TEMA 3: Independencia de variables aleatorias

Contenidos

- Definición y caracterizaciones.
- Propiedades de independencia.
- Reproductividad de distribuciones.
- Independencia de vectores aleatorios.

Independencia de las componentes de un vector aleatorio

Sea $\mathbf{X} = (X_1, \ldots, X_n)$ un vector aleatorio. Se dice que las variables $X_i : (\Omega, \mathcal{A}, P) :\longrightarrow (\mathbb{R}, \mathcal{B}, P_{X_i})$, $i = 1, \ldots, n$, son independientes si la función de distribución conjunta $F_{\mathbf{X}}$ factoriza en producto de las marginales, i.e.,

$$F_{\mathbf{X}}(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n).$$

Caso Discreto. Caracterización de independencia

Caracterización 1. Sea $\mathbf{X} = (X_1, \ldots, X_n)$ un vector aleatorio discreto. Se tiene que $X_i : (\Omega, \mathcal{A}, P) :\longrightarrow (E_{X_i}, \mathcal{B}, P_{X_i})$, $i = 1, \ldots, n$, son independientes si y sólo si, para cualquier $(x_1, \ldots, x_n) \in E_{X_1} \times \cdots \times E_{X_n}$,

$$P_{\mathbf{X}}(x_1, \ldots, x_n) = p_{X_1}(x_1) \cdots p_{X_n}(x_n) = P[X_1 = x_1] \cdots P[X_n = x_n].$$

Caracterización 2. Sea $\mathbf{X} = (X_1, \ldots, X_n)$ un vector aleatorio discreto. Se tiene que $X_i : (\Omega, \mathcal{A}, P) :\longrightarrow (E_{X_i}, \mathcal{B}, P_{X_i})$, $i = 1, \ldots, n$, son independientes si y sólo si, para cualquier $(x_1, \ldots, x_n) \in E_{X_1} \times \cdots \times E_{X_n}$,

$$P_{\mathbf{X}}(x_1, \ldots, x_n) = h_1(x_1) \cdots h_n(x_n),$$

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

Caso Continuo. Caracterización de independencia

Caracterización 1. Sea $\mathbf{X} = (X_1, \ldots, X_n)$ un vector aleatorio continuo. Se tiene que $X_i : (\Omega, \mathcal{A}, P) :\longrightarrow (\mathbb{R}, \mathcal{B}, P_{X_i})$, $i = 1, \ldots, n$, continuas, son independientes si y sólo si

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

Caracterización 2. Sea $\mathbf{X} = (X_1, \ldots, X_n)$ un vector aleatorio continuo. Se tiene que $X_i : (\Omega, \mathcal{A}, P) :\longrightarrow (\mathbb{R}, \mathcal{B}, P_{X_i})$, $i = 1, \ldots, n$, continuas, son independientes si y sólo si, para cualquier $(x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$f_{\mathbf{X}}(x_1, \ldots, x_n) = h_1(x_1) \cdots h_n(x_n),$$

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

Caracterización en términos de la Distribución de Probabilidad

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

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

Propiedades de la independencia de variables aleatorias

1. Si X sobre $(\Omega, \mathcal{A}, P)$ 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. Nótese que si Y se define sobre $(\Omega, \mathcal{A}, P)$ 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 $B_1, B_2 \in \mathcal{B}$,

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

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 $c \in \mathbb{R} \backslash B_2$.

2. Si $X_1, \ldots, X_n$ son independientes, cualquier subconjunto de ellas también lo son. 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.

3. Si $X_1, \ldots, X_n$ son independientes, todas las distribuciones condicionadas coinciden con las marginales correspondientes. 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

$$\begin{aligned} f_{X/Y}(x) &= \frac{f_{(X,Y)}(x,y)}{f_Y(y)} = \frac{f_X(x) f_Y(y)}{f_Y(y)} = f_X(x) \\ & \forall y \in \mathrm{Supp}(f_Y) \quad \forall x \text{ tal que } \quad (x,y) \in \mathrm{Supp}(f_{(X,Y)}) \end{aligned}$$

4. Si $X_1, \ldots, X_n$ son independientes y existe $M_{X_i}(t)$, para $t \in (-a_i, b_i)$, $a_i, b_i \in \mathbb{R}_+$, $i = 1, \ldots, n$. Entonces existe $M_{X_1,\ldots,X_n}(t_1, \ldots, t_n)$, para cualquier $(t_1, \ldots, t_n) \in (-a_1, b_1) \times \cdots \times (-a_n, b_n)$, y se tiene la siguiente igualdad:

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

para cualquier $(t_1, \ldots, t_n) \in (-a_1, b_1) \times \cdots \times (-a_n, b_n)$. Veamos la demostración en el caso continuo (el caso discreto se demuestra de forma análoga).

