GENERALIZACIONES DE LA LEY FUERTE DE LOS GRANDES NÚMEROS Y DEL LEMA DE BOREL-CANTELLI

Ramírez Ramírez Lilia Leticia

Octubre de 1998

Contenido

INTRODUCCIÓN

Como es bien sabido, la probabilidad tiene su origen en los problemas relacionados con juegos de azar y su objetivo primario fue el de calcular las probabilidades de ganar una partida en juegos tales como dados, cartas o ruleta; sin embargo, antes de que cualquier cálculo formal se hubiera realizado, las reglas de los juegos solían calificar con mayor número de puntos a aquellas jugadas con menor probabilidad. Estas reglas se pueden atribuir a la noción intuitiva que pudiera haberse tenido de lo que que después formó la definición de probabilidad de un evento en juegos con número finito de posibles resultados (número de casos favorables entre el número de casos totales) o a la observación del número de casos favorables en una colección de jugadas (frecuencia relativa de exitos).

A pesar de que para el ser humano es natural estimar la probabilidad de un evento como la frecuencia relativa de éste (como la probabilidad de lluvia en una tarde de verano), no fue sino hasta finales del siglo XVII que James Bernoulli formalizó esta idea intuitiva para fenómenos aleatorio cuyo espacio muestral tiene cardinalidad dos. Este resultado puede enunciarse de la siguiente manera:

Si $\{X_n\}_{n\geq 1}$ es una sucesión de v.a.'s independientes e idénticamente distribuidas con ley Bernoulli y parámetro p, entonces

$$\frac{X_1 + \cdots + X_n}{n} \to p \quad \text{en probabilidad.}$$

Después de la obtención de este Teorema, el estudio del comportamiento asintótico de suma de variables aleatorias ha conformado uno de los temas centrales de investigación en la probabilidad, el cual aún hasta nuestros días sigue dando nuevos resultados suceptibles a aplicarse en teorías tales como la llamada de Muestras Grandes, la que a su vez se relaciona directamente con la modelación matemática de fenómenos naturales.

El teorema obtenido por Bernoulli se conoce como la Ley Débil de los Grandes Números y a partir de su formulación se ha desarrollado y generalizado hacia diversas vertientes. Algunas de su generalizaciones o versiones han surgido a partir de la búsqueda de respuestas a preguntas como son:

El teorema es válido para algún otro tipo de distribución?

Qué condiciones deben cumplir las variables aleatorias para que su suma estandarizada converja en probabilidad?

Qué condiciones garantizan la convergencia casi segura?

Uno de los resultados más célebres, el cual da respuesta a las preguntas anteriores, es el llamado Ley Fuerte de los Grandes Números, demostrada por Kolmogórov en 1929 y el cual dice así:

Si $\{X_n\}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas, entonces

$$\frac{X_1 + \cdots + X_n}{n} \to L \quad \text{casi seguramente}$$

si y sólo si EX_1 existe, y en este caso, $L = EX_1$.

El Teorema de Kolmogórov, además de dar respuesta a las preguntas antes expresadas, abre nuevos rumbos hacia los cuales encaminar la investigación en la Teoría de la Probabilidad, a saber:

Si las variables aleatorias no tienen esperanza finita, existen constantes a_n, b_n tales que

$$\frac{X_1 + \cdots + X_n - b_n}{a_n} \to L$$

casi seguramente?. Algunas respuestas a esta pregunta tienen lugar en Loéve, M. (1955) páginas 252-272.

Otra pregunta que podemos formular, que es más general, es:

Dada una sucesión (arbitraria) de variables aleatorias, qué condiciones debe cumplir para que existan, a su vez, variables aleatorias $\{Y_n\}_{n\geq 1}$, $\{W_n\}_{n\geq 1}$ tales que

$$\frac{X_1 + \cdots + X_n - Y_n}{X_n} \to L \tag{1}$$

casi seguramente?.

En las demostraciones de los teoremas que dan respuesta a las preguntas anteriores, el Lema de Borel-Cantelli (o sus extensiones, conocidas como Leyes 0-1) son clave, ya que el resultado de Borel-Cantelli da condiciones bajo las cuales el límite superior de una sucesión de eventos tiene probabilidad 0 o 1, y su relación con la Ley Fuerte de los Grandes Números es la siguiente:

$$\frac{X_1 + \cdots + X_n - Y_n}{W_n} \to L \quad \text{casi seguramente si y sólo si} \quad P(\limsup_n E_n^\epsilon) = 0$$

para toda $\epsilon > 0$, donde

$$E_n^\epsilon = \left[\left|\frac{X_1 + \cdots + X_n - Y_n}{W_n} - L\right| \geq \epsilon\right].$$

En este trabajo se presentan varias versiones, tanto del Lema de Borel-Cantelli como de la Ley Fuerte de los Grandes Números. No es nuestra intención hacer un estudio exhaustivo, sino solamente presentar algunos resultados para variables aleatorias no necesariamente independientes y con esperanza finita.

En el primer Capítulo se exponen elementos de Teoría de la Probabilidad que se requieren a lo largo del trabajo. Se presupone que el lector tiene conocimientos de algunos conceptos y resultados de Teoría de la Medida tales como la integral de Lebesgue, la integral de Lebesgue-Stieltjes, los Teoremas de Cambio de Variable, de Convergencia Monótona y Convergencia Dominada.

El segundo Capítulo está dedicado a presentar los resultados clásicos del Lema de Borel-Cantelli y la Ley de los Grandes Números para eventos independientes e independientes por parejas. Se incluye la Ley Débil, la demostración de Kolmogórov de la Ley Fuerte, así como algunas otras versiones de este Teorema. En particular, es presentada una versión moderna más simple de la Ley Fuerte obtenida por Etemadi [8] que sólo requiere de independencia por parejas y existencia del primer momento .

El Capítulo tercero se refiere a la Teoría de martingalas. Se muestran algunos de sus principales propiedades y resultados que se utilizarán en el resto del trabajo. Los principales que se exponen son el Teorema de Muestreo Opcional o de Paro Opcional de Doob y un teorema de convergencia de martingalas.

En el Capítulo cuarto se estudia la Ley Fuerte de los Grandes Números en el caso en que las variables aleatorias no son necesariamente independientes y la suma $X_1 + \cdots + X_n$ se normaliza no necesariamente con constantes, sino con variables aleatorias (como en (1)). Se presenta un Lema de Borel-Cantelli demostrado por Lévy, el cual es clave para las demostraciones de las Leyes de Grandes Números de Lévy y de Chow. Por último se presenta un resultado que conjunta el Lema de Borel-Cantelli y la Ley Fuerte de los Grandes Números de Lévy.

A lo largo de este trabajo se considerará un espacio de probabilidad $(\Omega, \Im, P)$ y cualquier variable aleatoria se tendrá por definida en este espacio. También denotaremos a las variables aletorias, como es usual en la literatura, con las últimas letras del alfabeto, en mayúsculas.

Capítulo 1

CONCEPTOS BÁSICOS

En este Capítulo se sentarán los fundamentos de Teoría de Probabilidad necesarios para poder desarrollar los resultados que a lo largo del trabajo se exponen.

1.1 INDEPENDENCIA ESTOCÁSTICA Y VARIABLES ALEATORIAS

Un espacio de probabilidad es una terna $(\Omega, \Im, P)$, donde Ω es un conjunto no vacío que llamaremos espacio muestral, $\Im$ es una σ-álgebra de subconjuntos de Ω, llamada espacio de eventos y P es una medida de probabilidad con dominio en $\Im$, es decir una medida tal que $P(\Omega) = 1$.

Definición 1.1 *Lamaremos a E evento del espacio de probabilidad $(\Omega, \Im, P)$ si $E \in \Im$.*

Definición 1.2 *Sea T un conjunto de índices (diferente del vacío) y $E = \{E_t\}_{t\in T}$ una familia de eventos. Decimos que E es una familia de eventos independientes si para todo subconjuto finito $\{t_1, \ldots, t_n\}$ de T,*

$$P\left(\cap_{j=1}^{n} E_{t_j}\right) = \prod_{j=1}^{n} P(E_{t_j}).$$

Se dice que la clase finita de eventos $\{C_t\}_{t=1}^{n}$ son independientes si $P(\cap_1^n E_j) = \prod_1^n P(E_j)$ para todo $E_t \in C_t \quad (t = 1, 2, \ldots, n)$. Si $\{C_t\}_{t\in T}$ es una clase de eventos, se dice que es independiente si cada subclase finita es independiente.

Definición 1.3 *Una función $X : \Omega \to \mathcal{R}$ se llama variable aleatoria (v.a.) si $X^{-1}(B) \in \Im$ para todo $B \in \mathcal{B}(\mathcal{R})$ (Borelianos de $\mathcal{R}$).*

Proposición 1.1.1 *Si X es una v.a. y $g : \mathcal{R} \longrightarrow \mathcal{R}$ una función Borel medible ($g^{-1}(B) \in \mathcal{B}(\mathcal{R}) \ \ \forall B \in \mathcal{R}$) entonces $g(X)$ es también una variable aleatoria.*

Demostración

Sea $B \in \mathcal{B}(\mathcal{R})$. Como g es Borel medible, se tiene que $g^{-1}(B) \in \mathcal{B}(\mathcal{R})$ y como X es v.a. entonces $X^{-1}(g^{-1}(B)) \in \mathcal{B}(\mathcal{R})$

$\square$

Definición 1.4 *A la función* $P(X^{-1}(\cdot)) : \mathcal{B}(\mathcal{R}) \to \mathcal{R}^+$ *se le llama* distribución de la v.a. X *o* probabilidad inducida por la v.a. X *y se le denota como* $P \circ X^{-1}(\cdot)$ *o* $P_X(\cdot)$.

Proposición 1.1.2 *La distribución de la v.a.* X *es una medida de probabilidad en* $\mathcal{R}$.

Demostración

Sean $B, B_1, B_2, \ldots \in \mathcal{B}(\mathcal{R})$ tales que $B_i \cap B_j = \emptyset \ \forall i \neq j$. Entonces

(i) $P_X(B) = P(X^{-1}(B)) \geq 0$,

(ii) $P_X(\mathcal{R}) = P(X^{-1}(\mathcal{R})) = P(\Omega) = 1$, y

(iii) $P_X(\cup_1^\infty B_n)=P(X^{-1}(\cup_1^\infty B_n))=P(\cup_1^\infty X^{-1}(B_n))=\sum_1^\infty P(X^{-1}(B_n))=\sum_1^\infty P_X(B_n)$.

$\square$

Definición 1.5 *A la función* $F_X : \mathcal{R} \to \mathcal{R}^+$ *definida por*

$$F_X(x) = P_X(-\infty, x], \qquad x \in \mathcal{R}$$

se le llama función de distribución de la v.a. X.

Definición 1.6 *Llamamos a* $X = (X_1, \cdots, X_n)$ *vector aleatorio si* X_i *es variable aleatoria* $(i = 1, 2, \ldots, n)$.

El vector X es una función medible de $(\Omega, \Im, P)$ a $(\mathcal{R}^n, \mathcal{B}(\mathcal{R}^n))$ porque

$$X^{-1}(B_1 \times B_2 \times \cdots \times B_n) = X_1^{-1}(B_1) \cap \cdots \cap X_n^{-1}(B_n),$$

donde $B_i \in \mathcal{R}$ para $i = 1, 2, \ldots, n$ y la σ-álgebra generada por los rectángulos medibles $(B_1 \times B_2 \times \cdots \times B_n)$ genera a $\otimes_1^n \mathcal{B}(\mathcal{R})$, el cual es igual a $\mathcal{B}(\mathcal{R}^n)$ porque $\mathcal{R}$ es un espacio separable. Denotamos a la distribución del vector X como $P \circ X^{-1}(\cdot) = P_X(\cdot) = P_{(X_1,\ldots,X_n)}(\cdot)$.

Definición 1.7 *Sea* $\{X_i\}_{i=1}^n$ *una familia finita de v.a.'s. Decimos que las v.a.'s* $X_1, \ldots, X_n$ *son variables aleatorias independientes (v.a.i.) o* $\{X_i\}_{i=1}^n$ *es independiente si* $\{X_i^{-1}\}$ *forman una clase independiente de eventos.*

La definición anterior nos dice que que si $X_1, \ldots, X_n$ son v.a.i. entonces

$$P_{(X_1,\ldots,X_n)}(B_1, \ldots, B_n) = P_{X_1}(B_1) \cdot \cdots \cdot P_{X_n}(B_n),$$

donde $B_i \in \mathcal{B}(\mathcal{R})$ para $i = 1, 2, \ldots, n$.

Definición 1.8 *Sean* T *un conjunto de índices y* $\{X_i\}_{i\in T}$ *un familia de v.a.'s. Decimos que* $\{X_i\}_{i\in T}$ *es independientes si cada subfamilia finita de* $\{X_i\}_{i\in T}$ *es independiente.*

1.2 ESPERANZA MATEMÁTICA

Definición 1.9 *Definimos L_p como*

$$\begin{aligned} L_p &= \{f : \Omega \to \mathcal{R} | P(f = g) = 1, \text{ para alguna } g \text{ Borel medible,} \\ &\quad \text{definida para todo } \omega \in \Omega \text{ y tal que} \int |g|^p dP < \infty\}, \end{aligned}$$

donde $p \in \mathcal{N}$ y $\int |g|^p dP$ es la integral de Lebesgue de $|g|^p$ con respecto a la medida de probabilidad P.

Definición 1.10 *Si $g : \mathcal{R} \to \mathcal{R}$ es una función Borel medible y X es una v.a., definimos la esperanza de $g(X)$ como*

$$Eg(X) = E[g(X)] = \int_\Omega g(X(\omega))dP(\omega) = \int_\Omega g(X)dP,$$

siempre que $g(X) \in L_1$.

Por el Teorema de Cambio de Variable tenemos que si $Eg(X) \in L_1$, entonces

$$Eg(X) = \int_\mathcal{R} g(v)dP_X(v).$$

Como $P_X(a, b] = F_X(b) - F_X(a)$ y F_X es no decreciente (por monotonía de la probabilidad) y contínua por la derecha, entonces también podemos expresar a la esperanza de $g(X)$ como la integral de Lebesgue-Stieltjes:

$$Eg(X) = \int_\mathcal{R} g(v)dF_X(v).$$

Definición 1.11 *Si $X \in L_p$, $p = 1, 2, \ldots,$ denominaremos a EX^p* el momento de orden *p de X.*

Nota 1.12 *El momento de orden 1 siempre se denominará* la esperanza de *X.*

De la propiedades de la integral de Lebesgue se siguen inmediatamente el ssiguiente resultado

Corolario 1.2.1 *.*

(i) Si $X, Y \in L_1$ y $\alpha, \beta \in \mathcal{R}$, entonces $\alpha X + \beta Y \in L_1$ y $E(\alpha X + \beta Y) = \alpha EX + \beta EY$.

(ii) Si $X \in L_1$ entonces $|EX| \leq E|X|$.

(iii) Si $Y \in L_1$, X v.a. tal que $|X| \leq |Y|$ casi seguramente (c.s.), entonces $X \in L_1$ y $E|X| \leq E|Y|$.

(iv) $X \in L_1$, con $E|X| = 0$ si y sólo si $X = 0$ (c.s.).

(v) Si $E|X|^\lambda < \infty$ para algún $\lambda > 0$ entonces $E|X|^\nu < \infty$ para $0 \leq \nu \leq \lambda$

Proposición 1.2.2 (Desigualdad de Cauchy-Schwarz) *Sea X y Y dos v.a. tales que X, $Y \in L_2$. Entonces $XY \in L_1$ y*

$$(E|XY|)^2 \leq EX^2 EY^2.$$

Demostración

Cuando alguna de X e Y es igual a cero c.s. se cumple la desigualdad.

Sean $EX^2 > 0$, $EY^2 > 0$, definamos

$$a = \frac{X}{[EX^2]^{1/2}} \quad , \quad b = \frac{Y}{[EY^2]^{1/2}}.$$

Entonces aplicando la desigualdad $2|ab| \leq a^2 + b^2$, a, $b \in \mathcal{R}$, se obtiene que

$$|XY| \leq \frac{1}{2}[EX^2 EY^2]^{\frac{1}{2}} \left[\frac{X^2}{EX^2} + \frac{Y^2}{EY^2} \right],$$

de donde resulta que $E|XY| < \infty$ y $[E|XY|]^2 \leq EX^2 EY^2$.

$\square$

Proposición 1.2.3 *Sea X una v.a. y g una función Borel medible, no negativa tal que $g(X) \in L_1$. Supóngase además que g es par y no decreciente en $[0, \infty)$. Entonces $\forall \epsilon > 0$ se tiene*

$$P[|X| \geq \epsilon] \leq \frac{Eg(X)}{g(\epsilon)}.$$

Demostración

Si $F = \{|X| \geq \epsilon\}$, entonces $Eg(X) = \int_F g(X)dP + \int_{F^c} g(X)dP \geq g(\epsilon)P(F)$.

$\square$

Nota 1.13 *Si $g(x) = |x|^\lambda$, obtenemos la desigualdad de* **Markov**, *y si además $\lambda = 2$, entonces obtenemos la desigualdad de* **Chebyshev**.

