Leyes de los Grandes Números

Teoremas sobre Límites

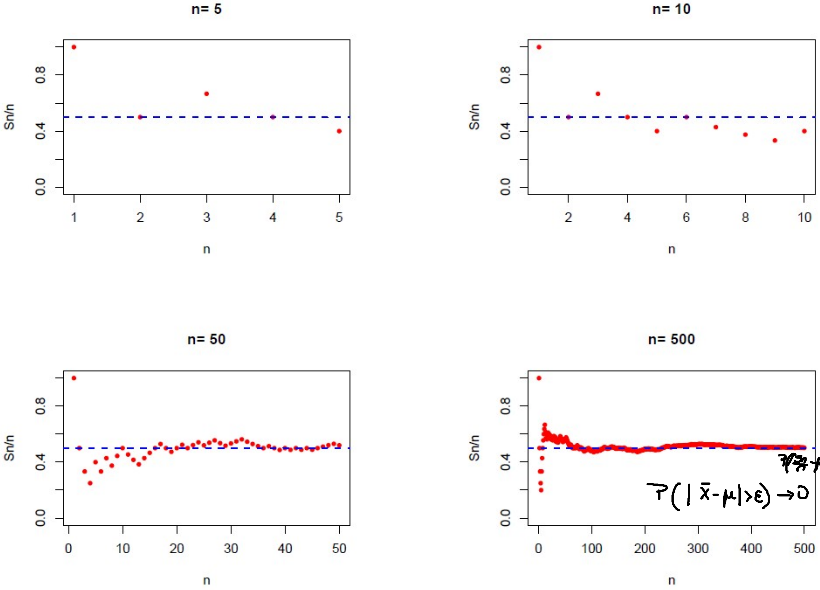
n	S_n
1	1
2	1
3	
$\vdots$	

Experimento: Lanzar una moneda honesta, n veces. Sea $Sn =$ número "caras" obtenidas en los n lanzamientos. Intuitivamente (definición frecuentista de la probabilidad) $S_n/n \cong 1/2$ y $\lim_{n\to\infty} S_n/n = 1/2$.

Este resultado es una consecuencia de una de las Leyes de los Grandes Números.

Estas leyes son una serie de teoremas donde se demuestra cuál es el *límite* o que propiedades satisfacen en el *límite* combinaciones de v.a.
Las hipótesis planteadas y el tipo de convergencia darán lugar a distintos teoremas.

1 si sale cara
0 si sale cruz



$P(|\bar{X}-\mu|<\varepsilon)\to 1$

$P(|\bar{X}-\mu|>\varepsilon)\to 0$

Tipos de Convergencia

$P(|Y_n - Y| \geq \varepsilon) = 1 - P(|Y_n - Y| < \varepsilon)$

Sean $Y, Y_1, Y_2, \ldots$ v.a. definidas sobre un mismo espacio de probabilidad $(\mathcal{E}, \mathcal{A}, P)$

- Se dice que Y_n converge a Y en **probabilidad** si $\forall \epsilon > 0$ $P(|Y_n - Y| \geq \epsilon) \to 0$ cuando $n \to \infty$, se denota por $Y_n \xrightarrow{P} Y$
- Se dice que Y_n converge a Y **casi ciertamente, casi en todo punto** (p.p, a.e) si $P(Y_n \to Y) = 1$ cuando $n \to \infty$. Esto es, $P(\{\omega : Y_n(\omega) \to Y(\omega)\}) = 1$
- Sean $Y, Y_1, Y_2, \ldots$ v.a. con funciones de distribución $F, F_1, F_2, \ldots$. Y_n converge en **distribución** a Y cuando $n \to \infty$ si $F_n(x) \to F(x)$ para todo x punto de continuidad de F. Se denota por $Y_n \xrightarrow{D} Y$.

Preliminares

Desigualdad de Chebyshev

Sea X va con $E(X) = \mu$, $Var(X) = \sigma^2$, y sea $a > 0$.
Entonces

$$P(|X - \mu| > a) \leq \frac{\sigma^2}{a^2}$$

Observaciones

- Esta desigualdad provee una cota para esa probabilidad que sólo depende de la varianza de X. No es necesario conocer la distribución de la variable aleatoria, sólo su media y su varianza.
- Esta cota puede ser grosera, o no informativa (considerar el caso $\sigma^2 > a^2$)

Desigualdad de Chebyshev

Sea X v.a, $E(X)=\mu$, $Var(X)=\sigma^2$. Sea $a>0 \Rightarrow$

$$P(|X-\mu|\geq a) \leq \frac{\sigma^2}{a^2}$$

Dem: Sea X v.a (supongamos que es continua).

