Kolmogorov y la teoría de la la probabilidad

David Nualart

Academia de Ciencias y Universidad de Barcelona

La axiomatización del cálculo de probabilidades

A. N. Kolmogorov: *Grundbegriffe des Wahrscheinlichkeitsrechnung* (1933)

• *Espacio de probabilidad*: $(\Omega, \mathcal{F}, P)$

Ω es un conjunto (conjunto de resultados de una experiencia aleatoria)

$\mathcal{F}$ es una σ-álgebra de partes de Ω

P es una medida en la σ-álgebra $\mathcal{F}$ tal que $P(\Omega) = 1$

• *Variable aleatoria*: Función $X : \Omega \to \mathbb{R}$ tal que para cada número a

$$\{\omega \in \Omega, X(\omega) < a\} \in \mathcal{F}$$

La ley de X es la medida imagen de P en la recta real:

$$P_X([a, b]) = P(X^{-1}([a, b])$$

Características de esta axiomàtica

1.- Era más práctica y útil que la *teoría de colectivos* de von Mises (1919)

2.- Proporciona un espacio de probabilidad preciso para cada experiencia aleatoria y permite eliminar la ambigüedad de las paradojas (Borel, Bertrand)

3.- Marco totalmente abstracto (sin estructura topológica)

Aportaciones de la monografía de Kolmogorov

a) *Construcción de una probabilidad en un producto infinito de espacios*

Objetivo: Construir una probabilidad P en $\Omega = \mathbb{R}^T$, donde $T = [0, \infty)$ o $T = \mathbb{N}$ a partir de sus leyes marginales $p_{t_1,\ldots,t_n}$, donde $p_{t_1,\ldots,t_n}$ es la imagen de P por la proyección

$$\pi_{t_1,\ldots,t_n} : \mathbb{R}^T \to \mathbb{R}^n$$

Teorema *La probabilidad P existe y es única siempre que las marginales sean compatibles*

Los elementos de $\mathbb{R}^T$ son funciones $x : T \to \mathbb{R}$ que se pueden considerar como trayectorias del proceso estocástico $X_t(x) = x(t)$

• Consecuencia: La ley de un proceso estocástico $(X_t)_{t \in T}$ está determinada por las leyes marginales

$$P_{(X_{t_1},\ldots,X_{t_n})} = P \circ (X_{t_1}, \ldots, X_{t_n})^{-1}$$

que pueden elegirse de manera arbitraria siempre que sean compatibles

b) *Contrucción de la probabilidad condicionada por una variable aleatoria X mediante el teorema de Lebesgue-Radon-Nikodym* (1930)

Objetivo: Para cada suceso $A \in \mathcal{F}$, encontrar una variable de la forma $f_A(X)$ que represente la probabilidad de A condicionada por una variable aleatoria X

Solución:

$$f_A(x) = \frac{dP(A \cap X^{-1}(\cdot))}{dP(X^{-1}(\cdot))}(x)$$

Por ejemplo, si $A = X^{-1}([a,b])$, entonces $f_A = \mathbf{1}_{[a,b]}$.

c) *La ley del 0-1*

Si $(X_n)_{n\geq 1}$ es una sucesión de variables aleatorias independientes, los sucesos de la σ-álgebra asintótica

$$\mathcal{G} = \bigcap_{n\geq 1} \mathcal{G}_n, \quad \mathcal{G}_n = \sigma(X_n, X_{n+1}, \ldots)$$

tienen probabilidad cero o uno.

Demostración:

Sea $A \in \mathcal{G}$ y supongamos $P(A) > 0$.

