) \to (P_t f)(x) \varepsilon R$ que é contínua, linear e positiva. A ela corresponde, segundo o Teorema de Riesz(1) , uma medida $p_t(x,dy)$, definida sôbre os borelianos(2) de F, medida esta positiva, de massa total igual a 1, e tal que:

(2.5) $$(P_t f)(x) = \int_F f(y) p_t(x,dy).$$

Com essa representação, fica claro o significado dos operadores de transição: se A é um boreliano de F, então

$$(P_t I_A)(x) = \int_F I_A(y)\, p_t(x,dy) = \int_A p_t(x,dy) = p_t(x,A)$$

é exatamente a probabilidade de que um processo que no instante inicial está no estado x , esteja, t instantes depois, na região A do espaço de fases.

Passamos agora à construção do processo de Markov "realizando" o semi-grupo $(P_t)_{t\geq 0}$. Definimos o espaço de trajetórias Ω como sendo exatamente $F^{[0,\infty)}$. Munido da topologia produto, o conjunto Ω torna-se um espaço topológico. Há, então, sentido em se considerar o conjunto $C(\Omega)$ das funções numéricas contínuas com domínio Ω.

O conjunto $C(\Omega)$ é um reticulado vetorial de funções, contendo I_Ω. Trata-se, agora, de definir um operador de Daniell sôbre $C(\Omega)$.

Para isso, consideramos, preliminarmente, o subconjunto $C_f(\Omega)$, das funções pertencentes a $C(\Omega)$ que dependem apenas de um número finito de coordenadas de Ω.

Seja x um ponto qualquer de F, fixado. Definirmos sôbre $C_f(F)$ a aplicação E^x, que a cada função f, dependendo apenas das coordenadas $t_1, \ldots t_n$, associa o número:

(2.6) $E^x f = \int_F \ldots \int_F f(x_1, \ldots, x_n)\, p_{t_1}(x, dx_1) p_{t_2 - t_1}(x_1, dx_2) \ldots p_{t_n - t_{n-1}}(x_{n-1}, dx_n)$

De agora em diante consideraremos sempre, salvo referência em contrário, $C_f(\Omega)$ e $C(\Omega)$ munidos da topologia da convergência uniforme.

(2.7) <u>Proposição</u>. A aplicação $E^x : C_f(\Omega) \to R$ é linear, contínua, de norma igual a 1 e positiva.

<u>Demonstração</u>: Vamos verificar a linearidade no caso em que cada uma das funções depende apenas de uma única coordenada; a verificação é análoga para o caso geral. Sejam, então, f dependendo apenas da coordenada t, e g dependendo apenas da coordenada t_2 (sendo $t_1 \neq t_2$). Definimos agora $\tilde{f}$ e $\tilde{g}$ como segue:

$$f(x_1) = \tilde{f}(x_1, x_2), \text{ para todo } x_2$$

$$g(x_2) = \tilde{g}(x_1, x_2), \text{ para todo } x_1.$$

Então:

$$E^x \tilde{f} = \int\int_{FF} \tilde{f}(x_1, x_2) p_{t_1}(x, dx_1) p_{t_2 - t_1}(x_1, dx_2) = \int\int_{FF} f(x_1) p_{t_1}(x, dx_1) p_{t_2 - t_1}(x_1, dx_2) =$$

$$= \int_F E^x f \; p_{t_2 - t_1}(x_1, dx_2) = E^x f$$

e, da mesma forma, $E^x \tilde{g} = E^x g$.

Por outro lado, pela linearidade da integral,

$E^x(\tilde{f} + \tilde{g}) = E^x \tilde{f} + E^x \tilde{g}$, o que encerra a verificação, já que $\tilde{f} + \tilde{g} = f + g$.

A continuidade decorre da limitação da integral:

$$|E^x h| = \left| \int_F \cdots \int_F h(x_1, \ldots, x_n)\, p_{t_1}(x, dx_1) \ldots p_{t_n - t_{n-1}}(x_{n-1}, dx_n) \right| \leq$$

$$\leq \int_F \cdots \int_F |h(x_1, \ldots, x_n)|\, p_{t_1}(x, dx_1) \ldots p_{t_n - t_{n-1}}(x_{n-1}, dx_n) \leq$$

$$\leq ||h||_\infty \int_F \cdots \int_F p_{t_1}(x, dx_1) \ldots (p_{t_n - t_{n-1}}(x_{n-1}, dx_n) = ||h||_\infty \quad .$$

Finalmente, a positividade de E^x é exatamente o que (2.2) está asseguran do.

(2.8) <u>Proposição</u> : a) O conjunto $C_f(\Omega)$ é denso em $C(\Omega)$;

b) E^x admite uma extensão, por continuidade, a $C(\Omega)$.

<u>Demonstração</u> : Decorre imediatamente do <u>Teorema de Stone e Weierstrass</u>(3) que $C_f(\Omega)$ é denso em $C(\Omega)$. Além disso:

i) E^x assume valores em R que é, óbviamente, métrico e completo;

ii) E^x é uniformemente contínua, pois é contínua e linear.

Verificados os itens a), i) e ii), estamos em condições de aplicar o <u>Teorema de Extensão das funções contínuas</u>(4); dêle decorre b).

Como é usual, continuaremos usando a notação E^x, para indicar, agora, a função estendida.

(2.9) <u>Proposição</u> :$(C(\Omega), E^x)$ é um esbôço de espaço de probabilidade.

<u>Demonstração</u> : Não há o que demonstrar, exceto, talvez, que o axioma(1.12), da "continuidade", está satisfeito. Mas isso decorre do lema clássico de Dini(5) e do fato de E^x ser linear e contínua.

Definido o espaço de probabilidades do processo, resta-nos apresentar a família $(X_t)_{t \geq 0}$, das funções de observações. Elas são definidas da ma neira óbvia, como sendo as projeções de Ω em F.
Isto é, para cada $t \geq 0$, definimos

$$X_t \; : \omega \in \Omega \to \omega(t) \in F.$$

Êsse processo está munido naturalmente da família de passados $(H_t)_{t>0}$, assim definida

$$H_t = L^1((g(X_s))_{\substack{g \in C(F) \\ 0 \leq s \leq t}}, E^x)$$

Como argumentos análogos, àqueles invocados nas proposições (2.8) e (2.9), mostra-se que $C_f(F^{[0,t]})$ é denso em $C(F^{[0,t]})$ e que

$H_t = L^1(C(F^{[0,t]}), E^x)$. Isso nos será útil em seguida.

(2.10) <u>Proposição</u> : O processo $(L^1(C(\Omega), E^x), (X_t)_{t\geq 0})$ é de Markov com respeito à família de passados $(H_t)_{t\geq 0}$.

<u>Demonstração</u> : Sejam s e t dois instantes fixados ao acaso, sendo $s\leq t$. Seja f uma função pertencente a $C(F)$ e seja g uma função pertencente a $C(F^n)$, sendo n um inteiro qualquer fixado. Sejam finalmente n instantes quaisquer $t_1,\ldots,t_n$, já escritos na ordem <u>crescente</u> , com $t_n = s$.

Vamos mostrar que:

(2.11) $E^x(f(X_t)g(X_{t_1},\ldots,X_{t_{n-1}},X_s)) = E^x((P_{t-s}f)(X_s)g(X_{t_1},\ldots,X_{t_{n-1}},X_s))$.

Nossa intenção é clara; feitas todas as extensões, $(P_{t-s}f)(X_s)$ **será** exatamente $E^x(f(X_t)|H_s)$ e como $(P_{t-s}f)(X_s)$ é um elemento de $H(h(X_s); h\in C(F))$, estará assim demonstrado que $E^x(f(X_t)|H_s) = E^x(f(X_t)|X_s)$.

Voltemos a (2.11):

$E^x(f(X_t)g(X_{t_1},\ldots,X_s)) \stackrel{def}{=} \int_F\ldots\int_F f(y)g(x_1,\ldots,x_n)P_{t_1}(x,dx_1)\ldots P_{s-t_{n-1}}(xn-1,dxn)$ $P_{t-s}(x_n,dy) = \int_F..\int_F (P_{t-s}f)(x_n)g(x_1,\ldots,x_n)P_{s-t_{n-1}}(x_n-1,dx_n) \stackrel{def}{=} E^x((P_{t-s}f)(X_s)g$ $(X_{t_1},\ldots,X_s))$.

