Más sobre convergencias

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

Algunas propiedades más de la convergencia en distribución

Recordamos una definición: Convergencia en distribución

Hoy vamos a seguir estudiando las nociones de convergencia.

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:**

$$X_n \xrightarrow{D} X$$

Unicidad del límite en distribución

Proposición

Si $X_n \xrightarrow{D} X$ *y* $X_n \xrightarrow{D} Y$, *entonces* $F_X = F_Y$ *(*X *e* Y *están idénticamente distribuidas)*

Demostración.

$F_X(x) = F_Y(x)$ en cada x que sea simultáneamente punto de continuidad de F_X y F_Y. Pero F_X y F_Y son crecientes, y tienen por lo tanto a lo sumo una cantidad numerable de discontinuidades. Deducimos que $F_X(x) = F_Y(x)$ para los x en un subconjunto denso de $\mathbb{R}$, y entonces para todo x ya que ambas son continuas por la derecha. $\square$

Parte II

El recíproco fuerte del teorema de Helly-Bray

Resultados que vimos en la clase anterior

Teorema (de Helly-Bray)

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.*

Teorema (Recíproco del teorema de Helly-Bray)

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* $X_n \xrightarrow{D} X$.

Recíproco fuerte del teorema de Helly-Bray

Con un poco más de trabajo, vamos a probar una versión más precisa:

Teorema (Recíproco fuerte del teorema de Helly)

Sean (X_n) es una sucesión de variables aleatorias finitas en casi todo punto, y X otra variable aleatoria finita en casi todo punto, tales que

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

para toda función $\psi : \mathbb{R} \to \mathbb{R}$ de clase C^∞ con soporte compacto, entonces $X_n \xrightarrow{D} X$.

Una observación

Vimos en la clase 15 que si X es una variable aleatoria finita en casi todo punto. está acotada en probabilidad: dado $\varepsilon > 0$ existe M_ε tal que

$$P\{|X| > M_\varepsilon\} < \varepsilon$$

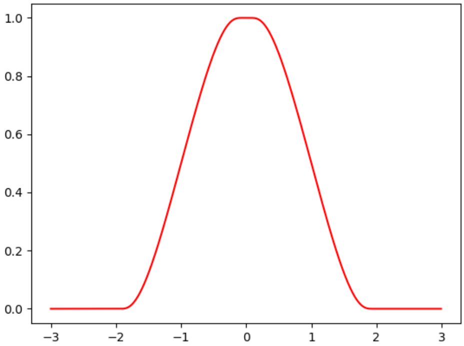
Lema

Si una sucesión de variables aleatorias (X_n) verifica la hipótesis del recíproco fuerte del teorema de Hellly-Bray, entonces es acotada en probabilidad, o sea: Dado, $\varepsilon > 0$ existe N_ε tal que

$$P\{|X_n| > N_\varepsilon\} < \varepsilon \text{ para todo } n \in \mathbb{N}$$

Demostración

Dado $\varepsilon > 0$, tomamos M_ε como en la diapositiva anterior. Y elegimos φ C^∞ con soporte en $K_\varepsilon = [-2M_\varepsilon, 2M_\varepsilon]$ tal que $\varphi \equiv 1$ en $J_\varepsilon = [-M_\varepsilon, M_\varepsilon]$.



Por ejemplo, acá vemos el gráfico de φ si $M_\varepsilon = 1$, con lo que $K_\varepsilon = [-2, 2]$ y $J_\varepsilon = [-1, 1]$.

Demostración (2)

Entonces $I_{K_\varepsilon} \geq \varphi$, luego

$$P\{|X_n| \leq 2M_\varepsilon\} = E[I_{K_\varepsilon}(X_n)] \geq E[\varphi(X_n)]$$

Y por lo tanto si $n \geq n_0(\varepsilon)$ tendremos

$$P\{|X_n| \leq 2M_\varepsilon\} \geq E[\varphi(X)] - \varepsilon$$

Pero $\varphi \geq I_{J_\varepsilon}$, luego

$$E[\varphi(X)] \geq E[I_{J_\varepsilon}(X)] = P\{|X| \leq M_\varepsilon\} \geq 1 - \varepsilon$$