Para todo $B \in \sigma(X_1, X_2, \ldots, X_n)$, se tiene

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = P(B)$$

ya que A y B son independientes

$P(\cdot|A)$ y P conciden en el álgebra $\bigcup_{n\geq 1} \sigma(X_1, X_2, \ldots, X_n)$ y por tanto coinciden en la σ-álgebra generada por todas las variables X_n

Luego, $P(A|A) = P(A)$ lo que implica $P(A) = 1$

Teoremas límite y series

Teorema (1928) $(X_n)_{n\geq 1}$ *variables aleatorias independientes y centradas*

$$\sum_{n\geq 1} E(X_n^2) < \infty \Rightarrow \sum_{n\geq 1} X_n \quad \text{converge c.s.}$$

Desigualdad de Kolmogorov $(X_n)_{n\geq 1}$ variables aleatorias independientes y centradas. Si $S_n = X_1 + \cdots + X_n$, entonces, para todo $\lambda > 0$

$$P(\max_{1\leq k\leq n} |S_k| \geq \lambda) \leq \frac{E(S_n^2)}{\lambda^2}$$

- Generalización de la desigualdad de Tchebychev
- Se extiende a martingalas (Doob)

Teorema de la Tres Series (1929) *$(X_n)_{n\geq 1}$ variables aleatorias independientes. Fijamos $A>0$ y definimos*

$$Y_n = X_n \mathbf{1}_{|X_n|\leq A}$$

Entonces, la serie $\sum_{n\geq 1} X_n$ converge casi seguramente si y solo si las tres series siguientes convergen

$$\textbf{(i)} \sum_{n\geq 1} P(|X_n|>A) = \sum_{n\geq 1} P(X_n \neq Y_n)$$

$$\textbf{(ii)} \sum_{n\geq 1} E(Y_n)$$

$$\textbf{(iii)} \sum_{n\geq 1} \mathrm{Var}(Y_n)$$

• La propiedad $\sum_{n\geq 1} P(X_n \neq Y_n) < \infty$ implica, por el lema de Borel Cantelli, que casi seguramente las dos sucesiones (X_n) y (Y_n) coinciden a partir de cierto lugar (sucesiones equivalentes)

• Consecuencia: La serie $\sum_{n\geq 1} X_n$ converge casi seguramente si y solo si converge en probabilidad

Leyes de los Grandes Números

$(X_n)_{n\geq 1}$ variables aleatorias independientes

Problema: Comportamiento asintótico de $\frac{X_1+\cdots+X_n}{n}$

Teorema (1930) $(X_n)_{n\geq 1}$ *variables aleatorias independientes y centradas* $S_n = X_1 + \cdots + X_n$.

$$\sum_{n\geq 1} \frac{E(X_n^2)}{n^2} < \infty \Rightarrow \frac{S_n}{n} \to 0 \quad \text{c.s.}$$

• La condición sobre las varianzas es óptima

Teorema (Ley Fuerte de los Grandes Números) $(X_n)_{n\geq 1}$ *variables aleatorias independientes y con la misma distribución*

$$\textbf{(i)} \qquad E(|X_1|) < \infty \Rightarrow \frac{S_n}{n} \to E(X_1) \quad \text{c.s.}$$

$$\textbf{(ii)} \qquad E(|X_1|) = \infty \Rightarrow \limsup_{n\to\infty} \frac{|S_n|}{n} = +\infty \quad \text{c.s.}$$

Ley del logaritmo iterado

Teorema (1929) $(X_n)_{n\geq 1}$ *variables aleatorias independientes, centradas.* $S_n = X_1 + \cdots + X_n$. *Se cumple*

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2B_n \log\log B_n}} = 1, \quad \text{c.s.}$$

si

$$B_n := \sum_{k=1}^{n} E(X_k^2) \to \infty$$

$$|X_n| \leq M_n = o\left(\sqrt{\frac{B_n}{\log\log B_n}}\right)$$

• En el caso identicamente distribuido, con $\sigma^2 = E(X_1^2)$, se obtiene

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n \log\log n}} = \sigma \quad \text{c.s.}$$

lo que precisa la ley fuerte de los grandes números: $\frac{S_n}{n} \to 0$ c.s.

