Lema de Borel-Cantelli y convergencia casi segura

Introducción

En el año 1909, se publicó un articulo de Borel el cual abrió una polémica acerca de las propiedades que debían pedirse a la función probabilidad en una formulación axiomática.En ese artículo, titulado *Les probabilités dénombrables et leurs applications arithmétiques*, decía Borel:

"Se distinguen generalmente, en los problemas de probabilidad, dos categorías principales, dependiendo de que el número de casos posibles sea finito o infinito: la primera categoría constituye lo que se llama las *probabilidades discontinuas*, o probabilidades en el dominio del discontinuo, mientras que la segunda categoría comprende las *probabilidades continuas o probabilidades geométricas*. Tal clasificación aparece como incompleta cuando se consideran los resultados de la Teoría de Conjuntos; entre la potencia de los conjuntos finitos y la potencia del continuo se encuentra la potencia de los conjuntos numerables; me propongo mostrar brevemente el interés respecto a las cuestiones de probabilidad en cuyo enunciado intervienen tales conjuntos; las llamaré, para abreviar, **probabilidades numerables**."

Enunciaremos el resultado de Borel utilizando el concepto de **ensayo de Bernoulli**, el cual se define como un experimento aleatorio que admite únicamente dos posibles resultados: éxito y fracaso.

Teorema de Borel. Consideremos una sucesión infinita numerable de ensayos de Bernoulli y sea p_n la probabilidad de éxito en el ensayo n. Denotemos por A_∞ al evento:

A_∞: Se obtiene una infinidad de éxitos.

Entonces:

Si la serie $\sum_{n=1}^{\infty} p_n$ es convergente, $P(A_\infty) = 0$.

Si los ensayos de Bernoulli son independientes y la serie $\sum_{n=1}^{\infty} p_n$ es divergente, $P(A_\infty) = 1$.

Obsérvese que el teorema de Borel rebasó el marco clásico ya que planteó el cálculo de la probabilidad de eventos cuya ocurrencia o no ocurrencia depende de los resultados de una infininidad de ensayos de Bernoulli.

El artículo de Borel causó un gran impacto en su época sobre todo por una aplicación de sus resultados para deducir una propiedad importante de los números reales.

Sea q es un número natural mayor que 1 y, dado $x \in (0,1)$, expresemos x en la base q:

$$x = \sum_{j=1}^{\infty} \frac{b_j}{q^j},$$

donde cada b_j es un entero no negativo menor que q.

Dado un número $b \in \{0, 1, 2, \dots, q-1\}$ denotemos por $f_n(b)$ a la fracción que resulta de dividir entre n el número de veces que aparece b en los primeros n términos del desarrollo de x en base q. Cuando $\lim_{n\to\infty} f_n(b)$ existe, llamemos a ese límite frecuencia total de b en x.

Se dice que $x = \sum_{j=1}^{\infty} \frac{b_j}{q^j}$ es normal con respecto a la base q si dado cualquier número $b \in \{0, \dots, q-1\}$, la frecuencia total de b en x existe y su valor es igual a $\frac{1}{q}$.

Se dice que $x \in (0,1)$ es absolutamente normal si es normal con respecto a cualquier base $q \in \{2, 3, \dots\}$.

Borel demostró entonces el siguiente resultado:

Para cada $j \in \mathbb{N}$, seleccionemos al azar un elemento del conjunto $\{0, \dots, q-1\}$ y definamos x como la serie $\sum_{j=1}^{\infty} \frac{b_j}{q^j}$. Entonces, la probabilidad de que x sea normal con respecto a la base q es igual a 1.

El resultado de Borel puede expresarse en la forma siguiente:

Sea $\mathcal{E}$ un experimento aleatorio y A un evento relativo a ese experimento, de probabilidad igual a p. Consideremos un nuevo experimento aleatorio consistente en la repetición indefinida del experimento $\mathcal{E}$, de tal manera que cada repetición es independiente de las otras. Sea X_n el número de veces que ocurre el evento A en las primeras n repeticiones del experimento, entonces $P\left[\lim_{n\to\infty} \frac{X_n}{n} = p\right] = 1$.

La forma general de este resultado se conoce como **Ley Fuerte de los Grandes Números**.

