Clase 24/05/2023.

Convergencia de variables aleatorias

La semana pasada introdujimos la noción de convergencia en probabilidad: Diremos que una sucesión $\{X_n\}_{n=1}^{\infty}$ de variables aleatorias $\boxed{\text{converge en probabilidad}}$ a la v.a X si se tiene que

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

En lenguaje ~~xxxxx~~ coloquial la convergencia en probabilidad nos dice que $X_n(\omega)$ se aproxima a X en un conjunto de probabilidad grande o lo que es lo mismo que X_n está separada de X a lo sumo por $\varepsilon > 0$ en un conjunto de probabilidad pequeña.

$\boxed{\text{Convergencia en media cuadrática}}$ (o en L^2)

Una sucesión de v.a. $\{X_n\}_{n=1}^{\infty}$ converge en media cuadrática si

$$E\,|X_n - X|^2 \xrightarrow[n\to\infty]{} 0 .$$

La desigualdad de Chebyshev nos permite demostrar la siguiente implicación

Convergencia en media cuadrática $\Rightarrow$ Converg en Proba

Otra noción importante es la convergencia casi segura c.s.

$\boxed{\text{Convergencia casi segura}}$

$X_n \to X$ c.s. si existe un subconjunto N tal que $P(N) = 0$ y si $\omega \notin N$ entonces $X_n(\omega) \to X(\omega)$

Aunque no lo demostraré se tiene que

$$X_n \xrightarrow{cs} X \implies X_n \to X \text{ en probabilidad}$$

Finalmente está la noción de convergencia en distribución

$X_n \xrightarrow{d} X$ (en distribución) si $F_n(x)$ y $F(x)$ denotan respectivamente las distribuciones de X_n y de X; esto es $F_n(x) = P(X_n \leq x)$ y $F(x) = P(X \leq x)$. Entonces

$X_n \xrightarrow{d} X$ si $\forall x$ de continuidad de $F(x)$ se verifica $F_n(x) \to F(x)$

Convergencia en probabilidad $\implies$ Convergencia en distribución

Tenemos el siguiente esquema

$$
\begin{array}{ccc}
 & & \text{C.S} \\
 & & \boxed{\text{Convergencia C.S}} \\
 & & \Downarrow \\
\text{CMC.} & & \text{P} \\
\boxed{\text{Convergencia en media cuadrática}} & \Rightarrow & \boxed{\text{Convergencia en probabilidad}} \\
 & & \Downarrow \\
 & & \text{d.} \\
 & & \boxed{\text{Convergencia en distribución}}
\end{array}
$$

Veamos que si $X_n \xrightarrow{MC} X \Rightarrow X_n \xrightarrow{P} X$. En efecto.

$$P(|X_n - X| > \varepsilon) \leq \frac{1}{\varepsilon^2} E(|X_n - X|^2) \longrightarrow 0.$$

pues el lado derecho tiende a cero por la convergencia en MC.

Las otras implicaciones son más laboriosas de demostrar

Ahora veamos un ejemplo muy sencillo sean $\{Y_n\}_{n=1}^{\infty}$ una colección de v.a. $N(0, \sigma_n^2)$ entonces si $\sigma_n^2 \to \sigma^2$

$$Y_n \xrightarrow{d} Z = N(0, \sigma^2).$$

En efecto

$$F_n(x) = \int_{-\infty}^{x} \frac{e^{-u^2/2\sigma^2}}{\sqrt{2\pi}\,\sigma}\,du$$

hacemos en la integral $\frac{u}{\sigma_n} = v$ obtenemos así

$$F_n(x) = \int_{-\infty}^{x/\sigma_n} \frac{e^{-v^2/2}}{\sqrt{2\pi}}\,v = \Phi\left(\frac{x}{\sigma_n}\right);$$

pero Φ (es una función continua y entonces

$$\Phi\left(\frac{x}{\sigma_n}\right) \longrightarrow \Phi\left(\frac{x}{\sigma}\right)$$

luego $\boxed{F_n(x) \longrightarrow Z = N(0,\sigma^2)}$

Nuevo tema: Estimación

Consideremos el siguiente problema tenemos que ~~un~~ modelo para modelar un fenómeno donde interviene el azar tenemos a disposición una densidad (o más bien una familia de densidades) que dependen de uno o varios parámetros. Por ejemplo nuestro modelo es la gaussiana

$N(\mu,\sigma^2)$ y supongamos que disponemos de un conjunto de observaciones tomadas de una población de manera independiente

$X_1(\omega), X_2(\omega), \dots, X_n(\omega) \dots$

y suponemos que dichas observaciones provienen de un mismo modelo. Entonces nos preguntamos ¿a partir de estas observaciones como inferir el valor de los parámetros?

Tomemos como ejemplo la $N(\mu, 1)$ una familia con densidad

$$\boxed{\frac{e^{-\frac{(x-\mu)^2}{2}}}{\sqrt{2\pi}}}$$

Podemos entonces definir a partir de las observaciones la media empírica

$$\boxed{\bar{X} = \frac{X_1(\omega) + X_2(\omega) + \dots + X_n(\omega)}{n}}$$

Sabemos también que si $Z \sim N(\mu, 1)$

$$E(Z) = \mu$$

Entonces por la LGN tenemos:

$\bar{X}_n \xrightarrow{P} E(X_i) = \mu$ luego sabemos que

Construimos a partir de $X_1, \dots, X_n$ la media empírica sabemos que nos aproximamos en probabilidad al verdadero valor del parámetro.

Otro ejemplo es el de la varianza de una gaussiana $N(2,\sigma^2)$ (conocemos de la esperanza que es igual a 2).

