PROBABILIDAD II

Grado en Matemáticas

Tema 3
Dependencia: esperanza condicional, martingalas.

Jesús Munárriz

Departamento de Matemáticas
Universidad Autónoma de Madrid
jesus.munarriz@uam.es

1. Esperanza condicional.
2. Martingalas.
3. Aplicaciones.

Definición: Esperanza condicional

Sea $(\Omega,\mathcal{F},P)$ un espacio de probabilidad, sea X una variable aleatoria con $E|X| < \infty$ y sea $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$. Entonces **la** esperanza condicional $E(X|\mathcal{G})$ es **la** variable aleatoria (única módulo conjuntos de medida cero) que satisface :

(1) $E(X|\mathcal{G})$ es $\mathcal{G}$-medible.

(2) Para cada conjunto $G \in \mathcal{G}$ tenemos:

$$\int_G E(X|\mathcal{G})dP = \int_G XdP.$$

Hemos visto que cuando $\mathcal{G}$ está generada por una partición de Ω, $E(X|\mathcal{G})$ se obtiene simplemente tomando el promedio de X sobre cada conjunto en la partición.

Cuando $E(X|\mathcal{G})$ existe, las siguientes propiedades son inmediatas a partir de la definición:

(1) Linealidad.

(2) Positividad.

(3) $E(E(X|\mathcal{G})) = E(X)$. Dem: tomamos $G = \Omega$ en (2) de la def.

¿Cuando existe $E(X|\mathcal{G})$?

Queremos probar que $E(X|\mathcal{G})$ existe para toda $X \in L^1(\Omega, \mathcal{F}, P)$ y toda sub-σ-álgebra de $\mathcal{G}$ de $\mathcal{F}$. Existen dos posibilidades.

La primera es definir directamente la esperanza condicional en $L^1(\Omega, \mathcal{F}, P)$, usando el Teorema de Radon-Nikodym: dada $X \in L^1(\Omega, \mathcal{F}, P)$, para todo $G \in \mathcal{G}$ definimos

$$\nu(G) = \int_G X \, dP.$$

Como ν es una medida con signo, absolutamente continua con respecto a la restricción de P a $\mathcal{G}$, existe una única función $\mathcal{G}$-medible $Y \in L^1(\Omega, \mathcal{G}, P)$ tal que para todo $G \in \mathcal{G}$,

$$\nu(G) = \int_G Y \, dP.$$

Por definición, $Y = E(X|\mathcal{G}) = d\nu/dP$.

La segunda posibilidad es definir $E(X|\mathcal{G})$ como la proyección ortogonal π de $L^2(\Omega, \mathcal{F}, P)$ sobre el subespacio cerrado $L^2(\Omega, \mathcal{G}, P)$.

Como las funciones simples son densas en L^1, también lo es $L^\infty \subset L^2 \subset L^1$. Como las proyecciones son lineales, $E(\cdot|\mathcal{G})$ se extiende a L^1 por linealidad y densidad.

Ciertas propiedades de las integrales se extienden de manera más o menos inmediata a la esperanza condicional. Como ésta ha sido definida en L^1, en general hay que asumir que los límites de funciones están en L^1.

Teorema: Convergencia Monótona de la Esperanza Condicional

Sea $\{X_n\}_{n\geq 1}$ una sucesión de variables aleatorias tales que $0 \leq X_n \uparrow X \in L^1$. Entonces

$$\lim_n E(X_n|\mathcal{G}) = E(\lim_n X_n|\mathcal{G}) = E(X|\mathcal{G}).$$

Teorema: Fatou Condicional

Si $0 \leq X_n \in L^1$ y $\liminf X_n \in L^1$, entonces c.s.

$$E(\liminf X_n|\mathcal{G}) \leq \liminf E(X_n|\mathcal{G}).$$

Teorema: Convergencia Dominada Condicional

Sea $\{X_n\}_{n\geq 1}$ una sucesión de variables aleatorias para las que existe una $\mathrm{V} \in L^1$ con $|X_n| \leq \mathrm{V}$. Si $X = \lim_n X_n$ c.s., entonces

$$\lim_n E(X_n|\mathcal{G}) = E(\lim_n \mathrm{X}_n|\mathcal{G}) = E(X|\mathcal{G}).$$

Teorema: Jensen Condicional

Sea $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ una función convexa y sea X una variable aleatoria tal que $X, \varphi(X) \in L^1$. Entonces c.s.

$$\varphi(E(X|\mathcal{G})) \leq E(\varphi(X)|\mathcal{G}).$$

Dem: Sea $L_{(a,b)}(x) = ax + b$, y sea $A = \{(a,b) \in \mathbb{R}^2 : L_{(a,b)} \leq \varphi\}$. Por convexidad, $\varphi = \sup\{L_{(a,b)} : (a,b) \in A\}$, luego

$$\varphi(E(X|\mathcal{G})) = \sup_{(a,b)\in A} \{L_{(a,b)}(E(X|\mathcal{G}))\} = \sup_{(a,b)\in A} \{aE(X|\mathcal{G}) + b\}$$

$$= \sup_{(a,b)\in A} \{E(aX + b|\mathcal{G})\} \leq E(\varphi(X)|\mathcal{G}),$$

ya que $aX(\omega) + b \leq \varphi(X(\omega))$, para todo $(a,b) \in A$ y todo $\omega \in \Omega$.

