
LEYES DE LOS GRANDES NUMEROS

Jorge Rodríguez B., Ana María Sanabria R.
Profesores Asistentes
Departamento de Matemáticas
Universidad del Valle
Cali-Colombia

Introducción

La idea de probabilidad está intimamente relacionada a la frecuencia relativa. Algunos resultados sobre esta relación son los llamados "Leyes de los Grandes Números", que establecen el tipo de convergencia, "débil" o "fuerte", en que las frecuencias relativas de un evento se aproximan a la probabilidad de este. Nos proponemos en este artículo hacer una presentación elemental de estas leyes con un enfoque conjuntista de la ley fuerte.

La *Ley Débil de los Grandes Números*, también llamada *Teorema de Bernoulli* antecedió históricamente a la *Ley Fuerte*, y en su enunciado clásico dice:

"Si la probabilidad de que ocurra un hecho en una prueba única es p , y se hacen varias pruebas, independientemente y en las mismas condiciones, la proporción más probable de que ocurran los hechos en el número total de pruebas es también p ; aún más, la probabilidad que la proporción en cuestión difiera de p en menos de una cantidad dada, por pequeña que sea, aumenta al mismo tiempo que aumenta el número de pruebas" [2].

Y la *Ley Fuerte de los Grandes Números*, llamada también *Teorema de Borel*, se puede enunciar así:

La frecuencia relativa con que ocurre un hecho en pruebas independientes y en las mismas condiciones converge a la probabilidad del hecho observado con probabilidad 1.

Teniendo en cuenta que las dos leyes hacen referencia a la frecuencia relativa con que ocurre un hecho en pruebas independientes y bajo las mismas condiciones, es útil precisar el espacio de probabilidad donde podemos describir estas observaciones.

El Espacio de las Muestras Infinitas

Representemos la ocurrencia de un evento por 1 y la no ocurrencia por 0. Así, el conjunto $\Omega_n = \{0, 1\}^n$, llamado el espacio de las muestras finitas de orden n , describe el conjunto de las posibles realizaciones en las n

pruebas. Esta idealización de las muestras finitas nos sugiere describir el espacio de muestras para un número arbitrario de pruebas por medio de la sucesiones infinitas de 0's y 1's:

$$\Omega = \{0, 1\}^N = \{\omega : \omega = (\omega_1, \omega_2, \dots, \omega_n, \dots); \omega_i \in \{0, 1\}, i \in N\}.$$

Consideremos ahora los conjuntos sobre los cuales definiremos la función de probabilidad en el espacio de muestras infinitas. Sean

$$C_{n,k}(l_n, \dots, l_{n+k-1}) = \{\omega \in \Omega : \omega_j = l_j, n \leq j \leq n+k-1\},$$

los *conjuntos cilíndricos*, donde $n, k \in N$ y $l_j \in \{0, 1\}$ para $j \in \{n, \dots, n+k-1\}$, y sea $\mathcal{F}$ la σ -álgebra generada por estos conjuntos. Puesto que las pruebas en que se observan las realizaciones del hecho son independientes y se efectúan bajo las mismas condiciones, se tiene que, la probabilidad sobre un conjunto cilíndrico viene dada por:

$$P(C_{n,k}(l_n, \dots, l_{n+k-1})) = \prod_{j=n}^{n+k-1} P(\omega_j = l_j),$$

con

$$P(\omega_j = l_j) = \begin{cases} p & \text{si } l_j = 1 \\ q & \text{si } l_j = 0 \end{cases}$$

donde $p + q = 1$ y $p, q > 0$.

Quedando entonces que el espacio de probabilidad de las muestras infinitas es $(\Omega, \mathcal{F}, P)$.

De otro lado, dada una muestra $\omega = (\omega_1, \omega_2, \dots) \in \Omega$, el resultado de la j -ésima observación de ésta, lo podemos definir por medio de la variable aleatoria X_j , $X_j(\omega) = \omega_j$ donde $\omega_j = 1$ significa que el evento se observó y $\omega_j = 0$ significa que el evento no se observó. De manera que la frecuencia relativa, con que ocurre el evento en las n primeras pruebas, viene dada por

$$\frac{S_n}{n} = \frac{\sum_{i=1}^n X_i(\omega)}{n}$$

y, la probabilidad de que S_n sea igual a k es :

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n.$$

La Ley Débil de los Grandes Números

La Ley Débil de los Grandes Números se puede probar directamente de la desigualdad de Chebyshev, sin embargo en esta sección la veremos como un caso particular del Modelo Poisson, corolario del Teorema de Chebyshev.