Entonces

$$P\{|X_n| \leq 2M_\varepsilon\} \geq 1 - 2\varepsilon$$

o sea:

$$P\{|X_n| > 2M_\varepsilon\} \leq \varepsilon$$

Demostración del recíproco fuerte de Helly-Bray

Sabemos que

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

para toda función $\psi : \mathbb{R} \to \mathbb{R}$ de clase C^∞ con soporte compacto. Vamos a probar que

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

para toda φ de clase C^∞ acotada. Con lo que la versión débil del teorema va implicar la versión fuerte.
Fijemos φ. Supongamos que $|\varphi(x)| \leq C$ para todo x.
Dado $\varepsilon > 0$, consideramos ρ de clase C^∞ tal que $\rho(x) = 1$ en $[-N_\varepsilon, N_\varepsilon]$ y $0 \leq \rho(x) \leq 1$ siempre, consideramos $\phi = \rho \cdot \varphi$ y escribimos

$$\begin{aligned} |E[\varphi(X)] - E[\varphi(X_n)]| &\leq |E[\varphi(X)] - E[\psi(X)]| \\ &+ |E[\psi(X)] - E[\psi(X_n)]| + |E[\varphi(X_n)] - E[\psi(X_n)]| \end{aligned}$$

Demostración del recíproco fuerte de Helly-Bray (2)

Acotemos:

$$
\begin{aligned}
|E[\varphi(X_n)] - E[\psi(X_n)]| &= \left| \int_{-\infty}^{\infty} [\varphi(x) - \psi(x)] \, dF_n(x) \right| \\
&\leq \int_{|x|>N_\varepsilon} (1-\rho(x)) \cdot |\varphi(x)| \, dF_n(x) \\
&\leq \int_{|x|>N_\varepsilon} C \, dF_n(x) = CP\{|X_n| > N_\varepsilon\} < C\varepsilon
\end{aligned}
$$

para todo n por el lema (¡Es una cota uniforme!). Similarmente:

$$|E[\varphi(X)] - E[\psi(X)]| < C\varepsilon$$

y finalmente por la hipótesis

$$|E[\psi(X)] - E[\psi(X_n)]| < \varepsilon \text{ si } n \geq n_0(\varepsilon)$$

Luego reemplazando en la desigualdad triangular que teníamos antes

$$|E[\varphi(X)] - E[\varphi(X_n)]| < (2C+1)\varepsilon \text{ si } n \geq n_0(\varepsilon)$$

Demostración del recíproco fuerte de Helly-Bray (3)

Entonces vemos que:

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

Como esto vale para toda φ de clase C^∞ acotada, por la vesión débil del terema deducimos que:

$$X_n \xrightarrow{D} X$$

como queríamos.

Un ejemplo

Ejercicio

Sea $D \subset \mathbb{R}$ un conjunto discreto (=sin puntos de acumulación) y sean $X_n, X : \Omega \to \mathbb{R}$ variables aleatorias concentradas en D. Llamemos

$$p_n(k) = P\{X_n = k\}, \quad p(k) = P\{X = k\}$$

Entonces

$$X_n \xrightarrow{D} X \Leftrightarrow p_n(k) \to p(k) \text{ para todo } k \in D$$

Esta enunciado generaliza el ejercicio 2 de la práctica 8, que corresponde al caso especial $D = \mathbb{Z}$. Vamos a resolverlo usando el teorema de Helly-Bray (aunque podría resolverse usando la definición de convergencia en distribución).

Solución

Supongamos primero que $X_n \xrightarrow{D} X$. Dado un $k \in D$ consideramos un entono abierto U de k donde k sea el único punto de D (existe por la hipótesis de que D es discreto).
Consideramos $\varphi : \mathbb{R} \to \mathbb{R}$ de clase C^∞ con soporte contenido en U tal que $\varphi(k) = 1$. Entonces por Helly-Bray

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

pero $\varphi(X_n)$ es una variable aleatoria discreta, pues X_n está concentrada en D. Luego

$$E[\varphi(X_n)] = \sum_{d \in D} \varphi(d) \cdot p_n(d) = p_n(k)$$

Similarmente

$$E[\varphi(X)] = p(k)$$

Se deduce que $p_n(k) \to p(k)$. Esto vale para todo $k \in D$.

