2
Noções sobre Martingais em tempo discreto e contínuo

Começamos recordando algumas noções básicas da Teoria de Probabilidade.

Definição 2.1 *Um* Espaço de Probabilidade *é uma tripla* $(\Omega, \mathcal{F}, \mathbb{P})$*, onde:*

1. $\Omega \neq \emptyset$ *é um conjunto arbitrário (dito "espaço amostral");*

2. $\mathcal{F}$ *é uma* σ*-álgebra de subconjuntos de* Ω *(os "eventos"), isto é:*

   (a) $\emptyset \in \mathcal{F}$

   (b) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

   (c) $\{A_n\}_{n\geq 1} \subset \mathcal{F} \Rightarrow \bigcup_{n\geq 1} A_n \in \mathcal{F}$

3. $\mathbb{P}$ *é uma função de conjunto*

$$\mathbb{P} : \mathcal{F} \to [0,1]$$
$$A \longmapsto \mathbb{P}(A),$$

tal que,

(i) $\mathbb{P}(\emptyset) = 0, \ \mathbb{P}(\Omega) = 1$

(ii) $\{A_n\}_{n\geq 1} \subset \mathcal{F}$*, disjuntos dois a dois* $(A_i \cap A_j = \emptyset, i \neq j)$
$$\Rightarrow \mathbb{P}(\bigcup_{n\geq 1} A_n) = \sum_{n\geq 1} \mathbb{P}(A_n) \ (\sigma - aditividade).$$

Obs: $\mathbb{P}(A)$ é a probabilidade do evento A. Em teoria da medida, o par $(\Omega, \mathcal{F})$ chama-se espaço mensurável, sendo os elementos de $\mathcal{F}$ chamados conjuntos mensuráveis.

Propriedades Básicas:

1. $A \in \mathcal{F} \Rightarrow \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$;
2. $A, B \in \mathcal{F}$ e $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$ (Monotonicidade);
3. $\mathbb{P}(\bigcup_{n\geq 1} A_n) \leq \sum_{n\geq 1} \mathbb{P}(A_n)$ ($\sigma-$sub aditividade).

Exemplo: Seja $\Omega = \{\omega_1, ..., \omega_n\}$ um espaço amostral finito, $\mathcal{F} = \mathcal{P}(\Omega)$ (conjunto das partes de Ω) e suponha que existam números reais $p_i \geq 0, i = 1, ...N$ com $\sum_{i=1}^{N} p_i = 1$. Então, se definirmos $\forall A \in \mathcal{F} : \mathbb{P}(A) \equiv \sum_{\omega_i \in A} p_i$, segue que $(\Omega, \mathcal{F}, \mathbb{P})$ e espaço de probabilidade.

Para lidar com o caso extremamente importante em que $\Omega = \mathbb{R}$, é conveniente a seguinte definição.

Definição 2.2 *A σ-álgebra gerada por uma coleção de subconjuntos de Ω ζ é a menor σ-álgebra contendo esta coleção, isto é, a interseção de todas as σ-álgebras contendo ζ.*

Exemplo: Se $\Omega = \mathbb{R}$, a σ-álgebra de borel é
$\mathcal{B}(\mathbb{R}) = \sigma(\{\text{intervalos semi-abertos } (a, b]\})$
$=\sigma(\{\text{intervalos } (a, b)\})$, isto é, a σ-álgebra gerada pelos abertos de $\mathbb{R}$
$=\sigma(\{\text{intervalos fechados } [a, b]\})$
$= \sigma(\{(-\infty, a]\})$

Para construir medidas de probabilidade em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, usa-se a noção de *função de distribuição*. Note que dada uma medida de probabilidade $\mathbb{P} : \mathcal{B}(\mathbb{R}) \longrightarrow [0, 1]$ em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ e definindo a função:

$$F : \mathbb{R} \longrightarrow \mathbb{R}$$

$$x \longmapsto F(x) \equiv \mathbb{P}(-\infty, x],$$

então tem-se que :

1. F é monótona não-decrescente;
2. F é contínua à direita com limites à esquerda;
3. $\lim_{x\to+\infty} F(x) = 1, \ \lim_{x\to-\infty} F(x) = 0.$

Definição 2.3 *Uma função $F : \mathbb{R} \longrightarrow \mathbb{R}$ satisfazendo $1-3$ é dita uma* função de distribuição.

Reciprocamente, o teorema seguinte afirma que uma função de distribuição determinha univocamente uma medida de probabilidade em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Teorema 2.4 *Dada uma função de distribuição $F : \mathbb{R} \longrightarrow \mathbb{R}$ existe uma única medida de probabilidade $\mathbb{P}$ em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tal que $\mathbb{P}(-\infty, x] = F(x)$ (ver [18]).*

Exemplo: $F(x) = \int_{-\infty}^{x} f(t)dt$ onde $f : \mathbb{R} \longrightarrow \mathbb{R}$, não negativa com $\int_{-\infty}^{\infty} f(t)dt = 1$, dita função densidade de probabilidade.

Por exemplo, $f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(t-m)^2}{2\sigma^2}}$ é uma densidade de probabilidade, dita normal/Gaussiana com parâmetros $m \in \mathbb{R}, \sigma > 0$.

Definição 2.5 *Dizemos que uma propriedade S em $(\Omega, \mathcal{F}, \mathbb{P})$ vale $\mathbb{P}$* quase-certamente *($\mathbb{P} - qc$ ou com $\mathbb{P}$ probabilidade um) se existe um conjunto $\Omega_0 \in \mathcal{F}$ com $\mathbb{P}(\Omega_0) = 1$ tal que a propriedade S é satisfeita $\forall \omega \in \Omega_0$.*

Obs: Em geral, o conjunto $\{\omega \in \Omega : \omega \text{ satisfaz } S\} \subset \Omega_0$ não é necessariamente mensurável. Porém, se $(\Omega, \mathcal{F}, \mathbb{P})$ for *completo* (ou seja, se $\mathcal{F}$ contém todo subconjunto de conjuntos de probabilidade zero), isto é verdade. Como todo espaço de probabilidade pode ser completado, supomos sempre que trabalhamos com $(\Omega, \mathcal{F}, \mathbb{P})$ completo (ver [18]).

Definição 2.6 *Dado um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, uma variável aleatória (va) é uma função*

$X : \Omega \longrightarrow \mathbb{R}$

$\omega \longmapsto X(\omega)$,

tal que $\forall B \in \mathcal{B}(\mathbb{R}), \{\omega \in \Omega : X(\omega) \in B\} \equiv X^{-1}(B) \in \mathcal{F}$.

Obs:Isto garante que faz sentido calcular $\mathbb{P}((X^{-1}(B)) \equiv \mathbb{P}_X(B)$; não é difícil checar que $\mathbb{P}_X(.)$ é uma medida de probabilidade em $\mathcal{B}(\mathbb{R})$ chamada de distribuição da va X.