Supongamos que f es su función de densidad.

$$P(|X-\mu|\geq a) = \int_{x:|x-\mu|\geq a} 1\cdot f(x)\,dx \leq \int_{x:|x-\mu|\geq a} \frac{(x-\mu)^2}{a^2}\cdot f(x)\,dx \leq (*)$$

si $|x-\mu|\geq a \Rightarrow |x-\mu|^2 \geq a^2$

$\Rightarrow \frac{|x-\mu|^2}{a^2} \geq 1$

$$(*) \leq \int_{-\infty}^{+\infty} \frac{(x-\mu)^2}{a^2}\cdot f(x)\,dx = \frac{1}{a^2}\int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\,dx = \frac{1}{a^2}\underbrace{E\left((X-\mu)^2\right)}_{Var(X)} = \frac{Var(X)}{a^2}$$

$$P(|X-\mu|\geq a) \overset{Var(X)=\sigma^2}{\leq} \frac{\sigma^2}{a^2}$$

Expresiones equivalentes: $P(|X-\mu|>a) \leq \sigma^2/a^2$

1. $\forall a > 0,\ P(|X-\mu| \leq a) \geq 1 - \frac{\sigma^2}{a^2}$
2. $\forall k \geq 0,\ P(|X-\mu| > k\sigma) \leq \frac{1}{k^2}$
3. $\forall k \geq 1,\ P(|X-\mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$

Observación

Las dos útimas expresiones nos informan cómo la desviación estandar mide cuan "concentradas" está la distribución alrededor de μ.

Ejemplo: El número de descargas diarias de una aplicación móvil es una variable aleatoria con media $\mu = 50$ y una desviación estándar $\sigma = 8$. ¿Cuál es la probabilidad que haya entre 34 y 66 descargas en un día? Sea X = nº de descargas diarias de esa aplicación.

$P(34 \leq X \leq 66)$. No sé cuál es la dist. de X.

Pero sí sé que $P(|X-\mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$

$P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - 1/k^2$

$$P(34 \leq X \leq 66) = ? \qquad \mu = 50 \qquad \sigma = 8$$

Cheby.

3) $$P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

$$\mu - k\sigma = 50 - k.8 = 34 \Rightarrow k.8 = 50 - 34 \Rightarrow k = \frac{16}{8} = 2$$

$$\mu + k\sigma = 50 + k.8 = 66 \Rightarrow k.8 = 66 - 50 \Rightarrow k = \frac{16}{8} = 2$$

$$P(34 \leq X \leq 66) \geq 1 - \frac{1}{2^2} = 1 - \frac{1}{4} = 0{,}75$$

La probabilidad que se realicen entre 34 y 66 descargas diarios es 0,75

$$P(|X-\mu|>a) \leq \frac{\sigma^2}{a^2}$$

$$\frac{\sigma^2}{a^2} \geq P(|X-\mu|>a) = 1 - P(|X-\mu| \leq a)$$

$$\Rightarrow P(|X-\mu| \leq a) \geq 1 - \frac{\sigma^2}{a^2}$$

Sea $a = k\sigma$

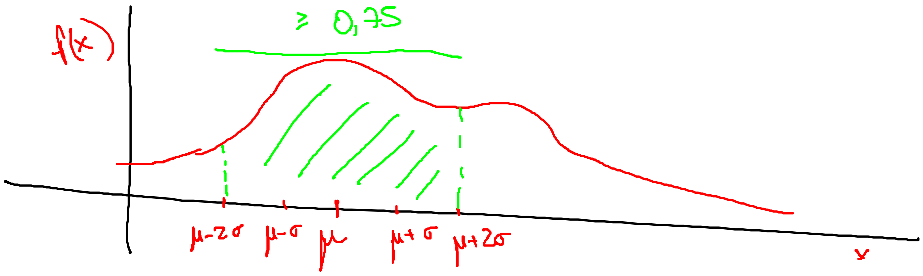
$$P(|X-\mu| \leq k\sigma) \geq 1 - \frac{\cancel{\sigma^2}}{k^2\sigma^2} = 1 - \frac{1}{k^2}$$

$$P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

$$P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

Sup $k=1 \Rightarrow P(\mu - \sigma \leq X \leq \mu + \sigma) \geq 1 - 1 = 0$ ✓

$k=2 \Rightarrow P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \geq 1 - \frac{1}{4} = 0{,}75$