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

Teorema de multiplicación de esperanzas:

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

(ii) Si $X_1, \ldots, X_n$ son independientes y $g_1, \ldots, g_n : \mathbb{R} \longrightarrow \mathbb{R}$ son funciones medibles, satisfaciendo $E[g_i(X_i)] < \infty$, para $i = 1, \ldots, n$, las variables aleatorias $g_1(X_1), \ldots, g_n(X_n)$ 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 $\mathrm{Cov}(X, Y) = 0$.

(iv) Si $X_1, \ldots, X_n$ son independientes con momento no centrado de orden dos finito
$$\mathrm{Var}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i), \quad a_1, \ldots, a_n \in \mathbb{R}.$$

Demostración La demostración de (i)–(iv) se hará en el caso continuo (de forma análoga se obtiene el caso discreto). En particular, (i) se deriva de forma directa de la siguiente expresión:

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

(ii) se obtiene considerando la siguiente identidad, para cualesquiera B_i, $i = 1, \ldots, n$, subconjuntos de la σ-álgebra de Borel $\mathcal{B}$ sobre la recta real $\mathbb{R}$,

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

(iii) Si X e Y son independientes, entonces

$$\mathrm{Cov}(X, Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0.$$

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

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

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

Distribuciones reproductivas

En este apartado se hace referencia a familias de distribuciones que son estables frente a la suma de variables aleatorias independientes. Específicamente,

- **Distribución Binomial**

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

donde $X_i \sim B(k_i, p)$, $i = 1, \ldots, n$, independientes. 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

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

Específicamente,

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

La ecuación (1) se obtiene 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.

De forma similar se demuestran las siguientes afirmaciones:

- $X_i \sim \mathcal{P}(\lambda_i)$, $i = 1, \ldots, n$, independientes

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

Para la demostración aplicar la independencia y utilizar que

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

- $X_i \sim BN(k_i, p)$, $i = 1, \ldots, n$, independientes

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

Para la demostración aplicar la independencia y utilizar que

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

- $X_i \sim G(p)$, $i = 1, \ldots, n$, independientes

$$\sum_{i=1}^n X_i \sim BN\left(n, p\right). \qquad (5)$$

Se obtiene de forma inmediata a partir de la propiedad anterior, dado que $X_i \sim G(p)$, coincide con $X_i \sim BN(1, p)$, i.e., $k_i = 1$, $i = 1, \ldots, n$.

- $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, $i = 1, \ldots, n$, independientes
$$\sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right). \tag{6}$$
se obtiene de forma inmediata a partir de la independencia y de la siguiente expresión:
$$M_{X_i}(t) = \exp\left(t\mu_i + \frac{t^2\sigma_i^2}{2}\right), \quad t \in \mathbb{R}, \quad i = 1, \ldots, n.$$
- $X_i \sim \Gamma(p_i, a)$, $i = 1, \ldots, n$, independientes
$$\sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} p_i, a\right). \tag{7}$$
Se obtiene de forma inmediata a partir de la independencia y de la siguiente expresión:
$$M_{X_i}(t) = \left[1 - \frac{t}{a}\right]^{-p_i}, \quad t < a, \quad i = 1, \ldots, n.$$
- *Consecuencia*: $X_i \sim \mathcal{E}(k_i, a)$, $i = 1, \ldots, n$, independientes
$$\sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} k_i, a\right) \equiv \mathcal{E}\left(\sum_{i=1}^{n} k_i, a\right). \tag{8}$$
- $X_i \sim \exp(\lambda)$, $i = 1, \ldots, n$, independientes
$$\sum_{i=1}^{n} X_i \sim \Gamma\left(n, \lambda\right) \equiv \mathcal{E}(n, \lambda). \tag{9}$$
Se obtiene directamente de la independencia y de la siguiente expresión:
$$M_{X_i}(t) = \left[1 - \frac{t}{\lambda}\right]^{-1}, \quad t < \lambda, \quad i = 1, \ldots, n.$$

Independencia para 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 $(\Omega, \mathcal{A}, P)$. Se considerarán ahora los conceptos de independencia dos a dos, e independencia mutua para dicha familia de variables aleatorias. Específicamente:

Independencia mutua

Las variables aleatorias $\{X_t,\ t \in T\}$ son mutuamente independientes, si para cualquier $k \in \mathbb{N}$, las variables aleatorias $X_{t_1}, \ldots, X_{t_k}$ son independientes, $t_1, \ldots, t_k \in T$.

