TEMA 6. Leyes de los grandes números y teorema central del límite

Contenidos

- Nociones de convergencia de variables aleatorias
- Leyes de los grandes números
- Introducción al problema central de límite clásico

Convergencia de sucesiones de variables aleatorias

Sea $\{X_n\}_{n\in\mathbb{N}}$ y X variables aleatorias sobre el mismo espacio de probabilidad con funciones de distribución F_{X_n}, $n \in \mathbb{N}$, y F_X, respectivamente. Se definen los siguientes tipo de convergencia:

- *Convergencia casi segura*:

$$X_n \rightarrow^{c.s.} X, \quad n \to \infty \;\Leftrightarrow\; P\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

- *Convergencia en probabilidad*:

$$X_n \rightarrow^{P} X, \quad n \to \infty \;\Leftrightarrow\; \forall \varepsilon > 0,\; \lim_{n\to\infty} P\left(|X_n - X| \leq \varepsilon\right) = 1.$$

- *Convergencia en Ley o en distribución*

$$X_n \rightarrow^{L} X, \quad n \to \infty \;\Leftrightarrow\; \lim_{n\to\infty} F_{X_n}(x) = F_X, \quad \forall x \in C(F_X),$$

  siendo $C(F_X)$ el conjunto de puntos de continuidad de F_X.

Nota 1 *Los conceptos de convergencia anteriores son invariantes frente a translaciones, mediante una constante fija c, de las variables aleatorias involucradas.*

Sobre los anteriores conceptos de convergencia destacaremos las siguientes relaciones:

- $X_n \rightarrow^{c.s.} X \;\Rightarrow\; X_n \rightarrow^{P} X \;\Rightarrow\; X_n \rightarrow^{L} X.$
- $X_n \rightarrow^{L} c \ (c \in \mathbb{R}) \Rightarrow\; X_n \rightarrow^{P} c.$

Teorema de continuidad para funciones generatrices de momentos $\{X_n\}_{n\in\mathbb{N}}$ y X variables aleatorias sobre el mismo espacio de probabilidad. Supongamos que existen las funciones generatrices de momentos M_{X_n}, $n \in \mathbb{N}$, y M_X sobre un intervalo común de $\mathbb{R}$ con límites finitos o infinitos, conteniendo al cero, i.e., sobre $(-a, b),\ a, b \in \mathbb{R}_+ \cup \{\infty\}$. Se tiene entonces

$$M_{X_n}(t) \to M_X(t), \quad n \to \infty, \quad \forall t \in (-a, b) \;\Rightarrow X_n \rightarrow^{L} X.$$

Leyes de los grandes números

Sean $X_n,\ n \in \mathbb{N}$, variables aleatorias sobre el mismo espacio de probabilidad $(\Omega, \mathcal{A}, P)$ independientes. Adicionalmente, sea S_n la variable aleatoria que define las sumas parciales de la secuencia $\{X_n\}_{n\in\mathbb{N}}$, i.e.,

$$S_n = \sum_{i=1}^{n} X_i, \quad n \geq 1.$$

Se consideran dos secuencias de números reales $\{a_n\}_{n\in\mathbb{N}}$ y $\{b_n\}_{n\in\mathbb{N}}$ tal que $b_n \uparrow \infty$. Asociadas a dichas secuencias se introducen las siguientes leyes:

- *Ley débil de los grandes números respecto $\{a_n\}_{n\in\mathbb{N}}$ y $\{b_n\}_{n\in\mathbb{N}}$.* Se dice que la secuencia aleatoria $\{X_n\}_{n\in\mathbb{N}}$ satisface la ley débil respecto a $\{a_n\}_{n\in\mathbb{N}}$ y $\{b_n\}_{n\in\mathbb{N}}$ si
$$\frac{S_n - a_n}{b_n} \rightarrow^{P} 0.$$
- *Ley fuerte de los grandes números respecto $\{a_n\}_{n\in\mathbb{N}}$ y $\{b_n\}_{n\in\mathbb{N}}$.* Se dice que la secuencia aleatoria $\{X_n\}_{n\in\mathbb{N}}$ satisface la ley fuerte respecto a $\{a_n\}_{n\in\mathbb{N}}$ y $\{b_n\}_{n\in\mathbb{N}}$ si
$$\frac{S_n - a_n}{b_n} \rightarrow^{c.s.} 0.$$

A continuación se formulan resultados que proporcionan condiciones suficientes o condiciones necesarias y suficientes para que una secuencia de variables aleatorias $\{X_n\}_{n\in\mathbb{N}}$ satisfaga la anteriores leyes respecto a secuencias $\{a_n\}_{n\in\mathbb{N}}$ y $\{b_n\}_{n\in\mathbb{N}}$ apropiadas.