Trata-se, agora, de estender (2.11) a H_s. Para começar, seja ϕ uma função qualquer em $C(F^{(0,s)})$; existe uma sequência $(\phi_n)_n$ de elementos de $C_f(F^{(0,s)})$, tais que $\|\phi_n - \phi\|_\infty \to 0$, quando $n\to+\infty$. Mas:

i) para cada n, vale (2.11), isto é, $E^x(f(X_t)\phi_n) = E^x((P_{t-s}f)(X_s)\phi_n)$

ii) $|E^x(f(X_t)\phi_n) - E^x(f(X_t)\phi)| \leq \|E^x\|\ \|f(X_t)\|_\infty \|\phi_n - \phi\|_\infty$

e

$|E^x((P_{t-s}f)(X_s)\phi_n) - E^x((P_{t-s}f)(X_s)\phi)| \leq \|E^x\|\ \|(P_{t-s}f)(X_s)\|_\infty \|\sigma_n - \sigma_n\|_\infty$ **e**

portanto que

$\lim_{n\to\infty}\|\phi_n - \phi\|_\infty = 0$ acarreta $\lim_{n\to\infty} E^x(f(X_t)\phi_n) = E^x(f(X_t)\phi)$ e

que $\lim_{n\to\infty} E^x((P_{t-s}f)(X_s)\phi_n) = E^x((P_{t-s}f)(X_s)\phi)$.

Fazendo-se a conexão dos dois limites através de i), obtemos a extensão de (2.11) para as funções de $C(F^{[0,s]})$.

Seja, agora, ϕ um elemento de $H_s = L^1(C(F^{[0,s]}), E^x)$ e seja $(\phi_n)_n$ uma sequência de elementos de $C(F^{[0,s]})$ convergindo na norma de L^1 *para* ϕ.

iii) Pela extensão já obtida de (2.11), nós sabemos que, qualquer que seja n, vale a igualdade $E^x(f(X_t)\phi_n) = E^x((P_{t-s}f)(X_s)\phi_n)$

iv) Além disso : $|E^x(f(X_t)(\phi_n-\phi))| \leq E^x|f(X_t)||\phi_n-\phi| \leq ||f||_\infty ||\phi_n-\phi||_1$

e $|E^x((P_{t-s}f)(X_s)(\phi_n-\phi))| \leq E^x|(P_{t-s}f)(X_s)||\phi_n-\phi| \leq ||P_{t-s}f||_\infty ||\phi_n-\phi||_1$

Então, $\lim_{n\to\infty} ||\phi_n-\phi||_1 = 0$ acarreta que $\lim_{n\to\infty} E^x(f(X_t)\phi_n) = E^x(f(X_t)\phi)$ e que

$\lim_{n\to\infty} E^x((P_{t-s}f)(X_s)\phi_n) = E^x((P_{t-s}f)\phi)$. Conectando-se os dois limites, através de iii), obtém-se a extensão de (2.11) a H_s. Isso encerra a demonstração.

Noto que procedimento análogo ao adotado na 2a. extensão, pode ser usado, para estendermos ainda uma vez a fórmula (2.11), dessa vez para para funções f, de Baire, em F.

Construindo o processo, nós notamos que o resultado é, de certa forma, superior às expectativas, pois o que nós obtivemos foi não <u>um</u> processo, e sim todo um sistema de processos, cada um dêles correspondendo a uma particular escolha do ponto x de F, que aparece indexando o operador de Daniell construído.

O significado disso é claro: o semi-grupo $(P_t)_{t\geq 0}$ fornece ao processo a "lei" que rege a sua passagem de um estado a outro. Mas nada é dito sôbre o comêço do processo, sôbre seu estado inicial. A escolha de um ponto x do espaço de fase na definição de E^x, corresponde a escolher x como estado inicial do processo. De fato, se A é um boreliano qualquer de F, então :

(2.12) $E^x \, I_{\{X_o \in A\}} \overset{def}{=} (P_o I_A)(x) \overset{(2.4)}{=} I_A(x)$

O evento $\{X_o = x\}$ é o conjunto das trajetórias do rpocesso que têm x como estado inicial.

Então o que (2.12) está dizendo é que a probabilidade definida por E^x, tem tôda a sua massa concentrada em $\{X_o=x\}$, isto é, $E^x I_{\{X_o = x\}} = 1$.

Se nós dispomos de um tal sistema $L^1(C(\Omega), E^x)_{x\varepsilon F}$ de processos de Markov, então a condição (2.1) pode ser formulada como:

(2.1)' $E^x(f(X_t)|H_s) = E^{X_s}(f(X_{t-s}))$.

Essa nova formulação talvez indique mais claramente que um processo de Markov é essencialmente um processo que recomeça a cada instante.

C. Sè nós trabalharmos com processo de Markov canonicos, isto é, se identificarmos o espaço amostral com o espaço das trajetórias, então é possível dar à propriedade de Markov, uma formulação mais flexível. Fazemos, isso, atráves da noção de <u>translação</u> no espaço das trajetórias.

Seja, então, $s \geq 0$, um real qualquer; definimos a translação θ_s, de amplitude s, como sendo a aplicação:

$$\theta_s : \omega \varepsilon \Omega \to \theta_s(\omega) \varepsilon \Omega$$

tal que: $\theta_s(\omega)(t) = \omega(t + s)$, para todo $t \geq 0$.

A idéia é a seguinte: $\theta_s(\omega)$ é a trajetória que se obtém percorrendo-se (ou considerando-se) a trajetória antiga ω, apenas a partir do instante s, tomado como origem da nova trajetória.

Lembramos, outra vez, que $\Omega = F^{[0,\infty)}$, de sorte que a definição acima faz sentido.

Voltamos, agora, à expressão (2.1)':

(2.1)' $E^x(f(X_t)|H_s) = E^{X_s}(f(X_{t-s}))$, sendo $0 \leq s \leq t$.

Com a ajuda de θ_s nós reescrevemos (2.1)' como

$$E^x(f(X_{t-s}) o \theta_s \; |H_s) = E^{X_s}(f(X_{t-s}))$$

Indicando $f(X_{t-s})$ pela letra ϕ, a fórmula pode ser, finalmente, reescrita como:

(2.13) $E^x(\phi o \theta_s|H_s) = E^{X_s}(\phi)$.

No caso, $\phi = f(X_{t-s})$, mas o fato é que se a condição vale para as funções contínuas, de Ω em R, dependendo apenas de uma coordenada, então ela vale para tôdos os elementos de $L^1(C(\Omega),E^x)$. Essa afirmação pode ser demonstrada, usando como roteiro as proposições (2.8),(2.9) e (2.10) e suas respectivas demonstrações.

Em suma, se $(L^1(C(\Omega),E^x)_{x\epsilon F}, (X_t)_{t\geq 0})$ é uma família de processos canônicos, munidos de uma família de passados $(H_t)_{t\geq 0}$, nós diremos que se trata de um processo de Markov, se (2.13) se verificar, para todo $s\geq 0$ e para tôda ϕ de L^1. Essa nova definição é equivalente à anterior, para os processos nos quais faz sentido definir-se translação.

NOTAS E REFERÊNCIAS

(1) Teorema de Riesz : "Se $\psi : C(K) \to R$ é uma aplicação linear, contínua e positiva, sendo K um espaço topológico compacto, então existe uma medida positiva μ sôbre a σ-álgebra de Baire de K, segundo a qual tôdas as funções pertencentes a $C(K)$ são integráveis e tal que:

$$\psi(f) = \int_K f(t)\mu(dt), \text{ qualquer que seja}$$

$f \,\epsilon C(K)$"

Definição : A σ-álgebra de Baire de um espaço topológico é a menor σ-álgebra que torna mensuráveis tôdas as funções contínuas do espaço.

Noto que o Teorema de Riesz nada mais é que um caso particular do Teorema de Daniell. É imediato que $(C(K),\psi)$ formam um sistema de Daniell, o axioma (1.12) decorrendo do famoso e utilíssimo lema de Dini, segundo o qual se $(f_n)_n$ é um sequência de funções contínuas sôbre um compacto, tal que $f_n \downarrow 0$, então $\|f_n\|_\infty \downarrow 0$.

Ora:

$\psi(f_n) = |\psi(f_n)| \leq \|\psi\| \, \|f_n\|_\infty$, donde a convergência monótona de $(f_n)_n$ a 0, acarreta a de $(\psi(f_n))_n$

Uma função ψ, do tipo descrito acima, é o que Bourbaki chama de medida de Radon, sôbre um espaço compacto.

No §2, ao citar o Teorema de Riesz, nós já usamos o fato de que, em espaços métricos, a σ-álgebra de Baire coincide com a σ-álgebra de Borel.

(2) Num espaço topológico, a σ-álgebra de Borel é a menor σ-álgebra contendo todos os abertos do espaço. Se X designa o espaço, nós a notaremos usualmente como $\mathcal{B}(X)$.

