MARTINGALAS
Rosario Romera
Febrero 2009

1. Nociones básicas

Definición: Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $T \neq \emptyset$ y sea $(\mathcal{F}_t)_{t \in T}$ una filtración en $\mathcal{F}$. Una familia $\{X_t\}_{t \in T}$ de v.a. reales definidas sobre $(\Omega, \mathcal{F}, P)$ se llama una **Martingala** (resp, super, sub-) con respecto a $\{\mathcal{F}_t\}_{t \geq 0}$ si:

(i) Para cada $t \in T$, X_t es medible con respecto a $\mathcal{F}_t$

(ii) $E[|X_t|] < \infty$ para todo t.

(iii) $E[X_t \mid \mathcal{F}_s] = X_s$ c.s. para todo s, t con $s < t$ (respectivamente "$\leq$", "$\geq$").

(iii') La condición (iii) es equivalente a $\int_D E[X_t \mid \mathcal{F}_s] dP = \int_D X_s dP$ (respectivamente "$\leq, \geq$"). Para todo $D \in \mathcal{F}_s, \quad s < t$.

Nota: En el caso especial $T = \mathbb{N}$ se habla de martingalas con parámetro de tiempo discreto. En tal caso, la martingala $\{X_n\}_{n \in \mathbb{N}}$ con respecto a $\{\mathcal{F}_n\}_n$ puede ser interpretada como un modelo de "juego equilibrado", donde la v.a. X_n denota la fortuna del jugador después de n jugadas y $\mathcal{F}_n$ describe la historia del juego en los primeros n juegos. La condición (iii) es entonces equivalente a $E[X_{n+1} \mid \mathcal{F}_n] = X_n$ (donde suponemos $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$) y estaría indicando que la competencia es justa.

Observación 1: Si $T = \mathbb{N}$ entonces $\{X_n\}_{n \in \mathbb{N}}$ es una martingala (respectivamente supermartingala o submartingala) si y sólo si $E[X_{n+1} \mid \mathcal{F}_n] = X_n$ ("$\leq$", "$\geq$") para todo $n \in \mathbb{N}$.

Observación 2: Al tomar $D = \Omega$, obtenemos que para cada supermartingala (respectivamente submartingala) se satisface que: $E[X_t] \leq E[X_s]$ (resp. $E[X_t] \geq E[X_s]$).

Observación 3: Si T es totalmente ordenado y $\{X_t, t \in T\}$ es una martingala entonces los valores esperados $E[X_t]$ no dependen de t.

Observación 4: Si Ω consta de un único punto entonces una supermartingala $\{X_n\}_{n \in \mathbb{N}}$ no es otra cosa que una sucesión decreciente de números reales. (analogamente: submartingala corresponde a sucesión creciente de números reales)

Desigualdad de Jensen: Sea φ una función convexa sobre un intervalo $I \subset \mathbb{R}$, es decir que:

$$\varphi\left(\alpha x + (1-\alpha)y\right) \leq \alpha\varphi(x) + (1-\alpha)\varphi(y), \quad \forall x, y \in I; \quad \forall \alpha \in [0,1]$$

entonces:

Si Y es una v.a. integrable definida sobre $(\Omega, \mathcal{F}, P)$ con valores en I y si $\varphi \circ Y$ es integrable entonces:

i) $E[Y \mid \xi] \in I$ c.s para toda sub-σ-álgebra ξ de $\mathcal{F}$.

ii) En particular: $\varphi(E[Y]) \leq E[\varphi \circ Y]$

Observación

a) Si $X = \{X_t, t \in T\}$ y $Y = \{Y_t, t \in T\}$ son martingalas (resp. super-, submartingalas) con respecto a $\{\mathcal{F}_t\}_t$ y $a, b \in \mathbb{R}$ (resp. $a, b \in [0, +\infty)$) entonces $\{aX_t + bY_t\}_t$ es una martingala con respecto a $\{\mathcal{F}_t\}_t$ (resp. super-, submartingala).

b) Si $X = \{X_t, t \in T\}$ es una martingala entonces $X = \{X_t^2, t \in T\}$ (si las v.a. son dos veces integrables), $X = \{X_t^+, t \in T\}$, $X = \{X_t^-, t \in T\}$, $X = \{|X_t|, t \in T\}$ son submartingalas.

c) $X = \{X_t, t \in T\}$ es una supermartingala entonces $X = \{X_t^-, t \in T\}$ es una submartingala.