Teorema 1 Ley débil de Bernoulli. *Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e idénticamente distribuidas según una Bernoulli con parámetro p, i.e., $X_n \sim \mathcal{B}(1,p)$, para cualquier $n \in \mathbb{N}$, entonces*

$$\frac{S_n - E[S_n]}{n} \to^P 0, \quad \frac{S_n}{n} \to^P p.$$

Teorema 2 Ley débil de Khintchine. *Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e idénticamente distribuidas y existe $E[X_n] = \mu$, para cualquier $n \in \mathbb{N}$, entonces*

$$\frac{S_n - E[S_n]}{n} \to^P 0, \quad \frac{S_n}{n} \to^P \mu.$$

Demostración. La demostración se derivará para el caso particular en el que existe $E[X_n^2]$, para cualquier $n \in \mathbb{N}$. A partir de la desigualdad de Tchebytschev, para cualquier $\varepsilon > 0$, teniendo en cuenta que X_i, $i = 1, \ldots, n$, son independientes y con la misma distribución, y por tanto,

$$\mathrm{Var}\,(S_n) = \mathrm{Var}\,(S_n - E[S_n]) = \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = n\mathrm{Var}(X_1),$$

$$\begin{aligned} P\left(\left|\frac{S_n - E[S_n]}{n}\right| \geq \varepsilon\right) &\leq \frac{\mathrm{Var}\,(S_n - E[S_n])}{\varepsilon^2 n^2} \\ &= \frac{n\mathrm{Var}(X_1)}{n^2\varepsilon^2} \to 0, \quad n \to \infty. \end{aligned}$$

En particular, dado que $E[S_n] = n\mu$, se obtiene $S_n/n \to^P \mu$, $n \to \infty$. El Teorema 1, se obtiene entonces considerando $\mu = p$.

Lema 1 Lemma de Borel-Cantelli. *Se considera una sucesión de eventos $\{A_n\}_{n\in\mathbb{N}}$, y se define el suceso:*

$$A_\infty = \bigcap_{k\in\mathbb{N}} \bigcup_{n\geq k} A_n.$$

Se tiene entonces:

(i) Si

$$\sum_{n=1}^{\infty} P(A_n) < \infty \;\Rightarrow\; P(A_\infty) = 0.$$

(ii) Si los sucesos A_n, $n \in \mathbb{N}$, son independientes y

$$\sum_{n=1}^{\infty} P(A_n) = \infty \;\Rightarrow\; P(A_\infty) = 1.$$

Lema 2 Criterio de Integrabilidad. *Sea $X : (\Omega, \mathcal{A}, P) \longrightarrow \mathbb{R}$, entonces X es absolutamente integrable, i.e.,*

$$E[|X|] < \infty \;\Leftrightarrow\; \sum_{n=1}^{\infty} P[|X| > n] < \infty.$$

Lema 3 *Sea X v.a. con $E[X] < \infty$, y sea*

$$A_n = \{\omega \in \Omega; |X(\omega)| \leq n\}.$$

Entonces,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} E[X^2 I_{A_n}] < \infty.$$

Lema 4 *Bajo las condiciones del resultado anterior, se considera una secuencia $\{X_n\}_{n\in\mathbb{N}}$ de variables aleatorias independientes e idénticamente distribuidas, absolutamente integrables, con media $\mu = E[X_1] = E[X_n]$, para cualquier $n \in \mathbb{N}$. Se define la secuencia de variables aleatorias truncadas*

$$Y_n = \begin{cases} X_n & si\ |X_n| \leq n \\ 0 & si\ |X_n| > n. \end{cases}$$

Las variables aleatorias Y_n, $n \in \mathbb{N}$, entonces satisfacen las siguientes propiedades:

$$\lim_{n\to\infty} E[Y_n] = \mu \tag{1}$$

$$\sum_{n=1}^{\infty} \frac{Var(Y_n)}{n^2} < \infty \tag{2}$$

$$\forall \omega \in \Omega \quad \exists n_0(\omega);\; X_n(\omega) = Y_n(\omega), \quad \forall n \geq n_0(\omega). \tag{3}$$

La prueba de (1) se obtiene de aplicar el Teorema de Convergencia Dominada. La prueba de (2) es directa a partir del Lema 3. Finalmente, (3) se deriva aplicando el Lema de Borel Cantelli. Como consecuencia directa del Lemma 4, se obtiene el siguiente corolario:

Corolario 1 *Bajo las condciones del lema anterior, el suceso*

$$B = \left\{ \omega \in \Omega;\ \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} |X_k(\omega) - Y_k(\omega)| = 0 \right\}.$$

se cumple con probabilidad uno, i.e., $P(B) = 1$.

