Práctico 8

martes, 17 de mayo de 2022

Desigualdad de Chebyshev:

X variable aleatoria

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

Ejercicio 1

Se realiza una encuesta para tratar de estimar el porcentaje de votos de determinado candidato. Al final de la misma se tendrán n respuestas, representadas en la muestra $X_1, X_2, \cdots, X_n$ de realizaciones de la variable de Bernoulli de parámetro p (la probabilidad de que una persona elegida al azar en la población vote por el candidato en cuestión).

1. A partir de la mencionada muestra ¿Qué valor usaría usted para estimar p?.
2. Usando la Desigualdad de Chevishoff (y antes de tomar la muestra) ¿Cuántos datos tomaría para que, con una probabilidad de al menos 0,95, el estimador no distara del valor de p más de un 0,03?. (Sugerencia: Demuestre y use que para cualquier valor de $p \in [0,1]$ se cumple $p(1-p) \leq 1/4$).

$X_1, X_2, \ldots, X_n$ iid $\sim Ber(p)$

$X_1 \sim Ber(p)$: X_1 vale 1 si la persona vota al candidato

X_1 vale 0 si no lo vota

$$P(X_1 = 1) = p$$

$$P(X_1 = 0) = 1-p$$

$$E(X_1) = 1\underbrace{P(X_1 = 1)}_{p} + 0\,P(X_1 = 0) = p$$

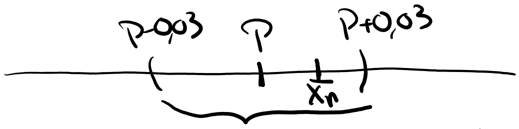
$$E(X_1^2) = 1^2\underbrace{P(X_1 = 1)}_{p} + 0^2 P(X_1 = 0) = p$$

$$Var(X_1) = E(X_1^2) - E(X_1)^2 = p - p^2 = p(1-p)$$

1) Para estimar p usaríamos $\dfrac{X_1 + X_2 + \dots + X_n}{n} = \overline{X_n}$

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

2)



queremos que $\overline{X_n}$ esté
en este intervalo con probabilidad $\geq 0,95$

$$P(|\overline{X_n} - p| \leq 0,03) \geq 0,95$$

Chebyshev: $P(|\overline{X_n} - E(\overline{X_n})| \geq \varepsilon) \leq \dfrac{Var(\overline{X_n})}{\varepsilon^2}$

$$E(\overline{X_n}) = E\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$$

$$= \frac{1}{n} E(X_1 + X_2 + \dots + X_n)$$

$$= \frac{1}{n}\left(\underbrace{E(X_1)}_{=p} + \underbrace{E(X_2)}_{=p} + \dots + \underbrace{E(X_n)}_{=p}\right)$$

$$= \frac{1}{n} np$$

$$= p$$

$$Var(\overline{X_n}) = Var\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$$

$$= \frac{1}{n^2} Var(X_1 + \dots + X_n)$$

porque las X_i son independientes

$$= \frac{1}{n^2}\left(\underbrace{Var(X_1)}_{p(1-p)} + \dots + \underbrace{Var(X_n)}_{p(1-p)}\right)$$

$$= \frac{1}{n^2}\, n\, p(1-p)$$

$$= \frac{p(1-p)}{n}$$

$$P(|\overline{X}_n - p| \leq 0{,}03) \geq 0{,}95 \iff P(|\overline{X}_n - p| > 0{,}03) \leq 0{,}05$$

$$P(|\overline{X}_n - p| > 0{,}03) \underset{\uparrow \text{Chebyshev}}{\leq} \frac{Var(\overline{X}_n)}{(0{,}03)^2} = \frac{p(1-p)}{n(0{,}03)^2} \leq \frac{1}{4n(0{,}03)^2}$$

si $p \in [0,1] \Rightarrow p(1-p) \leq \frac{1}{4}$

queremos $P(|\overline{X}_n - p| > 0{,}03) \leq 0{,}05$

entonces tomamos n tal que $\frac{1}{4n(0{,}03)^2} \leq 0{,}05$

$$\frac{1}{4n(0{,}03)^2} \leq 0{,}05 \iff \frac{1}{n} \leq 4(0{,}05)(0{,}03)^2$$

$$\iff n \geq \frac{1}{4(0{,}05)(0{,}03)^2} \approx 5555{,}5\ldots$$

entonces $\boxed{n \geq 5556}$

$p \in [0,1]$ entonces $p(1-p) \leq \frac{1}{4}$

$f(p) = p(1-p) = p - p^2$

$f'(p) = 1 - 2p \rightsquigarrow f$ tiene un máximo relativo en $\frac{1}{2}$

$f(\frac{1}{2}) = \frac{1}{2}(1 - \frac{1}{2}) = \frac{1}{4} \Rightarrow p(1-p) \leq \frac{1}{4}$

