Repaso de Probabilidad

Señales Aleatorias y Modulación

Instituto de Ingeniería Eléctrica
Facultad de Ingeniería
Universidad de la República

UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

2 de agosto de 2022

Probabilidad

- Un espacio de probabilidad está determinado por la terna $(\Omega, \mathcal{F}, P)$. Ω es el espacio muestral o universo de sucesos elementales, $\mathcal{F}$ el conjunto de sucesos de Ω (sigma álgebra) y $P : \mathcal{F} \to (0, 1)$ (una función de $\mathcal{F}$ en los reales).
- La probabilidad es un instrumento de medida que permite cuantificar la incertidumbre en la ocurrencia de los sucesos.

Probabilidad condicional

- **Def:** La probabilidad condicional de E dado F es (necesitamos que $P(F) > 0$)

$$P(E \mid F) = \frac{P(E \cap F)}{P(F)}$$

- Ley de probabilidad total

$$P(E) = P\left(\bigcup_{i=1}^{\infty}[E \cap F_i]\right) = \sum_{i=1}^{\infty} P(E \cap F_i) = \sum_{i=1}^{\infty} P(E \mid F_i)P(F_i)$$

F_i, $i = 1, 2, \ldots$ es una partición de Ω (posiblemente infinita), de conjuntos son disjuntos ($\Rightarrow F_i \cap F_j = \emptyset$ para $i \neq j$) y que cubren el espacio ($\Rightarrow \cup_{i=1}^{\infty} F_i = \Omega$)

Regla de Bayes

- Regla de Bayes

$$P(E \mid F) = \frac{P(F \mid E)P(E)}{P(F)}$$

La regla de Bayes permite la inversión temporal o de causalidad. Si F (futuro) viene después de E (pasado),

$\Rightarrow$ $P(E \mid F)$, prob. que el pasado (E) haya visto el futuro (F)

$\Rightarrow$ $P(F \mid E)$, proba. que el futuro (F) corresponda al pasado (E)

Independencia

- **Def:** los suceso E y F son independientes si $P(E \cap F) = P(E)P(F)$

  $\Rightarrow$ Sucesos que no son independientes son dependientes

- De la definición de probabilidad condicional

$$P(E \,|\, F) = \frac{P(E \cap F)}{P(F)} = \frac{P(E)P(F)}{P(F)} = P(E)$$

  Intuitivamente, conocer F no altera nuestra percepción sobre E, F no contiene información sobre E. El recíproco también es cierto: $P(F \,|\, E) = P(F)$

- **Def:** Los sucesos E_i, $i = 1, 2, \ldots$ se dicen mutuamente independientes si

$$P\left(\bigcap_{i \in I} E_i\right) = \prod_{i \in I} P(E_i)$$

Variable aleatorias (VA)

- **Def:** Una VA $X(\omega)$ es una función que le asigna un valor a un suceso elemental $\omega \in \Omega$

  $\Rightarrow$ Podemos pensar en VAs como medidas asociadas a un experimento

- Las probabilidades de las VAs se infieren de los sucesos elementales subyacentes

$$P(X(\omega) = x) = P(\{\omega \in \Omega : X(\omega) = x\})$$

$$P(X(\omega) \in (-\infty, x]) = P(\{\omega \in \Omega : X(\omega) \in (-\infty, x]\})$$

- Ejemplo: La función indicatriz de un suceso es una VA

  Sea $\omega \in \Omega$ una realización, y $E \subset \Omega$ un suceso

$$\mathbb{I}\{E\}(\omega) = \begin{cases} 1, & \text{if } \omega \in E \\ 0, & \text{if } \omega \notin E \end{cases}$$

  $\Rightarrow$ Indica que la realización ω pertenece al conjunto E, tomando el valor 1

