Variables Aleatorias (II). Esperanza matemática

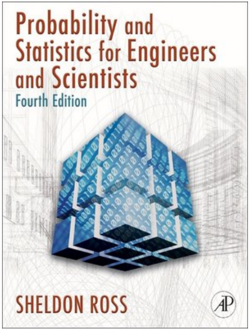
Estadística, Grado en Sistemas de Información

Constantino Antonio García Martínez

Universidad San Pablo Ceu

Contenidos

References



S. Ross. Introduction to Probability and Statistics for Engineers and Scientists. Chapter 4.

C.D. Barr, D.M. Diez, M. Çetinkaya-Rundel. OpenIntro Statistics. Chapters 2-3.

Esperanza matemática

Esperanza

La esperanza matemática de una VA representa el valor que esperamos obtener en media si el experimento aleatorio se repitiese indefinidamente.

Ejemplo: Valor esperado

- Se lanzan 100 monedas equilibradas. ¿Cuál es el número esperado de caras?
- Se lanza una moneda hasta que aparece la prima cara (cuya probabilidad es p). Sea la VA X: número de lanzamientos realizados. ¿Cuál es el valor esperado del número de lanzamientos?

Esperanza

Definimos la esperanza para una variable aleatoria discreta X como

$$E[X] = \sum_i x_i P(X = x_i).$$

Ejemplo: Media aritmética

¡Fíjate que si todas las probabilidades son iguales, obtenemos la fórmula de la media aritmética!

Definimos la esperanza para una variable aleatoria continua X como

$$E[X] = \int x f(x) \, dx.$$

La esperanza de X se suele llamar la **media** de la distribución μ.

Ejemplo: example

Un jugador gana 1 euro si al tirar un dado obtiene un 1 o un 3; pierde 2 euros si sale un 2, 4, 6; y gana 4 euros si sale un 5. ¿Cuál es la ganancia esperada? Sea X la VA "dinero ganado":

$$P(X = x) = \begin{cases} 3/6 & \text{si } x = -2, \\ 2/6 & \text{si } x = 1, \\ 1/6 & \text{si } x = 4, \\ 0 & \text{otherwise} \end{cases} .$$

$$\mathbb{E}[X] = -2\frac{3}{6} + 1\frac{2}{6} + 4\frac{1}{6} = 0 \text{ euros.}$$

Propiedades de la esperanza

- $\mathbb{E}[cX] = c \cdot \mathbb{E}[X]$.
- $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.
- Si X e Y son VAs independientes

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y].$$

Ejercicio: Propiedades de la esperanza
Demuestra las propiedades de la esperanza.

Propiedades de la esperanza

Ejemplo: Funciones de VAs

Si X se distribuye como $p(-2) = 0.15, p(0) = 0.2, p(1) = 0.5, p(2) = 0.15$, ¿cuál es la esperanza de $Y = X^2$?

$$\begin{aligned}\mathbb{E}[Y] &= 0p(0) + 1p(1) + 4\left(p(-2) + p(2)\right) \\ &= 4p(-2) + 0p(0) + 1p(1) + 4p(2) \\ &= \sum_{x_i \in \{-2,0,1,2\}} x_i^2 P(X = x_i).\end{aligned}$$

El ejemplo anterior sugiere que podemos calcular la esperanza de Y usando la distribución de X. Esta propiedad es, de hecho, cierta (y a veces se llama la **ley del estadísta inconsciente**)

$$\mathbb{E}[g(X)] = \sum g(x)p(x) \qquad \text{versión discreta}$$

$$\mathbb{E}[g(X)] = \int g(x)f(x)\,dx \qquad \text{versión continua}$$

Podemos generalizar este resultado a más de una VA, P.ej.:

$$\mathbb{E}[g(X, Y)] = \int g(x, y)f(x, y)\,dx\,dy.$$

Ejercicio:
Calcula la esperanza de la suma de dos dados.

Ejercicio:
Si X e Y tienen como distribución

$$f(x,y) = \begin{cases} xy/96 & 0 < x < 4, 1 < y < 5 \\ 0 & \text{en otro caso} \end{cases}$$

Calcula $\mathbb{E}[X], \mathbb{E}[Y], \mathbb{E}[XY], \mathbb{E}[2X+Y]$.

Varianza

Varianza

La esperanza es muy útil para "resumir" una VA, ya que $\mathbb{E}[X]$ nos permite saber cuál es la tendencia central de X.

