PROBABILIDAD II

Grado en Matemáticas

Tema 6
Modos de convergencia
de variables aleatorias

Javier Cárcamo

Departamento de Matemáticas
Universidad Autónoma de Madrid
javier.carcamo@uam.es

Tema 6: Modos de convergencia de variables aleatorias

Los resultados más célebres e importantes de la Teoría de la Probabilidad son los conocidos como **leyes de los grandes números**. Tales leyes no son otra cosa que teoremas que aseguran cierta convergencia de variables aleatorias bajo unas condiciones determinadas. Ahora bien, hay varios modos de convergencia de variables aleatorias

En este tema se analizan esos modos de convergencia y las relaciones que existen entre ellos. Todas las variables aleatorias que se consideran toman valores reales.

Definición: Sean $X, X_1, X_2, \ldots$ variables aleatorias definidas sobre el mismo espacio de probabilidad $(\Omega, \mathcal{F}, \mathrm{P})$. Entonces:

$$X_n \xrightarrow{\text{c.s.}} X \quad \text{si} \quad \mathrm{P}(X_n \to X) = 1.$$

$$X_n \xrightarrow{\text{P}} X \quad \text{si} \quad \forall \epsilon > 0,\ \mathrm{P}(|X_n - X| \geq \epsilon) \to 0.$$

$$X_n \xrightarrow{\text{m-}p} X \quad \text{si} \quad \mathrm{E}|X_n - X|^p \to 0 \qquad (p > 0).$$

Observación: La definición de $X_n \xrightarrow{\text{c.s.}} X$ tiene sentido.

Definición: Sean $X, X_1, X_2, \ldots$ variables aleatorias *no necesariamente definidas sobre el mismo espacio de probabilidad* y sean F y F_n las funciones de distribución de X y X_n, respectivamente. Entonces:

$$X_n \xrightarrow{\mathcal{D}} X \quad \text{si} \quad \forall x \in C_F,\ F_n(x) \to F(x),$$

donde C_F es el conjunto de puntos de continuidad de F.

Pregunta: ¿Por qué no $\forall x \in \mathbb{R}$ en la definición anterior?

Algunas observaciones

Observación 1: Si indicamos por $\to$ uno cualquiera de los tres primeros modos de convergencia, es claro que $X_n \to X$ es equivalente a $X_n - X \to 0$.

Observación 2: Si $X \equiv a$ (constante), las definiciones de $X_n \xrightarrow{\mathrm{P}} a$ y $X_n \xrightarrow{\text{m-}p} a$ tienen sentido también cuando las variables X_i *no* están definidas sobre el mismo espacio. En tal caso, es conveniente escribir P_n y E_n en lugar de P y E, siendo P_n la medida de probabilidad del espacio en el que está definida X_n (y E_n el operador esperanza correspondiente).

Observación 3: Cuando todas las variables involucradas son degeneradas, es decir constantes c.s., cada una de las cuatro convergencias coincide con la convergencia usual de sucesiones de números reales.

Observación 4: En cada modo de convergencia el límite es 'único' en un sentido preciso.

Esquema de las relaciones mutuas

Nota: Las convergencias casi segura y en media-p se llaman **modos de convergencia fuertes** (sobre todo la convergencia c.s.). Las convergencias en probabilidad y en distribución se denominan **modos de convergencia débiles** (sobre todo la convergencia en distribución).

Esquema de las relaciones mutuas

$$\begin{tikzcd}
& X_n \xrightarrow{\text{c.s.}} X \arrow[dr, Rightarrow] & & \\
& & X_n \xrightarrow{\mathrm{P}} X \arrow[r, Rightarrow] & X_n \xrightarrow{\mathcal{D}} X \\
X_n \xrightarrow{\text{m-}q} X\,(p<q) \arrow[r, Rightarrow] & X_n \xrightarrow{\text{m-}p} X \arrow[ur, Rightarrow] & &
\end{tikzcd}$$

Nota: Los recíprocos de las anteriores implicaciones no son ciertos en general.

