01

Clase 31/05/2023

Evaluación de estimadores.

En la clase pasada definimos el ECM (error cuadrático medio) de un estimador $\hat{\theta}$ de un parámetro θ como el número

$$ECM(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2].$$

S tiene la siguiente descomposición

$$\begin{aligned} ECM(\hat{\theta}) &= \mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}] - (\theta - \mathbb{E}[\hat{\theta}])^2] = \mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^2] - 2\mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])(\theta - \mathbb{E}[\hat{\theta}]) + (\hat{\theta} - \mathbb{E}[\hat{\theta}])^2 \\ &= \mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^2] + (\hat{\theta} - \mathbb{E}[\hat{\theta}])^2 = \mathrm{Var}(\hat{\theta}) + (\mathrm{Sesgo}(\hat{\theta}))^2 \end{aligned}$$

Aplicaremos esta evaluación en el caso de dos estimadores de la uniforme en $[0, \theta]$. En primer lugar consideremos una uniforme en $[0, \theta]$ sabemos que su densidad de $\frac{1}{\theta}$ para $0 \leq x \leq \theta$. Si calculamos la esperanza tenemos

$$\mathbb{E}[U_\theta] = \frac{1}{\theta} \int_0^\theta x dx = \frac{\theta}{2}.$$

Para conseguir el estimador de momentos de θ igualamos $\overline{X}_n = \frac{\theta}{2}$ así el estimador resulta $\hat{\theta}_1 = 2\overline{X}_n$. Podemos calcular el sesgo $\mathbb{E}[\hat{\theta}_1] = 2\mathbb{E}[\overline{X}_n] = \theta$. Obteniendo que el estimador es insesgado. Ahora podemos calcular la varianza

$$\mathrm{Var}(\hat{\theta}) = 4\mathrm{Var}(\overline{X}_n) = \frac{4}{n}\mathrm{Var}(U_\theta) = \frac{4}{n}[\frac{1}{\theta}\int_0^\theta x^2 dx - (\frac{\theta}{2})^2] = \frac{\theta^2}{3n}.$$

De esta forma obtenemos $ECM(\hat{\theta}_1) = \frac{\theta^2}{3n}$. Ahora consideremos el estimador de MV que conseguimos en la clase pasada $\hat{\theta}_2 = \max_{1\leq i\leq n} U_i$. Lo primero que tenemos que hacer es calcular la distribución de este estimador que llamaremos $F_n(x) = \mathbb{P}(\max_{1\leq i\leq n} U_i \leq x)$, pero

$$F_n(x) = \mathbb{P}(\cap_{i=1}^n \{U_i \leq x\}) = \prod_{i=1}^n \mathbb{P}(U_{\theta,i} \leq x) = (F_{U_\theta}(x))^n.$$

De esta forma si derivamos obtenemos la densidad

$$f_n(x) = n(F_{U_\theta}(x))^{n-1}\frac{1}{\theta} = n(\frac{x}{\theta})^{n-1}\frac{1}{\theta}, \text{ para } 0 \leq x \leq \theta.$$

Para calcular la esperanza tenemos

$$\mathbb{E}[\hat{\theta}_2] = \frac{n}{\theta^n}\int_0^\theta x^n dx = \frac{n}{n+1}\frac{\theta^{n+1}}{\theta} = \frac{n}{n+1}\theta \to \theta.$$

Igual para el segundo momento

$$\mathbb{E}[(\hat{\theta}_2)] = \frac{n}{\theta^n}\int_0^\theta x^2 x^{n-1} dx = \frac{n}{n+2}\theta^2.$$

Así

$$\mathrm{Var}(\hat{\theta}_2) = \frac{n}{n+2}\theta^2 - (\frac{n}{n+1}\theta)^2 = \frac{n}{(n+2)(n+1)^2}\theta^2.$$

Finalmente $ECM(\hat{\theta}_2) = \frac{n}{(n+2)(n+1)^2}\theta^2 + (\frac{\theta}{n+1})^2 = (\frac{2n+2}{n+2})\frac{\theta^2}{(n+1)^2}$.

Clase 31/05/2023

Hacia el TCL

Visto que podemos tener que un estimador converge al verdadero valor del parámetro que estamos estimando, nos interesa ahora evaluar la calidad de la aproximación. Para ello debemos entonces crear herramientas que permitan calibrar la aproximación, es por eso que recurrimos al Teorema Central del Límite, pero antes de motivarlo recordemos las nociones de convergencia que dimos hace dos clases.

