Capítulo 2

Leyes de Grandes Números

2.1. Introducción

Sea $(X_n, n \geq 1)$ una sucesión de variables aleatorias, definimos $S_n = X_1 + X_2 + \cdots + X_n$. Decimos que esta sucesión satisface una ley de grandes números si existen sucesiones de números reales $(a_n, n \geq 1)$ y $(b_n, n \geq 1)$ tales que

$$\frac{S_n - a_n}{b_n} \to 0 \quad \text{cuando } n \to \infty$$

donde los modos de convergencia de interés son convergencia en probabilidad, que corresponde a una ley débil y convergencia c.p. 1, que corresponde a una ley fuerte.

En este capítulo estudiaremos este tipo de leyes para sucesiones de variables aleatorias independientes y en general el problema de convergencia de series de variables aleatorias. Comenzamos por las leyes débiles.

2.2. Leyes Débiles de Grandes Números

Consideremos una sucesión de variables aleatorias no correlacionadas $(X_n, n \geq 1)$ con $\mathrm{E}(X_i) = \mu_i$, $\mathrm{Var}(X_i) = \sigma_i^2$. Sea $\bar{X}_n = S_n/n$, entonces

$$\mathrm{E}(\bar{X}_n) = \frac{1}{n}\sum_{i=1}^{n} \mu_i = \bar{\mu}_n, \qquad \mathrm{Var}(\bar{X}_n) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2.$$

El próximo resultado da condiciones bajo las cuales esta sucesión obedece una ley débil con $a_n = \sum_{i=1}^{n} \mu_i$ y $b_n = n$.

Teorema 2.1 (LDGN) *Si se satisface que*

$$\frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2 \to 0 \quad (n \to 0) \tag{2.1}$$

entonces

$$\bar{X}_n - \bar{\mu}_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_i) \xrightarrow{P} 0. \tag{2.2}$$

Demostración. Sea $\varepsilon > 0$, por la desigualdad de Chebychef tenemos

$$P(|\bar{X}_n - \bar{\mu}_n| > \varepsilon) \leq \frac{1}{\varepsilon^2}\mathrm{Var}(\bar{X}_n - \bar{\mu}_n) = \frac{1}{\varepsilon^2 n^2}\sum_{i=1}^{n} \sigma_i^2$$

y por (2.1), como ε está fijo, esta expresión tiende a 0 cuando $n \to \infty$. ■

Corolario 2.1 (LDGN para variables iid) *Sea* $(X_n, n \geq 1)$ *una sucesión de v.a.i.i.d. con* $\mathrm{E}(X_i^2) < \infty$. *Entonces*

$$\bar{X}_n \xrightarrow{P} \mathrm{E}(X_1).$$

Corolario 2.2 *Sea* $(X_n, n \geq 1)$ *una sucesión de v.a.i. con distribución de Bernoulli de parámetro* p *(Ber*(p)*), entonces*

$$\hat{p}_n \equiv \frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{P} p.$$

2.2.1. El Teorema de Aproximación de Weierstrass

El teorema de aproximación de Weierstrass dice que cualquier función continua f definida en un intervalo cerrado $[a, b]$ puede ser aproximada de manera uniforme por polinomios: Dado $\varepsilon > 0$ existe un polinomio g tal que

$$\sup_{x \in [a,b]} |f(x) - g(x)| < \varepsilon.$$

Serge Bernstein dió una demostración de este resultado usando la LDGN, que presentamos a continuación para funciones definidas en el intervalo $[0, 1]$. La extensión al caso general es sencilla.

Proposición 2.1 (Polinomios de Bernstein) *Sea* $f : [0, 1] \to \mathbb{R}$ *una función continua, entonces*

$$\sup_{x \in [0,1]} \Big| f(x) - \sum_{k=0}^{n} \binom{n}{k} f\big(\frac{k}{n}\big) x^k (1-x)^{n-k} \Big| \to 0$$

cuando $n \to \infty$.

Demostración. Sea $\varepsilon > 0$. Como f es uniformemente continua, existe $\delta > 0$ tal que si $0 \leq x, y \leq 1$ satisfacen $|x - y| < \delta$ entonces se tiene $|f(x) - f(y)| < \varepsilon$.

Sea $S_n, n \geq 1$ una sucesión de variables con distribución binomial de parámetros n y x y consideramos la variable aleatorias $f(S_n/n)$ cuyo valor esperado es

$$\begin{aligned} \mathrm{E}\Big[f\Big(\frac{S_n}{n} \Big) \Big] &= \sum_{k=0}^{n} f\big(\frac{k}{n}\big) P(S_n = k) \\ &= \sum_{k=0}^{n} \binom{n}{k} f\big(\frac{k}{n}\big) x^k (1-x)^{n-k}. \end{aligned}$$

Recordemos que $S_n = \sum_1^n \xi_i$ donde las ξ_i son independientes de Bernoulli con parámetro x, y por lo tanto $Var(\xi_i) = x(1-x) \leq 1/4$ para $x \in [0, 1]$. Usando la desigualdad de Chebychef tenemos

$$P\Big(\Big| \frac{S_n}{n} - x \Big| > \delta \Big) \leq \frac{1}{\delta^2 n^2} \operatorname{Var}(S_n) = \frac{1}{\delta^2 n} \operatorname{Var}(\xi_1) \leq \frac{1}{4\delta^2 n}$$

Por lo tanto, existe un entero n_0 que no depende de x, tal que, si $n \geq n_0$

$$P\Big(\Big| \frac{S_n}{n} - x \Big| > \delta \Big) < \varepsilon.$$

Usando la desigualdad triangular obtenemos

$$\begin{aligned}
\Big|\operatorname{E}\Big[f\Big(\frac{S_n}{n}\Big)\Big]-f(x)\Big| &= \Big|\sum_{k=0}^{n}\Big(f\Big(\frac{k}{n}\Big)-f(x)\Big)P(S_n=k)\Big| \\
&\le \sum_{k=0}^{n}\Big|f\Big(\frac{k}{n}\Big)-f(x)\Big|P(S_n=k) \\
&\le \sum_{|\frac{k}{n}-x|\le\delta}\Big|f\Big(\frac{k}{n}\Big)-f(x)\Big|P(S_n=k)+\sum_{|\frac{k}{n}-x|>\delta}\Big(\Big|f\Big(\frac{k}{n}\Big)\Big|+\big|f(x)\big|\Big)P(S_n=k) \\
&\le \sum_{|\frac{k}{n}-x|\le\delta}\varepsilon P(S_n=k)+\sum_{|\frac{k}{n}-x|>\delta}2\sup_{0\le x\le 1}|f(x)|P(S_n=k) \\
&= \varepsilon P\Big(\Big|\frac{S_n}{n}-x\Big|\le\delta\Big)+2\sup_{0\le x\le 1}|f(x)|P\Big(\Big|\frac{S_n}{n}-x\Big|>\delta\Big).
\end{aligned}$$

En consecuencia, para todo $n \geq n_0$,

$$\Big|\operatorname{E}\Big[f\Big(\frac{S_n}{n}\Big)\Big]-f(x)\Big| \le \varepsilon+2\varepsilon\sup_{0\le x\le 1}|f(x)|$$

y esto demuestra la proposición. ■

2.3. Leyes Fuertes de Grandes Números

Demostramos a continuación la LFGN de Etemadi, para sucesiones de v.a. que son independientes a pares y tienen igual distribución.