Lema 1.2.4 *Si Y es una v.a. no negativa entonces*

$$\sum_{n=1}^{\infty} P[Y \geq n] \leq EY \leq 1 + \sum_{n=1}^{\infty} P[Y \geq n]$$

Demostración

$$\begin{aligned}
\sum_{n=1}^{\infty} P[Y \geq n] &= \sum_{n=1}^{\infty}\sum_{k=n}^{\infty} P[k \leq Y < k+1] = \sum_{k=1}^{\infty}\sum_{n=1}^{k} P[k \leq Y < k+1] \\
&= \sum_{k=0}^{\infty} kP[k \leq Y < k+1] = \sum_{k=0}^{\infty}\int_{[k\leq Y<k+1]} kdP \\
&\leq \sum_{k=0}^{\infty}\int_{[k\leq Y<k+1]} YdP = EY \\
&\leq \sum_{k=0}^{\infty}(k+1)P[k \leq Y < k+1] \\
&= \sum_{n=1}^{\infty} P[Y \geq n] + \sum_{k=0}^{\infty} P[k \leq Y < k+1] \\
&= \sum_{n=1}^{\infty} P[Y \geq n] + 1
\end{aligned}$$

$\square$

Definición 1.14 *Decimos que la variable aletoria X está centrada en el punto c ($c \in \mathcal{R}$) si reemplazamos X por $X - c$.*

La selección de la constante c juega un papel importante en muchos problemas de probabilidad (como las Leyes de Grandes Números y el Teorema Central del límite).

Definición 1.15 *Si $X \in L_2$, al segundo momento de la v.a. X centrado en EX se le denomina como la varianza de X ($Var(X)$).*

Corolario 1.2.5 (Desigualdad de Kolmogórov) *Sean $X_1, X_2, \ldots, X_n$ v.a.i. tales que*

$$EX_k = 0$$

y

$$Var(X_k) = EX_k^2 = \sigma_k^2 < \infty, \quad k = 1, 2, \ldots, n.$$

Si $S_k = \sum_{j=1}^{k} X_j$, $k = 1, 2, \ldots, n$. Entonces para toda $\epsilon > 0$ se tiene que

$$P\{\max_{1\leq k\leq n} |S_k| \geq \epsilon\} \leq \frac{1}{\epsilon^2}\sum_{k=1}^{n}\sigma_k^2.$$

Demostración

Para $\epsilon > 0$ definamos los conjuntos

$$E_1 = \{|S_1| \geq \epsilon\}$$

y

$$E_k = \{|S_k| \geq \epsilon\} \cap \bigcap_{j=1}^{k-1} \{|S_j| < \epsilon\} \quad \text{para} \;\; k = 2, 3, \ldots, n.$$

Por construcción, tenemos que $E_i \cap E_j = \emptyset$ para toda $i \neq j$.

Sea $E = \{\max_{1\leq k\leq n} |S_k| \geq \epsilon\}$. Entonces $E = \cup_{k=1}^{n} E_k$. Como X_k son independientes y $EX_k = 0$, entonces

$$\begin{aligned} \sum_{k=1}^{n} \sigma_k^2 = \mathrm{Var}(S_n) &= \int_\Omega S_n^2 dP \geq \int_E S_n^2 dP \\ &= \sum_{k=1}^{n} \int_{E_k} S_n^2 dP. \end{aligned}$$

Por otro lado, si $1 \leq k \leq n$, tenemos que

$$\begin{aligned} \int_{E_k} S_n^2 dP &= \int_{E_k} (S_k + X_{k+1} + \cdots + X_n)^2 dP \\ &= \int_{E_k} S_k^2 dP + 2\int_{E_k} S_k(X_{k+1} + \cdots + X_n) dP + \int_{E_k} (X_{k+1} + \cdots + X_n)^2 dP, \end{aligned}$$

pero

$$2\int_{E_k} S_k(X_{k+1} + \cdots + X_n) dP = 2\int_\Omega I_{E_k} S_k dP \cdot \int_\Omega (X_{k+1} + \cdots + X_n) dP = 0$$

por la asociatividad de la independencia. Además, como

$$\int_{E_k} (X_{k+1} + \cdots + X_n)^2 dP \geq 0,$$

entonces

$$\begin{aligned} \mathrm{Var}(S_n) &= \int_\omega S_n^2 dP \geq \int_E S_n^2 dP \\ &= \sum_{k=1}^{n} \int_{E_k} S_n^2 dP \geq \sum_{k=1}^{n} \int_{E_k} S_k^2 dP \\ &\geq \sum_{k=1}^{n} \epsilon^2 P(E_k). \end{aligned}$$

Como

$$\epsilon^2 P(E) = \epsilon \sum_{k=1}^{n} P(E_k),$$

finalmente se tiene que

$$\epsilon^2 P(E) \geq \sum_{k=1}^{n} \mathrm{Var}(X_k)$$

$\square$

1.3 CONVERGENCIAS EN PROBABILIDAD, CASI SEGURA Y EN MEDIA

Definición 1.16 *Se dice que la sucesión de variables aleatorias* $\{X_n\}_{n\geq 1}$ converge casi seguramente *(c.s.) a la v.a.* X *si existe* $E \in \Im$ *con* $P(E) = 0$ *tal que* $\forall\omega \in E^c$ *se cumple*

$$|X_n(\omega) - X(\omega)| \to 0 \quad \textit{cuando} \;\; n \to \infty.$$

En este caso escribimos $X_n \overset{c.s.}{\to} X$.

La noción de convergencia casi donde sea, de Teoría de la Medida, es idéntica a la de convergencia casi segura en Probabilidad, pero cambia de nombre no sólo porque la medida de probabilidad es una medida con la propiedad de que $P(\Omega) = 1$, sino para hacer referencia a que una probabilidad es una medida de *posibilidad*.

Definición 1.17 *Se dice que la sucesión* $\{X_n\}_{n\geq 1}$ *es* Cauchy casi seguramente *si existe* $E \in \Im$ *tal que* $P(E) = 0$ *y* $|X_n(\omega) - X_m(\omega)| \to 0$ *cuando* $m, n \to \infty$ *para toda* $\omega \in E^c$.

Proposición 1.3.1 *Sea* $\{X_n\}_{n\geq 1}$ *una sucesión de v.a.'s. Entonces* $X_n \overset{c.s.}{\to} X$, X *v.a.* $\Leftrightarrow$ $\{X_n\}$ *es Cauchy c.s..*

Demostración

Si $X_n \overset{c.s.}{\to} X$ entonces existe $E \in \Im$ tal que $P(E) = 0$ y $X_n(\omega) \to X(\omega) \quad \forall\omega \in E^c$. Entonces para $\omega \in E^c$ y para m, n números enteros positivos se tiene que

$$|X_m(\omega) - X_n(\omega)| \leq |X_m(\omega) - X(\omega)| + |X_n(\omega) - X(\omega)| \to 0$$

cuando m y $n \to \infty$. Entonces $\{X_n\}$ es Cauchy c.s..

Por otra parte, si $\{X_n\}$ es Cauchy c.s. entonces existe $E \in \Im$ tal que $P(E) = 0$ y $|X_n(\omega) - X_m(\omega)|_{m,n\to\infty} \to 0 \;\; \forall\omega \in E^c$. Entonces la sucesión $\{X_n(\omega)\}$ es una sucesión Cauchy de números reales, lo que nos lleva a la conclusión de que existe un único número real $X(\omega)$ tal que $X(\omega) = \lim_{n\to\infty} X_n(\omega) \quad \forall\omega \in E^c$. Por tanto $X_n \overset{c.s.}{\to} X$. La función X es una variable aleatoria porque el límite de variables aleatorias es también una v.a..

□

La siguiente Proposición nos va a ser de mucha utilidad para demostrar la convergencia de sucesiones de v.a.'s.

Proposición 1.3.2 *La sucesión* $\{X_n\}_{n\geq 1}$ *de v.a.'s converge c.s. a la v.a.* X $\Leftrightarrow$

$$\lim_{n\to\infty} P\{\cup_{m=n}^{\infty}(|X_m - X| \geq \epsilon)\} = 0 \quad \forall\epsilon > 0.$$

Como $P\{\cup_{m=n}^{\infty}(|X_m - X| \geq \epsilon)\} \leq \sum_{m=n}^{\infty} P\{(|X_m - X| \geq \epsilon)\}$, entonces

Corolario 1.3.3

$$\sum_{m=1}^{\infty} P\{(|X_m - X| \geq \epsilon)\} < \infty \quad \forall \epsilon > 0 \Rightarrow X_n \stackrel{c.s.}{\to} X. \tag{1.1}$$

El Lema de Borel-Cantelli también nos da el resultado (1.1), pero no puede decirnos nada a partir de $\lim_{n\to\infty} P\{\cup_{m=n}^{\infty}(|X_m - X| \geq \epsilon)\} = 0$, que es una condicion más débil.

Demostración de la Proposición 1.3.2

Sea $\epsilon > 0$. Definimos $E_n(\epsilon) = [|X_n - X| \geq \epsilon]$ para $n \geq 1$.

Sea $D = \{\omega \in \Omega : X_n(\omega) \not\to X(\omega)$ cuando $n \to \infty\}$, entonces

$$\begin{aligned} D &= \cup_{\epsilon>0} \limsup_n E_n(\epsilon) \\ &= \cup_{\epsilon>0}(\cap_{n\geq 1} \cup_{m\geq n} E_m(\epsilon)) \\ &= \cup_{k\geq 1}(\cap_{n\geq 1} \cup_{m\geq n} E_m(1/k)). \end{aligned}$$

Por lo tanto, tenemos que

$$\begin{aligned} X_n \stackrel{c.s.}{\to} X &\Leftrightarrow P(D) = 0 \\ &\Leftrightarrow P(\limsup_n E_n(\epsilon)) = 0 \quad \forall \epsilon > 0. \end{aligned}$$

Como $\cup_{m=n} E_m(\epsilon) \downarrow \limsup E_n(\epsilon)$, entonces

$$P(\limsup_n E_n(\epsilon)) = \lim_n P(\cup_{m\geq n} E_m(\epsilon)),$$

lo cual implica que $X_n \stackrel{c.s.}{\to} X \Leftrightarrow \lim_{n\to\infty} P\{\cup_{m\geq n}(|X_m - X| \geq \epsilon)\} = 0 \quad \forall \epsilon > 0.$

□

Definición 1.18 *Sea* $\{X_n\}_{n\geq 1}$ *una sucesión de v.a.'s. Decimos que* X_n converge en probabilidad *a la v.a.* X *si* $\forall \epsilon > 0$

$$P\{|X_n(\omega) - X(\omega)| > \epsilon\} \to 0 \quad \textit{cuando } n \to \infty.$$

En este caso, escribimos $X_n \stackrel{P}{\to} X$.

Proposición 1.3.4 *Sea* $\{X_n\}_{n\geq 1}$ *una sucesión de v.a.'s tal que* $X_n \stackrel{c.s.}{\to} X$. *Entonces* $X_n \stackrel{P}{\to} X$.

Demostración

De la Proposición 1.3.2 tenemos que cuando $X_n \stackrel{c.s.}{\to} X$ se cumple

$$\lim_{n\to\infty} P\{\cup_{m\geq n}(|X_m - X|] \geq \epsilon)\} = 0.$$

Como $\{|X_n - X| \geq \epsilon\} \subset \cup_{m\geq n}\{|X_m - X| \geq \epsilon\}$, entonces se tiene que $X_n \stackrel{P}{\to} X$.

□

Definición 1.19 *Sea* $\{X_n\}_{n\geq 1}$*una sucesión de v.a.'s, tal que* $X_n \in L_1 \quad \forall n \geq 1$*. Decimos que* X_n converge en media *a la v.a.* $X \in L_1$ *si* $E|X_n - X| \to 0$ *cuando* $n \to \infty$*.*

En este caso escribimos $X_n \xrightarrow{L_1} X$*.*

Nota 1.20 *No es difícil demostrar que la convergencia en* L_p *(de Teoria de la Medida) implica la convergencia en media (o en* L_1*) cuando* $p \geq 1$*.*

Proposición 1.3.5 $X_n \xrightarrow{L_1} X \Rightarrow X_n \xrightarrow{P} X$*.*

La demostración se sigue de la desigualdad de Markov. Para cualquier $\epsilon > 0$,

$$P\{|X_n - X| \geq \epsilon\} \leq \frac{E|X_n - X|}{\epsilon}.$$

$\square$

1.4 PROBABILIDAD Y ESPERANZA CONDICIONALES

En los cursos básicos de probabilidad suele definirse a la esperanza condicional de la siguiente manera:

Definición 1.21 *Si* $A, B \in \Im$ *y* $P(B) > 0$ *entonces la probabilidad condicional de* A *dado* B *es igual a*

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

Sin embargo esta definición puede extenderse para el caso en el que $P(B) = 0$. El Teorema siguiente generaliza la definición anterior y el Teorema 1.4.2 constituye la definición general de la Esperanza Condicional.

Teorema 1.4.1 *Si* $\wp \subset \Im$ *es sigma-álgebra y* $B \in \Im$ *es fijo, entonces existe una función* $P[B|\wp] : (\Omega, \wp) \longrightarrow (\mathcal{R}, \mathcal{B}(\mathcal{R}))$*, llamada la probabilidad condicional de* B *dado* $\wp$*, tal que*

$$P(C \cap B) = \int_C P[B|\wp]dP \qquad \forall C \in \wp.$$

Si h *es una función con las mismas propiedades que* $P[B|\wp]$*, entonces* $h = P[B|\wp]$ *P-c.s..Cuando esto último ocurre, se dice que* $P[B|\wp]$ *es* esencialmente única*.*

Nota 1.22 *A partir de ahora basta con decir que* $\wp \subset \Im$ *para considerar a* $\wp$ *como sub* σ*-álgebra de* $\Im$*.*

Demostración

Sea $\lambda(C) = P(C \cap B)$, $C \in \wp$. Entonces λ es absolutamente contínua con respecto a P. Por el Teorema de Radon-Nikodym, existe una función f (esencialmente única respecto a P) que cumple $P(C \cap B) = \int_C f dP$ para todo $C \in \wp$.

$\square$

Teorema 1.4.2 *Sea Y una v.a. con valores en los reales extendidos y sea $\wp \subset \Im$. Si $Y \in L_1$, entonces existe una función (esencialmente única) $E[Y|\wp] : (\Omega, \wp) \longrightarrow \mathcal{R}, \mathcal{B}(\mathcal{R}))$, llamada la esperanza condicional de Y dado $\wp$, tal que*

$$\int_C Y dP = \int_C E[Y|\wp] dP \qquad \forall C \in \wp.$$

Nótese que debido a las definiciones anteriores se tiene que $P[B|\wp] = E[I_B|\wp]$ c.s..

Demostración

Si $\lambda(C) = \int_C Y dP$, para $C \in \wp$, se tiene que λ es absolutamente continua respecto a P . Por el Teorema de Radon-Nikodym, existe una función $f : \Omega \to \mathcal{R}$, $\wp$-medible, esencialmente única respecto a P e integrable, tal que

$$\lambda(C) = \int_C f dP \qquad \forall C \in \wp.$$

Definimos a f como $E[Y|\wp]$.

$\square$

Definición 1.23 *Si X es una v.a., definimos a la σ-álgebra generada por X como*

$$\sigma(X) = \{A \subset \Omega : X^{-1}(B) = A, \quad B \in \mathcal{B}(\mathcal{R})\}$$

Nota 1.24 *Alternativamente denotaremos como la esperanza condicional de una v.a. Y dada la σ-álgebra generada por una v.a. X como $E[Y|X]$, en lugar de escribir $E[Y|\sigma(X)]$.*

Así como hemos definido la esperanza condicional dada una σ-álgebra, podemos definir la esperanza $E[Y|X = x]$, la cual es una función con dominio en los reales, i.e., $E[Y|X = x] = h(x)$. La demostración del siguiente Teorema-Definición es muy similar a la del Teorema 1.4.2.

Teorema 1.4.3 *Sean X y Y una v.a.'s. Si $Y \in L_1$, entonces existe una función medible y esencialmente única $h : (\mathcal{R}, \mathcal{B}(\mathcal{R})) \to (\mathcal{R}, \mathcal{B}(\mathcal{R}))$ tal que para cada $B \in \mathcal{B}(\mathcal{R})$*

$$\int_{X \in B} Y dP = \int_B h(x) dP_X(x).$$

A h se le denota como $h(x) = E[Y|X = x]$.

Demostración

Sea $\lambda(C) = \int_{X \in C} Y dP$, entonces λ es absolutamente contínua con respecto a P_X. Por el Teorema de Radon-Nikodym se tiene la conclusión.

$\square$

Nótese que a partir de las dos defininciones anteriores de esperanza condicional se tiene que

$$E[Y|X] = h(X) \quad \text{c.s.} \quad \text{si} \qquad E[Y|X = x] = h(x).$$

Un ejemplo de esta observación es: si $E[Y|X = x] = x^2$, entonces $E[Y|X] = X^2$.

Algunas de las propiedades más importantes de la esperanza condicional se enuncian a continuación.

Propiedades

Si X, Y, $X_1, X_2, \ldots$ son variables aleatorias y $\wp \subset \Im$, se tienen las siguientes implicaciones.

(i) $X \in L_1 \Rightarrow E[X|\wp] = X \quad \text{c.s. en } \wp.$
(ii) $X \geq 0 \Rightarrow E[X|\wp] \geq 0. \quad \text{c.s.}$
(iii) $X, Y \in L_1 \Rightarrow E[aX + bY|\wp] \overset{c.s.}{=} aE[X|\wp] + bE[Y|\wp], \quad \forall a, b \in \mathcal{R}.$
(iv) Si $X, Y, X \cdot Y \in L_1$ y si X es $\wp -$ medible, entonces $E[X \cdot Y|\wp] \overset{c.s.}{=} XE[Y|\wp]$.
(v) $\wp' \subset \wp$ y $Y \in L_1 \Rightarrow$

$$E[Y|\wp'] = E[E[Y|\wp']|\wp] = E[E[Y|\wp]|\wp'] \;\; c.s.$$

(vi) Si $X \in L_1,$ y $\sigma(X)$ y $\wp$ son independientes, entonces $E[X|\wp] = EX \quad \text{c.s.}$
(vii) Si $X, Y \in L_1$ y X es independiente de Y, entonces $E[X|\sigma(Y)] = EX$.

Demostración

(i) $\int_C E[X|\wp]dP = \int_C XdP \;\; \forall C \in \wp \Rightarrow E[X|\wp] = X,\ P - c.s.$ en $\wp$.