(3) Teorema de Stone e Weierstrass : Dados um espaço compacto K e uma álgebra A de funções contínuas, contendo as constantes e separantes(isto é, tal que, se x e y são pontos distintos de K, então existe uma função f em A assumindo valôres distintos em x e em y), então A é denso em $\mathcal{C}$ (K)

(4) Teorema da Extensão de funções contínuas : Sejam S um subconjunto denso de um espaço métrico M e $f:S \to N$ uma aplicação uniformemente contínua onde N é um espaço métrico completp. Existe, então, uma única função $\bar{f}$: $M \to N$ que é contínua(e mesmo, é uniformemente contínua) e cuja restrição a S é f.

Noto que essa versão do teorema nos é suficiente, de vez que o espaço F com o qual estamos lidando, é compacto, com base enumerável de abertos, portanto, metrizável.

(5) lema de Dini : Seja K um espaço compacto. Se $(f_n)_n$ é uma sequência de funções contínuas sôbre K, tais que, $f_1(x) \geq f_2(x) \geq \ldots \geq f_n(x) \geq \ldots$, e $\lim_{n\to\infty} f_n(x) = 0$, para todo x em K, então $(f_n)_n$ converge uniformemente para zero.

"Para uma vasta classe de processos estocásticos, incluindo o processo de Wiener, a propriedade de Markov é preservada, se o processo é olhado não apenas como X_t, para t fixado, mas também como X_T, para algum T aleatório. Por exemplo, o movimento browniano sôbre a reta, começando em um ponto $x > 0$, se comporta, após atingir o zero pela primeira vez, exatamente como um processo iniciado no zero. Esta afirmação, apesar da sua aparente obviedade, necessita de demonstração..."(Dynkin).

§3 - Movimento Browniano

A. A propriedade forte de Markov

Doravante, desde que não haja perigo de confusão, nós nos referiremos a famílias de passados e processos, sem mencionar explicitamente o espaço de probabilidade que lhes serve de base, os processos sendo indicados por suas famílias de funções de observação.

Seja, então, $(\mathcal{H}_t)_{t\in R_+}$ uma família de passados e sejam S, T e T_n, $n\in N$, tempos aleatórios com relação a essa família de passados. É imediato que:

(3.1) se $S \leq T$, então $\mathcal{H}_S \subseteq \mathcal{H}_T$;

(3.2) se $(T_n)_n$ é uma sequência monótona decrescente, convergindo quase certamente para T, então:

$$\mathcal{H}_T = \bigcap_n \mathcal{H}_{T_n} .$$

Na verificação da propriedade forte de Markov, nós precisaremos de tempos aleatórios satisfazendo a uma condição mais forte que (1.21); nós queremos que:

(3.3) o conjunto $\{T \leq t\}$ seja um evento associado a $\mathcal{H}_t$, para todo instante t (veja bem: menor ou igual, e não estritamente menor, como em (1.21)).

(3.4) <u>Proposição</u> : Se $(\mathcal{H}_t)_t$ fôr contínua, então (3.3) será automaticamente verificado por todos os tempos aleatórios.

<u>Demonstração</u> : Seja t um instante qualquer e seja T um tempo aleatório.

É claro que $\{ T \leq t \} = \bigcap_{n=1}^{\infty} \{T<t + \frac{1}{n} \} = \bigcap_{n=m}^{\infty} \{T<t + \frac{1}{n}\}$ sendo m um inteiro positivo qualquer.

A última intersecção é, por (1.21) e pela monotonicidade da família de passados, um evento associado a $\mathcal{H}_{t+\frac{1}{m}}$. Como m é um inteiro positivo arbritrário, concluimos que $\{T\leq t\}$ é um evento associado a $\bigcap_{m=1}^{\infty} \mathcal{H}_{t+\frac{1}{m}} = \bigcap_{s>t} \mathcal{H}_s = \mathcal{H}_{t+}$.

Como, por hipótese, a família de passados é contínua, os subespaços $\mathcal{H}_t$ e $\mathcal{H}_{t+}$ coincidem, verificando-se, assim (3.3).

No que segue, $(X_t)_{t\geq 0}$ será sempre um processo de Markov, com respeito à família de passados $(\mathcal{H}_t)_{t\geq 0}$, tendo o espaço compacto F como espaço de fase e admitindo o semigrupo $(P_t)_{t\geq 0}$, de endomorfismos de $\mathcal{C}(F)$, como família de funções de transição.

(3.5) <u>Proposição</u> : Se T é um tempo aleatório com respeito a $(\mathcal{H}_t)_{t\geq 0}$, assumindo apenas um conjunto enumerável de valôres e satisfazendo (3.3), então:

(i) $g(X_T)$ é um elemento de $\mathcal{H}_T$, qualquer que seja $g \in \mathcal{C}(F)$;

(ii) $E(f(X_{T+s})(\mathcal{H}_T) = E^{X_T}(f(X_s)) = (P_s f)(X_T)$, qualquer que seja $f \in \mathcal{C}(F)$ e $s \in R_+$

<u>Demonstração</u> : Seja $\{t_0, t_1, \ldots\}$ o conjunto imagem de T

i) Seja t um instante qualquer; então

$$g(X_T)\, I_{\{T< t\}} = \sum_{t_j<t} g(X_{t_j})\, I_{\{T = t_j\}}$$

Mas essa é a expressão de uma variável aleatória associada a $\mathcal{H}_t$, já que $g(X_{t_j})$ é uma variável aleatória associada a $\mathcal{H}_{t_j}$ (pois $(\mathcal{H}_t)_{t\geq 0}$ é uma família de passados do processo), o mesmo acontecendo com $I_{\{T=t_j\}}$ (em virtude de (3.3)), de sorte que o produto das duas funções também é uma variável aleatória associada a $\mathcal{H}_{t_j}$; a função soma, para $t_j<t$, é uma variável

aleatória associada a H_t (em virtude da monotonicidade de $(H_t)_{t\geq 0}$).

Além disso, g é limitada (pois é uma função contínua sôbre um compacto), de sorte que a função $g(X_T)$ é integrável.

ii) Notemos inicialmente que $P_s f$ sendo um elemento de $C(F)$, a função $(P_s f)(X_T)$ será, por (i), um elemento de H_T, fazendo, pois sentido a afirmação (ii).

Sejam, então, um instante qualquer s e um elemento limitado ϕ de H_T; temos que:

$$E(f(X_{T+s})\ \phi) = E(f(X_{T+s}) \sum_{j=0}^{\infty} I_{\{T = t_j\}}\ \phi) =$$

$$= \sum_{j=0}^{\infty} E(f(X_{T+s})\ I_{\{T = t_j\}}\ \phi) \overset{*}{=} \sum_{j=0}^{\infty} E(f(X_{t_{j+s}})\ I_{\{T=t_j\}}\phi) =$$

$$= \sum_{j=0}^{\infty} E((P_s f)(X_{t_j})\ I_{\{T = t_j\}}\phi) = E(\sum_{j=0}^{\infty} (P_s f)(X_{t_j}) I_{\{T = t_j\}}\phi) =$$

$$= E((P_s f)(X_T)\phi).$$

A demonstração está assim encerrada. Na passagem assinalada com asterisco, usamos a propriedade de Markov e mais o fato de que se $\phi \varepsilon H_t$ e se T satisfaz (3.3), então $\phi I_{\{T = t_j\}} \varepsilon\ H_{t_j}$.

(3.6) <u>Proposição</u> : Na mesma situação de (3.5), seja S um outro tempo aleatório com respeito a $(H_t)_{t\geq 0}$, assumindo apenas um conjunto enumerável de valôres, <u>satisfazendo (3.3)</u> e <u>associado</u> a H_T. Então:

(3.7) $\quad E(f(X_{T+S})\ H_T) = (P_S f)(X_T)$.

<u>Demonstração</u> : Inicialmente, notamos que $(P_S f)(X_T)$ é um elemento associado a H_T, já que, por hipótese, S está associado a H_T. Faz, então, sentido a expressão (3.7).

A demonstração da igualdade é feita como em (3.5), só que dessa vez decompomos o espaço amostral em partes da forma $\{T = t_j \text{ e } S = s_i\}$.

(3.8) <u>Corolário</u> : Se $(H_t)_{t\geq 0}$ é contínua, então (3.7) vale para todos os tempos aleatórios discretos S e T, associados à família de passados, com a condição de S estar associado a H_T.

<u>Demonstração</u> : Basta aplicar (3.4).

A expressão (3.7) é a chamada <u>Propriedade Forte de Markov</u>. É a ela que Dynkim se refere no texto citado no começo do parágrafo. Nosso atual objetivo é estabelecer (3.7) em situações mais gerais. O corolário (3.8) nos sugere a conveniência de trabalhar sempre com famílias contínuas de passados, já que nesse caso a condição essencial (3.3) está automàticamente satisfeita, para todos os tempos aleatórios. A questão é saber se isso não limitará nosso campo de ação.