Teorema 2.2 (Etemadi) *Sea $(X_n, n \geq 1)$ una sucesión de v.a. independientes a pares con igual distribución y supongamos que $\operatorname{E}|X_1| < \infty$. Entonces*

$$\bar{X}_n \to \operatorname{E}(X_1) \quad c.p.1. \tag{2.3}$$

Demostración. Recordemos que $X_n = X_n^+ - X_n^-$ y como suponemos que $\operatorname{E}|X_1| < \infty$, (X_n^+) es una sucesión de v.a. independientes a pares con igual distribución que satisface $\operatorname{E}|X_1^+| < \infty$ y otro tanto es cierto para la sucesión (X_n^-). Teniendo en cuenta que

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n}X_i = \frac{1}{n}\sum_{i=1}^{n}X_i^+ - \frac{1}{n}\sum_{i=1}^{n}X_i^-$$

basta demostrar el teorema suponiendo que las variables son no-negativas.

Ahora definimos las variables truncadas

$$Y_n = X_n\mathbf{1}_{\{X_n\le n\}}, \quad n \geq 1.$$

Entonces

$$\begin{aligned}
\sum_{n=1}^{\infty}P(X_n\ne Y_n) = \sum_{n=1}^{\infty}P(X_n>n) &= \sum_{n=1}^{\infty}P(X_1>n) \\
&\le \sum_{n=1}^{\infty}\int_{n-1}^{n}P(X_1>t)\,dt \\
&= \int_0^{\infty}P(X_1>t)\,dt \\
&= \operatorname{E}(X_1) < \infty.
\end{aligned}$$

Por el lema de Borel-Cantelli tenemos que $P(X_n \neq Y_n \ i.v.) = 0$, lo que quiere decir que, con probabilidad 1, $X_n = Y_n$ para todos los índices n, excepto por una cantidad finita de ellos. En consecuencia basta demostrar que

$$\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i \to \mathrm{E}(X_1) \quad c.p.1 \tag{2.4}$$

ya que

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) \to 0 \quad c.p.1.$$

Observamos ahora que

$$\mathrm{E}(Y_n) = \mathrm{E}(X_n \mathbf{1}_{\{X_n \leq n\}}) = \mathrm{E}(X_1 \mathbf{1}_{\{X_1 \leq n\}}) \to \mathrm{E}(X_1)$$

por el TCM y en consecuencia

$$\mathrm{E}(\bar{Y}_n) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(Y_i) \to \mathrm{E}(X_1) \quad \text{cuando } n \to \infty. \tag{2.5}$$

Sea ahora $\alpha > 1$ y consideremos la sucesión de enteros $u_k = [\alpha^k]$. Supongamos que podemos demostrar que

$$\bar{Y}_{u_k} \xrightarrow{c.s.} \mathrm{E}(X_1) \quad \text{cuando } k \to \infty. \tag{2.6}$$

Para cualquier n existe k tal que $u_k \leq n < u_{k+1}$. Como las variables Y_n son no-negativas, tenemos

$$\frac{1}{n}\sum_{i=1}^{u_k} Y_i \leq \bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i \leq \frac{1}{n}\sum_{i=1}^{u_{k+1}} Y_i$$

y en consecuencia

$$\frac{u_k}{n}\bar{Y}_{u_k} \leq \bar{Y}_n \leq \frac{u_{k+1}}{n}\bar{Y}_{u_{k+1}}$$

de donde obtenemos que

$$\frac{1}{\alpha}\bar{Y}_{u_k} \leq \bar{Y}_n \leq \alpha \bar{Y}_{u_{k+1}}.$$

Haciendo $n \to \infty$ y usando (2.6) obtenemos, con probabilidad 1,

$$\frac{1}{\alpha}\,\mathrm{E}(X_1) \leq \liminf_{n\to\infty} \bar{Y}_n \leq \limsup_{n\to\infty} \bar{Y}_n \leq \alpha\,\mathrm{E}(X_1).$$

Como esta relación es cierta para todo $\alpha > 1$ obtenemos que (2.4) vale.

Ahora tenemos que demostrar que (2.6) es cierta y para esto basta ver que

$$\bar{Y}_{u_k} - \mathrm{E}(\bar{Y}_{u_k}) \to 0 \ c.p.1 \quad \text{cuando } k \to \infty. \tag{2.7}$$

Usamos la desigualdad de Chebychef y la independencia a pares de las variables $(Y_n, n \geq 1)$. Para cualquier $\varepsilon > 0$

$$\begin{aligned} P(|\bar{Y}_{u_k} - \mathrm{E}(\bar{Y}_{u_k})| > \varepsilon) &\leq \frac{1}{\varepsilon^2}\,\mathrm{Var}(\bar{Y}_{u_k}) \\ &= \frac{1}{\varepsilon^2 u_k^2}\sum_{i=1}^{u_k} \mathrm{Var}(Y_i) \\ &\leq \frac{1}{\varepsilon^2 u_k^2}\sum_{i=1}^{u_k} \mathrm{E}(Y_i^2). \end{aligned}$$

En consecuencia

$$\sum_{k=1}^{\infty} P(\bar{Y}_{u_k} - \mathrm{E}(\bar{Y}_{u_k})| > \varepsilon) \leq \frac{1}{\varepsilon^2} \sum_{k=1}^{\infty} \frac{1}{u_k^2} \sum_{i=1}^{u_k} \mathrm{E}(Y_i^2)$$
$$= \frac{1}{\varepsilon^2} \sum_{i=1}^{\infty} \mathrm{E}(Y_i^2) \sum_{k:u_k \geq i} \frac{1}{u_k^2}. \tag{2.8}$$

