Clase 2- Ley de los Grandes Números - Este material corresponde a los videos que siguen.

Probabilidades

Carrera de Especialización en Estadística,
Facultad de Ciencias Exactas y Naturales, UBA

2023

Cálculo de Probabilidades en el mundo normal

- Estandarización: Sea $W \sim \mathcal{N}(\mu_W, \sigma_W^2)$, entonces

$$\mathbb{P}\left(|W - \mu_W| > \varepsilon\right) = 2\left(1 - \phi\left(\frac{\varepsilon}{\sigma_W}\right)\right)$$

Cálculo de Probabilidades en el mundo normal

- Estandarización: Sea $W \sim \mathcal{N}(\mu_W, \sigma_W^2)$, entonces

$$\mathbb{P}\left(|W - \mu_W| > \varepsilon\right) = 2\left(1 - \phi\left(\frac{\varepsilon}{\sigma_W}\right)\right)$$

- Promedio de variables iid

$$X_i \sim \mathcal{N}(\mu, \sigma^2)\ , \quad \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \sigma^2/n)$$

$$P\left(|\overline{X}_n - \mu| > \varepsilon\right) =$$

Cálculo de Probabilidades en el mundo normal

- Estandarización: Sea $W \sim \mathcal{N}(\mu_W, \sigma_W^2)$, entonces

$$\mathbb{P}\left(|W - \mu_W| > \varepsilon\right) = 2\left(1 - \phi\left(\frac{\varepsilon}{\sigma_W}\right)\right)$$

- Promedio de variables iid

$$X_i \sim \mathcal{N}(\mu, \sigma^2)\ , \quad \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \sigma^2/n)$$

$$P\left(|\overline{X}_n - \mu| > \varepsilon\right) = 2\left(1 - \phi\left(\frac{\varepsilon * \sqrt{n}}{\sigma}\right)\right),$$

$$\lim_{n\to\infty} P\left(|\overline{X}_n - \mu| > \varepsilon\right) = 0\ , \forall\varepsilon > 0.$$

Convergencia en Probabilidad

Definition

Sean $(Y_n)_{n\geq 1}$, Y variables aleatorias. Diremos que $(Y_n)_{n\geq 1}$ converge a Y en probabilidad si para todo $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(|Y_n - Y| > \varepsilon\right) = 0$$

Notación: $Y_n \to Y$ en probabilidad

Convergencia en Probabilidad

Definition

Sean $(Y_n)_{n\geq 1}$, Y variables aleatorias. Diremos que $(Y_n)_{n\geq 1}$ converge a Y en probabilidad si para todo $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(|Y_n - Y| > \varepsilon\right) = 0$$

Notación: $Y_n \to Y$ en probabilidad

- Demostramos que si $(X_i)_{i\geq 1}$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$, entonces $\overline{X}_n \to \mu$ en probabilidad:

$$\lim_{n\to\infty} \mathbb{P}\left(|\overline{X}_n - \mu| > \varepsilon\right) = 0\ , \quad \forall\varepsilon > 0.$$

Próximo objetivo:

- $(X_i)_{i\geq 1}$ iid, $\mathbb{E}(X_i) = \mu$, $V(X_i) = \sigma^2$. $X_i \sim ?$.
- Promedio :
$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim ?$$
- Esperanza y Varianza del promedio
$$\mu_{\overline{X}_n} = \mathbb{E}(\overline{X}_n) = \mu \ , \quad \sigma^2_{\overline{X}_n} = V(\overline{X}_n) = \frac{\sigma^2}{n} \ .$$
- Queremos demostrar que $\overline{X}_n$ converge a μ en probabilidad:
$$\lim_{n\to\infty} P\left(|\overline{X}_n - \mu| > \varepsilon\right) = 0 \ , \quad \forall \varepsilon > 0.$$
- Problema: No podemos calcular probabilidades porque no conocemos la distribución de X_i.