Podemos también caracterizar la **dispersión** usando la **varianza**:

$$\sigma_x^2 = \mathrm{Var}[X] = \mathbb{E}[(X-\mu)^2],$$

o bien la **desviación estándar**:

$$\sigma_x = \sqrt{\mathrm{Var}[X]} = \sqrt{\mathbb{E}[(X-\mu)^2]},$$

Propiedades de la varianza

- $\text{Var}[X] = \mathbb{E}[X^2] - \mu^2$.
- $\text{Var}[cX] = c^2\text{Var}[X]$.
- Si X e Y son independientes:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y], \tag{1}$$

$$\text{Var}[X - Y] = \text{Var}[X] + \text{Var}[Y]. \tag{2}$$

Ejercicio: Varianza
Demuestra las propiedades de la varianza.

Varianza

Ejercicio:
Calcula la varianza de la suma de dos dados.

Ejercicio:
Si X se distribuye como

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{en otro caso} \end{cases}$$

Calcula $\mathbb{E}[X]$ y $\mathrm{Var}[X]$.

Covarianza y otros estadísticos

Covarianza

Hemos calculado $\mathrm{Var}[X + Y]$ para el caso en que X e Y son independientes, ¿Qué pasa si X e Y **no son independientes**?

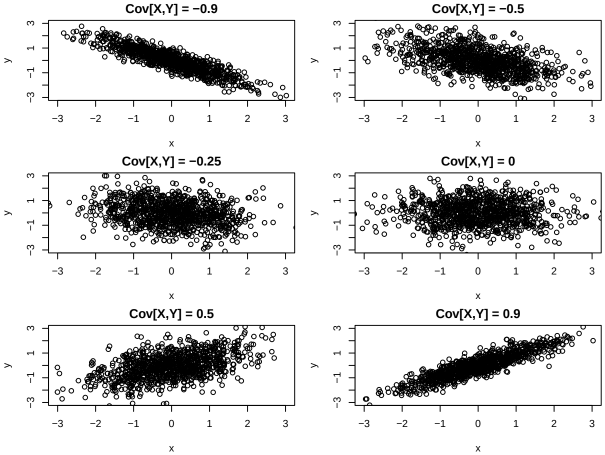
Debemos emplear la distribución conjunta de X e Y.

Definimos **la covarianza** para estudiar como varía X con Y:

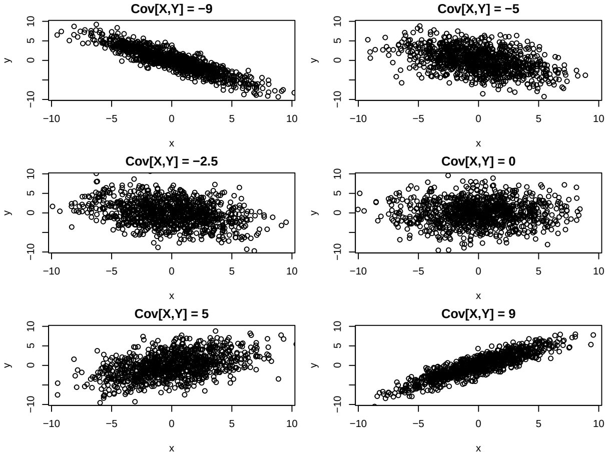
$$\sigma_{xy} = \mathrm{Cov}[X, Y] = \mathbb{E}[(X - \mu_x)(Y - \mu_y)].$$

¡Ojo! La covarianza sólo captura **relaciones lineales** y escala con la varianza de X e Y.

Covarianza


Cov[X,Y] = −0.9
Cov[X,Y] = −0.5
Cov[X,Y] = −0.25
Cov[X,Y] = 0
Cov[X,Y] = 0.5
Cov[X,Y] = 0.9
x
y
−3 −2 −1 0 1 2 3
−3 −1 1 3

Covarianza



Propiedades de la covarianza

1. $\sigma_{xy} = \mathrm{Cov}[X, Y] = \mathbb{E}[XY] - \mu_x \mu_y$.
2. Si X e Y son independientes entonces
$$\sigma_{xy} = 0.$$
3. $\mathrm{Var}[X \pm Y] = \sigma_x^2 + \sigma_y^2 \pm 2\sigma_{xy}$.
4. $|\sigma_{xy}| \leq \sigma_x \sigma_y$.

Ejercicio:
Si X es uniforme en $\{-1, 0, 1\}$ e $Y = X^2$, calcula σ_{xy}.

Correlación

Ejercicio:
Calcula la Covarianza de X e $Y = aX$.

Observamos:

- La Covarianza parece depender de las escalas de X e Y.
- Si X e Y son independientes, $\sigma_{xy} = 0$.
- Si X e Y son completamente dependientes, por ejemplo, $Y = X$, $\sigma_{xy} = \sigma_x \sigma_y$.