• La demostración de Kolmogorov hizo posible extender el resultado a variables no acotadas, mediante argumentos de truncación

Procesos de Markov

A. N. Kolmogorov: *Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung* Math.Ann. **104**, 415-458 (1931)

Un proceso estocástico $(X_t)_{t\geq 0}$ con valores en un espacio de estados E es un *proceso de Markov* si para todo $s < t$ y $A \subset \mathbb{R}$,

$$P(X_t \in A | X_r, 0 \leq r \leq s) = P(X_t \in A | X_s)$$

• *Futuro y pasado son independientes si se conoce el presente*

Podemos poner $P(X_t \in A | X_s) = P(s, X_s, t, A)$

• La función $P(s, x, t, A)$ representa la probabilidad de que el proceso esté en A en el instante t sabiendo que en el instante s está en x *(probabilidades de transición)*

Ecuación de Chapman-Kolmogorov: Si $s < u < t$

$$P(s, x, t, A) = \int_E P(s, x, u, dy) P(u, y, t, A)$$

En el caso $E = \mathbb{R}$ Kolmogorov obtiene ecuaciones diferenciales para las densidades

$$f(s,x,t,y) = \frac{P(s,x,t,dy)}{dy}$$

bajo ciertas hipótesis que garantizan la existencia de los límites siguientes denominados *(coeficientes de deriva y de difusión)*:

$$A(t,x) = \lim_{\delta\to 0} \frac{1}{\delta} \int_{\mathbb{R}} (y-x) P(t,x,t+\delta,dy)$$
$$B(t,x) = \lim_{\delta\to 0} \frac{1}{2\delta} \int_{\mathbb{R}} (y-x)^2 P(t,x,t+\delta,dy)$$

Ecuación "forward" de Kolmogorov:

$$\frac{\partial f}{\partial s} = -A(s,x)\frac{\partial f}{\partial x} - B^2(s,x)\frac{\partial^2 f}{\partial x^2}$$

Ecuación "backward" de Kolmogorov:

$$\frac{\partial f}{\partial t} = -\frac{\partial [A(t,y)f]}{\partial y} + \frac{\partial^2 [B^2(t,y)f]}{\partial y^2}$$

• Estas ecuaciones son el punto de partida de la interacción entre procesos de Markov y ecuaciones en derivadas parciales

Criterio de continuidad

Se dice que dos procesos estocásticos $X = (X_t)_{0\le t\le 1}$ y $Y = (Y_t)_{0\le t\le 1}$ son equivalentes (o que X es una modificacion de Y) si para cada t

$$P(X_t \neq Y_t) = 0$$

Dos procesos equivalentes pueden tener trayectorias diferentes

Teorema *Consideremos un proceso estocástico* $(X_t)_{0\le t\le 1}$. *Supongamos que existen constantes* $\gamma, c, \epsilon > 0$ *tales que*

$$E(|X_t - X_s|^\gamma) \le c|t-s|^{1+\epsilon}.$$

Entonces, existe una modificación de X *que tiene trayectorias continuas*

E. B. Slutsky: *Qualche proposizione relativa alla teoria delle funzioni aleatorie* . Giornale dell'Instituto Italiano dei Attuari **8**, 183-199 (1937)

• Se puede demostrar que la modiciación de X tiene trayectorias casi seguramente hölderianas de orden $\alpha < \frac{\epsilon}{\gamma}$, es decir

$$|X_t - X_s| \le G|t-s|^\alpha$$

• Ejemplo: El movimiento browniano cumple

$$E(|X_t - X_s|^{2k}) \le c_k|t-s|^k$$

y por lo tanto tiene trayectorias hölderianas de orden $\alpha < \frac{1}{2}$

Conclusiones

A) Kolmogorov es el fundador de la teoría de la probabilidad. La monografía de Kolmogorov transformó el carácter del cálculo de probabilidades, convirtiéndolo en una disciplina matemática

B) Los resultados sobre teoremas límite para sucesiones y series de variables independientes fueron definitivos y son todavía de actualidad

C) Las ideas de Kolmogorov influenciaron decisivamente casi todo el trabajo realizado sobre procesos de Markov e hicieron posible el desarrollo posterior del análisis estocástico