Em teoria da medida, uma va nada mais é que uma função $\mathcal{F}/\mathcal{B}(\mathbb{R})$ mensurável.

A noção de independência é crucial em teoria da probabilidade.

Definição 2.7 *:*

1. *Eventos $A, B \in \mathcal{F}$ são independentes se $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$; uma coleção $\{A_i\}_{i \in I}$ de eventos é independente se, para todo subconjunto finito $J \subset I$:*

   $\mathbb{P}\left(\cap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i)$.

2. *Uma coleção $\{\mathcal{Q}_i\}_{i \in I}$, onde cada $\mathcal{Q}_i$ é uma coleção de eventos, é independente se, para todo $J \subset I$ finito tem-se $\mathbb{P}\left(\cap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i)$, onde $A_i \in \mathcal{Q}_i$.*

3. *Uma coleção $\{X_i\}_{i \in I}$ de va's é independente quando $\{\sigma(X_i)\}_{i \in I}$ é independente.*

   ***Obs**: $\sigma(X_i)$ é a σ-álgebra gerada por todos os conjuntos da forma $X^{-1}(B),\ B \in \mathcal{B}(\mathbb{R})$.*

2.1 Esperança/Valor médio/Valor esperado

Lembramos o processo de construção da esperança de uma variável aleatória.

Seja X uma va simples, isto é, $X(\omega) = \sum_{i=1}^{n} a_i \mathbb{I}_{A_i}(\omega), A_i \in \mathcal{F}, \Omega = \sum_{i=1}^{n} A_i$(união disjunta), onde os $a_i's$ são todos distintos (pense $A_i = \{\omega \in \Omega : X(\omega) = a_i\}$), e $\mathbb{I}_{A_i}(\omega) = \begin{cases} 1, & \text{se } \omega \in A_i, \\ 0, & \text{caso contrário.} \end{cases}$, chamada a função indicadora de A_i.

Se X é simples definimos a esperança de X por: $E[X] \equiv \sum_{i=1}^{n} a_i \mathbb{P}(A_i)$.

Em seguida, se $X \geq 0$, existe sequência não-decrescente $\{X_n\}_{n \geq 1}$ de va's simples, não negativas, tal que $\forall \omega \in \Omega$:

$X(\omega) = \lim\limits_{n \to \infty} X_n(\omega)$. Então define-se nesse caso:

$E[X] = \lim\limits_{n \to \infty} E[X_n]$ (podendo ser $+\infty$).

Finalmente, se X é uma va qualquer, como $X = X^+ - X^-$, onde $X^+ = \max(X, 0)$ e $X^- = \max(-X, 0)$, define-se: $E[X] = E[X^+] - E[X^-]$, desde que $E[X^+] < \infty$ ou $E[X^-] < \infty$.

Obs: Se $E[|X|] < \infty,\ X$ é dita integrável.

Notação: $E[X] = \int_\Omega X(\omega)\mathbb{P}(d\omega) = \int_\Omega X d\mathbb{P}$ (integral de Lebesgue com respeito a $\mathbb{P}$). Note também que:

$E[X\mathbb{I}_A] = \int_\Omega X\mathbb{I}_A d\mathbb{P} \equiv \int_A X d\mathbb{P}$.

Propriedades básicas(ver [5]):

1. Se X, Y são integráveis, então $\alpha X + \beta Y$ é integrável e

   $E[\alpha X + \beta Y] = \alpha E[X] + \beta E[Y], \alpha, \beta \in \mathbb{R}$ (linearidade);

2. $|E[X]| \leq E|X|$;

3. Se X, Y são integráveis e independentes, então XY é integrável e

   $E[XY] = E[X]E[Y]$;

4. Se X é uma va e $g : \mathbb{R} \longrightarrow \mathbb{R}$ uma função mensurável tal que $E|g(X)| < +\infty$, então,

   $E[g(X)] = \int_\Omega g(X(\omega))\mathbb{P}(d\omega) = \int_\mathbb{R} g(x)\mathbb{P}_X(dx)$;

   (Se X tem densidade f, então $E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$).

5. Se $X(\omega) = 0, \forall \omega \in N^c$, onde $N \in \mathcal{F}, \mathbb{P}(N) = 0$, então X é integrável e $E[X] = 0$ (Em particular, se X, Y são integráveis e $X = Y$ $\mathbb{P} - qc$, então $E[X] = E[Y]$);

6. Se $g : \mathbb{R} \longrightarrow [0, \infty]$ é $\mathcal{B}(\mathbb{R})$-mensurável e não-decrescente, então: $g(c)\mathbb{P}(X \geq c) \leq E[g(X)],\ c \in \mathbb{R}$ (Desigualdade de Markov).

Observações:

1. $X = Y$ $\mathbb{P} - qc \Longleftrightarrow \mathbb{P}(\omega \in \Omega : X(\omega) = Y(\omega)) = 1$.

2. X é va, $E[X^2] < \infty \Longrightarrow E[|X|] < \infty$ e definimos a *variância* de X

   $Var(X) \equiv E[(X - E[X])^2] = E[X^2] - (E[X])^2$, e $\sqrt{Var(X)}$ chama-se desvio padrão.

   Exemplo: Se X é uma va normal com parâmetros $m \in \mathbb{R}, \sigma > 0$, temos:

   $$E[X] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} xe^{\frac{-(x-m)^2}{2\sigma^2}} dx = m,$$

   e

   $$Var(X) = E[X^2] - (E[X])^2 = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} xe^{\frac{-(x-m)^2}{2\sigma^2}} dx - m^2 = \sigma^2$$

   Notação: $X \stackrel{d}{=} N(m, \sigma^2)$.

3. Definimos $L^1(\Omega, \mathcal{F}, \mathbb{P})$ como sendo o espaço das va's integráveis (identificadas a menos de um conjunto de probabilidade zero), que é um espaço vetorial normado completo, isto é, um espaço de Banach (ver [5]) com norma:

$\|X\|_1 = E[|X|]$

(mais geralmente $L^p(\Omega, \mathcal{F}, \mathbb{P})$, $\|X\|_p = (E[|X|^p])^{\frac{1}{p}}, p \geq 1)$

Definição 2.8 (Noções de convergência de va's) *Seja $\{X_n\}_{n\geq 1}$ sequência de va's e X va. Então dizemos que $\{X_n\}_{n\geq 1}$ converge a X:*

1. *$\mathbb{P} - qc$ quando*

   $\mathbb{P}(\omega \in \Omega : X_n(\omega) \xrightarrow{n\uparrow\infty} X(\omega)) = 1.$

2. *Em probabilidade, se $\forall \epsilon > 0$*

   $\mathbb{P}(\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \epsilon) \xrightarrow{n\uparrow\infty} 0.$

3. *Em $L^p(\Omega, \mathcal{F}, \mathbb{P})$ se*

   $\|X - X_n\|_p^p = E[|X - X_n|^p] \xrightarrow{n\uparrow\infty} 0.$

2.2 Esperança Condicional

A noção de dependência pode ser formalizada pela noção de condicionamento.

