7: Conceptos y aplicaciones de convergencia de variables aleatorias

En la parte de inferencia estadística nos interesará describir algún fenómeno aleatorio.

① Primero asumiremos

$$\text{fenómeno aleatorio} \equiv \mathbb{X} \sim f_{\mathbb{X}}(x|\theta)$$

no conocemos θ, el parámetro que gobierna la forma del comportamiento agregado ideal que esperamos suponiendo para el fenómeno aleatorio.

② En general el interés se centra en

$$\mu = \mathbb{E}(\mathbb{X}) \qquad \text{"el comportamiento medio"}$$

③ Para estimar μ necesitamos datos, $x_1, x_2, \ldots, x_n$

$\Rightarrow$ Asumimos que esos datos son sólo una realización de

$\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n$ v.a. independientes en donde

$$\mathbb{X}_i \sim f_{\mathbb{X}}(x|\theta) \quad , \quad i = 1, 2, 3, \ldots, n.$$

$\equiv$ m.a. $\equiv$ v.a.i.i.d

④ Trataremos de estimar μ con

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

¿Por qué?

$$\mathbb{E}(\overline{X}_n) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i)$$

$$= \mu$$

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

$$\Rightarrow \lim_{n\to\infty} Var(\overline{X}_n) = 0$$

$\Rightarrow$ si $\underline{\underline{n}}$ es muy grande

$$\Rightarrow \underbrace{\overline{X}_n \xrightarrow[n\to\infty]{} \mu}_{\text{converge en algun sentido}}$$

Tenemos que formalizar estas ideas

El que $\overline{X}_n$ converge a μ (en algún sentido) nos da confianza para usar a $\overline{X}_n$ para estimar a μ. Pero y si buscamos una medida en términos de "probabilidades" para saber que tan buena o mala es nuestra estimación

$\Rightarrow$ necesitamos

$$\bar{X}_n \sim ¿?$$

y es fácil ver que si

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

$\mathbb{E}(Z_n) = 0$ y $Var(Z_n) = 1$

$\Rightarrow$ tanto la media y varianza permanecen constantes a medida que $\underline{n}$ aumenta

Esto nos da la oportunidad de estudiar la distribución acumulada de Z_n para $n \to \infty$, en otras palabras

¿?

$$\lim_{n\to\infty} F_{Z_n}(z) = F_Z(z)$$

El Teorema Central de Límite indica que bajo supuestos mínimos

de hecho $Z \sim N(0,1)$

En el primer caso bajo supuestos mínimos

$$\bar{X}_n \xrightarrow{P} \mu$$

En el segundo caso

$$Z_n \xrightarrow{D} Z$$

Para formalizar estos conceptos necesitamos dos desigualdades

$\boxed{\text{Desigualdad de Markov}}$

Si la v.a. X sólo toma valores positivos

$$\Rightarrow P(X \geq \varepsilon) \leq \mathbb{E}(X)/\varepsilon \quad , \quad \forall \varepsilon > 0$$

Dem

$$\mathbb{E}(X) = \int_0^\infty x f(x)\, dx$$

$$= \overbrace{\int_0^\varepsilon x f(x)\, dx}^{\geq 0} + \int_\varepsilon^\infty x f(x)\, dx$$

$$\geq \int_\varepsilon^\infty x f(x)\, dx$$

$$\geq \int_{\varepsilon}^{\infty} \varepsilon f(x)\,dx = \varepsilon P(X \geq \varepsilon)$$

$$\Rightarrow \quad P(X \geq \varepsilon) \leq \mathbb{E}(X)/\varepsilon \quad \square$$

Desigualdad de Chebyshev

Sea X una v.a. con $\mu = \mathbb{E}(X)$ y $\sigma^2 = Var(X)$

$$\Rightarrow P(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}, \quad \forall\, \varepsilon > 0$$

Dem

$$P(|X-\mu| \geq \varepsilon) = P((X-\mu)^2 \geq \varepsilon^2)$$

$$\leq \mathbb{E}((X-\mu)^2)/\varepsilon^2$$

$$= \sigma^2/\varepsilon^2 \quad \square$$

Ley débil de los grandes números (LDGN)

