PROBABILIDAD II

Grado en Matemáticas

Tema 4
Momentos y desigualdades

Javier Cárcamo

Departamento de Matemáticas
Universidad Autónoma de Madrid
javier.carcamo@uam.es

Tema 4: Momentos y desigualdades

1. Definiciones
2. Desigualdades de Markov y Chebyshev
3. Desigualdad de Jensen
4. Desigualdades de Hölder y Minkowski
5. Covarianza y correlación

Definiciones

Sea $X : \Omega \longrightarrow \mathbb{R}$ una v.a., $\alpha > 0$ y $a \in \mathbb{R}$. Se llama:

- **Momento absoluto de orden α:** $\mathrm{E}|X|^\alpha$.
- **Momento absoluto de orden α alrededor de a:** $\mathrm{E}|X-a|^\alpha$.
- **Momento de orden n:** $\mathrm{E}X^n$.
- **Momento de orden n alrededor de a:** $\mathrm{E}(X-a)^n$.

Espacios $\mathcal{L}_\alpha$

$$\mathcal{L}_\alpha = \{X : \Omega \longrightarrow \mathbb{R} : \mathrm{E}|X|^\alpha < \infty\}, \quad \alpha > 0.$$

Si $X \in \mathcal{L}_1$, la **varianza** de X es $\mathrm{Var}(X) = \mathrm{E}(X - \mathrm{E}X)^2$.

Lema: $a, b \in \mathbb{R}$. Para todo $\alpha > 0$, $|a+b|^\alpha \leq 2^\alpha(|a|^\alpha + |b|^\alpha)$.

Proposición: X v.a., $\alpha > 0$ y $a \in \mathbb{R}$.

$$\mathrm{E}|X|^\alpha < \infty \text{ si y solo si } \mathrm{E}|X-a|^\alpha < \infty.$$

Corolario: Sea $X \in \mathcal{L}_1$. $\mathrm{Var}(X) < \infty$ si y solo si $X \in \mathcal{L}_2$. En tal caso, $\mathrm{Var}(X) = \mathrm{E}X^2 - \mathrm{E}^2X$.

Teorema: Si $0 < \alpha < \beta$, entonces $\mathcal{L}_\beta \subset \mathcal{L}_\alpha$.

Desigualdades de Markov y Chebyshev

Teorema: Desigualdad de Markov

Sea X v.a. Para $\epsilon > 0$, $\alpha > 0$, se tiene

$$\mathrm{P}(|X| \geq \epsilon) \leq \frac{\mathrm{E}|X|^\alpha}{\epsilon^\alpha}.$$

Corolario: Desigualdad de Chebyshev

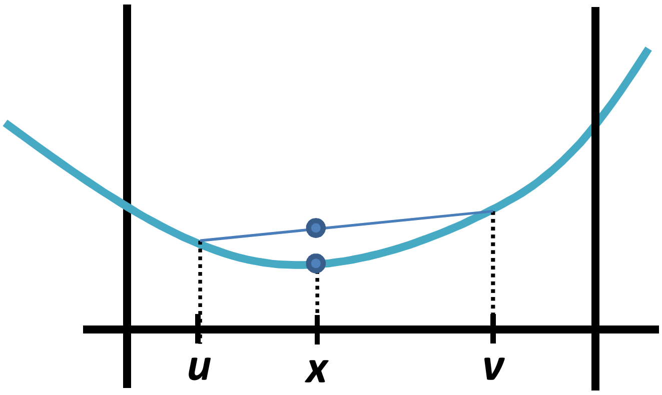
Sea $X \in \mathcal{L}_1$. Para todo $\epsilon > 0$, se tiene

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

Funciones convexas

Definición: Sea I intervalo real. Una función $f : I \longrightarrow \mathbb{R}$ se dice **convexa** si para todo $u, v \in I$, para todo $\lambda \in [0, 1]$, se tiene que

$$f(\lambda u + (1 - \lambda)v) \leq \lambda f(u) + (1 - \lambda)f(v).$$



Ejemplos: $f(x) = e^x$; $f(x) = |x|^\alpha$, $\alpha \geq 1$; $f(x) = 1/x$ $(x > 0)$; $f(x) = -\log(x)$ $(x > 0)$; etc.

Propiedades de las funciones convexas

1. f convexa si y sólo si para todos $u_1, \ldots, u_n \in I$, para todos $\lambda_1, \ldots, \lambda_n \geq$ con $\lambda_1 + \cdots + \lambda_n = 1$, se tiene que
$$f(\lambda_1 u_1 + \cdots + \lambda_n u_n) \leq \lambda_1 f(u_1) + \cdots + \lambda_n f(u_n).$$
2. Sea $f : I \longrightarrow \mathbb{R}$ derivable. f convexa si y sólo si f' es no decreciente.
3. Sea $f : I \longrightarrow \mathbb{R}$ dos veces derivable. f convexa si y sólo si $f'' \geq 0$.
4. Si f convexa, entonces f es continua salvo quizá en ∂I (extremos del intervalo I).
5. Si f es convexa, para todo $(u, f(u)) \in G_f$ tal que $u \in I^\circ$ existe una recta que pasa por $(u, f(u))$ (recta soporte) tal que su gráfica está totalmente por debajo de f. Es decir,
$$\forall u \in I^\circ, \exists a \in \mathbb{R} \quad : \quad a(x - u) + f(u) \leq f(x), \quad \forall x \in I.$$

Teorema: Desigualdad de Jensen

Sea I intervalo real, $X : \Omega \longrightarrow I$ v.a. y $f : I \longrightarrow \mathbb{R}$ convexa.
Si X, $f(X) \in \mathcal{L}_1$, entonces

$$f(\mathrm{E}X) \leq \mathrm{E}f(X).$$

Ejemplos:

1. $f(x) = e^x$.
2. $f(x) = |x|^\alpha$ $(\alpha \geq 1)$.
3. $f(x) = 1/x$ $(x \geq 0)$.