Dado $(\Omega, \mathcal{F}, \mathbb{P})$, $B \in \mathcal{F}$ tal que $\mathbb{P}(B) > 0$, definimos a probabilidade condicional de $A \in \mathcal{F}$ dado B por

$$\mathbb{P}_B(A) = \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Se A e B são independentes então $\mathbb{P}_B(A) = \mathbb{P}(A)$.

Note que $\mathbb{P}_B(.)$ é uma medida de probabilidade em $(\Omega, \mathcal{F})$ com $\mathbb{P}_B(B) = 1$ e $\mathbb{P}_B(B^c) = 0$. Dada uma va integrável X, definimos a esperança condicional de X dado B por: $E[X|B] \equiv \dfrac{E[X\mathbb{I}_B]}{\mathbb{P}(B)}$.

Obs: Não é difícil checar que se $\Omega = \displaystyle\sum_{n=1}^{+\infty} A_n$(união disjunta), $A_n \in \mathcal{F}$, $\mathbb{P}(A_n) > 0$, então $E[X] = \sum_{n=1}^{+\infty} E[X|A_n]\mathbb{P}(A_n)$.

A seguir, suponha que $\Omega = \sum_{i=1}^n B_i$, $\mathbb{P}(B_i) > 0$. Seja $\mathcal{D} = \sigma(\{B_i\}_{i=1}^n)$, então definimos a esperança condicional com respeito à σ-álgebra $\mathcal{D}$ denotada por $E[X|\mathcal{D}]$, como a va simples dada por:

$E[X|\mathcal{D}](\omega) = \sum_{i=1}^n E[X|B_i]\mathbb{I}_{B_i}(\omega), \ \forall \omega \in \Omega.$

Note que $E[X|\mathcal{D}]$ é $\mathcal{D}$-mensurável e $E[E[X|\mathcal{D}]\mathbb{I}_D] = E[X\mathbb{I}_D], \forall D \in\mathcal{D}$. Estas duas propriedades motivam a definição geral de esperança condicional.

Definição 2.9 *Dado um espaço de probabilidade* $(\Omega, \mathcal{F}, \mathbb{P})$*,* X *uma variável aleatória tal que* $E[|X|] < \infty$*, e* $\mathcal{G} \subset \mathcal{F}$ *uma sub-*σ*-álgebra de* $\mathcal{F}$*, então a* esperança condicional de X com respeito a $\mathcal{G}$ *é aquela variável aleatória, denotada por* $E[X|\mathcal{G}]$*, tal que*

a) $E[X|\mathcal{G}]$ *é* $\mathcal{G}$*-mensurável;*

b) $E[E[X|\mathcal{G}]\mathbb{I}_D] = E[X\mathbb{I}_D], \forall D \in \mathcal{G}$.

Obs: Pode-se mostrar que tal va existe (ver [18]).

Note que $E[X|\mathcal{G}]$ é definida a menos de um conjunto de probabilidade zero e qualquer va Z que satisfaz a definição acima chama-se uma *versão* de $E[X|\mathcal{G}]$.

Algumas propriedades da esperança condicional:

(i) Se Z é qualquer versão de $E[X|\mathcal{G}]$, então $E[X] = E[Z]$;

(ii) $E[\alpha X + \beta Y|\mathcal{G}] = \alpha E[X|\mathcal{G}] + \beta E[Y|\mathcal{G}]$ $\mathbb{P} - qc,\ \alpha,\ \beta \in \mathbb{R}$ (linearidade);

(iii) Se X é $\mathcal{G}$-mensurável, então $E[X|\mathcal{G}] = X\ \mathbb{P} - qc$.

(iv) Se $X \geq 0\ \mathbb{P} - qc$, então $E[X|\mathcal{G}] \geq 0\ \mathbb{P} - qc$ (positividade);

(v) Se X é independente de $\mathcal{G}$, então $E[X|\mathcal{G}] = E[X]\ \mathbb{P} - qc$.

(vi) Se X é $\mathcal{G}$-mensurável, Y e XY integráveis, então $E[XY|\mathcal{G}] = XE[Y|\mathcal{G}]$ $\mathbb{P}-$qc;

(vii) Se $\mathcal{A}$, $\mathcal{G}$ são sub-σ-álgebras com $\mathcal{A} \subset \mathcal{G} \subset \mathcal{F}$, então $E[E[X|\mathcal{A}]|\mathcal{G}] = E[E[X|\mathcal{G}]|\mathcal{A}] = E[X|\mathcal{A}]\ \mathbb{P} - qc$ (propriedade de iteração ou da torre);

(viii) Se $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ é convexa e $E[|\varphi(X)|] < \infty$, então $E[\varphi(X)|\mathcal{G}] \leq \varphi(E[X|\mathcal{G}])\ \mathbb{P} - qc$.(desigualdade de Jensen condicional)

2.3 Martingais em tempo discreto

Estamos agora em condições de definir o conceito fundamental de martingal. Precisamos da noção de filtração:

Definição 2.10 *Um espaço com filtração* $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ *consiste em um espaço de probabilidade* $(\Omega, \mathcal{F}, \mathbb{P})$ *munido de uma* filtração $\mathbb{F} = \{\mathcal{F}_n\}_{n\geq 1}$, *isto é, uma coleção crescente de sub-*σ*-álgebras de* $\mathcal{F}$*, isto é,*

$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \subseteq ... \subseteq \mathcal{F}$

Intuição: Em diversos modelos (por exemplo, em finanças), $\mathbb{F}$ representa a "estrutura informacional"ou como a informação é disponibilizada no tempo. Usualmente $\mathbb{F}$ é o *filtro natural* associado a uma sequência $\mathbb{X} = \{X_n\}_{n\geq 0}$, isto é, $\mathcal{F}_n = \sigma(X_0, X_1, ..., X_n)$, que é a "história"do processo até o instante n. A monotonicidade indica que a informação é acumulada sem ser perdida ao longo do tempo.

Obs: Uma sequência $\{X_n\}_{n\geq 1}$ de va's é chamada *processo estocástico em tempo discreto.*

Definição 2.11 *Um processo* $\mathbb{X} = \{X_n\}_{n\geq 0}$ *é dito* $\mathbb{F}$*-adaptado se para todo* $n \geq 0$, X_n *é* $\mathcal{F}_n$ *- mensurável.*

Intuição: Se $\mathbb{X}$ é $\mathbb{F}$-adaptado, $X_n(\omega)$ é "conhecido"no instante n.