Seja $(H_t)_{t\geq 0}$ uma família qualquer de passados, nós já encontramos na proposição (3.4) a família modificada $(H_{t+})_{t\geq 0}$, cuja definição, vamos lembrar, era a seguinte:

$$\text{para todo } t\geq 0, \quad H_{t+} = \bigcap_{s>t} H_s$$

Essa nova família é, não há dúvida, contínua, e se pudéssemos considerar sempre os processos munidos dela, nossos problemas estariam resolvidos. A questão que se coloca, então, é saber se processos markovianos, com respeito a $(H_t)_{t\geq 0}$, continuarão markovianos, com respeito a $(H_{t+})_{t\geq 0}$.

A proposição seguinte nos fornece uma condição <u>suficiente</u> para que isso aconteça.

(3.9) <u>Proposição</u> : Seja $(X_t)_{t\geq 0}$ um processo de Markov com respeito à família de passados $(H_t)_{t\geq 0}$, tendo como espaço de fase o espaço métrico compacto F e admitindo o semigrupo $(P_t)_{t\geq 0}$ de endomorfismos de $C(F)$, como família de operadores de transição. Uma condição suficiente para que $(X_t)_{t\geq 0}$ seja de Markov com respeito a $(H_{t+})_{t\geq 0}$ é que:

i) $\lim\limits_{s\downarrow t} X_s = X_t$ q.c.

ii) $\lim\limits_{s\downarrow t} P_s f = P_t f$

Demonstração :

Como H_{t+} contém H_t, qualquer que seja t, não há dúvida de que $f(X_t) \varepsilon H_t$, qualquer que seja f ε $C(F)$. O que temos que mostrar é que

$$E(f(X_t) \mid H_{s+}) = E^{X_s}(f(X_{t-s}) = (P_{t-s}f)(X_s) \text{ sempre que } s \leq t.$$

Seja $s_1, s_2, s_3, \ldots, s_n, \ldots$ uma sequência monótona decreseente de números reais, convergindo para s e limitada superiormente por t. Pela propriedade de Markov:

$$E(f(X_t)|H_{s_n}) = (P_{t-s_n}f)(X_{s_n}) \quad , \text{ para todo n.}$$

As hipóteses i) e ii) asseguram que

$$\lim_{n\to\infty}(P_{t-s_n}f)\ (X_{s_n}) = (P_{t-s}f)(X_s).$$

Por outro lado:

$\lim\limits_{n\to\infty} E(f(X_t)|H_{s_n}) = E(f(X_t)|H_{s+})$ q.c. (o cálculo dêste limite está feito na nota (1), no fim do parágrafo). Isso encerra a demonstração.

Observações a respeito da proposição (3.9):

(1º) Como $(P_t)_t$ forma um semigrupo, a hipótese(ii) é equivalente a

$$\lim_{t\downarrow 0} P_t f = f$$

(2º) Na hipótese (ii), supusemos convergência simples, ponto a ponto.Pois bem, pode-se mostrar que, diante das outras condições impostas ao es_ paço e ao semigrupo, supor convergência simples é equivalente a supor convergência uniforme(demonstração no texto de Meyer,"Processos de Markov" , Cap. XIII - T 1)

Se F é um espaço métrico compacto, diremos que um semigrupo$(P_t)_{t\geq 0}$ de operadores de transição, definidos sôbre $C(F)$ (ver §2) é de Feller , se êle satisfaz a hipótes (ii), ou a sua equivalente:

$$(3.10) \quad \lim_{t\downarrow 0} \|P_t f - f\|_\infty = 0 \quad , \text{ para todo } f\varepsilon C(F)$$

(3º) O conjunto de hipóteses da proposição(3.6) é, de certa forma, redun_ dante. De fato, Doob mostrou, como uma aplicação da teoria dos martingais, que um processo de Markov, cuja família de operadores de transição satisfaz

(ii) (ou(3.10)), admite uma modificação tendo trajetórias contínuas à direita (isto é, satisfazendo(i)).

Vamos, finalmente, estabelecer a propriedade forte de Markov, numa situação mais geral.

(3.11) <u>teorema</u> : Se $(H_t)_{t\geq 0}$ é contínua, se o processo tem trajetórias contínuas à direita e se o semigrupo de operadores de transição é de Feller, então, para todo par S e T de tempos aleatórios relativos a $(H_t)_{t\geq 0}$, nós temos que:

i) a aplicação X_T é uma variável aleatória associada a H_T;

ii) se S é uma variável aleatória associada a H_T, a propriedade forte de Markov (3.7) vale.

<u>Demonstração</u> : Se S e T são tempos aleatórios discretos o teorema se reduz às proposições (3.5), (3.6) e (3.8) já demonstradas Vamos, então, supor que S e T têm imagens não enumeráveis.

(i) Para cada inteiro positivo n, seja a aplicação:

$$(3.12) \qquad T_n = \sum_{k=1}^{\infty} k\, 2^{-n}\, I_{\{(k-1)2^{-n} \leq T < k2^{-n}\}}$$

Assim definido, T_n é um tempo aleatório com respeito a $(H_t)_{t\geq 0}$. Com efeito, se t é um instante qualquer:

$$\{T_n < t\} = \bigcup_{\substack{1\leq k<t2^n \\ k\in N}} \{(k-1)2^{-n} \leq T < k\,2^{-n}\}$$

e êsse é um evento associado a H_t, já que T sendo um tempo aleatório, o conjunto $\{(k-1)2^{-n} \leq T < k2^{-n}\}$ é um evento associado $H_{k2^{-n}}$, de sorte que a reunião enumerável, para $k < t\,2^n$, pertencerá a H_t, devido à monotonicidade da família de passados.

A se-uência $(T_n)_n$ é monótona decrescente, convergindo para T em todos os pontos; então, por (3.2),

$$H_T = \bigcap_n H_{T_n}$$

Finalmente, como T_n é discreto, qualquer que seja n, a primeira parte da proposição(3.5) assegura que $g(X_{T_n})$ é um elemento de $\mathcal{H}_{T_n}$, qualquer que seja g que nós consideremos em $\mathcal{C}(F)$.

Ora, $\lim_{n\to\infty} g(X_{T_n}) = g(\lim_{n\to\infty} X_{T_n}) = g(X_T)$, a primeira passagem estando assegurada pela continuidade da g e a segunda, pela continuidade à direita das trajetórias do processo; além disso

$\lim_{n\to\infty} g(X_{T_n}) = \lim_{\substack{n\to\infty \\ n\geq m}} g(X_{T_n})$, de sorte que $g(X_T)$ é uma variável aleatória associada a $\mathcal{H}_{T_m}$,qualquer que seja m, pois que é limite, quase certamente, de uma sequência de variáveis aleatórias associadas a $\mathcal{H}_{T_m}$ (aqui,mais uma vez,levo em consideração a monotonicidade,agora da família$(\mathcal{H}_{T_n})_n$).

Na verdade, $g(X_T)$ é mesmo um elemento de $\mathcal{H}_{T_m}$, qualquer que seja m,já que g é limitada, por ser contínua com domínio compacto.Então $g(X_T)$ pertence à intersecção dos $\mathcal{H}_{T_m}$, isto é, pertence a $\mathcal{H}_T$, o que encerra o item (i).

ii) Dado o tempo aleatório S, definimos a sequência $(S_n)_n$ por uma expressão análoga a (3.12). Sejam agora, dois inteiros positivos m e n quaisquer. O tempo S_m é uma variável aleatória associada a $\mathcal{H}_{T_n}$. Com efeito, dados dois instantes quaisquer s e t, nós temos que:

$$\{S_m<s\}\cap\{T_n<t\}=\bigcup_{\substack{1\leq k<s2^m \\ 1\leq j<t2^n \\ j,k\,\in N}}(\{(k-1)2^{-m}\leq S<k2^{-m}\}\cap\{(j-1)2^{-n}\leq T<j2^{-n}\}) ,$$

e essa é a expressão de um evento associado a $\mathcal{H}_t$, já que, S sendo associado a $\mathcal{H}_T$(por hipótese), cada um dos conjuntos intersecção está associado a $\mathcal{H}_{j2^{-n}}$ e a reunião enumerável, para $j2^{-n}< t$, está associada a $\mathcal{H}_t$.

Nessas condições, a proposição(3.6) garante que

$E(f(X_{T_n}+S_m)|\mathcal{H}_{T_n}) = (P_{S_m} f)(X_{T_n})$ qualquer que seja f pertencente a $\mathcal{C}(F)$.

A questão, agora, é passar ao limite em m e n. Quanto ao termo do lado direito, não há dúvida:

$\lim_{n\to\infty} (P_{S_m} f)(X_{T_n}) = (P_{S_m} f)(X_T)$, pois, para cada ponto ω do espaço amostral, a função $P_{S_m(\omega)} f$ é contínua, e $X_{T_n} \downarrow X_T$, quase certamente, pela continuidade à direita das trajetórias do processo.