Sea $X_1, X_2, \ldots, X_n$ una m.a. de X y $\mu = \mathbb{E}(X)$.
con $\sigma^2 = Var(X) < \infty$
Para cualquier $\varepsilon > 0$, se tiene que

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| \geq \varepsilon) = 0$$

en otras palabras para n suficientemente grande $\bar{X}_n$ estará concentrado en μ, i.e. $[\mu-\varepsilon, \mu+\varepsilon]$

Dem

$$P(|\bar{X}_n - \mu| \geq \varepsilon) = P((\bar{X}_n - \mu)^2 \geq \varepsilon^2)$$

$$\leq \mathbb{E}((\bar{X}_n - \mu)^2)/\varepsilon^2$$

$$= Var(\bar{X}_n)/\varepsilon^2$$

$$= \sigma^2/n\varepsilon^2$$

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| \geq \varepsilon) = 0$$

$\square$

Ejemplo

Sea p la proporción de votantes que apoyan a un determinado candidato a un puesto de elección popular. Si podemos seleccionar una m.a. de votantes

$$X_1, X_2, \ldots, X_n, \quad X_i \sim \text{Bernoulli}(p)$$

¿ qué tan grande debe ser n para que

$$P(|\bar{X}_n - \mu| \leq 0.01) \geq 0.95 \quad ?$$

$$\Leftrightarrow \quad P(|\bar{X}_n - \mu| \geq 0.01) \leq 0.05$$

$$P(|\bar{X}_n - \mu| \geq 0.01) \leq \frac{Var(\bar{X}_n)}{(0.01)^2} = \frac{p(1-p)}{n(0.01)^2} \leq \frac{1}{4n(0.01)^2}$$

$$\frac{1}{4n(0.01)^2} = 0.05$$

$$n = \frac{1}{4(0.01)^2(0.05)}$$

Convergencia en probabilidad

La convergencia que conocemos de cálculo está dada por

Sea $y_1, y_2, \ldots$ una sucesión de números reales y sea a otro número real. Decimos que

$$\lim_{n \to \infty} y_n = a$$

Si $\forall \varepsilon > 0$, existe n_0 t.q.

$$|y_n - a| \leq \varepsilon \quad \text{para todo } n \geq n_0$$

Convergencia en probabilidad

Sea $Y_1, Y_2, \ldots$ una sucesión de v.a. (no necesariamente independientes) y a un número real. Decimos que Y_n converge en probabilidad a $\underline{a}$, i.e.

$$Y_n \xrightarrow{P} a$$

Si $\forall \varepsilon > 0$, se tiene que

$$\lim_{n \to \infty} P(|Y_n - a| \geq \varepsilon) = 0$$

Desde esta definición, podemos ver que la LDGN simplemente es convergencia en probabilidad

$$\bar{X}_n \xrightarrow{P} \mu$$

Ejemplo

Sea $X_1, \dots, X_n$ una m.a. de $U(0,1)$ y definimos

$$Y_n = \min\{X_1, \dots, X_n\} = X_{(1)}^{n}$$

intuitivamente Y_n se aproxima a cero a medida que n crece. Pero

$$¿\, Y_n \xrightarrow{P} 0 \, ?$$

$$\begin{aligned}
P(|Y_n - 0| \geq \varepsilon) &= P(Y_n \geq \varepsilon) \\
&= P(X_1 \geq \varepsilon, X_2 \geq \varepsilon, \dots, X_n \geq \varepsilon) \\
&= \prod_{i=1}^{n} P(X_i \geq \varepsilon) \\
&= P(X \geq \varepsilon)^n \\
&= \left(1 - P(X \leq \varepsilon)\right)^n \\
&= (1-\varepsilon)^n
\end{aligned}$$

Como $\lim_{n \to \infty} (1-\varepsilon)^n = 0 \Rightarrow Y_n \xrightarrow{P} 0$