

Tema 3.- Independencia de variables aleatorias

Asignatura: PROBABILIDAD

Grado en Ingeniería Informática y Matemáticas

(3er Curso - 1er semestre)

©Prof. Dr. José Luis Romero Béjar

(Este material está protegido por la Licencia Creative Commons CC BY-NC-ND que permite "descargar las obras y compartirlas con otras personas, siempre que se reconozca su autoría, pero no se pueden cambiar de ninguna manera ni se pueden utilizar comercialmente").



Departamento de Estadística e Investigación Operativa
Facultad de Ciencias (Despacho XX)

Periodo de docencia: 13/09/2021 a 22/12/2021

- 1 Definición y caracterización
- 2 Propiedades de la independencia
- 3 Reproductividad de distribuciones
- 4 Independencia de vectores aleatorios

Outline

- 1 Definición y caracterización
- 2 Propiedades de la independencia
- 3 Reproductividad de distribuciones
- 4 Independencia de vectores aleatorios

Independencia de las componentes de un vector aleatorio

Sea $X = (X_1, \dots, X_n)$ un vector aleatorio. Se dice que las variables

$$X_i : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_{X_i}), \quad i = 1, \dots, n$$

son **independientes** si la **función de distribución conjunta** F_X **factoriza en producto de las marginales**, i.e.,

$$F_X(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n).$$

Caracterizaciones de independencia para el caso DISCRETO

- ❶ Sea $X = (X_1, \dots, X_n)$ un **vector aleatorio discreto**. Entonces,

$$X_i : (\Omega, \mathcal{A}, P) \longrightarrow (E_{X_i}, \mathcal{B}, P_{X_i}), \quad i = 1, \dots, n$$

son independientes si y sólo si, para cualquier $(x_1, \dots, x_n) \in E_{X_1} \times \dots \times E_{X_n}$,

$$P_X(x_1, \dots, x_n) = p_{X_1}(x_1) \cdots p_{X_n}(x_n) = P[X_1 = x_1] \cdots P[X_n = x_n].$$

- ❷ Sea $X = (X_1, \dots, X_n)$ un **vector aleatorio discreto**. Entonces,

$$X_i : (\Omega, \mathcal{A}, P) \longrightarrow (E_{X_i}, \mathcal{B}, P_{X_i}), \quad i = 1, \dots, n$$

son independientes si y sólo si, para cualquier $(x_1, \dots, x_n) \in E_{X_1} \times \dots \times E_{X_n}$,

$$P_X(x_1, \dots, x_n) = h_1(x_1) \cdots h_n(x_n),$$

donde $h_i : E_{X_i} \longrightarrow \mathbb{R}$, $i = 1, \dots, n$, son funciones arbitrarias.

Se puede decir que las componentes de un vector aleatorio **discreto** son **variables aleatorias independientes** si su **función masa de probabilidad** es **producto de las funciones masa de probabilidad marginales** o, **producto de funciones que solo dependen de las variables aleatorias marginales**, cada una de ellas.

Caracterizaciones de independencia para el caso CONTINUO

- ❶ Sea $X = (X_1, \dots, X_n)$ un vector aleatorio **continuo**. Se tiene que

$$X_i : (\Omega, \mathcal{A}, P) : \longrightarrow (\mathbb{R}, \mathcal{B}, P_{X_i}), \quad i = 1, \dots, n$$

continuas, son **independientes si y sólo si**

$$f_X(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n), \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n.$$

- ❷ Sea $X = (X_1, \dots, X_n)$ un vector aleatorio **continuo**. Se tiene que

$$X_i : (\Omega, \mathcal{A}, P) : \longrightarrow (\mathbb{R}, \mathcal{B}, P_{X_i}), \quad i = 1, \dots, n$$

continuas, son **independientes si y sólo si**, para cualquier $(x_1, \dots, x_n) \in \mathbb{R}^n$,

$$f_X(x_1, \dots, x_n) = h_1(x_1) \cdots h_n(x_n),$$

donde $h_i : \mathbb{R} \longrightarrow \mathbb{R}$, $i = 1, \dots, n$, son funciones arbitrarias.