Solución (2)

Recíprocamente, supongamos que $p_n(k) \to p(k)$ para todo $k \in D$. Queremos ver que $X_n \xrightarrow{D} X$. Para eso, vamos a usar el recíproco fuerte del teorema de Helly-Bray. Luego queremos probar que

$$E[\psi(X_n)] \to E[\psi(X)] \tag{1}$$

para toda $\psi : \mathbb{R} \to \mathbb{R}$ de clase C^∞ con soporte compacto. Llamemos K al soporte de ψ. Entonces $D \cap K$ es finito, y

$$E[\psi(X_n)] = \sum_{d \in D \cap K} \psi(d) \cdot p_n(d)$$

Similarmente

$$E[\psi(X)] = \sum_{d \in D \cap K} \psi(d) \cdot p(d)$$

Como son sumas finitas, es claro que (1) se va a cumplir, ya que el límite de una suma finita es igual a la suma de los límites.

Otra aplicación del teorema de Helly-Bray

Corolario

Si $X_n \xrightarrow{D} X$ *y* $g : \mathbb{R} \to \mathbb{R}$ *es una función continua, entonces* $g(X_n) \xrightarrow{D} g(X)$.

Demostración.

Por el recíproco del teorema de Helly-Bray nos bastará probar que

$$E[\psi(g(X_n))] \to E[\psi(g(X))] \text{ para toda } \psi \in C_c^\infty(R)$$

Esto es:

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

donde $\varphi = \psi \circ g$. Pero notamos que ψ es continua y acotada, por ser composición de continuas y ψ acotada. Deducimos que esto es cierto, en virtud del teorema de Helly-Bray. □

Una observación

Corolario

Si $X_n \xrightarrow{D} X$, *y* a, b *son constantes, entonces* $aX_n + b \xrightarrow{D} aX + b$.

Observación

Sin embargo, no es cierto en general que si

$$X_n \xrightarrow{D} X, Y_n \xrightarrow{D} Y \Rightarrow X_n + Y_n \xrightarrow{D} X + Y$$

Relación con la convergencia en probabilidad

Vimos en la clase anterior la siguiente

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 $X_n \xrightarrow{D} X$.

En general la recíproca no es cierta. Pero sí es cierta cuando las variables convergen a cero.

Proposición

Si $X_n \xrightarrow{D} 0$, entonces $X_n \xrightarrow{P} 0$.

Demostración

Fijemos $\delta > 0$.

$$\{|X_n| \geq \delta\} = \{X_n \leq -\delta\} \cup \{X_n \geq \delta\}$$

$$\begin{aligned} P\{|X_n| > \delta\} &= P\{X_n \leq -\delta\} + P\{X_n \geq \delta\} \\ &= P\{X_n \leq -\delta\} + 1 - P\{X_n < \delta\} \\ &\leq P\{X_n \leq -\delta\} + 1 - P\{X_n \leq \delta/2\} \\ &= F_{X_n}(-\delta) + 1 - F_{X_n}(\delta/2) \end{aligned}$$

Pero por la hipótesis

$$F_{X_n}(t) \to F_0(t) = \begin{cases} 1 & \text{si} \quad t > 0 \\ 0 & \text{si} \quad t < 0 \end{cases}$$

para todo $t \neq 0$. Luego,

$$P\{|X_n| > \delta\} \to 0$$

Como $\delta > 0$ es arbitrario, deducimos que $X_n \xrightarrow{P} 0$.

Parte III

Algunas propiedades más de la convergencia en probabilidad

La convergencia en probabilidad es la convergencia en un espacio métrico

Teorema

Dado un espacio de probabilidad $(\Omega, \mathcal{E}, P)$ *consideramos el espacio vectorial*

$$VA(\Omega) = \{X : \Omega \to \overline{\mathbb{R}} \text{ variable aleatoria finita en casi todo punto}\}$$

