Convergencia en distribución

Pablo L. De Nápoli

Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Probabilidades y Estadística para Matemática
Segundo cuatrimestre de 2021

Parte I

Repasamos las definición de la convergencia en distribución

Propiedades de las funciones de distibución (clase 4)

Teorema

Sea $X : \Omega \to \overline{\mathbb{R}}$ *una variable aleatoria y* $F = F_X$ *su función de distribución acumulada*

$$F_X(x) = P\{X \leq x\}$$

Entonces F *tiene las siguientes propiedades:*

i) $0 \leq F(x) \leq 1$ *y* F *es creciente.*

ii) F *es continua por la derecha.*

iii) $F(x_0) - \lim_{x \to x_0^-} F(x) = P\{X = x_0\}$ *En particular,* F *es continua en* $x = x_0$ *si y sólo si* $P\{X = x_0\} = 0$*.*

iv) *Si* X *es finita con probabilidad 1 (o sea* $P\{X = \pm\infty\} = 0$*), entonces:*

$$\lim_{x \to -\infty} F(x) = 0$$

$$\lim_{x \to +\infty} F(x) = 1$$

Convergencia en distribución

Hoy nos vamos a dedicar a estudiar la noción de convergencia en distribuciòn.

Definición

Sea $(X_n)_{n\in N} : \Omega_n \to \overline{\mathbb{R}}$ *una sucesión de variables aleatorias, definidas sobre espacios de probabilidad* $(\Omega_n, \mathcal{E}_n, P_n)$ *y finitas con probabilidad 1. Consideramos sus funciones de distribución acumulada* F_{X_n}, F_X*. Se dice que* (X_n) *converge en distribución a la variable* X *definida sobre un espacio de probabilidad* $(\Omega, \mathcal{E}, P)$ *si*

$$F_{X_n}(x) \to F_X(x)$$

para todo x *que sea un punto de continuidad de* F_X*.* **Notación:**

$$X_n \xrightarrow{D} X$$

Esta noción es más bien una noción de convegencia de distribuciones de probabilidad que de variables aleatorias. Por eso, también tiene sentido aunque las variables estuvieran definidas en distintos espacios de probabilidad.

Un ejemplo convergencia de en distribución que ya vimos

Distribuciones normales que se concentran

Si $X_n \sim N(0, \sigma_n^2)$ donde $\sigma_n \to 0$ y $X = 0$ con probabilidad 1. Entonces $X_n \xrightarrow{D} X$.

Aqui, como vimos en la clase 9,

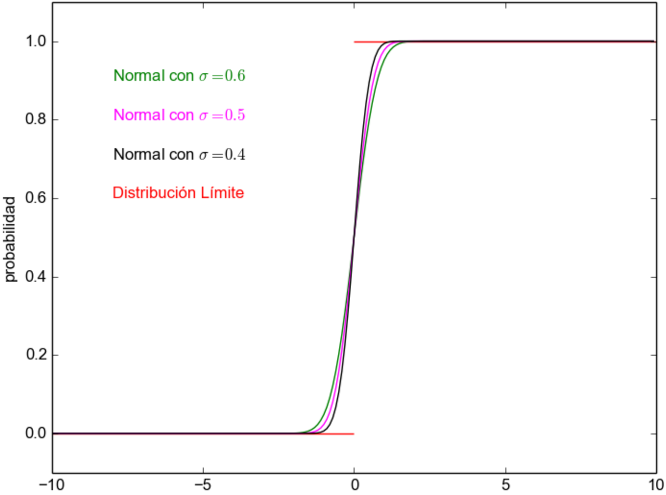
$$F_{X_n}(x) = \Phi\left(\frac{x}{\sigma_n}\right) \quad \text{donde } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

es la función de distribución de una normal estándar, mientras que F_X es el escalón (función de Heaviside). Vemos que

$$F_{X_n}(x) \to \begin{cases} 0 & \text{si} \quad x < 0 \\ 1 & \text{si} \quad x > 0 \\ 1/2 & \text{si} \quad x = 0 \end{cases}$$

En consecuencia, $F_{X_n}(x) \to F_X(x)$ para todo $x \neq 0$. Luego $X_n \xrightarrow{D} X$. Veremos más adelante, que converger en probabilidad implica converger en distribución.