2. Martingalas con parámetro discreto

En la sección anterior hemos definido en general el concepto de martingala y hemos dado algunas de sus propiedades más importantes. En este capítulo centraremos nuestra atención en el estudio de las martingalas con parámetro discreto, esto es, aquellas para las cuales $T = \mathbb{N}$. Retomando la definición dada en la sección anterior tenemos:

Definicion 1: Un proceso estocástico $\{X_n\}_{n\in\mathbb{N}}$ es una martingala (respectivamente super- ,sub-martingala) con respecto a la filtración $\{\mathcal{F}_n\}_n$ si se satisfacen las siguientes condiciones:

(1) X_n es $\mathcal{F}_n$-medible

(2) $E[|X_n|] < +\infty$

(3) $E[X_{n+1} \mid \mathcal{F}_n] = X_n$ c.s. (respectivamente "$\leq$", "$\geq$")

Definicion 2: Si $\{X_n\}_{n\in\mathbb{N}}$ y $\{Y_n\}_{n\in\mathbb{N}}$ son dos procesos estocásticos entonces decimos que $\{X_n\}_{n\in\mathbb{N}}$ es una martingala (respec. super-, submartingala) con respecto a $\{Y_n\}_n$ si $\{X_n\}_n$ es una martingala (respec. super-, submartingala) con respecto a $\{\mathcal{F}_n\}_n$ donde $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$.

Observación Si $\{X_n\}_n$ es una submartingala con respecto a $\{\mathcal{F}_n\}_n$ entonces $\{-X_n\}_n$ es una supermartingala con respecto a $\{\mathcal{F}_n\}_n$. Por lo tanto, se obtiene, en general, con muy pocas modificaciones, que todo lo que se demuestra para submartingalas es válido para supermartingalas y viceversa.

3. Ejemplos

Ejemplo 1

Sea $X_1, X_2, \ldots$ una sucesión de variables aleatorias independientes con $E[X_n] = 1$ para todo n. Sea $Y_n = \prod_{i=1}^{n} X_i$ entonces $\{Y_n : n \geq 1\}$ es una martingala con respecto a $\{X_n\}_n$. En efecto:

$$
\begin{aligned}
E[Y_{n+1} \mid X_1, X_2, \ldots, X_n] &= E[Y_n X_{n+1} \mid X_1, X_2, \ldots X_n] \\
&= Y_n E[X_{n+1} \mid X_1, \ldots, X_n] \\
&= Y_n E[X_{n+1}] = Y_n
\end{aligned}
$$

Ejemplo : Martingala de Doob

Sean $X, Y_1, Y_2, \ldots$ variables aleatorias arbitrarias tales que $E[|X|] < \infty$ y sea $X_n = E(X \mid Y_1, Y_2, \ldots, Y_n)$ Entonces $\{X_n, n \geq 1\}$ es una martingala con respecto a $\{Y_n\}_n$. En efecto:

$$
\begin{aligned}
E[|X_n|] &= E\left[|E[X \mid Y_1, Y_2, \ldots, Y_n]|\right] \\
&\leq E\left[E[|X| \mid Y_1, Y_2, \ldots, Y_n]\right] = E[|X|] < +\infty
\end{aligned}
$$

$$
\begin{aligned}
E[X_{n+1} \mid Y_1, \ldots, Y_n] &= E[(X \mid Y_1, Y_2, \ldots, Y_{n+1}) \mid Y_1, \ldots, Y_n] \\
&= E[X \mid Y_1, \ldots, Y_n] = X_n
\end{aligned}
$$

4. Propiedades Elementales

1. Si $\{X_n\}_n$ es una martingala (super-martingala) con respecto a $\{Y_n\}_n$ entonces $E[X_{n+k} \mid Y_0, \ldots, Y_n] = X_n \quad ("\leq")$ para todo $k \geq 0$.