Dadas $X, Y \in \Omega$, *definimos*

$$d(X, Y) = E[\varphi_0(X - Y)] = E\left[\frac{|X - Y|}{1 + |X - Y|}\right]$$

donde $\varphi_0 : \mathbb{R}_{\geq 0} \to \mathbb{R}$ *es la función*

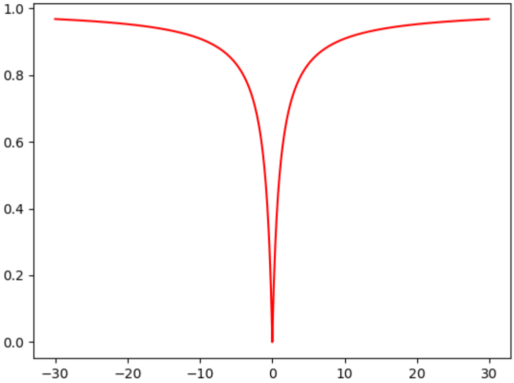
$$\varphi_0(x) = \frac{|x|}{1 + |x|}$$

Entonces d *es una métrica en* $VA(\Omega)$, *y*

$$X_n \xrightarrow{P} X \Leftrightarrow d(X_n, X) \to 0$$

Algunas obsevaciones

¿Cómo es el dibujo de φ_0 ?



Notamos φ_0 es continua y acotada, y que en $\mathbb{R}_{\geq 0}$, φ_0 es creciente.

Algunas observaciones

$$\begin{aligned} d(X,Y) = 0 &\Leftrightarrow E[\varphi_0(X-Y)] = 0 \\ &\Leftrightarrow \varphi_0(X-Y) = 0 \text{ con probabilidad } 1 \\ &\Leftrightarrow X = Y \text{ con probabilidad } 1 \end{aligned}$$

Si identificamos las variables que son iguales con probabilidad 1, $VA(\Omega)$ resulta un espacio métrico.

Es un ejemplo de espacio vectorial topológico, pero no un espacio normado pues

$$d(X,0) = E[\varphi_0(X)]$$

no es una norma. Además $VA(\Omega)$ resulta un espacio métrico acotado con diámetro 1 pues $|\varphi_0| \leq 1$.
Que $d(X,Y) = d(Y,X)$ es inmediato.

Desigualdad Triangular

Para probar que d cumple la desigualdad triangular, hay que usar que

$$\varphi_0(x+y) \leq \varphi_0(x) + \varphi_0(y) \text{ para todo } x, y \geq 0$$

Esto se prueba así:

$$\begin{aligned}
\varphi_0(x+y) &= \frac{x+y}{1+x+y} \\
&= \frac{x}{1+x+y} + \frac{y}{1+x+y} \\
&\leq \frac{x}{1+x} + \frac{y}{1+y} = \varphi_0(x) + \varphi_0(y)
\end{aligned}$$

Luego

$$\begin{aligned}
d(X,Y) &= E[\varphi_0(|X-Y|) \leq E[\varphi_0(|X-Z| + |Z-Y|)] \\
&\leq E[\varphi_0(|X-Z|)] + E[\varphi_0|Z-Y|)] \\
&= d(X,Z) + d(Z,Y)
\end{aligned}$$

Demostración del teorema

Si $X_n \xrightarrow{P} X$, entonces $X_n - X \xrightarrow{P} 0$, y entonces $X_n - X \xrightarrow{D} 0$. Como φ_0 es continua y acotada, el teorema de Helly-Bray nos dice que

$$d(X_n, X) = E[\varphi_0(X_n - X)] \to 0$$

Recíprocamente, si $d(X_n, X) \to 0$ notamos que para todo $\delta > 0$,

$$d(X_n, X) = E[\varphi_0(X_n - X)] \geq E[\varphi_0(\delta) I_{|X_n - X| \geq \delta}] = \frac{\delta}{1+\delta} P\{|X_n - X| \geq \delta\}$$

o sea:

$$P\{|X_n - X| \geq \delta\} \leq \frac{1+\delta}{\delta} \cdot d(X_n, X)$$

Por lo que $X_n \xrightarrow{P} X$.

