Clase Auxiliar 7 Probabilidades y Estadística Prim '09

FC; VR

Desigualdad de Markov: $X \geq 0$ v.a. $\Rightarrow \mathbb{P}(X \geq c) \leq \frac{\mathbb{E}(X)}{c}$

Desigualdad de Chebyshev: $\mathbb{P}(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$

Ley débil de los grandes números: $(X_i)_{i \in \mathbb{N}}$ iid, $\mathbb{E}(X_1) < \infty$. Entonces

$$\mathbb{P}\left(|\bar{X}_n - \mu| \geq \varepsilon\right) \xrightarrow{n \to \infty} 0$$

Teorema Central del Límite: $(X_i)_{i \in \mathbb{N}}$ iid, $\mathbb{E}(X_1) = \mu$, $Var(X_1) = \sigma^2$. Entonces

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{ley} N(0,1).$$

Ley fuerte de los grandes números: $(X_i)_{i \in \mathbb{N}}$ iid, $\mathbb{E}(X_1) = \mu < \infty$. Entonces

$$\mathbb{P}\left(\bar{X}_n \xrightarrow{n \to \infty} \mu\right) = 1$$

Problema 1 a) Pruebe que si X es v.a. con media 0 y varianza finita σ^2, entonces $\forall a > 0$, $\mathbb{P}(X \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$

b) Si el número de items producidos en una fábrica durante una semana es una v.a. con media 100 y varianza 400, calcule una cota superior para la probabilidad de que esta semana se produzcan al menos 120 items.

Sol a) Vemos que $X > a$ es equivalente a $X+b \geq a+b$, para $b>0$.

Entonces, $\mathbb{P}(X>a) = \mathbb{P}(X+b \geq a+b)$

$$\leq \mathbb{P}\left((X+b)^2 \geq (a+b)^2 \right)$$

Aplicando la desigualdad de Markov, se tiene que

$$\mathbb{P}\left((X+b)^2 \geq (a+b)^2 \right) \leq \frac{\mathbb{E}\left[(X+b)^2\right]}{(a+b)^2}$$

$$= \frac{\mathbb{E}(X^2+2bX+b^2)}{(a+b)^2}$$

$$= \frac{\mathbb{E}(X^2) + 2b\,\mathbb{E}(X) + b^2}{(a+b)^2} \qquad \text{con } \mathbb{E}(X)=0, \ \mathrm{Var}(X) = \mathbb{E}(X^2) = \sigma^2$$

$$\leq \frac{\sigma^2+b^2}{(a+b)^2}$$

$$\Rightarrow \mathbb{P}(X>a) \leq \underbrace{\frac{\sigma^2+b^2}{(a+b)^2}}_{f(b)} \qquad \forall b>0.$$

Minimizamos f: $\frac{\partial}{\partial b} f(b) = 0 \Rightarrow \frac{2b(a+b)^2 - (\sigma^2+b^2)\, 2(a+b)}{(a+b)^4} = 0$

$$\Rightarrow \cancel{2}b\,(a+b)^{\cancel{2}} = \cancel{2}(\sigma^2+b^2)\cancel{(a+b)}$$

$$ab + \cancel{b^2} = \sigma^2 + \cancel{b^2}$$

$$\Rightarrow \boxed{b = \frac{\sigma^2}{a}} \qquad \text{valor que minimiza } f.$$

Entonces, $\mathbb{P}(X \leq a) \leq \dfrac{\sigma^2 + \frac{\sigma^4}{a^2}}{\left(a + \frac{\sigma^2}{a}\right)^2} = \dfrac{\sigma^2 \cancel{\frac{a^2+\sigma^2}{a^2}}}{\frac{(a^2+\sigma^2)^{\cancel{2}}}{\cancel{a^2}}}$

$$\leq \frac{\sigma^2}{a^2+\sigma^2}$$

//

b) Sea $Y = \#$ de ítems producidos en 1 semana. Entonces nos piden una cota para

$$\mathbb{P}(Y \geq 120).$$

Notemos que $\mathbb{P}(Y \geq 120) = \mathbb{P}(Y - 100 > 120 - 100)$

$$= \mathbb{P}(Y - \mathbb{E}(Y) > 20)$$

Definiendo $X = Y - \mathbb{E}(Y)$, $\mathbb{E}(X) = 0$ y $\mathbb{V}ar(X) = \mathbb{V}ar(Y) = 400$.

Entonces $\mathbb{P}(Y \geq 120) = \mathbb{P}(X \geq 20)$

$$\leq \frac{\sigma^2}{a^2+\sigma^2} \qquad \text{con } \sigma^2 = \mathbb{V}ar(X), \; a = 20$$

$$\leq \frac{400}{400+400} = \frac{1}{2}$$

//

obs Notemos que usando desigualdad de Markov, se llega a

$$\mathbb{P}(Y \geq 120) \leq \frac{\mathbb{E}(Y)}{120} = \frac{100}{120} = \frac{5}{6}.$$

La cota obtenida antes es mejor.

Problema 2 Un conjunto de 200 personas, 100 hombres y 100 mujeres, es dividido al azar en 100 parejas. De una cota superior de la probabilidad que a lo más 30 pares sean mixtos.