Variables aleatorias discretas

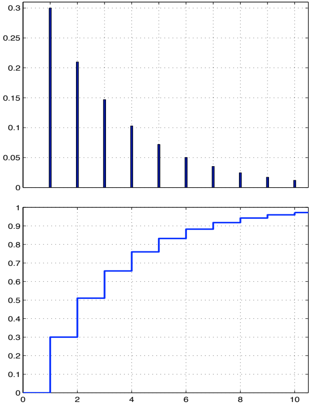
- Las VA discretas toman, a lo sumo, un conjunto numerable de valores

  Estan completamente caracterizadas por su Función de (masa de) probabilidad o probability mass function (pmf) $p_X(x) = P(X = x)$

- Si X tiene soporte $\{x_1, x_2, \ldots\}$, su pmf satisface
  - (i) $p(x_i) > 0$ para $i = 1, 2, \ldots$
  - (ii) $p(x) = 0$ para todo $x \neq x_i$
  - (iii) $\sum_{i=1}^{\infty} p(x_i) = 1$

- Función de distribución (acumulada) (cdf) $F_X(x) = P(X \leq x) = \sum_{i:x_i \leq x} p(x_i)$

  $\Rightarrow$ Función creciente con saltos en x_i



Variables aleatorias continuas

- Los valores posibles que toman las VAs continuas X forman un subconjunto denso $\mathcal{X} \subseteq \mathbb{R}$

  $\Rightarrow$ Cantidad infinita de valores no numerables

- La función de densidad de probabilidad (pdf) $f_X(x)$ es tal que para cualquier subconjunto $\mathcal{X} \subseteq \mathbb{R}$
  (A la derecha, pdf normal)

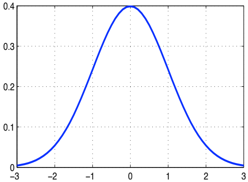
$$P(X \in \mathcal{X}) = \int_{\mathcal{X}} f_X(x) dx$$

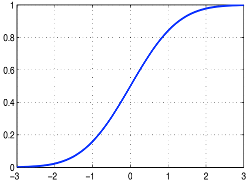
  $\Rightarrow$ Tenemos $P(X = x) = 0$ para todo $x \in \mathcal{X}$

- La cdf definida como antes y relacionada a la (A la derecha, cdf normal)

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(u)\, du$$

$$\Rightarrow P(X \leq \infty) = F_X(\infty) = \lim_{x \to \infty} F_X(x) = 1$$





Valores esperados

- Se nos pide que resumamos la información sobre una VA en un único valor,

  $\Rightarrow$ ¿cuál debería ser ese valor?

- Si se nos pidiera una descripción con unos pocos valores,

  $\Rightarrow$ ¿cuáles deberían ser esos valores?

- Los valores esperados (medias) son respuestas convenientes a estas preguntas

- **Atención:** Las esperanzas son descripciones condensadas

  $\Rightarrow$ No capturan todos los aspectos del fenómeno aleatorio

  $\Rightarrow$ La historia completa es contada por la distribución de probabilidad (cdf)

Valor esperado: Definición para VAs discretas

- La VA discreta X toma valores x_i, $i = 1, 2, \ldots$ con pmf $p(x)$
- **Def:** El valor esperado de la VA discreta X es

$$\mathbb{E}[X] := \sum_{i=1}^{\infty} x_i p(x_i) = \sum_{x:p(x)>0} xp(x)$$

- Es el promedio ponderado sobre los valores posibles x_i. Las probabilidades son pesos
- Es el promedio común si la VA toma valores x_i, $i = 1, \ldots, N$ equiprobables

$$\mathbb{E}[X] = \sum_{i=1}^{N} x_i p(x_i) = \sum_{i=1}^{N} x_i \frac{1}{N} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

Valor esperado: Definición para VAs continuas

- VA continua X toma valores en $\mathbb{R}$ con pdf $f(x)$
- **Def:** El valor esperado de la VA continua X es

$$\mathbb{E}[X] := \int_{-\infty}^{\infty} x f(x)\, dx$$

- Comparar con $\mathbb{E}[X] := \sum_{x:p(x)>0} x p(x)$ para el caso discreto
- Notar que la integral y la suma se suponen definidas

