Martingalas

Procesos Estocásticos

UC3M

Marzo 2012

Martingalas

Vamos a estudiar una clase de procesos que pueden verse como la fortuna de un jugador que juega repetidamente un *juego justo*. Así que, pensemos que M_n es la fortuna luego de n rondas.

Decimos que $M_0, M_1, \ldots$ es una **martingala** si para cualquier $n \geq 0$

1. $E|M_n| < \infty$
2. para cualquier sucesión de posibles valores $m_0, m_1, \ldots, m_n$

$$E[M_{n+1}|M_0 = m_0, M_1 = m_1, \ldots, M_n = m_n] = m_n$$

La segunda propiedad equivalente a

$$\begin{aligned} 0 &= E[M_{n+1} - m_n|M_0 = m_0, M_1 = m_1, \ldots, M_n = m_n] \quad (1) \\ &= E[M_{n+1} - M_n|M_0 = m_0, M_1 = m_1, \ldots, M_n = m_n] \end{aligned}$$

Es decir, condicionando al pasado, la ganancia neta esperada luego del turno siguiente es cero. Lo cual se entiende como que el juego es justo.

Esperanza condicional

Como el estudio de martingalas recae fuertemente en el concepto de esperanza condicional, es conveniente comprenderla en forma correcta.

En los cursos introductorios de probabilidad, si X, Y son variables aleatorias, la esperanza de X dado $Y = y$ se entiende como el valor esperado de la distribución de X dado $Y = y$,

$$E[X|Y = y] = \begin{cases} \sum_x x \; P(X = x|Y = y) & \text{caso discreto} \\ \int x \; f_{X|Y=y}(x|y)dx & \text{caso continuo} \end{cases}$$

En este curso, la **esperanza condicional de X dado Y** es una **variable aleatoria** que denotaremos por $E[X|Y]$, y que viene definido por

$$E[X|Y] = \psi(Y),$$

siendo

$$\psi(y) = E[X|Y = y].$$

Esperanza condicional con respecto a un evento

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y Y una variable aleatoria. Para cada $A \in \mathcal{F}$, la función

$$1_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

se llama función indicatriz del conjunto A.

Con la notación $E(Y, A)$ indicamos la esperanza de la variable aleatoria $Y 1_A$, i.e.

$$E(Y, A) = E(Y 1_A)$$

Note que si Y y A son independientes, es decir si para todo B se tiene $P(Y \in B, A) = P(Y \in B) P(A)$ entonces

$$E(Y|A) = E(Y)$$

Def. Sea $A \in \mathcal{F}$ un evento tal que $P(A) > 0$. La Esperanza Condicional con respecto a un evento A se define como

$$E(Y|A) = \frac{E(Y, A)}{P(A)}$$

Propiedades de la esperanza condicional con respecto a un evento

Propiedad 1 (Linealidad de la esperanza condicional)

$$E(\alpha X + Y|A) = \alpha E(X|A) + E(Y|A)$$

Propiedad 2 Si X es igual a una constante c en A entonces

$$E(XY|A) = cE(Y|A)$$

Propiedad 3 (desigualdad de Jensen) Si ϕ es convexa entonces

$$E(\phi(X)|A) \geq \phi(E(X|A))$$

Propiedad 4 Sea $A_1, \dots, A_k$ una familia disjunta de conjuntos tal que $B = \cup_{j=1}^{k} A_j$. Entonces

$$E(Y|B) = \sum_{j=1}^{k} E(Y|A_j) \frac{P(A_j)}{P(B)}$$

Esperanza condicional con respecto a una variable aleatoria

$$\psi(y) = E[X|Y = y]$$

es la función que asocia a cada posible valor y de Y la Esperanza Condicional respecto al evento $\{Y = y\}$. El evento puede ser de la forma

$$\{Y = y\} = \{Y_1 = y_1, \ldots, Y_n = y_n\}$$

así que que $\psi(y)$ puede tener la forma $\psi(y) = \psi(y_1, \ldots, y_n)$.

Mientras que $\psi(Y)$ es una variable aleatoria que hemos llamado esperanza condicional de X dado Y y convenido denotar por $E[X|Y]$. La variable Y puede ser n-dimensional, así que $E[X|Y]$ puede tener la forma

$$E[X|Y_1, \ldots, Y_n]$$

Hay un gran diferencia entre $E[X|Y = y]$ y $E[X|Y]$.

- $E[X|Y = y]$ es un número, es el valor esperado de la distribución condicional de X dado el evento $\{Y = y\}$
- $E[X|Y]$ es una variable aleatoria.