2. Si $\{X_n\}$ es una martingala (supermartingala) con respecto a $\{Y_n\}_n$ entonces para $0 \leq k \leq n$ se satisface

$$E[X_n] = E[X_k] \qquad (resp.\, E[X_n] \leq E[X_k])$$

3. Si $\{X_n\}_n$ es una martingala con respecto a $\{Y_n\}_n$ y f es una función convexa entonces $\{f(X_n) : n \geq 1\}$ es una submartingala con respecto a $\{Y_n\}_n$.

Demostraciones

Propiedad 1. (Por inducción)

Para $k = 0$ se tiene ya que X_n es medible con respecto a $\mathcal{F}_n = \sigma(Y_0, \ldots, Y_n)$. Supongamos válido para k y demostremos para $k+1$:

$$\begin{aligned} E[X_{n+k+1} \mid Y_0, \ldots, Y_n] &= E[E(X_{n+k+1} \mid Y_0, \ldots, Y_n, \ldots, Y_{n+k}) \mid Y_0, \ldots, Y_n] \\ &= E[X_{n+k} \mid Y_0, \ldots, Y_n] \\ &= X_n \end{aligned}$$

Propiedad 2.

$E[X_n \mid Y_0, \ldots, Y_k] = X_k$

Por lo tanto

$$E[X_n] = E\left[E[X_n \mid Y_0, \ldots, Y_k]\right] = E[X_k]$$

Propiedad 3.

$$\begin{aligned} E[f(X_{n+1}) \mid Y_1, Y_2, \ldots, Y_n] &\geq f(E[X_{n+1} \mid Y_1, \ldots, Y_n]) \quad \text{(desigualdad de Jensen)} \\ &= f(X_n) \end{aligned}$$

Observacion: Si se considera una martingala $\{X_n\}_n$ sin especificar la filtración se entenderá que se trata de una martingala con respecto a $\mathcal{F}_n = \sigma(X_2, X_2, \ldots, X_n)$.

5. Tiempo de Parada

Teorema 1: Si N es un tiempo de parada para la martingala $\{X_n\}_n$ entonces el proceso parado $\{\overline{X}_n\}$ definido por:

$$\overline{X}_n = \begin{cases} X_n & \text{si} \quad N \geq n \\ X_N & \text{si} \quad N < n \end{cases}$$

Es también una martingala

Observación

Como N es un tiempo de paro entonces $P(N < \infty) = 1$. Por lo tanto, para n suficientemente grande se tiene que $\overline{X}_n = X_N$. Por lo tanto, con probabilidad 1, $\overline{X}_n \rightarrow X_n$ cuando $n \to \infty$. Si N es acotado entonces se puede demostrar que lo anterior implica que $E[\overline{X}_n] \rightarrow E[X_N]$ cuando $n \to \infty$. Puesto que $E[\overline{X}_n] = E[X_1]$ para todo n entonces $E[X_N] = E[X_1]$

Es decir, si un jugador está participando en un juego equilibrado y decide salir del juego usando un tiempo de parada entonces su fortuna esperada final será igual a su fortuna esperada inicial.

Análogamente se tiene que:

Teorema 2: Si N es un tiempo de paro para $\{X_n : n \geq 1\}$ y N es acotado entonces:

$$\begin{aligned} &E[X_N] \geq E[X_1] \quad \text{si} \quad \{X_n : n \geq 1\} \quad \text{es una submartingala} \\ &E[X_N] \leq E[X_1] \quad \text{si} \quad \{X_n : n \geq 1\} \quad \text{es una supermartingala} \end{aligned}$$

Teorema 3 Si $\{X_n : n \geq 1\}$ es una submartingala y N es un tiempo de parada para $\{X_n : n \geq 1\}$ tal que $P(N \leq k) = 1$ para algún k entonces

$$E[X_1] \leq E[X_N] \leq E[X_k]$$

Teorema 4 Si $\{X_n : n \geq 1\}$ es una submartingala no-negativa entonces para todo $a > 0$, se tiene que:

$$P(\text{máx}(X_1, X_2, \ldots, X_k) > a) \leq \frac{E[X_k]}{a}$$

Corolario 1 Sea $\{X_n : n \geq 1\}$ una martingala. Entonces para $a > 0$ se satisface que:

$$P(\text{máx}\,(|X_1|, \ldots, |X_k|) > a) \leq \tfrac{E[|X_k|]}{a}$$
$$P(\text{máx}\,(|X_1|, \ldots, |X_k|) > a) \leq \tfrac{E[|X_k|^2]}{a^2}$$