Ahora, como $u_k = [\alpha^k] > \alpha^{k-1}$ para $1 < \alpha < \infty$,

$$\sum_{k:u_k \geq i} \frac{1}{u_k^2} \leq \sum_{k:\alpha^{k-1} \geq i} \frac{1}{\alpha^{(k-1)2}} \leq \frac{C_1}{i^2} \tag{2.9}$$

para alguna constante C_1, $0 < C_1 < \infty$. Usamos este resultado en (2.8) y recordamos que las variables X_i tienen igual distribución,

$$\sum_{k=1}^{\infty} P(\bar{Y}_{u_k} - \mathrm{E}(\bar{Y}_{u_k})| > \varepsilon) \leq \frac{C_1}{\varepsilon^2} \sum_{i=1}^{\infty} \frac{1}{i^2} \mathrm{E}(Y_i^2)$$
$$= \frac{C_1}{\varepsilon^2} \sum_{i=1}^{\infty} \frac{1}{i^2} \mathrm{E}(X_1^2 \mathbf{1}_{\{0 \leq X_1 \leq i\}})$$
$$= \frac{C_1}{\varepsilon^2} \sum_{i=1}^{\infty} \sum_{j=1}^{i} \frac{1}{i^2} \mathrm{E}(X_1^2 \mathbf{1}_{\{j-1 \leq X_1 \leq j\}})$$
$$= \frac{C_1}{\varepsilon^2} \sum_{j=1}^{\infty} \mathrm{E}(X_1^2 \mathbf{1}_{\{j-1 \leq X_1 \leq j\}}) \sum_{i=j}^{\infty} \frac{1}{i^2}$$
$$\leq \frac{C_1}{\varepsilon^2} \sum_{j=1}^{\infty} \mathrm{E}(jX_1 \mathbf{1}_{\{j-1 \leq X_1 \leq j\}}) C_2 j^{-1}$$
$$= \frac{C_1 C_2}{\varepsilon^2} \mathrm{E}(X_1) < \infty. \tag{2.10}$$

Como esta serie es convergente para cualquier $\varepsilon > 0$ obtenemos, por el lema de Borel-Cantelli, que (2.7) es cierto. ■

Antes de demostrar la LFGN de Kolmogorov, presentamos el siguiente lema.

Lema 2.1 *Sea $(X_n, n \geq 1)$ una sucesión de v.a.i.i.d. Las siguientes proposiciones son equivalentes:*

(a) $\mathrm{E}\,|X_1| < \infty$.

(b) Para todo $\varepsilon > 0$, $\sum_{n \geq 1} P(|X_1| \geq \varepsilon n) < \infty$.

(c) $\lim_{n\to\infty} \left|\frac{X_n}{n}\right| = 0$ *c.p. 1.*

Para su demostración recordamos el siguiente resultado: Sea X una v.a. no-negativa, entonces

$$\mathrm{E}(X) = \int_0^{\infty} P(X > x)\, dx$$

donde ambos lados de la ecuación convergen o divergen simultáneamente. Para demostrar esta relación, sea $A > 0$ y sea F la f.d. de X. Integrando por partes obtenemos

$$\int_0^A x dF(x) = -A(1 - F(A) + \int_0^A (1 - F(x))\, dx$$
$$= -A(P(X > A) + \int_0^A P(X > x)\, dx.$$

Si $\mathrm{E}(X) < \infty$ entonces

$$A(1 - F(A)) \leq \int_A^\infty x\, dF(x) \to 0 \quad \text{cuando } A \to \infty,$$

y en consecuencia la integral al lado derecho converge. Por otro lado, si la integral al lado derecho converge, la integral del lado izquierdo también, pues es más pequeña.

Observamos que como $P(X > x)$ y $P(X \geq x)$ difieren en una cantidad numerable de valores de x,

$$\mathrm{E}(X) = \int_0^\infty P(X > x)\, dx = \int_0^\infty P(X \geq x)\, dx. \tag{2.11}$$

Demostración del Lema 2.1. Partimos de la relación (2.11):

$$\mathrm{E}(|X_1|) = \int_0^\infty P(|X_1| \geq x)\, dx = \sum_{n=0}^\infty \int_n^{n+1} P(|X_1| \geq x)\, dx$$

y por lo tanto

$$\sum_{n=0}^\infty P(|X_1| \geq n+1) \leq \mathrm{E}(|X_1|) \leq \sum_{n=0}^\infty P(|X_1| \geq n),$$

es decir,

$$\mathrm{E}(|X_1|) < \infty \quad \Leftrightarrow \quad \sum_{n=0}^\infty P(|X_1| \geq n) < \infty.$$

Ponemos ahora $Y = X_1/\varepsilon$ y obtenemos

$$\begin{aligned} \mathrm{E}(|X_1|) < \infty \quad &\Leftrightarrow \quad \mathrm{E}(|Y|) < \infty \\ &\Leftrightarrow \quad \sum_{n=0}^\infty P(|Y| \geq n) < \infty \\ &\Leftrightarrow \quad \sum_{n=0}^\infty P(|X_1| \geq \varepsilon n) < \infty. \end{aligned}$$

Esto demuestra que las dos primeras proposiciones son equivalentes. Para ver la otra equivalencia comenzamos con (b): Dado $\varepsilon > 0$

$$\sum_{n\geq 1} P(|X_1| \geq \varepsilon n) = \sum_{n \geq 1} P(|X_n| \geq \varepsilon n) < \infty$$

y por el lema de Borel-Cantelli esto implica que

$$P\Big(\frac{|X_n|}{n} > \varepsilon \ i.v.\Big) = 0,$$

es decir que con probabilidad 1,

$$\limsup_{n\to\infty} \frac{|X_n|}{n} \leq \varepsilon \quad c.p.\ 1. \tag{2.12}$$

Como esto es válido para todo $\varepsilon > 0$ y la sucesión $|X_n|/n \geq 0$, obtenemos que (c) es cierto. El recíproco se demuestra de manera similar. ■

Teorema 2.3 (LFGN de Kolmogorov) *Sea $(X_n, n \geq 1)$ una sucesión de v.a.i.i.d. Existe $c \in \mathbb{R}$ tal que*

$$\bar{X}_n = \frac{1}{n} S_n \to c \quad c.p.\ 1 \tag{2.13}$$

si y solo si $\mathrm{E}(|X_1|) < \infty$*, y en este caso* $c = \mathrm{E}(X_1)$.

Demostración. Si $\mathrm{E}(|X_1|) < \infty$ el teorema de Etemadi implica el resultado. Veamos el recíproco, supongamos que (2.13) es válida. Tenemos

$$\frac{X_n}{n} = \frac{S_n - S_{n-1}}{n} = \frac{S_n}{n} - \Big(\frac{n-1}{n}\Big)\frac{S_{n-1}}{n-1} \to c - c = 0 \quad c.p.\ 1$$

y el lema 2.1 implica el resultado. ∎

2.3.1. Aplicaciones

Comenzamos por el teorema de Glivenko-Cantelli. Sea $(X_n, n \geq 1)$ una sucesión de v.a.i.i.d. con distribución común F. La función de distribución empírica (f.d.e.) basada en las primeras n variables de esta sucesión se denota $\widehat{F}_n(x)$ y se define para $x \in \mathbb{R}$ por

$$\widehat{F}_n(x, \omega) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(-\infty, x]}(X_i(\omega)).$$

Como vemos, la f.d.e. corresponde a la medida uniforme sobre los puntos $\{X_1(\omega), X_2(\omega), \dots, X_n(\omega)\}$ y por lo tanto $\widehat{F}_n$ tiene saltos de altura $1/n$ en los puntos de este conjunto. Si hay valores repetidos, la altura de los saltos es $1/n$ multiplicado por el número de repeticiones del valor.