Propiedades Esperanza Condicional con respecto a variables aleatorias

P1 : Sean $\alpha, \beta \in \mathbb{R}$ y X, Y, Z variables aleatorias. Entonces

$$E[\alpha X + \beta Y|Z] = \alpha E[X|Z] + \beta E[Y|Z]$$

P2: Sean X, Y variables aleatorias. Entonces

$$E[XY|X] = XE[Y|X]$$

P3: Si ϕ es convexa entonces

$$E[\phi(Y)|X] = \phi(E[Y|X])$$

P4 : Sean X, Y, Z variables aleatorias. Entonces

$$E[E[X|Y,Z]|Y] = E[X|Y]$$

P5 : Sean X, Y variables aleatorias independientes. Entonces

$$E[X|Y] = E[X]$$

Martingalas: Definición 2

Usando el concepto revisado de esperanza condicional, volvemos a la definición de martingala:

Def. : Decimos que $M_0, M_1, \ldots$ es una Martingala si para cualquier $n \geq 0$

1. $E|M_n| < \infty$
2. $E[M_{n+1}|M_0, M_1, \ldots, M_n] = M_n$

Si

$$E(M_{n+1}|M_0, M_1, \ldots, M_n) \leq M_n$$

decimos entonces que $\{M_n\}$ es una Súpermartingala.

Si

$$E[M_{n+1}|M_0, M_1, \ldots, M_n] \geq M_n$$

decimos que es una Submartingala.

Ejemplo: Paseos Aleatorios

Sean $X_0, X_1, \ldots, X_n$ variables aleatorias i.i.d. tal que $E[X_1] = \mu$ y sean

$$\begin{aligned} M_0 &= X_0 \\ M_1 &= X_0 + X_1 \\ \vdots &= \vdots \\ M_n &= X_0 + X_1 + \cdots + X_n \end{aligned}$$

La secuencia de variables aleatorias M_n es llamada Paseo aleatorio y es una súpermartingala si $\mu \leq 0$, una martingala si $\mu = 0$ y una submartingala si $\mu \geq 0$.

Es fáxil demostrar lo anterior, simplemente hay que usar el hecho de que

$$M_{n+1} = M_n + X_{n+1}$$

y que M_n y X_{n+1} son independientes.

Ejemplo: Precio de acciones

Sean $Y_0, Y_1, \ldots, Y_n, \ldots$ variables aleatorias independientes y positivas. Suponemos que una acción tiene precio M_0 a tiempo 0.

Un modelo popular para modelar el precio de la acción a tiempo n es

$$M_{n+1} = M_n Y_n$$

donde $(Y_n - 1) \times 100$ representa (en porcentaje) la variabilidad de la acción.

Usando las propiedades de la esperanza condicional, es muy sencillo demostrar que

$$E[M_{n+1}|M_0, \ldots, M_n] = M_n E[Y_n]$$

En particular, si $Y_1, \ldots, Y_n$ son identicamente distribuidas con $E[Y_1] = \mu$, tenemos que M_n es

- una martingala si $\mu = 1$
- una submartingala si $\mu > 1$
- una súpermartingala si $\mu < 1$

Ejemplo : Precio de acciones

Dos ejemplos famosos del modelo anterior son:

- **Black-Scholes discreto**. $Y_1, \ldots, Y_n, \ldots$ definidas por

$$Y_n = e^{Z_n}$$

siendo $Z_1, Z_2, \ldots$ variables aleatorias independientes normales $N(\mu, \sigma^2)$.

- **Modelo Binomial**. $Y_1, \ldots, Y_n, \ldots$ definidas por

$$P\left(Y_i = (1+t)e^{-r}\right) = p \text{ y } P\left(Y_i = (1+t)^{-1}e^{-r}\right) = 1 - p,$$

La constante r es la tasa de interés y los factores $(1+t)$ y $1/(1+t)$ modelan las variaciones del mercado y garantizan que el precio tiene la forma $M_0(1+t)^y e^{-nr}$, con $|y| \leq n$. La volatilidad del precio está asociada a p.

Martingalas respecto a un proceso: Definición 3

Def.: Decimos que $M_0, M_1, \dots$ es una Martingala respecto a un proceso $X_0, X_1, \dots, X_n, \dots$ si para cualquier $n \geq 0$,

1. $E|M_n| < \infty$
2. $E[M_{n+1}|X_0, X_1, \dots, X_n] = M_n$ (o equivalentemente $E[M_{n+1} - M_n|X_0, X_1, \dots, X_n] = 0$)

Dos ejemplos de martingalas respecto a un proceso son la martingala cuadrática y la martingala exponencial.

Martingala cuadrática. Sean $X_0, X_1, X_2, \dots$ variables aleatorias independientes con $E[X_i] = 0$ y $E[X_i^2] = \sigma^2$. Considere el paseo $S_n = X_0 + X_1 + \cdots + X_n$, y

$$M_n = S_n^2 - n\sigma^2$$

Entonces M_n es una martingala con respecto a $S_0, S_1, \dots$.

Martingala exponencial. Sean $X_0, X_1, X_2, \dots$ variables aleatorias i.i.d. tal que $\varphi(\theta) = E[\exp(\theta Z_i)]$ y sea $S_n = X_0 + X_1 + \cdots + X_n$. Entonces

$$M_n = \frac{e^{\theta S_n}}{\varphi^n(\theta)}$$

es una martingala con respecto a $S_0, S_1, \dots$.