Se puede decir que las componentes de un vector aleatorio **continuo** son **variables aleatorias independientes** si su **función de densidad** es **producto de las funciones de densidad marginales** o, **producto de funciones** que solo dependen de las **variables aleatorias marginales**, cada una de ellas.

Caracterización de independencia en términos de la Distribución de Probabilidad

Sea $X = (X_1, \dots, X_n)$ un **vector aleatorio con distribución de probabilidad conjunta** $P_X : \mathcal{B}^n \rightarrow [0, 1]$. Sean X_1, \dots, X_n , variables aleatorias con distribuciones de probabilidad, P_{X_1}, \dots, P_{X_n} , respectivamente, entonces, son **independientes si y sólo si, para cualesquiera subconjuntos de Borel** B_1, \dots, B_n de la recta real, se da la siguiente identidad:

$$\begin{aligned} P_X(B_1 \times \dots \times B_n) &= P(X_1 \in B_1, \dots, X_n \in B_n) \\ &= P_{X_1}(B_1) \cdots P_{X_n}(B_n) = P(X_1 \in B_1) \cdots P(X_n \in B_n). \end{aligned}$$

Outline

- 1 Definición y caracterización
- 2 Propiedades de la independencia**
- 3 Reproductividad de distribuciones
- 4 Independencia de vectores aleatorios

Propiedad 1

Si X sobre (Ω, \mathcal{A}, P) es una **variable aleatoria degenerada**, i.e. $P(X = c) = 1$, entonces X es **independiente de cualquier otra variable aleatoria** definida sobre el mismo espacio de probabilidad.

Demostración:

- Nótese que si Y se define sobre (Ω, \mathcal{A}, P) **discreta**

$$\begin{aligned} P(Y = y, X = c) &= P(Y = y) = p_Y(y) \times 1 \\ &= P(Y = y)P(X = c) = p_Y(y)p_X(c). \end{aligned}$$

- Si Y es **continua**, para cualquier $B_1, B_2 \in \mathcal{B}$,

$$P(Y \in B_1, X \in B_2) = \begin{cases} \int_{B_1} f_Y(y) dy & \text{si } c \in B_2 \\ 0 & \text{si } c \in \mathbb{R} \setminus B_2 \end{cases}$$

En ambos casos, se tiene la factorización de la probabilidad conjunta, teniendo en cuenta, que $P(X \in B_2) = 1$, si $c \in B_2$, y $P(X \in B_2) = 0$, si $c \in \mathbb{R} \setminus B_2$.

Propiedad 2

Si X_1, \dots, X_n son independientes, **cualquier subconjunto de ellas también lo son.**

Demostración:

- Esta propiedad se obtiene de forma inmediata, considerando la definición de la función masa de probabilidad marginal del subconjunto considerado, a partir de la función masa de probabilidad conjunta, en el caso discreto.
- Igualmente, se deduce, en el caso continuo, considerando la definición de la densidad de probabilidad marginal del subconjunto considerado, a partir de la densidad de probabilidad conjunta.

Propiedad 3

Si X_1, \dots, X_n son independientes, **todas las distribuciones condicionadas coinciden con las marginales** correspondientes.

Demostración:

- Esta propiedad se obtiene también de forma directa, a partir de la definición de función masa de probabilidad y función de densidad condicionadas.
- Por ejemplo, en el caso bidimensional, si X e Y son v.a. continuas independientes, entonces

$$f_{X/Y}(x) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)} = \frac{f_X(x)f_Y(y)}{f_Y(y)} = f_X(x),$$

$$\forall y \in \text{Supp}(f_Y), \quad \forall x \text{ tal que } (x,y) \in \text{Supp}(f_{(X,Y)}).$$

$\text{Supp}(\cdot)$ se refiere al **soporte de la función de densidad**, es decir, donde toma valores no nulos.