Comentario: tomando $\varphi(\cdot) = |\cdot|$, Jensen condicional nos dice que $|E(X|\mathcal{G})| \leq E(|X||\mathcal{G})$. Es un sencillo ejercicio probar dicha desigualdad directamente. Otra consecuancia inmediata de Jensen condicional es el siguiente

Corolario: $E(\cdot|\mathcal{G})$ es una contracción en L^p, $1 \leq p \leq \infty$

Sea $1 \leq p \leq \infty$ y sea $X \in L^p$. Entonces

$$\|E(X|\mathcal{G})\|_p \leq \|X\|_p \quad \textit{para} \quad p \geq 1.$$

Dem: El caso $p = \infty$ es bastante obvio. Para los demás, por Jensen condicional tenemos $|E(X|\mathcal{G})|^p \leq E(|X|^p|\mathcal{G})$. A continuación integramos y tomamos raíces p-ésimas.

Proposición. Sea X una variable aleatoria con $E|X| < \infty$. Si X es $\mathcal{G}$-medible, entonces $E(X|\mathcal{G}) = X$.

Teorema: Propiedad de la Torre

Si $\mathcal{H}$ es una sub-σ-álgebra de $\mathcal{G}$ y $X \in L^1$, entonces

$$E[E(X|\mathcal{G})|\mathcal{H}] = E(X|\mathcal{H}) = E[E(X|\mathcal{H})|\mathcal{G}]$$

Dem: la primera igualdad es inmediata a partir de la definición de esperanza condicional. La segunda es un caso especial de la proposición anterior.

Sabemos que si c es una constante, entonces $E(cX) = cE(X) < \infty$. El siguiente resultado generaliza este hecho a sub-σ-álgebras no necesariamente triviales.

Teorema

Si $X, Z \in L^1$ y Z es $\mathcal{G}$-medible, entonces $E(\mathrm{Z}X|\mathcal{G}) = \mathrm{Z}E(X|\mathcal{G})$.

Dem: mediante las reducciones habituales, basta considerar el caso $Z = \mathbf{1}_A$, donde $A \in \mathcal{G}$. Pero entonces el resultado es inmediato a partir de la definición de esperanza condicional.

En el caso de la independencia, la nueva información es irrelevante.

Teorema

Si la v.a. $X \in L^1$ es independiente de $\mathcal{G}$, entonces $E(X|\mathcal{G}) = E(X)$.

Dem: por definición de esperanza condicional, basta probar que para todo $G \in \sigma(\mathcal{G})$

$$\int_G E(X|\sigma(\mathcal{G}))dP = \int_G E(X)dP,$$

lo que se deduce de la independencia y del resultado anterior.

Definición: una **filtración** es cualquier familia creciente $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq ... \subseteq \mathcal{F}$ de sub-σ-álgebras de $\mathcal{F}$. Escribimos

$$\mathcal{F}_\infty := \sigma\left(\cup_n \mathcal{F}_n\right) \subseteq \mathcal{F}.$$

Llamamos **procesos estocásticos** a colecciones de variables aleatorias indexadas por un conjuntos T, que generalmente se interpretan como "el tiempo". Para nosotros, típicamente $T = \mathbb{N}$ ó $T = \mathbb{N} \setminus \{0\}$ ó $T = -\mathbb{N}$. No consideraremos procesos estocásticos en tiempo continuo (típicamente $T = [0, \infty)$).

Definición: un proceso $X = \{X_n\}_{n\geq 0}$ se dice **adaptado** a la filtración $\{\mathcal{F}_n\}_{n\geq 0}$ si para cada n, la v.a. X_n es $\mathcal{F}_n$-medible.

Interpretamos $\mathcal{F}_n$ como la información disponible en tiempo n. Tal información crece con n si tenemos una filtración, y captura todo lo que necesitamos saber sobre $X_0, \ldots, X_n$ si el proceso es adaptado. Con frecuencia tomamos $\mathcal{F}_n := \sigma(X_0, \ldots, X_n)$.

Definición: un proceso $C = \{C_n : n \in \mathbb{N} \setminus \{0\}\}$ es **previsible** o **predecible** si C_n es $\mathcal{F}_{n-1}$-medible.

Ejemplo: en un juego de azar, la decisión de apostar o no en tiempo n, y cuanto apostar, debemos tomarla con la información disponible en tiemplo $n - 1$.

Las martingalas modelizan matemáticamente la noción de "juego de azar justo" o equilibrado.

Definición: Martingala

Un proceso $X = \{X_n\}_{n\geq 0}$ es una **martingala** relativa a la filtración $\{\mathcal{F}_n\}_{n\geq 0}$ si:

(i) X es adaptado, es decir, para cada n la v.a. X_n es $\mathcal{F}_n$-medible.

(ii) Para toda $n \geq 0$, $E|X_n| < \infty$. Es decir, $X_n \in L^1$.