(ii) $\int_C XdP \geq 0$ si $X \geq 0 \quad \forall C \in \wp$.

(iii) Se sigue de la linealidad de la integral de Lebesgue.

(iv) Si $X = I_A$ con $A \in \wp$,

$$\int_C E[XY|\wp]dP = \int_{C\cap A} YdP = \int_{C\cap A} E[Y|\wp]dP = \int_C XE[Y|\wp]dP.$$

Si $X = \sum_{i=1}^{n} a_i I_{B_i}$ con $a_i \geq 0$ y $B_i \in \wp$ $\forall i$, se demuestra el resultado gracias a la linealidad de la integral. Utilizando el Teorema de Convergencia Monótona, se obtiene el resultado para funciones medibles no negativas, y como $X = X^+ - X^-$ (donde X^+ y X^- son la la parte positiva y negativa de X, respectivamente), por la linealidad de la integral se obtiene la conclusión.

(v) Nótese que $\int_A E[Y|\wp']dP = \int_A YdP = \int_A E[Y|\wp]dP \quad \forall A \in \wp'$. Se sigue $E[Y|\wp'] = E[E[Y|\wp]|\wp']$ c.s..

Por otra parte, $E[Y|\wp']$ es una función $\wp'$-medible, y entonces $\wp-$medible, por lo que $E[Y|\wp'] = E[E[Y|\wp']|\wp]$ c.s..

(vi) Si $X = I_A$ con $A \in \sigma(X)$, $\int_C E[I_A|\wp]dP = \int_C I_A dP = \int_{C\cap A} dP = P(C\cap A) = P(C)P(A) = \int_C E(X)$. Por un argumento similar al anterior, se puede concluir que la igualdad es cierta para funciones simples, para funciones medibles no negativa y por ultimo para cualquier v.a..

(vii) Si $A \in \sigma(X)$, entonces existe $C \in \mathcal{B}(\mathcal{R})$ tal que $X^{-1}(C) = A$, similarmente, existe $D \in \mathcal{B}(\mathcal{R})$ tal que $Y^{-1}(D) = B$. Por lo tanto tenemos que $P(A\cap B) = P[X^{-1}(C) \cap Y^{-1}(D)] = P[X^{-1}(C)] \cdot P[Y^{-1}(D)] = P(A)P(B)$. La conclusión se sigue del inciso anterior.

$\square$

Capítulo 2

LEMA DE BOREL-CANTELLI Y LEY FUERTE DE LOS GRANDES NÚMEROS

En este Capítulo se exponen los resultados clásicos del Lema de Borel-Cantelli (B-C) y la Ley Fuerte de los Grandes Números (LFGN). Algunos de éstos se refieren a variables aleatorias que son independientes o independientes por parejas.

En la primera Sección se demuestra el Lema de Borel Cantelli y algunas extensiones. Además se exponen algunas de sus aplicaciones.

En la segunda Sección se estudian las Ley Fuerte de Grandes Números y al igual que en la primera, se exponen algunas de sus extensiones. Algunas de éstas tienen demostraciones sencillas pero no pueden considerarse más generales que la de Kolmogórov porque requieren de condiciones de existencia de los momentos de orden superior o igual a uno.

A pesar de que la LFGN de Kolmogórov sigue siendo un resultado muy fuerte, debido a que nos puede proporcionar la existencia del primer momento y la convergencia de la serie de v.a.'s a éste, requiere que las variables aleatorias sean independienteS e idénticamente distribuidas. Las versiones que se presentan en la segunda Sección, a pesar de no ser más generales, nos dan algunas ventajas en cuanto a la reduccion de estas hipótesis.

2.1 LEMA DE BOREL-CANTELLI

Son múltiples los resultados de límites en la teoría de la probabilidad que se comprueban utilizando el Lema de Borel-Cantelli, el cual consta de dos partes, una que se refiere a la convergencia y otra a la divergencia de una sucesión de eventos. La parte de divergencia del Lema requiere de una hipótesis de independencia de eventos, la cual puede ser refinada para poder generalizar el resultado.

Lema 2.1.1 (DE BOREL-CANTELLI) *Sea* $(\Omega, \Im, P)$ *un Espacio de Probabilidad y sea* $\{E_n\}$ *una sucesión de eventos del espacio.*

$$a) \quad \sum_{n=1}^{\infty} P(E_n) < \infty \Longrightarrow P(\limsup_n E_n) = 0. \tag{2.1}$$

Si además los eventos son independientes

$$b) \quad \sum_{n=1}^{\infty} P(E_n) = \infty \Longrightarrow P(\limsup_n E_n) = 1. \tag{2.2}$$

Demostración

Para demostrar la primera parte definimos a E como

$$E = \limsup_n E_n = \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} E_m$$

y podemos escribir $E = \cap_{n=1}^{\infty} F_n$, donde $F_n = \cup_{m=n}^{\infty} E_m$. Entonces

$$P(F_n) = P(\cup_{m=n}^{\infty} E_m) \leq \sum_{m=n}^{\infty} P(E_m).$$

Como $\sum_{n=1}^{\infty} P(E_n) < \infty$ se cumple $\lim_{n \to \infty} \sum_{m=n}^{\infty} P(E_m) = 0$

Notemos que $F_n \downarrow E$ cuando $n \to \infty$, lo cual implica

$$P(E) = \lim_{n \to \infty} P(F_n) = 0.$$

Con lo que se demuestra (2.1).

Para probar la segunda parte es suficiente con mostrar que

$$P(\cup_{n=1}^{\infty} \cap_{m=n}^{\infty} E_m^c) = 0,$$

pero para esto basta con verificar que

$$P(\cap_{m=n}^{\infty} E_m^c) = 0 \ \ \forall n \in \mathcal{N}.$$

Como $1 - x \leq \exp(-x) \ \ \forall x \in \mathcal{R}$ entonces

$$\begin{aligned} P(\cap_{m=n}^{\infty} E_m^c) &\leq P(\cap_{m=n}^{N} E_m^c) = \prod_{m=n}^{N} P(E_m^c) \\ &= \prod_{m=n}^{N} (1 - P(E_m)) \\ &\leq \exp[-\sum_{m=n}^{N} P(E_m)] \end{aligned}$$

A partir de que $\sum_m P(E_m)$ diverge, se tiene $\lim_{j\to\infty} \exp[-\sum_{m=n}^{n=j} P(E_m)] = 0$ entonces

$$P(\cap_{m=n}^{\infty} E_m^c) \leq \lim_{N\to\infty} \exp[-\sum_{m=n}^{N} P(E_m)] = 0.$$

Asi tenemos (2.2)

$\square$

A partir del Lema de Borel-Cantelli podemos decir que si la coleccion de eventos $\{E_n\}$ son independientes, entonces

$$P(\limsup_n E_n) = 0 \Leftrightarrow \sum_{n=1}^{\infty} P(E_n) < \infty.$$

Observemos que que es justamente la parte referente a la convergencia de eventos del Lema la que nos permite enunciar el Corolario 1.3.3, ya que $X_n \to X$ c.s. si y sólo si $P(\limsup |X_n - X| \geq \epsilon) = 0 \ \ \forall \epsilon > 0$.

Algunas otras aplicaciones del Lema de Borel-Cantelli, las constituyen los siguientes ejemplos y el Teorema 2.1.2.

Ejemplo 2.1 *Si $\{X_n\}$ es una sucesión de variables aleatorias en L_1, existen constantes positivas $c_n \to \infty$ tales que*

$$\frac{X_n}{c_n} \overset{c.s.}{\to} 0.$$

La demostración se sigue del Corolario 1.3.3, ya que para cada $n \in \mathcal{N}$ existe c_n tal que $P[|X_n|/c_n > \epsilon] < 1/n$, lo cual implica que $\sum_{n=1}^{\infty} P[|X_n/c_n| > \epsilon] < \infty$ y por lo tanto $X_n/c_n \overset{c.s.}{\to} 0$.

Ejemplo 2.2 *Sea A_n el evento de que salga cara en el n-ésimo y (n+1)-ésimo lanzamiento de una moneda. Si definimos $A = \limsup_n A_n$, entonces A es el evento de que salgan dos caras consecutivas infinitas veces.*

Como $P(A_n) = 1/4 \ \ \forall n \in \mathcal{N}$, entonces $\sum_{n=1}^{\infty} P(A_{2n}) = \infty$. Por el Lema de B-C, se tiene que $P(A) = 1$.

Teorema 2.1.2 *Sea $\{X_n\}$ una sucesión de v.a. tal que $X_n \overset{P}{\to} X$, donde X es v.a., entonces existe una subsucesión $\{X_{n_k}\} \subset \{X_n\}$ que cumple*

$$X_{n_k} \overset{c.s.}{\to} X \quad \textit{cuando} \quad k \to \infty$$

Se observa que el Teorema anterior, es un tipo de complemento a la Proposición 1.3.4.

Demostración

Como para toda $\epsilon > 0$ tenemos que $P[|X_n - X| > \epsilon] \to 0$, entonces podemos escoger enteros positivos $n_k \to \infty$ que cumplan que $P[|X_{n_k} - X| > \epsilon] \leq 2^{-k}$ para cada $k \geq 1$. Entonces $\sum_{k=1}^{\infty} P[|X_{n_k} - X| > \epsilon] < \infty$. Por el Lema de Borel-Cantelli se obtiene que

$$P[\limsup |X_{n_k} - X| > \epsilon] = 0$$

lo que a su vez nos indica que $X_{n_k} \overset{c.s.}{\to} X$.

$\square$

A continuación se presenta un primer refinamiento de la parte divergente de eventos del Lema de B-C.

Teorema 2.1.3 *Sea $\{E_n\}$ una sucesión arbitraria de eventos tal que $\sum_{n=1}^{\infty} P(E_n) = \infty$ y*

$$\liminf_n \frac{\sum_{j,k \leq n} P(E_j \cap E_k)}{(\sum_{k \leq n} P(E_k))^2} \leq 1. \tag{2.3}$$

entonces $P(\limsup_n E_n) = 1$.

A pesar de que este refinamiento elimina la condición de independencia de los eventos, es difícil de manejar cuando se quiere verificar sus condiciones.

Con el objeto de realizar la demostración de este Teorema se realizarán las siguientes definiciones.

Si tenemos la sucesión de eventos $\{E_i\}$ consideremos $N_n = I_1 + I_2 + \cdots + I_n$, donde $I_i = I_{E_i}$ (Función Indicadora el evento E_i) para $1 \leq i$. Como

$$\limsup_n E_n = \{w : \sup_n N_n(w) = \infty\},$$

entonces $P(\limsup_n E_n)$ puede ser estudiado por medio de las variables aleatorias N_n.

Supongamos que $p_k = P(E_k)$ y $m_n = p_1 + p_2 + \cdots + p_n$.

Como $E[I_k] = p_k$ entonces $E[N_n] = m_n$ y si $m_n > x$, por la desigualdad de Chebyshev se tiene que

$$P[N_n \leq x] \leq P[|N_n - m_n| \geq m_n - x] \leq \frac{\operatorname{Var}(N_n)}{(m_n - x)^2}. \tag{2.4}$$

Con estas definiciones, procedemos entonces a la demostración.

Demostración

$$Sea \qquad \theta_n = \frac{\sum_{j,k \leq n} P(E_j \cap E_k)}{(\sum_{k \leq n} P(E_k))^2}.$$

Utilizando la notación anterior tenemos que

$$\begin{aligned} \operatorname{Var}(N_n) &= EN_n^2 - m_n^2 = \sum_{j,k \leq n} E[I_j I_k] - m_n^2 \\ &= \sum_{j,k \leq n} P(E_j \cap E_k) - m_n^2 = (\theta_n - 1) m_n^2 y \end{aligned}$$

por (2.4) resulta que

$$P[N_n \leq x] \leq \frac{(\theta_n - 1)m_n^2}{(m_n - x)^2} \quad \text{para } x < m_n.$$

Como $m_n^2/\{(m_n - x)^2\} \to 1$, entonces $\liminf_n P[N_n \leq x] \leq 0$.

A partir de que $P[\sup_k N_k < x] \leq P[N_n \leq x]$ obtenemos que $P[\sup_k N_k < x] = 0$. Si tomamos la unión sobre $x = 1, 2, \ldots$ entonces

$$P[\sup_n N_n < \infty] = 0,$$

por lo que

$$P[\sup_n N_n = \infty] = 1$$

□

El siguiente Teorema nos permite generalizar la parte divergente del Lema de Borel-Cantelli a eventos independientes por parejas.

Teorema 2.1.4 *Sea $\{E_n\}$ una sucesión de eventos independientes por pares tal que*

$$\sum_{n=1}^{\infty} P(E_n) = \infty,$$

entonces $P(\limsup_n E_n) = 1$.

A pesar de que este Teorema nos permite considerar solamente independencia por parejas en lugar de independencia entre los eventos, su condición de independencia sigue siendo una condición muy fuerte.

Demostración

Con el fin de realizar la demostración utilizaremos la notación anterior.

Como $\operatorname{Var}(N_n) = EN_n^2 - m_n^2$ y

$$\begin{aligned}
EN_n^2 &= E\left[\sum_{i=1}^{n} I_i^2 + 2\sum_{1\leq j\leq i\leq n} I_i I_j\right] \\
&= \sum_{i=1}^{n} EI_i^2 + 2\sum_{1\leq j\leq i\leq n} EI_i EI_j \\
&= \sum_{i=1}^{n} EI_i + 2\sum_{1\leq j\leq i\leq n} EI_i EI_j + \sum_{i=1}^{n}(EI_i)^2 - \sum_{i=1}^{n}(EI_i)^2 \\
&= \sum_{i=1}^{n}(EI_i)^2 + 2\sum_{1\leq j\leq i\leq n} EI_i EI_j + \sum_{i=1}^{n}(EI_i - (EI_i)^2) \\
&= m_n^2 + \sum_{i=1}^{n}(p_i - p_i^2),
\end{aligned}$$

entonces $\mathrm{Var}(N_n) = \sum_{i=1}^n (p_i - p_i^2)$. Si tomamos $\theta_n = 1 + \sum_{i \leq n} \frac{(p_i - p_i^2)}{m_n^2}$ tenemos, utilizando (2.4), que

$$P[N_n \leq x] \leq \frac{\sum_{i=1}^n (p_i - p_i^2)}{(m_n - x)^2} = \frac{(\theta_n - 1)m_n^2}{(m_n - x)^2}.$$

Como $\theta \geq 1 - (m_n/m_n^2)$ y $m_n \to \infty$ cuando $n \to \infty$ por hipótesis, entonces $\liminf_n(\theta_n - 1) \leq 0$. Con los argumentos utilizados para el caso anterior podemos concluir que $P(\limsup_n E_n) = 1$.

□

Es inmediato reconocer que el Teorema 2.1.4 es una generalización del Lema de B-C original y a partir de la demostración de éste, tenemos que si $\{E_n\}$ es una sucesión de eventos independientes por pares, entonces $P(\limsup_n E_n) = 0 \Leftrightarrow \sum P(E_n) < \infty$.

Si en el Lema de Borel-Cantelli hacemos a un lado la hipótesis de independencia en la parte divergente del mismo, entonces se puede demostrar mediante contraejemplo que $\sum_{n=1}^\infty P(E_n) = \infty \not\Longrightarrow P(\limsup_n E_n) = 1$.

Ejemplo 2.3 *Sean $\Omega = [0,1]$, $\Im$ la $\sigma-$ álgebra de Borel de Ω, P la medida de Lebesgue y $E_n = (0, 1/n)$ para $n = 1, 2, \ldots$.*

Claramente vemos que $\{E_n\}$ no es una sucesión de eventos independientes y que

$$\limsup_n E_n = \lim_n E_n = \cap_{n=1}^\infty E_n = \emptyset.$$

Entonces $P(\limsup_n E_n) = 0$, pero $\sum_{n=1}^\infty P(E_n) = \sum_{n=1}^\infty 1/n = \infty$.

2.2 LEYES DE GRANDES NÚMEROS

Las Leyes de Grandes Números son resultados dentro de los llamados Teoremas Límite, y su definición (la cual se ha modificado debido a la evolución que estos han tenido) es la siguiente: las Leyes de Grandes Números se refieren a la convergencia de la suma de variables aleatorias. La Ley Débil considera convergencia en probabilidad y la Fuerte convergencia casi segura.

Algunas versiones de la Ley Débil de los grandes Números tienen como consecuencia que si observamos n realizaciones de una variable aleatoria (n grande) y dichas realizaciones son independientes unas de otras, entonces el promedio de los valores obtenidos será muy cercano a la media de la variable aleatorias, en un alto porcentaje. La Proposición 2.2.1 es una Ley Débil que no requiere que las observaciones efectuadas provengan de variables aleatorias con misma distribución.

Proposición 2.2.1 *Sea $X_1, X_2, \ldots, X_n$ una sucesión de v.a.'s independientes tal que*

$$\lim_{n\to\infty} n^{-2} \sum_{k=1}^n VarX_k = 0,$$

entonces

$$g_n = (\sum_1^n X_k - \sum_1^n EX_k)/n \xrightarrow{P} 0.$$

Demostración

Por la desigualdad de Chebyshev, basta demostrar que g_n converge en L_2.

$$E(g_n^2) = n^{-2}\mathrm{Var}(X_1 + \cdots + X_n) = n^{-2}\sum_{k=1}^{n}\mathrm{Var}(X_k) \to 0 \quad \text{cuando} \ \ n \to \infty.$$

□

Como la Ley Fuerte de Grandes Números se refiere a la convergencia de la suma de v.a's en forma casi segura, entonces por la Proposicion 1.3.4, se tiene que la Ley Débil de Grandes Números es un caso particular de ésta.

El Teorema siguiente es una versión de la Ley Fuerte de los Grandes Números, la cual requiere que las variables aleatorias pertenezcan a L_4.