Propiedad 4. Caracterización de la independencia por medio de la FGM

Sean X_1, \dots, X_n v.a. tal que existen $M_{X_i}(t)$, para $t \in (-a_i, b_i)$, $a_i, b_i \in \mathbb{R}^+$, $i = 1, \dots, n$. Entonces, X_1, \dots, X_n **son independientes** si y solo si para cualquier $(t_1, \dots, t_n) \in (-a_1, b_1) \times \dots \times (-a_n, b_n)$:

$$M_{X_1, \dots, X_n}(t_1, \dots, t_n) = M_{X_1}(t_1) \cdots M_{X_n}(t_n).$$

Demostración:

Veamos la demostración en el caso continuo (el caso discreto se demuestra de forma análoga).

$$\begin{aligned} M_{X_1, \dots, X_n}(t_1, \dots, t_n) &= \int_{\mathbb{R}^n} \exp\left(\sum_{i=1}^n t_i X_i\right) f_{X_1, \dots, X_n}(x_1, \dots, x_n) dx_1 \cdots dx_n \\ &= \int_{-\infty}^{\infty} \exp(t_1 X_1) f_{X_1}(x_1) dx_1 \cdots \int_{-\infty}^{\infty} \exp(t_n X_n) f_{X_n}(x_n) dx_n = M_{X_1}(t_1) \cdots M_{X_n}(t_n), \\ &\forall (t_1, \dots, t_n) \in (-a_1, b_1) \times \dots \times (-a_n, b_n). \end{aligned}$$

Teorema de multiplicación de esperanzas

(i) Si X_1, \dots, X_n son independientes, y existe $E[X_i]$ para $i = 1, \dots, n$, se tiene que

$$\exists E[X_1 \cdots X_n] = E[X_1] \cdots E[X_n].$$

(ii) Si X_1, \dots, X_n son independientes y $g_1, \dots, g_n : \mathbb{R} \rightarrow \mathbb{R}$ son funciones medibles, las variables aleatorias $g_1(X_1), \dots, g_n(X_n)$ son independientes. Se tiene entonces

$$E[g_1(X_1) \cdots g_n(X_n)] = E[g_1(X_1)] \cdots E[g_n(X_n)].$$

(iii) Si X e Y son independientes entonces $\text{Cov}(X, Y) = 0$.

(iv) Si X_1, \dots, X_n son independientes con momento no centrado de orden dos finito

$$\text{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \text{Var}(X_i), \quad a_1, \dots, a_n \in \mathbb{R}.$$

Teorema de multiplicación de esperanzas (demostración caso continuo)

En particular, (i) se deriva de forma directa de la siguiente expresión:

$$\begin{aligned} E[X_1 \cdots X_n] &= \int_{\mathbb{R}^n} x_1 \cdots x_n f_{X_1, \dots, X_n}(x_1, \dots, x_n) dx_1 \cdots dx_n \\ &= \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1) dx_1 \cdots \int_{-\infty}^{\infty} x_n f_{X_n}(x_n) dx_n = E[X_1] \cdots E[X_n] \end{aligned}$$

(ii) se obtiene considerando que para cualesquiera B_i , $i = 1, \dots, n$, subconjuntos de la σ -álgebra de Borel \mathcal{B} de \mathbb{R}^n ,

$$\begin{aligned} P_{g_1(X_1), \dots, g_n(X_n)}(B_1, \dots, B_n) &= P(g_1(X_1) \in B_1, \dots, g_n(X_n) \in B_n) = P(X_1 \in g_1^{-1}(B_1), \dots, X_n \in g_n^{-1}(B_n)) \\ &= P_{X_1, \dots, X_n}(g_1^{-1}(B_1), \dots, g_n^{-1}(B_n)) = \prod_{i=1}^n P_{X_i}(g_i^{-1}(B_i)) = \prod_{i=1}^n P(g_i(X_i) \in B_i) = \prod_{i=1}^n P_{g_i(X_i)}(B_i). \end{aligned}$$

(iii) Si X e Y son independientes, entonces $\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$.