Para x fijo, $\widehat{F}_n(x, \cdot)$ es una variable aleatoria acotada y es el estimador natural de la función de distribución $F(x)$. Podemos hallar su valor esperado

$$\begin{aligned} \mathrm{E}(\widehat{F}_n(x, \omega)) &= \frac{1}{n} \sum_{j=1}^{n} \mathrm{E}(\mathbf{1}_{\{X_j(\omega) \leq x\}}) \\ &= \frac{1}{n} \sum_{j=1}^{n} P(X_j(\omega) \leq x) = F(x) \end{aligned}$$

de modo que, para cada x, $\widehat{F}_n(x)$ es un estimador insesgado de $F(x)$. Mas aún, por la Ley Fuerte de los Grandes Números, para cada x existe un conjunto nulo A_x tal que

$$\lim_{n \to \infty} \widehat{F}_n(x, \omega) = F(x) \tag{2.14}$$

siempre que $\omega \notin A_x$.

El teorema de Glivenko-Cantelli nos dice que un resultado más fuerte es cierto. Nos dice que la convergencia en (2.14) es cierta simultáneamente para todo x con probabilidad 1 y además que la convergencia es uniforme.

Teorema 2.4 (Glivenko-Cantelli) *Sea $X_1, \dots, X_n$ una colección de v.a.i. con distribución común F y sea $\widehat{F}_n(x) = \widehat{F}_n(x, \omega)$ la función de distribución empírica correspondiente. Entonces*

$$\sup_x |\widehat{F}_n(x) - F(x)| \to 0$$

con probabilidad 1.

Antes de ver la demostración de este teorema introducimos la función de cuantiles asociada a la f.d. F. La función de cuantiles (f.c.) Q es la inversa (generalizada) de F: para cualquier $u \in (0, 1)$

$$Q(u) = F^{\leftarrow}(u) = \inf\{s : F(s) \geq u\}.$$

Propiedades

1. Q es creciente (en sentido amplio) y continua por la izquierda.

   Supongamos $u < v$ ambos en $(0, 1)$, entonces

   $$Q(u) = \text{ínf}\{s : F(s) \geq u\} \leq \text{ínf}\{s : F(s) \geq v\} = Q(v).$$

   Para ver que es continua por la izquierda supongamos que $u \in (0, 1)$, $u_n \uparrow u$ pero $Q(u_n) \uparrow Q(u^-) < Q(u)$. Entonces existen $\delta > 0$ e y tales que para todo n,

   $$Q(u_n) < y < Q(u) - \delta$$

   La primera desigualdad y la definición de Q dicen que $F(y) \geq u_n$ para todo n, y haciendo $n \to \infty$ obtenemos $F(y) \geq u$, de donde, por la definición de Q, obtenemos $y \geq Q(u)$ pero $y < Q(u) - \delta$, lo cual es una contradicción.

2. $F(Q(u)) \geq u$.

   Comenzamos por ver que el conjunto $A(u) = \{s : F(s) \geq u\}$ es cerrado. Si $s_n \in A(u)$ y $s_n \downarrow s$ entonces $u \leq F(s_n) \downarrow F(s)$, de modo que $F(s) \geq u$ y $s \in A(u)$. Si $s_n \uparrow s$ y $s_n \in A(u)$ entonces $u \leq F(s_n) \uparrow F(s^-) \leq F(s)$ y $F(s) \geq u$, de modo que $s \in A(u)$ de nuevo y $A(u)$ es cerrado. Como $A(u)$ es cerrado, ínf $A(u) \in A(u)$, es decir, $Q(u) \in A(u)$, lo que implica que $F(Q(u)) \geq u$.

3. $Q(u) \leq t$ sii $u \leq F(t)$; $Q(u) > t$ sii $u > F(t)$.

   Esto es consecuencia de la definición de Q.

4. $F(Q(u)^-) \leq u \leq F(Q(u))$.

   La segunda desigualdad es la propiedad 2. Para ver la primera tomamos $t < Q(u)$, por la propiedad anterior esto equivale a $F(t) < u$. Haciendo $t \uparrow Q(u)$ obtenemos la primera desigualdad.

5. Sea $X \sim F$ y definamos $Y = aX + b$. La f.d. de Y es

   $$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P(X \leq \frac{y-b}{a}) = F_X(\frac{y-b}{a})$$

   Una transformación de este tipo se conoce como un *cambio de ubicación y escala*; a es el parámetro de escala y b el de ubicación. La función de cuantiles de Y es

   $$Q_Y(u) = aQ_X(u) + b$$

   (Esto es fácil de ver si F_X es continua y estrictamente creciente, de modo que Q_X es la inversa de F_X). Por lo tanto, una transformación lineal de la variable X, produce una transformación lineal del mismo tipo en la función de cuantiles.

6. Sea $([0, 1], \mathcal{B}, m)$ el espacio de Lebesgue y sea U una v.a. con distribución uniforme (su f.d. es la función identidad en $[0, 1]$). Si F es una f.d. y Q es la f.c. correspondiente, entonces $Q(U(\cdot))$ es una variable aleatoria en $[0, 1]$ con f.d. F:

   $$m(Q(U) \leq t) = m(U \leq F(t)) = F(t).$$

La función de cuantiles empírica (f.c.e.) $\widehat{Q}_n$ es la inversa generalizada de $\widehat{F}_n$, es decir, es la función definida sobre $(0, 1)$ con valores en $\mathbb{R}$ tal que, dado u, $0 < u < 1$, $\widehat{Q}_n(u)$ es el menor valor a la izquierda del cual se encuentra una proporción u de los datos.

Si llamamos $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ a los estadísticos de orden de la muestra, entonces

$$\widehat{Q}_n(u) = X_{(i)} \quad \text{cuando } \frac{i-1}{n} < u \leq \frac{i}{n}.$$

Regresamos ahora a la demostración del teorema

Demostración del teorema de Glivenko-Cantelli. Al igual que en la demostración de (2.14) pero con los conjuntos $\{X_j(\omega) < x\}$ en lugar de $\{X_j(\omega) \leq x\}$, la LFGN implica que

$$\lim_{n\to\infty} F_n(x^-, \omega) = F(x^-)$$

salvo en un conjunto B_x de medida 0.