Un ejemplo convergencia de en distribución (2)



El Teorema de De Moivre-Laplace (clase 7)

Teorema (De Moivre-Laplace)

Sea S_n una sucesión de variables aleatorias con distribución binomial $Bi(n, p)$. Consideramos las variables estandarizadas

$$S_n^* = \frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}[S_n]}} = \frac{S_n - np}{\sqrt{npq}}$$

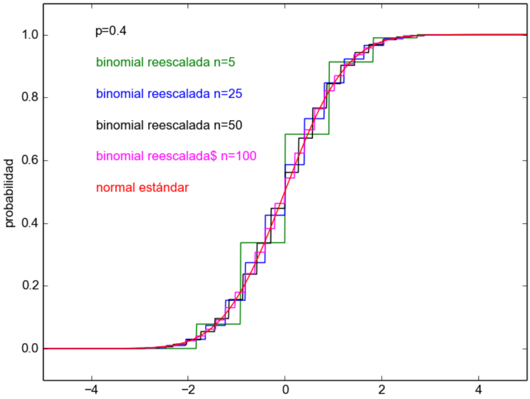
Entonces

$$P\{a < S_n^* \leq b\} \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, dx = F_N(b) - F_N(a)$$

uniformemente en a y en b cuando $n \to +\infty$, donde F_N es la función de distribución de una variable aleatoria con distribución normal estándar. En particular,

$$S_n^* \xrightarrow{D} N$$

Ilustración del teorema de De Moivre-Laplace



Para $p = 0{,}4$ y distintos valores de n, dibujamos la función de distribución de la distribución binomial, junto con la de la normal estándar.

Parte II

Relación con la convergencia en probabilidad

La convergencia en probabilidad implica la convergencia en distribución

Recordamos la definición

Definición

Sea $(X_n)_{n \in N} : \Omega \to \overline{\mathbb{R}}$ *una sucesión de variables aleatorias, definidas sobre un mismo espacio de probabilidad* $(\Omega, \mathcal{E}, P)$ *y finitas con probabilidad 1. Se dice que* (X_n) *converge en probabilidad a la variable* X *si para todo* $\delta > 0$*, tenemos que*

$$P\{|X - X_n| > \delta\} \to 0 \text{ cuando } n \to +\infty$$

Notación:

$$X_n \xrightarrow{P} X$$

Proposición

Sea (X_n) *es una sucesión de variables aleatorias definidas sobre un mismo espacio de probabilidad* $(\Omega, \mathcal{E}, P)$ *y finita son probabilidad 1. Si* $X_n \xrightarrow{P} X$*, entonces* $X_n \xrightarrow{D} X$.

Demostración (1)

Sea $x_0 \in \mathbb{R}$ tal que F_X sea continua en x_0. Entonces

$$X - |X - X_n| \leq X_n$$

Si $X > x_0 + \varepsilon$ y $|X_n - X| < \varepsilon \Rightarrow X_n > x_0$. Lo podemos traducir en una inclusión de conjuntos:

$$\{X > x_0 + \varepsilon\} \cap \{|X_n - X| < \varepsilon\} \subset \{X_n > x_0\}$$

Tomamos complemento. La inclusión se da vuelta, y usamos las leyes de De Morgan.

$$\{X_n \leq x_0\} \subset \{X \leq x_0 + \varepsilon\} \cup \{|X_n - X| \geq \varepsilon\}$$