Convergencia casi segura

$X_1, X_2, X_3, \ldots$ sucesión de variables aleatorias

$a \in \mathbb{R}$

Decimos que X_n converge casi seguramente a a si

$$P(\{X_n \to a\}) = 1$$

escribimos $X_n \xrightarrow{cs} a$

$$\begin{array}{cccccc} X_1 & X_2 & X_3 & X_4 & \cdots & X_n \cdots \\ \downarrow & \downarrow & \downarrow & \downarrow & & \downarrow \\ a_1 & a_2 & a_3 & a_4 & & a_n \end{array}$$

la probabilidad de que el límite de a_n sea a es 1

Ejercicio 2
Demostrar que si $X_n \xrightarrow[n]{c.s.} a$ e $Y_n \xrightarrow[n]{c.s.} b$ y $g : \mathbb{R} \to \mathbb{R}$ continua entonces

1. $X_n + Y_n \xrightarrow[n]{c.s.} a + b$
2. $X_n Y_n \xrightarrow[n]{c.s.} ab$
3. $g(X_n) \xrightarrow[n]{c.s.} g(a)$.

- si el evento A implica el evento B y $P(A) = 1$ entonces $P(B) = 1$

  $A \subset B$

- si $P(A) = 1$ y $P(B) = 1 \Rightarrow \begin{cases} P(A \cup B) = 1 \\ P(A \cap B) = 1 \end{cases}$

$$\underbrace{P(A \cup B)}_{1} = \underbrace{P(A)}_{1} + \underbrace{P(B)}_{1} - P(A \cap B)$$

$$P(A\cap B) = -\underbrace{P(A\cup B)}_{=1} + \underbrace{P(A)}_{=1} + \underbrace{P(B)}_{=1}$$

$$\Rightarrow P(A\cap B) = 1$$

1. $X_n \xrightarrow{cs} a$

$Y_n \xrightarrow{cs} b$

queremos probar que $X_n + Y_n \xrightarrow{cs} a+b$

o sea queremos ver que $P(X_n + Y_n \longrightarrow a+b) = 1$

$X_n \xrightarrow{cs} a \quad \Rightarrow \quad P(X_n \longrightarrow a) = 1$

$Y_n \xrightarrow{cs} b \quad \Rightarrow \quad P(Y_n \longrightarrow b) = 1$

entonces $P(\{X_n \rightarrow a\} \cap \{Y_n \rightarrow b\}) = 1$

tenemos que

$\{X_n \rightarrow a\} \cap \{Y_n \rightarrow b\} \subset \{X_n + Y_n \longrightarrow a+b\}$

entonces $P(\{X_n + Y_n \longrightarrow a+b\}) = 1$

$$\Rightarrow X_n + Y_n \xrightarrow{cs} a+b$$

2. $X_n \xrightarrow{cs} a$

$Y_n \xrightarrow{cs} b$

queremos ver que $X_n Y_n \xrightarrow{cs} ab$

o sea queremos ver que $P(X_n Y_n \longrightarrow ab) = 1$

$X_n \xrightarrow{cs} a \quad \Rightarrow \quad P(X_n \rightarrow a) = 1$

$Y_n \xrightarrow{cs} b \quad \Rightarrow \quad P(Y_n \longrightarrow b) = 1$

entonces $P(\{X_n \to a\} \cap \{Y_n \to b\}) = 1$

tenemos que

$\{X_n \to a\} \cap \{Y_n \to b\} \subset \{X_n Y_n \longrightarrow ab\}$

entonces $P(X_n Y_n \longrightarrow ab) = 1$

$\Rightarrow \quad X_n Y_n \xrightarrow{cs} ab$

(3) $X_n \xrightarrow{cs} a$

$g: \mathbb{R} \to \mathbb{R}$ función continua

queremos ver que $g(X_n) \xrightarrow{cs} g(a)$

$X_n \xrightarrow{cs} a \quad \Rightarrow \quad P(X_n \longrightarrow a) = 1$

$\left.\begin{array}{l} X_n \longrightarrow a \\ g \text{ función continua} \end{array}\right\} \Rightarrow \quad g(X_n) \longrightarrow g(a)$