Sea $x_{m,k} = Q(k/m)$, $1 \leq k \leq m, m \geq 1$. Por la propiedad 4 de las funciones de cuantil $F(x_{m,k}^-) \leq (k/m) \leq F(x_{m,k})$ y en consecuencia

$$F(x_{m,k}^-) - F(x_{m,k-1}) \leq \frac{1}{m}, \qquad F(x_{m,1}^-) \leq \frac{1}{m}, \qquad 1 - F(x_{m,m}) \leq \frac{1}{m}.$$

Definimos

$$D_n(\omega) = \sup_x |\widehat{F}_n(x,\omega) - F(x)|$$
$$D_{m,n}(\omega) = \max_{1\leq k\leq m} \left\{ |\widehat{F}_n(x_{m,k},\omega) - F(x_{m,k})|, |\widehat{F}_n(x_{m,k}^-,\omega) - F(x_{m,k}^-)| \right\}$$

Si $x_{m,k-1} \leq x < x_{m,k}$ tenemos

$$\widehat{F}_n(x,\omega) \leq \widehat{F}_n(x_{m,k}^-,\omega) \leq F(x_{m,k}^-) + D_{m,n}(\omega) \leq F(x) + \frac{1}{m} + D_{m,n}(\omega)$$
$$\widehat{F}_n(x,\omega) \geq \widehat{F}_n(x_{m,k-1},\omega) \geq F(x_{m,k-1}) - D_{m,n}(\omega) \geq F(x) - \frac{1}{m} - D_{m,n}(\omega)$$

Es posible demostrar desigualdades similares para los casos $x < x_{m,1}$ y $x \geq x_{m,m}$ y a partir de todas estas desigualdades podemos concluir que

$$D_n(\omega) \leq D_{m,n}(\omega) + \frac{1}{m}. \tag{2.15}$$

Si ω no está en el conjunto $\cup_{m,k}(A_{x_{m,k}} \cup B_{x_{m,k}})$, que tiene probabilidad 0, entonces $\lim_n D_{m,n}(\omega) = 0$ y en consecuencia $\lim_n D_n(\omega) = 0$ por (2.16). ■

La segunda aplicación de la LFGN es al método MonteCarlo, inventado en la década de 1940 por Stanislav Ulam. Aunque el método puede ser aplicado en muchos contextos, vamos a presentarlo considerando el problema de aproximar una integral definida. Para simplificar supongamos que f es una función acotada, definida en el intervalo $[a, b]$ y sea U una v.a. con distribución uniforme en el mismo intervalo. La función $f(U)$ también es una variable aleatoria cuyo valor esperado es

$$\mathrm{E}(f(U)) = \int_a^b f(x)\, dF_U(x) = \int_a^b f(x)\frac{1}{b-a}\, dx$$

ya que U tiene distribución uniforme. Por lo tanto, si $(U_n, n \geq 1)$ es una sucesión de v.a.i. con distribución uniforme en $[a, b]$, la LFGN nos dice que, con probabilidad 1,

$$\frac{b-a}{n}\sum_{i=1}^n f(U_i) \to (b-a)\,\mathrm{E}(f(U)) = \int_a^b f(x)\, dx.$$

El método MonteCarlo consiste en simular las v.a. U_n y calcular la suma en la expresión anterior para aproximar el valor de la integral.

En general, si A es una cantidad que deseamos aproximar numéricamente y existe una función f y una variable aleatoria X cuya distribución es fácil de simular en una computadora y se satisface que

$$A = \mathrm{E}(f(X))$$

entonces podemos aplicar el método MonteCarlo para aproximar A.

Ejemplo 2.1
Supongamos que queremos hallar la probabilidad de que una variable X con distribución normal de parámetros $\mu = 0$ y $\sigma^2 = 2$ tome valores en el intervalo $[0, 2]$, esto es

$$P(X \in [0,2]) = \int_0^2 \frac{1}{\sqrt{4\pi}} \exp\{-\frac{x^2}{4}\}\, dx$$

Para esto generamos variables uniformes en $[0, 2]$, evaluamos el integrando en estos valores, calculamos su promedio y lo multiplicamos por 2. El código de `R` que sigue hace esto.

```
n <- 1000
x <- runif(n)*2
y <- exp(-x^2/4)/(2*sqrt(pi))
prob <- 2*sum(y)/n
```

Para tener una idea de la precisión de la aproximación, comparamos con el valor que obtenemos usando la función `pnorm`, que calcula los valores de la función de distribución normal. La instrucción sería `pnorm(2,sd=sqrt(2))-0.5`. La siguiente tabla presenta los resultados para diversos valores de n

Cuadro 2.1: Ejemplo del método MonteCarlo para la aproximación de una probabilidad Gaussiana. La tabla presenta los valores de la aproximación para diferentes tamaños de la muestra y compara con el valor que se obtiene usando la función `pnorm` en `R`.

n	Valor	Error
100	0.41598	-0.00537
1,000	0.41636	-0.004989
10,000	0.42223	0.000879
100,000	0.42140	0.0000465
1,000,000	0.42130	-0.0000599

▲

Supongamos ahora, para simplificar, que la función f está definida en el intervalo $[0, 1]$. Teniendo en cuenta que

$$\mathrm{E}\Big(\frac{1}{n}\sum_1^n f(U_i)\Big) = \int_0^1 f(x)\, dx$$

es natural considerar la varianza (o la desviación típica) del promedio como una medida del error que cometemos en la aproximación:

$$\mathrm{Var}\Big(\frac{1}{n}\sum_1^n f(U_i)\Big) = \frac{1}{n^2}\mathrm{Var}\Big(\sum_1^n f(U_i)\Big)$$

y teniendo en cuenta la independencia de las variables

$$\mathrm{Var}\Big(\frac{1}{n}\sum_1^n f(U_i)\Big) = \frac{1}{n^2}\sum_1^n \mathrm{Var}(f(U_i)) = \frac{1}{n}\mathrm{Var}(f(U_1)).$$

Recordando que $\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X)^2)$, podemos estimar el valor de la varianza del promedio por

$$\frac{1}{n}\Big[\frac{1}{n}\sum_{i=1}^n f^2(U_i) - \Big(\frac{1}{n}\sum_{i=1}^n f(U_i)\Big)^2\Big].$$

Es usual tomar la desviación típica como medida del error y los resultados anteriores muestran que la desviación típica decrece como $1/\sqrt{n}$. Por lo tanto, para reducir el error en un orden de magnitud, que equivale a dividir el error por 10, hay que incrementar el tamaño de la muestra dos órdenes de magnitud, que equivale a multiplicar el tamaño de la muestra por 100.