(iii) $E(X_n|\mathcal{F}_{n-1}) = X_{n-1}$ para toda $n \geq 1$.

La propiedad (iii) de las martingalas es la realmente definitoria, la que captura la idea de juego justo.

Las submartingalas modelizan la noción de "juego de azar favorable".

Definición: una **submartingala** se define de manera idéntica a la de martingala, salvo que la propiedad (iii) de las martingalas se reemplaza por

$(iii)'$ $E(X_n|\mathcal{F}_{n-1}) \geq X_{n-1}$ para toda $n \geq 1$.

Las supermartingalas modelizan la noción de "juego de azar desfavorable".

Definición: una **supermartingala** se define de manera idéntica a la de martingala, salvo que la propiedad (iii) de las martingalas se reemplaza por

$(iii)''$ $E(X_n|\mathcal{F}_{n-1}) \leq X_{n-1}$ para toda $n \geq 1$.

Ejemplos: 1) Suma de variables aleatorias independientes con media 0.

Sean X_1, $X_2, \ldots$ variables aleatorias independientes con $E(|X_k|) < \infty$ y $E(X_k) = 0$, $\forall k$. Tomamos $\mathcal{F}_0 := \{\emptyset, \Omega\}$ y $\mathcal{F}_n := \sigma\left(X_1, X_2, \cdots, X_n\right)$.

Las v.a.'s $S_0 := 0$ y $S_n := X_1 + X_2 + \cdots + X_n$ están en L^1 y cada S_n es $\mathcal{F}_n$-medible. $S = \{S_n\}_{n\geq 0}$ es una martingala: para $n \geq 1$

$$E(S_n|\mathcal{F}_{n-1}) = E(S_{n-1}|\mathcal{F}_{n-1}) + E(X_n|\mathcal{F}_{n-1}) = S_{n-1} + E(X_n) = S_{n-1}.$$

2) Acumulación de información sobre una variable aleatoria.

Sea $\{\mathcal{F}_n\}$ una filtración y sea $\zeta \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$. Definamos $M_n := E(\zeta|\mathcal{F}_n)$. Por la Propiedad de la Torre, tenemos

$$E(M_n|\mathcal{F}_{n-1}) = E(E(\zeta|\mathcal{F}_n)|\mathcal{F}_{n-1}) = E(\zeta|\mathcal{F}_{n-1}) = M_{n-1}.$$

Por tanto $M = \{M_n\}_{n\geq 0}$ es una martingala.

Comentario: si X es una submartingala, entonces $-X$ es una supermartingala. Si X es una submartingala y una supermartingala, entonces , entonces X es una martingala.

Proposición: Sea $Y = \{Y_n\}_{n\geq 0}$ una martingala con respecto a $\{\mathcal{F}_n\}_{n\geq 0}$. Entonces:
a) $E(Y_{n+m}|\mathcal{F}_n) = Y_n$ para todo $n, m \geq 0$.
b) $E(Y_m) = E(Y_0)$ para todo m.

Dem: a) se obtiene a partir de la propiedad de la torre aplicando la propiedad de las martingalas m veces, y b), usando a) con $n = 0$ y tomando esperanzas.

Ejercicio: Enunciar y probar el resultado análogo para submartingalas y supermartingalas.

La decisión de dejar de participar en un juego aleatorio, ¿en qué información puede basarse?

Definición: Tiempo de parada

Sea $T : \Omega \to \{0, 1, 2, ...\} \cup \{\infty\}$ una variable aleatoria. Decimos que T es un tiempo de parada con respecto a la filtración $\{\mathcal{F}_n\}$, si para todo $n \geq 0$, el evento $\{T = n\} \in \mathcal{F}_n$, o equivalentemente, si

$$\{T \leq n\} = \{\omega \in \Omega : T(\omega) \leq n\} \in \mathcal{F}_n, \ \forall n \leq \infty.$$

Ejemplo: Tiempo de primera visita a un conjunto.

Sea $\{Y_n\}_{n\geq 0}$ un proceso adaptado a la filtración $\{\mathcal{F}_n\}_{n\geq 0}$, y sea $[a,b]\subset\mathbb{R}$. Entonces

$$T_1 = \text{ínf}\{n : Y_n \in [a,b]\}$$

es un tiempo de parada con respecto a la filtración $\{\mathcal{F}_n\}$. Basta observar observar que para cada n,

$$\{T_1 = n\} = \{Y_0 \notin [a,b]\} \cap ... \cap \{Y_{n-1} \notin [a,b]\} \cap \{Y_n \in [a,b]\}.$$

¿Existen estrategias que nos permitan transformar un juego equilibrado en uno favorable? ¿o uno desfavorable en uno equilibrado?

Definición: Proceso parado en T

Sea T un tiempo de parada para el proceso $\{Y_n\}_{n\geq 0}$, y sea $T \wedge n := \min\{T, n\}$. Al proceso $\{Y_{T\wedge n}\}_{n\geq 0}$ dado por

$$Y_{T\wedge n} := \begin{cases} Y_n & \text{si } n \leq T \\ Y_T & \text{si } n > T \end{cases}$$

se le denomina **proceso parado en** T.