(iv) Si X_1, \dots, X_n son independientes con momento no centrado de orden dos finito

$$\begin{aligned} \text{Var} \left(\sum_{i=1}^n a_i X_i \right) &= E \left(\left[\sum_{i=1}^n a_i X_i \right]^2 \right) - \left[\sum_{i=1}^n a_i E[X_i] \right]^2 \\ &= \sum_{i,j} a_i a_j E[X_i X_j] - \sum_{i,j} a_i a_j E[X_i] E[X_j] = \sum_{i,j} a_i a_j [E[X_i X_j] - E[X_i] E[X_j]] \\ &= \sum_{i,j} \delta_{i,j} a_i^2 [E[X_i^2] - [E[X_i]]^2] = \sum_{i=1}^n a_i^2 [E[X_i^2] - [E[X_i]]^2] = \sum_{i=1}^n a_i^2 \text{Var}(X_i) \end{aligned}$$

donde $\delta_{i,j}$ representa la delta de Kronecker que vale 1 cuando $i = j$ y vale cero cuando $i \neq j$.

Matriz de covarianzas (apéndice)

Dado $X = (X_1 \cdots X_n)$ un vector aleatorio n -dimensional. Se define su **matriz de covarianzas** como la matriz formada por todas las covarianzas de sus componentes. Se denota por Σ_X y tiene la expresión:

$$\Sigma_X = (\text{cov}(X_i, X_j))_{i,j=1}^n = \begin{pmatrix} \sigma_{X_1}^2 & \cdots & \text{cov}(X_1, X_n) \\ \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \cdots & \sigma_{X_n}^2 \end{pmatrix}.$$

Propiedades de la matriz de covarianzas

- Es una matriz simétrica y por tanto diagonalizable.
- Si hay independencia dos a dos de las componentes del vector aleatorio se tiene que Σ_X es una matriz diagonal.
- La matriz de covarianzas es el punto de partida para las distintas **técnicas de aprendizaje supervisado y no supervisado en machine learning** (Análisis de componentes principales, análisis factorial, análisis discriminante, etc).

Outline

- 1 Definición y caracterización
- 2 Propiedades de la independencia
- 3 Reproductividad de distribuciones**
- 4 Independencia de vectores aleatorios

Reproductividad de la Binomial

Dadas $X_i \sim B(k_i, p)$, $i = 1, \dots, n$, v.a. independientes, entonces:

$$\sum_{i=1}^n X_i \sim B\left(\sum_{i=1}^n k_i, p\right). \quad (1)$$

La **demostración** de esta afirmación se obtiene de forma inmediata a partir de la propiedad 4 de independencia, considerando la expresión de la función generatriz de momentos M_{X_i} de X_i , que viene dada por

$$M_{X_i}(t) = (pe^t + (1-p))^{k_i}, \quad \forall t \in \mathbb{R}, \quad i = 1, \dots, n.$$

Específicamente,

$$\begin{aligned} M_{\sum_{i=1}^n X_i}(t) &= E\left[\exp\left(t \sum_{i=1}^n X_i\right)\right] = E\left[\prod_{i=1}^n \exp(tX_i)\right] = \prod_{i=1}^n E[\exp(tX_i)] \quad (2) \\ &= \prod_{i=1}^n (pe^t + (1-p))^{k_i} = (pe^t + (1-p))^{\sum_{i=1}^n k_i}. \end{aligned}$$

La expresión (1) se deduce de forma directa de (2), dado que $(pe^t + (1-p))^{\sum_{i=1}^n k_i}$ es la función generatriz de momentos de una Binomial con parámetros $\sum_{i=1}^n k_i$ y p .

Reproductividad de la Poisson

Dadas $X_i \sim \mathcal{P}(\lambda_i)$, $i = 1, \dots, n$, independientes, entonces:

$$\sum_{i=1}^n X_i \sim \mathcal{P}\left(\sum_{i=1}^n \lambda_i\right) \quad (3)$$

Para la **demostración** se aplica la independencia y se utiliza que

$$M_{X_i}(t) = \exp(\lambda_i(e^t - 1)), \quad t \in \mathbb{R}, \quad i = 1, \dots, n.$$