Obs: Se $\mathcal{F}_n = \sigma(Y_0, Y_1, ..., Y_n)$ para algum processo $\mathbb{Y} = \{Y_n\}_{n\geq 0}$, então nesse caso X_n é $\mathcal{F}_n$ - mensurável se, e somente se, X_n=$f_n(Y_0, Y_1, ..., Y_n)$ para alguma função $f_n : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$ que é $\mathcal{B}(\mathbb{R}^{n+1})$ - mensurável (ver [18]). Note que um processo $\mathbb{X}$ é automaticamente adaptado ao seu filtro natural.

Definição 2.12 *Um processo* $\mathbb{X} = \{X_n\}_{n\geq 0}$ *em* $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ *é um* martingal *(ou um* $(\mathbb{P}, \mathbb{F})-$*martingal) quando,*

1. $\mathbb{X}$ *é* $\mathbb{F}$ *-adaptado;*

2. $E[\mid X_n \mid] < \infty$ $\forall n \geq 0$;

3. $E[X_{n+1} \mid \mathcal{F}_n] = X_n$ $\mathbb{P}$-*q.c*, $\forall n \geq 0$.

$\mathbb{X}$ *é um sub (respectivamente super) - martingal se em 3. a igualdade é trocada por* $\geq$ *(resp.* $\leq$*).*

Observações:

(i) Martingais são processos "constantes em média", no sentido de que $E[X_1] = E[X_2] = E[X_3] = ...$, enquanto que sub (resp. super) martingais são "crescentes (resp. decrescentes) em média";

(ii) $\mathbb{X}$ é super-martingal $\Longleftrightarrow -\mathbb{X} = \{-X_n \colon n \geq 0\}$ é submartingal;

(iii) $\mathbb{X}$ é martingal $\Longleftrightarrow$ $\mathbb{X}$ é sub e super martingal (simultaneamente);

(iv) $\mathbb{X}$ com $X_0 \in L^1(\Omega, \mathcal{F}_0, \mathbb{P})$ é martingal (sub/super) $\Longleftrightarrow \mathbb{X} - X_0 \equiv \{X_n - X_0 : n \geq 1\}$ é martingal (sub/super); Logo podemos supor $X_0 \equiv 0$;

(v) Usando a propriedade da iteração da esperança condicional mostra-se que a propriedade *3.* na definição 2.12 equivale a $E[X_n \mid \mathcal{F}_m] = X_m$, $n > m$.

Exemplos:

1. Se $\{X_n\}_{n \geq 1}$ são v.a's i.i.d com $E[|X_n|] < \infty$ e $E[X_n] = 0$, $n \geq 1$, $\mathcal{F}_n = \sigma(X_0, X_1, ..., X_n)$, $\mathcal{F}_0 = \{\phi, \Omega\}$ e $S_n \equiv \sum_{k=1}^{n} X_k$, então $\mathbb{S} = \{S_n : n \geq 0\}$ é $(\mathbb{P}, \mathbb{F})$- martingal.

Intuição: O exemplo acima pode ser generalizado. Considere o seguinte jogo bastante simples. Uma moeda, não necessariamente honesta, é lançada sucessivamente e observa-se a face resultante. Se der cara, o jogador ganha a aposta (que é, digamos, paga pela banca de um cassino) e se der coroa ele perde (e paga à banca).

Supondo que os lançamentos sucessivos da mesma moeda sejam completamente independentes, podemos modelar a evolução do jogo através de uma sequência de variáveis aleatórias $\{X_n\}_{n \geq 0}$ independentes e com distribuição de Bernoulli:

$$\mathbb{P}(X_n = 1) = 1 \text{ e } \mathbb{P}(X_n = -1) = q,$$

onde $p + q = 1$ e associamos o valor 1 (respectivamente -1) quando der cara (respectivamente coroa). Isto é, $\{X_n\}_{n \geq 0}$ é um processo estocástico em tempo discreto. Seja $\mathcal{F}_n = \sigma(X_0, ..., X_n)$ um filtro natural.

No instante $n-1$ (já feito o lançamento correspondente) não sabemos o resultado do lançamento seguinte. Portanto, se H_n é o valor apostado sobre o resultado do instante n, H_n deve depender somente da informação disponível até o instante imediatamente anterior, $n-1$. Isto é,

$$H_n = H_n(X_0, X_1, ..., X_{n-1})$$

é uma função mensurável com respeito à $\mathcal{F}_{n-1}$, para todo $n \geq 1$. Um processo $\{H_n\}_{n\geq 1}$ tal que $H_n \in \mathcal{F}_{n-1}$ é dito *predizível*. No nosso exemplo $\{H_n\}_{n\geq 1}$ pode se pensado como uma *estratégia de aposta*.

Ora, se V_{n-1} é o montante que o apostador possui no instante $n-1$, então, seguindo sua estratégia, o saldo no instante n será dado por:

$$V_n = V_{n-1} + H_n X_n = V_0 + \sum_{k=1}^{n} H_k X_k,$$

e o seu ganho esperado na n-ésima jogada é

$$\begin{aligned} E[V_n - V_{n-1} \mid \mathcal{F}_{n-1}] &= E[H_n X_n \mid \sigma(X_0, X_1, ..., X_{n-1})] \\ &= H_n E[X_n \mid X_0, X_1, ..., X_{n-1}] = H_n(p-q). \end{aligned}$$

Note que para $p = q = \frac{1}{2}$, correspondendo a um jogo "honesto", tem-se

$$E[V_n \mid \mathcal{F}_{n-1}] = V_{n-1},$$

e o processo $\{V_n\}_{n\geq 0}$ é um martingal. Quando $p \leq q$ (respectivamente $p \geq q$), correspondendo a um jogo "desfavorável"(respectivamente "favorável") tem-se $E[V_n \mid \mathcal{F}_{n-1}] \leq V_{n-1}$, (respectivamente $E[V_n|\mathcal{F}_{n-1}] \geq V_{n-1}$) o processo é chamado **super**martingal (respectivamente **sub**martingal).

2. Seja $\mathbb{F}$ filtração e $X \in L^1(\Omega, F, \mathbb{P})$, então se $M_n \equiv E[X|\mathcal{F}_n]$, $\{M_n\}_{n\geq 1}$ é $(\mathbb{P}, \mathbb{F})$- martingal.

A seguinte noção é extremamente importante para o Teorema de Doob:

Definição 2.13 *Seja* $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ *espaço com filtração. Uma função* $\tau : \Omega \to \mathbb{N} \bigcup \{\infty\}$ *é um* Tempo de Parada *(TP) quando* $\forall n \in \mathbb{N} = \{0, 1, 2, ...\}$,

$$\{\tau = n\} \equiv \{\omega \in \Omega : \tau(\omega) = n\} \in \mathcal{F}_n.$$

Observações:

1. τ pode tomar valor $+\infty$, por exemplo, modelando o tempo de espera de um evento que nunca ocorre.