Teorema 2.2.2 *Sea $\{X_n\}_{n\geq 1}$ una sucesión de v.a.i., con cuarto momento finito.*

Supongamos que $EX_n = \mu$, $Var(X_n) = \sigma^2$ y $E[(X_n - \mu)^4] = \rho$ para $n = 1, 2, \ldots$. Entonces

$$\frac{S_n}{n} \xrightarrow{c.s.} \mu,$$

donde $S_n = X_1 + X_2 + \cdots + X_n$.

Demostración

Por la desigualdad de Markov (Proposición 1.2.3) con $g(x) = x^4$, se tiene que

$$\begin{aligned}
P\left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right] &= P[|S_n - n\mu| > n\epsilon] \\
&= P\left\{\left|\sum_{i=1}^{n}(X_i - \mu)\right| > n\epsilon\right\} \\
&\leq \frac{1}{(n\epsilon)^4} E\left[\left(\sum_{i=1}^{n}(X_i - \mu)\right)^4\right] \\
&= \frac{1}{(n\epsilon)^4}[nE(X_1 - \mu)^4 + \\
&\quad +n(n-1)(E(X_1 - \mu)^2)^2] \qquad (2.5) \\
&\leq \frac{1}{(n\epsilon)^4} n^2[\rho + (\sigma^2)^2] \\
&\leq \frac{K}{n^2}.
\end{aligned}$$

La igualdad (2.5) se obtiene a partir de que las v.a.'s son independientes, de que $EX_n = \mu \ \forall n$ (lo cual a su vez implica que términos de la forma $E[(X_i - \mu)^3(X_j - \mu)] = 0$, $\forall i \neq j$) y de que la sucesión de v.a.'s tienen segundo momento central y cuarto momento centrales constantes.

Como K es una constante y $\sum_{n=1}^{\infty} 1/n^2 < \infty$, por el Lema de B-C o Corolario 1.3.3, concluimos

$$\frac{S_n}{n} \overset{c.s.}{\to} \mu.$$

□

Proposición 2.2.3 *Sea $\{X_n\}$ una sucesión de variables aleatorias independientes que cumplen*

$$\sigma_n^2 = Var(X_n) < \infty \qquad n = 1, 2, \ldots$$

y $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$, entonces $\sum_{n=1}^{\infty}(X_n - EX_n)$ converge c.s..

Demostración

Sea $S_n = \sum_{k=1}^{n} X_k$, por el Corolario 1.2.5 aplicado a las variables aleatorias $X_{n+1} - EX_{n+1}, \ldots, X_{n+m} - EX_{n+m}$, donde m y $n \in \mathcal{N}$, se tiene que para toda $\epsilon > 0$

$$P[\max_{1 \leq k \leq m} |S_{n+k} - S_n - ES_{n+k} + ES_n| \geq \epsilon] \leq \frac{1}{\epsilon^2} \sum_{k=1}^{m} \sigma_{n+k}^2 \tag{2.6}$$

Definimos $T_k = S_k - ES_k$, $\Delta_k = \sup_{v \geq 1}[|T_{k+v} - T_k|]$ y $\Delta = \inf_{k \geq 1} \Delta_k$.

Si tomamos límites en ambos lados de (2.6) cuando $m \to \infty$, se tiene

$$P[\Delta_n \geq \epsilon] \leq \frac{1}{\epsilon^2} \sum_{k=n+1}^{\infty} \sigma_k^2 \qquad \forall n \geq 1,$$

por lo que

$$P[\Delta \geq \epsilon] \leq \frac{1}{\epsilon^2} \sum_{k=n+1}^{\infty} \sigma_k^2.$$

Como $\sum_{k=1}^{\infty} \sigma_k^2 < \infty$, entonces $P[\Delta \geq \epsilon] = 0$; lo cual implica, a su vez, que la sucesión $\{S_n\}$ es Cauchy c.s..

La conclusión se sigue de la Proposición 1.3.1.

□

Antes de enunciar algunas otras Leyes de Grandes Números, se realizarán algunas definiciones y se verificarán algunos resultados importantes.

Lema 2.2.4 (de Toeplitz) *Sea $\{a_n\}$ una sucesión de números reales tal que $a_n \to a$ cuando $n \to \infty$. Entonces $n^{-1} \sum_{k=1}^{n} a_k \to a$ cuando $n \to \infty$.*

Lo que el Lema de Toeplitz nos dice, es que si una sucesión de números reales converge, entonces su media aritmética converge al mismo límite que la sucesión original.

Demostración
Como $a_n \to a$ para $\epsilon > 0$ sabemos que existe $n_0(\epsilon)$ tal que $\forall n \geq n_0(\epsilon)$

$$|a_n - a| < \frac{\epsilon}{2}.$$

Escogemos n_0^* tal que

$$\frac{1}{n_0^*}\sum_{k=1}^{n_0} |a_n - a| < \frac{\epsilon}{2}.$$

Si $n > n_0^*$

$$\begin{aligned} \left|\frac{1}{n}\sum_{k=1}^{n} a_k - a\right| &\leq \frac{1}{n_0^*}\sum_{k=1}^{n_0} |a_k - a| + \frac{1}{n}\sum_{k=n_0+1}^{n} |a_k - a| \\ &< \frac{\epsilon}{2} + \frac{n-n_0}{n}\cdot\frac{\epsilon}{2} < \epsilon \end{aligned}$$

□

Lema 2.2.5 (de Kronecker) *Sea $\{a_n\}$ una sucesión de reales tales que $\sum_{n\geq 1} a_n < \infty$. Entonces*

$$n^{-1}\sum_{k=1}^{\infty} k a_k \to 0$$

cuando $n \to \infty$.

Demostración
Sea $s_0 = 0$ y $s_n = \sum_1^n a_k$, $n = 1, 2, \ldots$. Se tiene entonces

$$\begin{aligned} \frac{1}{n}\sum_{k=1}^{n} k a_k &= \frac{1}{n}\sum_{k=1}^{n} k(s_k - s_{k-1}) \\ &= s_n - \frac{1}{n}\sum_{k=1}^{n-1} s_k. \end{aligned} \tag{2.7}$$

La serie $\{s_n\}$ converge a un límite finito s y a partir del Lema de Toeplitz, $n^{-1}\sum_{k=1}^{n-1} s_k \to s$ cuando $n \to \infty$

□

Teorema 2.2.6 *Sea $\{X_n\}$ una sucesión de v.a.i. tales que $Var(X_n) = \sigma_n^2 < \infty$, $n = 1, 2, \ldots$. Supongamos que $\sum_{n=1}^{\infty} \sigma_n^2/n^2 < \infty$ y sea $S_n = \sum_{k=1}^{\infty} X_k$, $n = 1, 2, \ldots$, entonces la sucesión $\{n^{-1}(S_n - ES_n)\}$ converge c.s. a cero.*

Este Teorema se deriva de la Proposición 2.2.3 y constituye una Ley Fuerte de los Grandes Números menos restrictiva que el Teorema 2.2.2, ya que solo requiere que exista el segundo momento centrado y no requiere que el primer y segundo momento central sean iguales para cada X_n $n \in \mathcal{N}$.

Demostración

Definimos $Y_n = (X_n - EX_n)/n$, $n = 1, 2, \ldots$. Se sigue que $EY_n = 0$ y $\mathrm{Var}(Y_n) = \sigma_n^2/n^2$. Por la hipótesis, tenemos $\sum_{n=1}^{\infty} \mathrm{Var}(Y_n) < \infty$. De la Proposición 2.2.3 se obtiene que $\sum_{n=1}^{\infty} Y_n$ converge c.s. y aplicando el Lema de Kronecker a la sucesión $\{Y_n\}$ se obtiene el resultado.

□

Definición 2.4 *Decimos que dos sucesiones de variables aleatorias $\{X_n\}$ y $\{X'_n\}$ son equivalentes (tail-equivalent) si difieren c.s. en un número finito de términos.*

La condición anterior también se puede expresar como

$$P[\limsup(X_n - X'_n) \neq 0] = P[X_n \neq X'_n \ \ i.o.] = 0.$$

Definición 2.5 *Si las sucesiones de variables aleatorias $\{X_n\}$ y $\{Y_n\}$ convergen en el mismo evento (excepto para eventos de medida de probabilidad cero), entonces decimos que $\{X_n\}$ y $\{Y_n\}$ son equivalentes en convergencia.*

Proposición 2.2.7 *Si $\{X_n\}$ y $\{X'_n\}$ son dos sucesiones de variables aleatorias tales que $\sum_{n=1}^{\infty} P[X_n \neq X'_n] < \infty$, entonces las sucesiones $\{X_n\}$ y $\{X'_n\}$ son equivalentes y las series $\{b_n^{-1} \sum_{k=1}^{\infty} X_k\}$ y $\{b_n^{-1} \sum_{k=1}^{\infty} X'_k\}$, donde $b_n \uparrow \infty$, convergen en el mismo evento al mismo límite, excluyendo al evento nulo.*

Demostración

Por el Lema de Borel-Cantelli, se tiene que

$$P[X_n \neq X'_n \ \ i.o.] = 0.$$

La segunda parte se sigue inmediantamente.

□

A continuación se presenta la Ley Fuerte de Grandes Números de Kolmogorov. A diferencia de las versiones de LGN presentadas hasta ahora, la que desarrolló Kolmogórov tiene una doble implicación, dándonos condiciones necesarias y suficientes (referente a la convergencia de la suma estandarizada de v.a.'s) para la existencia del primer momento. Cuando la suma estandarizada de v.a. converge o el primer momento existe, también nos dice cuál es el límite de convergencias de la suma de v.a.'s.

Teorema 2.2.8 (Ley de los Grandes Números de Kolmogórov) *Sea* $\{X_n\}$ *una sucesión de variables aleatorias independientes e idénticamente distribuidas (v.a.i.i.d.) Sea* $S_n = \sum_{k=1}^{n} X_k$, $n = 1, 2, \ldots$. *Entonces la sucesión* $\{n^{-1}S_n\}$ *converge a un límite finito* α $\Leftrightarrow$ $E|X_1| < \infty$. *Mas aún, en ese caso* $EX_1 = \alpha$.

Demostración

Supongamos $E|X_1| < \infty$ y X v.a. con la misma distribución que X_1.

Sea $E_0 = \Omega$ y $E_n = [|X| \geq n]$, $n = 1, 2, \ldots$.

Como $E_n \downarrow \emptyset$ entonces $E_n = \cup_{k=n}^{\infty}(E_k - E_{k+1})$, $n = 1, 2, \ldots$ es una unión ajena de eventos y

$$P(E_n) = \sum_{k=n}^{\infty} P(E_n - E_{n+1}).$$

Lo que a su vez nos lleva a que

$$\sum_{n=1}^{\infty} P(E_n) = \sum_{n=1}^{\infty} nP(E_n - E_{n+1}).$$

Como $E_0 = \Omega$ entonces

$$\begin{aligned} 1 + \sum_{n=1}^{\infty} P(E_n) &= \sum_{n=0}^{\infty} P(E_n) \\ &= \sum_{n=1}^{\infty} nP(E_{n-1} - E_n). \end{aligned}$$

Por el Lema 1.2.4 se tiene que

$$\sum_{n=1}^{\infty} P(E_n) \leq E|X| \leq 1 + \sum_{n=1}^{\infty} P(E_n),$$

relación de la cual se llega a que

$$\sum_{n=1}^{\infty} P(E_n) < \infty. \tag{2.8}$$

Para todo $n \in \mathcal{N}$ definimos X_n^* como

$$X_n^* = X_n I_{[|X_n| < n]}.$$

Entonces

$$\begin{aligned} \mathrm{Var}(X_n^*) &\leq EX_n^{*2} \\ &= \int_{[|X_n| < n]} X_n^2 dP \\ &= \int_{E_n^c} X^2 dP \\ &\leq \sum_{k=1}^{n} k^2 P(E_{k-1} - E_k), \end{aligned}$$

a partir de lo cual tenemos

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{\mathrm{Var}(X_n^*)}{n^2} &\leq \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{k^2}{n^2}P(E_{k-1}-E_k)\\ &= \sum_{k=1}^{\infty}k^2P(E_{k-1}-E_k)\sum_{n=k}^{\infty}\frac{1}{n^2}.\end{aligned}$$

Para $k \geq 1$ se tiene que

$$\begin{aligned}\sum_{n=k}^{\infty}\frac{1}{n^2} &\leq \frac{1}{k^2}+\int_k^{\infty}\frac{dx}{x^2}\\ &= \frac{1}{k^2}+\frac{1}{k}\\ &\leq \frac{2}{k}.\end{aligned}$$

Entonces

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{\mathrm{Var}(X_n^*)}{n^2} &\leq 2\sum_{k=1}^{\infty}kP(E_{k-1}-E_k)\\ &= 2[1+\sum_{n=1}^{\infty}P(E_n)]\\ &< \infty \qquad \text{por (2.8)}.\end{aligned}$$

Definimos $S_n^* = \sum_{k=1}^{n} X_k^*$, $n = 1, 2, \ldots$. A partir del Teorema 2.2.6, aplicado a $\{X_n^*\}$ podemos concluir que

$$\frac{S_n^* - ES_n^*}{n} \xrightarrow{c.s.} 0 \quad \text{cuando } n \to \infty.$$

Ahora se mostrará que $n^{-1}ES_n^* \to EX$. Para todo $n \in \mathcal{N}$ tenemos que

$$EX_n^* = \int_{[|X_n|<n]} X_n dP = \int_{E_n^c} XdP = E(XI_{E_n^c}).$$

Como $E_n^c \uparrow \Omega$, se tiene que $XI_{E_n^c} \xrightarrow{c.s.} X$ cuando $n \to \infty$ y usando el Teorema de Convergencia Dominada de Lebesgue se concluye que $EX_n^* \to EX$.

Por el Teorema de Toeplitz se tiene que $n^{-1}ES_n^* \to EX$.

Notemos ahora que $\{X_n\}$ y $\{X_n^*\}$ son equivalentes, por lo que

$$\frac{S_n - S_n^*}{n} \xrightarrow{c.s.} 0.$$

Entonces $n^{-1}X_n \xrightarrow{c.s.} EX$ cuando $n \to \infty$.

Supongamos que la sucesión $\{n^{-1}S_n\}$ converge c.s. a un limite finito α. Como

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{n-1}{n}\cdot\frac{S_{n-1}}{n-1},$$

podemos concluir que $\frac{X_n}{n} \xrightarrow{c.s.} 0$ cuando $n \to \infty$. Gracias al Lema de B-C tenemos que $\sum_{n=1}^{\infty} P[|X_n| \geq \epsilon] < \infty$ para $\epsilon > 0$. En particular, $\sum_{n=1}^{\infty} P(E_n) < \infty$ y entonces $E|X| < \infty$. La conclusión sigue de la primera parte del Teorema

$\square$

El siguiente Teorema es una variación de la Ley Fuerte de los Grandes Números, la cual no solo es elemental, en el sentido de que no utiliza la desigualdad de Kolmogórov, sino que es más aplicable porque solamente se requiere que las variables aleatorias sean independientes por pares. Sin embargo, requiere que $X_1 \in L_1$. Por otro lado, a diferencia de las LFGN correspondientes a los Teoremas 2.2.2 y 2.2.6, sólo requiere que exista el primer momento, pero necesita que las v.a.'s tengan la misma distribución.

Teorema 2.2.9 (Etemadi) *Sea $\{X_n\}$ una sucesión de v.a. independientes por pares e idénticamente distribuidas. Si $S_n = \sum_{i=1}^{n} X_i$ entonces tenemos*

$$E|X_1| < \infty \Longrightarrow \lim_{n\to\infty} \frac{S_n}{n} = EX_1 \quad c.s.$$

Demostración

Como $E|X_1| < \infty \Leftrightarrow EX_1^+$ y EX_1^- son finitas y $X_i = X_i^+ - X_i^-$ para $i = 1, 2, \ldots$, sin pérdida de generalidad se puede asumir que $X_i \geq 0$. Sea X v.a. con la misma distribución que X_1.

Sea $Y_i = X_i I_{[X_i \leq i]}$ y sea $S'_n = \sum_{i=1}^{\infty} Y_i$. Para $\epsilon > 0$ sea $k_n = [\![\alpha^n]\!]$, $\alpha > 1$, donde $[\![x]\!]$ representa a la función *mayor entero menos o igual que x*. Usando la desigualdad de Chebyshev tenemos

$$\begin{aligned}
\sum_{n=1}^{\infty} P\left[\left|\frac{S'_{k_n} - ES'_{k_n}}{k_n}\right| > \epsilon\right] &\leq \sum_{n=1}^{\infty} \frac{1}{k_n^2} \cdot \frac{\mathrm{Var} S'_{k_n}}{\epsilon^2} = c\sum_{n=1}^{\infty} \frac{\mathrm{Var} S'_{k_n}}{k_n^2} \\
&= c\sum_{n=1}^{\infty} \frac{1}{k_n^2} \sum_{i=1}^{k_n} \mathrm{Var} Y_i \leq c\sum_{i=1}^{\infty} \frac{EY_i^2}{i^2} \\
&= c\sum_{i=1}^{\infty} \frac{1}{i^2} \int_0^i X^2 dF(x) \\
&= c\sum_{i=1}^{\infty} \frac{1}{i^2} \left(\sum_{k=0}^{i-1} \int_k^{k+1} X^2 dF(x)\right) \\
&\leq c\sum_{k=0}^{\infty} \frac{1}{k+1} \int_k^{k+1} X^2 dF(x) \\
&\leq c\sum_{k=0}^{\infty} \int_k^{k+1} X dF(x) \\
&= cEX < \infty,
\end{aligned}$$

donde $F(x)$ es la función de distribución de X_1 y c es una constante positiva.

