Tema 2.- Vectores aleatorios

Asignatura: PROBABILIDAD

Grado en Ingeniería Informática y Matemáticas

(3er Curso - 1er semestre)

©Prof. Dr. José Luis Romero Béjar

Departamento de Estadística e Investigación Operativa
Facultad de Ciencias (Despacho 10)

Periodo de docencia: 12/09/2022 a 20/12/2022

Esquema de contenidos

Esquema de contenidos

σ-álgebra de Borel sobre $\mathbb{R}^n$

- En este tema se considerará la σ-álgebra $\mathcal{B}^n \subset \mathcal{P}(\mathbb{R}^n)$ de **Borel** sobre $\mathbb{R}^n$, definida como la mínima σ-álgebra que contiene todos los intervalos de $\mathbb{R}^n$.
- De forma similar al caso unidimensional, se demuestra que, en particular, dicha σ-álgebra se **puede generar a partir de la colección de todos los intervalos** de $\mathbb{R}^n$ de la forma

$$(-\infty, \mathrm{x}] = (-\infty, x_1] \times \cdots \times (-\infty, x_n], \quad \mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n. \tag{1}$$

Definición

Un **vector aleatorio** $\mathrm{X} = (X_1, \ldots, X_n)$ sobre un espacio probabilístico base $(\Omega, \mathcal{A}, P)$ se define como una función $\mathrm{X} : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n)$ de modo que:

$$\mathrm{X}^{-1}(B) = \{\omega \in \Omega;\ \mathrm{X}(\omega) \in B\} \in \mathcal{A}, \quad \forall B \in \mathcal{B}^n,$$

donde $X_i : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B})$, $i = 1, \ldots, n$.

Caracterización

$\mathrm{X} = (X_1, \ldots, X_n)$ es un **vector aleatorio si y sólo** si para cualquier $\mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$\mathrm{X}^{-1}((-\infty, \mathrm{x}]) = \{\omega \in \Omega / X_1(\omega) \leq x_1, \ldots, X_n(\omega) \leq x_n\} \in \mathcal{A}$$
$$\{\mathrm{X} \leq \mathrm{x}\} = \{X_1 \leq x_1, \ldots, X_n \leq x_n\} \in \mathcal{A}$$

Esta caracterización se obtiene de forma inmediata a partir del resultado que prueba que $\mathcal{B}^n$ se puede generar a partir de todos los intervalos de $\mathbb{R}^n$ de la forma (1).

Caracterización en términos de sus componentes

Teorema de medibilidad.
$X = (X_1, \ldots, X_n)$ es un vector aleatorio sobre $(\Omega, \mathcal{A}, P)$ si y sólo si X_i es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$, para $i = 1, \ldots, n$.

Demostración

(i) Supongamos que $\mathrm{X} = (X_1, \ldots, X_n)$ es un vector aleatorio sobre $(\Omega, \mathcal{A}, P)$. Se tiene entonces, en particular, para todo $x_i \in \mathbb{R}$, y para todo $i = 1, \ldots, n$, que

$$\begin{aligned} X_i^{-1}((-\infty, x_i]) &= \{\omega \in \Omega; X_i(\omega) \in (-\infty, x_i],\ X_j(\omega) \in \mathbb{R},\ i \neq j\} \\ &= \mathrm{X}^{-1}((-\infty, \infty) \times \cdots \times (-\infty, x_i] \times \cdots \times (-\infty, \infty)) \in \mathcal{A}, \end{aligned}$$

donde la última identidad se obtiene a partir de la expresión (1) anterior.
(ii) Si X_i, $i = 1, \ldots, n$, son variables aleatorias unidimensionales sobre $(\Omega, \mathcal{A}, P)$, entonces para cualquier vector $\mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, se tiene

$$\begin{aligned} \mathrm{X}^{-1}((-\infty, \mathrm{x}]) &= \{\omega \in \Omega / X_1(\omega) \leq x_1, \ldots, X_n(\omega) \leq x_n\} \\ &= \bigcap_{i=1}^{n} X_i^{-1}((-\infty, x_i]) \in \mathcal{A}, \end{aligned}$$

donde en la obtención de la última ecuación se ha aplicado que X_i es v.a., por tanto, $X_i^{-1}((-\infty, x_i]) \in \mathcal{A}$, para todo $i = 1, \ldots, n$, así como el hecho de que, por definición, una σ-álgebra, en particular $\mathcal{A}$, es cerrada para la intersección de conjuntos.

Esquema de contenidos

Distribución de probabilidad

Sea $\mathrm{X} = (X_1, \ldots, X_n)$ un vector aleatorio sobre $(\Omega, \mathcal{A}, P)$. Se define entonces su **distribución de probabilidad** $P_{\mathrm{X}} : \mathcal{B}^n \longrightarrow [0, 1]$ como una función de conjuntos sobre $\mathcal{B}^n$ satisfaciendo

$$P_{\mathrm{X}}(B) = P\{\omega \in \Omega;\ \mathrm{X}(\omega) \in B\} = P(\mathrm{X} \in B), \quad \forall B \in \mathcal{B}^n. \tag{2}$$

Proposición

La función de conjuntos $P_{\mathrm{X}}(B)$ definida en (2) ***es una medida de probabilidad****.*

Demostración

Se comprueba que $P_{\mathbf{X}}$ satisface los tres axiomas que determinan una medida de probabilidad. Es decir,

- **A1** $P_{\mathbf{X}}(B) = P(\mathbf{X} \in B) \geq 0, \quad \forall B \in \mathcal{B}^n$, dado que P es una medida de probabilidad.
- **A2** $P_{\mathbf{X}}(\mathbb{R}^n) = P(\Omega) = 1$, dado que P es una medida de probabilidad.
- **A3** Sea $\{B_i\}_{i\in\mathbb{N}} \subset \mathcal{B}^n$; $B_i \cap B_j =, \emptyset, \; i \neq j$. Se tiene entonces

$$P_{\mathbf{X}}\left(\bigcup_{n\in\mathbb{N}} B_n\right) = P\left(\mathbf{X} \in \bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n\in\mathbb{N}} P\left(\mathbf{X} \in B_n\right) = \sum_{n\in\mathbb{N}} P_{\mathbf{X}}(B_n). \tag{3}$$

donde para la obtención de estas identidades se ha aplicado la definición de $P_{\mathbf{X}}$ y la aditividad de P como medida de probabilidad.