2. Note que $\{\tau = +\infty\} = \{\tau < +\infty\}^c = (\bigcup_{n\in\mathbb{N}}\{\tau = n\})^c = \bigcap_{n\in\mathbb{N}}\{\tau = n\}^c \in \mathcal{F}_\infty$. Assim, $\{\tau = n\} \in \mathcal{F}_n, \forall n \in \mathbb{N}\bigcup\{\infty\}$;

Exemplo: O primeiro instante que um conjunto é atingido: Seja $\{X_n\}_{n\geq 0}$ processo $\mathbb{F}$-adaptado e $A \in \mathcal{B}(\mathbb{R})$, então $\tau \equiv \inf\{n \in \mathbb{N} : X_n \in A\}$ (onde $\inf\emptyset \equiv 0$) é um tempo de parada pois, $\forall n \in \mathbb{N}$:

$$\{\tau = n\} = \{X_0 \in A^c, X_1 \in A^c, ..., X_{n-1} \in A^c\} \in \mathcal{F}_n.$$

Definição 2.14 *Seja τ um TP então definimos*

$$\mathcal{F}_\tau \equiv \{A \in \mathcal{F}_\infty : A \cap \{\tau = n\} \in \mathcal{F}_n, \forall n \in \mathbb{N}\} \subset \mathcal{F}_\infty$$

chamada σ-álgebra da "informação disponível até o instante τ".

Alguns fatos básicos (ver [5]):

1. $\mathcal{F}_\tau$ é σ-álgebra;

2. Se $\tau \equiv k \in \mathbb{N}$, então τ é TP e $\mathcal{F}_\tau = \mathcal{F}_k$;

3. Se τ é TP e $B \in \mathcal{F}_\tau$, então $B \cap \{\tau = \infty\} \in \mathcal{F}_\infty$ e portanto $B \cap \{\tau = n\} \in \mathcal{F}_n \ \forall n \in \mathbb{N}\bigcup\{\infty\}$;

4. τ é TP $\Longleftrightarrow \{\tau \leq n\} \in \mathcal{F}_n, \ \forall n \in \mathbb{N} \Longleftrightarrow \{\tau > n\} \in \mathcal{F}_n, \forall n \in \mathbb{N}$;

5. Se B $\in \mathcal{F}_\infty$ então $B \in \mathcal{F}_\tau \Longleftrightarrow B \cap \{\tau \leq n\} \in \mathcal{F}_n \ \forall n \in \mathbb{N}$;

6. Se τ_1 e τ_2 são tempos de parada, então $\tau_1 + \tau_2$ é tempo de parada;

7. Se $\{\tau_k\}_{k\geq 1}$ são tempos de parada, então $\sup_k \tau_k$ e $\inf_k \tau_k$ são tempos de parada (exemplo, $\tau_n \equiv min(\tau, n) = \tau \wedge n$);

8. Se $\{\tau_k\}_{k\geq 1}$ é sequência monótona de tempos de parada então $\lim_{k\to\infty} \tau_k$ é TP.

9. $\tau_1 \leq \tau_2 \Rightarrow \mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$;

10. Se $\{X_n\}_{n\in\mathbb{N}}$ é $\mathbb{F}$-adaptado e τ é TP , $\mathbb{P}(\tau < +\infty) = 1$, definimos

$$X_\tau(\omega) \equiv \sum_{m=0}^{+\infty} X_m \mathbb{I}_{\{\tau=m\}}(\omega).$$

Então X_τ é v.a e além disso é $\mathcal{F}_\tau$- mensurável;

11. Em particular se $\{X_n\}_{n\in\mathbb{N}}$ é martingal (super/sub) então o "processo truncado em τ ", $\{X_n^\tau \equiv X_{\tau\wedge n}\}_{n\in\mathbb{N}}$ é um (super/sub) $\mathbb{F}-$martingal.

De fato,

$$\begin{aligned} X_{\tau\wedge n} &= X_{\tau\wedge n}\mathbb{I}_{\{\tau\leq n-1\}} + X_{\tau\wedge n}\mathbb{I}_{\{\tau\geq n\}} \\ &= X_\tau \mathbb{I}_{\{\tau\leq n-1\}} + X_n \mathbb{I}_{\{\tau\geq n\}} \\ &= \sum_{m=0}^{n-1} X_m \mathbb{I}_{\{\tau=m\}} + X_n \mathbb{I}_{\{\tau\geq n\}}. \end{aligned}$$

Logo $X_{\tau\wedge n}$ é $\mathcal{F}_n$- mensurável e integrável Assim:

$$X_{n+1}^\tau \ - X_n^\tau = (X_{n+1} - X_n)\mathbb{I}_{\{\tau\geq n+1\}}.$$

Lembrando que $\{\tau \geq n+1\} = \{\tau \leq n\}^c \in \mathcal{F}_n$ e das propriedades da esperança condicional, temos:

$$\begin{aligned} E[X_{n+1}^\tau - X_n^\tau | \mathcal{F}_n] &= E[(X_{n+1} - X_n)\mathbb{I}_{\{\tau\geq n+1\}} | \mathcal{F}_n] \\ &= \mathbb{I}_{\{\tau\geq n+1\}} E[X_{n+1} - X_n | \mathcal{F}_n] = 0. \end{aligned}$$

Na última parte da tese precisaremos do seguinte resultado, uma versão do Teorema da Parada Ótima de Doob (ou Teorema da Amostragem Opcional). Este resultado afirma que sob certas condições a propriedade de (super) martingal é preservada entre instantes dados por tempos de parada.

Teorema 2.15 *Seja* $\mathbb{X} = \{X_n\}_{n\in\mathbb{N}}$ *um supermartingal em* $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ *e sejam* τ_1 *e* τ_2 *tempos de parada limitados* $\mathbb{P}$ *- qc com* $\mathbb{P}(\tau_1 \leq \tau_2) = 1$*. Então* X_{τ_1} *e* X_{τ_2} *são integráveis e*

$$E[X_{\tau_2}|\mathcal{F}_{\tau_1}] \leq X_{\tau_1} \ \mathbb{P} - qc.$$

Demonstração: Seja k $\in \mathbb{N}$ um majorante de τ_2. Então

$$|X_{\tau_i}| \leq \sum_{j=1}^{k} |X_j|,$$

e portanto $X_{\tau_i},\ i = 1, 2$ é integrável. Agora, suponha que fosse $\tau_2 \equiv k \in \mathbb{N}$. Seja $A \in \mathcal{F}_{\tau_1}$. Como $A \cap \{\tau_1 = j\} \in \mathcal{F}_j$ temos para $j \leq k$, usando a propriedade de supermartingal, que:

$$\int_{A\cap\{\tau_1=j\}} X_{\tau_1} d\mathbb{P} = \int_{A\cap\{\tau_1=j\}} X_j d\mathbb{P} \geq \int_{A\cap\{\tau_1=j\}} X_k d\mathbb{P}$$