Más tarde, Hausdorff formuló y demostró el resultado de Borel,acerca de los números normales, utilizando La Teoría de la Medida:

Sea q es un número natural mayor que 1, entonces la medida del conjunto de puntos en el intervalo $(0, 1)$ que son normales con respecto a la base q, es igual a 1.

Como corolario se tiene que la medida del conjunto de puntos en el intervalo $(0, 1)$ que son absolutamente normales es igual a 1.

En su razonamiento, Borel utilizó algunas de las propiedades que son equivalentes a la σ-aditividad; sin embargo él consideraba que la σ-aditividad no podía considerarse como una propiedad de cualquier función de probabilidad. Para fundamentar su afirmación, daba el siguiente ejemplo:

"**Supongamos, por ejemplo, que existe una manera de elegir de entre la colección infinita de números enteros, uno de ellos al azar, de manera que cada uno de ellos tenga la misma probabilidad, esta probabilidad deberá entonces ser nula, pero su suma debe ser igual a 1**."

Lema de Borel-Cantelli

El teorema de Borel tiene ahora una formulación más general, conocida como lema de Borel-Cantelli.

Recordemos que, dada una sucesión $A_1, A_2, \ldots$ de subconjuntos de un conjunto $\mathbb{F}$, se define el límite inferior (lím ínf) y el límite superior (lím sup) de esa sucesión de la siguiente manera:

$$\liminf A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$$

$$\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

Obsérvese que $\liminf A_n$ está formado por todos los elementos $x \in \mathbb{F}$ para los cuales existe $N \in \mathbb{N}$ tal que $x \in A_n$ para cualquier $n \geq N$, mientras que $\limsup A_n$ está formado por todos los elementos $x \in \mathbb{F}$ que pertenecen a una infinidad de conjuntos de la sucesión. Así que se tiene siempre $\liminf A_n \subset \limsup A_n$.

Teorema 1 (**Lema de Borel-Cantelli - 1a. parte**). *Sea $A_1, A_2, \ldots$ una sucesión de eventos tales que $\sum_{n=1}^{\infty} P(A_n) < \infty$ y sea:*

$$A = \{\omega \in \Omega : \omega \in A_n \text{ para una infinidad de valores de } n\}$$

Entonces $P(A) = 0$.

Demostración

Para cada $m \in \mathbb{N}$, sea $B_m = \bigcup_{n=m}^{\infty} A_n$. Entonces la sucesión de eventos B_m es decreciente y $A = \bigcap_{m=1}^{\infty} B_m$, así que:

$$P(A) = P\left[\bigcap_{m=1}^{\infty} B_m\right] = \lim_{m\to\infty} P\left[B_m\right]$$

Pero:

$$P(B_m) = P\left(\bigcup_{n=m}^{\infty} A_n\right) \leq \sum_{n=m}^{\infty} P(A_n)$$

Por lo tanto:

$$P(A) \leq \lim_{m\to\infty} \sum_{n=m}^{\infty} P(A_n) = 0$$

■

Lema 1. *Sea* $(p_n)_{n\in\mathbb{N}}$ *una sucesión de números reales en el intervalo* $[0,1]$ *tal que la serie* $\sum_{n=1}^{\infty} p_n$ *es divergente, entonces* $\lim_{n\to\infty}\prod_{j=1}^{n}(1-p_j)=0$.

Demostración

Por el teorema del valor medio, se tiene, para cada $x\in[0,1)$:

$$\ln(1-x)=-\tfrac{x}{1-\theta x},\text{ donde }\theta\in(0,1)$$

De manera que, para cualquier $x\in[0,1)$, $\ln(1-x)\leq -x$, es decir, $1-x\leq e^{-x}$, resultado que también es válido para $x=1$.

En particular, para cualquier $n\in\mathbb{N}$, se tiene:

$$\prod_{j=1}^{n}(1-p_j)<e^{-\sum_{j=1}^{n}p_n}$$

Así que:

$$\lim_{n\to\infty}\prod_{j=1}^{n}(1-p_j)\leq \lim_{n\to\infty}e^{-\sum_{j=1}^{n}p_n}=0$$

■

Teorema 2 (Lema de Borel-Cantelli - 2a. parte). *Sea* $A_1, A_2, \ldots$ *una sucesión de eventos independientes tales que* $\sum_{n=1}^{\infty} P(A_n)=\infty$ *y sea:*

$$A=\{\omega\in\Omega:\omega\in A_n\text{ para una infinidad de valores de }n\}.$$

Entonces $P(A)=1$.