Observación:

Por tanto, a partir de la Proposición anterior, se tiene que **un vector aleatorio** X **se define como una función medible** entre los espacios $(\Omega, \mathcal{A}, P)$ y $(\mathbb{R}^n, \mathcal{B}^n, P_{\mathrm{X}})$, i.e.,

$$\mathrm{X} : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n, P_{\mathrm{X}}).$$

Función de distribución de un vector aleatorio

La **función de distribución de un vector aleatorio** X se define como una función

$$F_{\mathrm{X}} : \mathbb{R}^n \longrightarrow [0,1]$$

dada por

$$\begin{aligned} F_{\mathrm{X}}(\mathrm{x}) = P_{\mathrm{X}}\left((-\infty, \mathrm{x}]\right) &= P(\mathrm{X} \in (-\infty, \mathrm{x}]) \\ &= P(X_1 \leq x_1, \ldots, X_n \leq x_n), \quad \forall \mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n. \end{aligned}$$

Propiedades de la función de distribución de un vector aleatorio

F_X satisface las siguientes propiedades (que **caracterizan a las funciones de distribución**):

(i) $\forall i = 1, \ldots, n$, y para cualesquiera $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \in \mathbb{R}$,

$$x_i < x_i' \Rightarrow F_X(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) \leq F_X(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n).$$

(ii) $\forall i = 1, \ldots, n$, y para cualesquiera $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \in \mathbb{R}$,

$$\lim_{x_i' \to x_i; x_i' > x_i} F_X(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n) = F_X(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n).$$

(iii) $\forall i = 1, \ldots, n$, y para cualesquiera $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \in \mathbb{R}$,

$$\begin{aligned} \exists \lim_{x_i \to -\infty} & F_X(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) \\ = \ & F_X(x_1, \ldots, x_{i-1}, -\infty, x_{i+1}, \ldots, x_n) = 0. \end{aligned}$$

(iv) Se tiene el siguiente límite:

$$\exists \lim_{x_i \to \infty,\ i=1,\ldots,n} F_X(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) = F_X(\infty, \ldots, \infty) = 1.$$

(v) $\forall x_1, \ldots, x_n \in \mathbb{R}$, $\forall i = 1, \ldots, n$,

(i) $$\lim_{\epsilon \to 0, \epsilon > 0} F_X(x_1, \ldots, x_i - \epsilon, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_i < x_i, \ldots, X_n \leq x_n)$$
$$= F_X(x_1, \ldots, x_i^-, \ldots, x_n)$$

(ii) $$P(X_1 \leq x_1, \ldots, X_{i-1} \leq x_{i-1}, X_i = x_i, X_{i+1} \leq x_{i+1}, \ldots, X_n \leq x_n)$$
$$= F_X(x_1, \ldots, x_i, \ldots, x_n) - F_X(x_1, \ldots, x_i^-, \ldots, x_n)$$

La última igualdad es nula cuando la función de distribución F_X es continua en el argumento i-ésimo en el punto x_i.

(vi) $\forall x_1, \ldots, x_n \in \mathbb{R}$, y para cualquier $\varepsilon_1, \ldots, \varepsilon_n \in \mathbb{R}_+$:

$$\begin{aligned} & F_X(x_1 + \varepsilon_1, \ldots, x_n + \varepsilon_n) \\ & - \sum_{i=1}^{n} F_X(x_1 + \varepsilon_1, \ldots, x_{i-1} + \varepsilon_{i-1}, x_i, x_{i+1} + \varepsilon_{i+1}, \ldots, x_n + \varepsilon_n) \\ & + \sum_{i=1}^{n} \sum_{j=1}^{n} F_X\,(x_1 + \varepsilon_1, \ldots, x_{i-1} + \varepsilon_{i-1}, x_i, x_{i+1} + \varepsilon_{i+1}, \\ & \qquad \ldots, x_{j-1} + \varepsilon_{j-1}, x_j, x_{j+1} + \varepsilon_{j+1}, \ldots, x_n + \varepsilon_n) - \\ & \cdots + (-1)^n F_X(x_1, \ldots, x_n) \geq 0 \end{aligned}$$

Demostración de la monotonía de la función de distribución n-dimensional

A continuación se derivará la prueba de (vi), considerando el caso $n = 2$, a partir del cual, el caso de un n general se obtendría por inducción.

$$\begin{aligned}
&F_{\mathbf{X}}(x_1+\varepsilon_1, x_2+\varepsilon_2) - F_{\mathbf{X}}(x_1, x_2+\varepsilon_2) - F_{\mathbf{X}}(x_1+\varepsilon_1, x_2) + F_{\mathbf{X}}(x_1, x_2)\\
&= P(X_1 \le x_1+\varepsilon_1, X_2 \le x_2+\varepsilon_2) - P(X_1 \le x_1, X_2 \le x_2+\varepsilon_2) - (P(X_1 \le x_1+\varepsilon_1, X_2 \le x_2) + P(X_1 \le x_1, X_2 \le x_2))\\
&= P(x_1 < X_1 \le x_1+\varepsilon_1, X_2 \le x_2+\varepsilon_2) - P(x_1 < X_1 \le x_1+\varepsilon_1, X_2 \le x_2)\\
&= P(x_1 < X_1 \le x_1+\varepsilon_1, x_2 < X_2 \le x_2+\varepsilon_2) \ge 0 \quad \text{(p.q } P \text{ es una medida de probabilidad).}
\end{aligned}$$

Observación:

- Todas las propiedades anteriores se demuestran de forma similar al caso unidimensional.
- Esta propiedad (vi) se puede ver como **una extensión al caso n-dimensional de la propiedad de monotonía** (no decreciente) de la función de distribución. De hecho, para el caso $n = 1$ se formularía como $F_X(x+\varepsilon) - F_X(x) \ge 0$, para cualquier $\varepsilon > 0$, que es cierto por ser F_X una función monótona no decreciente.
- Las propiedades (i)-(iv) y (vi) **caracterizan a la función de distribución de un vector aleatorio**, que determina de forma unívoca la distribución de probabilidad de dicho vector, según se enuncia en el siguiente resultado.