Então, somando em $j = 0, 1, ..., k$,

$$\int_A X_{\tau_1} d\mathbb{P} = \sum_{j=0}^{k} \int_{A\cap\{\tau_1=j\}} X_j d\mathbb{P} \geq \sum_{j=0}^{k} \int_{A\cap\{\tau_1=j\}} X_k d\mathbb{P} = \int_A X_{\tau_2} d\mathbb{P}$$

o que demonstra o teorema no caso de $\tau_2 \equiv$ k. Suponha agora que $\tau_2 \leq$k (não constante). Aplicamos o resultado anterior para o supermartingal truncado $\{X_n^{\tau_2}\}_{n\geq 0}$ e para tempos de parada τ_1 e k. Então temos, $\forall A \in \mathcal{F}_{\tau_1}$:

$$\int_A X_{\tau_1} d\mathbb{P} = \int_A X_{\tau_1}^{\tau_2} d\mathbb{P} \geq \int_A X_k^{\tau_2} d\mathbb{P} = \int_A X_{\tau_2} d\mathbb{P}.\square$$

Corolário Se $\mathbb{X}$ é martingal, sob as hipóteses acima então $E[X_{\tau_2}|\mathcal{F}_{\tau_1}] = X_{\tau_1}$ $\mathbb{P}$-qc.

Demonstração: Basta aplicar o resultado anterior a $\mathbb{X}$ e $-\mathbb{X}$.□

2.4
Integrabilidade Uniforme e Convergência em L^1

Afim de generalizar o teorema de Doob para o caso de tempo contínuo, a ser utilizado no cap. 3, vamos precisar de alguns resultados sobre integrabilidade uniforme e convergência de martingais.

Definição 2.16 *Uma coleção ζ de variáveis aleatórias em $(\Omega, \mathcal{F}, \mathbb{P})$ é dita* Uniformemente Integrável *(UI) se*

$$\lim_{a\to\infty} \sup_{X\in\zeta} E[|X|\mathbb{I}_{\{|X|>a\}}] = 0.$$

Equivalentemente, dado $\epsilon > 0$, existe $k \geq 0$ tal que $\forall X \in \zeta$:

$$E[|X|\mathbb{I}_{\{|X|>k\}}] < \epsilon.$$

Observações:

1. Se ζ é coleção UI então tomando k_1 associada a $\epsilon = 1$ temos $\forall X \in \zeta$:

$$E|X| = E[|X|\mathbb{I}_{\{|X|>k_1\}}] + E[|X|\mathbb{I}_{\{|X|\leq k_1\}}] \leq 1 + k_1 < +\infty,$$

ou seja, $\forall X \in \zeta : X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Isto é, uma família UI de variáveis aleatórias é necessariamente limitada em $L^1(\Omega, F, \mathbb{P})$. Porém, não vale a recíproca. Por exemplo, seja $(\Omega, \mathcal{F}, \mathbb{P})$ =$([0, 1], \mathcal{B}[0, 1], \lambda)$ e considere $\{X_n\}_{n\geq 1}$, onde $X_n \equiv n\mathbb{I}_{B_n}$, onde $B_n = (0, \frac{1}{n})$, mas dado $k > 0, \forall n > k$

$$E[|X|\mathbb{I}_{\{|X|>k\}}] = n\mathbb{P}(B_n) = 1,$$

portanto ζ não é UI.

2. Um critério simples para UI: dominação.

Se existe $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, tal que $\forall X \in \zeta : |X| \leq Y$ $\mathbb{P}-$qc , então ζ é UI. De fato:

$$\sup_{X\in\zeta} E[|X|\mathbb{I}_{\{|X|>a\}}] \leq E[Y\mathbb{I}_{\{|X|>a\}}] \xrightarrow{a\uparrow+\infty} 0$$

Em particular, se $X_i \in L^1(\Omega, F, \mathbb{P}), i = 1, ..., n$ então $\zeta = \{X_1, ..., X_n\}$ é UI pois ∀i $= 1, ..., n$

$$|X_i| \leq \sum_{j=1}^{n} |X_j| = Y \in L^1(\Omega, F, \mathbb{P})$$

Proposição 2.17 *Seja $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$; então a coleção $\zeta = \{E[X|\mathcal{G}] : \mathcal{G}$ é sub-σ- álgebra de $\mathcal{F}\}$ é UI.*

Demonstração : Seja $\mathcal{G}$ uma sub-σ-álgebra de $\mathcal{F}$ e Y uma versão de $E[X|\mathcal{G}]$. Pela desigualdade de Jensen condicional: $|Y| \leq E[|X||\mathcal{G}]$ $\mathbb{P}$-q.c e então:

$$\int_{\{|Y|\geq a\}} |Y| d\mathbb{P} \leq \int_{\{|Y|\geq a\}} E[|X||\mathcal{G}] d\mathbb{P} \leq$$

$$\leq \int_{\{|Y|\geq a;|X|\leq b\}} |X| d\mathbb{P} + \int_{\{|Y|\geq a;|X|\geq b\}} |X| d\mathbb{P} \leq b\mathbb{P}(|Y| \geq a) + \int_{\{|X|>b\}} |X| d\mathbb{P}$$

$$\leq \frac{b}{a} E[|Y|] + \int_{\{|X|>b\}} |X| d\mathbb{P} \leq \frac{b}{a} E[|X|] + E[|X|\mathbb{I}_{\{|X|>b\}}]$$

que vai a zero quando fazemos $a \to +\infty$ e então $b \to +\infty$, ou seja,

$$\lim_{a\to\infty} \sup_{Y\in\zeta} E[|Y|\mathbb{I}_{|Y|\geq a}] = 0,$$

ou seja, ζ é uma família UI.