  $\Rightarrow$ De no ser así decimos que la esperanza no existe

Valor esperado de una función de una VA

- Consideremos una función $g(X)$ de una VA X. ¿Valor esperado de $g(X)$?
- g(X) es una VA, por lo que tiene una pmf $p_{g(X)}(g(x))$

$$\mathbb{E}[g(X)] = \sum_{g(x):p_{g(X)}(g(x))>0} g(x)p_{g(X)}(g(x))$$

$\Rightarrow$ Requiere calcular la pmf de $g(X)$. Hay una forma simple de hacerlo

Teorema

Sea $g(X)$ una función de una VA discreta X con pmf $p_X(x)$. Entonces

$$\mathbb{E}[g(X)] = \sum_{i=1}^{\infty} g(x_i)p_X(x_i)$$

- Suma ponderada de valores funcionales. No es necesario calcular la pmf de $g(X)$
- Lo mismo vale para X va continua

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)\,dx$$

Valor esperado para una función lineal de una VA

- Consideremos la función afín $g(X) = aX + b$

$$\begin{aligned}
\mathbb{E}\left[aX + b\right] &= \sum_{i=1}^{\infty}(ax_i + b)p_X(x_i) \\
&= \sum_{i=1}^{\infty} ax_i p_X(x_i) + \sum_{i=1}^{\infty} bp_X(x_i) \\
&= a\sum_{i=1}^{\infty} x_i p_X(x_i) + b\sum_{i=1}^{\infty} p_X(x_i) \\
&= a\mathbb{E}\left[X\right] + b1
\end{aligned}$$

- Se puede intercambiar la esperanza con constantes aditivas o multiplicativas

$$\mathbb{E}\left[aX + b\right] = a\mathbb{E}\left[X\right] + b$$

$\Rightarrow$ Vale lo mismo para VAs continuas (linealidad de la integral)

Valor esperado de una función indicatriz

- Sean X una VA y $\mathcal{X}$ un conjunto

$$\mathbb{I}\{X \in \mathcal{X}\} = \begin{cases} 1, & \text{if } x \in \mathcal{X} \\ 0, & \text{if } x \notin \mathcal{X} \end{cases}$$

- El valor esperado de $\mathbb{I}\{X \in \mathcal{X}\}$ en el caso discreto es

$$\mathbb{E}[\mathbb{I}\{X \in \mathcal{X}\}] = \sum_{x:p_X(x)>0} \mathbb{I}\{x \in \mathcal{X}\} p_X(x) = \sum_{x \in \mathcal{X}} p_X(x) = \mathrm{P}(X \in \mathcal{X})$$

- De la misma forma, para el caso continuo

$$\mathbb{E}[\mathbb{I}\{X \in \mathcal{X}\}] = \int_{-\infty}^{\infty} \mathbb{I}\{x \in \mathcal{X}\} f_X(x) dx = \int_{x \in \mathcal{X}} f_X(x) dx = \mathrm{P}(X \in \mathcal{X})$$

- Valor esperado de la indicatriz = Probabilidad del suceso correspondiente

Momentos, momentos centrados y varianza

- **Def:** El momento de orden n ($n \geq 0$) de una VA es

$$\mathbb{E}\left[X^n\right] = \sum_{i=1}^{\infty} x_i^n p(x_i)$$

- **Def:** El momento centrado de orden n corrige el sesgo de la media, esto es

$$\mathbb{E}\left[\left(X - \mathbb{E}\left[X\right]\right)^n\right] = \sum_{i=1}^{\infty} \left(x_i - \mathbb{E}\left[X\right]\right)^n p(x_i)$$

- Momento de orden 0: $\mathbb{E}\left[X^0\right] = 1$; momento de orden 1: es la media $\mathbb{E}\left[X\right]$
- El momento centrado de orden 2 es la varianza. Mide el ancho de la pmf

$$\mathrm{var}\left[X\right] = \mathbb{E}\left[\left(X - \mathbb{E}\left[X\right]\right)^2\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}^2[X]$$