Nota: Una función f es **cóncava** si $-f$ es convexa. La desigualdad de Jensen para funciones cóncavas es $f(\mathrm{E}X) \geq \mathrm{E}f(X)$.

Ejemplos:

1. $f(x) = -e^x$.
2. $f(x) = |x|^\alpha$ $(0 < \alpha \leq 1)$.
3. $f(x) = \log(x)$ $(x > 0)$.

Lema: Para todo $a, b \in \mathbb{R}$ y $p, q > 1$ con $1/p + 1/q = 1$ (p y q conjugados), se tiene

$$|a \cdot b| \leq \frac{|a|^p}{p} + \frac{|b|^q}{q}.$$

Teorema: Desigualdad de Hölder

$p, q > 1$ con $1/p + 1/q = 1$. Si $X \in \mathcal{L}_p$ e $Y \in \mathcal{L}_q$, entonces

$$\mathrm{E}|XY| \leq (\mathrm{E}|X|^p)^{1/p}(\mathrm{E}|Y|^q)^{1/q}.$$

En particular $XY \in \mathcal{L}_1$.

Corolario: Desigualdad de Cauchy-Schwarz

Si $X \in \mathcal{L}_2$ e $Y \in \mathcal{L}_2$, entonces

$$\mathrm{E}|XY| \leq (\mathrm{E}|X|^2)^{1/2}(\mathrm{E}|Y|^2)^{1/2}.$$

Teorema: Desigualdad de Minkowski

Sea $p \geq 1$. Si $X, Y \in \mathcal{L}_p$, entonces $X + Y \in \mathcal{L}_p$ y

$$(\mathrm{E}|X+Y|^p)^{1/p} \leq (\mathrm{E}|X|^p)^{1/p} + (\mathrm{E}|Y|^p)^{1/p}.$$

Nota (Espacios L_p): Para $p \geq 1$,

$$\mathcal{L}_p = \{X : \Omega \longrightarrow \mathbb{R} : \mathrm{E}|X|^p < \infty\} \quad \text{es un espacio vectorial.}$$

$$\|X\|_p = (\mathrm{E}|X|^p)^{1/p}.$$

1. $\|X+Y\|_p \leq \|X\|_p + \|Y\|_p$ (Minkowski).
2. $\|\lambda X\|_p = |\lambda| \|X\|_p$.

Es decir, $\| \bullet \|_p$ es una **seminorma**. No es norma ya que $\|X\|_p = 0$ si y sólo si $X = 0$ c.s.

Si definimos la relación de equivalencia $X \sim Y$ si y sólo si $X = Y$ c.s., entonces $\mathcal{L}_p / \sim = L_p$ es un espacio normado.

Definición: Dada una variable $X \in \mathcal{L}_2$, se llama **desviación típica** de X a

$$\sigma_X = \sqrt{\mathrm{Var}(X)}.$$

Definición: Si $X, Y \in \mathcal{L}_2$, se llama **covarianza** de X, Y a

$$\mathrm{Cov}(X, Y) = \mathrm{E}\left[(X - \mathrm{E}X)(Y - \mathrm{E}Y)\right].$$

Nota:

1. Si $X, Y \in \mathcal{L}_2$, entonces existe $\mathrm{Cov}(X, Y)$ y
$$|\mathrm{Cov}(X, Y)| \leq \sigma_X \sigma_Y.$$
2. Si $X, Y \in \mathcal{L}_2$, entonces
$$\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}X\mathrm{E}Y.$$
3. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$ y $\mathrm{Cov}(Y, X) = \mathrm{Cov}(X, Y)$.
4. $\mathrm{Cov}(aX, bY) = ab\mathrm{Cov}(X, Y)$.

Definición: Sean $X, Y \in \mathcal{L}_2$. Se dice que X e Y están **incorreladas** (o **incorrelacionadas**) si $\text{Cov}(X, Y) = 0$, es decir, si $\mathrm{E}(XY) = \mathrm{E}X\mathrm{E}Y$.

Observación: X, Y independientes $\Rightarrow X$, Y incorreladas.
Sin embargo, X, Y incorreladas $\not\Rightarrow X$, Y independientes.

1. $\text{Var}(aX + b) = a^2\text{Var}(X)$.
2. Varianza de una suma de v.a.: Si $X_1, \ldots, X_n \in \mathcal{L}_2$,
$$\text{Var}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \text{Var}(X_i) + 2 \sum_{1\leq i<j\leq n} \text{Cov}(X_i, X_j).$$
3. Si $X_1, \ldots, X_n$ incorreladas dos a dos (en particular si son independientes), entonces
$$\text{Var}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \text{Var}(X_i).$$

Definición: Dadas X, Y v.a. no degeneradas, el **coeficiente de correlación (lineal de Pearson)** se define mediante
$$\rho_{X,Y} = \frac{\text{Cov(X,Y)}}{\sigma_X \sigma_Y} \quad \text{(es adimensional)}.$$

Nota: $\rho_{X,Y} \in [-1, 1]$.

Observación: Análisis de los extremos

1. $\rho_{X,Y} = 0 \iff X$, Y incorreladas.
2. $\rho_{X,Y} = 1 \iff Y = aX + b$ c.s. con $a > 0$, $b \in \mathbb{R}$.
3. $\rho_{X,Y} = -1 \iff Y = aX + b$ c.s. con $a < 0$, $b \in \mathbb{R}$.