Demostración

Para cada $m\in\mathbb{N}$, sea $B_m=\bigcup_{n=m}^{\infty}A_n$. Entonces la sucesión de eventos B_m^c es decreciente y $A^c=\bigcap_{m=1}^{\infty}B_m^c$, así que:

$$P(A^c)=P\left[\bigcap_{m=1}^{\infty}B_m^c\right]=\lim_{m\to\infty}P\left[B_m^c\right]$$

Pero, $B_m^c=\bigcap_{n=m}^{\infty}A_n^c\subset\bigcap_{n=m}^{m+k}A_n^c$ para cualquier $k\in\mathbb{N}$, así que, $P(B_m^c)\leq\prod_{n=m}^{m+k}\left[1-P(A_n)\right]$ para cualquier $k\in\mathbb{N}$. Por lo tanto:

$$P(B_m^c)\leq\lim_{k\to\infty}\prod_{n=m}^{m+k}\left[1-P(A_n)\right]=0$$

Se concluye entonces que:

$$P(A)=1-P(A^c)=1-\lim_{m\to\infty}P\left[B_m^c\right]=1$$

■

Convergencia casi segura

En el contexto de la Teoría de la Probabilidad, la convergencia casi en todas partes será denominada **convergencia casi segura**.

Definición 1 (**Convergencia casi segura**). *Diremos que una sucesión de variables aleatorias* $(X_n)_{n\in\mathbb{N}}$ *converge casi seguramente a una variable aleatoria* X *si* $P\left[\lim_{n\to\infty} X_n = X\right] = 1$. *Si éste es el caso, se escribirá* $X_n \xrightarrow{c.s.} X$.

Las siguientes propiedades se siguen inmediatamente de las correspondientes propiedades para las sucesiones de números reales:

i. Si una sucesión de variables aleatorias $(X_n)_{n\in\mathbb{N}}$ converge casi seguramente a X, entonces cualquier subsucesión de $(X_n)_{n\in\mathbb{N}}$ también converge casi seguramente a X.

ii. Si $(X_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias tal que $X_n \xrightarrow{c.s.} X$ y $X_n \xrightarrow{c.s.} Y$, entonces $P\left[X = Y\right] = 1$.

iii. Si c es una constante y $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias tal que $X_n \xrightarrow{c.s.} X$, entonces $cX_n \xrightarrow{c.s.} cX$.

iv. Si $(X_n)_{n\in\mathbb{N}}$ y $(Y_n)_{n\in\mathbb{N}}$ son dos sucesiones de variables aleatorias tales que $X_n \xrightarrow{c.s.} X$ y $Y_n \xrightarrow{c.s.} Y$, entonces $X_n + Y_n \xrightarrow{c.s.} X + Y$ y $X_nY_n \xrightarrow{c.s.} XY$.

Teorema 3. *Sea* $(X_n)_{n\in\mathbb{N}}$ *una sucesión de variables aleatorias. Entonces* $X_n \xrightarrow{c.s.} 0$ *si y sólo si para cualquier* $\varepsilon > 0$*, se tiene:*

$$P\left[\{\omega \in \Omega : |X_n(\omega)| > \varepsilon \text{ para una infinidad de valores de } n\}\right] = 0$$

Demostración

Supongamos primero que $\lim_{n\to\infty} X_n = 0$ casi seguramente. Entonces existe un conjunto $\Omega_0 \subset \Omega$ de probabilidad 0 tal que si $\omega \in \Omega_0^c$ entonces $\lim_{n\to\infty} X_n(\omega) = 0$. Así que, dado $\omega_0 \in \Omega_0^c$ y $\varepsilon > 0$, existe $N \in \mathbb{N}$ tal que $|X_n(\omega_0)| \leq \varepsilon$ para cualquier $n \geq N$; esto significa que:

$$\omega_0 \in \bigcap_{n=N}^{\infty} \{\omega \in \Omega : |X_n(\omega)| \leq \varepsilon\}$$