Tomamos probabilidad. Usamos que es creciente y subaditiva:

$$P\{X_n \leq x_0\} \leq P\{X \leq x_0 + \varepsilon\} + P\{|X_n - X| \geq \varepsilon\}$$

Esto establece la desigualdad:

$$F_{X_n}(x_0) \leq F_X(x_0 + \varepsilon) + P\{|X_n - X| \geq \varepsilon\}$$

Demostración (2)

Similarmente

$$X_n \leq X + |X_n - X|$$

Si $X \leq x_0 - \varepsilon$ y $|X_n - X| \leq \varepsilon \Rightarrow X_n \leq x_0$. Lo podemos traducir en una inclusión de conjuntos:

$$\{X \leq x_0 - \varepsilon\} \cap \{|X_n - X| \leq \varepsilon\} \subset \{X_n \leq x_0\}$$

Tomamos complemento. La inclusión se da vuelta, y usamos las leyes de De Morgan.

$$\{X_n > x_0\} \subset \{X > x_0 - \varepsilon\} \cup \{|X_n - X| > \varepsilon\}$$

Tomamos probabilidad. Usamos que es creciente y subaditiva:

$$P\{X_n > x_0\} \leq P\{X > x_0 - \varepsilon\} + P\{|X_n - X| > \varepsilon\}$$

Esto establece la desigualdad:

$$1 - F_{X_n}(x_0) \leq 1 - F_X(x_0 - \varepsilon) + P\{|X_n - X| > \varepsilon\}$$

o

$$F_X(x_0 - \varepsilon) - P\{|X_n - X| > \varepsilon\} \leq F_{X_n}(x_0)$$

Entonces juntando todo tenemos que

$$F_X(x_0-\varepsilon) - P\{|X_n - X| > \varepsilon\} \leq F_{X_n}(x_0) \leq F_X(x_0+\varepsilon) + P\{|X_n - X| \geq \varepsilon\}$$

Entonces como $X_n \xrightarrow{P} X$ por hipótesis,

$$F_X(x_0-\varepsilon) \leq \liminf_{n\to+\infty} F_{X_n}(x_0) \leq \limsup_{n\to+\infty} F_{X_n}(x_0) \leq F_X(x_0+\varepsilon)$$

Y cuanndo $\varepsilon \to 0$, como F_X es continua en x_0,

$$\liminf_{n\to+\infty} F_{X_n}(x_0) = \limsup_{n\to+\infty} F_{X_n}(x_0) = F_X(x_0)$$

Es decir que

$$\lim_{n\to+\infty} F_{X_n}(x_0) = F_X(x_0)$$

como queríamos probar.

Parte III

Convergencia de la esperanza: El Teorema de Helly-Bray

Planteo del problema

Nos gustaría analizar cuando es válido que

$$X_n \to X \Rightarrow E[X_n] \to E[X]$$

suponiendo que las X_n y X tengan esperanza finita, para alguna de las nociones de convergencia que estudiamos.

¡Pero hay que tener cuidado! Esto no es cierto en general, incluso para la convergencia casi segura (la más fuerte de las nociones de convegencia que usamos en la toería de probabilidades).

Un Ejemplo

Consideramos el espacio muestral $\Omega = [0,1]$ con la probabilidad dada por la medida de Lebesgue, y miremos as funciones $X_n : [0,1] \to \mathbb{R}$,

$$X_n(x) = n \cdot x^{n-1}.$$

Recordamos que usando este espacio muestral, estas funciones se pueden ver como variables aleatorias. Entonces

$$X_n \xrightarrow{c.s.} 0$$

ya que

$$X_n(x) \to 0 \text{ para todo } x \neq 1$$

y $m(\{1\}) = 0$. Pero

$$E[X_n] = \int_0^1 nx^{n-1}\, dx = x^n\big|_0^1 = 1$$

En general, la integral del límite puede no ser igual al límite de las integrales.