Ej: Para funciones afines

$$\mathrm{var}\left[aX + b\right] = a^2 \mathrm{var}\left[X\right]$$

Distribuciones de probabilidad conjunta: cdf conjunta

- Queremos estudiar problemas con más de una VA, e.g. X e Y
- Las distribuciones de probabilidad de X e Y no son suficientes

  $\Rightarrow$ La distribución de probabilidad conjunta (cdf) of (X, Y) se define como

$$F_{XY}(x, y) = \mathrm{P}\left(X \leq x, Y \leq y\right)$$

- Si X, Y claras del contexto, omitimos el subíndice: $F_{XY}(x, y) = F(x, y)$
- Podemos recuperar $F_X(x)$ considerando todos los valores posibles de Y

$$F_X(x) = \mathrm{P}\left(X \leq x\right) = \mathrm{P}\left(X \leq x, Y \leq \infty\right) = F_{XY}(x, \infty)$$

  $\Rightarrow$ $F_X(x)$ y $F_Y(y) = F_{XY}(\infty, y)$ se llaman cdfs marginales

Distribuciones de probabilidad conjunta: pmf conjunta

- Consideremos las VAs discretas X e Y
  X toma valores en $\mathcal{X} := \{x_1, x_2, \ldots\}$ e Y en $\mathcal{Y} := \{y_1, y_2, \ldots\}$
- La pmf conjunta de (X, Y) se define como

$$p_{XY}(x, y) = \mathrm{P}\left(X = x, Y = y\right)$$

- Los valores posibles de (x, y) son elementos del producto Cartesiano $\mathcal{X} \times \mathcal{Y}$
  - (x_1, y_1), (x_1, y_2), ..., (x_2, y_1), (x_2, y_2), ..., (x_3, y_1), (x_3, y_2), ...
- La pmf marginal $p_X(x)$ se obtiene sumando sobre todos los valores posibles de Y

$$p_X(x) = \mathrm{P}\left(X = x\right) = \sum_{y \in \mathcal{Y}} \mathrm{P}\left(X = x, Y = y\right) = \sum_{y \in \mathcal{Y}} p_{XY}(x, y)$$

$\Rightarrow$ De la misma forma $p_Y(y) = \sum_{x \in \mathcal{X}} p_{XY}(x, y)$. Marginalizamos sumando

pdf conjunta

- Sean X, Y VAs continuas y $\mathcal{A} \subset \mathbb{R}^2$ un conjunto
- La pdf conjunta es una función $f_{XY}(x,y) : \mathbb{R}^2 \to \mathbb{R}^+$ tal que

$$P\left((X,Y) \in \mathcal{A}\right) = \iint_{\mathcal{A}} f_{XY}(x,y)\, dxdy$$

- Marginalización. Hay dos formas de escribir $P\left(X \in \mathcal{X}\right)$

$$P\left(X \in \mathcal{X}\right) = P\left(X \in \mathcal{X}, Y \in \mathbb{R}\right) = \int_{X \in \mathcal{X}} \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dy\, dx$$

$\Rightarrow$ Definición de $f_X(x)$ $\Rightarrow$ $P\left(X \in \mathcal{X}\right) = \int_{X \in \mathcal{X}} f_X(x)\, dx$

- Tenemos entonces

$$\Rightarrow f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dy, \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dx$$

Vectores aleatorios

- Por conveniencia es común agrupar las VAs en un vector

  $\Rightarrow$ La distribución de probabilidad de un vector es la distribución conjunta de sus elementos

- Sean por ejemplo dos VAs X e Y. El vector aleatorio es $\mathrm{X} = [X, Y]^T$
- Si X e Y son discretas, el vector X es discreto con pmf

$$p_{\mathrm{X}}(\mathrm{x}) = p_{\mathrm{X}}\left([x, y]^T\right) = p_{XY}\left(x, y\right)$$

- Si X e Y son continuas, X es continuo con pdf

$$f_{\mathrm{X}}(\mathrm{x}) = f_{\mathrm{X}}\left([x, y]^T\right) = f_{XY}\left(x, y\right)$$