Também : $\lim_{m\to\infty} (P_{S_m} f)(X_T) = (P_S f)(X_T)$, pois $S_m \downarrow S$, quase certamente e a família de operadores de transição é, por hipótese, de Feller.

Finalmente, é verdade que:

$$\lim_{n\to\infty} E(f(X_{T_n} + S_m) | H_{T_n}) = E(f(X_T + S_m) | H_T)$$

e que

$$\lim_{m\to\infty} E(f(X_T + S_m) | H_T) = E(f(X_T + S) | H_T).$$

Essas duas passagens ao limite são mais delicadas e serão justificadas na nota (2), no final do parágrafo. Mas isso era tudo que restava para demonstrar.

Por todo o § , até aqui, nós escrevemos simplesmente E, em vez de E^x, porque era totalmente indiferente qual o particular ponto de partida das trajetórias. No corolário seguinte, continua sendo indiferente êsse ponto de partida, mas nós o mencionaremos, numa tentativa de tornar mais claros os cálculos da demonstração.

(3.13) <u>Corolário</u>(lei 0-1 de Blumenthal) : Seja $(X_t)_{t\geq 0}$ um processo de Markov, com respeito a uma família $(H_t)_{t\geq 0}$ de passados, tendo trajetórias contínuas à direita e realizando um semigrupo de Feller. Nessas condições, H_{0+} é trivial, isto é, $P^x(A) = 0$ ou 1, para todo evento A associado a H_{0+}.

<u>Demonstração</u> : $P^x(A) := E^x(I_A) = E^x(I_A . I_A) =$

$$= E^x\{(I_A . \theta_0) I_A\} \overset{*}{=} E^x\{E^x o(I_A) . I_A\} \overset{**}{=} E^x\{E^x(I_A)\ I_A\} \overset{***}{=} E^x(I_A) E^x\{I_A\} =$$

$$= \{P^x(A)\}^2.$$

A passagem marcada por(*) é justificada pela propriedade forte de Markov; (**) é porque $P^x\{X_0 = x\} = 1$; (***) é a passagem, para fora da integral, da constante $E^x(I_A)$.

B. O movimento browniano

O movimento browniano tem como espaço de fase o R^n que nós compactificamos, pelo método de Alexandroff, acrescentando mais um ponto no infinito, seja ∞ êsse ponto. O compactificado de R^n será notado $\overline{R^n}$. Como é habitual, $C(\overline{R^n})$ indica a classe das funções reais contínuas, cujo domínio é $\overline{R^n}$. Podemos identificar $C(\overline{R^n})$ como a classe

$$C_\infty(R^n) = \{f \in C(R^n) : \text{existe} \lim_{|x|\to\infty} f(x) < \infty\} ;$$

assim se f pertence a $C_\infty(R^n)$ nós a identificaremos com o elemento f' de $C(\overline{R^n})$ tal que : $f'(x) = f(x)$, para todo $x \in R^n$

$$\text{e} \qquad f'(\infty) = \lim_{|x|\to\infty} f(x) .$$

Se $x = (x_1, \ldots, x_n)$ é um elemento do R^n, nós escreveremos $|x|$, para indicar o número real positivo $(x_1^2 + \ldots + x^2n)^{1/2}$.

<u>Definição</u> : Para todo $t>0$, seja $n_t : R^n \to R$ a função que a cada ponto x do R^n, atribui o valor:

$$n_t(x) := \frac{1}{(2\pi t)^{n/2}} e^{-\frac{|x|^2}{2t}}$$

Essa família de funções tem as seguintes propriedades:

i) Qualquer que seja $t>0$, a função n_t é integrável e $\int_{R^n} n_t(x)\,dx = 1$.

ii) Em consequência,quaisquer que sejam $f \in C_\infty(R^n)$, $t>0$ e $x \in R^n$, está bem definida (e é finita) a integral $\int_{R^n} f(y)\, n_t(x-y)dy$. Vamos indicar por $(N_t f)(x)$ essa integral.

iii) Considerada como função de x,$(N_t f)(x)$ pertence a $C_\infty(R^n)$. De fato, seja $(x_n)_n$ uma sequência de pontos em R^n, tais que $|x_n - x| \to 0$ ($|x_n| \to \infty$) quando $n\to\infty$. Então, $f(y)\, n_t(x_n - y) \to f(y) n_t(x-y)$ ($f(y) n_t(x-y) \to 0$) e, como $|f(y) n_t(x_n - y)| \leq |f(y)|$, qualquer que seja y, o Teorema da Convergência Dominada garante que:

$(N_t f)(x_n) \to (N_t f)(x)$ $((N_t f)(x_n) \to 0)$, que é exatamente o que havíamos afirmado.

Notemos que em ii) e iii) não era necessário que f fôsse um elemento de $C_\infty(R^n)$. Bastava que f fôsse limitado e que houvesse sentido em se integrar $f(y)n_t(x-y)$; isto é, bastava que f fôsse uma função de Baire, limitada. Então :

(3.15) qualquer que seja f, de Baire e limitada, a função $N_t f$ pertence a $C_\infty(R^n)$; mais precisamente, pertence a $C_0(R^n) = \{ g \in C_\infty(R^n) : \lim_{|x|\to\infty} g(x) = 0\}$

iii) Qualquer que seja $t>0$, a aplicação

$$N_t : f \in C_\infty(R^n) \to N_t f \in C_\infty(R^n).$$

é um endomorfismo de $C_\infty(R^n)$.

Com efeito, a linearidade decorre, imediatamente, da definição de N_t. A continuidade também é imediata, já que $|(N_t f)(x)| \leq \| f \|_\infty$, qualquer que seja $x \in R^n$.

iv) $N_t(N_s f) = N_{t+s} f$, qualquer que seja $f \in C_\infty(R^n)$. Vamos fazer o cálculo para $n = 1$. Seja x um ponto qualquer de R^n:

$$\{N_t(N_s f)\}(x) = \int_R (N_s f)(y)\, n_t(x-y)dy = \int_R f(z) \int_R n_t(y-z) n_s(x-y) dy\, dz.$$

Fazemos agora as substituições: $u = x-z$ e $v = x-y$ e notamos que a integral

$$\int_R n_t(u-v)n_s(v)dv = \frac{1}{\sqrt{2\pi t}\sqrt{2\pi s}} \int_R e^{-\frac{(u-v)^2}{2t} - \frac{v^2}{2s}} dv =$$

$$= \frac{1}{\sqrt{2\pi t}\sqrt{2\pi s}} \int_R e^{-\frac{u^2}{2(s+t)} - \{(\frac{s+t}{2st})^{1/2} v - (\frac{2st}{s+t})^{1/2} \frac{u}{2t}\}^2} dv =$$

$$= n_{s+t}(u) \frac{1}{\sqrt{2\pi}} (\frac{s+t}{2st})^{1/2} \int_R e^{-\{(\frac{s+t}{2st})^{1/2} v - (\frac{2st}{s+t})^{1/2} \frac{u}{2t}\}^{1/2}} dv =$$

$$= n_{s+t}(u) \int_R n_1(w)dw = - n_{s+t}(u) \quad \text{, onde}$$

$$w = -\{(\frac{s+t}{2st})^{1/2} v - (\frac{2st}{s+t})^{1/2} \frac{u}{2t}\}^{1/2}.$$

Então: $\int_R f(z) \int_R n_t(y-z)\, n_s(x-y)\, dy\, dz = - \int_R f(x-u) \int_R n_t(u-v) n_s(v)\, dv(-du)$

$= \int_R f(x-u)\, n_{t+s}(u)\, du = \int_R f(z) n_{t+s}(x-z)dz = (N_{t+s} f)(x)$ que é o que queríamos demonstrar.

Vamos completar a família $(N_t)_{t>0}$, definindo N_0 como a aplicação identidade sôbre $C(\overline{R^n})$.

Definição : O processo de Markov realizando o semigrupo $(N_t)_{t\geq 0}$, de endomorfismos de $C(\overline{R^n})$, definido acima, será chamado de

Movimento Browniano (no R^n).

Vamos indicar por $(X_t)_{t\geq 0}$ êsse processo; o subespaço de probabilidades gerado por $(X_s)_{0\leq s\leq t}$ será notado por H_t.

3. Existência de uma versão com trajetórias contínuas do movimento browniano.

Até o fim do parágrafo, se não houver referência em contrário, $\{L^1(\overline{R^n}^{R_+}, E^x), (X_t)_{t\geq 0}\}$ será o movimento browniano com origem no ponto x, munido da sua família natural de passados $(H_t)_{t\geq 0}$.