El Teorema de Helly-Bray (en intervalos finitos)

Teorema (Helly-Bray (en intervalos finitos))

Supongamos que $F_n : [a,b] \to \mathbb{R}$ *es una sucesión de funciones de distribución tales que* $F_n(x) \to F(x)$ *en cada punto de continuidad de* $F(x)$*, entonces:*

$$\int_a^b \varphi(x)\, dF_n(x) \to \int_a^b \varphi(x)\, dF(x) \tag{1}$$

para toda función continua $\varphi \in C[a,b]$.

Corolario (Reformulación en el lenguaje probabilítico)

Si $X_n \xrightarrow{D} X$ *y* X_n *son variables aleatorias concentradas en un intervalo* $[a,b]$*, entonces*

$$E[\varphi(X_n)] \to E[\varphi(X)]$$

para toda función continua $\varphi \in C[a,b]$.

Repasamos: Particiones y sumas de Riemann-Stieltjes

Sea

$$\pi : a = x_0 < x_1 < x_2 < \ldots < x_n = b$$

una partición del intervalo $(a, b]$ (Dar una partición no es otra cosa que elegir finitos puntos del intervalo en orden creciente) y elijamos puntos intermedios $\xi_i \in (x_i, x_{i+1}]$ en cada intervalito de la partición (En realidad, estamos trabajando con particiones con puntos marcados, pero no lo haremos explícito en la notación). Consideramos entonces las sumas de Riemann-Stieltjes

$$S_\pi(\varphi, F) = \sum_{i=0}^{n-1} \varphi(\xi_i) \cdot (F(x_{i+1}) - F(x_i))$$

Repasamos: Definición de la integral de Riemann-Stieltjes

Definición

Diremos que la integral

$$\int_a^b \varphi(x)\, dF(x) = \int_{(a,b]} \varphi(x)\, dF(x)$$

existe y toma el valor $I \in \mathbb{R}$ si las sumas $S_\pi(\varphi, F)$ tienden al valor I cuando la norma

$$|\pi| = \max_{0 \leq i \leq n-1} |x_{i+1} - x_i|$$

de la partición π tiende a cero, es decir si dado $\varepsilon > 0$, existe $\delta > 0$ tal que $|I - S_\pi(\varphi, F)| < \varepsilon$ para toda partición π con $|\pi| < \delta$.

Recordamos que la integral siempre existe si F es creciente y $\varphi \in C[a, b]$.

Convergencia uniforme de las integrales de Stieltjes

Para poder probar este teorema, necesitamos un resultado sobre convergencia uniforme de las integrales de Stieltjes que mencionamos en la clase 8 (Pueden ver la demostración en uno de los apéndices de mis notas).

Teorema

Sea $\varphi \in C[a,b]$. Dados $\varepsilon > 0$ y $C > 0$, existe un $\delta > 0$ (que depende de $\varepsilon > 0$ y C pero es independiente de F) tal que si F es cualquier función $F : [a,b] \to \mathbb{R}$ creciente tal que

$$m_F((a,b]) = F(b) - F(a) \leq C$$

entonces

$$\left| \int_a^b \varphi(x)\, dF(x) - S_\pi(\varphi, F) \right| < \varepsilon$$

para toda partición con puntos marcados con $|\pi| < \delta$.

Demostración del teorema de Helly-Bray en intervalos finitos

Dado $\varepsilon > 0$, por el resultado anterior, existirá un $\delta > 0$ tal que:

$$\left| \int_a^b \varphi(x)\, dF_n(x) - S_\pi(\varphi, F_n) \right| < \varepsilon$$

para todo n, y también

$$\left| \int_a^b \varphi(x)\, dF(x) - S_\pi(\varphi, F) \right| < \varepsilon$$

para cualquier partición π de $[a, b]$ que verifique que $|\pi| < \delta$ (Pues $F_n(1) - F_n(0) \leq 1$).

