Una demostración alternativa de la ley fuerte de los grandes números

José R. Berrendero*

Departamento de Matemáticas, Universidad Autónoma de Madrid, 28049 Madrid

Resumen

Estas notas contienen una demostración completa de la ley fuerte de los grandes números (LFGN) para variables aleatorias independientes e idénticamente distribuidas (v.a.i.i.d.) con media finita. La demostración es esencialmente autocontenida. Además de las propiedades elementales de las medidas de probabilidad y la esperanza, solo se usan sin demostración tres resultados muy conocidos: la desigualdad de Markov, el primer lema de Borel-Cantelli y el teorema de la convergencia dominada. La demostración está tomada de Steele (2015) y se han completado los detalles para que pueda ser comprendida por los alumnos de la asignatura Probabilidad II del Grado en Matemáticas de la Universidad Autónoma de Madrid.

1. Preliminares

Sean X y X_n, $n = 1, 2, \ldots$, variables aleatorias definidas en un espacio de probabilidad común $(\Omega, \mathcal{F}, \mathrm{P})$. Se dice que la sucesión X_n converge casi seguro a X, y se denota $X_n \xrightarrow{\text{c.s.}} X$, si $\mathrm{P}(\omega : X_n(\omega) \to X(\omega)) = 1$.

Teniendo en cuenta la definición de límite superior de una sucesión de sucesos ($\limsup A_n = \bigcap_{n=1}^\infty \bigcup_{k=n}^\infty A_k$) es muy fácil comprobar que

$$X_n \xrightarrow{\text{c.s.}} X \Leftrightarrow \mathrm{P}(\limsup\{|X_n - X| \geq \epsilon\}) = 0, \quad \text{para todo } \epsilon > 0.$$

En todos los resultados que siguen suponemos que $X_1, X_2, \ldots$ son v.a.i.i.d. Denotamos por $\mu_k = \mathrm{E}(X_1^k)$ al momento (común a todas las variables) de orden k y $\mu = \mu_1$ a la esperanza. También usamos la notación habitual $S_n = X_1 + \cdots + X_n$. Finalmente, I_A denota la variable aleatoria que vale 1 si $\omega \in A$ y 0 en caso contrario.

2. Lemas previos

El primer lema es una versión de la LFGN de Cantelli para variables con momento de cuarto orden finito. El objetivo final es probar el mismo resultado, pero suponiendo únicamente esperanzza finita.

Lema 1. *Sean $X_1, X_2, \ldots$ v.a.i.i.d. con $E(X_1^4) = \mu_4 < \infty$. Entonces*

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

*joser.berrendero@uam.es

Demostración. Sin perder generalidad se puede suponer que $\mu = 0$. Si fuese $\mu \neq 0$, aplicamos el resultado a las variables centradas $X_i - \mu$, que sí tienen esperanza igual a cero. Tenemos que probar que para todo $\epsilon > 0$, $\mathrm{P}(\limsup\{|S_n/n| \geq \epsilon\}) = 0$. Para ello, se aplica la desigualdad de Markov. Dado $\epsilon > 0$,

$$\mathrm{P}\left(\left|\frac{S_n}{n}\right| \geq \epsilon\right) = \mathrm{P}\left(\frac{S_n^4}{n^4} \geq \epsilon^4\right) \leq \frac{\mathrm{E}(S_n^4)}{n^4\epsilon^4}.$$

Por el primer lema de Borel-Cantelli, basta demostrar $\sum_{n=1}^{\infty} \mathrm{E}(S_n^4)/n^4 < \infty$. Vamos a calcular $\mathrm{E}(S_n^4)$. Se observa que

$$S_n^4 = \left(\sum_{i=1}^{n} X_i\right)^4 = \sum_{i=1}^{n} X_i^4 + 6\sum_{i<j} X_i^2 X_j^2 + \text{Resto}.$$

Al tomar esperanzas, $\mathrm{E}(\text{Resto}) = 0$ porque como las variables son independientes, la esperanza de los productos es igual al producto de las esperanzas y el término de resto siempre incluye algún factor $\mathrm{E}(X_i) = 0$. Por lo tanto, $\mathrm{E}(S_n^4) = n\mu_4 + 6n(n-1)/2\mu_2^2$. Al dividir esta expresión por n^4 se obtienen los términos de una serie convergente. □