Independencia dos a dos

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 $\mathbf{X}_1, \ldots, \mathbf{X}_m$, definidos sobre el mismo espacio probabilístico base $(\Omega, \mathcal{A}, P)$, con dimensiones posiblemente diferentes, denotadas por $n_1, \ldots, n_m$, se dice que $\mathbf{X}_1, \ldots, \mathbf{X}_m$ son independientes si su función de distribución de probabilidad conjunta, $n_1 + \cdots + n_m$-dimensional, factoriza en producto de las funciones de distribución de probabilidad marginales n_i dimensionales, $i = 1, \ldots, m$.

$$F_{\mathbf{X}_1,\ldots,\mathbf{X}_m}(\mathbf{x}_1, \ldots, \mathbf{x}_m) = F_{\mathbf{X}_1}(\mathbf{x}_1) \times \cdots \times F_{\mathbf{X}_m}(\mathbf{x}_m), \quad \forall (\mathbf{x}_1, \ldots, \mathbf{x}_m),$$

donde

$$\mathbf{x}_i = (x_{1,i}, \ldots, x_{n_i,i}) \in \mathbb{R}^{n_i}, \quad i = 1, \ldots, m.$$

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 $n_1 + \cdots + n_m$-dimensionales, en términos de las marginales n_i-dimensionales, $i = 1, \ldots, m$. Específicamente, en el caso discreto y continuo, $\mathbf{X}_1, \ldots, \mathbf{X}_m$ son independientes si y sólo si,

$$\begin{aligned} P_{\mathbf{X}_1,\ldots,\mathbf{X}_m}(\mathbf{x}_1,\ldots,\mathbf{x}_m) &= \prod_{i=1}^{m} P_{\mathbf{X}_i}(\mathbf{x}_i), \quad \forall (\mathbf{x}_1,\ldots,\mathbf{x}_m) \in E_{\mathbf{X}_1} \times \cdots \times E_{\mathbf{X}_m} \\ f_{\mathbf{X}_1,\ldots,\mathbf{X}_m}(\mathbf{x}_1,\ldots,\mathbf{x}_m) &= \prod_{i=1}^{m} f_{\mathbf{X}_i}(\mathbf{x}_i), \quad \forall (\mathbf{x}_1,\ldots,\mathbf{x}_m) \in \mathbb{R}^{n_1+\cdots+n_m}, \end{aligned}$$

respectivamente.

Adicionalmente, en términos de la distribución de probabilidad, $\mathbf{X}_1, \ldots, \mathbf{X}_m$ son independientes si y sólo si, para cualesquiera $B_i \in \mathcal{B}^{n_i}$, $i = 1, \ldots, m$, se tienen las siguientes identidades:

$$\begin{aligned} P_{\mathbf{X}_1,\ldots,\mathbf{X}_m}(B_1 \times \cdots \times B_m) &= P(\mathbf{X}_1 \in B_1, \ldots, \mathbf{X}_m \in B_m) \\ &= \prod_{i=1}^{m} P(\mathbf{X}_i \in B_i) = \prod_{i=1}^{m} P_{\mathbf{X}_i}(B_i). \end{aligned}$$

Caracterización en términos de la función generatriz de momentos

Supongamos que existen las funciones generatrices de momentos de $\mathbf{X}_1, \ldots, \mathbf{X}_m$, respectivamente definidas en los intervalos $\mathbf{I}_i = \prod_{j=1}^{n_i}(-a_{j,i}, b_{j,i})$, $a_{j,i}, b_{j,i} > 0$, $\mathbf{I}_i \subseteq \mathbb{R}^{n_i}$, $i = 1, \ldots, m$. Entonces, $\mathbf{X}_1, \ldots, \mathbf{X}_m$ son independientes si y sólo si, para cualesquiera $(\mathbf{t}_1, \ldots, \mathbf{t}_m) \in \mathbf{I}_1 \times \cdots \times \mathbf{I}_m$,

$$\begin{aligned} M_{\mathbf{X}_1,\ldots,\mathbf{X}_m}(\mathbf{t}_1,\ldots,\mathbf{t}_m) &= E\left[\exp\left(\sum_{i=1}^{m} \langle \mathbf{t}_i, \mathbf{X}_i \rangle\right)\right] \\ \prod_{i=1}^{m} E\left[\exp\left(\langle \mathbf{t}_i, \mathbf{X}_i \rangle\right)\right] &= \prod_{i=1}^{m} M_{\mathbf{X}_i}(\mathbf{t}_i). \end{aligned}$$

Propiedades de independencia entre vectores aleatorios

Se comprueba de forma inmediata la reformulación de las propiedades de independencia estudiadas para variables aleatorias unidimensionales. En particular, se tiene

- Si $\mathbf{X}_1, \ldots, \mathbf{X}_m$ son independientes, cualquier subconjunto $\mathbf{X}_{i_1}, \ldots, \mathbf{X}_{i_k}$, $0 < k < n$, de $\mathbf{X}_1, \ldots, \mathbf{X}_m$ está constituido por vectores aleatorios independientes.

- Si $\mathbf{X}_1, \ldots, \mathbf{X}_m$, son vectores aleatorios independientes, para cualesquiera g_i, $i = 1, \ldots, m$, aplicaciones medibles, los vectores aleatorios $g_1(\mathbf{X}_1), \ldots, g_m(\mathbf{X}_m)$, son también independientes.