Proposición: Sea $\{Y_n\}_{n\geq 0}$ un proceso estocástico adaptado a la filtración $\{\mathcal{F}_n\}_{n\geq 0}$, con $Y_n \in L^1\ \forall n$, y sea T un tiempo de parada con respecto a $\{\mathcal{F}_n\}_{n\geq 0}$. Entonces $\{Y_{T\wedge n}\}_{n\geq 0}$ es un proceso adaptado a $\{\mathcal{F}_n\}_{n\geq 0}$, y $\forall n$, $E(|Y_{T\wedge n}|) < \infty$.

Dem.: es fácil ver que $\{Y_{T\wedge n}\}_{n\geq 0}$ es un proceso adaptado a $\{\mathcal{F}_n\}_{n\geq 0}$, y además, $E(|Y_{T\wedge n}|) \leq \sum_{i=0}^{n} E(|Y_i|) < \infty$.

Ejercicio: convencerse de que el proceso parado en T es adaptado y de que $|Y_{T\wedge n}| \leq \sum_{i=0}^{n} |Y_i|$.

Preview of coming attractions: veremos que bajo *hipótesis razonables*, no existen estrategias que nos permitan convertir un juego desfavorable en uno equilibrado, o éste en uno favorable.

El siguiente teorema nos dice que usar un tiempo de parada no altera la naturaleza, favorable, equilibrada o desfavorable, del juego.

Teorema de Parada Opcional de Doob

Sea T un tiempo de parada para el proceso $\{Y_n\}_{n\geq 0}$ con respecto a la filtración $\{\mathcal{F}_n\}$.

Si $\{Y_n\}_{n\geq 0}$ es una martingala (respectivamente una submartingala, o respectivamente, una supermartingala) entonces el proceso $\{Y_{T\wedge n}\}_{n\geq 0}$ también es una martingala (respectivamente una submartingala, o respectivamente, una supermartingala).

Sugerencia: aprenderse el resultado y su prueba (incluyendo (i) y (ii), no solo (iii)).

Dem.: Basta probarlo para submartingalas (¿por qué?), es decir, si $\{Y_n\}_{n\geq 0}$ es una submartingala, entonces

$$E(Y_{T\wedge n}|\mathcal{F}_{n-1}) \geq Y_{T\wedge(n-1)}.$$

El proceso

$$\mathbf{1}_n = \mathbf{1}_{\{T\geq n\}} = \begin{cases} 1 & \text{si } n \leq T \\ 0 & \text{si } n > T \end{cases}$$

es previsible:

$$\mathbf{1}_n^{-1}\{1\} = \{T \geq n\} = \{T < n\}^c = \{T \leq n-1\}^c \in \mathcal{F}_{n-1}.$$

Como $Y_{T\wedge n} = Y_{T\wedge(n-1)} + \mathbf{1}_n \cdot (Y_n - Y_{n-1})$, tenemos

$$\begin{aligned} E(Y_{T\wedge n}|\mathcal{F}_{n-1}) &= E(Y_{T\wedge(n-1)} + \mathbf{1}_n\ (Y_n - Y_{n-1})|\mathcal{F}_{n-1}) \\ &= Y_{T\wedge(n-1)} + \mathbf{1}_n\ E(Y_n - Y_{n-1}|\mathcal{F}_{n-1}) \\ &\geq Y_{T\wedge(n-1)}. \end{aligned}$$

Al siguiente teorema también se le conoce como teorema de parada opcional de Doob. Incide en la idea de que bajo condiciones razonables, utilizar un tiempo de parada no nos permite adquirir ventaja.

Teorema (parada opcional de Doob)

Sea $Y = \{Y_n\}_{n\geq 0}$ una martingala (resp. sub., resp. super.) con respecto a la filtración $\{\mathcal{F}_n\}_{n\geq 0}$, y sea T un tiempo de parada (con respecto a $\{\mathcal{F}_n\}_{n\geq 0}$). Entonces

$$Y_T \ \text{ es integrable y } \ E(Y_T) = E(Y_0)$$

(resp. $\geq$, resp. $\leq$) si al menos una de las siguientes condiciones se cumple:
a) $\|T\|_\infty < \infty$.
b) $P(T < \infty) = 1$ y $\sup_n \|Y_n\|_\infty < \infty$.
c) $E(T) < \infty$ y $\sup_n \|Y_{n+1} - Y_n\|_\infty < \infty$.