Las demostraciones de la LFGN habitualmente involucran el uso de alguna desigualdad maximal. En la demostración clásica que suele aparecer en los libros de texto de probabilidad, se aplica la desigualdad maximal de Kolmogorov. En nuestro caso, usamos una desigualdad maximal debida a Hopf (1954):

Lema 2. *Sean $X_1, X_2, \ldots$ v.a. independientes e idénticamente distribuidas con $E|X_1| < \infty$. Sea $M_n = \max\{0, S_1, S_2, \ldots, S_n\}$ Entonces,*

$$E(X_1 I_{\{M_n>0\}}) \geq 0.$$

Demostración. Si $M_n > 0$, tenemos

$$M_n = \max\{S_1, S_2, \ldots, S_n\} = X_1 + \max\{0, X_2, X_2 + X_3, \ldots, X_2 + \cdots + X_n\} = X_1 + M_{n-1}^+,$$

donde definimos $M_n^+ = \max\{0, X_2, X_2 + X_3, \ldots, X_2 + \cdots + X_{n+1}\}$. Como consecuencia,

$$M_n I_{\{M_n>0\}} = X_1 I_{\{M_n>0\}} + M_{n-1}^+ I_{\{M_n>0\}} \leq X_1 I_{\{M_n>0\}} + M_n^+ I_{\{M_n>0\}},$$

ya que $M_{n-1}^+ \leq M_n^+$. Reordenando términos,

$$(M_n - M_n^+) I_{\{M_n>0\}} \leq X_1 I_{\{M_n>0\}}.$$

Dado que $(M_n - M_n^+) I_{\{M_n=0\}} = -M_n^+ I_{\{M_n=0\}} \leq 0$, usando las dos últimas desigualdades tenemos que

$$M_n - M_n^+ = (M_n - M_n^+) I_{\{M_n>0\}} + (M_n - M_n^+) I_{\{M_n=0\}} \leq X_1 I_{\{M_n>0\}}.$$

Si observamos las definiciones de M_n y M_n^+ vemos que tienen la misma distribución (puesto que las variables $X_1, X_2, \ldots$ son i.i.d.) Como consecuencia, al tomar esperanzas en la última desigualdad tenemos $\mathrm{E}(X_1 I_{\{M_n>0\}}) \geq 0$. □

Observación: Si se analiza la demostración anterior, se ve que no es necesario que las variables sean i.i.d. Lo único que hace falta es $\mathrm{E}(M_n) = \mathrm{E}(M_n^+)$, y esta condición se cumple si la sucesión X_n es estacionaria.

En el siguiente lema expresamos la desigualdad maximal de forma más conveniente para nuestros propósitos:

Lema 3. *Sean* $X_1, X_2, \dots$ *v.a. independientes e idénticamente distribuidas con* $E(X_1) = 0$. *Entonces, para todo* $\epsilon > 0$,

$$\mathrm{P}\left(\sup_{n\geq 1}\frac{S_n}{n} > \epsilon\right) \leq \frac{\mathrm{E}|X_1|}{\epsilon}.$$

Demostración. Sea $\epsilon > 0$. Consideramos las variables $X_i' = X_i - \epsilon$. Se observa que

$$M_n' = \max\{0, S_1', S_2', \dots, S_n'\} > 0 \Leftrightarrow \max\{S_1, S_2/2, \dots, S_n/n\} > \epsilon.$$

Al aplicar el lema 2 a las variables X_i' resulta

$$0 \leq \mathrm{E}(X_1' I_{\{M_n'>0\}}) = \mathrm{E}((X_1 - \epsilon) I_{\{\max\{S_1, S_2/2, \dots, S_n/n\} > \epsilon\}}) \leq \mathrm{E}|X_1| - \epsilon \mathrm{P}(\max\{S_1, S_2/2, \dots, S_n/n\} > \epsilon),$$

de donde

$$\mathrm{P}(\max\{S_1, S_2/2, \dots, S_n/n\} > \epsilon) \leq \frac{\mathrm{E}|X_1|}{\epsilon}.$$