Esquema de contenidos

Teorema de correspondencia

Existe una correspondencia biunívoca entre funciones de probabilidad P sobre $(\mathbb{R}^n, \mathcal{B}^n)$ y las funciones $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ satisfaciendo las propiedades (i)-(iv) y (vi) anteriores. Dicha correspondencia está determinada por la relación:

$$P((-\infty, \mathrm{x}]) = F(\mathrm{x}), \quad \forall \mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Más concretamente,

(a) Si P es una medida de probabilidad sobre $(\mathbb{R}^n, \mathcal{B}^n)$, la función $F : \mathbb{R}^n \longrightarrow \mathbb{R}$, definida mediante la igualdad $F(\mathrm{x}) = P((-\infty, \mathrm{x}])$, para cualquier $\mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, satisface las propiedades (i)-(iv) y (vi).

(b) Recíprocamente, si $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función que satisface (i)-(iv) y (vi), entonces existe una única medida de probabilidad P sobre $(\mathbb{R}^n, \mathcal{B}^n)$, verificando $P((-\infty, \mathrm{x}]) = F(\mathrm{x})$, para cualquier $\mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Demostración (indicaciones)

(a) Esta demostración consiste en probar las propiedades (i)-(iv) y (vi), que, como ya se ha comentado anteriormente, se obtienen de forma análoga al caso unidimensional, para el caso de las propiedades (i)-(iv). (En relación con la propiedad (vi), dicha propiedad se ha derivado en la sección anterior).

(b) Se requiere probar que conocida $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ se puede calcular $P(B)$, para cualquier $B \in \mathcal{B}^n$. Es decir, para cualquier $B \in \mathcal{B}^n$, $P(B)$ se puede calcular a partir del conocimiento de $P((-\infty, \mathrm{x}]) = F(\mathrm{x})$, para cualquier $\mathrm{x} \in \mathbb{R}^n$. Esta última afirmación se puede demostrar a partir de los conceptos de **conjunto elemental**, i.e., conjunto que se puede expresar como unión finita de intervalos disjuntos de $\mathbb{R}^n$, y de conjunto **σ-elemental**, i.e., conjunto que se puede expresar como la unión numerable de conjuntos elementales disjuntos.

Corolario

(a) **La distribución de probabilidad de un vector aleatorio determina y es determinada por su función de distribución.**

(b) Las propiedades (i)-(iv) y (vi) **caracterizan a las funciones de distribución de vectores aleatorios** (toda función que las cumpla es la función de distribución de un vector aleatorio con la dimensión correspondiente).

Cálculo de probabilidades de vectores bidimensionales

Nos centraremos ahora en el **cálculo de probabilidades de intervalos de** $\mathbb{R}^n$, ya que se puede abordar de forma más directa y sencilla, a partir de la función de distribución $F_{\mathbf{X}}$ de un vector aleatorio **X**.

$$P_{\mathbf{X}}(I_1 \times I_2) = P(\mathbf{X} \in I_1 \times I_2) = P(X_1 \in I_1, X_2 \in I_2),$$

siendo I_1, I_2 intervalos de $\mathbb{R}$, e $I_1 = (a,b), (a,b], [a,b), [a,b],\ a,b \in \mathbb{R} \bigcup \{\infty, +\infty\},\ a \leq b$.

- $P(X_1 \in (a,b), X_2 \in I_2) = P(X_1 < b, X_2 \in I_2) - P(X_1 \leq a, X_2 \in I_2), \quad a < b$
- $P(X_1 \in (a,b], X_2 \in I_2) = P(X_1 \leq b, X_2 \in I_2) - P(X_1 \leq a, X_2 \in I_2)$
- $P(X_1 \in [a,b), X_2 \in I_2) = P(X_1 < b, X_2 \in I_2) - P(X_1 < a, X_2 \in I_2)$
- $P(X_1 \in [a,b], X_2 \in I_2) = P(X_1 \leq b, X_2 \in I_2) - P(X_1 < a, X_2 \in I_2)$

Se considera ahora

$$P(X_1 \leq x, X_2 \in I_2),\ \ I_2 = (c,d), (c,d], [c,d), [c,d],\ \ x,c,d \in \mathbb{R},\ \ c \leq d.$$

- $P(X_1 \leq x, X_2 \in (c,d)) = P(X_1 \leq x, X_2 < d) - P(X_1 \leq x, X_2 \leq c) = F_{\mathbf{X}}(x, d^-) - F_{\mathbf{X}}(x,c), \quad c < d$
- $P(X_1 \leq x, X_2 \in (c,d]) = P(X_1 \leq x, X_2 \leq d) - P(X_1 \leq x, X_2 \leq c) = F_{\mathbf{X}}(x, d) - F_{\mathbf{X}}(x,c)$
- $P(X_1 \leq x, X_2 \in [c,d)) = P(X_1 \leq x, X_2 < d) - P(X_1 \leq x, X_2 < c) = F_{\mathbf{X}}(x, d^-) - F_{\mathbf{X}}(x,c^-)$
- $P(X_1 \leq x, X_2 \in [c,d]) = P(X_1 \leq x, X_2 \leq d) - P(X_1 \leq x, X_2 < c) = F_{\mathbf{X}}(x, d) - F_{\mathbf{X}}(x,c^-)$

De forma ánaloga se calcularía

$$P(X_1 < x, X_2 \in I_2),\ \ I_2 = (c,d), (c,d], [c,d), [c,d],\ \ x,c,d \in \mathbb{R},\ \ c \leq d.$$

Esquema de contenidos

Definición

Un vector aleatorio X es **discreto** si su conjunto de valores es numerable. Es decir, si existe $E_X \subset \mathbb{R}^n$, numerable, tal que

$$P_X(E_X) = P(X \in E_X) = 1.$$

Función masa de probabilidad

Se define la **función masa de probabilidad** de un vector aleatorio discreto como:

$$p_{\mathrm{X}} : E_{\mathrm{X}} \longrightarrow [0,1]$$
$$p_{\mathrm{X}}(\mathrm{x}) = P(\mathrm{X} = \mathrm{x}) = P(X_1 = x_1, \ldots, X_n = x_n)$$

Se verifica:

- $P(\mathrm{X} = \mathrm{x}) \geq 0, \quad \forall \mathrm{x} \in E_{\mathrm{X}}$
- $\sum_{\mathrm{x} \in E_{\mathrm{X}}} P(\mathrm{X} = \mathrm{x}) = 1$

Caracterización

Toda función real-valuada y no negativa, definida sobre un subconjunto numerable de $\mathbb{R}^n$, tal que la **suma de sus valores es uno**, es la función masa de probabilidad de un vector aleatorio n-dimensional con valores en dicho conjunto.