Convergencia c.s. versus convergencia en probabilidad

Lema: Son equivalentes:

(a) $X_n \xrightarrow{\text{c.s.}} X$.

(b) Para todo $\epsilon > 0$, $\mathrm{P}\Big(\limsup_{n\to\infty}\{|X_n - X| \geq \epsilon\}\Big) = 0$.

(c) Para todo $\epsilon > 0$, $\lim_{n\to\infty} \mathrm{P}\Big(\sup_{k\geq n} |X_k - X| \geq \epsilon\Big) = 0$.

Corolario: Se tiene:

$$\sum_{n\geq 1} \mathrm{P}(|X_n - X| \geq \epsilon) < \infty, \quad \epsilon > 0 \quad \Longrightarrow \quad X_n \xrightarrow{\text{c.s.}} X.$$

Definición: Cuando la serie anterior converge, se dice que hay **convergencia completa** de X_n a X.

Pregunta: ¿Bajo qué condiciones es la convergencia completa equivalente a la convergencia c.s.?

Convergencia c.s. versus convergencia en probabilidad

Teorema: Si $X_n \xrightarrow{\text{c.s.}} X$, entonces $X_n \xrightarrow{\text{P}} X$.

Observación:

$$X_n \xrightarrow{\text{P}} X \nRightarrow X_n \xrightarrow{\text{c.s.}} X.$$
$$X_n \xrightarrow{\text{m-}p} X \nRightarrow X_n \xrightarrow{\text{c.s.}} X.$$

Teorema: Si $X_n \xrightarrow{\text{P}} X$, entonces $X_{n_k} \xrightarrow{\text{c.s.}} X$, para alguna subsucesión (X_{n_k}).

Corolario: Son equivalentes:

(a) $X_n \xrightarrow{\text{P}} X$.

(b) Toda subsucesión de X_n tiene a su vez una subsucesión que converge a X c.s.

Convergencia en media-p versus convergencia en media-q Convergencia en media-p vs convergencia en probabilidad

Teorema: Si $X_n \xrightarrow{\text{m-}q} X$, entonces $X_n \xrightarrow{\text{m-}p} X$, para $0 < p < q$.

Observación: $X_n \xrightarrow{\text{m-}p} X \not\Rightarrow X_n \xrightarrow{\text{m-}q} X$, para $0 < p < q$.

Teorema: Si $X_n \xrightarrow{\text{m-}p} X$, entonces $X_n \xrightarrow{\text{P}} X$.

Observación:

$$X_n \xrightarrow{\text{P}} X \not\Rightarrow X_n \xrightarrow{\text{m-}p} X.$$
$$X_n \xrightarrow{\text{c.s.}} X \not\Rightarrow X_n \xrightarrow{\text{m-}p} X.$$

Algunos recíprocos parciales

Teorema: Si $|X_n - X| \leq K$ (constante) $n \geq 1$ y $X_n \xrightarrow{\text{P}} X$, entonces $X_n \xrightarrow{\text{m-}p} X$.

Teorema: Si $X_n \xrightarrow{\text{c.s.}} X$ y $|X_n| \leq U$ v.a. integrable $(n \geq 1)$, entonces $X_n \xrightarrow{\text{m-}1} X$.

Convergencia en probabilidad versus convergencia en distribución

Teorema: Si $X_n \xrightarrow{\text{P}} X$, entonces $X_n \xrightarrow{\mathcal{D}} X$.

Demostración: Para $x \in C_F$, bastará demostrar que:

(1) $\limsup F_n(x) \leq F(x)$.

(2) $F(x) \leq \liminf F_n(x)$.