Finalmente,

$$\mathrm{P}\left(\sup_{n\geq 1}\frac{S_n}{n} > \epsilon\right) = \lim_{k\to\infty} \mathrm{P}(\max\{S_1, S_2/2, \dots, S_k/k\} > \epsilon) \leq \frac{\mathrm{E}|X_1|}{\epsilon}.$$

□

3. Demostración de la LFGN

Teorema 1. *Sean* $X_1, X_2, \dots$ *v.a.i.i.d. con* $E(X_1) = \mu$ *finita. Entonces*

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

Demostración. De nuevo sin perder generalidad se puede suponer que $\mu = 0$. Tenemos que probar $S_n/n \xrightarrow{\text{c.s.}} 0$. Aplicamos a continuación la técnica de truncamiento en su versión más simple. Dado $K > 0$, $X_n = X_n I_{\{|X_n|\leq K\}} + X_n I_{\{|X_n|>K\}} \equiv X_n'(K) + X_n''(K)$. Se observa que las variables $X_i'(K)$ son i.i.d. y además

$$\lim_{K\to\infty} \mathrm{E}(X_1'(K)) = \lim_{K\to\infty} \int_\Omega X_1 I_{\{|X_1|\leq K\}} dP = \mathrm{E}(X_1) = 0,$$

por el teorema de la convergencia dominada (que es aplicable dado que X_1 es integrable). También, las variables $X_i''(K)$ son i.i.d. y aplicando de nuevo el teorema de la convergencia dominada,

$$\lim_{K\to\infty} \mathrm{E}(|X_1''(K)|) = \lim_{K\to\infty} \int_\Omega |X_1| I_{\{|X_1|>K\}} dP = 0.$$

Por el lema 1 se tiene $S_n'(K)/n \xrightarrow{\text{c.s.}} \mathrm{E}(X_1'(K))$. Es aplicable porque las variables están acotadas por K y tienen momentos finitos de todos los órdenes. Por tanto, con probabilidad igual a uno,

$$\limsup \frac{S_n}{n} = \lim_{n\to\infty} \frac{S_n'(K)}{n} + \limsup \frac{S_n''(K)}{n} \leq \mathrm{E}(X_1'(K)) + \sup_{n\geq 1} \frac{S_n''(K)}{n}.$$

Como consecuencia, dado $\epsilon > 0$ arbitrario,

$$\mathrm{P}\left(\limsup \frac{S_n}{n} > 2\epsilon\right) \leq \mathrm{P}\left(\mathrm{E}(X_1'(K)) + \sup_{n\geq 1}\frac{S_n''(K)}{n} > 2\epsilon\right) \leq \mathrm{P}\left(\sup_{n\geq 1}\frac{S_n''(K)}{n} > \epsilon\right), \quad (1)$$

donde la segunda desigualdad es cierta para K suficientemente grande ya que $\lim_{K\to\infty} \mathrm{E}(X_1'(K)) = 0$. La ecuación (1) y el lema 3 implican que para $\epsilon > 0$ arbitrario

$$0 \leq \mathrm{P}\left(\limsup \frac{S_n}{n} > 2\epsilon\right) \leq \frac{\mathrm{E}|X_1''(K)|}{\epsilon} \to 0, \text{ si } K \to \infty.$$

Es decir, con probabilidad 1 se tiene $\limsup(S_n/n) \leq 0$. Sustituyendo X_i por $-X_i$, deducimos que también con probabilidad 1, $\liminf(S_n/n) = -\limsup(-S_n/n) \geq 0$. Como consecuencia $0 \leq \liminf(S_n/n) \leq \limsup(S_n/n) \leq 0$ c.s., es decir, $S_n/n \to 0$, c.s. □

Referencias

Hopf, E. (1954). The general temporally discrete Markoff process. *Journal of Rational Mechanics and Analysis*, **3**, 13-45.

Steele, J.M. (2015). Explaining a Mysterious Maximal Inequality—and a Path to the Law of Large Numbers. *The American Mathematical Monthly*, **122**, 490-494.