Además tenemos por el Lema de Toeplitz (Lema 2.2.4),

$$EX = \lim_{n\to\infty} \int_0^n X dF(x) = \lim_{n\to\infty} EY_n = \lim_{n\to\infty} \frac{ES'_{k_n}}{k_n}.$$

Por el Lema de Borel-Cantelli y la Proposición 1.3.2 o el Corolario 1.3.3 obtenemos entonces

$$\lim_{n\to\infty} \frac{S'_{k_n}}{k_n} = EX \quad c.s.$$

Además

$$\begin{aligned}
\sum_{n=1}^{\infty} P[Y_n \neq X_n] &= \sum_{n=1}^{\infty} P[X_n > n] = \sum_{n=1}^{\infty} \int_n^{\infty} dF(x) \\
&= \sum_{n=1}^{\infty} \sum_{i=n}^{\infty} \int_i^{i+1} dF(x) = \sum_{i=1}^{\infty} i \int_i^{i+1} dF(x) \\
&\leq \sum_{i=1}^{\infty} \int_i^{i+1} x dF(x) \leq EX_1 \leq \infty. \qquad (2.9)
\end{aligned}$$

Por la Proposición 2.2.7 se tiene que $Y_n = X_n$ c.s. en un número finito de términos y además

$$\lim_{n\to\infty} \frac{S_{k_n}}{k_n} = EX_1 \quad c.s.$$

Como $X_i \geq 0$, entonces S_n es monótona y por esta razón se puede concluir que

$$\frac{1}{\alpha} EX_1 \leq \liminf_n \frac{S_n}{n} \leq \limsup_n \frac{S_n}{n} \leq \alpha EX_1 \quad c.s. \ \forall \ \alpha > 1.$$

□

Capítulo 3

MARTINGALAS

A principios del presente siglo, Bernstein y Lévy introdujeron el concepto de martingalas en forma de sumas consecutivas de variables aleatorias y aportaron algunas generalizaciones de resultados que eran relacionados a límites de suma de variables aletorias independientes.

El nombre de martingala apareció en la literatura de la probabilidad moderna en el año de 1939 y fué el tema sobre el cual el matemático Doob dió valiosos resultados durante la década de los cuarentas y principios de los cincuentas.

La teoría de martingalas, como la misma teoría de probabilidad, tiene sus orígenes en la teoría de juegos (evidencia de ésto es que la idea de martingala expresa el concepto de un juego justo), pero el trabajo del matemático J. L. Doob fué el que cambió totalmente la dirección de la materia con su libro (1953) [5], el cual ha sido básico para el estudio del tema por casi tres décadas.

Los resultados de martingalas ahora representan una importante herramienta, con la cual pueden generalizarse resultados que requieren de condiciones de independencia.

3.1 INTRODUCCIÓN

Definición 3.1 *Sean* $\{X_n\}_{n\geq 1}$ *una sucesión de v.a.'s y* $\{\Im_n\}_{n\geq 1} = \{\Im_n\}$ *una familia de* σ*-álgebras tal que* $\Im_n \subset \Im_m \subset \Im$, $\forall n \leq m$ *(en este caso se dice que* $\{\Im_n\}$ filtración de $\Im$*), para las cuales* X_i *es* $\Im_i$*-medible* $\forall i \in \mathcal{N}$. *Entonces* $\{X_n\}_{n\in\mathcal{N}}$, *se denomina* martingala, *si y sólo si* $X_i \in L_1 \quad \forall i \in \mathcal{N}$ *y*

$$X_n = E[X_m|\Im_n] \qquad c.s. \ \ \forall n \leq m.$$

La sucesión $\{X_n\}$ *se llama supermartingala o submartingala si el signo '=' es reemplazado por '*$\leq$*' o '*$\geq$*', respectivamente.*

Nota 3.2 *Diremos que que una sucesión de v.a.,* $\{X_n\}$, *es martingala (sub o supermartingala), sin mencionar las* σ*-álgebras* $\{\Im_n\}$, *cuanto éstas sean la inducidas por las primeras* n *variables aleatorias, es decir, cuando* $\Im_n = \sigma(X_1, \ldots, X_n)$.

Definición 3.3 *Diremos que la sucesión de v.a.'s* $\{X_n\}$ *es una sucesión* adaptada a la filtración $\{\Im_n\}$ *si* X_n *es* $\Im_n$*-medible para* $n = 1, 2, \ldots$. *En este caso denotaremos a* $\{X_n\}$ *como* $\{X_n, \Im_n\}_{n\geq 1}$.

Proposición 3.1.1 *Si* $X_n \in L_1 \ \ \forall n \in \mathcal{N}$, *la sucesión de variables aleatorias* $\{X_n, \Im_n\}$ *es una martingala, si y sólo si* $X_n = E[X_{n+1}|\Im_n] \ \ \forall n$. *Análogamente, el resultado es cierto para supermatingalas y submartingalas.*

Demostración

Si $\{X_n, \Im_n\}$ es una martingala entonces, es inmediato que $\Im_n \subset \Im_{n+1}$, $X_n \in L_1$ y $X_n = E[X_{n+1}|\Im_n]$.

Supongamos ahora que $m > n$. Entonces $E[X_m|\Im_n] = E[E[X_m|\Im_{m-1}]|\Im_n]{=}E[X_{m-1}|\Im_n]$. Inductivamente se obtiene que

$$E[X_m|\Im_n] = X_n \quad \text{c.s.}$$

$\square$

Si X_n es la ganancia en la jugada $n-$ésima, en una partida de juego, entonces $S_n = \sum_1^n X_i$ es la ganancia total después de n jugadas. La fortuna esperada, dada la información de las ganancias anteriores, después de $n+1$ jugadas es de

$$E[S_{n+1}|S_1, \ldots, S_n] \quad (= E[S_{n+1}|\sigma(S_1, \ldots, S_n)]).$$

El juego es *justo* si $E[S_{n+1}|S_1, \ldots, S_n] = S_n$ (martingala), ya que la ganancia esperada en el juego $n+1$, dada la ganancia hasta la jugada $n-$ésima, es igual a

$$E[S_{n+1} - S_n|S_1, \ldots, S_n] = S_n - S_n = 0.$$

Por la misma observación, se tienen que el juego es favorable si $\{S_n, \sigma(S_1, \ldots, S_n)\}$ es submartingala y es desfavorable si es supermartingala.

Como la $\sigma-$ álgebra $\Im_n = \sigma(S_1, S_2, ..., S_n)$ representa la influencia del pasado al tiempo n, intuitivamente $\Im_n$ debe tener la misma información que $\sigma(X_1, X_2, ..., X_n)$. Este resultado se demuestra en el Lema siguiente

Lema 3.1.2 $\sigma(X_1, X_2, ..., X_n) = \sigma(S_1, S_2, ..., S_n)$

Demostración

Como $S_1 = X_1$ y $S_n = \sum_{i=1}^n X_i$, $S_1, S_2, S_3, \ldots$ son $\sigma(X_1, X_1 + X_2....)$ -medibles,

$$\sigma(S_1, S_2, S_3, ...) \subset \sigma(X_1, X_1 + X_2....).$$

Similarmente, $X_1, X_1 + X_2, X_1 + X_2 + X_3$ son $\sigma(S_1, S_2, S_3, ...)$-medibles, por lo que $\sigma(X_1, X_1 + X_2....) \subset \sigma(S_1, S_2, S_3, ...)$

□

Una propiedad importante que cumplen las X_i (definidas como anteriormente) es que son ortogonales, es decir, $EX_iX_j = 0 \ \ \forall \ i \neq j$ si $S_n \in L_2 \quad \forall n \in \mathcal{N}$.

Teorema 3.1.3 *Si $\{S_n\}$ es una martingala y $S_n \in L_2$ $\forall n \geq 1$, entonces la sucesión $\{X_n\}$ de v.a.'s definidas como $X_1 = S_1$, $X_n = S_n - S_{n-1}$, para $n \geq 2$, es una sucesión de v.a. ortogonales.*

Demostración

Si $j < k$ y $\Im_j = \sigma\left(S_1, ..., S_j\right)$,

$$\begin{aligned} EX_jX_k &= E[(S_j - S_{j-1})(S_k - S_{k-1})] \\ &= E\left[[E((S_j - S_{j-1})(S_k - S_{k-1}))|\Im_j]\right] \\ &= E\left[(S_j - S_{j-1})E[S_k - S_{k-1}|\Im j]\right], \end{aligned}$$

pero $E\left[S_k - S_{k-1}|\Im_j\right] = S_j - S_j = 0$.

□

Nótese que $\{X_n, \Im_n\}$ es martingala si y sólo si $\int_A X_n = \int_A X_{n+1} dP$ para toda $A \in \Im_n$, $n = 1, 2, \ldots$. Esto se sigue a partir de la Proposición 3.1.1 y de la definición de esperanza condicional. Similarmente, $\{X_n, \Im_n\}$ es submartingala si y sólo si

$$\int_A X_n dP \leq \int_A X_{n+1} dP \quad \forall A \in \Im_n$$

y es supermartingala si y sólo si

$$\int_A X_n dP \geq \int_A X_{n+1} dP \quad \forall A \in \Im_n.$$

Nota 3.4 *Si $A = \Omega$, tenemos que EX_n es constante en una martingala, creciente en una submartingala y decreciente en una supermartingala.*

Como ejemplo de martingala tenemos: sean $Y \in L_1$ y $\{\Im_n\}$ una filtración de $\Im$ y $X_n = E[Y|\Im_n]$, entonces $\{X_n, \Im_n\}$ es una martingala. Esta afirmación se verifica fácilmente a partir de las propiedades de esperanza condicional

$$\begin{aligned} E[X_{n+1}|\Im_n] &= E[E[Y|\Im_{n+1}]|\Im_n] \\ &= E[Y|\Im_n] = X_n. \end{aligned}$$

Teorema 3.1.4 (Desigualdad de Jensen) *Si $g : I \to \mathcal{R}$, donde I es un intervalo abierto de $\mathcal{R}$, g es convexa y X es una v.a. tal que $X(\omega) \in I$ para $\omega \in \Omega$, entonces se tiene que si $EX < \infty$ y $\wp \subset \Im$,*

$$E[g(X)|\wp] \geq g(E[X|\wp]) \quad c.s..$$

En particular $E[g(X)] \geq g(E[X]$.

Demostración

Nótese que $E[X|\wp](\omega) \in I$ c.s. y que si existe $a \in \mathcal{R}$ tal que $X > a$, entonces $E[X|\wp] > a$ c.s. porque

$$0 \geq \int_{\{E[X|\wp] \leq a\}} E[X-a|\wp]dP = \int_{\{E[X|\wp] \leq a\}} (X-a)dP \geq 0$$

por definición de esperanza condicional, lo cual implica que $X = a$ en $\{E[X|\wp] \leq a\}$. Se tiene entonces que $P\{E[X|\wp] \leq a\} = 0$.

Como g es convexa, para toda $y \in I$, existen dos sucesiones de reales, $\{a_n\}$ y $\{b_n\}$, tales que $g(y) = \sup_n(a_n y + b_n)$. Entonces $g(X) \geq a_n X + b_n$ para toda n.

Aplicando lo que se acaba de mostrar, tenemos que

$$E[g(X)|\wp] \geq a_n E[X|\wp] + b_n \quad \text{c.s..}$$

Si tomamos el supremo sobre n, tenemos la demostración completa.

□

Gracias a la desigualdad de Jensen, podemos afirmar que si $\{X_n\}$ es una submartingala, también lo es $\{X_n^+\}$ y que si $\{X_n\}$ es una martingala, entonces $\{|X_n|^r\}$, para $r \geq 1$, es una submartingala.

Teorema 3.1.5 (de salto de Halmos) *Sea $\{X_n, \Im_n\}$ una submartingala y sea*

$$\epsilon_1, \epsilon_2, \ldots$$

v.a.'s definidas como

$$\epsilon_k = \begin{cases} 1 & si\ (X_1, \ldots X_k) \in B_k \\ 0 & si\ (X_1, \ldots X_k) \notin B_k, \end{cases}$$

donde $B_k \subset \mathcal{B}(\mathcal{R}^n)$, $k = 1, 2, \ldots$. Sea $Y_1 = X_1$ y $Y_n = X_1 + \epsilon_1(X_2 - X_1) + \cdots + \epsilon_{n-1}(X_n - X_{n-1})$, para $n \geq 2$. Entonces $\{Y_n, \Im_n\}$ es también una submartingala y $EY_n \leq EX_n$ para toda n. Si $\{X_n, \Im_n\}$ es una martingala, entonces $\{Y_n, \Im_n\}$ también lo es y $EY_n = EX_n$.

Si reemplazamos a X_n por S_n, la ganancia después de n jugadas, se tiene que, Y_n representa la ganancia después de n jugadas siguiendo una estrategia de juego. Después de observar $S_1, \ldots, S_{n-1}$ podemos escoger el jugar o no en el siguiente turno ($\epsilon_{n-1} = 1, \epsilon_{n-1} = 0$, respectivamente). En caso de que $\epsilon_{n-1}(S_1, \ldots, S_{n-1}) = 1$, la ganancia en la jugada n es de $S_n - S_{n-1}$. Lo que el Teorema nos dice, es que no importa la estrategia que se siga, el juego seguirá siendo favorable (justo).

Demostración

$$\begin{aligned} E[Y_{n+1}|\Im_n] &= E[Y_n + \epsilon_n(X_{n+1} - X_n)|\Im_n] \\ &= Y_n + \epsilon_n E[X_{n+1} - X_n|\Im_n]. \end{aligned}$$

Entonces $E[Y_{n+1}|\Im_n]$ es mayor a Y_n cuando $\{X_n, \Im_n\}$ es submartingala y es igual a Y_n cuando $\{X_n, \Im_n\}$ es martingala.

La segunda parte se demostrará por inducción sobre k.

Como $Y_1 = X_1$ entonces $EX_1 = EY_1$.

Supongamos $E(X_k - Y_k) \geq 0$ (e igual a cero, cuando $\{X_n, \Im_n\}$ es martingala), entonces

$$\begin{aligned} X_{k+1} - Y_{k+1} &= X_{k+1} - Y_k - \epsilon_k(X_{k+1} - X_k) \\ &= (1 - \epsilon_k)(X_{k+1} - X_k) + X_k - Y_k. \end{aligned}$$

La esperanza de $X_{k+1} - Y_{k+1}$ dado $\Im_n$, está determindada así, por la siguiente expresión

$$\begin{aligned} E[X_{k+1} - Y_{k+1}|\Im_k] &= (1 - \epsilon_k)E[X_{k+1} - X_k|\Im_k] + E[S_k + Y_k]|\Im_k \\ &\geq E[X_k - Y_k|\Im_k] = X_k - Y_k. \end{aligned} \tag{3.1}$$

Obtenemos entonces que

$$E[X_{k+1} - Y_{k+1}] \geq E(X_k - Y_k) \geq 0. \tag{3.2}$$

Si $\{X_n, \Im_n\}$ es martingala, entonces se tene igualdad en (3.1) y en (3.2).

□

Teorema 3.1.6 (de cruce) *Sea $\{X_k, \Im_k, k = 1, \ldots, n\}$ una submartingala. Si a y b son dos números reales tales que $a < b$, definimos U_{ab} como el número de veces que las v.a.'s salen del intervalo (a, b).*

Sea $T_1 = T_1(\omega)$ el primer entero en $\{1, 2, \ldots, n\}$ tal que $X_{T_1} \leq a$, T_2 el primer entero mayor que T_1 tal que $X_{T_2} \geq b$, T_3 el primer entero mayor que T_2 tal que $X_{T_3} \leq a$, y así sucesivamente definimos T_j, con $j \in \mathcal{N}$. Definimos a $T_i = \infty$ si la condición no se satisface.

Si N es la cantidad de T_i que son finitas, definimos a U_{ab} como $U_{ab} = N/2$ si N es par y $U_{ab} = (N-1)/2$ si N es impar, entonces

$$EU_{ab} \leq \frac{1}{b-a} E[(X_n - a)^+].$$

Demostración

Supongamos $a = 0$ y que $X_i \geq 0$ para toda i en $\{1, \ldots, n\}$. Definimos a T_i como antes y a ϵ_i como

$$\epsilon_i = \begin{cases} 0 & \text{Si } i < T_1 \\ 1 & T_1 \leq i < T_2 \\ 0 & T_2 \leq i < T_3 \\ 1 & T_3 \leq i < T_4 \\ \vdots & \ddots \end{cases}$$

Entonces

$$X_1 + \epsilon_1(X_2 - X_1) + \cdots + \epsilon_{n-1}(X_n - X_{n-1}) = X_1 + X_{T_2} - X_{T_1} + X_{T_4} - X_{T_3} + \cdots.$$

Si $Y_j = X_1 + \epsilon_1(X_2 - X_1) + \cdots + \epsilon_{j-1}(X_j - X_{j-1})$, $j = 1, \ldots, n$, entonces Y_j la suma de los incrementos desde que se es igual a cero hasta que se es igual o se rebasa b, más X_1 . Entonces

$$Y_n \geq b \cdot U_{ab}.$$

Como ϵ_j puede ser expresado en términos de $X_1, \ldots, X_j$, $\{Y_k, \Im_k, k = 1, \ldots, n\}$ es una submartingala (por el Teorema 3.1.5) y $EY_n \leq EX_n$. Se tiene así que

$$EU_{ab} \leq \frac{1}{b} EY_n \leq \frac{1}{b} EX_n.$$

En general, $\{(X_k - a)^+, \Im_k, k = 1, \ldots, n\}$ es una submartingala y el número de veces que $\{X_k, \Im_k, k = 1, \ldots, n\}$ *sale* de (a, b) es el mismo de que veces que $\{(X_k - a)^+, \Im_k, k = 1, \ldots, n\}$ sale de $(0, b-a)$.

$\square$

A continuación se presenta un Teorema de convergencia de martingalas.

Teorema 3.1.7 *Sea $\{X_n\}$ es una martingala, si*

$$\lim E|X_n| = K < \infty,$$

entonces $\lim_{n\to\infty} X_n = X_\infty$ *existe c.s. y* $E|X_\infty| \leq K$.