Proposição 2.18 *Se $\{X_n\}_{n\geq 1}$ é uma sequência UI e $\lim_{n\to\infty} X_n = X$ $\mathbb{P}-$ qc, então X é integrável e $\lim_{n\to\infty} E[|X_n - X|] = 0$*

Demonstração : Pelo Lema de Fatou (ver [5]),

$$E|X| = E[\lim_{n\to+\infty} |X_n|] \leq \lim_{n\to+\infty} E[|X_n|] \leq \sup_n E[|X_n|]$$

$$\leq \sup_n E[|X_n \mathbb{I}_{|X_n|>a}|] + a < +\infty$$

$$\Rightarrow X \in L^1(\Omega, F, \mathbb{P})$$

Agora,

$$|X_n - X| \leq |X_n - X|\mathbb{I}_{\{|X_n|\leq a\}} + |X|\mathbb{I}_{|X_{\{n\}}|>a} + |X_{\{n\}}|\mathbb{I}_{\{|X_n|>a\}} \ (*)$$

O primeiro termo é limitado por $(a + |X|) \in L^1$ e portanto novamente pelo Teorema da Convergência Dominada (TCD): $\lim_{n\to+\infty} E[|X_n - X|\mathbb{I}_{\{|X_n|\leq a\}}] = 0$. O segundo termo é dominado por $|X| \in L^1$ e portanto novamente pelo

TCD temos: $\lim_{n\to+\infty} E[|X|\mathbb{I}_{\{|X_n|>a\}}] = E[|X|\mathbb{I}_{\{|X|>a\}}]$. A esperança do terceiro termo é $E[|X|\mathbb{I}_{\{|X_n|>a\}}] \le \sup_n E[|X|\mathbb{I}_{|X_n|>a}]$. Assim, tomando a esperança em $(*)$ temos:

$$\lim_{n\to\infty} E|X_n - X| \le E[|X|\mathbb{I}_{|X|>a}] + \sup_n E[|X_n|\mathbb{I}_{\{|X_n|>a\}}]$$

que vai a zero quando $a \to +\infty$. Logo:

$$\|X_n - X\|_1 \xrightarrow{n\uparrow\infty} 0.$$

Proposição 2.19 *Se* $\{X_n\}_{n\ge1}$ *é uma sequência UI e* $\lim_{n\to\infty} X_n = X$ $\mathbb{P}-q.c$, *então* $\lim_{n\to+\infty} E[X_n|\mathcal{G}] = E[X|\mathcal{G}]$ *em* L^1 *e em probabilidade.*

Demonstração : Da proposição anterior sabemos que $X_n \xrightarrow{n\uparrow\infty} X$ em L^1. Agora, pela desigualdade de Jensen condicional:

$$\|E[X|\mathcal{G}]\|_1 \le \|X\|_1$$

$$\|E[X_n|\mathcal{G}] - E[X|\mathcal{G}]\|_1 \le \|X_n - X\|_1 \xrightarrow{n\uparrow\infty} 0$$

e portanto $E[X_n|\mathcal{G}] \xrightarrow{n\uparrow\infty} E[X|\mathcal{G}]$ em L^1. Pela desigualdade de Markov: convergência $L^1 \Rightarrow$ convergência em probabilidade.

2.5 Processos Estocásticos

Dado $(\Omega, \mathcal{F}, \mathbb{P})$ e $\mathbb{T} = [0, +\infty)$ ou $[0, 1]$, então um processo estocástico é uma coleção $\mathbb{X} = \{X_t\}_{t\in\mathbb{T}}$, onde cada X_t é uma variável aleatória.

Note que um processo estocástico pode ser pensado como uma função de duas variáveis

$$\mathbb{X} : \mathrm{T} \otimes \Omega \longrightarrow \mathbb{R}$$

$$(t, \omega) \mapsto \mathbb{X}(t, \omega),$$

onde fixando $t \in \mathbb{T}$, então

$$\mathbb{X}(t, .) : \Omega \longrightarrow \mathbb{R}$$

$$\omega \mapsto \mathbb{X}(t, \omega) = X_t(\omega)$$

é uma va enquanto que fixando $\omega \in \Omega$,

$$\mathbb{X}(.,\omega) : \mathbb{T} \longrightarrow \mathbb{R}$$

$$t \mapsto \mathbb{X}(t,\omega) = X_t(\omega),$$

é a “trajetória”(ou realização) do processo associada a ω.

Várias noções vistas anteriormente para sequências de va's são introduzidas para processos estocásticos. Assim, uma filtração é uma coleção crescente ($\mathcal{F}_t \subseteq \mathcal{F}_s\ \forall t \leq s$) $\mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathbb{T}}$ de sub-σ-álgebras de $\mathcal{F}$ e um processo $\mathbb{X} = \{X_t\}_{t\in\mathbb{T}}$ é dito $\mathbb{F}$ adaptado se $\forall t \in \mathbb{T},\ X_t$ é $\mathcal{F}_t$-mensurável. Usualmente $\mathbb{F}$ é o filtro natural, $\mathcal{F}_t = \sigma(X_s; s \leq t)$ [1]. Da mesma forma, um *tempo de parada* é uma va $\tau : \Omega \longrightarrow \mathbb{T}$ (possivelmente tomando valor $+\infty$) tal que $\forall t \in \mathbb{T},\ \{\tau \leq t\} \in \mathcal{F}_t$. Define-se também $\mathcal{F}_\tau = \{A \in \mathcal{F} : \forall t \in \mathbb{T},\ A \cap \{\tau \leq t\} \in \mathcal{F}_t\}$, que é uma σ-álgebra.

Várias propriedades de tempos de parada continuam valendo em tempo contínuo, por exemplo se τ é tempo de parada, então τ é $\mathcal{F}_\tau$-mensurável, se $\tau,\ \sigma$ são tempos de parada, então $\tau \vee \sigma,\ \tau \wedge \sigma$ são tempos de parada, se $\tau \leq \sigma$ são t.p, então $\mathcal{F}_\tau \subseteq \mathcal{F}_\sigma$, etc. (ver [9]).

Exemplo O *processo de Wiener* (ou movimento Browniano) (standard) em $(\Omega, \mathcal{F}, \mathbb{P})$, munido do filtro natural $\mathbb{F}$, é um processo estocástico $\{W_t\}_{t\in\tau}$ tal que

1. $\mathbb{P}(W_0 = 0) = 1$
2. Se $0 \leq s < t$ a va $W_t - W_s = \eta(0, t-s)$ é independente da $\sigma(W_u; u \leq s) = \mathcal{F}_s$ (variação é independente do passado).
3. $t \longrightarrow W_t(\omega)$ são contínuas $\mathbb{P}$-qc

Obs: Pode-se mostrar que um tal processo existe (ver [5]).

Definição 2.20 *Dado espaço de probabilidade* $(\Omega, \mathcal{F}, \mathbb{P})$ *e uma filtração* $\mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathbb{T}}$, *um processo* $\mathbb{M} = \{M_t\}_{t\in\mathbb{T}}$ *é um* $(\mathbb{P},\mathbb{F})$-martingal *se:*

- $\{M_t\}$ *é* $\mathbb{F}$ *adaptado;*
- $E|M_t| < \infty,\ \forall t \in \mathbb{T}$*;*
- $\forall s \leq t,\ E[M_t|\mathcal{F}_s] = M_s\ \mathbb{P} - qc$

[1] Por razões técnicas usualmente supõe-se que a filtração $\mathbb{F}$ satisfaz as chamadas condições usuais, a saber $\forall t \in \mathbb{T},\ \mathcal{F}_t$ é *completa* e *contínua á direita*, isto é, $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$, ver [9]

(análogo para sub/super)

Exemplo: Não é difícil checar que o processo de Wiener é um martingal com respeito ao filtro natural.