Distribución de probabilidad y función de distribución

Se definen la **distribución de probabilidad** y la **función de distribución** de un vector aleatorio discreto como:

$$\begin{aligned}
P_{\mathrm{X}}(B) &= P(\mathrm{X} \in B) = \sum_{\mathrm{x} \in B \cap E_{\mathbf{X}}} P(\mathrm{X} = \mathrm{x}) \\
&= \sum_{\mathrm{x} \in B \cap E_{\mathbf{X}}} P(X_1 = x_1, \ldots, X_n = x_n), \quad \forall B \in \mathcal{B}^n \\
F_{\mathrm{X}}(\mathrm{x}) &= F_{\mathrm{X}}(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n) \\
&= \sum_{\mathrm{x}' \in E_{\mathbf{X}};\ x_i' \leq x_i,\ i=1,\ldots,n} P(\mathrm{X} = \mathrm{x}') \\
&= \sum_{\mathrm{x}' \in E_{\mathbf{X}};\ x_i' \leq x_i,\ i=1,\ldots,n} P(X_1 = x_1', \ldots, X_n = x_n') \\
&\qquad \forall \mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n
\end{aligned}$$

Caracterización de vector aleatorio discreto

Un vector aleatorio $\mathbf{X} = (X_1, \ldots, X_n) : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n)$ **es discreto** si y sólo si, para $i = 1, \ldots, n$, $X_i : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B})$ es **una variable aleatoria discreta unidimensional**.

Demostración ($\Rightarrow$)

Supongamos que $\mathbf{X}$ es un vector aleatorio discreto, i.e., $\exists E_{\mathbf{X}} \subset \mathbb{R}^n$ tal que $P(\mathbf{X} \in E_{\mathbf{X}}) = 1$. Definamos el subconjunto

$$\Psi^i_{\mathbb{R}}(E_{\mathbf{X}}) = \{y \in \mathbb{R};\ \exists \mathbf{x} \in E_{\mathbf{X}},\ \mathbf{x} = (x_1, \ldots, x_i = y, \ldots, x_n)\}.$$

Dicho conjunto representa la proyección de $E_{\mathbf{X}}$ sobre la i-ésima componente. Dado que $E_{\mathbf{X}}$ es numerable, $\Psi^i_{\mathbb{R}}(E_{\mathbf{X}})$ es también numerable. Adicionalmente,

$$\begin{aligned} P(X_i \in \Psi^i_{\mathbb{R}}(E_{\mathbf{X}})) &= P\left(X_1 \in \mathbb{R}, \ldots, X_i \in \Psi^i_{\mathbb{R}}(E_{\mathbf{X}}), \ldots, X_n \in \mathbb{R}\right) \\ &\geq P\left(\mathbf{X} \in E_{\mathbf{X}}\right) = 1. \end{aligned}$$

Por tanto, X_i es una variable aleatoria unidimensional discreta, para $i = 1, \ldots, n$.

Demostración ($\Leftarrow$)

Supongamos ahora que X_i es una variable aleatoria unidimensional discreta, para $i = 1, \ldots, n$, i.e., existe un subconjunto numerable $E_{X_i} \subset \mathbb{R}$, tal que $P(X_i \in E_{X_i}) = 1$, para $i = 1, \ldots, n$. Por tanto el subconjunto de $\mathbb{R}^n$ dado por $E_{X_1} \times \cdots \times E_{X_n}$ es también numerable. Se tiene entonces,

$$\begin{aligned} P\left(\mathbf{X} = (X_1, \ldots, X_n) \in E_{X_1} \times \cdots \times E_{X_n}\right) &= P(X_1 \in E_{X_1}, \ldots, X_n \in E_{X_n}) \\ = P\left(\bigcap_{i=1}^{n} \left\{X_i \in E_{X_i}\right\}\right) &\underset{\text{Desigualdad de Boole}}{\geq} 1 - \sum_{i=1}^{n} P\left(\left\{X_i \in \mathbb{R} \backslash E_{X_i}\right\}\right) \\ &= 1 - \sum_{i=1}^{n} P(X_i \in E_{X_i}^c) = 1 \end{aligned}$$

Observación

- El resultado anterior nos indica que cualquier conjunto finito de variables aleatorias discretas puede definir un vector aleatorio discreto de dimensión dada por el número de variables.
- Recíprocamente, un vector aleatorio discreto determina un conjunto finito de variables aleatorias discretas.

A continuación se van a ilustrar unos cuantos **ejemplos de cálculo de la función de distribución** de un vector aleatorio discreto.

Ejemplo 1

Sea un vector aleatorio bidimensional discreto, $\mathrm{X} = (X_1, X_2)$ con función masa de probabilidad definida mediante las siguientes igualdades

$$\begin{aligned} P((1,2)) &= P((1,3)) = P((2,2)) = P((2,3)) = 1/6 \\ P((3,3)) &= 2/6 \end{aligned}$$

Función de distribución bivariante asociada

$$F_{\mathrm{X}}(x_1, x_2) = \begin{cases} 0 & \text{si} \quad x_1 < 1 \quad \text{ó} \quad x_2 < 2 \\ 1/6 & \text{si} \quad 1 \leq x_1 < 2,\ 2 \leq x_2 < 3 \\ 1/3 & \text{si} \quad 1 \leq x_1 < 2; x_2 \geq 3 \text{ ó } 2 \leq x_2 < 3;\ 2 \leq x_2 < 3 \text{ ó } x \geq 3; 2 \leq x_2 < \\ 2/3 & \text{si} \quad 2 \leq x_1 < 3,\ x_2 \geq 3 \\ 1 & \text{si} \quad x_1 \geq 3,\ x_2 \geq 3 \end{cases}$$

Para **calcular**, por ejemplo, $P((x_1, x_2);\ x_1 + x_2 = 4)$.

$$P((x_1, x_2);\ x_1 + x_2 = 4) = P((1,3)) + P((2,2)) = 1/3.$$

a) Probar que es un vector aleatorio discreto

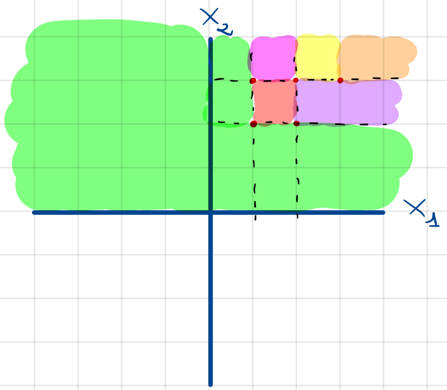
$X_1 \backslash X_2$	2	3
1	1/6	1/6
2	1/6	1/6
3	0	2/6