Martingalas respecto a Cadenas de Markov

Teorema 1 Sea $\{X_n\}$ una Cadena de Markov con espacio de estados S discreto y matriz de transición $P = \{p(i,j)\}_{i,j \in S}$. Sea $g : S \times \mathbb{N} \to \mathbb{R}$ tal que

$$g(i,n) = \sum_{j \in S} p(i,j) g(j, n+1)$$

Entonces $M_n = g(X_n, n)$ es una martingala.

Ejemplo : Sean $X_0, X_1, \ldots, X_n, \ldots$ independientes y tal que $P(X_i = 1) = p$ y $P(X_i = -1) = 1 - p$, $0 < p < 1$, $i = 1, \ldots$. Sea $S_n = X_0 + \cdots + X_n$ y sea

$$g(x) = \left(\frac{1-p}{p} \right)^x$$

Entonces, se puede demostrar con el teorema anterior que

$$M_n = g(S_n)$$

es una Martingala.

Propiedades elementales

Antes de discutir los resultados centrales de la teoría de martingalas, es conveniente aclarar algunos resultados elementales.

Recordemos que $\{X_n\}$ es una martingala (super-martingala) con respecto a $\{Y_n\}$ si

$$E[X_{n+k}|Y_0, \ldots, Y_n] = X_n \quad (resp. \leq)$$

para todo $k \geq 0$.

De manera que:

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

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

2. Si $\{X_n\}$ es una martingala con respecto a $\{Y_n\}$ y ϕ es una función convexa entonces $\{\phi(X_n)\}$ es una submartingala con respecto a $\{Y_n\}$.

Tiempos de parada

Def.: Decimos que la variable aleatoria T es un Tiempo de parada para el proceso $\{X_n\}$ si la ocurrencia o no del evento $\{T = n\}$ puede ser determinado conociendo sólo los valores $X_0, X_1, \ldots, X_n$ (no se requiere conocer ni X_{n+1}, ni $X_{n+2}, \ldots$).

Ejemplo. Si X_n es una CM que representa nuestro capital en euros luego de jugar n veces, el instante (aleatorio) T en el que por primera vez tenemos m euros es un tiempo de parada. De hecho

$$\{T = n\} = \{X_0 \neq m, \ldots, X_{n-1} \neq m, X_n = m\}$$

Hay muchos teoremas famosos sobre martingalas. Posiblemente el más famoso de ellos establece que

no se puede vencer un juego desfavorable

Y para formalizar este importante resultado requerimos el concepto de tiempo de parada.

Teorema del muestreo opcional

Teorema 2. Sea M_n una supermartingala con respecto a $X_1, \ldots, X_n, \ldots$ y T un tiempo de parada. Entonces, el proceso parado $M_{\min(T,n)}$ es una superrmartingala con respecto a $X_1, \ldots, X_n, \ldots$. En particular,

$$M_{\min(T,n)} \leq M_0$$

Un resultado análogo para submartingalas ($M_{\min(T,n)} \geq M_0$) y para martingalas ($M_{\min(T,n)} = M_0$) también vale.

El teorema puede resultar decepcionante para muchos jugadores. Si la estrategia del juego es parar, según un histórino determinado, en promedio no puedes vencer al casino. No importa tu estrategia de parada, el juego sigue siendo desfavorable.

Pero las aplicaciones del teorema pueden ser más sorpendentes.

Comportamiento asintótico

Teorema 4. Si M_n es una martingala tal que para cualquier n

$$E\,|M_n| \leq k$$

para alguna costante $k < \infty$ entonces $\lim_{n\to\infty} M_n$ existe y es una variable aleatoria M, tal que $P(M < \infty) = 1$.

El próximo Teorema es un Teorema Central del Límite para martingalas.

Teorema 5. Sea M_n una martingala con incrementos acotados, es decir, para todo n, $|M_{n+1} - M_n| < k$ para algún $k < \infty$. Sea

$$\sigma_n^2 = E[(M_{n+1} - M_n)^2 | M_0, M_1 \ldots, M_n]$$

y definamos $n(s) = \min\left\{n : \sum_{i=1}^n \sigma_i^2 \geq s\right\}$. Entonces

$$\lim_{s\to\infty} \frac{M_{n(s)}}{\sqrt{s}} \Rightarrow \mathcal{N}(0,1)$$

Desigualdad de Azuma

Uno de las teoremas más explotados sobre martingalas es la siguiente desigualdad de concentración:

Supongamos que

$$|M_{n+1} - M_n| < k_n$$

Entonces, para cualquier entero N y real $t > 0$ se cumple que

$$P(|M_n - E(M_n)| \geq t) \leq 2 \exp\left(\frac{-t^2}{2\sum_{n=1}^{N} k_n^2}\right)$$

Una aplicación de la desigualdad de Azuma al Paseo Aleatorio simple establece que

$$P(S_n > t) \leq \exp\left(\frac{-t^2}{2n}\right)$$