Fijemos una partición cualquiera π de $[a, b]$ tal que $|\pi| < \delta$. Claramente podemos elegir los puntos de subdivisión de esta partición π para que sean puntos de continuidad de F (pues el conjunto de puntos de discontinuidad de F es a lo sumo numerable, y por lo tanto su conjunto de puntos de continuidad es denso en $[a, b]$).

Demostración del teorema de Helly-Bray en intervalos finitos (2)

Entonces notamos que como hay finitos puntos en la partición, claramente tendremos que:

$$\lim_{n \to +\infty} S_\pi(\varphi, F_n) = S_\pi(\varphi, F)$$

Es decir, que dado $\varepsilon > 0$, existirá un n_0, tal que si $n \geq n_0$,

$$|S_\pi(\varphi, F_n) - S_\pi(\varphi, F)| < \varepsilon$$

En consecuencia, si $n \geq n_0$,

$$\left| \int_a^b \varphi(x)\, dF_n(x) - \int_a^b \varphi(x)\, dF(x) \right| \leq \left| \int_a^b \varphi(x)\, dF_n(x) - S_\pi(\varphi, F_n) \right|$$
$$+ |S_\pi(\varphi, F_n) - S_\pi(\varphi, F)| + \left| S_\pi(\varphi, F) - \int_a^b \varphi(x)\, dF(x) \right| < 3\varepsilon$$

Como $\varepsilon > 0$ es arbitrario, esto prueba el teorema.

El Teorema de Helly-Bray

Un resultado análogo se verifica para integrales en intervalos infinitos:

Teorema

Supongamos que $F_n : \mathbb{R} \to [0,1]$ es una sucesión de funciones de distribución tales que $F_n(x) \to F(x)$ en cada punto de continuidad de $F(x)$, entonces:

$$\int_{-\infty}^{\infty} \varphi(x)\, dF_n(x) \to \int_{-\infty}^{\infty} \varphi(x)\, dF(x) \tag{2}$$

para toda función continua acotada $\varphi : \mathbb{R} \to \mathbb{R}$.

Corolario (Reformulación en el lenguaje probabilítico)

Si $X_n \xrightarrow{D} X$, entonces

$$E[\varphi(X_n)] \to E[\varphi(X)]$$

para toda función continua $\varphi : \mathbb{R} \to \mathbb{R}$ continua y acotada.

Demostración

Supongamos que $|\varphi(x)| \leq M \;\forall\, x \in \mathbb{R}$. Dado $\varepsilon > 0$, podemos elegir $R > 0$ tal que:

$$1 - F(R) + F(-R) = \int_{x \leq -R \vee X > R} dF(x) < \frac{\varepsilon}{M}$$

y por lo tanto

$$\left| \int_{x \leq -R \vee X > R} \varphi(x)\, dF_n(x) \right| < 2\varepsilon$$

Además, podemos suponer que R y $-R$ son puntos de continuidad de F. Entonces, como $F_n(R) \to F(R)$ y $F_n(-R) \to F(-R)$ cuando $n \to +\infty$, podemos elegir n_1 tal que para $n \geq n_1$ se verifique

$$1 - F_n(R) + F_n(-R) = \int_{x \leq -R \vee X > R} dF_n(x) < \frac{2\varepsilon}{M}$$

y por lo tanto:

$$\left| \int_{|x| > R} \varphi(x)\, dF_n(x) \right| < 2\varepsilon$$

Demostración (2)

Entonces,en virtud del teorema anterior, podemos elegir un n_2 tal que si $n \geq n_2$ se verifica:

$$\left| \int_{-R}^{R} \varphi(x)\, dF_n(x) - \int_{-R}^{R} \varphi(x)\, dF(x) \right| < \varepsilon$$