(3.17) Proposição :Sejam s e t dois instantes e f um elemento de $C_\infty(R^n)$, quaisquer. Então: $E^x\{f(X_{t+s}-X_t)|H_t\}=E^0\{f(X_s)\}$.

Demonstração:

$$E^x\{f(X_{t+s}-X_t)|H_t\} = E^x\{f(X_s-X_0)\circ\theta_t|H_t\} \overset{(*)}{=}$$

$$= E^{X_t}\{f(X_s-X_0)\}.$$

Mas, qualquer que seja $y \in R^n$:

$$E^y\{f(X_s-X_0)\} \underset{**}{\overset{qc}{=}} E^y\{f(X_s-y)\} := \int_{R^n} f(z-y)n_s(z-y)dz =$$

$$\underset{(***)}{=} \int_{R^n} f(u)n_s(u)du := E^0\{f(X_s)\},$$ o que encerra a demonstração.

Vamos justificar, brevemente, as passagens em(*), usamos a propriedade de Markov; em (**), usamos o fato que $P^y\{X_0 = y\}= 1$; em(***), simplesmente fizemos a substituição $u = z-y$

(3.18) Corolário : Sejam os instantes $0 \leq t_1 \leq t_2 \leq t_3 \leq t_4$ e as funções f e g pertencentes a $C_\infty(R^n)$, quaisquer, então:

$$E^x\{f(X_{t_4} - X_{t_3})g(X_{t_2} - X_{t_1})\} = E^o\{f(X_{t_4-t_3})\}E^o\{g(X_{t_2-t_1})\}$$

Demonstração: $E^x\{f(X_{t_4} - X_{t_3})g(X_{t_2} - X_{t_1})\} =$

$$= E^x\{E^x\{f(X_{t_4} - X_{t_3})|H_{t_2}\}g(X_{t_2} - X_{t_1})\} \overset{(3.17)}{=}$$

$$= E^x\{E^o\{f(X_{t_4-t_3})\}\ g(X_{t_2}-X_{t_1})\} = E^o\{f(X_{t_4-t_3})\}E^x\{g(X_{t_2}-X_{t_1})\} =$$

$= E^o\{f(X_{t_4-t_3})\}\ E^o\ \{\ g(X_{t_2-t_1})\}$, como queríamos demonstrar.

No próximo corolário, tomaremos n = 1 (isto é, consideraremos o movimento browniano tendo $\overline{R}$, como espaço de fase).

(3.19) Corolário : Seja $i : \overline{R} \to R$ a seguinte função:

$$i(x) = x \quad , \text{ qualquer que seja } x \in R$$
$$i(\infty) = 0$$

Então, o processo $\{L^1(\overline{R}^{R_+}, E^x), (i(X_t))_{t \geq 0}\}$ é um martingal com respeito à família $(H_t)_{t \geq 0}$.

Demonstração : A aplicação $i(X_t)$ é contínua quase certamente, de sorte que não há dúvida de que $i(X_t)$ pertence a H_t, qualquer que seja t .

Por outro lado:

$$E^x\{i(X_{t+s})|H_t\} \overset{qc}{=} E^x\{i(X_t) + i(X_{t+s} - X_t)|H_t\} =$$

$$= E^x\{i(X_t)|H_t\} + E^x\{i(X_{t+s} - X_t)|H_t\} =$$

$= i(X_t) + E^o\{i(X_s)\} \overset{qc}{=} i(X_t) + \int_R x\ n_s(x)dx = i(X_t)$, encerrando a demonstração.

Notemos que $i(X_t) = X_t$ q.c. Assim, tôda versão de $i(X_t)$ será também uma versão de X_t.

A existência de uma versão com trajetórias contínuas será provada apenas para o movimento browniano uni-dimensional. E isso basta, já que o caso geral, do movimento em n dimensões, pode ser reduzido ao caso unidimensional, pela decomposição do movimento nas suas n componentes.

Mostraremos a existência de uma versão contínua do movimento browniano em duas etapas : na primeira, usando o fato de que $(i(X_t))_{t\geq 0}$ é um martingal , nós recorreremos ao teorema(1.27), e mostraremos a existência de uma versão com trajetórias contínuas à direita.

(3.20) <u>Proposição</u> : O processo $(i(X_t))_{t\geq 0}$ é contínuo em probabilidade.

<u>Demonstração</u> : queremos mostrar que $\lim\limits_{s\to t} P^x\{|i(X_t) - i(X_s)|\geq\delta\}=0$

qualquer que seja $\delta=0$.

Mas: $P^x\{|i(X_t) - i(X_s)|\geq\delta\} = \int\limits_{|y|\geq\delta} n_{t-s}(y)dy$

Fazendo a substituição $u = \dfrac{y}{t-s}$ (estamos supondo $t\geq s$), a última integral adquire a seguinte forma:

$$\int\limits_{u^2\geq\frac{\delta^2}{|t-s|}} n_1(u)du$$

Como n_1 é uma função integrável, essa última integral tende a zero, quando $|t-s|\to 0$.

(3.21) <u>Corolário</u> : Vamos definir, para todo t, a função

$$i(X_{t+}) = \lim_{n\uparrow\infty} i(X_{t+\frac{1}{n}}) \ .$$

Então: $i(X_t) = i(X_{t+})$ q.c.

(3.22) <u>Corolário</u> : O processo$(i(X_{t+}))_{t\geq 0}$ é um supermartingal com respeito à família modificada $(\mathcal{H}_{t+})_{t\geq 0}$ (3)

(3.23) <u>Proposição</u> : Seja $\{L^1(\overline{R}^{R_+}, E^x) , (X_t)_{t\geq 0}\}$ o movimento browniano unidimensional, munido da sua família natural de passados $(\mathcal{H}_t)_{t\geq 0}$.

Então, existe um processo $\overline{(X_t)}_{t\geq 0}$ definido sôbre $L^1\overline{(R}^{R_+}, E^x)$ e tal que :

i) $(\overline{X}_t)_{t\geq 0}$ tem trajetórias contínuas à direita

ii) $\overline{X}_t = X_t$ qc , para todo $t\geq 0$.

<u>Demonstração</u> : O processo $(i(X_{t+}))_{t\geq 0}$ é um supermartingal com respeito à família contínua $(H_{t+})_{t\geq 0}$. Além disso,

$$i(X_{t+}) \overset{qc}{=} i(X_t) \quad \text{, de sorte que}$$

$$E^x\{i(X_{t+})\} = E^x\{i(X_t)\}\}$$

Mas , $E^x\{i(X_t)\} = \int_R y\ n_t(x-y)dy = x$, qualquer que seja $t\geq 0$, o que garante a continuidade da aplicação $t\varepsilon R_+ \to E^x\{i(X_t)\}$ e a aplicabilidade do teorema (1.27).

Existe, então, uma versão com trajetórias contínuas à direita de $(i(X_{t+}))_{t\geq 0}$. Seja $\overline{(X_t)}_{t\geq 0}$ essa versão.

Mas, êsse processo $(\overline{X}_t)_{t\geq 0}$ é exatamente o processo prometido, já que:

$$\overline{X}_t \overset{qc}{=} i(X_{t+}) \overset{qc}{=} i(X_t) \overset{qc}{=} X_t.$$

Esta assim demonstrada a proposição .

Daqui, por diante, sempre que mencionarmos o movimento browniano, estaremos nos referindo à sua versão possuindo trajetórias contínuas à direita.

O que nós mostraremos até o fim do parágrafo é que, na verdade, essa versão tem, só pode ter, trajetórias contínuas. Essa revelação nos chegará a partir da propriedade (3.26) que nós apresentaremos logo abaixo.

Vamos definir uma distância ρ , compatível com a topologia do espaço com a topologia do espaço compacto $\overline{R}$.

Sejam as funções:

$$\alpha(x,y) = \begin{cases} |x-y| / (2 + |x|^2 + |y|^2) & , \text{ se } x \in R \text{ e } y \in R \\ 0 & , \text{ se } x = \infty \text{ ou } y = \infty \end{cases}$$

$$\beta(x) = \begin{cases} |x|/(2 + |x|) & , \text{ se } x \in R \\ 1/2 & , \text{ se } x = \infty \end{cases}$$

Então, $\rho(x,y) := \alpha(x,y) + |\beta(x) - \beta(y)|$, quaisquer que sejam x e y pertencentes a $\bar{R}$.

Assim definida, ρ é uma distância sôbre $\bar{R}$. Vê-se fàcilmente que:

(3.24) $\rho(x,y) \leq |x-y|$, se x e y pertencem a R.

