La desigualdad de Chebyschev y la Ley de los Grandes Nùmeros

Pablo L. De Nápoli

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

Probabilidades y Estadística para Computación
Primer Cuatrimestre de 2023

La desigualdad de Chebyschev

Proposición (desigualdad de Chebyschev)

Sea X una variable aleatoria con esperanza y varianza finitas entonces

$$P\{|X - E(X)| \geq \delta\} \leq \frac{Var(X)}{\delta^2}$$

Intuitivamente, la desigualdad de Tchebschev dice que la varianza de la variable X nos da una estimación de la probabilidad de que X tome valores alejados de su esperanza. Si $\mathrm{Var}(X)$ es pequeña, entonces es poco probable que X tome un valor alejado de $E(X)$.

Demostración - caso v.a. discreta

Consideremos primero el caso en que X es una variable aleatoria discreta. Recordamos que

$$Var(X) = \sum_k (x_k - \mu)^2 \cdot p_k$$

donde $\mu = E[X]$ y $p_k = P\{X = x_k\}$
Si sumamos sobre los valores de k que satisfacen que $|x_k - \mu| > \delta$, vemos que:

$$\begin{aligned} Var(X) &\geq \sum_{|x_k - x| \geq \delta} (x_k - \mu)^2 \cdot p_k \\ &\geq \sum_{|x_k - x| \geq \delta} \delta^2 p_k \\ &= \delta^2 \sum_{|x_k - x| \geq \delta} p_k = \delta^2 \, P\{|X - \mu| > \delta\} \end{aligned}$$

Despejando obtenemos la desigualdad buscada.

Demostración - caso v.a. continua

Si X es una variable continua, la demostración es análoga Recordamos que

$$Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\, dx$$

donde $\mu = E[X]$ y f es la densidad de probabilidad de X.
Si integramos sobre los valores de x que satisfacen que $|x-\mu| > \delta$, vemos que:

$$\begin{aligned} Var(X) &\geq \int_{|x-\mu|>\delta} (x-\mu)^2 \cdot f(x)\, dx \\ &\geq \int_{|x-\mu|>\delta} \delta^2 p_k \cdot f(x)\, dx \\ &= \delta^2 \sum_{||x-\mu|>\delta} \cdot f(x)\, dx = \delta^2\, P\{|X-\mu| > \delta\} \end{aligned}$$

Despejando obtenemos la desigualdad buscada.

Bernoulli, siempre Bernoulli

Una vez más volvemos a los ensayos de Bernoulli, donde considerábamos un experimento aleatorio con dos resultados que convencionalmente se llaman

- éxito (1) con probabilidad p.
- fracaso(0) con probabilidad $q = 1 - p$.

Introdujimos las variables aleatorias de Bernoulli X_i dadas por

$$X_i = \begin{cases} 1 & \text{si la i-ésima realización del experimento es un éxito} \\ 0 & \text{si la i-ésima realización del experimento es un fracaso} \end{cases}$$

Las X_i son variables aleatorias discretas. También lo es el número de éxitos en n ensayos.

$$S_n = X_1 + X_2 + \ldots + X_n$$

Como ya vimos S_n tiene distribución binomial; $S_n \sim Bi(n, p)$

$$P\{S_n = k\} = b(k, n, p) = \binom{n}{k} p^k q^{n-k}, \quad 0 \leq k \leq n$$

El teorema de Bernoulli

Imaginemos que realizamos una sucesión ilimitada de ensayos de Bernoulli.. El siguiente teorema debido a Jacques Bernoulli, y publicado en 1713 en su libro *Ars Conjectandi*, formaliza la idea de que $f_n \to p$ cuando $n \to +\infty$.

Teorema (Teorema de J. Bernoulli)

Sea

$$f_n = \frac{S_n}{n}$$

la frecuencia relativa de éxitos en los n primeros ensayos de una sucesión ilimitada de ensayos de Bernoulli. Entonces dado cualquier $\delta > 0$*,*

$$P\{|f_n - p| > \delta\} \to 0 \text{ conforme } \; n \to \infty$$

Demostración del teorema de Bernoulli

En una clase anterior, vimos que que $E[S_n] = np$ y $\text{Var}(S_n) = npq$. Entonces

$$E[f_n] = p, \text{Var}(f_n) = \text{Var}\left(\frac{S_n}{n}\right) = \frac{pq}{n}$$

.

Luego, por la desigualdad de Chebyschev,

$$P\{|f_n - p| > \delta\} \leq \frac{\text{Var}(f_n)}{\delta^2}$$

En consecuencia:

$$P\{|f_n - p| \geq \delta\} \leq \frac{pq}{n\delta^2} \to 0 \text{ cuando } n \to +\infty \tag{1}$$

Ejemplo

Enunciado

Supongamos que tiramos un dado (equilibrado) 10 veces, y que en 8 de las tiradas sale un 6. ¿cómo podríamos acotar la probabilidad de que esto ocurra usando la desigualdad de Chebyschev ?

Aquí $n = 10$, $p = 1/6$, $q = 5/6$ $S_n = 8$, $f_n = \frac{8}{10}$.
Tomamos

$$\delta = |f_n - p| = \frac{19}{30} \approx 0{,}633333333$$

La cota anterior dice que la probabilidad de observar esto se puede acotar por

$$\frac{pq}{n\delta^2} = 0{,}034$$

O sea: esto puede ocurrir a lo sumo un $3,4\,\%$ de las veces.

Ley débil de los grandes números

Una generalización del teorema de Bernoulli (que se prueba con el mismo argumento) es la siguiente, conocida (al igual que a veces el teorema de Bernoulli) como la ley débil de los grandes números:

Teorema (Ley débil de los grandes números - caso de variancia finita)

Sean $X_1, X_2, \ldots, X_n, \ldots$ *una secuencia infinita de variables aleatorias independientes e idénticamente distribuidas, con*

$$E[X_i] = \mu$$

$$Var(X_i) = \sigma^2 < +\infty$$

Entonces si llamamos

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

y tomamos cualquier $\delta > 0$, *tenemos que*

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

Demostración

Por linealidad de la esperanza, $E[\overline{X}_n] = \mu$, y por otro lado

$$\text{Var}(\overline{X}_n) = \frac{\sigma^2}{n}$$

ya que las X_i son independientes. La desigualdad de Tchebyschev, dice entonces que:

$$P\{|\overline{X}_n - \mu| > \delta\} \leq \frac{\sigma^2}{n\delta^2} \to 0 \text{ cuando } n \to +\infty$$

Algunas observaciones sobre el teorema de Bernoulli

- Si bien la prueba del teorema de Bernoulli, resulta muy sencilla hoy en día, J. Bernoulli dice en su libro que estuvo pensando en este teorema durante más de 20 años, lo cuál muestra que el resultado no es para nada trivial.

- Como todo teorema matemático, el teorema de Bernoulli no afirma nada sobre la realidad, es solamente una afirmación sobre el modelo matemático

  (La cuestión de la validez práctica de un modelo matemático sólo se puede decidir sobre bases empíricas, es decir contrastándolo con la experiencia). Sin embargo, podemos interpretarlo como una muestra de la consistencia interna de nuestro modelo matemático.

- La ley débil de los grandes números recibe este nombre, porque existe otro teorema conocido como la ley fuerte de los grandes números, que afirma que en realidad $f_n \to p$ (o $\overline{X}_n \to \mu$) con probabilidad 1. Este teorema no lo vamos a ver en este curso.