1.- Convergencia en probabilidad: una sucesión de v.a. $\{X_n\}$ converge en probabilidad (que denotaremos por $\mathbb{P}$)

$$X_n \xrightarrow{\mathbb{P}} X \iff \lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0 \quad \forall \varepsilon > 0$$

2.- Diremos que $\{X_n\}_{n=1}^{\infty}$ converge en media cuadrática a X si se tiene

$$E[X_n - X]^2 \longrightarrow 0 \quad \text{cuando } n \to \infty$$

Por la desigualdad de Chebyshev tenemos que

$$E|X_n - X|^2 \longrightarrow 0 \implies X_n \xrightarrow{\mathbb{P}} X$$

3.- Una sucesión de v.a. $\{X_n\}_{n=1}^{\infty}$ converge casi seguramente a X, si existe un conjunto $N \subseteq \Omega$ tal que $\mathbb{P}(N) = 0$ y si $\omega \notin N$ entonces

$$X_n(\omega) \to X(\omega).$$

4.- Finalmente tenemos la convergencia en distribución

Diremos que X_n converge en distribución a X que denotaremos

$X_n \xrightarrow{d} X$ si $\forall x$ punto de continuidad de $F_X(x)$ la función de distribución de X, se tiene que $F_{X_n}(x) \longrightarrow F_X(x)$.

Se tiene las implicaciones siguientes

$$\begin{tikzcd}
 & \boxed{\text{Conv. c.s}} \arrow[d, Rightarrow] \\
\boxed{\text{CMC}} \arrow[r] & \boxed{\text{Probabilidad}} \arrow[d, Rightarrow] \\
 & \boxed{\text{Convergencia en d.}}
\end{tikzcd}$$

Sabemos que si $\{X_n\}_{n=1}^{\infty}$ es una muestra aleatoria simple de una v.a. con $E|X_n|^2 < +\infty$ entonces

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n} \xrightarrow{LGN} E(X) = \mu.$$

Esta ley admite una generalización, pero la demostración excede el nivel de este curso.

Si $E(|X_n|^2) < \infty$ entonces.

$$\overline{X}_n \xrightarrow{c.s.} E(X)$$

Este teorema es debido a Kolmogorov y constituye una de los éxitos de la teoría axiomática de la probabilidad.

Sabemos lo siguiente sobre la media muestral

$$E(\overline{X}_n) = \frac{n\,E(X)}{n} = \mu .$$

y podemos calcular

$$Var(\overline{X}_n) = E(\overline{X}_n - \mu)^2 = \frac{1}{n^2} E\left[\sum_{j=1}^{n} (X_j - \mu)\right]^2$$

$$= \frac{1}{n^2}\; n\, Var(X_j) .$$

Pero cada una de las variables tiene la misma varianza que denotaremos por σ^2; luego

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

y esta igualdad es la que nos permite concluir que

$\overline{X}_n \xrightarrow{P} E(X) = \mu$ y como mencionamos antes el siguiente resultado también es cierto $\overline{X}_n \xrightarrow{cs} E(X) = \mu$.

Podemos escribir

$$\sqrt{n}\,(\overline{X}_n - \mu) = \sqrt{n}\left(\frac{1}{n}\sum_{j=1}^{n} X_j - \mu\right)$$

$$= \sqrt{n}\left(\frac{1}{n}\sum_{j=1}^{n} (X_j - \mu)\right)$$

$$= \frac{1}{\sqrt{n}}\sum_{j=1}^{n} (X_j - \mu) := \frac{1}{\sqrt{n}} S_n$$

El teorema central del límite tiene que ver con la convergencia de

$$F_n(x) = P\left(\frac{1}{\sqrt{n}}\frac{S_n}{\sigma} \leq x\right)$$

La motivación para estudiar el comportamiento de $F_n(x)$ proviene históricamente de la Teoría de los errores.

En diferentes disciplinas (Agrimensura, Joyería, construcciones, Navegación) cuando se necesitaba medir algo con precisión, se acostumbraba a medir el objeto o el individuo, etc. varias veces y se obtenían un número n de mediciones

$X_1, X_2, \ldots, X_n$ se consideraba que para "disminuir el error" se debía tomar el promedio de las medidas

$\bar{X}_n = \dfrac{X_1 + \ldots + X_n}{n}$ y se postulaba que ese promedio daba un número más cercano al verdadero valor.

Se justificaba esta afirmación recurriendo a la obra de dios.

De Moivre en el siglo $\boxed{XVII}$ estableció, por primera vez, el teorema en el caso en que las v.a. son Bernoulli con parámetro p, en este caso

$$S_n = B_1 + \cdots + B_n \text{ es una Binomial con}$$

parámetro (n, p)

$$y \quad \frac{S_n}{n} \to p \quad (LGN),$$

de esta manera $\left(\dfrac{S_n}{n} - p\right) = \dfrac{S_n - np}{n}$

De Moivre demostró que

$\cancel{F_n(x)\ \sqrt{n}\left(\frac{S_n}{n}-p\right)\leq x}$

$$F_n(x) = \sqrt{n}\left(\frac{S_n}{n}-p\right) \to \Phi\left(\frac{x}{\sigma^2}\right)$$

donde Φ es la distribución gaussiana y $\boxed{\sigma^2 = p(1-p)}$.