(3.25) <u>Proposição</u> : Qualquer que seja $\varepsilon > 0$,

$$\lim_{t \downarrow 0} \frac{1}{t} P^x\{|X_t - X_o| \geq \varepsilon\} = 0$$

<u>Demonstração</u> : Notemos inicialmente que, qualquer que seja $x \in R$:

$$P^x\{|X_t - X_o| \geq \varepsilon\} = P^o\{|X_t| \geq \varepsilon\} \quad \int_{|x| \geq \varepsilon} n_t(x)dx.$$

Fazemos a substituição $u = \frac{x}{\sqrt{t}}$. Tudo que temos a mostrar, então, é que:

$$\lim_{t \downarrow 0} \frac{1}{t} \int_{u^2 \geq \frac{\varepsilon^2}{t}} n_1(u)_2\, du = \lim_{t \downarrow 0} \frac{1}{2 \pi t} \int_{u^2 \geq \frac{\varepsilon^2}{t}} e^{-\frac{u^2}{2}} du = 0$$

Mas, isso decorre, imediatamente, da seguinte desigualdade(4):

$$\int_{|u| \geq h} e^{-\frac{u^2}{2}} du \leq \frac{2}{h} e^{-\frac{h^2}{2}}$$

Está, assim, terminada a demonstração.

(3.26) <u>Corolário</u> : Qualquer que seja $\varepsilon \geq 0$,

$$\lim_{t \downarrow 0} \frac{1}{t} P^x\{\rho(X_t, X_o) \geq \varepsilon\} = 0$$

<u>Demonstração</u>: Pela observação(3.24), sabemos que o evento

$\{\rho(X_t, X_o) < \varepsilon\}$ contém o evento $\{|X_t - X_o| < \varepsilon\}$.

Logo, $P^x\{\rho(X_t,X_o)< \varepsilon\} \geq P^x\{|X_t - X_o|< \varepsilon\}$ e, portanto,

$$\lim_{t\downarrow 0} \frac{1}{t} P^x\{\rho(X_t,X_o)\geq \varepsilon\} = \lim_{t\downarrow 0} \frac{1}{t} (1 - P^x\{\rho(X_t,X_o)< \varepsilon\}) \leq$$

$$\leq \lim_{t\downarrow 0} \frac{1}{t} (1 - P^x\{|X_t - X_o|< \varepsilon\}) = \lim_{t\downarrow 0} \frac{1}{t} P^x\{|X_t - X_o|\geq \varepsilon\}=0,$$

que encerra a demonstração.

Vamos mostrar, de passagem, que $(N_t)_{t\geq 0}$ é um semigrupo de Feller:

(3.27) Corolário : $\lim_{t\downarrow 0} \|N_t f - f\|_\infty = 0$, para todo $f \in C(\bar{R})$.

Demonstração : Tendo domínio compacto, a função f, mais que contínua, uniformemente contínua. Portanto, dado $\varepsilon > 0$, existe $\delta \geq 0$ tal que :

$$|f(x) - f(y)| < \varepsilon \quad , \text{ sempre que } \rho(x,y)\leq \delta .$$

Por outro lado, o corolário (3.26) garante a existência de um instante t_o, tal que:

$$P^x\{\rho(X_t)> \delta\} < \frac{\varepsilon}{\|f\|_\infty} \quad , \text{ sempre que } t_o>t>0$$

Então, se $t_o>t>0$:

$$|(N_t f)(x) - f(x)|: = |\int_R \{f(y) - f(x)\}n_t(x-y)\, dy| \leq$$

$$\leq \int_{\rho(x,y)\leq \delta} |f(x) - f(y)|\, n_t(x-y)dy + \int_{\rho(x,y)>\delta} |f(x) - f(y)|n_t(x-y)dy \leq$$

$\leq \varepsilon + 2\|f\|_\infty P^x\{\rho(X_t)> \delta\} < 3\varepsilon$, como queríamos demonstrar.

Na verdade, nós poderíamos ter mostrado, a partir de (3.26), que, para o movimento browniano, vale mesmo o seguinte resultado:

$$\lim_{t\downarrow 0} \frac{1}{t} \|N_t f - f\|_\infty = 0 .$$

(3.28) Teorema : O movimento browniano tem a propriedade forte de Markov.

Demonstração : O teorema já está pràticamente demonstrado:(3.23) assegura a continuidade à direita das trajetórias do movimento browniano; (3.27) garante que seu semigrupo de operadores de transição é de Feller. Aplicamos, agora, o teorema (3.11) e encerramos o assunto

Chegou o momento do golpe de misericórdia na questão da continuidade do movimento browniano.

(3.29) <u>Teorema</u> : Seja $(X_t)_{t\geq 0}$ um processo de Markov, com trajetórias contínuas à direita e com limite à esquerda, realizando um semigrupo $(P_t)_{t\geq 0}$ de endomorfismos de $\mathcal{C}(F)$, sendo F um espaço compacto. Seja d uma distância compatível com a topologia de F. Para todo $x\varepsilon F$ e $\varepsilon>0$, seja $B^d(x,\varepsilon)=\{y\varepsilon F \mid d(x,y)\leq\varepsilon\}$

Se $\lim_{t\downarrow 0} \sup_{x\varepsilon F} \frac{1}{t}\{1-(P_t I_{B^d(x,\varepsilon)})(x)\}=0$, então as trajetórias do processo são quase certamente contínuas.

<u>Demonstração</u> : Tendo trajetórias contínuas à direita e com limite à esquerda, só resta ao processo a possibilidade de ser descontínuo aos saltos. Trata-se, então, de mostrar que o conjunto das trajetórias que saltam, é desprezível.

Chamemos de A o conjunto dos pontos amostrais cujas trajetórias saltam; $A = \bigcup_{n\varepsilon N} A_n$, sendo A_u o conjunto des pontos amostrais, cujas trajetórias saltam ao menos uma vez, durante o intervalo de tempo $[0,n]$.

Invocando outra vez a continuidade à direita e a existência de limite à esquerda das trajetórias do processo, escrevemos:

$$A_u = \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{l\geq k} \bigcup_{m=1}^{ln} \left\{ d\left(X_{\frac{m}{l}}, X_{\frac{m-1}{l}}\right) \geq \frac{1}{j} \right\}$$

Ora :

$$P^x\left(\bigcup_{m=1}^{ln} \left\{d\left(X_{\frac{m}{l}}, X_{\frac{m-1}{l}}\right) \geq \frac{1}{j} \right\}\right) \leq \sum_{m=1}^{ln} P^x\left\{d\left(X_{\frac{m}{l}}, X_{\frac{m-1}{l}}\right) \geq \frac{1}{j}\right\} \leq$$

$$\leq \sup_{x\varepsilon F} l_n\{1-(P_{\frac{1}{l}} I_{B^d(x, \frac{1}{j})})(x)\} \to 0, \text{ quando } l\to\infty.$$

. Portanto, $P^x(A_n) = 0$; $P^x(A) \leq \sum_{n=1}^{\infty} P^x(A_n) = 0$, como queríamos demonstrar.

(3.30) <u>Corolário</u> : A versão com trajetórias contínuas à direita, do movimento browniano, na verdade, tem quase certamente, trajetórias contínuas.

<u>Demonstração</u> : Não há o que demonstrar; basta reconhecer em (3.26) a condição suficiente do teorema acima. De fato, por definição, se $(X_t)_{t\geq 0}$ é o movimento browniano, então:

$P^x\{\rho(X_{t+s}, X_t) \leq \varepsilon\} = (N_s I_B \rho_{(x,\varepsilon)})(x)$. A condição de uniformidade do limite está satisfeita, já que $(N_s I_B \rho_{(x,\varepsilon)})(x) = (N_s I_B \rho_{(0,\varepsilon)})(0)$, qualquer que seja $x \varepsilon R$.

Notas e referências

(1) $\lim\limits_{n \to \infty} E(f(X_t) | H_{s_n}) = E(f(X_t) | H_{s+})$ q.c.

Para fazer justiça à citação de Hunt, nós obteremos o resultado acima, como consequência de um dos teoremas básicos da teoria dos martingais (referência: Meyer, "Probabilités et potentiel" - V - T21).

Teorema : Seja $(Y_n)_{n \varepsilon -N}$ um supermartingal com respeito à família de passados $(G_n)_{n \varepsilon -N}$; designemos por $G_{-\infty} = \bigcap\limits_{n \varepsilon -N} G_n$.

Suponhamos que valha a condição $\sup\limits_n E(Y_n) < +\infty$.

As seguintes propriedades são, então, verdadeiras:

a) as variáveis aleatórias Y_n são uniformemente integráveis;

b) existe uma variável aleatória $Y_{-\infty}$ integrável, pertencente a $G_{-\infty}$, para a qual convergem, quase certamente e na norma de L^1, as variáveis Y_n, quando $n \to -\infty$;

c) o processo $(Y_n)_{n \varepsilon \{-\infty\} \cup (-N)}$ é um supermartingal com respeito à família $(G_n)_{n \varepsilon \{-\infty\} \cup (-N)}$.