Desigualdades

- Markov: $X \geq 0$, entonces para todo $\delta > 0$ vale que

$$\mathbb{P}(X \geq \delta) \leq \frac{\mathbb{E}[X]}{\delta}$$

Demostración:

Desigualdades

- Markov: $X \geq 0$, entonces para todo $\delta > 0$ vale que

$$\mathbb{P}(X \geq \delta) \leq \frac{\mathbb{E}[X]}{\delta}$$

- Tchebycheff : Sea W una v.a. y $\varepsilon > 0$

$$\mathbb{P}\left(|W - \mathbb{E}[W]| \geq \varepsilon\right) \leq \frac{V[W]}{\varepsilon^2}$$

Demostración:

El mundo normal vs. Desigualdad Tchebycheff

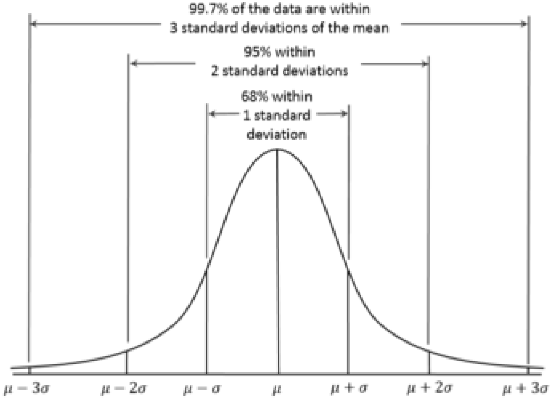
- Probabilidades exactas en el mundo normal

$$\mathbb{P}\left(|W-\mu_W|>\varepsilon\right)=2\left(1-\phi\left(\frac{\varepsilon}{\sigma_W}\right)\right)$$

- Cotas superiores universales (asumiendo $V(W)$ finita)

$$\mathbb{P}\left(|W-\mu_W|>\varepsilon\right)\leq\frac{\sigma_W^2}{\varepsilon^2}$$

Regla Normal - $W \sim \mathcal{N}(\mu, \sigma^2)$



$$\begin{aligned} \mathbb{P}\left(|W-\mu| \geq \sigma\right) &= 0.32\ , \\ \mathbb{P}\left(|W-\mu| \geq 2\sigma\right) &= 0.05\ , \\ \mathbb{P}\left(|W-\mu| \geq 3\sigma\right) &= 0.03 \end{aligned}$$

Desigualdad de Tchebycheff

Sea W una v.a. con media $\mathbb{E}(W) = \mu$ y $V(W) = \sigma^2$. Luego, $\varepsilon > 0$

$$\mathbb{P}\left(|W - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2}$$

Tomando $\varepsilon = k\sigma$,

$$\mathbb{P}\left(|W - \mu| \geq k\sigma\right) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

Desigualdad de Tchebycheff

Sea W una v.a. con media $\mathbb{E}(W) = \mu$ y $V(W) = \sigma^2$. Luego, $\varepsilon > 0$

$$\mathbb{P}\left(|W - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2}$$

Tomando $\varepsilon = k\sigma$,

$$\mathbb{P}\left(|W - \mu| \geq k\sigma\right) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

En particular, para $k = 1, 2, 3$, obtenemos las las siguientes cotas:

$$\begin{aligned}
\mathbb{P}\left(|W - \mu| \geq \sigma\right) &\leq 1\,, \\
\mathbb{P}\left(|W - \mu| \geq 2\sigma\right) &\leq \frac{1}{4}\,, \\
\mathbb{P}\left(|W - \mu| \geq 3\sigma\right) &\leq \frac{1}{9}
\end{aligned}$$

El mundo normal vs. Desigualdad Tchebycheff

- Probabilidades exactas en el mundo normal

$$
\begin{aligned}
\mathbb{P}\left(|W-\mu| \geq \sigma\right) &= 0.32\,, \\
\mathbb{P}\left(|W-\mu| \geq 2\sigma\right) &= 0.05\,, \\
\mathbb{P}\left(|W-\mu| \geq 3\sigma\right) &= 0.03
\end{aligned}
$$

- Cotas superiores universales (asumiendo $V(W)$ finita)

$$
\begin{aligned}
\mathbb{P}\left(|W-\mu| \geq \sigma\right) &\leq 1\,, \\
\mathbb{P}\left(|W-\mu| \geq 2\sigma\right) &\leq \frac{1}{4}\,, \\
\mathbb{P}\left(|W-\mu| \geq 3\sigma\right) &\leq \frac{1}{9}
\end{aligned}
$$

Desigualdad de Tchebycheff aplicada a promedios

- $(X_i)_{i\geq 1}$ iid, $\mathbb{E}(X_i) = \mu$, $V(X_i) = \sigma^2$.
- Promedio :