Lema 2.1 *Seja τ tempo de parada. Então existe uma sequência decrescente $\{\tau_n\}_{n\geq 1}$ de tempo de parada tomando valores num conjunto discreto, tal que $\tau_n \downarrow \tau$ e $\{\tau_n = +\infty\} = \{\tau_n = \infty\}$.*

Demonstração: Defina $\forall n \geq 1$; $\omega \in \Omega$:

$$\tau_n(\omega) = \begin{cases} \frac{k+1}{2^n} & \text{se } \dfrac{k}{2^n} < \tau(\omega) \leq \dfrac{k+1}{2^n},\ k = 0, 1, ..., \\ +\infty & \text{se } \tau(\omega) = +\infty. \end{cases}$$

Claramente $\tau_n \downarrow \tau$ e se $\dfrac{k}{2^n} < t \leq \dfrac{k+1}{2^n}$ temos:

$$\{\tau_n \leq t\} = \{\tau_n \leq \frac{k}{2^n}\} = \{\tau \leq \frac{k}{2^n}\} \in \mathcal{F}_{\frac{k}{2^n}} \subset \mathcal{F}_t$$

$$\Longrightarrow \tau_n \text{ é tempo de parada.}\square$$

Seja $\{X_t\}_{t\in\mathbb{T}}$ processo estocástico em $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{T})$ e um tempo de parada τ. Seja

$$X_\tau : \ \Omega \longrightarrow \mathbb{R} \tag{2-1}$$

$$\omega \longmapsto X_\tau(\omega) = \begin{cases} X(\tau(\omega), \omega) & \text{se } \tau(\omega) < \infty, \\ X & \text{se } \tau(\omega) = +\infty, \end{cases} \tag{2-2}$$

onde X é uma variável aleatória fixa. ($X = X(\infty) = \lim_{t\to\infty} X_t$ se este limite existe).

Lema 2.2 *Se $\tau < \infty$ ou $X(\infty)$ existe, então X_τ é variável aleatória.*

Demonstração: Assumindo $\{X_t\}_{t\in\mathbb{T}}$ contínuo (ou mesmo contínuo à direita com limites à esquerda), então $X_\tau(\omega) = \lim_{n\to\infty} X_{\tau_n}(\omega)$, onde $\{\tau_n\}_{n\geq 1}$ é sequência aproximadora para τ. Portanto basta provar que $\forall n \geq 1$, X_{τ_n} é mensurável.

Mas, $\forall B \in \mathcal{B}(\mathbb{R})$, temos:

$$\{X_{\tau_n} \in B\} = \{X_\infty \in B, \tau_n = +\infty\} \bigcup_{k=0}^{+\infty} \{X_{\frac{k+1}{2^x}} \in B\} \cap \{\tau_n = \frac{k+1}{2^n}\} \in \mathcal{F}. \ \square$$

Podemos agora demonstrar a versão em tempo contínuo do teorema da Parada Ótima:

Teorema 2.3 (Teorema da Parada Ótima, tempo contínuo) *Seja* $\{X_t\}_{t\in\mathbb{T}}$ *um* $\mathbb{F}$*-martingal contínuo (ou cadlag) e* τ_1,τ_2 *tempos de parada com* $\mathbb{P}(\tau_1 \leq \tau_2) = 1$, τ_2 *limitado* $\mathbb{P}$*-qc. Então,*

$$\mathbb{E}[X_{\tau_2}|\mathcal{F}_{\tau_1}] = X_{\tau_1}\ \mathbb{P} - qc$$

Demonstração: Primeiro, se $b > 0$ e τ um tempo de parada tomando valores num conjunto finito J, majorados por b, então $\{(M_u, \mathcal{F}_u)\}_{u\in J}$, onde $M_u(\omega) = X_u(\omega),\ u \in J$ é um martingal discreto. Então, pelo teorema de Doob em tempo discreto com $\sigma_1 = \tau$ e $\sigma_2 = b$ segue que

$$\mathbb{E}[M_{\sigma_2}|\mathcal{F}_{\sigma_1}] = M_{\sigma_1}.$$

Agora, $\mathbb{E}[M_{\sigma_2}|\mathcal{F}_{\sigma_1}] = \mathbb{E}[X_b|\mathcal{F}_\tau]$ e

$$M_{\sigma_1}(\omega) = M_\tau(\omega) = \sum_{u\in J} M_u(\omega)\mathbb{I}_{\{\tau=u\}}(\omega) = \sum_{u\in J} X_u(\omega)\mathbb{I}_{\{\tau=u\}} = X_\tau(\omega).$$

Assim, $\mathbb{E}[X_b|\mathcal{F}_\tau] = X_\tau$.

Segue então pela proposição 2.17 que a coleção de va's $\{X_\tau\}_\tau$ onde τ varia sobre a coleção de tempo de parada equilimitados por b e assumindo valores num conjunto finito, é UI.

Agora, sejam $\{\tau_n^{(1)}\}_{n\geq 1}$ e $\{\tau_n^{(2)}\}_{n\geq 1}$, tempos de parada tomando valores num conjunto finito e tais que $\tau_n^{(1)} \downarrow \tau_1$, $\tau_n^{(2)} \downarrow \tau_2$, como construídos no lema 2.1 e tal que $\forall n \geq 1\ \tau_n^1 \leq \tau_n^2$.

Novamente pelo teorema de Doob em tempo discreto obtemos

$$\mathbb{E}[X_{\tau_n^{(2)}}|\mathcal{F}_{\tau_n^{(1)}}] = X_{\tau_n^{(1)}}\quad \mathbb{P} - qc.$$

e condicionando com respeito a $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_n^{(1)}}$ vem que

$$\mathbb{E}[X_{\tau_n^{(2)}}|\mathcal{F}_{\tau_1}] = \mathbb{E}[X_{\tau_n^{(1)}}|\mathcal{F}_{\tau_1}]\quad \mathbb{P} - qc.\ (*)$$

Como trajetórias são contínuas (ou contínuas á direita) temos que para $n \uparrow +\infty$, $X_{\tau_n^{(1)}} \to X_{\tau_1}$ e $X_{\tau_n^{(2)}} \to X_{\tau_2}$ em toda parte. Como as sequências $\{X_{\tau_n^{(1)}}\}_{n\geq 1}$ e $\{X_{\tau_n^{(2)}}\}_{n\geq 1}$ são UI, segue pela proposição 2.19, passando ao limite em $(*)$ quando $n \uparrow \infty$ que, em L^1 e em probabilidade,

$$\mathbb{E}[X_{\tau_2}|\mathcal{F}_{\tau_1}] = X_{\tau_1}\ \mathbb{P} - qc. \quad \square$$