$E = E_1 \times E_2 = \{1,2,3\} \times \{2,3\} \longrightarrow P[X \in E] = 1$

$\Sigma = 1$

b) Función de distribución

$\forall (a,b) \in \mathbb{R}^2 ; F(a,b) = P[X_1 \leq a, X_2 \leq b] \quad \forall a,b \in \mathbb{R}$



$$F(a,b) = \begin{cases} 0 & X_1 < 1 \text{ ó } X_2 < 2 \\ 1/6 & 1 \leq X_1 < 2 \text{ y } 2 \leq X_2 < 3 \\ 2/6 & 2 \leq X_1 \text{ y } 2 \leq X_2 < 3 \\ 2/6 & 1 \leq X_1 < 2 \text{ y } X_2 \geq 3 \\ 4/6 & 2 \leq X_1 < 3 \text{ y } X_2 \geq 3 \\ 1 & X_1 \geq 3 \text{ y } X_2 \geq 2 \end{cases}$$

c) $P[(x_1, x_2) / x_1 + x_2 = 4]$

$\hookrightarrow P[(1,3),(2,2)] = 1/6 + 1/6 = 1/3$

Hallar la función de distribución

Ejemplo 2

Sea X $= (X, Y)$ el vector aleatorio discreto con función masa de probabilidad definida mediante la siguiente tabla:

$X \backslash Y$	-1	1
1	1/6	1/3
2	1/12	1/4
3	1/12	1/12

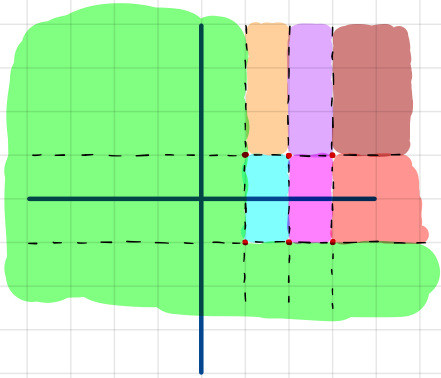
Se calculan las siguientes probabilidades:

- $P(X \leq 2, Y > 0) = P(X = 1, Y = 1) + P(X = 2, Y = 1) = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$
- $P(X \geq 2) = P(X = 2, Y = -1) + P(X = 2, Y = 1) + P(X = 3, Y = -1) + P(X = 3, Y = 1) = \frac{1}{12} + \frac{1}{4} + \frac{1}{12} + \frac{1}{12} = \frac{6}{12} = \frac{1}{2}$
- $P(Y < 0) = P(X = 1, Y = -1) + P(X = 2, Y = -1) + P(X = 3, Y = -1) = \frac{1}{6} + \frac{1}{12} + \frac{1}{12} = \frac{1}{3}$

Función de distribución

$\forall (a,b) \in \mathbb{R}^2 ;\ F(a,b) = P[X_1 \leq a, X_2 \leq b]$

$E_x = E_{x_1} \times E_{x_2} = \{1,2,3\} \times \{-1,1\}$



$$F(a,b) = \begin{cases} 0 & X_1 < 1 \text{ ó } X_2 < -1 \\ 1/6 & 1 \leq X_1 < 2 \text{ y } -1 \leq X_2 < 1 \\ 1/2 & 1 \leq X_1 < 2 \text{ y } X_2 \geq 1 \\ 1/4 & 2 \leq X_1 < 3 \text{ y } -1 \leq X_2 < 1 \\ 5/6 & 2 \leq X_1 < 3 \text{ y } X_2 \geq 1 \\ 1/3 & X_1 \geq 3 \text{ y } -1 \leq X_2 < 1 \\ 1 & X_1 \geq 3 \text{ y } X_2 \geq 1 \end{cases}$$

Esquema de contenidos

Definición

Un vector aleatorio $\mathrm{X} : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n)$ es de tipo **continuo** si existe una función $f_{\mathrm{X}} : \mathbb{R}^n \longrightarrow \mathbb{R}$, tal que

$$F_{\mathrm{X}}(\mathrm{x}) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f_{\mathrm{X}}(t_1, \ldots, t_n) dt_1, \ldots, dt_n, \quad \forall \mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Función de densidad

La función f_{X} recibe el nombre de **función de densidad de probabilidad**.
Como consecuencia se tiene que la función de distribución F_{X} es:

- continua sobre $\mathbb{R}^n$, derivable salvo un conjunto de medida nula de $\mathbb{R}^n$.
- tiene derivada continua sobre el dominio donde se define F_{X}.

Propiedades de la función de densidad

La función de densidad presenta las siguientes propiedades:

- $\lim_{x_i \to -\infty,\ i=1,\ldots,n} f_X(x_1, \ldots, x_n) = \lim_{x_i \to \infty,\ i=1,\ldots,n} f_X(x_1, \ldots, x_n) = 0.$
- El conjunto de discontinuidades de f_X en $\mathbb{R}^n$ es numerable, es decir, tiene medida nula.
- Si $x = (x_1, \ldots, x_n)$ es un punto de continuidad de f_X, entonces

$$\exists \frac{d^n F_X(x_1, \ldots, x_n)}{dx_1 \ldots dx_n} = f_X(x_1, \ldots, x_n).$$

- Los valores de f_X pueden modificarse en un conjunto de medida nula sin afectar a F_X (como primitiva de f_X en sus puntos de continuidad).
- Puesto que f_X determina F_X, también determina la distribución de probabilidad. Es decir,

$$\forall B \in \mathcal{B}^n,\ P_X(B) = \int_B f_X(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

  En particular, si E es un subconjunto numerable de $\mathbb{R}^n$, entonces

$$P_X(E) = \int_E f_X(x_1, \ldots, x_n) dx_1 \ldots dx_n = 0.$$

Caracterización de función de densidad

- Si $f_X : \mathbb{R}^n \longrightarrow \mathbb{R}$ es la función de densidad de un vector aleatorio, entonces es no negativa e integrable y su integral sobre $\mathbb{R}^n$ vale uno, i.e.,

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(t_1, \ldots, t_n) dt_1 \ldots dt_n = 1.$$

- Recíprocamente, si f_X es no negativa e integrable y su integral sobre $\mathbb{R}^n$ vale uno, entonces es la función de densidad de un vector aleatorio n-dimensional.