2.4. Convergencia c.s. de Sumas de v.a.i.

En esta sección no supondremos que las variables tienen igual distribución.

Proposición 2.2 (Desigualdad de Skorohod) *Sea $\{X_n, n \geq 1\}$ una sucesión de v.a.i. y sea $\alpha > 0$ fijo. Para $n \geq 1$ definimos $S_n = \sum_{i=1}^n X_i$ y*

$$c = \sup_{j \leq N} P(|S_N - S_j| > \alpha).$$

Suponemos también que $c < 1$, entonces

$$P(\sup_{j \leq N} |S_j| > 2\alpha) \leq \frac{1}{1-c} P(|S_N| > \alpha).$$

Demostración. Definimos

$$J = \inf\{j : |S_j| > 2\alpha\},$$

con la convención de que $\inf \varnothing = \infty$. Observamos que

$$\{\sup_{j \leq N} |S_j| > 2\alpha\} = \{J \leq N\} = \bigcup_{j=1}^N \{J = j\},$$

donde la unión es sobre conjuntos disjuntos. Tenemos

$$\begin{aligned} P(|S_N| > \alpha) &\geq P(|S_N| > \alpha, J \leq N) \\ &= \sum_{j=1}^N P(|S_N| > \alpha, J = j) \\ &\geq \sum_{j=1}^N P(|S_N - S_j| \leq \alpha, J = j). \end{aligned} \tag{2.16}$$

Para justificar el último paso supongamos que

$$|S_N(\omega) - S_j(\omega)| \leq \alpha, \quad \text{y} \quad J(\omega) = j$$

de modo que $|S_j(\omega)| > 2\alpha$. Si fuese el caso que $|S_N(\omega)| \leq \alpha$, entonces sería cierto que $|S_N(\omega) - S_j(\omega)| > \alpha$, lo cual es imposible, de modo que $|S_N(\omega)| > \alpha$. Hemos verificado que

$$\{|S_N - S_j| \leq \alpha, J = j\} \subset \{|S_N| > \alpha, J = j\}$$

lo que justifica (2.16), es decir,

$$P(|S_N| > \alpha) \geq \sum_{j=1}^N P(|S_N - S_j| \leq \alpha, J = j).$$

Además

$$S_N - S_j = \sum_{i=j+1}^N X_j \in \sigma(X_{j+1}, \dots, X_N)$$

y

$$\{J = j\} = \{\sup_{i<j} |S_i| \leq 2\alpha, |S_j| > 2\alpha\} \in \sigma(X_1, \dots, X_j)$$

y por lo tanto estos eventos son independientes. En consecuencia

$$\begin{aligned}P(|S_N| > \alpha) &\geq \sum_{j=1}^{N} P(|S_N - S_j| \leq \alpha)P(J = j)\\ &\geq \sum_{j=1}^{N}(1-c)P(J=j)\\ &= (1-c)P(J \leq N)\\ &= (1-c)P(\sup_{j\leq N}|S_j| > 2\alpha).\end{aligned}$$

■

Para la demostración del teorema de Lévy necesitamos el siguiente resultado

Lema 2.2 *Sea $(X_n, n \geq 1)$ una sucesión monótona de v.a. Si $X_n \to X$ en probabilidad entonces $X_n \to X$ c.s.*

Demostración. Ver problema 21 de la lista 1. ■

Teorema 2.5 (Lévy) *Si $\{X_n, n \geq 1\}$ es una sucesión de v.a.i. entonces*

$$\sum_n X_n \quad \textit{converge en probabilidad} \;\Leftrightarrow\; \sum_n X_n \quad \textit{converge c.p.1.}$$

Si $S_n = \sum_{i=1}^n X_i$ el enunciado anterior dice que las siguientes afirmaciones son equivalentes

1. *(S_n) es de Cauchy en probabilidad.*
2. *(S_n) converge en probabilidad.*
3. *(S_n) converge con probabilidad 1.*
4. *(S_n) es de Cauchy con probabilidad 1.*

Demostración. Supongamos que (S_n) converge en probabilidad, de modo que es de Cauchy en probabilidad. Demostraremos que es de Cauchy c.s. y en consecuencia converge c.s. Para ver que es de Cauchy c.s. necesitamos demostrar que

$$\xi_N = \sup_{m,n\geq N}|S_m - S_n| \to 0 \quad c.s.,$$

cuando $N \to \infty$. Observamos que $(\xi_N,\ N \geq 1)$ es una sucesión decreciente y por lo tanto basta mostrar que $\xi_N \xrightarrow{P} 0$ cuando $N \to \infty$. Como

$$\begin{aligned}\xi_N &= \sup_{m,n\geq N}|S_m - S_N + S_N - S_n| \leq \sup_{m\geq N}|S_m - S_N| + \sup_{n\geq N}|S_n - S_N|\\ &= 2\sup_{n\geq N}|S_n - S_N| = 2\sup_{j\geq 0}|S_{N+j} - S_N|,\end{aligned}$$

basta probar que

$$\sup_{j\geq 0}|S_{N+j} - S_N| \xrightarrow{P} 0. \tag{2.17}$$

Dados $\varepsilon > 0$ y $0 < \delta < 1/2$, la hipótesis de que (S_n) es de Cauchy en probabilidad implica que existe $N_{\varepsilon,\delta}$ tal que

$$P(|S_m - S_{m'}| > \frac{\varepsilon}{2}) \leq \delta \tag{2.18}$$

siempre que $m, m' \geq N_{\varepsilon,\delta}$, y entonces

$$P(|S_{N+j} - S_N| > \frac{\varepsilon}{2}) \leq \delta, \quad \forall j \geq 0, \tag{2.19}$$

siempre que $N \geq N_{\varepsilon,\delta}$. Tenemos

$$\begin{aligned} P(\sup_{j\geq 0}|S_{N+j} - S_N| > \varepsilon) &= P(\lim_{N'\to\infty}[\sup_{N'\geq j\geq 0}|S_{N+j} - S_N| > \varepsilon]) \\ &= \lim_{N'\to\infty} P(\sup_{N'\geq j\geq 0}|S_{N+j} - S_N| > \varepsilon). \end{aligned}$$