Nota 3.5 *Siempre podemos hablar del* $\lim E|X_n|$ *(aunque sea a infinito) ya que* $\{E|X_n|\}$ *es una sucesión creciente (Por Nota 3.4 y Desigualdad de Jensen).*

Demostración

El subconjuto A, de Ω donde $\{X_n\}$ no converge es

$$\{\limsup_n X_n(\omega) - \liminf_n X_n(\omega) \neq 0\} = \cup_{r_1,r_2 \in \mathcal{Q}} \{\limsup_n X_n(\omega) > r_2 > r_1 > \liminf_n X_n(\omega)\}.$$

Si $A_{r_1,r_2} = \{\limsup_n X_n > r_2 > r_1 > \liminf_n X_n\}$ para $r_1, r_2 \in \mathcal{Q}$ y $U_{r_1 r_2,n}$ es el número de veces que las v.a. $X_1, \ldots, X_n$ salen del intervalo $[r_1, r_2]$, entonces $U_{r_1 r_2,n} \to \infty$ cuando $n \to \infty$, en A_{r_1,r_2}. Por otro lado, el Teorema 3.1.6 nos dice

$$EU_{r_1 r_2,n} \leq \frac{1}{r_2 - r_1} E[(X_n - r_1)^+] \leq \frac{1}{r_2 - r_1} (K + |r_1|).$$

Estas dos afirmaciones, junto con el hecho de que $K < \infty$, señalan que $P(A_{r_1,r_2}) = 0$. Entonces $\limsup_n X_n = \liminf_n X_n$ c.s., es decir que el $\lim_{n\to\infty} X_n$ existe (finito o infinito) con probabilidad 1.

Por el Lema de Fatou,

$$E|X_\infty| \leq \lim_{n\to\infty} E|X_n| = K,$$

donde $\lim_{n\to\infty} X_n = X_\infty$. Entonces X_∞ es integrable y finito c.s..

$\square$

Como caso particular, podemos observar que si las v.a. son no negativas o no positivas, entonces

$$E|X_n| = EX_n = constante$$

y

$$E|X_n| = -EX_n = constante,$$

respectivamente y para toda n.

3.2 TIEMPO DE PARO Y TEOREMA DE PARO OPCIONAL

Sea $\{S_n, \Im_n\}$ una martingala. Si nuevamente consideramos a S_n como el capital total de un jugador después de n jugadas y el jugador puede decidir dejar la partida después de la τ-ésima jugada, queremos conocer lo que se puede decir de S_τ, que es el capital final.

Nótese que τ debe ser una v.a. con la propiedad de que dados $S_1, \ldots, S_n$, se puede tomar la desición de seguir o no jugando.

Definición 3.6 *Sea $\{\Im_n\}_{n\geq 1}$ una filtración de $\Im$. Decimos que τ es un tiempo de paro para $\{\Im_n\}$ si es un mapeo $\tau : \Omega \to \{0, 1, \ldots, \infty\}$ tal que $\{\tau \leq n\} \in \Im_n$ para cada $n \in N$.*

Como $\{\tau = n\} = \{\tau \leq n\} - \{\tau \leq n-1\}$ y $\{\tau \leq n\} = \cup_0^n \{\tau = k\}$, la definición es equivalente al requerimiento de que $\{\tau = n\} \in \Im_n$ para $n = 0, 1, \ldots$.

Si $\{S_n\}$ es una sucesión de v.a.'s, un *tiempo de paro para* $\{S_n\}_{n\geq 1}$ es el tiempo de paro relativo a $\{\sigma(S_0, S_1, \ldots, S_n)\}_{n\geq 0}$.

Uno de los ejemplos más importantes de tiempo de paro es *the hitting time* o tiempo de *llegada* a un conjunto B.

Sea $\{X_n\}$ es una sucesión de v.a.'s y $B \in \mathcal{B}(\mathcal{R})$, definimos $\tau(\omega) = min\{n \in \mathcal{N} | X_n(\omega) \in B\}$ si $X_n(\omega) \in B$ para algún n, $\tau(\omega) = \infty$ si $X_n(\omega)$ nunca *llega* a B. τ definida así, es un tiempo de paro para $\{X_n\}_{n\geq 1}$, ya que

$$\{\tau \leq n\} = \cup_{k\leq n} \{X_k \in B\} \in \sigma(X_k, k \leq n).$$

Si $\{S_n\}$ es una sucesión de v.a.'s, donde S_n representa la fortuna del jugador después de la n-ésima jugada, τ es un tiempo de paro para $\{S_n\}_{n\geq 1}$ y

$$S_j^*(\omega) = \begin{cases} S_j(\omega) & \text{si } j \leq \tau(\omega) \\ S_{\tau(\omega)}(\omega), & \text{si } j > \tau(\omega), \end{cases} \tag{3.3}$$

entonces $S_j^*(\omega)$ representa la jugada con la estrategia de *salida* o tiempo de paro τ.

La trasformación (3.3) se denomina como la *transformación bajo el Sistema de Paro Opcional*.

Lo que el siguiente Teorema nos dice, es que no importa la estrategia de salida que se siga, el juego permanaecerá siendo favorable o justo es que así era en un principio.

El Teorema se enunciará, fuera del contexto de juegos, para $\{X_n, \Im_n\}$ martingala.

Teorema 3.2.1 (Paro Opcional) *Supongamos que la submartingala (martingala)* $\{X_n\}$ *se trasfoma según (3.3), es decir, sea* τ *un tiempo de paro para* $\{X_n\}_{n\geq 1}$ *y*

$$Y_j(\omega) = \begin{cases} X_j(\omega) & \textit{si } j \leq \tau(\omega) \\ X_{\tau(\omega)}(\omega), & \textit{si } j > \tau(\omega). \end{cases} \tag{3.4}$$

Entonces $\{Y_n\}$ *es también una submartingala (martingala) y*

$$EX_1 \leq EY_n \leq EX_n \quad (EX_1 = EY_n) \quad \textit{para } n \geq 1.$$

Demostración

Como Y_n es igual, por partes, a $X_1, \ldots, X_n$, entonces

$$E|Y_n| \leq \sum_1^n E|X_j| < \infty.$$

Supongamos que $\{X_n\}$ es una submartingala (martingala). Se tiene que demostrar que

$$\int_F Y_{n+1} dP \overset{\geq}{(=)} \int_F Y_n dP \tag{3.5}$$

para todo $F \in \sigma(Y_1, \ldots, Y_n)$.

Como $Y_j = Y_{\tau(\omega)}$ si $j \geq \tau(\omega)$, entonces

$$\int_{F\cap\{\tau(\omega)\leq n\}} Y_{n+1}dP = \int_{F\cap\{\tau(\omega)\leq n\}} Y_n dP.$$

Ahora notemos que $Y_1, \ldots, Y_n$ son v.a.'s definidas en términos de $X_1, \ldots, X_n$, por lo que $F \in \sigma(X_1, \ldots, X_n)$. Entonces $F \cap \{\tau(\omega) \leq n\} \in \sigma(X_1, \ldots, X_n)$. Utilizando la propiedad de submartingala (martingala) de $\{X_n\}$ obtenemos

$$\int_{F\cap\{\tau(\omega)>n\}} X_{n+1}dP \overset{\geq}{(=)} \int_{F\cap\{\tau(\omega)>n\}} X_n dP.$$

A partir de que $X_{n+1} = Y_{n+1}$ y $X_n = Y_n$ en $\{\tau(\omega) > n\}$, se tiene, conjuntando ambos resultados, que (3.5) se cumple.

Procederemos ahora con la demostración de la segunda parte.

Supongamos que $\{X_n\}$ es submartingala (martingala). Como $\{Y_n\}$ es submartingala (martingala) y $X_1 = Y_1$, entonces

$$EX_1 = EY_1 \overset{\leq}{(=)} EY_n \qquad \text{para toda } n \geq 1.$$

Además

$$\begin{aligned} EY_n &= \int_{\{\tau(\omega)\geq n\}} X_n dP + \sum_{j=1}^{n-1} \int_{\{\tau(\omega)=j\}} X_j dP \\ &\leq EX_n. \end{aligned}$$

□

Definición 3.7 *Definimos a la σ-álgebra $\Im_\tau$ como*

$$\Im_\tau = \{E \in \Im | E \cap [\tau = k] \in \Im_k, 1 \leq k < \infty\}.$$

$\Im_\tau$ se denomina como la σ-álgebra parada en τ o la σ-álgebra de eventos hasta el tiempo aleatorio τ.

Para tener una idea de lo que significa $\Im_\tau$, consideremos el siguiente caso. Si $\tau(\omega) < \infty$ para toda ω y $\Im_n = \sigma(X_1, \ldots, X_n)$, entonces $I_E(\omega) = I_E(\omega')$ para cualquier E en $\Im_\tau$ si y sólo si $X_i(\omega) = X_i(\omega')$ para $i \leq \tau(\omega) = \tau(\omega')$. Es decir, la información en $\Im_\tau$ consiste de los valores $\tau(\omega), X_1(\omega), \ldots, \Im_{\tau(\omega)}(\omega)$.

Proposición 3.2.2 *Si $\{X_n\}$ es una sucesión adaptada a $\{\Im_n\}$, $a \in \mathcal{R}$ y τ es un tiempo de paro con respecto a la misma filtración, entonces la v.a. $X_{min\{\tau(\omega),a\}} = X_{\tau\wedge a}$ es $\Im_\tau$ medible.*

Demostración Basta ver que para cualquier real c, $\{X_{\tau\wedge a} \leq c\} \in \Im_n$. Por demostrar que

$$\{X_{\tau\wedge a} \leq c\} \cap \{\tau = n\} \in \Im_n \quad \forall n.$$

Si $a \leq n$

$$\{X_{\tau\wedge a} \leq c\} \cap \{\tau = n\} = \{X_a \leq c\} \cap \{\tau = n\} \in \Im_n,$$

y si $a > n$

$$\{X_{\tau\wedge a} \leq c\} \cap \{\tau = n\} = \{X_n \leq c\} \cap \{\tau = n\} \in \Im_n.$$

Entonces

$$\{X_{\tau\wedge a} \leq c\} \cap \{\tau = n\} \in \Im_n \quad \forall n.$$

□

A continuación se generalizará el concepto de Paro Opcional al de Muestreo Opcional.

Nota 3.8 *El resultado de Muestreo Opcional que se enuncia en el Teorema 3.2.3, también se conoce como el Teorema de Paro Opcional de Doob.*

El *Muestreo Opcional* transforma el proceso $\{X_n, \Im_n\}$ al proceso $\{Y_n\}$ de la manera siguiente. Sea $\tau_1, \tau_2, \ldots$ una sucesión finita o infinita de tiempos de paro, con la propiedad de que

$$1 \leq \tau_1 \leq \tau_2 \leq \cdots < \infty \qquad \text{c.s.,} \tag{3.6}$$

entonces definimos $Y_n(\omega) = X_{\tau_n(\omega)}(\omega)$, $n \geq 1$.

En particular, si consideramos $\Im_n = \sigma(X_1, \ldots, X_n)$ y $\tau(\omega)$ tiempo de paro finito que determina la transformación de Paro Opcional, y si definimos a $\tau_j(\omega)$ como

$$\tau_j(\omega) = \min[\tau(\omega), j] = \tau \wedge j,$$

entonces las τ_j's satisfacen la condición impuesta de Muestreo Opcional y el Muestreo Opcional determinado por las τ_j's es el mismo que determina τ con Paro Opcional.

Teorema 3.2.3 (Doob) *Sea $\{X_n, \Im_n\}$ es una submartingala (martingala) que es transformada al proceso $\{Y_n\}$ por Muestreo Opcional. Si se cumplen las siguientes condiciones:*

$$\begin{aligned} &(a) \quad E|Y_n| < \infty \\ &(b) \quad \liminf_k \int_{\tau_n > k} |X_n| dP = 0, \end{aligned}$$

para toda n, entonces $\{Y_n, \Im_{\tau_n}\}$ es una submartingala (martingala) y

$$EY_n \overset{\geq}{(=)} EX_1 \qquad n \geq 1.$$

Demostración

Si $A \in \Im_{\tau_n}$, entonces

$$A \cap \{\tau_{n+1} \leq k\} = \left(\cup_{i=1}^{k}[A \cap \{\tau_n = i\}]\right) \cap \{\tau_{n+1} \leq k\}.$$

Pero $A \cap \{\tau_n = i\} \in \Im_i \subset \Im_k$ para $i = 1, \ldots, k$ y $\{\tau_{n+1} \leq k\} \in \Im_k$, por lo que $A \in \Im_{\tau_{n+1}}$. Es decir, $\Im_{\tau_n} \subset \Im_{\tau_{n+1}}$.

Si $A \in \Im_{\tau_n}$, se tiene que demostrar que

$$\int_A Y_{n+1} dP \overset{\geq}{(=)} \int_A Y_n dP \qquad \forall n. \tag{3.7}$$

Como $A = \cup_j [A \cap \{\tau_n = j\}]$, es suficiente demostrar (3.7) para $D_j = A \cap \{\tau_n = j\}$, el cual pertenece a $\Im_j$.

Si $k > j$ y $\tau_n = j$ se tiene que $\tau_{n+1} \geq j$. Entonces

$$\int_{D_j} Y_{n+1} dP = \sum_{i=j}^{k} \int_{D_j \cap \{\tau_{n+1} = i\}} Y_{n+1} dP + \int_{D_j \cap \{\tau_{n+1} > k\}} Y_{n+1} dP,$$

por lo que

$$\int_{D_j} Y_{n+1} dP = \sum_{i=j}^{k} \int_{D_j \cap \{\tau_{n+1} = i\}} X_i dP + \int_{D_j \cap \{\tau_{n+1} > k\}} X_k dP$$

$$- \int_{D_j \cap \{\tau_{n+1} > k\}} (X_k - Y_{n+1}) dP. \tag{3.8}$$

Combinando el $i = k$ término en (3.8) con la $\int X_k dP$, obtenemos

$$\begin{aligned} \int_{D_j \cap \{\tau_{n+1} = k\}} X_k dP + \int_{D_j \cap \{\tau_{n+1} > k\}} X_k dP &= \int_{D_j \cap \{\tau_{n+1} \geq k\}} X_k dP \\ &\overset{\geq}{(=)} \int_{D_j \cap \{\tau_{n+1} \geq k\}} X_{k-1} dP, \end{aligned}$$

pues se cumple que $\{\tau_{n+1} \geq k\} = \{\tau_{n+1} \leq k-1\}^c \in \Im_{k-1}$ y $D_j \in \Im_j \subset \Im_{k-1}$. Por otro lado,

$$\int_{D_j \cap \{\tau_{n+1} \geq k\}} X_{k-1} dP = \int_{D_j \cap \{\tau_{n+1} > k-1\}} X_{k-1} dP,$$

el cual puede combinarse con el termino $i = k-1$ de (3.8) para obtener

$$\int_{D_j \cap \{\tau_{n+1} > k-2\}} X_{k-2} dP.$$

Si se continua de forma inductiva, se obtiene

$$\int_{D_j} Y_{n+1} dP \geq \int_{D_j \cap \{\tau_{n+1} \geq j\}} X_j dP - \int_{D_j \cap \{\tau_{n+1} > k\}} (X_k - Y_{n+1}) dP.$$

Ahora, por hipótesis tenemos que $\int_{D_j\cap\{\tau_{n+1}>k\}} X_k dP \to 0$ cuando $k \to \infty$ a través de una subsucesión apropiada, y $\int_{D_j\cap\{\tau_{n+1}>k\}} Y_{n+1} dP \to 0$ cuando $k \to \infty$, esto último como consecuencia de que $\{\tau_{n+1} > k\} \downarrow \emptyset$. Finalmente, $D_j \cap \{\tau_{n+1} \geq j\} = D_j$ porque $D_j \subset \{\tau_n = j\}$ y como $X_j = Y_n$ en D_j, entonces

$$\int_{D_j} Y_{n+1} dP \overset{\geq}{(=)} \int_{D_j} Y_n dP.$$

La segunda parte es consecuencia de que $EX_1 = EY_1$ y de que Y_n es submartingala (martingala).

□

Proposición 3.2.4 *Sean $\{X_n\}$ una submartingala (martingala), $\{\tau_n\}$ una sucesión creciente de tiempos de paro finitos c.s., entonces las condiciones (a) y (b) del Teorema anterior, se cumplen si*

$$\begin{array}{ll} (i) & \textit{cada } \tau_n \textit{ es acotado} \quad \textit{c.s.} \\ (ii) & E(\sup_n |X_n|) < \infty \end{array}$$

Nota 3.9 *La condición (i) siempre se satisface por el Paro Opcional y (ii) se cumple si $\{X_n\}$ es uniformemente acotada.*

Demostración

Se demostrará que $(i) \Rightarrow (a)$ y (b) del Teorema 3.2.3.

Supongamos que $\tau_n \leq k_n$ c.s. para $k_n \in \mathcal{N}$, entonces

$$\int_\Omega |X_{\tau_n}| dP = \sum_{j \leq k_n} \int_{\{\tau_n = i\}} |X_i| dP \leq \sum_{i \leq k_n} E|X_i| < \infty,$$

por lo que $E|X_n| < \infty$. Por otra parte, se tiene que $\{\tau_n > k\} \downarrow \emptyset$ cuando $k \to \infty$, cumpliendose (b).

Ahora se demostrará que $(ii) \Rightarrow (a)$ y (b).