Aplicação : Definimos $Y_{-n} = E(f(X_t) | H_{s_n})$

$$G_{-n} = H_{s_n}.$$

O processo $(Y_{-n})_{-n}$, assim definido, é um martingal com respeito a $(G_{-n})_{-n}$. De fato, qualquer que seja n , Y_{-n} pertence a G_{-n}, pela definição de esperança condicional. Além disso, dados dois inteiros quaisquer m e n, tais que $-m \leq -n$,

$$E(Y_{-n}) | G_{-m}) = E(E(f(X_t) | H_{s_n}) | H_{s_m}) =$$

$$= E(f(X_t) | H_{s_m}) = Y_{-m} \quad \text{q.c.}$$

Sendo um martingal, $E(Y_{-n}) = E(Y_{-1})$ qualquer que seja $n \in \mathbb{N}$, de fato :

$E(Y_{-n}) = E(E(f(X_t)|\mathcal{H}_{s_n})) = E(f(X_t))$, outra vez, pela definição de esperança condicional.

Nessas condições o teorema que acabamos de enunciar, garante a existência de uma variável aleatória $Y_{-\infty}$, pertencente a $\mathcal{G}_{-n} = \bigcap_n \mathcal{H}_{s_n} = \mathcal{H}_{s+}$, para a qual convergem, quase certamente e fortemente, os Y_{-n}.

Resta mostrar que $Y_{-\infty} = E(f(X_t)|\ \mathcal{H}_{s+})$ q.c.

Seja ϕ um elemento limitado, qualquer de $\mathcal{H}_{s+}$; a convergência forte de $(Y_{-n})_{-n}$ a $Y_{-\infty}$ garante que :

$$E(Y_{-\infty}\ \phi) = \lim_{-n \to -\infty} E(Y_{-n}\phi)$$

Por outro lado, pela definição de esperança condicional :

$$E(Y_{-n}\ \phi) = E\{E(f(X_t)|\mathcal{H}_{s_n})\phi\} = E\{f(X_t)\phi\} =$$

$$= E\{E(f(X_t)|\mathcal{H}_{s+})\phi\} \quad \text{qualquer que seja } n.$$

Ou seja, $E(Y_{-\infty}\ \phi) = E\{E(f(X_t)|\mathcal{H}_{s+})\phi\}$, qualquer que seja ϕ limitada pertencente a $\mathcal{H}_{s+}$.

Tomamos, então, $\phi_1 = I_{\{Y_{-\infty} < E(f(X_t)|\mathcal{H}_{s+})\}}$ e $\phi_2 = I_{\{Y_{-\infty} > E(f(X_t)|\mathcal{H}_{s+})\}}$, e concluimos que, forçosamente, $\phi_1 = 0$ q.c. e $\phi_2 = 0$ q.c., donde

$$Y_{-\infty} = E(f(X_t)|\mathcal{H}_{s+}) \text{ q.c.}$$

(2) i) $\lim_{n\to\infty} E(f(X_{T_n+S_m})|\mathcal{H}_{T_n}) = E(f(X_{T+S_m}|\mathcal{H}_T))$

Vamos obter i) como consequência de um resultado mais geral; êsse é o bom caminho.

No que segue estaremos sempre lidando com um espaço de probabilidades $L^1(S,E)$; ϕ, ϕ_n, ψ, λ serão elementos dêsse espaço e G, G_n serão subespaços de probabilidade.

a) <u>Proposição</u> : $|E(\phi|G)| \leq E(|\phi||G)$, quaisquer que sejam ϕ e G.

- Teorema da Convergência Monótona para esperanças condicionais

Seja $(\phi_n)_n$ uma sequência monótona, convergindo quase certamente para ϕ e tal que o conjunto

$\{E(\phi_n) : n \in N\}$ seja limitado. Nessas condições, qualquer que seja G,

$$E(\phi_n|G) \to E(\phi|G) \text{ q.c.}$$

Demonstração

Por hipótese, a sequência (ϕ_n) é monótona; vamos supô-la crescente. Então, se $m \leq n$

$$\|E(\phi_n|G) - E(\phi_m|G)\|_1 = E\{|E(\phi_n - \phi_m|G)|\} \leq$$

$$\leq E\{E(|\phi_n - \phi_m||G)\} := E\{|\phi_n - \phi_m|\} = E\{\phi_n - \phi_m\} \to 0$$

Isso mostra que $(E(\phi_n|G))_n$ é uma sequência de Cauchy de elementos de G. Como G é fechado (e L^1 é completo), existe um elemento ψ de G, para o qual a sequência de esperanças converge fortemente. A monotonicidade da sequência de esperanças garante que a convergência se dá, também, quase certamente.

Tudo o que resta a mostrar é que $\psi = E(\phi|G)$ q.c. Mas isso é uma decorrência da igualdade

$E(\psi\lambda) = E(E(\phi|G)\lambda)$, para todo λ limitado, pertencente a G.

) Teorema : Se $(\phi_n)_n$ é uma sequência monótona crescente, convergindo quase certamente para ϕ, e se $(G_n)_n$ é uma sequência monótona decrescente, tal que,

$$G = \bigcap_n G_n \quad \text{, então}$$

$$E(\phi_n|G_n) \to E(\phi|G) \text{ q.c.}$$

Demonstração:

Consideremos as seguintes desigualdades

$$\overline{\lim_n} E(\phi_n|G_n) = \lim_{m\to\infty} \sup_{n \geq m} E(\phi_n|G_n) \overset{(*)}{\leq}$$

$$\overset{(*)}{\leq} \lim_{m\to\infty} \sup_{n\geq m} E(\phi|G_n) = \overline{\lim_n} E(\phi|G_n) \overset{(**)}{=} E(\phi|G).$$

A passagem assinalada com (*) é justificada pelo fato de que a sequência $(\phi_n)_n$ é monótona crescente, monotonicidade essa que é preservada pela esperança condicional; a nota (1) justifica a passagem(**).

Analogamente, mostra-se que

$\underline{\lim}_n E(\phi_n|G_n) \geq \lim_n E(\phi_m|G_n) = E(\phi_m|G)$, qualquer que seja m. Ora, pelo teorema b),

$E(\phi_m| G) \uparrow E(\phi|G)$. Juntamos agora as duas desigualdades e obtemos a tese do teorema.

Corolário: O Teorema continua válido se a hipótese da monotonicidade de $(\phi_n)_n$ fôr substituída pela seguinte hipótese:

existe um elemento λ de L^1, tal que, qualquer que seja n, $|\phi_n| \leq \lambda$, q.c.

Demonstração : Definimos uma nova sequência

$$\phi_n' = \sup_{m \geq n} |\phi_n - \phi|$$

e recaimos no teorema.

d) O resultado i) sai agora como consequência direta do corolário acima; basta fazermos as seguintes identificações:

$$\phi_n = f(X_{T_n} + S_m) \;;\; \phi = f(X_{T+S_m})$$

$G_n = \mathcal{H}_{T_n}$, notando-se além disso que $|\phi_n| \leq \|f\|_\infty$, qualquer que seja n.

i) $\lim_{m\to\infty} E(f(X_{T+S_m})|\mathcal{H}_T) = E(f(X_{T+S})| \mathcal{H}_T)$.

O resultado decorre do seguinte toerema:

Teorema da Convergência Dominada para esperanças condicionais :

Seja $(\phi_n)_n$ uma sequência de elementos de um espaço de probabilidades L^1, convergindo quase certamente para ϕ.

Se existir um elemento ψ de L^1, tal que,

$|\phi_n| \leq \psi$ q.c., qualquer que seja n, então $E(\phi_n|G) \to E(\phi|G)$ q.c., qualquer que seja o subespaço de probabilidades que consideremos.

Demonstração: Basta definirmos uma nova sequência

$\phi'_n = \sup_{m\geq n} |\phi_m - \phi|$ e aplicarmos o teorema da convergência monótona para esperanças condicionais.

3) No corolário(3.22), nós utilizamos o seguinte teorema básico da teoria dos martingais (ver Meyer "Probabilités et Potentiel" VI - T4)

Teorema : Seja $(X_t)_{t\geq 0}$ um supermartingal, com respeito à família depas passados $(H_t)_{t\geq 0}$.

a) Valem as seguintes desigualdades:

$X_t \geq E(X_{t+} | H_t)$ q.c.

$X_{t-} \geq E(X_t | H_{t-})$ q.c.

) o processo $(X_{t+})_{t\geq 0}$ é um supermartingal com respeito à família de passados $(H_{t+})_{t\geq 0}$.

) ver Feller, "An Introduction to Probability Theory and its applications ol I, pag.166.