Dicho de otra forma, si $\omega_0 \in \Omega_0^c$, entonces, dada cualquier $\varepsilon > 0$:

$$\omega_0 \in \bigcup_{m=1}^{\infty} \left[\bigcap_{n=m}^{\infty} \{\omega \in \Omega : |X_n(\omega)| \leq \varepsilon\}\right]$$

Así que:

$$\Omega_0^c \subset \bigcup_{m=1}^{\infty} \left[\bigcap_{n=m}^{\infty} \{\omega \in \Omega : |X_n(\omega)| \leq \varepsilon\}\right]$$

Por lo tanto:

$$P\left(\bigcap_{m=1}^{\infty} \left[\bigcup_{n=m}^{\infty} \{\omega \in \Omega : |X_n(\omega)| > \varepsilon\}\right]\right) = 0$$

Inversamente, supongamos que, para cualquier $\varepsilon > 0$, se tiene:

$$P\left(\bigcap_{m=1}^{\infty} \left[\bigcup_{n=m}^{\infty} \{\omega \in \Omega : |X_n(\omega)| > \varepsilon\}\right]\right) = 0$$

Para cada $r \in \mathbb{N}$, sea:

$$B_r = \bigcap_{m=1}^{\infty} \left[\bigcup_{n=m}^{\infty} \left\{\omega \in \Omega : |X_n(\omega)| > \tfrac{1}{r}\right\}\right]$$

Se tiene $P(B_r) = 0$ para cualquier $r \in \mathbb{N}$ y la sucesión de conjuntos $B_1, B_2, \ldots$ es creciente, así que:

$$P\left(\bigcup_{r=1}^{\infty} B_r\right) = \lim_{r\to\infty} P(B_r) = 0$$

Pero:

$$B_r^c = \left\{\omega \in \Omega : \text{Existe } N(\omega) \in \mathbb{N} \text{ tal que } |X_n(\omega)| \leq \tfrac{1}{r} \text{ para cualquier } n \geq N(\omega)\right\}$$

De manera que si $\omega \in \bigcap_{r=1}^{\infty} B_r^c$, entonces para cualquier $r \in \mathbb{N}$ existe $N(\omega) \in \mathbb{N}$ tal que $|X_n(\omega)| \leq \frac{1}{r}$ para cualquier $n \geq N(\omega)$. En particular, dada $\varepsilon > 0$ sea $r \in \mathbb{N}$ tal que $\frac{1}{r} < \varepsilon$ y $N(\omega) \in \mathbb{N}$ tal que $|X_n(\omega)| \leq \frac{1}{r}$ para cualquier $n \geq N(\omega)$, entonces $|X_n(\omega)| < \varepsilon$ para cualquier $n \geq N(\omega)$, lo cual significa que $\lim_{n\to\infty} X_n(\omega) = 0$. Es decir:

$$\bigcap_{r=1}^{\infty} B_r^c \subset [\lim_{n\to\infty} X_n = 0]$$

Sea $\Omega_0 = \bigcup_{r=1}^{\infty} B_r$. Entonces $P(\Omega_0) = 0$ y si $\omega \in \Omega_0^c = \bigcap_{r=1}^{\infty} B_r^c$, entonces $\lim_{n\to\infty} X_n(\omega) = 0$. Así que $\lim_{n\to\infty} X_n = 0$ casi seguramente.

∎

Corolario 1. *Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias. Entonces $X_n \xrightarrow{c.s.} X$ si y sólo si para cualquier $\varepsilon > 0$, se tiene:*

$$P\left[\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon \text{ para una infinidad de valores de } n\}\right] = 0$$

Teorema 4. *Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias tales que:*

$$\sum_{n=1}^{\infty} P\left[|X_n| > \varepsilon\right] < \infty \text{ para cualquier } \varepsilon > 0$$

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

Demostración

Dada $\varepsilon > 0$, sea $A(\varepsilon) = \limsup \{\omega \in \Omega : |X_n(\omega)| > \varepsilon\}$.

Por el lema de Borel-Cantelli, $P[A(\varepsilon)] = 0$ para cualquier $\varepsilon > 0$. Así que el resultado se sigue aplicando el corolario 1.

∎

Corolario 2. *Sea X una variable aleatoria y $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias tales que $\sum_{n=1}^{\infty} P\left[|X_n - X| > \varepsilon\right] < \infty$ para cualquier $\varepsilon > 0$. Entonces $X_n \xrightarrow{c.s.} X$.*