Lema 5 Primera Ley fuerte de Kolmogorov. *Sea* $\{X_n\}_{n\in\mathbb{N}}$ *una secuencia de variables aleatorias independientes con esperanza finita, y supongamos que*

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

Entonces

$$\frac{S_n - \sum_{i=1}^{n} \mu_i}{n} \to^{c.s.} 0, \quad \mu_i = E[X_i], \quad i = 1, \dots, n.$$

Teorema 3 Segunda Ley fuerte de Kolmogorov. *Sea* $\{X_n\}_{n\in\mathbb{N}}$ *una secuencia de variables aleatorias independientes e idénticamente distribuidas (i.i.d.), con* $E[|X_n|] < \infty$. *Se tienen entonces*

$$\frac{S_n - E[S_n]}{n} \to^{c.s} 0.$$

En realidad, $\{X_n\}_{n\in\mathbb{N}}$ *i.i.d.*

$$\frac{S_n - E[S_n]}{n} \to^{c.s} 0 \ \Leftrightarrow\ E[|X_n|] < \infty.$$

Demostración. Se consideran los sucesos:

$$\begin{aligned} C &= \{\omega \in \Omega;\ \frac{S_n(\omega)}{n} \to \mu, \quad n \to \infty\} \\ D &= \{\omega \in \Omega;\ \frac{Y_1(\omega) + \cdots + Y_n(\omega)}{n} \to \overline{\mu}_n = \frac{\sum_{i=1}^{n} \mu_i}{n}, \quad n \to \infty\}, \end{aligned}$$

donde $\mu_i = E[Y_i]$, $i = 1, \dots, n$, e Y_i se define como en el Lemma 4, para $i = 1, \dots, n$. La secuencia $\{Y_n\}_{n\in\mathbb{N}}$ satisface la hipótesis del Lemma 5 (primera ley fuerte de Kolmogorov). Por tanto, $P(D) = 1$. Además, a partir de la ecuación (1) del Lema 4, $\mu_k \to \mu$, $k \to \infty$, por tanto, $\overline{\mu}_n \to \mu$. Como consecuencia, considerando el suceso B definido en el Corolario 1

$$B \cap D \subset C.$$

Dado que

$$P(B) = P(D) = 1 \ \Rightarrow P(C) = 1$$

con lo que se concluye la prueba.

Corolario 2 Ley Fuerte de Borel. *Sea* $\{X_n\}_{n\in\mathbb{N}}$ *una secuencia de variables aleatorias independientes e idénticamente distribuidas según una Bernoulli con parámetro* p, *i.e.,* $X_n \sim \mathcal{B}(1,p)$, *para cualquier* $n \in \mathbb{N}$, *entonces*

$$\frac{S_n - E[S_n]}{n} \to^{c.s.} 0, \quad \frac{S_n}{n} \to^{c.s.} p.$$

Problema clásico del límite central

Teorema 4 Primer teorema límite (Bernoulli). *Sea* $\{X_n\}_{n\in\mathbb{N}}$ *una secuencia de variables aleatorias independientes e idénticamente distribuidas según una Bernoulli con parámetro* p, *i.e.,* $X_n \sim \mathcal{B}(1,p)$, *para cualquier* $n \in \mathbb{N}$, *entonces*

$$\frac{S_n - E[S_n]}{n} \to^{L} 0.$$

Se obtiene de forma directa del Corollario 2, dado que la convergencia casi segura implica la convergencia en Ley.

Teorema 5 Segundo teorema límite (De Moivre y Laplace). *Sea* $\{X_n\}_{n\in\mathbb{N}}$ *una secuencia de variables aleatorias independientes e idénticamente distribuidas según una Bernoulli con parámetro* p, *i.e.,* $X_n \sim \mathcal{B}(1,p)$, *para cualquier* $n \in \mathbb{N}$, *entonces*

$$\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} = \frac{S_n - np}{\sqrt{np(1-p)}} \to^{L} Z \sim \mathcal{N}(0,1).$$

Teorema 6 *Para cada* $n \in \mathbb{N}$, $X_{n1}, \ldots, X_{nn}$ *son variables aleatorias independientes e idénticamente distribuidas según una Bernoulli con parámetro* p_n, *y* $S_{nn} = \sum_{i=1}^{n} X_{ni}$. *Si*

$$\lim_{n\to\infty} np_n = \lambda(>0) \;\Rightarrow\; S_{nn} \to^{L} Y \sim \mathcal{P}(\lambda)$$

Pasamos ahora a formular un resultado sobre el problema clásico de límite central, que consiste en derivar condiciones suficientes que garanticen, que para una secuencia de variables aleatorias independientes, sobre el mismo espacio probabilístico base $(\Omega, \mathcal{A}, P)$, se dan los siguientes límites:

$$\begin{aligned} &\frac{S_n - E[S_n]}{n} \to^{L} 0, \quad \exists E[X_n],\ \forall n \in \mathbb{N} \\ &\frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}(S_n)}} \to^{L} Z \sim \mathcal{N}(0,1), \quad \exists E[X_n^2], \quad \forall n \in \mathbb{N}. \end{aligned} \tag{4}$$