Entonces, tendremos que:

$$\begin{aligned}
\left| \int_{-\infty}^{\infty} \varphi(x)\, dF_n(x) - \int_{-\infty}^{\infty} \varphi(x)\, dF(x) \right| &\leq \left| \int_{-\infty}^{\infty} \varphi(x)\, dF_n(x) - \int_{-R}^{R} \varphi(x)\, dF(x) \right| \\
&+ \left| \int_{-R}^{R} \varphi(x)\, dF_n(x) - \int_{-R}^{R} \varphi(x)\, dF(x) \right| \\
&+ \left| \int_{-R}^{R} \varphi(x)\, dF(x) - \int_{-\infty}^{\infty} \varphi(x)\, dF(x) \right| \\
&< 4\varepsilon
\end{aligned}$$

Como $\varepsilon > 0$ es arbitrario, esto prueba el teorema.

Parte IV

Un disgresión técnica: Funciones de prueba

Funciones de Prueba

Para el siguiente teorema, vamos a usar el espacio de funciones de prueba

$$\mathcal{D} = \mathcal{C}_c^\infty(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : f \text{ es } C^\infty \text{ y tiene soporte compacto}\}$$

La condición de que f es C^∞ dice que todas las derivadas $f^{(k)}$ de f existen y son continuas en todo $\mathbb{R}$.
La condición de que f tiene soporte compacto, dice que

$$\text{soporte}(f) = \overline{\{x \in \mathbb{R} : f(x) \neq 0\}}$$

es un conjunto compacto de $\mathbb{R}$, o equivalentemente: existe un intervalo $[a, b]$ tal que $f(x) = 0$ si $x \notin [a, b]$.
A primera vista, parece un espacio muy pequeño. Uno podría pensar que $\mathcal{D} = \{0\}$.
¡Sin embargo vamos a ver que esto no es así!

La función de Cauchy

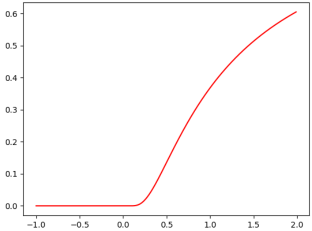
Consideramos $f : \mathbb{R} \to \mathbb{R}$ dada por

$$f(x) = \begin{cases} e^{-1/x} & \text{si} \quad x > 0 \\ 0 & \text{si} \quad x \leq 0 \end{cases}$$

Esta función es C^∞ (no tiene soporte compacto). Notamos que

$$f^{(k)}(0) = 0 \text{ para todo } k$$

por lo que el polinomio de Taylor de f de grado k en el origen es el polinomio nulo para todo k, aunque la función f no es idénticamente nula.

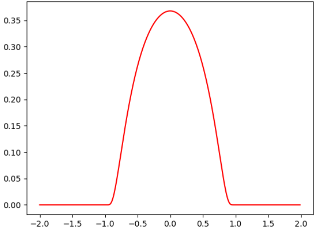


Construyendo una función de prueba

Consideremos ahora: $g : \mathbb{R} \to \mathbb{R}$ dada por

$$g(x) = \begin{cases} e^{-1/(1-x^2)} & \text{si} \quad |x| < 1 \\ 0 & \text{si} \quad |x| \geq 1 \end{cases}$$

Vemos que $g \in \mathcal{D}$ y soporte$(f) = [-1, 1]$. Reescalando, dado un intervalo cualquiera $[a, b]$ podríamos construir $g \in \mathcal{D}$ tal que soporte$(g) = I = [a, b]$



Escalones suaves

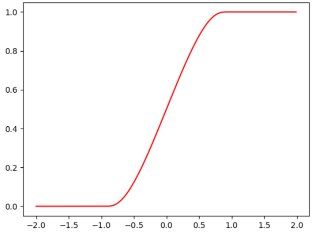
Consideremos ahora la función $h : \mathbb{R} \to \mathbb{R}$ dada por

$$h(x) = \frac{1}{c} \int_{-1}^{x} g(x)dx \text{ para } -1 \leq x \leq 1$$

donde g es la del ejemplo anterior

$$c = \int_{-1}^{1} g(x)dx$$

Podemos $h(x) = 0$ si $x < -1$ y $h(x) = 1$ si $x > 1$. Resulta que h es C^{∞} y $0 \leq h(x) \leq 1$. Notemos que $g'(x) = h(x)$ si $x \in (-1, 1)$ por el teorema fundamental del cálculo.