Una aplicación del lema de Borel Cantelli

Proposición (Criterio para la convergencia casi segura)

Sea $(X_n) : \Omega \to \overline{\mathbb{R}}$ *una sucesión de variables aleatorias, y* $X : \Omega \to \mathbb{R}$ *otra variables aleatoria. Supongamos que para para todo* $\varepsilon > 0$

$$\sum_{n=1}^{\infty} P(\{|X_n - X| > \varepsilon\}) < +\infty \tag{2}$$

(o sea, esta serie converge). Entonces

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

Este teorema lo vimos en la clase 16, aunque allí lo enuncié con la hipótesis más fuerte de que

$$\sum_{n=1}^{\infty} E[|X_n - X|^p] < +\infty \text{ para algún } p > 0$$

Pero si miran la demostración, (2) fue lo que realmente usamos.

Relación entre las convergencias en probabilidad y casi segura

Anteriormente demostramos que

Teorema

Si $X_n \xrightarrow{c.s.} X$, $X_n \xrightarrow{P} X$.

y vimos con con un contraejemplo que la recíproca no es cierta. Sin embargo vamos a probar:

Teorema

Si $X_n \xrightarrow{P} X$, *existe una subsucesión* (X_{n_k}) *de* X_n *tal que* $X_{n_k} \xrightarrow{c.s.} X$.

Demostración

Como $P\{|X_n - X| \geq \delta) \to 0$ para todo $\delta > 0$, existirá un n_1 tal que

$$P\{|X_{n_1} - X| \geq 1\} \leq \frac{1}{2}$$

Luego elegimos $n_2 > n_1$ tal que

$$P\{|X_{n_2} - X| \geq \frac{1}{2}\} \leq \frac{1}{4}$$

y procediendo inductivamente podemos elegir una sucesión creciente de índices

$$n_1 < n_2 < \ldots < n_k < n_{k+1} \ldots$$

tal que:

$$P\{|X_{n_k} - X| \geq \frac{1}{k}\} \leq \frac{1}{2^k}$$

Entonces

$$\sum_{k=k_0}^{\infty} P\{|X_{n_k} - X| \geq \varepsilon\} \leq \sum_{k=k_0}^{\infty} P\left\{|X_{n_k} - X| \geq \frac{1}{k}\right\} < +\infty$$

para $k_0 \geq \frac{1}{\varepsilon}$. Por el teorema anterior, $X_{n_k} \xrightarrow{c.s.} X$.

Parte IV

Convergencia en L^p

El Espacio L^p

Consideramos un espacio de probabilidad $(\Omega, \mathcal{E}, P)$. Para $p \geq 1$, consideramos el espacio

$$L^p(\Omega) = \{\text{variables aleatorias } X : \Omega \to \overline{\mathbb{R}} : E(|X|^p) < \infty\}$$

$L^p(\Omega)$ es un espacio vectorial normado con la norma

$$\|X\|_p = E(|X|^p)^{1/p}$$

Esta norma induce una nueva noción de convergencia de variables aleatorias: la convergencia en L^p:

$$X_n \xrightarrow{Lp} X \Leftrightarrow \|X_n - X\|_p = E(|X_n - X|^p)^{1/p} \to 0$$

Para $p = 2$, esta norma proviene del producto interno

$$\langle X, Y \rangle = E(X \cdot Y)$$

Es $L^2(\Omega)$ es un espacio con producto interno o espacio pre-Hilbert.

Convergencia en norma p

- Por la desigualdad de Markov,

$$P\{|X_n - X| \geq \lambda\} \leq \frac{1}{\lambda^p} E(|X_n - X|^p)$$

  Se deduce que si $X_n \xrightarrow{Lp} X$, entonces $X_n \xrightarrow{P} X$.

- El siguiente es un resultado fudamental que se ve en los cursos de análisis real:

Teorema (Teorema de Riesz-Fischer)

Para toodo $p \geq 1$, $L^p(\Omega)$ es un un espacio normado completo (En particular, $L^2(\Omega)$ es un espacio de Hilbert).