Para ver (1), comprobar que para $\epsilon > 0$, se tiene que

$$\{X_n \leq x\} \subset \{X \leq x + \epsilon\} \cup \{|X_n - X| \geq \epsilon\}.$$

Para mostrar (2), comprobar que

$$\{X \leq x - \epsilon\} \subset \{X_n \leq x\} \cup \{|X_n - X| \geq \epsilon\}.$$

Observación: $X_n \xrightarrow{\mathcal{D}} X \not\Rightarrow X_n \xrightarrow{\text{P}} X$.

Teorema: Si $X_n \xrightarrow{\mathcal{D}} a$ (constante), entonces $X_n \xrightarrow{\text{P}} a$.

Leyes fuertes y débiles de los grandes números

Las llamadas **leyes de los grandes números** son teoremas de convergencia que tienen la forma siguiente:

Sea (X_n) una sucesión de v.a.que cumplen las hipótesis $H_1, H_2, \ldots$. Entonces se tiene

$$\frac{S_n - a_n}{b_n} \xrightarrow{?} 0.$$

- $S_n = \sum_{i=1}^{n} X_i$.
- $\xrightarrow{?}$ indica uno de los modos de convergencia anteriores.
- Las sucesiones de constantes (a_n) y (b_n) se llaman **constantes de normalización**.
- La ley se llama **fuerte** cuando la convergencia es casi segura, y **débil** cuando la convergencia es en probabilidad o en distribución. (Toda ley fuerte implica la correspondiente ley débil.)

Leyes fuertes y débiles de grandes números

Sea (X_n) una sucesión de v.a.que cumplen las hipótesis $H_1, H_2, \ldots$. Entonces se tiene

$$\frac{S_n - a_n}{b_n} \xrightarrow{?} 0.$$

Las hipótesis H_i, suelen ser de varios tipos:

(a) Sobre los momentos. Es decir si las variables son integrables o de cuadrado integrable, etc. Cuanto menor sea la exigencia de integrabilidad mayor generalidad tendrá el teorema (o la ley).

(b) Sobre la dependencia de las variables X_i. Es decir, si se trata de variables mutuamente independientes, o basta con que sean independientes dos a dos, o incorrelacionadas, etc.

(c) Sobre la distribución de probabilidad de las X_i. Es decir si las variables tienen o no la misma distribución, etc.

(d) Sobre los requisitos que tienen que cumplir las sucesiones de constantes de normalización.

La ley débil de Bernoulli

Históricamente, la primera ley de grandes números fue obtenida por Bernoulli en el siglo XVIII. La ley se refiere a la proporción de caras que se obtienen en n lanzamientos de una p-moneda (moneda que cae de cara con probabilidad p) y al modo en que esa proporción se aproxima a p cuando el número de lanzamientos es grande. En términos más o menos coloquiales dice lo siguiente:

Haciendo un número suficientemente grande de lanzamientos, podemos conseguir que la probabilidad de que dicha proporción se diferencie de p en menos de una cantidad prefijada esté tan próxima a 1 como queramos.

Teorema: Ley débil de Bernoulli

Si $X_1, X_2, \ldots$ son independientes y tienen la misma distribución de Bernoulli de parámetro p, entonces

$$\frac{S_n}{n} \xrightarrow{\mathrm{P}} p.$$

La ley débil de Bernoulli: Extensiones

Ejercicio: Sean $X_1, X_2, \ldots$ variables aleatorias tales que $\mathrm{Var}(X_n) \leq K$ $(n \geq 1)$ y $\mathrm{Cov}(X_i, X_j) \leq 0$, para $i \neq j$. Entonces, si $a > 1/2$, se tiene

$$\frac{S_n - \mathrm{E}S_n}{n^a} \xrightarrow{\text{m-2}} 0.$$

Ejercicio: En las condiciones del ejercicio anterior, si las X_i tienen la misma media μ, se tiene

$$\frac{S_n}{n} \xrightarrow{\text{m-2}} \mu.$$

Pregunta: Estos resultados se pueden generalizar/extender de muchas formas, ¿se te ocurre alguna?