- La cdf del vector es $\Rightarrow F_{\mathrm{X}}(\mathrm{x}) = F_{\mathrm{X}}\left([x, y]^T\right) = F_{XY}\left(x, y\right)$
- En general podemos definir VAs n-dimensionales $\mathrm{X} := [X_1, X_2, \ldots, X_n]^T$

  $\Rightarrow$ Es solo notación, las definiciones siguen del caso $n = 2$

Esperanzas conjuntas

- X e Y VAs, $g(X, Y)$ función es también una VA
- El valor esperado de $g(X, Y)$ para X, Y VAs discretas se escribe

$$\mathbb{E}\left[g(X, Y)\right] = \sum_{x,y:p_{XY}(x,y)>0} g(x, y)p_{XY}(x, y)$$

- Si X, Y VAs continuas

$$\mathbb{E}\left[g(X, Y)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f_{XY}(x, y)\, dxdy$$

Valor esperado de una suma de variables aleatorias

- Ejemplo: Valor esperado de suma de dos VAs

$$\mathbb{E}\left[X+Y\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{XY}(x,y)\,dxdy$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{XY}(x,y)\,dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f_{XY}(x,y)\,dxdy$$

- Luego,

$$\mathbb{E}\left[X+Y\right] = \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy\,dx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy$$
$$= \int_{-\infty}^{\infty} x f_X(x)\,dx + \int_{-\infty}^{\infty} y f_Y(y)\,dy$$
$$= \mathbb{E}\left[X\right] + \mathbb{E}\left[Y\right]$$

$\Rightarrow$ Usamos expresiones marginales

- Esperanza $\leftrightarrow$ suma $\Rightarrow \mathbb{E}\left[\sum_i X_i\right] = \sum_i \mathbb{E}\left[X_i\right]$

La esperanza es un operador lineal

- Combinando $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ y $\mathbb{E}[aX+b] = a\mathbb{E}[X] + b$ probamos que

$$\mathbb{E}[a_x X + a_y Y + b] = a_x \mathbb{E}[X] + a_y \mathbb{E}[Y] + b$$

- En notación vectorial ($\mathrm{a} \in \mathbb{R}^n$, $\mathrm{X} \in \mathbb{R}^n$, b escalar)

$$\mathbb{E}\left[\mathrm{a}^T \mathrm{X} + b\right] = \mathrm{a}^T \mathbb{E}[\mathrm{X}] + b$$

- La esperanza conmuta con operadores lineales

Independencia de VAs

- Los suceso E y F son independientes si $\mathrm{P}(E \cap F) = \mathrm{P}(E)\,\mathrm{P}(F)$
- **Def:** Las VAs X e Y son independientes si los sucesos $\{X \leq x\}$ y $\{Y \leq y\}$ son independiente cualquiera sea x e y, i.e.

$$\mathrm{P}(X \leq x, Y \leq y) = \mathrm{P}(X \leq x)\,\mathrm{P}(Y \leq y)$$

$\Rightarrow$ Por definición, esto es equivalente a $F_{XY}(x,y) = F_X(x)F_Y(y)$

- Para VAs discretas, es equivalente a la relación análoga con pmfs

$$p_{XY}(x,y) = p_X(x)p_Y(y)$$

- Para VAs continuas vale el análogo para pdfs

$$f_{XY}(x,y) = f_X(x)f_Y(y)$$

- Independencia $\Leftrightarrow$ La distribución conjunta se factoriza como producto de marginales

Valor esperado de un producto de VAs independientes

Teorema

Para VAs X, Y independientes y funciones $g(X)$, $h(Y)$arbitrarias:

$$\mathbb{E}\left[g(X)h(Y)\right] = \mathbb{E}\left[g(X)\right]\mathbb{E}\left[h(Y)\right]$$

El valor esperado del producto es el producto de los valores esperados

- Se demuestra que $g(X)$ y $h(Y)$ también son independientes. Intuitivo

Ej: Caso especial cuando $g(X) = X$ y $h(Y) = Y$ se tiene

$$\mathbb{E}\left[XY\right] = \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$$