Si $Z = \sup_n |X_n|$ se cumple que $\int_\Omega |X_{\tau_n}| dP \leq EZ < \infty$, probando (a). Por otro lado,

$$\int_{\{\tau > k\}} |X_k| dP \leq \int_{\{\tau_n > k\}} Z dP \to 0 \quad \text{cuando } k \to \infty$$

□

Capítulo 4

EXTENSIONES DEL LEMA DE BOREL-CANTELLI Y LA LEY FUERTE DE LOS GRANDES NÚMEROS

Como mencionamos en el Capítulo anterior, la teoría de martingalas ha permitido generalizar una gran variedad de resultados que requerían de independencia entre las variables aleatorias, y algunos de los primeros Teoremas generalizados fueron el Lema de Borel-Cantelli y la Ley Fuerte de los Grandes Números. En éste Capítulo estudiamos algunas de sus extensiones que se derivan de las propiedades de martingalas.

En la primera Sección se estudia el Lema de Borel-Cantelli de Lévy y un par de resultados que se pueden obtener directamente de éste. En la segunda Sección tratamos la Ley Fuerte de los Grandes Números de Lévy y de Chow para compararlas después con el resultado que se encuentra en la Sección tercera y el cual conjunta el Lema de B-C y la Ley Fuerte de Grandes Números de Lévy.

4.1 LEMA DE BOREL-CANTELLI DE LÉVY

En el Ejemplo 2.3 se verifica que la parte divergente del Lema de B-C puede no ser cierta si los eventos no son independientes, pero las extensiones del Capítulo 2 dan pie a pensar que basta con un *grado de independencia* adicional para que la conclusión sea cierta cuando los eventos no son independientes.

Definición 4.1 *Definimos al supremo esencial de la v.a.* $|X|$ *como*

$$sup^*|X| = \inf\{M > 0 | P(|X| > M) = 0\}.$$

El Teorema siguiente es clave para la demostración de la Forma Condicional del Lema de B-C y la Ley de los Grandes Números de Lévy.

Teorema 4.1.1 *Si $\{X_n\}$ es martingala y*

$$E\{\sup_{n\geq 0}(X_{n+1} - X_n)\} < \infty \qquad (X_0 = 0), \tag{4.1}$$

entonces $\lim_{n\to\infty} X_n$ *existe y es finito para casi toda* $\omega \in \Omega$ *en* $[\limsup_n X_n(\omega) < \infty]$.

Demostración

Supongamos que la condición (4.1) se satisface. Sea N un número positivo y sea

$$\tau_N(\omega) = \min\{j \in \mathcal{N} : X_j(\omega) > N\}$$

y $\tau_N(\omega) = \infty$ si no existe $n \in \mathcal{N}$ tal que $X_n(\omega) > N$. Definimos $X_n^{(N)}$ como

$$X_n^{(N)}(\omega) = \begin{cases} X_n(\omega), & \text{si } n < \tau_N(\omega), \\ X_{\tau_N(\omega)}(\omega), & \text{si } n \geq \tau_N(\omega), \end{cases}$$

$(X_n^{(N)}(\omega) = X_{\tau_n \wedge n}(\omega))$ y sea $m(\omega) = \sup^*_{n\geq 0}(X_{n+1}(\omega) - X_n(\omega))$. La condición $\tau_N(\omega) = k$ es una condición sobre $X_1, X_2, ..., X_k$, y $\left\{X_n^{(N)}\right\}$ es el proceso modificado obteniendo de $\{X_n\}$ a partir de τ_N . De acuerdo con el Teorema 3.2.1, $\left\{X_n^{(N)}\right\}$ es una martingala. Más aún $X_n^{(N)} \leq N + m$ c.s. para toda n, por lo que $X_n^{(N)} \in L_1 \ \ \forall n \in \mathcal{N}$. Por el Teorema 3.1.7, el límite $\lim_{n\to\infty} X_n^{(N)}$ existe y es finito c.s..

Como $X_n^{(N)}(\omega) = X_n(\omega)$ si $\sup_n X_j(\omega) \leq N$, entonces el $\lim_{n\to\infty} X_n(\omega)$ existe y es finito para casi toda ω tal que $\sup_n X_n(\omega) \leq N$, esto es, tal que $\limsup_n X_n(\omega) < \infty$.

□

Corolario 4.1.2 *Sea $Y_1, Y_2, \ldots$ una sucesión de v.a. unifomemente acotadas y sea $p_j = E[Y_j|Y_1, \ldots, Y_{j-1}]$. Entonces la serie $\sum_1^\infty Y_j < \infty$ $(= \infty)$ c.s. en el conjunto de $\omega \in \Omega$ donde $\sum_1^\infty p_j(\omega)$ converge (diverge).*

Demostración

Sea $M \in \mathcal{N}$ tal que $|Y_i| \leq M \ \forall i \in \mathcal{N}$. Si $X_n = \sum_1^n (Y_j - p_j)$, el proceso $\{X_n\}$ es una martingala. Por otro lado

$$p_i = E[Y_i|Y_1, \ldots, Y_{i-1}] \geq E[-M|Y_1, \ldots, Y_{i-1}] = -M$$

y

$$p_i = E[Y_i|Y_1, \ldots, Y_{i-1}] \leq M$$

entonces $|Y_i - p_i| \leq 2M \ \forall i \in \mathcal{N}$.

Como

$$E(\sup_{n\geq 0}|X_{n+1} - X_n|) = E(\sup_{n\geq 0}|Y_{n+1} - p_{n+1}|) < \infty,$$

se puede aplicar el Teorema anterior a los procesos $\{X_n\}$ y $\{-X_n\}$. Entonces $\lim_{n\to\infty} X_n(\omega)$ existe c.s. y es finito donde

$$\limsup_n X_n(\omega) < \infty \quad \text{o} \quad \liminf_n X_n(\omega) > -\infty.$$

Por lo tanto,

$$P\{\lim_{n\to\infty} X_n = \infty\} = P\{\lim_{n\to\infty} X_n = -\infty\} = 0.$$

Es así como se concluye que las series deben converger y diverger *juntas* con probabilidad 1. Es decir, que la medida del conjunto donde $\sum Y_i < \infty$ y $\sum p_i$ diverge o $\sum Y_i = \infty$ y $\sum p_i$ converge, es cero.

$\square$

El siguiente resultado se conoce como la forma condicional del Lema de Borel-Cantelli de Lévy y se sigue del Corolario anterior, definiendo a $Y_n(\omega) = I_{E_n}$ para toda $n \geq 1$.

Corolario 4.1.3 (Lévy) *Si $\{E_n\}$ es una sucesión de eventos, entonces $P(\limsup E_n) = 0$ $(=1)$ si*

$$\sum_{n\geq 1} P(E_n|\Im_{n-1}) < \infty (= \infty) \qquad c.s..$$

Considerando $\Im_n = \sigma(E_1, \ldots, E_n)$.

La conclusión del Corolario se sigue a partir de que $\limsup_n E_n = \{\omega : \sum I_{E_n}(\omega) = \infty\}$.

El resultado de Lévy reduce las condiciones de independencia de los eventos a la convergencia de la serie de probabilidades condicionales. Además refina la parte convergente del Lema de B-C porque $\sum P(E_n) < \infty$ implica $\sum P(E_n|\Im_{n-1}) < \infty$ c.s. (esto se deduce a partir de que $\sum \int_\Omega P(E_n|\Im_{n-1}) = \sum P(E_n)$ y $\int \sum P(E_n|\Im_{n-1}) = \sum \int P(E_n|\Im_{n-1})$ por el Teorema de Convergencia Monótona). Notemos que si los eventos son independientes, su resultado se reduce al Lema de B-C original.

Definición 4.2 *Sea $\{E_n\}$ una sucesión de eventos. Definimos $\phi_1 = 0$ y*

$$\phi_n = \sup_{F\in\Im_{n-1}} |P(E_n|F) - P(E_n)| \quad \text{para } n \geq 2,$$

donde $\Im_n = \sigma(E_1, E_2, \ldots, E_n)$.

La función ϕ_n nos dice que tan diferente es $P(E_n)$ de la probabilidad dada la información de los eventos anteriores. Cuando los eventos $\{E_n\}$ son independientes, se tiene que

$$\phi_n = \sup_{F\in\Im_{n-1}} |P(E_n|F) - P(E_n)| = 0.$$

Por otro lado, es fácil de verificar que $0 \leq \phi_n \leq 1$.

Antes de enunciar el resultado obtenido por Iosifescu y Theodorescu (Teorema 4.1.5), se demostrará el siguiente Lema.

Lema 4.1.4 *Para toda $F \in \Im_{n-1}$ tenemos que*

$$|P(E_n \cap F) - P(E_n)P(F)| \leq \phi_n P(F)$$

Demostración

$|P(E_n \cap F) - P(E_n)P(F)| = |\int_A [P(E_n|\Im_{n-1}) - P(E_n)]dP| \leq \phi_n P(A)$.

□

Teorema 4.1.5 *Supongamos que $\sum_{n=1}^{\infty} \phi_n < \infty$ y que $\sum P(E_n) = \infty$ entonces*

$$P(\limsup_n E_n) = 1.$$

El Teorema anterior es un refinamiento de la parte divergente del Lema de B-C y a diferencia del resultado de Lévy, nos pide la misma condición que el Lema de Borel-Cantelli original más otra que mide el grado de dependencia entre las variables. Si E_n se vuelve cada vez más independiente de los eventos anteriores, lo suficientemente rápido para que la serie de ϕ_n's converja, y a su vez la serie de sus probabilidades diverge, entonces se cumple el resultado de Borel-Cantelli.

Si los eventos $\{E_n\}$ son independientes, el resultado se reduce al original de B-C.

El Teorema 4.1.5 es un caso particular del Lema de B-C de Lévy, ya que si $\sum_1^{\infty} P(E_n) = \infty$ y $\sum_1^{\infty} \phi_n < \infty$, entonces se tiene que $\sum_1^{\infty} P(E_n|\Im_{n-1}) = \infty$ porque $\sum_1^{\infty} |P(E_n|\Im_{n-1}) - P(E_n)| \leq \sum_1^{\infty} \phi_n < \infty$ c.s.. Una demostración alternativa que no utiliza el Corolario 4.1.3, es la siguiente y se encuentra en [10] páginas 1 y 2.

Demostración

Supongamos que $P(\limsup_n E_n) < 1$, entonces existe $m \in \mathcal{N}$ tal que

$$P(\cup_{i=m}^{\infty} E_i) < 1.$$

Entonces $\forall n > m$ tenemos que

$$P(\cup_{i=m}^{n} E_i) = P(E_m) + P(E_m^c \cap E_{m+1}) + \cdots + P(E_m^c \cap E_{m+1}^c \cap \cdots \cap E_{n-1}^c \cap E_n).$$

Haciendo uso del Lema 4.1.4 tenemos

$$\begin{aligned} P(\cup_{i=m}^{n} E_i) &\geq P(E_m) + P(E_m^c)[P(E_{m+1} - \phi_m)] + \cdots \\ &\quad \cdots + P(E_m^c \cap \ldots \cap E_{n-1}^c)[P(E_n) - \phi_{n-1}] \\ &\geq P((\cup_{i=m}^{n} E_i)^c) \sum_{i=m}^{n} P(E_i) - \sum_{i=m}^{n-1} \phi_i, \end{aligned}$$

donde $P((\cup_{i=m}^{n} E_i)^c) > 0$. Tomando $n \to \infty$ tenemos una contradicción, por lo que $P(\limsup_n E_n) = 1$.

□

Otro resultado, el cual requiere una condición adicional para que dada la divergencia de $\sum P(E_n)$, se tenga $P(\limsup_n E_n) = 1$, es el Teorema 4.1.6; enunciado por Serfling en el año de 1975, generaliza el resultado del Teorema 4.1.5, puesto que si $\sum_{i\geq 1} \phi_i < \infty$, se tiene que $\sum_{n\geq 1} E|p_n - P(E_n)| < \infty$, debido a que $|p_n - P(E_n)| \leq \phi_n$ c.s.. Una demostración alternativa se expone en [13] páginas 728-730.

Teorema 4.1.6 *Para una sucesión de eventos $\{E_n\}$ arbitraria tal que*

$$\sum_{n=1}^{\infty} E|p_n - P(E_n)| < \infty \tag{4.2}$$

y $\sum_1^\infty P(E_n) = \infty$ entonces $P(\limsup_n E_n) = 1$, donde $p_1 = P(E_1)$ y $p_n = P(E_n|\Im_{n-1})$ para $n \geq 2$.

La prueba del Teorema 4.1.6 que se encuentra en [13] se realiza a través de algunas propiedades y resultados de la función distancia entre variables aleatorias ($d(X,Y) = \sup_F |P(X \in F) - P(Y \in F)|$) aplicada a la serie de variables aleatorias $\{X_n\}$ $(= \{I_{E_n}\})$ y la serie de variables aleatorias $\{Y_n\}$ ($Y_i \sim$ Bernoulli$(P(E_i))$ $i = 1, 2, \ldots$) independientes.

Demostración
Por el Teorema de Convergencia Monótona se tiene

$$\begin{aligned}
\infty > \sum_{n=1}^{m} E|p_n - P(E_n)| &= \sum_{n=1}^{m} \int |p_n - P(E_n)| dP \\
&= \int \sum_{n=1}^{m} |p_n - P(E_n)| dP \\
&\geq \int \left| \sum_{n=1}^{m} (p_n - P(E_n)) \right| dP.
\end{aligned}$$

Entonces $|\sum_1^m (p_n - P(E_n))| < \infty$ c.s. para toda $m > 1$. Como $\sum^\infty P(E_n) = \infty$, entonces $\sum^\infty p_n = \infty$.
Por el Lema de B-C de Lévy (Corolario 4.1.3), concluimos que $P(\limsup_n E_n) = 1$

□

4.2 LEY FUERTE DE GRANDES NÚMEROS DE LÉVY Y CHOW

Como se ha mencionado en la sección 2 del Capítulo 2, la Ley de los Grandes Números se refiere a la convergencia casi segura de la serie $\{\sum X_i\}$, pero el término utilizado de *se* *refiere*, no solo se limita a darnos condiciones suficientes sobre $\{X_i\}$ para que

$$\frac{S_n - ES_n}{n} \to 0 \quad \text{c.s.},$$

sino que podemos considerar el caso más general de dar condiciones suficientes sobre las sucesiónes de v.a.'s $\{X_i\}$ y $\{W_i\}$ para que

$$\frac{\sum_1^n X_i - W_i}{n} \to 0 \quad \text{c.s.}$$

(Teorema 4.2.1), o aún más, podemos considerar el proporcionar condiciones suficientes sobre las sucesiones de v.a.'s $\{X_i\}$, $\{W_i\}$ y $\{Y_i\}$ para que se cumpla que

$$\frac{\sum_1^n X_i - W_i}{Y_n} \to 0 \quad \text{c.s.}$$

(Teorema 4.2.7).

Teorema 4.2.1 (Ley Fuerte de los Grandes Números de Lévy) *Sea $\{X_n\}$ una sucesión de v.a.'s uniformemente acotadas. Definamos $p_1 = P(X_1 = 1)$ y*

$$p_i = E[X_i | X_1, \ldots, X_{i-1}]$$

para $i \geq 2$. Entonces

$$\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} (X_i - p_i) = 0$$

c.s. donde $\sum p_i$ converge.

Demostración

Si $\sum_1^\infty p_i < \infty$, por el Corolario 4.1.2, se tiene que $\sum_1^\infty X_i < \infty$. Entonces

$$\lim_{n\to\infty} n^{-1} \sum (X_i - p_i) = 0 \quad \text{c.s donde } \sum p_i \text{ converge.}$$

□

Algunos resultados importantes se enunciarán antes de formular la Ley Fuerte de Grandes Números de Chow.

Teorema 4.2.2 *Sean $\{X_n, \Im_n\}$ una submartingala tal que $E(X_1^-) < \infty$; $\{Z_n, \Im_{n-1}, n \geq 2\}$ y $\{Y_n, \Im_n, n \geq 2\}$ dos sucesiones estocásticas tal que Y_n toma valores finitos en los reales para cada n.*

Sea $Z_1 = Y_1 = 0$ y $X_n \leq Z_n + Y_n$ para $n \geq 2$.

Para $b > 0$, sea $t(\omega)$ el primer índice para el cual $Y_n(\omega) > b$ y

$$E[Y_t I_{[t<\infty]}] < \infty \tag{4.3}$$

Entonces X_n converge c.s. donde $\sup_n Z_n < \infty$ y $\sup_n Y_n < b$

Demostración

Para cualquier $a > 0$ fijo, sea $t'(\omega)$ el primer índice para el cual $Z_n(\omega) > a$ y sea

$$\tau_n(\omega) = \min(t'(\omega) - 1, t(\omega), n).$$

Entonces $\tau_n(\omega)$ es una sucesión acotada de tiempos de paro tal que $\tau_n \leq \tau_{n+1}$. Por la Proposición 3.2.4, tenemos que X_{τ_n} es una submartingala. Como $E(X_1^-) < \infty$, se tienen $E(X_n^-) < \infty \; \forall n$ (Nota 3.4) y

$$\begin{aligned} E(X_{\tau_n}^+) &\leq E(Z_{\tau_n}^+) + \sum_{j=1}^{n} E(Y_j^+ I_{[\tau_n = j]}) \\ &\leq a + \sum_{j=1}^{n} E(Y_j^+ (I_{[t'-1 \geq j = t]} + I_{[t'-1 = j < t]})) + \\ &\quad + E(Y_n^+ I_{[t'-1 > n < t]}) \\ &\leq a + b + \sum_{j=1}^{n} E(Y_j^+ I_{[t'-1 \geq j = t]}) \leq a + b + E(Y_t^+ I_{[t < \infty]}). \end{aligned}$$

Por estos dos resultados, se tiene $E(|X_{\tau_n}|) < \infty$ y por el Teorema 3.1.7, X_{τ_n} converge c.s.. Entonces X_n converge c.s. donde $t = t' = \infty$, es decir, c.s. donde $\sup_n Z_n < a$ y $\sup_n Y_n < b$. Como a es arbitrario, se tiene el fin de la prueba.