Se considera el siguiente lema que se aplicará en la demostración del Teorema límite de Lévy que posteriormente se enuncia:

Lema 6 *Sea* $\{c_n\}_{n\in\mathbb{N}}$ *una sucesión de números complejos tal que* $c_n \to c$, $n \to \infty$. *Entonces*

$$\left(1 + \frac{c_n}{n}\right)^n \to \exp(c), \quad n \to \infty.$$

Teorema 7 Teorema límite de Lévy. *Sea* $\{X_n\}_{n\in\mathbb{N}}$ *una secuencia de variables aleatorias independientes e idénticamente distribuidas. Se obtienen los siguientes límites en ley:*

(i) Si $\exists E[X_1] \Rightarrow \exists E[X_n]$, $\forall n$, *se tiene* $\frac{S_n - E[S_n]}{n} \to^L 0$.

(ii) Si $\exists E[X_1^2] \Rightarrow \exists E[X_n^2]$, $\forall n$, *se tiene* $\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \to^L Z \sim \mathcal{N}(0,1)$.

Demostración. La afirmación (i) se obtiene como consecuencia directa del Teorema 2 o bien, del Teorema 3.

La afirmación (ii) se demostrará para el caso particular, donde existe la función generatriz de momentos de las variables aleatorias de la sucesión $\{X_n\}_{n\in\mathbb{N}}$, definidas, todas ellas, sobre un intervalo común. Se considera la secuencia

$$S_n^* = \frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}(S_n)}} = \frac{S_n - n\mu}{\sqrt{n}\sigma},$$

donde $\sigma = \sqrt{\mathrm{Var}(X_n)}$, para cualquier $n \in \mathbb{N}$. Se asume que $\mu = E[X_n] = 0$, para cualquier $n \in \mathbb{N}$. En caso contrario, se considera $\widetilde{X}_n = X_n - \mu$, $n \in \mathbb{N}$.

Dado que X_n, $n \in \mathbb{N}$, son independientes e idénticamente distribuidas, con $\mu = 0$, la función generatriz de momentos

$$M_{S_n^*}(t) = \left[M_{X_1}\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n,$$

donde $M_{X_1} = M_{X_n} = M_X$, para todo $n \in \mathbb{N}$. Se considera ahora el desarrollo de Taylor de segundo orden de $M_X = E[\exp(tX)]$, $t \in (-a, b)$, $a, b \in \mathbb{R}_+$, que relaciona las derivadas en cero de la función generatriz de momentos con los momentos de X. Es decir,

$$\begin{aligned} M_X(t) &= 1 + \frac{dM_X(0)}{dt}t + \frac{1}{2}\frac{d^2M_X(0)}{dt^2}t^2 + t^2 e_2(t) \\ &= 1 + \frac{\sigma^2}{2}t^2 + t^2 e_2(t) = 1 + t^2\left[\frac{\sigma^2}{2} + e_2(t)\right], \quad \lim_{t\to 0} e_2(t) = 0. \end{aligned} \tag{5}$$

Se considera ahora la correspondiente aproximación para $M_{S_n^*}(t)$. Más concretamente,

$$
\begin{aligned}
M_{S_n^*}(t) &= \left(1+\left(\frac{t}{\sqrt{n}\sigma}\right)^2\left[\frac{\sigma^2}{2}+e_2\left(\frac{t}{\sqrt{n}\sigma}\right)\right]\right)^n \\
&= \left(1+\frac{t^2}{n}\left[\frac{1}{2}+\frac{e_2\left(\frac{t}{\sqrt{n}\sigma}\right)}{\sigma^2}\right]\right) \qquad (6)
\end{aligned}
$$

Además,

$$t^2\left[\frac{1}{2}+\frac{e_2\left(\frac{t}{\sqrt{n}\sigma}\right)}{\sigma^2}\right] \to \frac{t^2}{2}, \quad n\to\infty.$$

Aplicando el Lemma 6, fijado un t, siendo

$$c_n = t^2\left[\frac{1}{2}+\frac{e_2\left(\frac{t}{\sqrt{n}\sigma}\right)}{\sigma^2}\right], \quad c=\frac{t^2}{2}$$

se obtiene

$$M_{S_n^*}(t)\to \exp\left(\frac{t^2}{2}\right), \quad n\to\infty,$$

que corresponde a la función generatriz de momentos de una variable aleatoria normal con media cero y varianza uno. Por el teorema de continuidad para funciones generatrices de momentos se obtiene que

$$S_n^* \to^L Z \sim \mathcal{N}(0,1).$$

Nota 2 *El nombre del* teorema del límite central *se debe a que proporciona una buena aproximación en el centro de la distribución, pero no tan buena en las colas de la misma.*