- Esperanzas y productos se pueden intercambiar si las VAs son independientes

Varianza de una suma de VAs independientes

- Sean X_n, $n = 1, \ldots N$ independientes, con $\mathbb{E}[X_n] = \mu_n$, $\text{var}[X_n] = \sigma_n^2$
- Q: ¿Varianza de la suma $X := \sum_{n=1}^{N} X_n$?
- Nótese que la media de X es $\mathbb{E}[X] = \sum_{n=1}^{N} \mu_n$. Luego

$$\text{var}[X] = \mathbb{E}\left[\left(\sum_{n=1}^{N} X_n - \sum_{n=1}^{N} \mu_n\right)^2\right] = \mathbb{E}\left[\left(\sum_{n=1}^{N}(X_n - \mu_n)\right)^2\right]$$

- Desarrollando e intercambiando suma y esperanza

$$\text{var}[X] = \sum_{n=1}^{N} \sum_{m=1}^{N} \mathbb{E}\Big[(X_n - \mu_n)(X_m - \mu_m)\Big]$$

Varianza de una suma de VAs independientes (cont.)

- Separamos auto-productos y productos cruzados. Luego usamos independencia, y $\mathbb{E}(X_n - \mu_n) = 0$

$$
\begin{aligned}
\text{var}[X] &= \sum_{n=1, n\neq m}^{N} \sum_{m=1}^{N} \mathbb{E}\Big[(X_n - \mu_n)(X_m - \mu_m)\Big] + \sum_{n=1}^{N} \mathbb{E}\Big[(X_n - \mu_n)^2\Big] \\
&= \sum_{n=1, n\neq m}^{N} \sum_{m=1}^{N} \mathbb{E}(X_n - \mu_n)\mathbb{E}(X_m - \mu_m) + \sum_{n=1}^{N} \sigma_n^2 = \sum_{n=1}^{N} \sigma_n^2
\end{aligned}
$$

- Si las VAs son independientes $\Rightarrow$ La varianza de la suma es la suma de las varianzas

- Hay resultados más generales para VAs independientes X_i, $i = 1, \ldots, n$

$$
\text{var}\left[\sum_i (a_i X_i + b_i)\right] = \sum_i a_i^2 \text{var}[X_i]
$$

Media muestral

Ej: Sean Y_i, $i = 1, \ldots n$ VAs independientes con $\mathbb{E}[Y_i] = \mu$, $\text{var}[Y_i] = \sigma^2$

- La media muestral es $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$. ¿Cuánto valen $\mathbb{E}\left[\bar{Y}\right]$ y $\text{var}\left[\bar{Y}\right]$?
- Valor esperado $\Rightarrow \mathbb{E}\left[\bar{Y}\right] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[Y_i] = \mu$
- Varianza $\Rightarrow \text{var}\left[\bar{Y}\right] = \frac{1}{n^2}\sum_{i=1}^{n} \text{var}[Y_i] = \frac{\sigma^2}{n}$ (independencia)

Covarianza

- **Def:** La covarianza de X and Y es (generaliza la varianza a pares de VAs)

$$\text{cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

- Si $\text{cov}(X, Y) = 0$ las VAs X, Y se dicen no correlacionadas
- Si X, Y son independientes entonces $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ y $\text{cov}(X, Y) = 0$

  $\Rightarrow$ La independencia implica no correlación

- El inverso no vale, puede ser $\text{cov}(X, Y) = 0$ para X, Y dependientes
  - Ej: X uniforme en $[-a, a]$ y $Y = X^2$

  $\Rightarrow$ Pero si X, Y son normales, correlación nula implica independencia

- Si $\text{cov}(X, Y) > 0$ entonces X e Y tienden a moverse en la misma dirección

  $\Rightarrow$ Correlación positiva

- Si $\text{cov}(X, Y) < 0$ entonces X e Y tienden a moverse en direcciones opuestas

  $\Rightarrow$ Correlación negativa