Demostraciones

Teorema 1

Sea

$$I_n = \begin{cases} 1 & \text{si} \quad N \geq n \\ 0 & \text{si} \quad N < n \end{cases}$$

Tenemos:

Si $N \geq n$ entonces $\overline{X}_n = X_n$, $\overline{X}_{n-1} = X_{n-1}$, $I_n = 1$ y por lo tanto

$$\overline{X}_n = \overline{X}_{n-1} + I_n(X_n - X_{n-1}) \quad (*)$$

Si $N < n$ entonces $\overline{X}_{n-1} = \overline{X}_n = X_N$, $I_n = 0$ y de nuevo $(*)$ es válido.

Por lo tanto,

$$\begin{aligned} E[\overline{X} \mid X_1, X_2, \ldots, X_{n-1}] &= E[\overline{X}_{n-1} + I_n(X_n - X_{n-1}) \mid X_1, \ldots, X_{n-1}] \\ &= E[\overline{X}_{n-1} \mid X_1, \ldots, X_{n-1}] + I_n E[X_n - X_{n-1} \mid X_1, \ldots, X_{n-1}] \\ &= \overline{X}_{n-1} + I_n\left(E[X_n \mid X_1, \ldots, X_{n-1}] - E[X_{n-1} \mid X_1, \ldots, X_{n-1}]\right) \\ &= \overline{X}_{n-1} + I_n(X_{n-1} - X_{n-1}) = \overline{X}_{n-1} \quad \square \end{aligned}$$

Teorema 2

Como N es acotado y $\overline{X}_n \to X_N$, con probabilidad 1 cuando $n \to \infty$ entonces $\overline{X}_n \to X_N$ cuando $n \to \infty$, puesto que $\{\overline{X}_n : n \geq 1\}$ es una submartingala (resp. una supermartingala) entonces $E[\overline{X}_n] \geq E[X_1]$ para todo n y por lo tanto $E[X_N] \geq E[X_1]$ (respectivamente $E[\overline{X}_n] \leq E[X_1]$ para todo n).

Teorema 3

Como N es acotado entonces por el teorema anterior (Teorema 2) $E[X_N] \geq E[X_1]$.

Por otra parte,

$$\begin{aligned} E[X_k \mid X_1, X_2, \ldots, X_N = l, N = l] &= E[X_k \mid X_1, X_2, \ldots, X_l = l, N = l] \\ &= E[X_k \mid X_1, X_2, \ldots, X_l] \\ &\ \text{Por ser } N \text{ tiempo de paro con respecto a } \{X_n\}_n \\ &\geq X_l = X_N \quad \text{por ser } \{X_n\}_n \text{ una submartingala} \end{aligned}$$

Tomando valores esperados se obtiene:

$$E[X_k] \geq E[X_N] \quad \square$$

Teorema 4
Sea N el menor valor de i con $i \leq k$ tal que $X_i > a$.
Si $X_i \leq a$ para todo $i = 1, \ldots, k$ entonces definimos $N = k$.
N es un tiempo de paro con respecto a $\{X_n : n \geq 1\}$.
Además $\max(X_1, X_2, \ldots, X_k > a)$ si y sólo si $X_N > a$, por lo tanto

$$\begin{aligned} P\left(\max(X_1, X_2, \ldots, X_K) > a\right) &= P(X_N > a) \\ &= E[1_{\{X_N > a\}}] \\ &\leq E\left[\frac{1}{a}.X_N\right] \\ &= \frac{1}{a}E[X_N] \leq \frac{1}{a}E[X_K] \end{aligned}$$

pues $N \leq k \qquad \square$

Corolario 1
Como $\varphi(x) = |x|$ es una función convexa y $\{X_n : n \geq 1\}$ es una martingala, se sigue de la desigualdad de Jensen que $\{|X_n| : n \geq 1\}$ es una submartingala no negativa. Por el teorema anterior obtenemos que:

$$P\left(\max(|X_1|, \ldots, |X_k|) > a\right) \leq \frac{E[|X_k|]}{a}$$

Por otra parte, como $\Psi(x) = x^2$ es una función convexa y $\{X_n : n \geq 1\}$ es una martingala obtenemos de nuevo por la desigualdad de Jensen que $\{X_n^2 : n \geq 1\}$ es una submartingala no-negativa, por lo tanto,

$$P(\max(X_1^2, \ldots, X_k^2) > a^2) \leq \frac{E[X_k^2]}{a^2}$$

Esto es:

$$P(\max(|X_1|, \ldots, |X_k|) > a) \leq \frac{E[X_k^2]}{a^2} \qquad \square$$

6. Teorema de Convergencia

El siguiente teorema es básico en la teoría de martingalas

Teorema 1: Si $\{X_n\}_{n \geq 1}$ es una martingala tal que para algún $M < \infty$ se satisface que $E[|X_n|] \leq M$ para todo n entonces con probabilidad 1, $\lim_{n\to\infty} X_n$ existe y es finito. Para la demostración, ver Ross página 315.

Corolario 1: Si $\{X_n : n \geq 0\}$ es una martingala no negativa, entonces, con probabilidad 1, $\lim_{n\to\infty} X_n$ existe y es finito.

Vamos a dar a continuación dos aplicaciones del teorema de convergencia para martingalas, para hacerlo introduciremos el siguiente concepto:

Definición: Un **proceso de Galton-Watson** $\{Z_n\}_{n\in\mathbb{N}}$ es una cadena de Markov homogénea con conjunto de estados S los enteros no negativos. Sus probabilidades de transición Π_{ij} con $i, j \in S$ están expresadas en términos de una sucesión dada $\{p_k, X_k = 0, 1, 2, \ldots\}$, $p_k \geq 0$ y $\sum p_k = 1$ mediante

$$\Pi_{ij} = P(Z_{n+1} = j \mid Z_n = i) = \begin{cases} p_j^{*i} = 1, & i \geq 1, \quad j \geq 0 \\ \delta_{0j}, & i = 0, \quad j \geq 0 \end{cases}$$

donde δ_{ij} es el símbolo de Kronecker y $p_j^{*i}, j = 0, 1, 2, \ldots$ es la j-ésima componente de la i-ésima convolución de $\{p_k, k = 0, 1, 2, \ldots\}$. Para evitar trivialidades se supone que $p_0 + p_1 < 1$ y que $p_j \neq 1$ para todo j, pues en caso contrario el proceso se extingue con probabilidad 1.

El proceso de Galton-Watson $\{Z_n\}_n$ es un modelo matemático que permite describir el desarrollo de una población que se comporta de la siguiente manera:
Supongamos que en el tiempo $t = 0$ hay Z_0 individuos cada uno de los cuales, al cabo de una unidad de tiempo, es reemplazado por un número aleatorio de nuevos individuos (sus hijos) de acuerdo a una distribución común de descendencia $\{p_k, k = 0, 1, 2, \ldots\}$ e independientemente de los demás individuos. El número de individuos en la primera generación es entonces una suma de Z_0 variables aleatorias cada una con distribución $\{p_k, k = 0, 1, 2, \ldots\}$.
Los individuos de la primera generación son reemplazados, de manera análoga, por nuevos individuos los cuales constituyen la segunda generación y así sucesivamente. El número de hijos que cada individuo tiene es independiente de su historia familiar y de los demás individuos. La variable aleatoria Z_n representa el número de individuos en la n-ésima generación.
Sin pérdida de generalidad podemos suponer $Z_0 = 1$.

