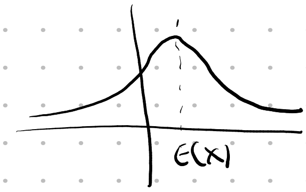
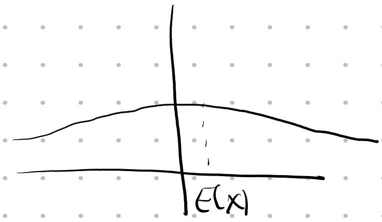
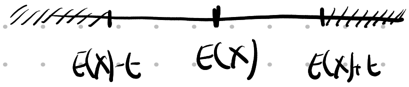
Desigualdad de Chebyshev

X variable aleatoria



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





$$|X-E(X)|\geq t \begin{cases} X \geq E(X)+t \\ X \leq E(X)-t \end{cases}$$

→ "cuanto más pequeña es la varianza menos probable es que X caiga lejos de la esperanza"

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 (que asumimos independientes), 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 $p(1-p) \leq 1/4$).

$X_1 \sim Ber(p)$ $\quad p$ = probabilidad de que la persona 1 vote al candidato

X_1 → vale 1 si la persona 1 vota al candidato
→ vale 0 si no lo vota

$X_i \sim Ber(p)$ $\quad p$ = probabilidad de votar al candidato

X_i → vale 1 si la persona i vota al candidato
→ vale 0 si no

$$p_{X_i}(1) = P(X_i = 1) = p$$

$$p_{X_i}(0) = P(X_i = 0) = 1-p$$

$$E(X_i) = 1\, p_{X_i}(1) + 0\, p_{X_i}(0) = p$$

$$E(X_i^2) = 1^2 p_{X_i}(1) + 0^2 p_{X_i}(0) = p$$

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

$X_i \sim Ber(p)$ $\quad E(X_i) = p$ y $Var(X_i) = p(1-p)$

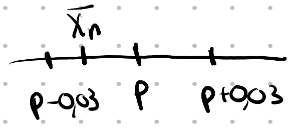
1. $X_1 + X_2 + X_3 + \dots + X_n$ = cantidad de personas que votan al candidato

para estimar p vamos a usar:

$$\overline{X_n} := \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

2. buscamos n para que nuestra estimación esté a menos de 0,03 del valor real con una probabilidad $\geq 0{,}95$

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



<u>Chebyshev</u>: $P(|Y - E(Y)| \geq t) \leq \dfrac{Var(Y)}{t^2}$

vamos a calcular $E(\overline{X_n})$ y $Var(\overline{X_n})$:

linealidad de la esperanza

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

$X_i \sim Ber(p)$

$E(X_i) = p$

$Var(X_i) = p(1-p)$

linealidad de la esperanza $\rightarrow$

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

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

$$= p$$

$$Var(\overline{X_n}) = Var\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n^2} Var(X_1 + X_2 + \dots + X_n)$$

si X e Y son <u>independientes</u>

$$\boxed{\begin{array}{c} Var(X+Y) = Var(X) + Var(Y) \\ Var(cX) = c^2 Var(X) \end{array}}$$

$$= \frac{1}{n^2}\left(\underbrace{Var(X_1)}_{p(1-p)} + \underbrace{Var(X_2)}_{p(1-p)} + \dots + \underbrace{Var(X_n)}_{p(1-p)}\right)$$ porque las X_i son independientes

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

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

$$E(\overline{X}_n) = p \quad y \quad Var(\overline{X}_n) = \frac{p(1-p)}{n}$$

Chebyshev: $P(|Y - E(Y)| \geq t) \leq \frac{Var(Y)}{t^2}$

buscamos n tal que

$$P(|\overline{X}_n - \underset{\substack{\| \\ E(\overline{X}_n)}}{p}| \leq 0{,}03) \geq 0{,}95$$

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

entonces por Chebyshev

$$P(|\overline{X}_n - \underbrace{p}_{E(\overline{X}_n)}| \geq 0{,}03) \leq \frac{Var(\overline{X}_n)}{(0{,}03)^2} = \frac{p(1-p)}{n(0{,}03)^2} \leq \frac{1}{4n(0{,}03)^2}$$

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

buscamos n tal que

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

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

$$n \geq \frac{1}{0{,}05 \cdot (0{,}03)^2 \cdot 4} = 5555{,}5$$

tomamos 5556

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

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

$f'(p) = 1 - 2p \rightsquigarrow f'(\frac{1}{2}) = 0 \Rightarrow f$ tiene un extremo en $\frac{1}{2}$

$f''(p) = -2 \rightsquigarrow f''(\frac{1}{2}) = -2 < 0 \Rightarrow f$ tiene un máximo en $\frac{1}{2}$

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

Convergencia casi segura

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

decimos que X_n converge casi seguramente a $a \in \mathbb{R}$ si

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

escribimos $X_n \xrightarrow{cs} a$

sucesión de va $\to$ X_1 X_2 X_3 $\ldots$ X_n $\ldots$

sucesión de números reales $\to$ $a_1 \in R_{X_1}$ $a_2 \in R_{X_2}$ $a_3 \in R_{X_3}$ $\ldots$ $a_n \in R_{X_n}$

$\hookrightarrow$ número real

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. $g(X_n) \xrightarrow[n]{c.s.} g(a)$.