Teorema de Chebyshev

En la desigualdad de Chebyshev tenemos que si la media μ_x y la varianza σ_x^2 de una variable aleatoria X son finitas, la probabilidad de tener observaciones distintas de la media en un valor mayor a un número δ está controlada por la varianza y δ :

$$P(|X - \mu_x| > \delta) \leq \frac{\sigma_x^2}{\delta^2} \quad ^1.$$

¹ $P(|X - \mu_x| > \delta) = P(\{\omega : |X(\omega) - \mu_x| > \delta\})$.

Su demostración es una consecuencia directa de la definición de varianza.

$$\begin{aligned}\text{Var}(X) &= \int_{\Omega} (x - \mu_x)^2 dP(x) \\ &\geq \int_{\{\omega: |X(\omega) - \mu_x| > \delta\}} (x - \mu_x)^2 dP(x) \\ &\geq \delta^2 * P(|X - \mu_x| > \delta) .\end{aligned}$$

Y de esta desigualdad podemos obtener el llamado

Teorema 1 (Chebyshev) Si se tiene una sucesión de variables aleatorias X_n , con medias μ_n y varianzas σ_n^2 finitas, σ_n^2 convergiendo a 0, entonces, para todo $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} P(|X_n - \mu_n| > \epsilon) = 0.$$

En el caso particular de que las variables aleatorias X_n , sean independientes y Bernoulli, con $P(X_n = 1) = p_n$, las variables aleatorias $S_n = \sum_{i=1}^n X_i/n$ cumplen las hipótesis del Teorema de Chebyshev², pudiendo concluirse entonces que:

$$\left\{ \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n p_i}{n} \right\} \rightarrow 0, \text{ en probabilidad,}$$

resultado llamado *Modelo de Poisson*, que bajo la hipótesis $p_n = p$ para todo n , se reduce al

Teorema 2 (Bernoulli) Si $\{X_n\}$ es una sucesión de variables aleatorias Bernoulli independientes e igualmente distribuidas con $P(X_n = 1) = p$ para todo n , entonces,

$$\left\{ \frac{\sum_{i=1}^n X_i}{n} - p \right\} \rightarrow 0, \text{ en probabilidad.}$$

La Ley Fuerte de los Grandes Números

En esta sección, demostraremos que la convergencia del teorema de Bernoulli se puede fortalecer a convergencia casi segura, obteniéndose el llamado *Teorema de Borel*.

Veamos primero un par de resultados:

Teorema 3 Sea $\{X_n\}$ una sucesión de variables aleatorias. Si para todo $k \in N$,

$$\sum_{n=1}^{\infty} P\left(|X_n| \geq \frac{1}{k}\right) < \infty,$$

entonces

$$X_n \rightarrow 0, \text{ en convergencia casi segura.}$$

Demostración:

Mostrar que $X_n \rightarrow 0$ en convergencia casi segura es equivalente a demostrar que $P(C) = 1$ o lo que es lo mismo que $P(C^c) = 0$, donde C es el llamado conjunto de convergencia de la sucesión $\{X_n\}$,

$$C = \bigcap_{k \geq 1} \bigcup_{m \geq 1} \bigcap_{n \geq m} \left\{ |X_n| < \frac{1}{k} \right\}.$$

² Obsérvese que $\mu_{S_n} = \frac{\sum p_i}{n}$ y $\sigma_{S_n}^2 = \frac{\sum p_i q_i}{n}$.

Veamos que $P(C^c) = 0$. Obsérvese que

$$C^c = \bigcup_{k \geq 1} \bigcap_{m \geq 1} \bigcup_{n \geq m} \left\{ |X_n| \geq \frac{1}{k} \right\}.$$

Puesto que,

$$P \left(\bigcup_{n \geq m} \left\{ |X_n| \geq \frac{1}{k} \right\} \right) \leq \sum_{n \geq m} P \left(|X_n| \geq \frac{1}{k} \right) < \infty,$$

dado $\epsilon > 0$ y k fijo, existe $m_0 = m_0(\epsilon, k)$ tal que

$$\sum_{n \geq m_0} P \left\{ |X_n| \geq \frac{1}{k} \right\} < \frac{\epsilon}{2^k}$$

y puesto que

$$\bigcap_{m \geq 1} \bigcup_{n \geq m} \left\{ |X_n| \geq \frac{1}{k} \right\} \subset \bigcup_{n \geq m_0} \left\{ |X_n| \geq \frac{1}{k} \right\},$$

se tiene que

$$P \left(\bigcap_{m \geq 1} \bigcup_{n \geq m} \left\{ |X_n| \geq \frac{1}{k} \right\} \right) < \frac{\epsilon}{2^k},$$

y por lo tanto

$$\begin{aligned} P \left(\bigcup_{k \geq 1} \bigcap_{m \geq 1} \bigcup_{n \geq m} \left\{ |X_n| \geq \frac{1}{k} \right\} \right) &\leq \sum_{k=1}^{\infty} P \left(\bigcap_{m \geq 1} \bigcup_{n \geq m} \left\{ |X_n| \geq \frac{1}{k} \right\} \right) \\ &\leq \sum_{k=1}^{\infty} \frac{\epsilon}{2^k} \\ &= \epsilon. \end{aligned}$$