□

Corolario 4.2.3 *La condición (4.3) del Teorema anterior puede ser reemplazada por*

$$E(\sup_{n \geq 1}(Y_{n+1} - Y_n^+)) < \infty \tag{4.4}$$

Demostración

Sea t como antes, entonces $t \geq 2$ y (4.4) implica que

$$\begin{aligned} E[Y_t I_{[t < \infty]}] &\leq E[\sup(Y_{n+1} - Y_n^+)] + E[Y_{t-1}^+ I_{[t < \infty]}] \\ &< \infty. \end{aligned}$$

□

Corolario 4.2.4 *Sea $\{X_n, \Im_n\}$ una submartingala y sea $\{Y_n, \Im_n\}$ una sucesión estocástica tal que $E|Y_n| < \infty$, $E[\sup(Y_{n+1}^- - Y_n^-)] < \infty$ y $X_n \leq Y_n$, entonces X_n converge c.s. donde*

$$\sup \sum_{n=1}^{m} (E[Y_{n+1} | \Im_n] - Y_n) < \infty \tag{4.5}$$

$$\sup Y_n^- < \infty \tag{4.6}$$

Demostración

Sea $Z_m = \sum_2^m (E[Y_n|\Im_{n-1}] - Y_n)$. Entonces $\{Z_n, \Im_n, n \geq 2\}$ es una martingala y

$$\begin{aligned} Z_m &= -Y_m + \sum_{n=3}^{m}(E[Y_n|\Im_{n-1}] - Y_{n-1}) + E[Y_2|\Im_1] \\ &\leq Y_m^- + \sum_{n=3}^{m}(E[Y_n|\Im_{n-1}] - Y_{n-1}) + E[Y_2|\Im_1] \\ &= Y_m^- + U_m. \end{aligned}$$

Entonces U_m es $\Im_{m-1}$-medible, $m \geq 2$. Por el Corolario 4.2.3 se tiene que Z_m converge c.s. donde $\sup_{n\geq 2} Y_n^- < \infty$ y $\sup_{n\geq 2} U_n < \infty$. Como $Z_n + X_n$ es una submartingala y $Z_n + X_n \leq Z_n + Y_n \leq U_n$, $Z_n + X_n$ converge c.s. donde $\sup_n U_n < \infty$. Entonces X_n converge c.s. donde (4.5) y (4.6) se cumplen.

□

Corolario 4.2.5 *Sea $\{X_n, \Im_n\}$ una submartingala y $p \geq 1$.*

(a) Si $E(X_n^+)^p < \infty$, entonces X_n converge c.s. donde

$$\sum_{n=2}^{\infty}\left(E[(X_n^+)^p|\Im_{n-1}] - (X_{n-1}^+)^p\right) < \infty$$

(b) Si $E|X_n|^p < \infty$, entonces X_n convege c.s. donde

$$\sum_{n=2}^{\infty}(E[|X_n|^p|\Im_{n-1}] - |X_{n-1}|^p) < \infty$$

Demostración

Como $X_n \leq X_n^+ \leq (X_n^+)^p + 1$ y $X_n \leq |X_n| \leq |X_n|^p + 1$, la demostración se sigue inmediatamente del Corolario anterior.

□

Teorema 4.2.6 *Sea $\{X_n\}$ una sucesión de v.a.'s adaptadas a la filtración $\{\Im_n\}$, tal que $p_n = E[X_n|\Im_{n-1}] < \infty$ c.s. para toda n. Además, sea A el conjunto donde*

$$\sum_{n=2}^{\infty} E\left[|X_n - p_n|^2 I_{[|X_n - p_n| \leq a_n]} + |X_n - p_n| I_{[|X_n - p_n| > a_n]} |\Im_{n-1}\right] < \infty,$$

para algunas constantes $a_n \geq c > 0$. Entonces $S_n^* = \sum_1^n (X_i - p_i)$ converge c.s. en A.

Demostración

Sin pérdida de generalidad, supongamos que $p_n = 0$ para toda n. Sea $X_n' = X_n I_{[|X_n| \leq a_n]}$. Por la propiedad de martingala, se tiene

$$|E[X_n'|\Im_{n-1}]| = \left|E[X_n I_{[|X_n|>a_n]}|\Im_{n-1}]\right| \leq E[|X_n| I_{[|X_n|>a_n]}|\Im_{n-1}],$$

c.s.. Entonces

$$\sum_{n=2}^{\infty} |E[X_n'|\Im_{n-1}]| < \infty \quad \text{c.s. en A.} \tag{4.7}$$

Como

$$\begin{aligned} \sum_{n=2}^{\infty} P(X_n' \neq X_n|\Im_{n-1}) &= \sum_{n=2}^{\infty} P(|X_n| > a_n|\Im_{n-1}) \\ &\leq \frac{1}{c}\sum_{n=2}^{\infty} E[|X_n| I_{[|X_n|>a_n]}|\Im_{n-1}] \\ &< \infty \qquad \text{c.s. en A.} \end{aligned}$$

Por el Corolario 4.2.5, se tiene $\sum_2^\infty I_{[X_n \neq X_n']} < \infty$ c.s. en A y

$$P(A \cap (\limsup[X_n \neq X_n'])) = 0. \tag{4.8}$$

Si definimos a $E[X_1'|\Im_0] = 0$ y a $Y_n = \sum_{i=1}^n (X_i' - E[X_i'|\Im_{i-1}])$, entonces $\{Y_n, \Im_n\}$ es una martingala y

$$\begin{aligned} E[Y_n^2|\Im_{n-1}] - Y_{n-1}^2 &= E[(X_n')^2|\Im_{n-1}] - E^2[X_n|\Im_{n-1}] \\ &\leq E[(X_n')^2|\Im_{n-1}] \leq E[X_n^2 I_{[|X_n|\leq a_n]}|\Im_{n-1}]. \end{aligned}$$

Se tiene así que

$$\sum_{n=2}^{\infty} (E[Y_n^2|\Im_{n-1}] - Y_{n-1}^2) < \infty \qquad \text{c.s. en A.}$$

Por *(b)* del Corolario 4.2.5, obtenemos que Y_n converge c.s. en A y por (4.7) y (4.8) se concluye que S_n^* converge c.s. en A.

□

Teorema 4.2.7 (Ley Fuerte de los Grandes Números de Chow) *Sea* $\{X_n, \Im_n\}$ *una sucesión de v.a.'s tal que* $p_n = E[X_n|\Im_{n-1}] < \infty$. *Si* $\{Y_n\}$ *es una sucesión de v.a.'s positivas adaptada a la filtración* $\{\Im_{n-1}\}$ *que cumple con:*

$$E[(X_n - p_n)Y_n^{-1}] < \infty \qquad \textit{para toda } n \geq 2.$$

Si además, $q \in [1, 2]$, *entonces* $\lim_{n\to\infty} S_n^*/Y_n = \sum_1^n (X_i - p_i)/Y_n = 0$ *c.s. donde*

$$\sum_{n=2}^{\infty} E[|X_n - p_n|^q|\Im_{n-1}]Y_n^{-q} < \infty \qquad Y_n \to \infty. \tag{4.9}$$

Demostración

Sea $V_n = X_n - p_n/Y_n$ y $U_n = V_2 + \cdots + V_n$. Entonces $\{U_n, \Im_n, n \geq 2\}$ es una martingala. Del Teorema 4.2.6 se obtiene que U_n converge c.s. donde (4.9) se cumple, si $q \in [1, 2]$.

□

4.3 REFINAMIENTO DEL LEMA DE B-C Y LA LFGN

En esta sección consideraremos a $\{X_i\}$ una sucesión de v.a. que toman valores 0 y 1, adaptada a una filtración $\{\Im_n\}$ y $p_1 = P[X_1 = 1]$ y $p_n = P[X_n = 1|\Im_{n-1}]$, para $n \geq 2$.

Teorema 4.3.1

$$R_n = \frac{X_1 + \cdots + X_n}{p_1 + \cdots + p_n} \overset{c.s.}{\to}_{n\to\infty} L < \infty$$

y $L = 1$ *c.s. donde* $\sum_{i=1}^{\infty} p_i$ *es infinita*

Este Teorema unifica la forma condicional del Lema de B-C y la Ley Fuerte de los Grandes Números de Lévy, ya que $\sum X_i = \infty$ c.s. donde $\sum p_i = \infty$ y $\sum X_i < \infty$ c.s. donde $\sum p_i < \infty$, y por otra parte,

$$\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} (X_i - p_i) = 0$$

c.s. donde $\sum p_i < \infty$. Además, también se encuentra relacionado con el Teorema 4.2.7, en el sentido de que si solo consideramos el conjunto donde $\sum_1^\infty p_i$ diverge, entonces este Teorema es un caso particular del de Chow. Para verificar esto último, solo basta definir a

$$Y_n = \sum_1^n p_i \qquad \forall n \geq 1 \quad \text{y} \qquad q = 1,$$

entonces

$$E\left(\frac{X_n - p_n}{\sum_1^n p_i}\right) \leq E\left(\frac{X_n}{\sum_1^n p_i}\right) < \infty.$$

Por el Teorema 4.2.7, se tiene que

$$\frac{\sum_1^n (X_i - p_i)}{\sum_1^n p_i} \to 0$$

c.s. donde

$$\sum_1^\infty E[|X_n - p_n||\Im_{n-1}](\sum_1^n p_i)^{-1} < \infty \quad \text{y} \qquad \sum_1^\infty p_i \uparrow \infty.$$

Pero

$$\begin{aligned}\sum_1^\infty E[|X_n - p_n||\Im_{n-1}](\sum_1^n p_i)^{-1} &= \sum_1^\infty E\left[\frac{|X_n - p_n|}{\sum_1^n p_i}|\Im_{n-1}\right] \\ &= \sum_1^\infty E\left[\left|\frac{X_n - p_n}{\sum_1^n p_i}\right||\Im_{n-1}\right] \\ &< \infty.\end{aligned}$$

Para realizar la demostración del Teorema 4.3.1, requerimos de hacer algunas definiciones y obtener algunos resultados.

Consideraremos a $\{Y_n', \Im_n\}$ una sucesión estocástica tal que

$$p_n' = E[Y_n'|\Im_{n-1}] < \infty \quad \text{c.s.}$$

y

$$V_n = E[(Y_n' - p_n')^2|\Im_{n-1}] < \infty \quad \text{c.s.}$$

Definimos $Y_n = Y_n' - p_n'$ para $n \geq 1$.

Teorema 4.3.2 *Si a y b son dos números positivos, entonces la probabilidad de que para alguna n, $Y_1 + \cdots + Y_n \geq a(V_1 + \cdots + V_n) + b$, es menor que $1/1 + ab$.*

Corolario 4.3.3 *Si a y b son dos números positivos entonces la probabilidad de que para alguna n, $|Y_1 + \cdots + Y_n| \geq a(V_1 + \cdots + V_n) + b$, es menor que $2/1 + ab$.*

La demostración del Corolario es inmediata del Lema anterior.

Lema 4.3.4 *Si $D = \{\omega \in \Omega : \sum_i^\infty V_i = \infty\}$ entonces*

$$\lim \frac{Y_1 + \cdots + Y_n}{V_1 + \cdots + V_n} = 0 \quad \textit{c.s. en } D$$

Demostración

Sea

$$D(a,b) = \{\omega \in \Omega : |Y_1 + \cdots + Y_n| < a(V_1 + \cdots + V_n) + b \;\; \forall n \in \mathcal{N}\}$$

Notemos que

$$\begin{aligned}&\{|Y_1 + \cdots + Y_n| < a(V_1 + \cdots + V_n) + b \;\; \forall n \in \mathcal{N}\} = \\ &= \left\{\frac{|Y_1 + \cdots + Y_n|}{V_1 + \cdots + V_n} < a + \frac{b}{V_1 + \cdots + V_n} \;\; \forall n \in \mathcal{N}\right\} \\ &= \left\{\frac{|Y_1 + \cdots + Y_n|}{|V_1 + \cdots + V_n|} < a + \frac{b}{V_1 + \cdots + V_n} \;\; \forall n \in \mathcal{N}\right\} \qquad (4.10)\end{aligned}$$

de donde se concluye que

$$\limsup_n \left|\frac{Y_1 + \cdots + Y_n}{V_1 + \cdots + V_n}\right| \leq a \tag{4.11}$$

en $D \cap D(a,b)$

Si b es creciente a ∞ (tomando valores enteros), aplicando el Corolario 4.3.3 se prueba que (4.11) se cumple c.s. en D.

Finalmente si a es decreciente a 0 (tomando recíprocos de enteros positivos), se concluye que (4.11) es cierto para $a = 0$ c.s. en D. Entonces

$$\liminf_n \left|\frac{Y_1 + \cdots + Y_n}{V_1 + \cdots + V_n}\right| = \limsup_n \left|\frac{Y_1 + \cdots + Y_n}{V_1 + \cdots + V_n}\right| = 0$$

por lo tanto

$$\lim \frac{Y_1 + \cdots + Y_n}{V_1 + \cdots + V_n} = 0 \quad \text{c.s. en } D$$

□

Lema 4.3.5 *El límite de* $(Y_1 + \cdots + Y_n)$ *cuando* $n \to \infty$*, existe y es finito c.s. en* $G = \{\omega \in \Omega : \sum_1^\infty V_i < \infty\}$.

Demostración

Sea $D_k(a,b) = [|Y_{k+1} + \cdots + Y_{k+n}| < a(V_{k+1} + \cdots + V_{k+n}) + b \ \ \forall n]$. El Corolario 4.3.3 implica que la probabilidad de $D_k(a,b) \geq 1 - 2/1 + ab$.

Sea $G_k(c)$ el conjunto donde $V_{k+1} + \cdots + V_{k+n} \leq c \ \forall n$. Para cada número positivo c se tiene que

$$\lim_{k\to\infty} P[G - G_k(c)] = 0 \qquad \text{en } G_k(c) \cap D_k(a,b),$$

$$|Y_{k+1} + \cdots + Y_{k+n}| < ac + b \qquad \forall n$$

Para completar la prueba se demostrará que $\forall \epsilon > 0$ existe un conjunto medible $C(\epsilon)$ de G con las siguientes propiedades:

(i) Y_1, $Y_1 + Y - 2$, $Y_1 + Y_2 + Y_3$,... es una sucesión de Cauchy en $C(\epsilon)$.

(ii) $P[G - C(\epsilon)] < \epsilon$

Sean $0 < b_i \to 0$ y $a_i > 0$ tal que $\sum_{i=1}^\infty 1/1 + a_i b_i < \epsilon/4$.

Escogemos $c_i > 0$ tal que $a_i c_i \to 0$, k_i lo suficientemente grande para que

$$\sum_{i-1}^\infty P[G - G_{k_i}(c_i)] < \frac{\epsilon}{2}$$

entonces

$$C(\epsilon) = \cap_{i=1}^\infty [G_{k_i}(c_i) \cap D_{k_i}(a_i, b_i)]$$

□

Demostración del Teorema 4.3.1

Por el Lema 4.3.4 tenemos que

$$\lim_{n\to\infty} \frac{(X_1 - p_1) + \cdots + (X_n - p_n)}{p_1(1-p_1) + \cdots + p_n(1-p_n)} \stackrel{c.s.}{=} 0 \tag{4.12}$$

y

$$\lim_{n\to\infty} \frac{(X_1 - p_1) + \cdots + (X_n - p_n)}{p_1 + \cdots + p_n} \stackrel{c.s.}{=} 0 \tag{4.13}$$

donde $\sum_{j=1}^{\infty} p_j(1-p_j) = \infty$. Más aún

$$\lim_{n\to\infty} \sum_{i=1}^{n} (X_i - p_i) \tag{4.14}$$

existe y es finito casi seguramente donde

$$\sum_{j=1}^{\infty} p_j(1-p_j) < \infty,$$

por el Lema 4.3.5.

Entonces (4.13) se cumple en $[\sum_1^\infty p_j = \infty]$. Finalmente, $\sum_1^\infty p_j < \infty$ implica que $\sum_1^\infty p_j(1-p_j) < \infty$; de (4.14) $\sum_1^\infty X_i < \infty$ casi seguramente en $\sum_1^\infty p_j < \infty$

$\square$

BIBLIOGRAFÍA

[**1**] Ash, Robert, B. (1972) *Real Analysis and Probability.* Academic Press, Inc.

[**2**] Billingsley, P. (1986) *Probability and Measure.* John Wiley & Sons.

[**3**] Chow, Y. S. (1965) *Local Convegence of Martingales and Law or Large Number.* Annals of Mathematical Statistics **36** 552-558

[**4**] Chung, Kai Lai (1974) *A Course in Probability Theory Real Analysis and Probability.* Academic Press, New York.

[**5**] Doob, J. L. (1953) *Stochastic Processes.* John Wiley & Sons.

[**6**] Dubins, L. and Freedman, D. A. (1965) *A Sharper Form of the Borel-Cantelli lemma and the Strong Law.* Annals of Mathematical Statistics **36** 800-807.

[**7**] Dudley R. M. (1989) *Real Analysis and Probability.* Wadsworth & Brooks/Cote.

[**8**] Etemadi, N. *An elementary Proof of the Strong Law or Large Numbers.* Z. Wahrscheinlichkitstheorie Vervv. Bebiete **55** 119-122.

[**9**] Feller, W. (1968) *Introducción a la teoría de Probabilidades y sus Aplicaciones.* Limusa.

[**10**] Iosifescu, M. and Theodorescu, R. (1969) *Random Processes and Learning.* Springer-Verlag, New York.

[**11**] Laha, R. G. and Rohatgi, V. K. (1979) *Probability Theory.* John Wiley & Sons, Inc.

[**12**] Lévy, P. (1937) *Theorie de l ´addition des variables aléatoires.* Gauthier-Villars, Paris.

[**13**] Loéve, M. (1955) *Probability Theory.* Vol I. Springer-Verlag.

[**14**] Serfling R. J. (1975) *A General Poisson Approximation Theorem.* Annals of Probability **3** 726-731. Real Analisys and Probability