$P(|X-\mu| < \sigma) = 0{,}68$

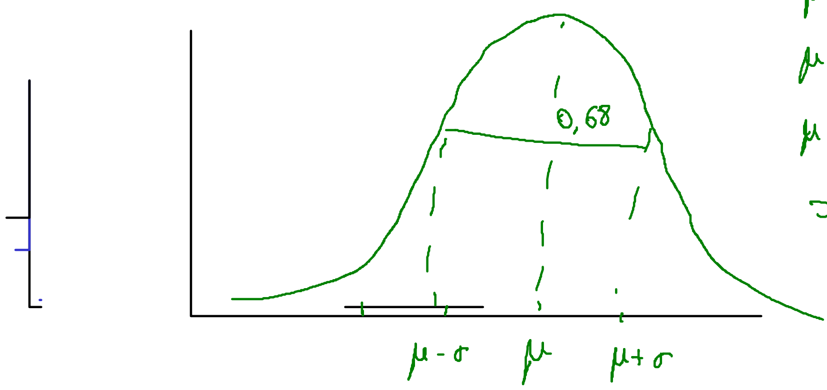
$\mu \pm \sigma \rightarrow 68\%$

$\mu \pm 2\sigma \rightarrow 95\%$

$\mu \pm 3\sigma \rightarrow 99{,}7\%$

$P(|X-\mu| \leq 2\sigma) = 0{,}95$

Si $X \sim N(\mu, \sigma^2)$



Teorema 9.1: Ley débil de Chebyshev

Sean $X_1, X_2, \ldots$ v.a. independientes, supongamos que $\exists c < \infty$ tal que $\forall n,\ \mathrm{Var(X_n) < c}$. Si $S_n = \sum_{i=1}^{n} X_i$, entonces S_n satisface la Ley Débil de los grandes números:

$$\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0$$

Ej. para Matemática.

Ley débil de los grandes números (Teo. de Khintchin)

Teorema 9.2: Teorema de Khintchin

Si $X_1, X_2, \ldots$ son v.a. independientes, idénticamente distribuidas, tales que $E(X_i) = \mu$, $\mathrm{Var(X_i)} = \sigma^2$. Sea $S_n = \sum_{i=1}^{n} X_i$. Entonces

$S_n/n \to \mu$ en probabilidad.

Esto dice que la probabilidad que la media muestral se acerque a la media poblacional es alta, siempre que n sea lo suficientemente grande.

Dem: Quiero probar que dado $\varepsilon > 0$, $P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0$

$$P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) = P\left(n \cdot \left|\frac{S_n}{n} - \mu\right| > n\varepsilon\right) = P\left(\left|S_n - n\mu\right| > n\varepsilon\right) = (1)$$

$$E(S_n) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \mu = n\mu \Rightarrow \begin{cases} E(S_n) = n\mu \\ \mathrm{Var}(S_n) = n\sigma^2 \end{cases}$$

$E(X_i) = \mu$

$$\mathrm{Var}(S_n) = \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = n\sigma^2$$

$$P\left(\left|\frac{S_n}{n}-\mu\right|>\varepsilon\right)=P\left(|S_n-n\mu|>n\varepsilon\right)=P\left(|S_n-E(S_n)|>n\varepsilon\right)\leq$$

$E(S_n)=n\mu$

$Var(S_n)=n\sigma^2$

des. Cheb.

$$\leq\frac{Var(S_n)}{n^2\varepsilon^2}=\frac{n\sigma^2}{n^2\varepsilon^2}=\frac{\sigma^2}{n\varepsilon^2}$$

$$\lim_{n\to\infty}\quad 0\leq\lim_{n\to\infty}P\left(\left|\frac{S_n}{n}-\mu\right|>\varepsilon\right)\leq\lim_{n\to\infty}\frac{\sigma^2}{n\varepsilon^2}=0$$

$$\therefore\quad \lim_{n\to\infty}P\left(\left|\frac{S_n}{n}-\mu\right|>\varepsilon\right)=0\quad y\quad \frac{S_n}{n}\xrightarrow[n\to\infty]{P}\mu$$ ∎

Observación

Este resultado nos dice que el promedio de un número finito de observaciones (realizaciones de las X_i) es una buena aproximación a la estimación de la media poblacional μ.

Corolario 9.1: Teorema de Bernoulli (1713)

Consideremos una sucesión de ensayos Bernoulli independientes, con la misma probabilidad p de éxito en cada ensayo. Sea S_n la variable aleatoria que indica el número de éxitos en los primeros n ensayos. Entonces

$S_n/n \to p$ en probabilidad.

Teorema 9.3: Ley Fuerte de Kolmogorov

Si $X_1, X_2, \ldots$ son v.a. independientes, idénticamente distribuidas, tales que $E(X_i) = \mu$, entonces

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \to \mu \text{ casi ciertamente}$$

(sin demostración)