Dem.: Basta probarlo para submartingalas. Suponemos que o a) se cumple, o b) se cumple, o c) se cumple, en ese orden.
a) Sea N un número natural tal que $T \leq N$ c.s.; entonces

$$E(|Y_T|) = E(|Y_{T\wedge N}|) \leq \sum_{i=0}^{N} E(|Y_i|) < \infty.$$

El el proceso $\{Y_{T\wedge n}\}_{n\geq 0}$ es una submartingala por el Teorema de Parada Opcional, luego $E(Y_{\mathrm{T\wedge N}}) \geq E(Y_0)$.

b) Suponemos que $P(T<\infty)=1$ y $K := \sup_n \|Y_n\|_\infty < \infty$. Como $P(T<\infty)=1$, $\text{lím}_n Y_{T\wedge n} = Y_T$ c.s., y como para todo n, $|Y_n| \leq K$, tenemos que $|Y_{T\wedge n}| \leq K$, luego por el Teor. de la convergencia dominada de Lebesgue (TCDL),

$$E(Y_0) \leq \lim_n E(Y_{T\wedge n}) = E(Y_T) \quad \text{y} \quad E(|Y_T|) \leq K.$$

c) Suponemos que $E(T) < \infty$ y $K := \sup_n \|Y_{n+1} - Y_n\|_\infty < \infty$. Como en b) tenemos que $P(T < \infty) = 1$ y por tanto $\lim_n Y_{T\wedge n} = Y_T$ c.s.; si el TCDL es aplicable, entonces

$$E(Y_0) \leq \lim_n E(Y_{T\wedge n}) = E(Y_T) \quad \text{y} \quad E(|Y_T|) < \infty.$$

Para justificar el uso del TCDL, notamos que

$$|Y_{T\wedge n}| \leq |Y_0| + |Y_{T\wedge n} - Y_0| \leq |Y_0| + \sum_{i=1}^{T\wedge n} |Y_i - Y_{i-1}|$$

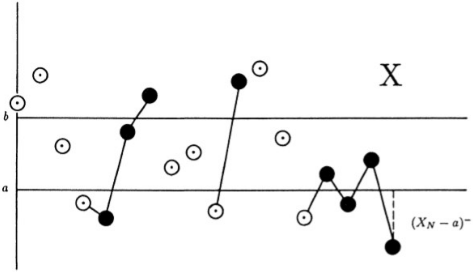
$$\leq |Y_0| + K \cdot (T \wedge n) \leq |Y_0| + KT \in L^1.$$

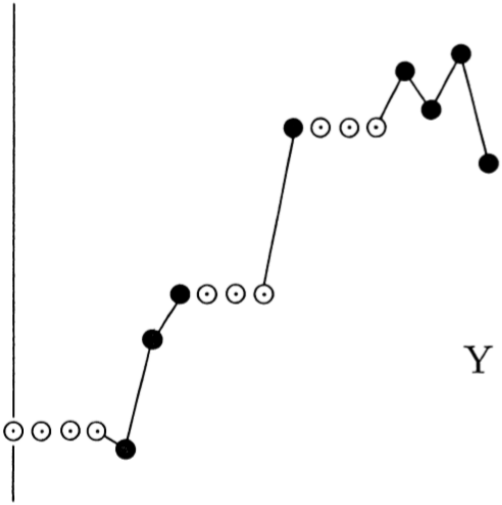
Teorema de la Convergencia de las Martingalas de Doob

Sea X una martingala (o una submartingala, o una supermartingala) tal que $\sup_n E(|X_n|) < \infty$. Entonces $\lim X_n$ existe casi seguro, y además, $P(|\lim X_n| < \infty) = 1$.

Antes de demostrar el resultado, mencionamos la idea de la prueba: si dados $a < b$ un para conjunto de medida positiva de trayectorias cruzamos de a a b infinitas veces, podemos diseñar una estrategia que nos garantice ganancias infinitas, lo que lleva a una contradicción. Tomando a muy próximo a b, concluimos que las oscilaciones son pocas y pequeñas.

Ilustraciones tomadas del libro *Probability with Martingales*, Cambridge University Press, 1991, de David Williams.




Y

Sea $X = \{X_n\}_{n\geq 0}$ un proceso que contabiliza nuestras "ganancias" (pueden ser negativas) en tiempo n, en un juego de azar en el que apostamos 1 en cada partida. La diferencia $X_n - X_{n-1}$ representa la ganancia en la partida n-ésima.

Supongamos ahora que decidimos apostar cantidades quizá distintas de 1. Si C_n representa la apuesta en la n-ésima partida, entonces $C_n(X_n - X_{n-1})$ nos proporciona las ganancias en la partida n-ésima, y por tanto las ganancias totales hasta ese momento son $Y_0 = 0$, y en tiempo n,

$$Y_n = \sum_{1\leq k\leq n} C_k(X_k - X_{k-1}).$$

Observación: la apuesta C_n debe realizarse con la información dada por $X_1, \ldots, X_{n-1}$, pero sin conocer X_n. Es decir, el proceso $\{C_n\}_{n\geq 1}$ debe ser previsible.

Teorema: transformadas de martingalas. Muestreo Opcional de Doob

Sea $X = \{X_n\}_{n\geq 0}$ una martingala (resp. sub., resp. super.) con respecto a la filtración $\{\mathcal{F}_n\}_{n\geq 0}$, y sea $C = \{C_n\}_{n\geq 1}$ un proceso previsible con $C_n \in L^\infty$ para todo $n \geq 1$. Entonces el proceso $Y = \{Y_n\}_{n\geq 0}$ definido por

$$Y_n = \sum_{1\leq k\leq n} C_k(X_k - X_{k-1})$$

es una martingala (resp. sub., resp. super.).