Sol: Primero para la mujer i, definimos X_i la va que indica si le tocó por pareja un hombre o una mujer:

$$X_i = \begin{cases} 0 & \text{si le toca una mujer} \quad (\text{prob } 1-p) \\ 1 & \text{si le toca un hombre.} \quad (\text{prob } p) \end{cases}$$

Entonces $X = \sum_{i=1}^{100} X_i$ es el total de parejas mixtas.

Ahora, $p = \mathbb{P}(X_i = 1) = \frac{100}{199} = \frac{\#\text{ de hombres}}{\#\text{ total de opciones.}}$

(notamos que X_i es Bernoulli de parametro $p = \mathbb{E}(X_i)$)

Para $i \neq j$ calculamos

$$\mathbb{E}(X_i X_j) = \mathbb{P}(X_i=1, X_j=1) = \mathbb{P}(X_i=1)\mathbb{P}(X_j=1 \mid X_i=1) + 0$$

$$= \frac{100}{199} \times \frac{99}{197}$$

Entonces, $\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} \mathbb{E}(X_i) = 100 \cdot \frac{100}{199} \approx 50,25.$

$$\mathrm{Var}(X) = \sum_{i=1}^{100} \mathrm{Var}(X_i) + \sum_{i \neq j}^{100} \mathrm{Cov}(X_i, X_j) \qquad \mathrm{Var}(X_i) = p(1-p)$$

$$= 100 \cdot \frac{100}{199} \frac{99}{199} + (100^2 - 100)\left(\frac{100}{199} \cdot \frac{99}{197} - \left(\frac{100}{199}\right)^2\right)$$

$$\approx 25,126.$$

Usando la desigualdad de Chevishev,

$$\mathbb{P}(X \leq 30) \leq \mathbb{P}(|X - 50,25| \geq 20,25)$$

$$\leq \frac{25,126}{(20,25)^2} \approx 0,61$$

Problema 3 Si X es una v.a. Gamma $(n,1)$ ¿cuan grande se necesita n para que

$$\mathbb{P}\left(\left|\frac{X}{n} - 1\right| > 0,01\right) < 0,01 ?$$

Sol Como $X \sim \text{Gamma}(n,1)$, entonces X se puede escribir como suma de v.a. exponenciales iid.

$$X \overset{ley}{=} \sum_{i=1}^{n} X_i \quad , \text{con } X_i \sim \text{Exp}(1), \quad X_i \text{ iid.}$$

Entonces, $\mathbb{P}\left(\left|\frac{X}{n} - 1\right| > 0,01\right) = \mathbb{P}\left(\left|\overline{X}_n - \mathbb{E}(X_1)\right| > 0,01\right) < 0,01.$

Por las leyes de grandes números, tal n existe.

Ahora, si ocupamos Chevishev, imponemos

$$\mathbb{P}\left(|\overline{X}_n - 1| > 0,01\right) \leq \frac{\mathbb{V}ar(\overline{X}_n)}{(0,01)^2} \overset{\substack{\text{imposición} \\ \downarrow}}{<} 0,01$$

con $\mathbb{V}ar(\overline{X}_n) = \frac{1}{n^2} \mathbb{V}ar\left(\text{Gamma}(n,1)\right) = \frac{1}{n^2} \frac{n}{1} = \frac{1}{n}$

Entonces $\frac{1}{n} < (0,01)^3 \Rightarrow n > 100^3 = 1.000.000$ //

Problema 4 El # de días que cierto tipo de componente dura hasta fallar es una v.a. con distribución F. Una vez que la componente falla es inmediatamente reemplazada por otra del mismo tipo.

Si X_i denota el tiempo de vida de la componente i-ésima en uso, y

$S_n = \sum_{i=1}^{n} X_i$ representa el instante de la n-ésima falla,

se define la tasa de fallas en el largo plazo por $r = \lim_{n\to\infty} \frac{n}{S_n}$

Suponiendo que las X_i son iid, calcule r.

Sol Por la ley fuerte de los grandes números,

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{S_n}{n} = \mathbb{E}(X_1) \right) = 1$$

Entonces, casi seguramente

$$\lim_{n\to\infty} \frac{S_n}{n} = \mu = \mathbb{E}(X_1) = \int_0^\infty (1 - F(s))\, ds$$

y entonces, $\lim_{n\to\infty} \frac{n}{S_n} = \frac{1}{\mu}$

Ejercicio: calcular el caso en que $F(t) = t^2 \mathbb{1}_{(0,1)}(t)$.

obs: si X va, $X \geq 0$

$$\mathbb{E}(X) = \int_{x=0}^{\infty} x f_X(x)\, dx = \int_{x=0}^{\infty} \mu([0,x))\, f_X(x)\, dx$$

con $\mu(A) = \int_A dx$

$$= \int_{x=0}^{\infty} \int_{s=0}^{\infty} \mathbb{1}_{[0,x)}(s)\, ds\, f_X(x)\, dx \overset{\text{Fubini}}{=} \int_{s=0}^{\infty} \int_{x=0}^{\infty} \mathbb{1}_{[0,x)}(s)\, f_X(x)\, dx\, ds$$

$$= \int_{s=0}^{\infty} \int_{x=0}^{\infty} \mathbb{1}_{(s,\infty)}(x)\, f_X(x)\, dx\, ds = \int_{s=0}^{\infty} \int_{x\in(s,\infty)} f_X(x)\, dx\, ds = \int_{s=0}^{\infty} \mathbb{P}(X>s)\, ds = \int_{s=0}^{\infty} (1-F(s))\, ds$$