Veamos un ejemplo particular de este teorema que tiene el interés de que el resultado es exacto.

Supongamos que cada v.a. $X_j \sim N(\mu,\sigma^2)$, entonces tenemos que $Z_j = \frac{X_j-\mu}{\sigma} \sim N(0,1)$

y sabemos que

$$S_n = Z_1 + \cdots + Z_n \quad \text{es gaussiana}$$

con media cero y varianza n esto es

$$Var\left(\frac{S_n}{\sqrt{n}}\right) = \frac{1}{n}\, Var(S_n) = 1.$$

Esto implica que

$$\frac{S_n}{\sqrt{n}} \sim N(0,1) \quad \text{es decir}$$

$$F_n(x) = P\left(\frac{S_n}{\sqrt{n}} \leq x\right) = P\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n} \frac{X_j-\mu}{\sigma} \leq x\right)$$

$= \Phi(x)$ o lo que es lo mismo

$$P\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(X_j-\mu)\leq \sigma x\right)=\Phi(x)\ (*)$$

También

$$(*) = P\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}X_j-\sqrt{n}\mu\leq \sigma x\right)$$

$$= P\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}X_j \leq \sigma x+\sqrt{n}\mu\right)=\Phi(x)$$

si hacemos el cambio de variable

$$\sigma x+\sqrt{n}\mu = y$$

$$x = \frac{y-\sqrt{n}\mu}{\sigma}$$

tenemos

$$\boxed{P\left(\frac{S_n}{\sqrt{n}}\leq y\right)=\Phi\left(\frac{y-\sqrt{n}\mu}{\sigma}\right)}$$

y en este caso esta igualdad vale para todo n siempre que las v.a. que sumamos sean $N(\mu,\sigma^2)$ e independientes.

Lo interesante es que si suponemos que

1. Las X_j tienen la misma distribución
2. Son independientes entre si
3. $E(X_j)=\mu$ y $Var(X_j)=\sigma^2$

Entonces Lindeberg demostró, extendiendo resultados anteriores como el de De Moivre, que.

Si definimos $F_n(x) = P\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{(X_j-\mu)}{\sigma}\leq x\right)$

entonces

$$\boxed{F_n(x)\rightarrow\Phi(x)}$$

o usando nuestra definición de convergencia en distribución

$$\boxed{\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{(X_j-\mu)}{\sigma}\xrightarrow{d}N(0,1)}$$

También se tiene que si $a\leq b$.

$$P\left(a\leq\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\frac{X_j-\mu}{\sigma}\right)\leq b\right)\rightarrow\Phi(b)-\Phi(a)$$

Lo que podemos escribir también

$$=P\left(\sqrt{n}\sigma a\leq\sum_{j=1}^{n}(X_j-\mu)\leq\sqrt{n}\sigma b\right)$$

$$=P\left(\sqrt{n}\sigma a+n\mu\leq\sum_{j=1}^{n}X_j\leq\sqrt{n}\sigma b+n\mu\right)$$

$$=P\left(\frac{\sigma a}{\sqrt{n}}+\mu\leq\bar{X}_n\leq\frac{\sigma b}{\sqrt{n}}+\mu\right)$$

$$=\Phi(b)-\Phi(a)$$

~~[illegible]~~

Lo que es equivalente a que

$$\mathbb{P}\left(\overline{X}_n - \frac{\sigma b}{\sqrt{n}} \leq \mu \leq \overline{X}_n - \frac{\sigma}{\sqrt{n}} a \right) = \Phi(b) - \Phi(a)$$

$$\frac{\sigma b}{\sqrt{n}} = b_1 \Rightarrow b = \frac{\sqrt{n}\, b_1}{\sigma}$$

$$\frac{\sigma a}{\sqrt{n}} = a_1 \Rightarrow a = \frac{\sqrt{n}\, a_1}{\sigma}$$

Luego

$$\mathbb{P}\left(\overline{X}_n - b_1 \leq \mu \leq \overline{X}_n - a_1 \right) = \Phi\left(\frac{\sqrt{n}\, b_1}{\sigma}\right) - \Phi\left(\frac{\sqrt{n}\, a_1}{\sigma}\right)$$

Aquí se ve la génesis de los intervalos de confianza un tema que debemos discutir el próximo lunes.