Observación: asumimos que las apuestas C_n están acotadas para asegurar la integrabilidad. Podemos modificar esta condición, suponiendo que $C_n \in L^p$, si exigimos que $X_j \in L^q$, donde p y q son exponentes conjugados.

Dem.: Sea X es una supermartingala adaptada a $\{\mathcal{F}_n\}$. Al ser C previsible,

$$\begin{aligned}
E(Y_n|\mathcal{F}_{n-1}) &= \sum_{1\leq k\leq n} E(C_k(X_k - X_{k-1})|\mathcal{F}_{n-1}) \\
&= \sum_{1\leq k\leq n} C_k E(X_k - X_{k-1}|\mathcal{F}_{n-1}) \\
&= \sum_{1\leq k\leq n} C_k(E(X_k|\mathcal{F}_{n-1}) - X_{k-1}) \\
&\leq \sum_{1\leq k< n} C_k(E(X_k|\mathcal{F}_{n-1}) - X_{k-1}) \\
&= \sum_{1\leq k< n} C_k(X_k - X_{k-1}) = Y_{n-1}.
\end{aligned}$$

Si X es una submartingala o una martingala, el resultado puede reducirse al caso anterior, o bien puede probarse del mismo modo.

Definimos $C_1 := \mathbf{1}_{\{X_0 < a\}}$ y, para $n \geq 2$,

$$C_n := \mathbf{1}_{\{C_{n-1}=1\}}\mathbf{1}_{\{X_{n-1}\leq b\}} + \mathbf{1}_{\{C_{n-1}=0\}}\mathbf{1}_{\{X_{n-1}<a\}}.$$

Sea s_i la i-ésima vez que empezamos a jugar bajo dicha estrategia previsible, y sea t_i la i-ésima vez que paramos de jugar. Tenemos que

$$X_{s_i}(\omega) < a, \quad X_{t_i}(\omega) > b \quad (1 \leq i \leq k).$$

El número de cruzamientos ascendentes (up)

$$U_N[a,b](\omega)$$

del intervalo $[a,b]$ realizado por la trayectoria $n \mapsto X_n(\omega)$ hasta el tiempo N, es el mayor $k \in \mathbb{Z}^+$ tal que

$$0 \leq s_1 < t_1 < s_2 < t_2 < \ldots < s_k < t_k \leq N.$$

Por tanto, en tiempo N tenemos

$$Y_N(\omega) \geq (b-a)U_N[a,b](\omega) - (X_N(\omega)-a)^-.$$

Lema de Doob: Cruzamientos

Sea X una supermartingala y sea $U_N[a,b]$ el número de cruzamientos ascendentes del intervalo $[a,b]$ en tiempo N. Entonces

$$(b-a)E(U_N[a,b]) \leq E((X_N-a)^-).$$

Dem.: sabemos que la transformada $\{Y_n\}_{n=0}^{\infty}$ dada por $Y_0 = 0$, $Y_n = \sum_{1\leq k\leq n} C_k(X_k - X_{k-1})$ es una supermartingala, luego

$$E(Y_N) \leq E(Y_0) = 0.$$

Por tanto, integrando ambos lados de

$$Y_N(\omega) \geq (b-a)U_N[a,b](\omega) - (X_N(\omega)-a)^-$$

y usando $E(Y_N) \leq E(Y_0) = 0$ hallamos la cota

$$0 \geq E(Y_N) \geq E((b-a)U_N[a,b]) - E((X_N-a)^-);$$

reordenando términos tenemos que

$$E((X_N-a)^-) \geq (b-a)E(U_N[a,b]).$$

$\square$

Del lema de los cruzamientos deducimos el siguiente

Lema

Sea X una supermartingala tal que $\sup_n E(|X_n|) < \infty$. Dados $a < b$ y $U_\infty[a,b] := \lim_N U_N[a,b]$, tenemos que $P(U_\infty[a,b] = \infty) = 0$.

Dem.: sea $K := |a| + \sup_n E(|X_n|) < \infty$. Por el lema de los cruzamientos,

$$(b-a)E(U_N[a,b]) \leq E((X_N - a)^-) \leq E(|X_N|) + |a| \leq K.$$

Por el Teorema de la Convergencia monótona,

$$(b-a)E(U_\infty[a,b]) = (b-a)\lim_N E(U_N[a,b]) \leq K < \infty,$$

luego $P(U_\infty[a,b] = \infty) = 0$. $\square$

Recordamos el

Teorema de la Convergencia de las Martingalas de Doob

Sea X una martingala (o una submartingala, o una supermartingala) tal que $\sup_n E(|X_n|) < \infty$. Entonces $\lim X_n$ existe casi seguro, y además, $P(|\lim X_n| < \infty) = 1$.