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim ?$$

- Esperanza y Varianza del promedio

$$\mu_{\overline{X}_n} = \mathbb{E}(\overline{X}_n) = \mu \ , \quad \sigma^2_{\overline{X}_n} = V(\overline{X}_n) = \frac{\sigma^2}{n} \ .$$

- Tchebycheff dice:

$$\mathbb{P}\left(|\overline{X}_n - \mu_{\overline{X}_n}| \geq \varepsilon\right) \ \leq \ \frac{\sigma^2_{\overline{X}_n}}{\varepsilon^2}$$

$$\mathbb{P}\left(|\overline{X}_n - \mu| \geq \varepsilon\right) \ \leq \ \frac{\sigma^2/n}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

Desigualdad de Tchebycheff aplicada a promedios de Variables Bernoulli

- $X_i \sim \mathcal{B}(1, p)$
- $\mathbb{E}(X_i) = p$, $V(X_i) = p(1-p)$.
- $\overline{X}_n$ es la frecuencia relativa de éxitos en las n repeticiones.

$$\mathbb{P}\left(|\overline{X}_n - p| \geq \varepsilon\right) \ \leq \ \frac{p(1-p)}{n\varepsilon^2} \leq$$

Desigualdad de Tchebycheff aplicada a promedios de Variables Bernoulli

- $X_i \sim \mathcal{B}(1, p)$
- $\mathbb{E}(X_i) = p$, $V(X_i) = p(1-p)$.
- $\overline{X}_n$ es la frecuencia relativa de éxitos en las n repeticiones.

$$\mathbb{P}\left(|\overline{X}_n - p| \geq \varepsilon\right) \ \leq \ \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1/4}{n\varepsilon^2}$$

Convergencia en Probabilidad

Definition

Sean $(Y_n)_{n\geq 1}$, Y variables aleatorias. Diremos que $(Y_n)_{n\geq 1}$ converge a Y en probabilidad si para todo $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(|Y_n - Y| > \varepsilon\right) = 0$$

Notación: $Y_n \to Y$ en probabilidad

Convergencia en Probabilidad

Definition

Sean $(Y_n)_{n\geq 1}$, Y variables aleatorias. Diremos que $(Y_n)_{n\geq 1}$ converge a Y en probabilidad si para todo $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(|Y_n - Y| > \varepsilon\right) = 0$$

Notación: $Y_n \to Y$ en probabilidad

- Demostramos que si $(X_i)_{i\geq 1}$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$, entonces $\overline{X}_n \to \mu$ en probabilidad:

$$\lim_{n\to\infty} \mathbb{P}\left(|\overline{X}_n - \mu| > \varepsilon\right) = 0\ , \quad \forall\varepsilon > 0.$$

Teorema: Ley de los Grandes Números (LGN)

Sean $(X_i)_{i\geq 1}$ i.i.d., con $\mathbb{E}(X_i) = \mu$ y $V(X_i) = \sigma^2$, para todo i.
Entonces, el promedio converge a μ en probabilidad:

es decir para todo $\varepsilon > 0$ vale que

$$\lim_{n\to\infty} \mathbb{P}\left(|\overline{X}_n - \mu| > \varepsilon\right) = 0$$

$$\overline{X}_n \to \mu \text{ en probabilidad}$$

Demostración de la Ley de los Grandes Números

Demostración: Tchebycheff prueba que

$$\mathbb{P}\left(|W - \mathbb{E}[W]| \geq \varepsilon\right) \leq \frac{V[W]}{\varepsilon^2}$$

Vamos a invocar la Desigualdad de Tchebicheff, pero con $W = \overline{X}_n$. En tal caso,

$$\mu_{\overline{X}_n} = \mathbb{E}(\overline{X}_n) = \mu$$

$$\sigma^2_{\overline{X}_n} = V(\overline{X}_n) = \frac{\sigma^2}{n}\,.$$

Tenemos entonces que

$$\mathbb{P}\left(|\bar{X}_n - \mu| > \varepsilon\right) = \mathbb{P}\left(|\bar{X}_n - \mu_{\overline{X}_n}| > \varepsilon\right) \leq \frac{V(\overline{X}_n)}{\varepsilon^2} = \frac{1}{\varepsilon^2}\frac{\sigma^2}{n} \to 0$$

Modelo de medición - Errores UNIFORMES

$$X_i = \mu + \varepsilon_i \ , \quad \varepsilon_i \sim \mathcal{U}(-1, 1)$$

1. Calcular $\mathbb{E}(X_i)$ y $V(X_i)$
2. Obtener una cota para la probabilidad de que el promedio de n mediciones diste de la verdadera magnitud μ en menos de 0.1 unidades en función de n. Estudie su monotonía y el límite cuando n tiende a infinito de esta probabilidad.
3. Determinar cuán grande debe ser n para que $P\left(|\overline{X}_n - \mu| < 0.1\right) \geq 0.99$.