tenemos $\{X_n \to a\} \subset \{g(X_n) \to g(a)\}$

entonces $P(g(X_n) \to g(a)) = 1$

$\Rightarrow \quad g(X_n) \xrightarrow{cs} g(a)$

Ley fuerte de los grandes números

$X_1, X_2, X_3, \ldots$ sucesión de variables aleatorias iid

Definimos una nueva sucesión de variables aleatorias

$\sum^{n}$

$$\overline{X_n} := \frac{1}{n}\sum_{i=1}^{n} X_i$$

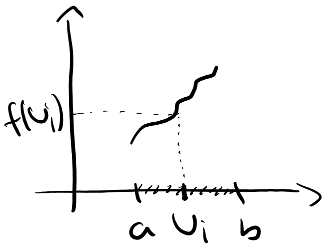
$$\text{LFGN: } \overline{X_n} \xrightarrow{cs} E(X_1)$$

Ejercicio 3
Este ejercicio describe el *método de Montecarlo* para el cálculo de integrales.

1. Sean $(U_i)_{i\in\mathbb{N}} \sim \mathcal{U}[a,b]$ *iid* y $f \in R[a,b]$ (f es integrable Riemann en $[a,b]$), mostrar que:

$$\frac{1}{n}\sum_{i=1}^{n} f(U_i) \xrightarrow[n]{c.s.} \frac{1}{b-a}\int_a^b f(x)\,dx.$$

$$\frac{1}{n}\sum_{i=1}^{n} f(U_i)$$



$$f_{U_1}(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a,b] \\ 0 & \text{si no} \end{cases}$$

Definimos $Y_i = f(U_i)$

$$\frac{1}{n}\sum_{i=1}^{n} f(U_i) = \frac{1}{n}\sum_{i=1}^{n} Y_i = \overline{Y_n} \xrightarrow{cs} E(Y_1)$$

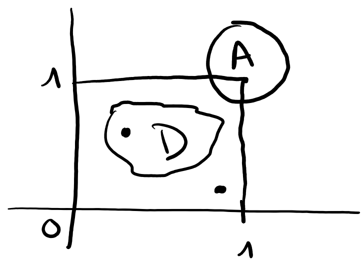
$$E(Y_1) = E(f(U_1)) = \int_{-\infty}^{+\infty} f(x) f_{U_1}(x)\,dx = \int_a^b f(x)\frac{1}{b-a}\,dx = \frac{1}{b-a}\int_a^b f(x)\,dx$$

$$\frac{1}{n}\sum_{i=1}^{n} f(U_i) = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow[\uparrow \text{ LFGN}]{cs} E(Y_1) = \frac{1}{b-a}\int_a^b f(x)\,dx$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} f(U_i) \xrightarrow{cs} \frac{1}{b-a}\int_a^b f(x)\,dx$$

2. Sea D una región arbitraria de $[0,1]\times[0,1]$ y sean $U_1,\ U_2,\ldots,U_n$ variables aleatorias independientes e idénticamente distribuidas con distribución $\mathcal{U}([0,1]\times[0,1])$, es decir que se cumple que $\mathbf{P}(U\in A) = \text{área}(A\cap[0,1]\times[0,1])$.

Si $a_n = \dfrac{\#\{i : 1\leqslant i\leqslant n \text{ y } U_i\in D\}}{n}$ probar que $a_n \xrightarrow[n]{c.s.} \text{área}(D)$.



$U_1,\ldots,U_n$ iid $\sim \mathcal{U}([0,1]\times[0,1])$

$P(U_1\in A) = \text{area}(A\cap([0,1]\times[0,1]))$

$$a_n = \frac{\#\{i : 1\leq i\leq n \ y\ U_i\in D\}}{n}$$

queremos ver que $a_n \xrightarrow{cs} \text{area}(D)$

$$X_i = \begin{cases} 1 & \text{si } U_i\in D \\ 0 & \text{si no} \end{cases}$$

$X_i \sim Ber(P(U_i\in D))$

$P(U_i\in D) = \text{area}(D)$

$E(X_i) = \text{area}(D)$

$$a_n = \frac{X_1+X_2+\cdots+X_n}{n} = \overline{X}_n \xrightarrow{cs} E(X_1) = \text{area}(D)$$

LFGN