Escalones suaves (2)

Finalmente, tomando

$$\varphi_\delta(x) = 1 - h\left(\frac{x - x_0}{\delta} - 1\right)$$

podemos probar el siguiente lema que afirma que podemos aproximar la función indicadora $I_{\leq x_0} = I_{(-\infty, x_0]}$ del la semirrecta $(-\infty, x_0]$ por la derecha, por funciones suaves.

Lema

Para cada $x_0 \in \mathbb{R}$ y cada $\delta > 0$, existe φ_δ de clase C^∞ tal que:

- $0 \leq \varphi_\delta(x) \leq 1$.
- $\varphi_\delta(x) = 1$ *si* $x \leq x_0$.
- $\varphi_\delta(x) = 0$ *si* $x \geq x_0 + \delta$.

En particular, $\varphi_\delta(x) \geq I_{\leq x_0}(x)$ *para todo* x.

Similarmente, podemos aproximar $I_{\leq x_0}$ por la izquierda, por funciones suaves, definiendo

$$\varphi_{-\delta}(x) = 1 - h\left(\frac{x - x_0}{\delta} + 1\right)$$

Lema

Para cada $x_0 \in \mathbb{R}$ y cada $\delta > 0$, existe $\varphi_{-\delta}$ de clase C^∞ tal que:

- $0 \leq \varphi_{-\delta}(x) \leq 1$.
- $\varphi_{-\delta}(x) = 1$ *si* $x \leq x_0 - \delta$.
- $\varphi_{\delta}(x) = 0$ *si* $x \geq x_0$.

En particular, $\varphi_{\delta}(x) \leq I_{\leq x_0}(x)$ *para todo* x.

Parte V

Caracterización de la convergencia en distribución: El Recíproco del tereorema de Helly

El Recíproco del tereorema de Helly

Teorema (Recíproco del teorema de Helly)

Si (X_n) es una sucesión de variables aleatorias tales que $E[\varphi(X_n)] \to E[\varphi(X)]$ para toda función $\varphi : \mathbb{R} \to \mathbb{R}$ de clase C^∞ acotada, entonces $X_n \xrightarrow{D} X$.

Demostración

Tenemos que probar que $F_{X_n}(x_0) \to F_X(x_0)$ cuando $n \to +\infty$, para cada punto de continuidad x_0 de F_X. Para ello, la idea es usar los escalones suaves φ_δ. Primero por la derecha

Dado $\varepsilon > 0$, afirmamos que si δ es suficientemente pequeño,

$$|E[\varphi_\delta(X)] - F_X(x_0)| < \frac{\varepsilon}{2} \tag{3}$$

Notamos que:

$$\begin{aligned}
E[\varphi_\delta(X)] &= \int_{-\infty}^{\infty} \varphi_\delta(x)\, dF_X(x) \\
&= \int_{-\infty}^{x_0} \varphi_\delta(x)\, dF_X(x) + \int_{x_0}^{x_0+\delta} \varphi_\delta(x)\, dF_X(x) + \int_{x_0+\delta}^{\infty} \varphi_\delta(x)\, dF_X(x) \\
&= \int_{-\infty}^{x_0} 1\, dF_X(x) + \int_{x_0}^{x_0+\delta} \varphi_\delta(x)\, dF_X(x) + \int_{x_0+\delta}^{\infty} 0\, dF_X(x) \\
&= F_X(x_0) + \int_{x_0}^{x_0+\delta} \varphi_\delta(x)\, dF_X(x)
\end{aligned}$$