Utilizamos la Ley de los Grandes Números para Estimación

Estimación de la Esperanza

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- $\mathbb{E}(X_i) = \mu$, $V(X_i) = \sigma^2$
- Por la ley de los grandes números

$$\overline{X}_n \to \mathbb{E}(X_1) = \mu \quad \text{en probabilidad}$$

- En tal caso, definimos

$$\widehat{\mu} = \widehat{\mu}_n = \overline{X}_n$$

.

- Propiedades de $\widehat{\mu}_n$:
  - $\mathbb{E}(\widehat{\mu}_n) = \mu$
  - $\widehat{\mu}_n \to \mu$ en probabilidad.

Variables Bernoulli

- $X_i \sim \mathcal{B}(1, p)$, i.i.d.
- $\mathbb{E}(X_i) = p$, $V(X_i) = p(1-p)$
- Por la ley de los grandes números

$$\overline{X}_n \to \mathbb{E}(X_1) = p \quad \text{en probabilidad}$$

- En tal caso, definimos

$$\widehat{p} = \widehat{p}_n = \overline{X}_n$$

.

- Propiedades de $\widehat{p}_n$:
  - $\mathbb{E}(\widehat{p}_n) = p$
  - $\widehat{p}_n \to p$ en probabilidad.

Estimación de Probabilidades - Ejemplo

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- Nos interesa estimar $F(3) = \mathbb{P}(X \leq 3)$.
- Definimos
$$Y_i = I_{X_i \leq 3} = I_{(-\infty,3]}(X_i)$$
- $Y_i \sim \mathcal{B}(1,p)$, con
$$p = \mathbb{E}(Y_1) = \mathbb{P}(Y_1 = 1) = \mathbb{P}(X_1 \leq 3) = F(3)$$
- Por la ley de los grandes números (caso Bernoulli)
$$\overline{Y}_n \to \mathbb{E}(Y_1) = F(3) \quad \text{en probabilidad}$$
- En tal caso, definimos
$$\widehat{F}(3) = \widehat{F}_n(3) = \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} I_{X_i \leq 3}$$

.

Estimación de $F(t)$

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- Nos interesa estimar $F(t) = \mathbb{P}(X \leq t)$.
- Definimos
$$Y_i = I_{X_i \leq t} = I_{(-\infty, t]}(X_i)$$
- $Y_i \sim \mathcal{B}(1, p)$, con
$$p = \mathbb{E}(Y_1) = \mathbb{P}(Y_1 = 1) = \mathbb{P}(X_1 \leq t) = F(t)$$
- Por la ley de los grandes números (caso Bernoulli)
$$\overline{Y}_n \to \mathbb{E}(Y_1) = F(t) \quad \text{en probabilidad}$$
- En tal caso, definimos
$$\widehat{F}(t) = \widehat{F}_n(t) = \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} I_{X_i \leq t}$$

.

Estimación de Probabilidades.

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- Nos interesa estimar $\mathbb{P}(X \in A)$.
- Definimos
$$Y_i = I_{X_i \in A} = I_A(X_i)$$
- $Y_i \sim \mathcal{B}(1,p)$, con
$$p = \mathbb{E}(Y_1) = \mathbb{P}(Y_1 = 1) = \mathbb{P}(X_1 \in A)$$
- Por la ley de los grandes números (caso Bernoulli)
$$\overline{Y}_n \to \mathbb{E}(Y_1) = \mathbb{P}(X_1 \in A) \quad \text{en probabilidad}$$
- Es decir,
$$\frac{1}{n}\sum_{i=1}^{n} I_{X_i \in A} \longrightarrow \mathbb{P}(X_1 \in A)$$

Demostramos que la frecuencia relativa converge a la probabilidad.