Podemos proponer una **medida de dependencia lineal** entre X e Y:

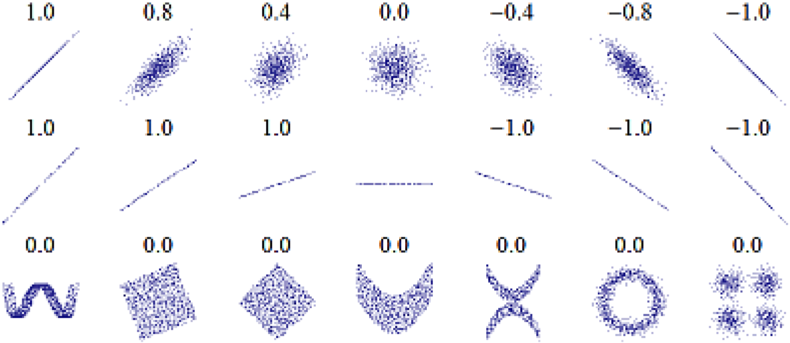
$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}.$$

LLamamos a esta medida, **coeficiente de correlación [de Pearson]**. Si $\rho = 0$, decimos que las variables están **incorreladas**.

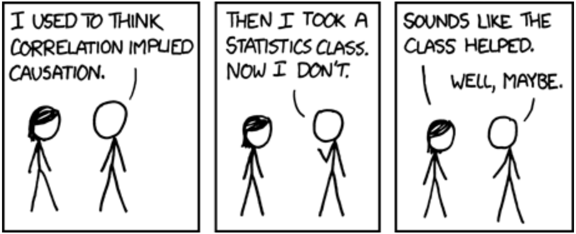
Ejercicio:
Demuestra que $-1 \leq \rho \leq 1$.

Correlación

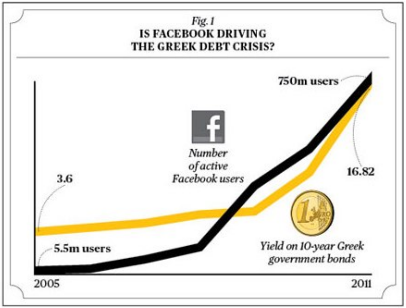
La correlación **solo captura relaciones lineales**.

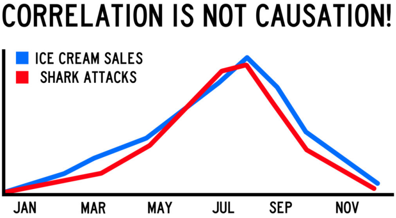


¡Correlación no implica causalidad!



¡Correlación no implica causalidad!





Both ice cream sales and shark attacks increase when the weather is hot and sunny, but they are not caused by each other (they are caused by good weather, with lots of people at the beach, both eating ice cream and having a swim in the sea)

Ejercicio:
Sean D_1 y D_2 los resultados de tirar el dado 1 y el dado 2. Sean $X = D_1 + D_2$ e $Y = D_1 - D_2$. Demuestra que X e Y están incorreladas pero que no son independientes.

Momentos

Hemos visto que las expresiones siguientes juegan un papel fundamental en el cálculo de estadísticos resumen:

$$\mu_r = \mathbb{E}[(X-\mu)^r] \qquad \text{Momento central de orden r} \tag{3}$$

$$\mu'_r = \mathbb{E}[X^r] \qquad \text{Momento de orden r en torno al origen.} \tag{4}$$

Podemos emplear los momentos para definir más estadísticos. Por ejemplo:

- **Coeficiente de asimetría** (skewness):

$$\alpha_3 = \frac{\mu_3}{\sigma^3}.$$

- **Curtosis**: Nos permite medir cómo de picuda es una distribución. Se define como

$$\alpha_4 = \frac{\mu_4}{\sigma^4}.$$

Nos referimos a α_k como el **momento estandarizado de orden k**:

$$\alpha_k = \frac{\mu_k}{\sigma^k}.$$

Desigualdad de Chebyshev y ley débil de los números grandes

Desigualdad de Chebyshev

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

o, para el caso especial $\epsilon = k\sigma$:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Ejemplo:
Desigualdad de Chebyshev Para una distribución **cualquiera** con media y varianza finitas, sabemos que

$$P(|X - \mu| \leq 2\sigma) \geq 0.75.$$

Como consecuencia de la desigualdad de Chebyshev obtenemos:

Ley débil de los grandes números
Sea $X_1, X_2, \cdots, X_n$ VAs mutuamente independientes con media μ y varianza σ^2 finitas. Si $S_n = \sum_{i=1}^n X_i$

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0.$$

Intuitivamente, un resultado más fuerte que cabría esperar que fuera cierto es,

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

pero **NO** es así. Sin embargo si es cierto que

Ley fuerte de los grandes números

$$P\left(\lim_{n\to\infty} \frac{S_n}{n} = \mu\right) = 1.$$

Cuando un sucesos tiene Prob. 1, decimos que es **casi seguro**. Anteponemos el “casi” para distinguirlos de los sucesos **necesarios**.