Observación

En el caso de vectores aleatorios continuos, **no se tiene la equivalencia estudiada para el caso de vectores aleatorios discretos**, en relación con el carácter continuo de sus componentes. Más concretamente, **se tiene sólo la implicación formulada en el siguiente resultado**.

Caracterización de vector aleatorio continuo

Sea $\mathrm{X} = (X_1, \ldots, X_n) : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n, P_{\mathrm{X}})$ es un vector aleatorio continuo. **Entonces,** $X_i : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n, P_{X_i})$ es una variable aleatoria continua, para $i = 1, \ldots, n$.

Demostración

Para cada $i = 1, \ldots, n$,

$$\begin{aligned} F_{X_i}(x_i) &= P(X_i \leq x_i) = P(X_1 \in \mathbb{R}, \ldots, X_i \leq x_i, \ldots, X_n \in \mathbb{R}) \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{x_i} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(t_1, \ldots, t_n) dt_1 \ldots dt_n \\ &= \int_{-\infty}^{x_i} \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(t_1, \ldots, t_n) dt_1 \ldots dt_{i-1} dt_{i+1} \ldots dt_n \right] dt_i \end{aligned}$$

Definimos entonces la función $f_{X_i} : \mathbb{R} \longrightarrow \mathbb{R}$ dada por

$$f_{X_i}(t_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(t_1, \ldots, t_i, \ldots, t_n) dt_1 \ldots dt_{i-1} dt_{i+1} \ldots dt_n.$$

Dicha función es no negativa e integrable y su integral sobre $\mathbb{R}$ vale uno. Adicionalmente, se tiene la siguiente igualdad:

$$F_{X_i}(x_i) = \int_{-\infty}^{x_i} f_{X_i}(t_i) dt_i.$$

Por tanto, X_i es de tipo continuo con densidad de probabilidad f_{X_i}.

Contraejemplo

Se considera seguidamente un contraejemplo, donde **se observa que un conjunto finito de variables continuas no definen necesariamente un vector aleatorio continuo**.

Sea (X_1, X_2) un vector aleatorio tal que X_1 es de tipo continuo y $X_2 = 2X_1 + k$, para un cierto $k > 0$.

Entonces (X_1, X_2) no es un vector aleatorio continuo, ya que si lo fuera

$$P(X_2 = 2X_1 + k) = \int_{\underbrace{\{(x_1,x_2)\in\mathbb{R}^2;\ x_2=2x_1+k\}}} f_{(X_1,X_2)}(t_1, t_2)dt_1dt_2 = 0.$$

Este conjunto se trata de una recta (medida nula), de ahí que sea 0 la integral

Sin embargo, a partir de la definición de dicho vector se tiene:

$$P(X_2 = 2X_1 + k) = 1.$$

A continuación se van a ilustrar unos cuantos **ejemplos de cálculo de la función de distribución** de un vector aleatorio continuo.

Ejemplo 1

Se considera la siguiente función:

$$f_{\mathbf{X}}(x,y) = \begin{cases} k & 0 < y < x < 1 \\ 0 & \text{resto de valores} \end{cases}$$

Calcular k para que $f_{\mathbf{X}}$ sea la función de densidad de probabilidad de un vector aleatorio bidimensional $\mathbf{X} = (X, Y)$ y hallar la función de distribución asociada.

- Para que $f_{\mathbf{X}}$ sea una función de densidad de probabilidad se debe verificar $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbf{X}}(x,y)dxdy = 1$. En efecto $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbf{X}}(x,y)dxdy = \int_0^1\int_0^x kdydx = \int_0^1 k\,[y]_{y=0}^{y=x}\,dx = k\int_0^1 xdx = k\left[\frac{x^2}{2}\right]_{x=0}^{x=1} = \frac{k}{2}$, de donde se deduce que $k = 2$ para que $f_{\mathbf{X}}$ sea la función de densidad de probabilidad.

El profe aconseja integrar primero respecto de y

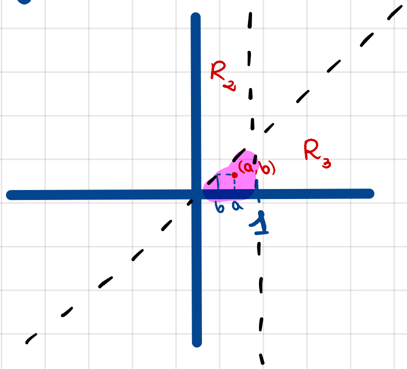
- Función de distribución de probabilidad

¡ESTÁ MAL!

Ver abajo

$$F_{\mathbf{X}}(x,y) = \begin{cases} 0 & x < 0 \quad \text{ó} \quad x > 0\ y < 0 \\ 2\left[\int_y^x\int_0^y dvdu + \int_0^y\int_0^u dvdu\right] & 0 < y < x < 1,\ 0 < x \le 1 \\ 2\int_0^x\int_0^u dvdu = x^2 & 0 < x < y < 1,\ 0 < y \le 1 \\ 1 & x > 1\ y > 1 \end{cases}$$

•) $f_{(x,y)}(x,y) \geq 0 \Rightarrow k \geq 0$



$$f_{(x,y)}(x,y) = \begin{cases} 2 & 0<y<x<1 \\ 0 & \end{cases}$$

Función de distribución

$$F_{(x,y)}(a,b) = P[X \leq a, Y \leq b] = \begin{cases} 0 & a \leq 0 \text{ ó } b \leq 0 \\ \int_0^b\int_0^x 2\,dy\,dx + \int_b^a\int_0^b 2\,dy\,dx & 0<b<a<1 \\ \int_0^a\int_0^x 2\,dy\,dx & b>a \quad 0<a<1 \\ \int_0^b\int_0^x 2\,dy\,dx + \int_b^1\int_0^b 2\,dy\,dx & 0<a<1 \quad b>a \\ 1 & a \geq 1 \text{ y } b \geq 1 \end{cases}$$

Ejemplo 2

Sea $f_{\mathrm{X}}(x,y) = \begin{cases} x+y & 0<x<1,\ 0<y<1 \\ 0 & \text{en el resto} \end{cases}$. Calcular la función de distribución de probabilidad F_{X} del vector aleatorio $\mathrm{X} = (X, Y)$, con función de densidad de probabilidad f_{X}.