Queremos ahora usar la desigualdad de Skorohod. Sea $X_i' = X_{N+i}$ y

$$S_j' = \sum_{i=1}^{j} X_i' = \sum_{i=1}^{j} X_{N+i} = S_{N+j} - S_N.$$

Con esta notación tenemos

$$\begin{aligned} P(\sup_{N'\geq j\geq 0}|S_{N+j} - S_N| > \varepsilon) &= P(\sup_{N'\geq j\geq 0}|S_j'| > \varepsilon) \\ &\leq \frac{1}{1 - \sup_{j\leq N'} P(|S_{N'}' - S_j'| > \frac{\varepsilon}{2})} P(|S_{N'}'| > \frac{\varepsilon}{2}) \\ &\leq \frac{1}{1-\delta}\,\delta \leq 2\delta \end{aligned}$$

teniendo en cuenta la selección de δ. Observamos ahora que usando (2.18) obtenemos

$$\sup_{j\leq N'} P(|S_{N'}' - S_j'| > \frac{\varepsilon}{2}) = \sup_{j\leq N'} P(|S_{N+N'} - S_{N+j}| > \frac{\varepsilon}{2}) \leq \delta.$$

Como podemos escoger δ arbitrariamente pequeño, esto demuestra el resultado. ■

Teorema 2.6 (Criterio de Convergencia de Kolmogorov) *Supongamos que* $(X_n,\ n \geq 1)$ *es una sucesión de v.a.i. Si*

$$\sum_n \mathrm{Var}(X_n) < \infty,$$

entonces

$$\sum_{n\geq 1}(X_n - \mathrm{E}(X_n)) \quad \textit{converge c.p.1.}$$

Demostración. Sin pérdida de generalidad podemos suponer que $\mathrm{E}(X_j) = 0$. Entonces la suma de las varianzas es

$$\sum_{n\geq 1} \mathrm{E}\, X_n^2 < \infty.$$

Esto implica que (S_n) es de Cauchy en L^2 porque para $m < n$

$$||S_n - S_m||_2^2 = \mathrm{Var}(S_n - S_m) = \sum_{j=m+1}^{n} \mathrm{E}\, X_j^2 \to 0,$$

cuando $m, n \to \infty$. En consecuencia (S_n) también es de Cauchy en probabilidad porque

$$P(|S_n - S_m| \geq \varepsilon) \leq \varepsilon^{-2}\,\mathrm{Var}(S_n - S_m) = \varepsilon^{-2} \sum_{j=m+1}^{n} \mathrm{Var}(X_j) \to 0$$

cuando $n, m \to \infty$. Por el Teorema de Lévy (S_n) es convergente c.p.1. ■

2.5. Leyes Fuertes de Grandes Números II

Lema 2.3 (Kronecker) *Sean x_k y a_k dos sucesiones de números reales tales que $0 < a_k \uparrow \infty$. Si*

$$\sum_{k=1}^{\infty} \frac{x_k}{a_k} \quad \textit{converge},$$

entonces

$$\lim_{n\to\infty} \frac{1}{a_n} \sum_{k=1}^{n} x_k = 0.$$

Demostración. Sea $r_n = \sum_{k=n+1}^{\infty} x_k/a_k$ de modo que $r_n \to 0$ cuando $n \to \infty$. Dado $\varepsilon > 0$ existe $N_0 = N_0(\varepsilon)$ tal que si $n \geq N_0$ tenemos $|r_n| \leq \varepsilon$. Además

$$\frac{x_k}{a_k} = r_{k-1} - r_k$$

de modo que

$$\begin{aligned}\sum_{k=1}^{n} x_k &= \sum_{k=1}^{n} (r_{k-1} - r_k) a_k \\ &= \sum_{j=1}^{n-1} (a_{j+1} - a_j) r_j + a_1 r_0 - a_n r_n.\end{aligned}$$

Entonces, para $n \geq N_0$,

$$\begin{aligned}\frac{1}{a_n} \Big|\sum_{k=1}^{n} x_k\Big| &\leq \frac{1}{a_n} \Big(\sum_{j=1}^{N_0-1} (a_{j+1} - a_j)|r_j| + \sum_{j=N_0}^{n-1} (a_{j+1} - a_j)|r_j| \Big) + \Big|\frac{a_1 r_0}{a_n}\Big| + \Big|\frac{a_n r_n}{a_n}\Big| \\ &= \frac{Const}{a_n} + \frac{\varepsilon}{a_n} \sum_{j=N_0}^{n-1} (a_{j+1} - a_j) + |r_n| \\ &= o(1) + \frac{\varepsilon(a_n - a_{N_0})}{a_n} + \varepsilon \\ &\leq 2\varepsilon + o(1).\end{aligned}$$

y esto demuestra el resultado. ■

El lema de Kronecker permite demostrar la siguiente ley fuerte.

Corolario 2.3 *Sea $(X_n,\ n \geq 1)$ una sucesión de v.a.i. con $\mathrm{E}(X_n^2) < \infty$. Supongamos que existe una sucesión $b_n \uparrow \infty$ tal que*

$$\sum_k \mathrm{Var}\Big(\frac{X_k}{b_k}\Big) < \infty,$$

entonces

$$\frac{S_n - \mathrm{E}(S_n)}{b_n} \to 0 \quad \textit{c.p.}1.$$

Demostración. Como la suma de las varianzas es finita, por el criterio de convergencia de Kolmogorov tenemos que

$$\sum_n \frac{X_n - \mathrm{E}(X_n)}{b_n} \quad \text{converge c.p.1.}$$

y por el lema de Kronecker

$$\frac{1}{b_n}\sum_{k=1}^{n}(X_k - \mathrm{E}(X_k)) \to 0$$

con probabilidad 1. ■

Ejemplo 2.2
Sea $(X_n,\ n \geq 1)$ una sucesión i.i.d. con distribución común continua F. Definimos

$$\mu_n = \sum_{j=1}^{n} \mathbf{1}_{\{X_j \text{ es un record}\}} = \sum_{j=1}^{n} \mathbf{1}_j$$

donde $\mathbf{1}_j = \mathbf{1}_{\{X_j \text{ es un record}\}}$, de modo que μ_n es el número de records en las primeras N observaciones.

Proposición 2.3 *El número de records en una sucesión i.i.d. crece logaritmicamente y se tiene con probabilidad 1*

$$\lim_{n\to\infty} \frac{\mu_n}{\log n} = 1.$$

Para la demostración usaremos el siguiente resultado de Análisis Real:

$$\sum_{j=1}^{n} \frac{1}{j} - \log n \to c$$

cuando $n \to \infty$ donde $c > 0$ es la constante de Euler.