Reproductividad de la Binomial Negativa

Dadas $X_i \sim BN(k_i, p)$, $i = 1, \dots, n$, independientes, entonces:

$$\sum_{i=1}^n X_i \sim BN\left(\sum_{i=1}^n k_i, p\right) \quad (4)$$

Para la **demostración** se aplica la independencia y se utiliza que

$$M_{X_i}(t) = \left[\frac{p}{1 - (1-p)e^t} \right]^{k_i}, \quad t \in \mathbb{R}, \quad i = 1, \dots, n.$$

Reproductividad de la Geométrica

Dadas $X_i \sim G(p)$, $i = 1, \dots, n$, independientes, entonces:

$$\sum_{i=1}^n X_i \sim BN(n, p) \quad (5)$$

La **demostración** es inmediata a partir de la propiedad anterior, dado que $X_i \sim G(p)$, coincide con $X_i \sim BN(1, p)$, i.e., $k_i = 1$, $i = 1, \dots, n$.

Reproductividad de la Normal

Dadas $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, $i = 1, \dots, n$, independientes, entonces:

$$\sum_{i=1}^n X_i \sim \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right) \quad (6)$$

La **demostración** es inmediata a partir de la independencia y de la siguiente expresión:

$$M_{X_i}(t) = \exp\left(t\mu_i + \frac{t^2\sigma_i^2}{2}\right), \quad t \in \mathbb{R}, \quad i = 1, \dots, n.$$

Reproductividad de la Gamma

Dadas $X_i \sim \Gamma(p_i, a)$, $i = 1, \dots, n$, independientes, entonces:

$$\sum_{i=1}^n X_i \sim \Gamma\left(\sum_{i=1}^n p_i, a\right) \quad (7)$$

La **demostración** es inmediata a partir de la independencia y de la siguiente expresión:

$$M_{X_i}(t) = \left[1 - \frac{t}{a}\right]^{-p_i}, \quad t < a, \quad i = 1, \dots, n.$$

En **consecuencia**, si $X_i \sim \mathcal{E}(k_i, a)$, $i = 1, \dots, n$, independientes, entonces:

$$\sum_{i=1}^n X_i \sim \Gamma\left(\sum_{i=1}^n k_i, a\right) \equiv \mathcal{E}\left(\sum_{i=1}^n k_i, a\right). \quad (8)$$

Reproductividad de la Exponencial

Dadas $X_i \sim \exp(\lambda)$, $i = 1, \dots, n$, independientes, entonces:

$$\sum_{i=1}^n X_i \sim \Gamma(n, \lambda) \equiv \mathcal{E}(n, \lambda) \quad (9)$$

La **demostración** se obtiene de la independencia y de la siguiente expresión:

$$M_{X_i}(t) = \left[1 - \frac{t}{\lambda}\right]^{-1}, \quad t < \lambda, \quad i = 1, \dots, n.$$

Outline

- 1 Definición y caracterización
- 2 Propiedades de la independencia
- 3 Reproductividad de distribuciones
- 4 Independencia de vectores aleatorios**

Independencia de familias de variables aleatorias

Se considera un conjunto T arbitrario de índices, normalmente, un conjunto infinito numerable. Se introduce entonces una familia de variables aleatorias $\{X_t, t \in T\}$, asociada a los índices del conjunto T , todas ellas definidas sobre el mismo espacio probabilístico base (Ω, \mathcal{A}, P) . Se dice entonces que:

- Las variables aleatorias $\{X_t, t \in T\}$ son **mutuamente independientes**, si para cualquier $k \in \mathbb{N}$, las variables aleatorias X_{t_1}, \dots, X_{t_k} son independientes, $t_1, \dots, t_k \in T$.
- Las variables aleatorias $\{X_t, t \in T\}$ son **independientes dos a dos**, si para cualesquiera índices t_i, t_j , del conjunto T , con $t_i \neq t_j$, $i \neq j$, las variables aleatorias X_{t_i} y X_{t_j} , son independientes.

Como caso particular, cuando $T = \mathbb{N}$, se obtienen las caracterizaciones de **independencia mutua** e **independencia dos a dos para sucesiones de variables aleatorias** $\{X_n, n \in \mathbb{N}\}$.

Independencia de vectores aleatorios

La definición y caracterizaciones de independencia para vectores aleatorios se pueden formular de forma sencilla, como extensión directa de las estudiadas para variables aleatorias unidimensionales.