Parte V

Un ejemplo: la serie armónica aleatoria

Planteo del problema

Sabemos que la serie armónica

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n} + \ldots$$

diverge, mientras que la serie armónica alternada

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \cdots + (-1)^n \frac{1}{n} + \ldots$$

convege. ¿Qué sucederá si elegimos los signos al azar?
Más precisamente: Sea (X_n) una sucesión de variables aleatorias independientes con distribución de Rademacher

$$P\{X_n = 1\} = 1/2, P\{X_n = -1\} = 1/2$$

¿Qué podemos decir de la convergencia de la siguiente serie?

$$\sum_{n=1}^{\infty} \frac{X_n}{n}$$

Convergencia de la serie armónica aleatoria (1)

Vamos a ver que la serie armónica aleatoria converge en L^2. Para ello consideramos sus sumas parciales:

$$Y_k = \sum_{n=1}^{k} \frac{X_n}{n}$$

Afirmamos que (Y_k) es una sucesión de Cauchy en L^2. Usando la independencia de las X_k, vemos que si $k > j$

$$\|Y_k - Y_j\|_2^2 = E(|Y_k - Y_j|^2) = \operatorname{Var}\left(\sum_{n=j}^{k} \frac{X_n}{n}\right)$$

$$= \sum_{n=j}^{k} \frac{\operatorname{Var}(X_n)}{n^2} = \sum_{n=j}^{k} \frac{1}{n^2} \leq \sum_{n=j}^{\infty} \frac{1}{n^2} \to 0$$

cuando $j, k \to +\infty$, por ser la cola de una serie convergente. Por el teorema de Riesz-Fischer, Y_k converge en L^2 hacia una variable aleatoria $Y \in L^2$.

Convergencia de la serie armónica aleatoria (2)

Además la desigualdad triangular implica que $\|Y_k\|_{L^2} \to \|Y\|_{L^2}$, luego

$$\text{Var}(Y) = \|Y\|_{L^2}^2 \leq \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

Pregunta: ¿también tendremos convergencia casi segura?. La respuesta será afirmativa. Para ello nuevamente vamos a aplicar el criterio para la convergencia casi segura que vimos antes, con $p = 4$. Hemos de comprobar

$$\sum_{k=1}^{\infty} E[|Y - Y_k|^4] < +\infty$$

Notemos que

$$Y - Y_k = \sum_{n=k+1}^{\infty} \frac{X_n}{n}$$

Convergencia de la serie armónica aleatoria (3)

Ahora repetimos el argumento pero en L^4. Vamos a usar

Lema (Un caso especial de la desigualdad de Khinchine,clase 16)

Sean (X_k) *una sucesión de variables aleatorias independientes con* $E[X_k] = 0$ *y cuarto momento acotado*

$$E[|X_k|^4] \leq M \text{ donde } M \in \mathbb{R}$$

Entonces si los (a_j) *son reales,*

$$E\left[\left(\sum_{i=1}^{n} a_i X_i\right)^4\right] \leq M \left(\sum_{i=1}^{n} a_i^2\right)^2$$

Convergencia de la serie armónica aleatoria (4)

Usando este lema, tenemos que:

$$\|Y_k - Y_j\|_4^4 = E(|Y_k - Y_j|^4) = E\left[\left(\sum_{n=j}^{k} \frac{X_i}{n}\right)^4\right] \leq M\left(\sum_{n=j}^{k} \frac{1}{n^2}\right)^2$$

$$\leq M\left(\sum_{n=j}^{\infty} \frac{1}{n^2}\right)^2 \to 0 \text{ cuando } j, k \to +\infty$$

Entonces Y_k también es de Cauchy en L^4. Y de nuevo, por el teorema de Riesz-Fischer converge en L^4 (al mismo límite, porque converger en L^4 implica conveger en L^2, como consecuencia de la desigualdad de Jensen).

Convergencia de la serie armónica aleatoria (5)

Haciendo $k \to +\infty$, la cuenta anterior da

$$\|Y - Y_j\|_4^4 = E(|Y - Y_j|^4) \leq M \left(\sum_{n=j}^{\infty} \frac{1}{n^2} \right)^2 \leq \frac{C}{j^2}$$

Se deduce que

$$\sum_{k=1}^{\infty} E[|Y - Y_j|^4] < \infty$$

y por lo tanto, por el criterio para la convergencia casi segura,

$$Y_j \xrightarrow{c.s.} Y$$

es decir que la serie

$$\sum_{n=1}^{\infty} \frac{X_n}{n}$$

converge en forma casi segura.