Se tiene

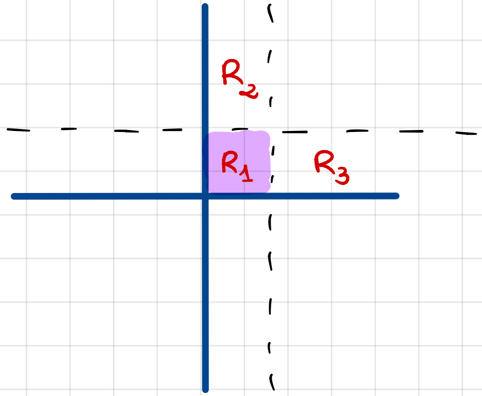
$$F_{\mathrm{X}}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{\mathrm{X}}(u,v)\,dvdu$$

$$= \int_0^x \int_0^y (u+v)\,dvdu = \int_0^x \left(u[v]_{v=0}^{v=y} + \left[\frac{v^2}{2}\right]_{v=0}^{v=y} \right) du$$

$$= \int_0^x \left[uy + \frac{y^2}{2} \right] du = \frac{1}{2}\left(yx^2 + y^2x\right), \quad (x,y) \in [0,1) \times [0,1)$$

Por tanto,

$$F_{\mathrm{X}}(x,y) = \begin{cases} 0 & x<0 \text{ ó } y<0 \\ \frac{1}{2}\left(yx^2+y^2x\right) & 0\le x<1, \quad 0 \le y < 1 \\ \frac{1}{2}(x^2+x) & 0 \le x < 1, \quad y \ge 1 \\ \frac{1}{2}(y^2+y) & x \ge 1, \quad 0 \le y < 1 \\ 1 & x \ge 1,\ y \ge 1 \end{cases}$$



Función de distribución

$$F_{(x,y)}(a,b) = \begin{cases} 0 & a \le 0 \text{ ó } b \le 0 \\ \int_0^a \int_0^b (x+y)\,dy\,dx & 0<a<1 \text{ y } 0<b<1 \\ \int_0^a \int_0^1 (x+y)\,dy\,dx & 0<a<1 \text{ y } b \ge 1 \\ \int_0^1 \int_0^b (x+y)\,dy\,dx & a \ge 1 \text{ y } 0<b<1 \\ 1 & a \ge 1 \text{ y } b \ge 1 \end{cases}$$

Ejemplo 3

Sea $f_{\mathbf{X}}(x,y) = \begin{cases} k\left[\frac{xy}{2}+1\right] & 0<x<1,\ -1<y<1 \\ 0 & \text{en otro caso} \end{cases}$

- Calcular k para que $f_{\mathbf{X}}$ sea una función de densidad de probabilidad.
  Se tiene que $f_{\mathbf{X}}(x,y) \geq 0 \Leftrightarrow k \geq 0$.
  Adicionalmente, debe verificarse:

$$\int_{-\infty}^{\infty} f_{\mathbf{X}}(x,y)dxdy = 1 = \int_0^1 \int_{-1}^1 k\left[\frac{xy}{2}+1\right] dydx = \int_0^1 k\left[x\left[\frac{y^2}{4}\right]_{y=-1}^{y=1} + [y]_{y=-1}^{y=1}\right] dx = \int_0^1 2kdx = 2k$$

  de donde se deduce que $k = 1/2$.

- Calcular la función de distribución de probabilidad $F_{\mathbf{X}}$.

$$\begin{aligned} F_{\mathbf{X}}(x,y) &= \int_0^x \int_{-1}^y \left[\frac{uv+2}{4}\right] dvdu \\ &= \frac{1}{4}\int_0^x \left(\left[u\frac{v^2}{2}\right]_{v=-1}^{v=y} + 2[v]_{v=-1}^{v=y}\right) du = \frac{1}{4}\int_0^x \left(\frac{uy^2}{2} - \frac{u}{2} + 2y + 2\right) du \\ &= \frac{x^2(y^2-1)}{16} + \frac{x(y+1)}{2}, \quad 0 \leq x < 1, \quad -1 \leq y < 1 \end{aligned}$$

Se tiene entonces que $F_{\mathbf{X}}(x,y) = \begin{cases} 0 & x<0 \text{ ó } y<-1 \\ \frac{x^2(y^2-1)}{16} + \frac{x(y+1)}{2} & 0 \leq x < 1,\ -1 \leq y < 1 \\ x & 0 \leq x < 1,\ y \geq 1 \\ \frac{y^2+8y+7}{16} & x \geq 1,\ -1 \leq y < 1 \\ 1 & x \geq 1,\ y \geq 1 \end{cases}$

Ejemplo 4

Dada la v.a. bidimensional continua (X, Y), cuya función de densidad de probabilidad conjunta viene dada por:

$$f(x,y) = \begin{cases} cx^2y & x^2 \leq y \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

Calcular el valor de la constante c para que f sea una función de densidad de probabilidad y, en tal caso, calcular la función de distribución de probabilidad asociada.

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy \\ &= \int_{-1}^{1}\left[\int_{x^2}^{1} cx^2ydy\right]dx = \int_{-1}^{1} cx^2\left[\frac{y^2}{2}\right]_{y=x^2}^{y=1} dx \\ &= \int_{-1}^{1}\left[\frac{cx^2}{2} - \frac{cx^6}{2}\right]dx = \left[\frac{cx^3}{6} - \frac{cx^7}{14}\right]_{x=-1}^{x=1} = \frac{4c}{21}. \end{aligned}$$

Por tanto, $c = 21/4$. Por tanto,

$$f(x,y) = \begin{cases} \frac{21}{4}x^2y & x^2 \leq y \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

Ejemplo 5

Calcular la función de distribución de probabilidad asociada a la función de densidad $f_{\mathbf{X}}$ del vector aleatorio $\mathbf{X} = (X_1, X_2)$, cuyos valores se encuentran en el recinto delimitado por las rectas $x = 0$, $y = 0$, y $x + y = 1$.