Ahora, como este último resultado se tiene para todo $\epsilon > 0$, concluimos que:

$$P(C^c) = 0.$$

Otro resultado que interviene en la demostración del Teorema de Borel es el siguiente:

Lema 1 Para todo $\delta > 0$,

$$P(|X - \mu| \geq \delta) \leq \frac{E(X - \mu)^4}{\delta^4}.$$

Con demostración similar a la de la Desigualdad de Chebyshev.

Teorema 4 (Borel) Sea $\{X_n\}$ una sucesión de variables aleatorias Bernoulli independientes con $P(X_n = 1) = p$. Si $S_n = \sum_{i=1}^n X_i$, entonces

$$\frac{S_n}{n} \longrightarrow p \text{ en convergencia casi segura.}$$

Demostración:

Es suficiente, por el Teorema 3, mostrar que para todo $k \in N$ la serie,

$$\sum_{n=1}^{\infty} P \left(\left| \frac{S_n}{n} - p \right| \geq \frac{1}{k} \right)$$

converge. Por el Lema 1, se tiene que para todo $k \in N$,

$$P \left(\left| \frac{S_n}{n} - p \right| \geq \frac{1}{k} \right) \leq k^4 E \left\{ \left(\frac{S_n}{n} - p \right)^4 \right\}.$$

Resolviendo $E\left\{\left(\frac{S_n}{n} - p\right)^4\right\}$, podemos encontrar una cota para ésta expresión:

$$\begin{aligned} E \left\{ \left(\frac{S_n}{n} - p \right)^4 \right\} &= \frac{1}{n^4} E \left\{ \left[\sum_{i=1}^n (X_i - p) \right]^4 \right\} \\ &= \frac{1}{n^4} \sum_{i,j,k,l=1}^n E \{ (X_i - p)(X_j - p)(X_k - p)(X_l - p) \} \end{aligned}$$

Debido a la independencia de las variables y al hecho de que $E(X_i - p) = 0$, el valor

$$E_{ijkl} = E[(X_i - p)(X_j - p)(X_k - p)(X_l - p)]$$

es diferente de cero solo en los siguientes casos:

(i) $i = j = k = l$

$$\begin{aligned} E_{ijkl} = E((X_i - p)^4) &= (1-p)^4 p + (-p)^4 q \\ &= pq(p^3 + q^3), \end{aligned}$$

(ii) $(i = j, k = l, j \neq k)$, ó, $(i = k, j = l, k \neq l)$, ó, $(j = k, i = l, k \neq i)$

$$\begin{aligned} E_{ijkl} = E\{(X_i - p)^2(X_j - p)^2\} &= E\{(X_i - p)^2\} E\{(X_j - p)^2\} \\ &= \{(1-p)^2 p + (-p)^2 q\}^2 \\ &= p^2 q^2, \quad \text{para } i \neq j. \end{aligned}$$

Teniendo n terminos $E((X_i - p)^4)$ y $3n(n-1)$ terminos $E\{(X_i - p)^2(X_j - p)^2\}$. Así que,

$$E \left\{ \left[\sum_{i=1}^n (X_i - p) \right]^4 \right\} = npq(p^3 + q^3) + 3n(n-1)p^2 q^2 < \frac{n^2}{4}$$

y por tanto,

$$P \left\{ \left| \frac{S_n}{n} - p \right| \geq \frac{1}{k} \right\} \leq \frac{k^4}{4n^2}, \quad \text{para todo } k \in N,$$

teniéndose que la serie

$$\sum_{n=1}^{\infty} P \left\{ \left| \frac{S_n}{n} - p \right| \geq \frac{1}{k} \right\} \leq \sum_{n=1}^{\infty} \frac{k^4}{4n^2}$$

converge.

Con el desarrollo de la Teoría de la Medida, la *Ley Débil* se puede ver como una consecuencia de la *Ley Fuerte*, debido al resultado de que convergencia casi segura (convergencia con probabilidad 1) implica convergencia en probabilidad.

Referencias

- [1] Chung, K. L., A Course in Probability Theory, New York, Harcourt, Brace and World, Inc., 1968.
- [2] Gnedenko, B. V., The Theory of Probability, MIR Publisher, Moscu.
- [3] Newman James, Editor. "La Ley de los Grandes Números", El mundo de las Matemáticas Vol. 3. Sigma. Ediciones Grijalbo.