Más concretamente, para m **vectores** X_1, \dots, X_m , definidos sobre el mismo espacio probabilístico base (Ω, \mathcal{A}, P) , con dimensiones posiblemente diferentes, denotadas por n_1, \dots, n_m , se dice que X_1, \dots, X_m **son independientes** si su **función de distribución de probabilidad conjunta**, $n_1 + \dots + n_m$ -dimensional, **factoriza en producto de las funciones de distribución de probabilidad marginales** n_i dimensionales, $i = 1, \dots, m$.

$$F_{X_1, \dots, X_m}(x_1, \dots, x_m) = F_{X_1}(x_1) \times \dots \times F_{X_m}(x_m), \quad \forall (x_1, \dots, x_m),$$

donde

$$x_i = (x_{1,i}, \dots, x_{n_i,i}) \in \mathbb{R}^{n_i}, \quad i = 1, \dots, m.$$

Caracterización de independencia de vectores aleatorios

De forma similar se **caracteriza la independencia**, en el caso **discreto y continuo**, en términos de la factorización de la función masa de probabilidad y función de densidad de probabilidad conjuntas $n_1 + \dots + n_m$ -dimensionales, en términos de las marginales n_i -dimensionales, $i = 1, \dots, m$.

Es decir, en el caso **discreto y continuo**, X_1, \dots, X_m son independientes si y sólo si,

$$P_{X_1, \dots, X_m}(x_1, \dots, x_m) = \prod_{i=1}^m P_{X_i}(x_i), \quad \forall (x_1, \dots, x_m) \in E_{X_1} \times \dots \times E_{X_m}$$

$$f_{X_1, \dots, X_m}(x_1, \dots, x_m) = \prod_{i=1}^m f_{X_i}(x_i), \quad \forall (x_1, \dots, x_m) \in \mathbb{R}^{n_1 + \dots + n_m},$$

respectivamente.

Adicionalmente, en **términos de la distribución de probabilidad**, X_1, \dots, X_m son **independientes si y sólo si**, para cualesquiera $B_i \in \mathcal{B}^{n_i}$, $i = 1, \dots, m$, se tienen las siguientes identidades:

$$P_{X_1, \dots, X_m}(B_1 \times \dots \times B_m) = P(X_1 \in B_1, \dots, X_m \in B_m) = \prod_{i=1}^m P(X_i \in B_i) = \prod_{i=1}^m P_{X_i}(B_i).$$

Caracterización en términos de la función generatriz de momentos

Supongamos que **existen las funciones generatrices de momentos** de X_1, \dots, X_m , definidas, respectivamente, en los intervalos $I_i = \prod_{j=1}^{n_i} (-a_{j,i}, b_{j,i})$, $a_{j,i}, b_{j,i} > 0$, $I_i \subseteq \mathbb{R}^{n_i}$, $i = 1, \dots, m$.

Entonces, X_1, \dots, X_m son **independientes si y sólo si**, $\forall (t_1, \dots, t_m) \in I_1 \times \dots \times I_m$,

$$M_{X_1, \dots, X_m}(t_1, \dots, t_m) = E \left[\exp \left(\sum_{i=1}^m \langle t_i, X_i \rangle \right) \right]$$

$$\prod_{i=1}^m E [\exp (\langle t_i, X_i \rangle)] = \prod_{i=1}^m M_{X_i}(t_i).$$

Propiedades de independencia entre vectores aleatorios

Se **pueden reformular** de forma inmediata las **propiedades de independencia** estudiadas para variables aleatorias unidimensionales.

En particular, se tiene

- Si X_1, \dots, X_m son **independientes**, cualquier **subconjunto** X_{i_1}, \dots, X_{i_k} , $0 < k < n$, de X_1, \dots, X_m está **constituido por vectores aleatorios independientes**.
- Si X_1, \dots, X_m , son vectores aleatorios **independientes**, para cualesquiera g_i , $i = 1, \dots, m$, aplicaciones medibles, los vectores aleatorios $g_1(X_1), \dots, g_m(X_m)$, son **también independientes**.