Demostración. Tenemos que las $\mathbf{1}_j$ son independientes,

$$\mathrm{E}(\mathbf{1}_j) = P(\mathbf{1}_j = 1) = \frac{1}{j},$$

$$\mathrm{Var}(\mathbf{1}_j) = \mathrm{E}(\mathbf{1}_j^2) - (\mathrm{E}\,\mathbf{1}_j)^2 = \frac{1}{j} - \frac{1}{j^2} = \frac{j-1}{j^2}.$$

Usando esto obtenemos

$$\sum_{j=2}^{\infty} \mathrm{Var}\left(\frac{\mathbf{1}_j}{\log j}\right) = \sum_{j=2}^{\infty} \frac{1}{(\log j)^2}\,\mathrm{Var}(\mathbf{1}_j) = \sum_{j=2}^{\infty} \frac{j-1}{j^2(\log j)^2} < \infty$$

Por lo tanto el criterio de convergencia de Kolmogorov implica que

$$\sum_{j=2}^{\infty}\left(\frac{\mathbf{1}_j - \mathrm{E}(\mathbf{1}_j)}{\log j}\right) = \sum_{j=2}^{\infty} \frac{\mathbf{1}_j - 1/j}{\log j}$$

converge y por el lema de Kronecker

$$\frac{1}{\log n}\sum_{j=1}^{n}\left(\mathbf{1}_j - \frac{1}{j}\right) \to 0 \quad c.p.1,$$

pero

$$\frac{1}{\log n}\sum_{j=1}^{n}\left(\mathbf{1}_j - \frac{1}{j}\right) = \frac{1}{\log n}\Big(\sum_{j=1}^{n}\mathbf{1}_j - \sum_{j=1}^{n}\frac{1}{j}\Big) = \frac{1}{\log n}\Big(\mu_n - \sum_{j=1}^{n}\frac{1}{j}\Big).$$

Por lo tanto

$$\frac{\mu_n}{\log n} - 1 = \frac{\mu_n - \sum_{j=1}^{n} j^{-1}}{\log n} + \frac{\sum_{j=1}^{n} j^{-1} - \log n}{\log n} \to 0.$$

■

2.6. El Teorema de las Tres Series de Kolmogorov

Teorema 2.7 (Teorema de la Tres Series) *Sea* $\{X_n,\ n \geq 1\}$ *una sucesión de v.a. independientes. Para que* $\sum_n X_n$ *converja con probabilidad 1 es necesario y suficiente que exista* $c > 0$ *tal que*

(a) $\sum_n P(|X_n| > c) < \infty$.

(b) $\sum_n \mathrm{Var}(X_n \mathbf{1}_{\{|X_n| \leq c\}}) < \infty$.

(c) $\sum_n \mathrm{E}(X_n \mathbf{1}_{\{|X_n| \leq c\}})$ *converge.*

Demostración de la suficiencia. Supongamos que las tres series convergen y definamos

$$X'_n = X_n \mathbf{1}_{\{|X_n| \leq c\}}.$$

Por (a)

$$\sum_n P(X_n \neq X'_n) = \sum_n P(|X_n| > c) < \infty,$$

de modo que (X_n) y (X'_n) son asintóticamente equivalentes. Por lo tanto $\sum_n X_n$ converge c.s. si y sólo si $\sum_n X'_n$ converge c.s.

Usando (b) tenemos

$$\sum_n \mathrm{Var}(X'_n) < \infty$$

y por el criterio de convergencia de Kolmogorov

$$\sum_n (X'_n - \mathrm{E}(X'_n)) \quad \text{converge c.s.}$$

Por (c) tenemos que $\sum_n \mathrm{E}(X'_n)$ converge y en consecuencia $\sum_n X'_n$ converge. ■

La demostración de la necesidad puede verse en el libro de Resnick, o en el de Chung.

Corolario 2.4 *Si* (X_n) *son v.a.i. con* $\mathrm{E}\, X_n = 0$, $n \geq 1$ *y*

$$\sum_{n=1}^{\infty} \mathrm{E}\left(X_n^2 \mathbf{1}_{\{|X_n| \leq 1\}} + |X_n| \mathbf{1}_{\{|X_n| > 1\}}\right) < \infty \tag{2.20}$$

entonces $\sum_n X_n$ *converge c.s.*

Demostración. Ejercicio (Ver problema 22, lista de problemas 3). ■

Corolario 2.5 (Loève) *Si* $(X_n,\ n \geq 1)$ *son v.a.i. y para ciertas constantes* $0 < \alpha_n \leq 2$ *se tiene que* $\sum_n \mathrm{E}\, |X_n|^{\alpha_n} < \infty$, *con* $\mathrm{E}(X_n) = 0$ *cuando* $1 \leq \alpha_n \leq 2$, *entonces* $\sum_n X_n$ *converge c.s.*

Demostración. Podemos considerar por separado los casos $1 \leq \alpha_n \leq 2$, $n \geq 1$ y $0 < \alpha_n < 1$, $n \geq 1$. En el primer caso

$$X_n^2 \mathbf{1}_{\{|X_n| \leq 1\}} + |X_n| \mathbf{1}_{\{|X_n| > 1\}} \leq |X_n|^{\alpha_n},$$

de donde se obtiene (2.20). En el segundo

$$\sum_1^{\infty} \mathrm{E}(X_n^2 + |X_n|) \mathbf{1}_{\{|X_n| \leq 1\}} \leq 2 \sum_1^{\infty} \mathrm{E}\, |X_n|^{\alpha_n} < \infty,$$

y en ambos casos

$$\sum_1^\infty P(|X_n| \geq 1) \leq \sum_1^\infty \mathrm{E}\,|X_n|^{\alpha_n} < \infty,$$

de modo que las tres series del teorema 2.7 convergen. ■

Ejemplo 2.3
Sea $\{X_n,\ n \geq 1\}$ una sucesión de v.a.i.i.d. con

$$P(X_1 = 1) = \frac{1}{2} = P(X_1 = -1)$$

y consideremos la serie

$$\sum_n \frac{X_n}{n}, \tag{2.21}$$

que es la serie armónica con signos aleatorios. La pregunta que nos hacemos es si esta serie es convergente c.s.

Podemos usar el Teorema de las Tres Series con $c = 1$ y todos los términos de las series (a) y (c) valen cero, de modo que estas convergen trivialmente. Por otro lado

$$\mathrm{Var}(X_n/n) = \frac{1}{n^2}$$

de modo que la serie (b) también converge y en consecuencia (2.21) converge c.s.

Finalmente, enunciamos sin demostración un resultado conocido como la LFGN de Marcinkiewicz-Zygmund. La demostración se puede ver en el libro de Chow & Teicher.

Teorema 2.8 *Si $(X_n,\ n \geq 1)$ es una sucesión i.i.d. y $S_n = \sum_1^n X_i$ entonces para cualquier $p \in (0,2)$*

$$\frac{S_n - nc}{n^{1/p}} \to 0 \quad c.p.1$$

para alguna constante c si y sólo si $\mathrm{E}\,|X_1|^p < \infty$ y en este caso $c = \mathrm{E}(X_1)$ si $1 \leq p < 2$ mientras que c es arbitraria (y la podemos tomar igual a 0) para $0 < p < 1$.