Demostración (2)

Entonces

$$|E[\varphi_\delta(X)] - F_X(x_0)| = \left| \int_{x_0}^{x_0+\delta} \varphi_\delta(x)\, dF_X(x) \right| \leq F_X(x_0+\delta) - F_X(x_0)$$

acotando la integral de Stieltjes usando que $0 \leq \varphi_\delta(x) \leq 1$.
Entonces, la afirmación (3) se deduce de la continuidad (por la derecha) de la función de distribución F_X.
Fijamos un $\delta = \delta(\varepsilon)$ tal que se verifique (3). Entonces, por la hipótesis, existirá un n_0 tal que si $n \geq n_0$ tenemos que,

$$|E[\varphi_\delta(X_n)] - E[\varphi_\delta(X)]| < \frac{\varepsilon}{2} \tag{4}$$

Demostración (4)

Usando que $\varphi_\delta(x) \geq I_{\leq x_0}(x)$ deducimos que si $n \geq n_0$, tenemos que:

$$\begin{aligned} F_{X_n}(x_0) &= P\{X \leq x_0\} = E[I_{\leq x_0}(X)] \\ &\leq E[\varphi_\delta(X_n)] \\ &\leq E[\varphi_\delta(X)] + \frac{\varepsilon}{2} \quad \text{por (4)} \\ &\leq F_X(x_0) + \varepsilon \quad \text{por (3)} \end{aligned}$$

si $n \geq n_0(\varepsilon)$. Como $\varepsilon > 0$ es arbitrario, hemos probado que

$$\limsup_{n \to +\infty} F_{X_n}(x_0) \leq F_X(x_0) \tag{5}$$

Demostración (3)

Para probar que $F_{X_n}(x_0) \to F_X(x_0)$, necesitamos demostrar también una desigualdad en el sentido contrario.
Para ello, aproximamos $I_{(-\infty,x_0]}$ por escalones suaves desde la izquierda.
El argumento entonces es similar. Usando la continuidad de F en x_0 por la izquierda tendremos que si δ es suficientemente pequeño,

$$|E[\varphi_{-\delta}(X)] - F_X(x_0)| < \frac{\varepsilon}{2} \tag{6}$$

ya que

$$|E[\varphi_{-\delta}(X)] - F_X(x_0)| = \left| \int_{x_0-\delta}^{x_0} \varphi_{-\delta}(x)\, dF_X(x) \right| \leq F_X(x_0) - F_X(x_0 - \delta)$$

Fijamos un $\delta = \delta(\varepsilon)$ tal que se verifique (6). Usando la hipótesis, dado $\varepsilon > 0$, existirá un n_0 tal que si $n \geq n_0$ tenemos que,

$$|E[\varphi_{-\delta}(X_n)] - E[\varphi_{-\delta}(X)]| < \frac{\varepsilon}{2} \tag{7}$$

Demostración (4)

Ahora notamos que $\varphi_{-\delta} \leq I_{\leq x_0}$, luego

$$\begin{aligned} F_{X_n}(x_0) &= P\{X_n \leq x_0\} = E[I_{(-\infty, x_0-\delta]}(X_n)] \\ &\geq E[\varphi_{-\delta}(X_n)] \\ &\geq E[\varphi_{-\delta}(X)] - \frac{\varepsilon}{2} \quad \text{por (7)} \\ &\leq F_X(x_0) - \varepsilon \quad \text{por (6)} \end{aligned}$$

si $n \geq n_0(\varepsilon)$. Como $\varepsilon > 0$ es arbitrario, hemos probado que

$$\liminf_{n \to +\infty} F_{X_n}(x_0) \geq F_X(x_0) \tag{8}$$

Juntando (5) y (8), hemos probado que:

$$F_{X_n}(x_0) \to F_X(x_0) \quad \text{cuando } n \to +\infty$$

como queríamos.