Dem.: Dados a y b con $a < b$, sea

$$\Lambda_{a,b} := \{\omega \in \Omega : \liminf X_n(\omega) < a < b < \limsup X_n(\omega)\},$$

y sea

$$\Lambda := \{\omega \in \Omega : X_n(\omega) \text{ no tiene límite en } [-\infty, \infty]\}.$$

Entonces

$$\Lambda = \bigcup_{\{a,b \in \mathbb{Q} : a < b\}} \Lambda_{a,b}.$$

Como

$$\Lambda_{a,b} \subseteq \{\omega : U_\infty[a,b](\omega) = \infty\},$$

concluimos que $P(\Lambda_{a,b}) = 0$ y por tanto, $P(\Lambda) = 0$.
Definiendo

$$X_\infty(\omega) := \liminf X_n(\omega),$$

tenemos que $\lim X_n$ existe y es igual a X_∞ casi seguramente.
Por Fatou,

$$E(|X_\infty|) = E(\liminf |X_n|) \leq \liminf E(|X_n|) \leq \sup E(|X_n|) < \infty,$$

luego $P(|X_\infty| < \infty) = 1$. $\square$

Consideramos filtraciones $\{\mathcal{F}_n\}_{n\leq 0}$ indexadas hacia atrás: $\mathcal{F}_0 \supset \mathcal{F}_{-1} \supset \mathcal{F}_{-2} \supset \dots$, y definimos $\mathcal{F}_{-\infty} := \cap_{n\leq 0}\mathcal{F}_n$.

Definición: Martingala reversa

Un proceso $X = \{X_n\}_{n\geq 0}$ es una **martingala reversa** relativa a la filtración $\{\mathcal{F}_n\}_{n\geq 0}$ si:

(i) X es adaptado, es decir, para cada n la v.a. X_n es $\mathcal{F}_n$-medible.

(ii) $X_0 \in L^1$.

(iii) $E(X_n|\mathcal{F}_{n-1}) = X_{n-1}$ para toda $n \geq 1$.

N. B.: Para toda $n \leq 0$, $E(X_0|\mathcal{F}_n) = X_n$. Luego $E|X_n| = E|E(X_0|\mathcal{F}_n)| \leq E\left(E(|X_0||\mathcal{F}_n)\right) = E|X_0| < \infty$.

Teorema: Convergencia de las Martingalas Reversas

Sea $X = \{X_n\}_{n\geq 0}$ una martingala reversa. Entonces,

$$X_{-\infty} := \lim_{n\to-\infty} X_n \text{ existe casi seguramente.}$$

Comentario: puede demostrarse que adicionalmente

$$\lim_{n\to-\infty} X_n = X_{-\infty} = E(X_0|\mathcal{F}_{-\infty}) \text{ c.s. y en } L^1.$$

La razón radica en que para todo $n \leq 0$, tenemos $E(X_0|\mathcal{F}_n) = X_n$. Como la esperanza condicional es una contracción, las martingalas reversas tienen una propiedad de compacidad en L^1, denominada *integrabilidad uniforme*, en la que no entraremos, pero que explica la convergencia en L^1.

Dem.: para N arbitrario, aplicamos el Lema de los Cruzamientos de Doob a la martingala (directa y finita) $X_{-N}, X_{-N+1}, \dots, X_0$, obteniendo la cota

$$(b-a)E(U_N[a,b]) \leq E((X_0-a)^-).$$

Tomando $N \to \infty$ concluimos que $P(U_\infty[a,b]=\infty)=0$, y el resto de la prueba de la convergencia c.s. es como en el caso anterior.

Aplicaciones: variables aleatorias con varianza finita

Ley Fuerte de los Grandes Números de Kolmogorov, para $X_i \in \mathcal{L}^2$

Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes con varianzas finitas, las cuales satisfacen la condición de Kolmogorov

$$\sum_{k=1}^{\infty} \frac{Var(X_k)}{k^2} < \infty.$$

Entonces para casi todo $\omega \in \Omega$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (X_i - E(X_i))(\omega) = 0.$$

Comentario: La hipótesis $X_i \in \mathcal{L}^2$ facilita la demostración de la convergencia. Por ello, no es necesario pedir que las variables estén idénticamente distribuidas. El caso $X_i \in \mathcal{L}^1$ es muy distinto.

Dem.: usamos el siguiente caso especial del Lema de Kronecker: sea $\{x_n\}_{n\geq 1}$ una sucesión de números reales. Si la serie $\sum_{i=1}^{\infty} \frac{x_i}{i}$ es convergente, entonces $\lim_n \frac{1}{n}\sum_{i=1}^{n} x_i = 0$.

Definimos $Y_i := X_i - E(X_i)$, $M_n = \sum_{i=1}^{n} \frac{Y_i}{i}$ y $\mathcal{F}_n = \sigma(Y_1, ..., Y_n)$. Entonces $M = \{M_n\}_{n=1}^{\infty}$ es una martingala con $E(M_n) = 0$. Por Doob y Kronecker, para casi todo $\omega \in \Omega$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} Y_i(\omega) = 0.$$

Nos falta justificar que la hipótesis del TCMD se cumple, es decir, que $\sup_n E|M_n| < \infty$, o equivalentemente, que $\sup_n (E|M_n|)^2 < \infty$.