Teorema 2: Sean $\{Z_n\}_n$ un proceso de Galton-Watson, $m = E[Z_1]$ $\quad W_n := \frac{Z_n}{m^n}$ $\quad$ y $\mathcal{F}_n = \sigma(Z_0, Z_1, \ldots, Z_n)$. Si $0 < m < \infty$ entonces $\{W_n\}_n$ es una martingala con respecto a $\{\mathcal{F}_n\}_n$.
Más aún como $W_n \geq 0$ entonces existe una variable aleatoria no negativa W tal que con probabilidad 1,

$$\lim_{n\to\infty} W_n = W$$

Teorema 3: Sea $\{X_n\}_n$ una sucesión de variables aleatorias independientes e igualmente distribuidas con $\mu = E[X_1] < +\infty$.
Sea $S_n = \sum_{i=1}^{n} X_i$ $\quad$ entonces:

$$P\left(\lim_{n\to\infty} \frac{S_n}{n} = \mu\right) = 1$$

Demostraciones

Corolario 1 Como X_n es no negativo,

$$E[|X_n|] = E[X_n] = E[X_1]$$

Por lo tanto, del teorema 1 se concluye la existencia con probabilidad 1, del $\lim_{n\to\infty} X_n$. $\square$

Teorema 2 Por el Corolario 1, tenemos que sólo necesitamos probar que $\{W_n\}_n$ es una martingala.
En efecto:

$$\begin{aligned} E\left[W_{n+1} \mid \mathcal{F}_n\right] &= E\left[\frac{Z_{n+1}}{m^{n+1}} \mid \mathcal{F}_n\right] \\ &= \frac{1}{m^{n+1}} E\left[Z_{n+1} \mid \mathcal{F}_n\right] \\ &= \frac{1}{m^{n+1}} E[Z_{n+1} \mid Z_n] \quad \text{por ser } \{Z_n\} \text{ una cadena de Markov} \\ &= \frac{1}{m^{n+1}}.mZ_n = W_n \qquad \square \end{aligned}$$

La segunda aplicación que presentamos del teorema de convergencia para martingalas, es una demostración de la ley fuerte de lo grandes números.

Teorema 3 Supondremos que la función generadora de momentos de las variables aleatorias X_n existe.
Sea $\varepsilon > 0$ fijo pero arbitrario y consideremos la función

$$g(t) = e^{t(\mu+\varepsilon)}/m(t)$$

donde $m(t) = E[e^{tX_1}]$.

Es claro que:

$$\begin{aligned} g(0) &= 1 \\ g'(0) &= \left.\frac{m(t).(\mu+\varepsilon)e^{t(\mu+\varepsilon)} - e^{t(\mu+\varepsilon)}.m'(t)}{[m(t)]^2}\right|_{t=0} = \varepsilon > 0 \end{aligned}$$

Por lo tanto existe $t_0 > 0$ tal que $g(t_0) > 1$.
Vamos a probar que:

$$P\left(\frac{S_n}{n} > \mu + \varepsilon \,\text{para un número infinito de } n\right) = 0$$

y que

$$P\left(\frac{S_n}{n} \leq \mu - \varepsilon \,\text{para un número infinito de } n\right) = 0$$

Con lo cual se tendría que

$P(\mu - \varepsilon \leq \frac{S-n}{n} \leq \mu + \varepsilon$ para todos, salvo un número finito de n)= 1.

Tenemos

$$(*) \qquad \frac{S_n}{n} \geq \mu + \varepsilon \quad \to \quad \frac{e^{t_0 S_n}}{m^n(t_0)} \geq \left[\frac{e^{t_0(\mu+\varepsilon)}}{m(t_0)}\right]^n = [g(t_0)]^n$$

Es fácil verificar (ejercicio!) que

$$M_n := \frac{e^{t_0 S_n}}{m^n(t_0)} \qquad \text{es una martingala no negativa}$$

por lo tanto, por el teorema de convergencia para martingalas se tiene que, con probabilidad 1,

$$\lim_{n\to\infty} M_n \qquad \text{existe} \quad \text{y esto es finito.}$$

Puesto que $g(t_0) > 1$ se sigue de $(*)$ que

$$P\left(\frac{S_n}{n} > \mu + \varepsilon \,\text{para un número infinito de}\, n\right) = 0$$

De manera análoga se prueba que

$$P\left(\frac{S_n}{n} \leq \mu + \varepsilon \,\text{para un número infinito de}\, n\right) = 0. \qquad \square$$