Podemos entonces definir

$$\overline{m}_2 = \frac{X_1^2 + X_2^2 + \dots + X_n^2}{n}$$

momento empírico de segundo orden si

$Z \sim N(2,\sigma^2)$ su densidad es

$$\frac{e^{-\frac{(x-2)^2}{2\sigma^2}}}{\sqrt{2\pi}\,\sigma}$$

Al calcular el segundo momento

$$m_2 = \int_{-\infty}^{\infty} x^2 \frac{e^{-\frac{(x-2)^2}{2\sigma^2}}}{\sqrt{2\pi}\,\sigma}\, dx = \int_{-\infty}^{\infty} (\sigma v + 2)^2 \frac{e^{-v^2/2}}{\sqrt{2\pi}}\, dv$$

$m_2 = \sigma^2 + 4$ y entonces tenemos

$$\hat{m}_{2,n} = \frac{X_1^2 + \dots + X_n^2}{n}$$

$$E\,\hat{m}_{2,n} = \frac{n\sigma^2 + n4}{n} = \sigma^2 + 4$$

Luego $$\hat{m}_{2,n} - 4 = \frac{(X_1^2 - 4) + (X_2^2 - 4) + \dots + (X_n^2 - 4)}{n}$$

$$\boxed{E\hat{m}_{2,n} - 4 = \sigma^2}$$

pero además la LGN implica que

$$\boxed{\hat{m}_{2,n} - 4 \longrightarrow \sigma^2}$$

Estadístico y estimador

Un estadístico T es una variable aleatoria que se calcula en función de las variables $X_1, \dots, X_n$ y en general lo escribimos

$$\boxed{T(X_1, X_2, \dots, X_n)}.$$

Un estimador es un estadístico, que creemos que contiene relevante θ de la distribución común de la $\{X_i\}$ que denotamos por X.

Método de estimación

1 Método de los momentos

Para ilustrarlo trabajemos el siguiente ejemplo.
Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria simple (o muestra aleatoria) que proviene de una densidad

$$f(x;\theta) = (\theta+1)x^{\theta} \quad \text{para } x \in [0,1]$$

Calculemos la esperanza

$$E(X) = (\theta+1)\int_0^1 x^{\theta+1}dx = \frac{(\theta+1)}{(\theta+2)}$$

Esto quiere decir que la esperanza de X es una función del parámetro $g(\theta)$ en nuestro caso $\boxed{g(\theta) = \frac{(\theta+1)}{(\theta+2)}}$

Si calculamos la media muestral $\overline{X}_n$ el estimador de momentos consiste en igualar la media empírica a $g(\theta)$ y despejar. Esto es

$$\overline{X}_n = \frac{\theta+1}{\theta+2}$$

y el estimador será entonces

$$\hat{\theta} = g^{-1}(\overline{X}_n)$$

que en este caso podemos calcular sin problemas.

$$\overline{X}_n(\hat{\theta}+2) = \hat{\theta}+1$$

$$\overline{X}_n\hat{\theta} - \hat{\theta} = 1 - 2\overline{X}_n$$

luego $\hat{\theta}(\overline{X}_n - 1) = 1 - 2X_n$

$$\Rightarrow \boxed{\hat{\theta} = \frac{1-2X_n}{X_n-1}}$$

El método funciona igual cuando hay más parámetros

$$\overline{m}_1 = \overline{X}_n$$

$$\overline{m}_2 = \frac{X_1^2 + \cdots + X_n^2}{n}$$

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

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

$$\overline{m}_2 = \frac{X_1^2 + \cdots + X_n^2}{n} = \sigma^2 + \mu^2$$

$$\Rightarrow \quad \hat{\sigma}^2 = \overline{m}_2 - \hat{\mu}^2 = \overline{m}_2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2$$

$$\boxed{\hat{\sigma}^2 = \overline{m}_2 - \overline{m}_1^2}.$$

Noción de consistencia:

Diremos que un estimador es consistente en probabilidad si

$$T_n(X_1, \dots, X_n) \xrightarrow{P} \theta \quad \left(\begin{array}{c}\text{verdadero valor} \\ \text{del parámetro}\end{array}\right)$$

y será consistente en norma cuadrática si

$$E\,(T_n - \theta)^2 \xrightarrow[n \to \infty]{} 0.$$

Todo estimador consistente en media cuadrática lo es en probabilidad, pero el recíproco no siempre es cierto.

Ejemplo:

$$\boxed{\bar{X}_n \xrightarrow{P} E(X) = \mu}$$

caso gaussiano

$$\bar{m}_2 = \frac{X_1^2 + \cdots + X_n^2}{n} \xrightarrow{P} E(X^2) = \sigma^2 + \mu^2$$

luego como se tiene que si $X_n \xrightarrow{P} X$ e $Y_n \xrightarrow{P} Y$

$$\boxed{X_n + Y_n \xrightarrow{P} X + Y}$$

y si f es una función continua y $X_n \xrightarrow{P} X$

$$\Rightarrow \boxed{f(X_n) \xrightarrow{P} f(X)}$$

Entonces:

En el caso Gaussiano

$$\begin{cases} \overline{m}_2 \xrightarrow{P} \sigma^2 + \mu^2 \\ (\overline{m}_1)^2 \xrightarrow{P} \mu^2 \end{cases}$$

Entonces usando el teorema

$$\boxed{\overline{m}_2 - (\overline{m}_1)^2 \longrightarrow \sigma^2}$$

Veamos una igualdad interesante. Definamos el estadístico

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

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

$$= \frac{1}{n}\sum_{j=1}^{n} X_j^2 - \overline{X}_n^2 = \overline{m}_2 - (\overline{m}_1)^2$$