Para toda $n \geq 1$,

$$(E|M_n|)^2 \leq E(|M_n|^2) = E(M_n^2) - (EM_n)^2 = Var(M_n).$$

Por independencia,

$$Var(M_n) = \sum_{i=1}^{n} \frac{Var(\mathrm{Y}_i)}{i^2} \leq \sum_{i=1}^{\infty} \frac{Var(\mathrm{X}_i)}{i^2} < \infty.$$

$\square$

Ley Fuerte de los Grandes Números de Kolmogorov, para $X_i \in \mathcal{L}^1$

Sea $\{X_n\}_{n=0}^{\infty}$ una sucesión de variables aleatorias independientes idénticamente distribuidas. Entonces para casi todo $\omega \in \Omega$,

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{i=0}^{n} X_i(\omega) = E(X_0).$$

Comentario: por distribución idéntica, todas las medias son iguales: $E(X_0) = E(X_1) = E(X_2) = \cdots$. Usamos el Teorema de la Convergencia de las Martingalas Reversas y la Ley 0-1 de Kolmogorov.

Dem.: para $n \geq 0$, sea $S_n := \sum_{i=0}^{n} X_i$ y sea
$\mathcal{F}_{-n} = \sigma(S_n, X_{n+1}, X_{n+2}, \dots) = \sigma(S_n, S_{n+1}, S_{n+2}, \dots)$.

Ejercicio: sea $M_{-n} = \frac{1}{n+1} \sum_{i=0}^{n} X_i = \frac{S_n}{n+1}$. Probar que por simetría, para $n \leq 0$

$$E(X_0|\mathcal{F}_n) = E(X_1|\mathcal{F}_n) = \cdots = E(X_{-n}|\mathcal{F}_n) = M_n.$$

Usando el ejercicio anterior, concluimos que $M = \{M_n\}_{n=0}^{-\infty}$ es una martingala reversa. Luego

$$\lim_{n \to -\infty} M_n = M_{-\infty} \text{ existe c.s..}$$

Como $M_{-\infty}$ es medible con respecto a la σ-álgebra asintótica $\cap_{n=0}^{-\infty} \mathcal{F}_n$, la v.a. $M_{-\infty}$ es constante c.s., digamos $M_{-\infty} = c$. Debemos comprobar que $E(M_{-\infty}) = E(X_0)$.

Si todas las variables $|X_n|$ están acotadas por la misma constante, digamos, T, entonces $|M_{-n}| \leq T$ y por el TCDL,

$$c = EM_{-\infty} = E(\lim_{n \to -\infty} M_n) = \lim_{n \to -\infty} E(M_n) = E(X_0).$$

En el caso general, fijamos $\varepsilon > 0$ y escogemos T tal que

$$E|X_0|\mathbf{1}_{\{|X_0|>T\}} < \varepsilon.$$

Esto es posible por el TCDL. Definimos $X_n^T := X_n \mathbf{1}_{\{|X_n|>T\}}$ y notamos que estas variables son independientes (¿por qué?). Ahora, el límite de los promedios M_{-n}^T existe en casi todo punto, es constante, digamos, C^T , y $C^T = E(X_0^T)$ por convergencia reversa y el TCDL. Además, $M_0 = X_0$, luego $E|M_0|\mathbf{1}_{\{|M_0|>T\}} < \varepsilon$.

Sabemos que existe un conjunto A de probabilidad 1 tal que para toda $\omega \in A$,

$$|c - c^T| = \left| \lim_{n \to -\infty} M_n(\omega) - \lim_{n \to -\infty} M_n^T(\omega) \right|$$

$$= \lim_{n \to -\infty} |E(M_0 \mathbf{1}_{\{|M_0| > T\}} | \mathcal{F}_{-n})(\omega)|.$$

Convergencia c.s. implica convergencia en probabilidad, luego existe una $N \gg 1$ y un conjunto $B \subset A$ con $P(B) > 1/2$ tal que para todo $\omega \in B$,

$$|c - c^T| < |E(M_0 \mathbf{1}_{\{|M_0| > T\}} | \mathcal{F}_{-N})(\omega)| + \varepsilon$$

$$\leq E(|M_0| \mathbf{1}_{\{|M_0| > T\}} | \mathcal{F}_{-N})(\omega) + \varepsilon.$$

Integrando sobre B, tenemos

$$|c - c^T| \leq 2 \int_B |c - c^T| dP$$

$$\leq 2 \int_B (E(|M_0| \mathbf{1}_{\{|M_0|>T\}} | \mathcal{F}_{-N})(\omega) + \varepsilon) dP(\omega)$$

$$\leq 2 \int_\Omega (E(|M_0| \mathbf{1}_{\{|M_0|>T\}} | \mathcal{F}_{-N})(\omega) + \varepsilon) dP(\omega)$$

$$\leq 2 \int_\Omega (|M_0| \mathbf{1}_{\{|M_0|>T\}}(\omega) + \varepsilon) dP(\omega) < 4\varepsilon.$$

Luego por el TCDL,

$$c = \lim_{T\to\infty} c^T = \lim_{T\to\infty} E(X_0^T) = E(\lim_{T\to\infty} X_0^T) = E(X_0).$$

$\square$