* $A \subset B$ y $\overbrace{P(A) = 1}^{A\ cs} \Rightarrow P(B) = 1$
* sean A y B eventos tales que $P(A) = 1$ y $P(B) = 1$

$P(A \cup B) = 1$ porque $A \subset A \cup B$ y $P(A) = 1$

$P(A \cap B) = 1$

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

① $X_n \xrightarrow{cs} a$, $Y_n \xrightarrow{cs} b$

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

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

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

$Y_n \xrightarrow{cs} b \implies P(\{Y_n \to b\}) = 1$

entonces $P(\underbrace{\{X_n \to a\}}_{\text{sucesión de números reales}} \cap \underbrace{\{Y_n \to b\}}_{\text{sucesión de números reales}}) = 1$

$$\left.\begin{array}{l} X_n \to a \\ Y_n \to b \end{array}\right] \implies X_n + Y_n \to a+b$$ (sucesiones de números reales)

entonces $\{X_n \to a\} \cap \{Y_n \to b\} \subset \{X_n + Y_n \to a+b\}$

como $P(\{X_n \to a\} \cap \{Y_n \to b\}) = 1$ tenemos que

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

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

② $X_n \xrightarrow{cs} a$, $Y_n \xrightarrow{cs} b$

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

osea queremos ver que $P(\{X_n Y_n \to ab\}) = 1$

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

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

Sucesiones de números reales
$$\begin{cases} X_n \to a \\ Y_n \to b \end{cases} \Rightarrow X_n Y_n \longrightarrow ab$$

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

$$\left.\begin{array}{l} P(\{X_n \to a\}) = 1 \\ P(\{Y_n \to b\}) = 1 \end{array}\right\} \Rightarrow P(\{X_n \to a\} \cap \{Y_n \to b\}) = 1$$

$$\Rightarrow P(\{X_n Y_n \longrightarrow ab\}) = 1$$

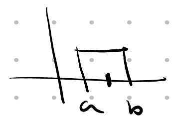
Entonces $X_n Y_n \xrightarrow{cs} ab$

Primer parcial octubre 2020

Ejercicio 7

Sean X_1, X_2, …, X_n variables aleatorias $\overbrace{\text{i.i.d.}}^{\text{independientes identicamente distribuidas}}$ con distribución uniforme en el intervalo $[-2, 2]$. Sea $\bar{X}_n$ el promedio. Aplicando la desigualdad de Chebyshev, hallar el menor valor de n que cumple $\mathbb{P}(|\bar{X}_n| > 0.2) \leq 0.05$.

(A) 334 **(B)** 67 **(C)** 667 **(D)** 22 **(E)** 2000



$X_i \sim U[-2,2]$ $\qquad X \sim U[a,b] \to E(X) = \frac{a+b}{2}$

$\to Var(X) = \frac{(b-a)^2}{12}$

$\overline{X_n}$

buscamos n para que $P(|\overline{X_n}| > 0{,}2) \leq 0{,}05$

$X_i \sim U[-2,2]$

$$E(X_i) = \frac{-2+2}{2} = 0$$

$$Var(X_i) = \frac{(2-(-2))^2}{12} = \frac{16}{12} = \frac{4}{3}$$

$$E(\overline{X_n}) = E\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \frac{1}{n}\left(\overbrace{E(X_1)}^{0} + \overbrace{E(X_2)}^{0} + \dots + \overbrace{E(X_n)}^{0}\right) = 0$$

$$Var(\overline{X_n}) = Var\left(\frac{X_1+X_2+\dots+X_n}{n}\right) = \frac{1}{n^2} Var(X_1+X_2+\dots+X_n)$$

$$= \frac{1}{n^2}\left(\underbrace{Var(X_1)}_{=4/3} + \underbrace{Var(X_2)}_{=4/3} + \dots + \underbrace{Var(X_n)}_{=4/3}\right)$$

$$= \frac{1}{n^2} n \frac{4}{3}$$

$$= \frac{4}{3n}$$

$$E(\overline{X_n}) = 0 \qquad Var(\overline{X_n}) = \frac{4}{3n}$$

Chevyshev: $$P(|\overline{X_n} - \underbrace{E(\overline{X_n})}_{=0}| > t) \leq \frac{Var(\overline{X_n})}{t^2}$$

$$P(|\overline{X_n}| > t) \leq \frac{Var(\overline{X_n})}{t^2}$$

tomamos $t = 0,2$:

$$P(|\overline{X_n}| > 0,2) \leq \frac{Var(\overline{X_n})}{(0,2)^2} = \frac{4}{3n(0,2)^2}$$

buscamos n tal que

$$\frac{4}{3n(0,2)^2} \leq 0,05$$

$$\frac{1}{n} \leq \frac{0,05 \cdot 3(0,2)^2}{4}$$

$$n \geq \frac{4}{0,05 \cdot 3 \cdot (0,2)^2} = 666,66\dots$$

entonces tomamos $n = 667$