Dicho recinto lo llamaremos R_1 y viene dado por:

$$R_1 = \{(x, y) \in \mathbb{R}^2;\ 0 < x < 1,\ 0 < y < 1 - x < 1\}.$$

Para los puntos dentro de este recinto, la función de distribución asociada $F_{\mathbf{X}}$ se calcula como sigue:

$$F_{\mathbf{X}}(x, y) = \int_0^x \int_0^y f_{\mathbf{X}}(u, v)dvdu, \quad \forall (x, y) \in R_1.$$

Adicionalmente, para el recinto R_2 definido por

$$R_2 = \{(x, y) \in \mathbb{R}^2;\ x < 0,\ \text{ó } x > 0 \text{ e } y < 0\},$$

se obtiene

$$F_{\mathbf{X}}(x, y) = 0, \quad \forall (x, y) \in R_2.$$

El siguiente dominio R_3 vendría dado por:

$$R_3 = \{(x, y) \in \mathbb{R}^2;\ 1 - x < y < 1;\ 0 < x < 1\}$$

Ejemplo 5 (continuación)

En este recinto,

$$\begin{aligned} F_{\mathbf{X}}(x,y) &= \int_{\mathbf{0}}^{x}\int_{\mathbf{0}}^{\mathbf{1}-x} f_{\mathbf{X}}(u,v)dvdu \\ &+ \int_{\mathbf{0}}^{\mathbf{1}-y}\int_{\mathbf{1}-x}^{y} f_{\mathbf{X}}(u,v)dvdu \\ &+ \int_{\mathbf{1}-y}^{x}\int_{\mathbf{1}-x}^{\mathbf{1}-u} f_{\mathbf{X}}(u,v)dvdu, \quad \forall (x,y)\in R_{\mathbf{3}} \end{aligned}$$

Se considera ahora el recinto

$$R_{\mathbf{4}} = \{(x,y);\ x \geq 1;\ 0 < y < 1\}.$$

En este recinto,

$$\begin{aligned} F_{\mathbf{X}}(x,y) &= \int_{\mathbf{0}}^{\mathbf{1}-y}\int_{\mathbf{0}}^{y} f_{\mathbf{X}}(u,v)dvdu \\ &+ \int_{\mathbf{1}-y}^{\mathbf{1}}\int_{\mathbf{0}}^{\mathbf{1}-u} f_{\mathbf{X}}(u,v)dvdu \end{aligned}$$

Finalmente, se considera el recinto

$$R_{\mathbf{5}} = \{(x,y);\ 0 < x < 1;\ y \geq 1\}.$$

$$\begin{aligned} F_{\mathbf{X}}(x,y) &= \int_{\mathbf{0}}^{x}\int_{\mathbf{0}}^{\mathbf{1}-x} f_{\mathbf{X}}(u,v)dvdu \\ &+ \int_{\mathbf{1}-x}^{\mathbf{1}}\int_{\mathbf{0}}^{\mathbf{1}-v} f_{\mathbf{X}}(u,v)dudv \end{aligned}$$

Esquema de contenidos

Ejemplo marginal - Caso continuo

$$f(x,y) = \frac{1}{400} \qquad 0<x<20 \,;\, 0<y<20$$

$$f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy = \int_0^{20} \frac{1}{400}\,dy = \frac{1}{20} \Longrightarrow f_X(x) = \frac{1}{20} \qquad 0<x<20$$

$$f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx = \int_0^{20} \frac{1}{400}\,dx = \frac{1}{20} \Longrightarrow f_Y(y) = \frac{1}{20} \qquad 0<y<20$$

Distribuciones marginales

Sea $\mathrm{X} = (X_1, \ldots, X_n) : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n, P_{\mathrm{X}})$, un vector aleatorio. Para cualquier subconjunto de subíndices $\{i_1, \ldots, i_k\}$ del conjunto $\{1, \ldots, n\}$, se define **la función de distribución marginal** del subvector aleatorio $(X_{i_1}, \ldots, X_{i_k})$ como sigue:

$$F_{X_{i_1}, \ldots, X_{i_k}}(x_{i_1}, \ldots, x_{i_k}) = \lim_{x_l \to \infty;\ l \neq i_1, \ldots, i_k} F_{\mathrm{X}}(x_1, \ldots, x_n). \tag{4}$$

Distribuciones marginales para el caso discreto

Sea $\mathrm{X} = (X_1, \ldots, X_n)$ un vector aleatorio discreto tal que $P(\mathrm{X} \in E_{\mathrm{X}}) = 1$, para un cierto subconjunto E_{X} numerable de $\mathbb{R}^n$, si y sólo si, para $i = 1, \ldots, n$, $\exists E_{X_i} \subset \mathbb{R}$, numerable, tal que, $P(X_i \in E_{X_i}) = 1$. A partir de la definición de f.m.p (diapositiva 20), **la función masa de probabilidad del subvector aleatorio discreto** $(X_{i_1}, \ldots, X_{i_k})$ se calcularía como sigue:

$$\begin{aligned} p_{X_{i_1},\ldots,X_{i_k}}(x_{i_1}, \ldots, x_{i_k}) &= \sum_{x_l \in E_{X_l};\ l \neq i_1,\ldots,i_k} p_{X_1,\ldots,X_n}(x_1 \ldots x_n) \\ &= \sum_{x_l \in E_{X_l};\ l \neq i_1,\ldots,i_k} P(X_1 = x_1, \ldots, X_n = x_n) \\ &= P(X_{i_1} = x_{i_1}, \ldots, X_{i_k} = x_{i_k}), \end{aligned}$$

para cualquier subconjunto de subíndices $\{i_1, \ldots, i_k\}$ del conjunto $\{1, \ldots, n\}$.

Distribuciones marginales para el caso continuo

Para vectores aleatorios continuos, **la función de densidad de probabilidad marginal** de cualquier subvector o variable aleatoria k- dimensional $(X_{i_1}, \ldots, X_{i_k})$, con $k < n$, $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$, construida a partir de k componentes del vector $\mathrm{X} = (X_1, \ldots, X_n)$, se calcularía mediante integración de la función de densidad de probabilidad conjunta f_{X} con respecto a los argumentos, cuyos subíndices no coinciden con los elementos del conjunto $\{i_1, \ldots, i_k\}$, es decir,

$$f_{X_{i_1}, \ldots, X_{i_k}}(x_{i_1}, \ldots, x_{i_k}) = \int_{\mathbb{R}^{n-k}} f_{\mathrm{X}}(x_1, \ldots, x_n) \prod_{l \in \{1, \ldots, n\};\; l \neq i_1, \ldots, i_k} dx_l.$$

Distribuciones condicionadas

- En diversas áreas de aplicación surgen problemas relacionados con el estudio del comportamiento de un subconjunto de variables aleatorias desconocidas (o no observables), dados los valores observados o conocidos de otro subconjunto de variables aleatorias relacionadas.
- Para describir y caracterizar dicho comportamiento se introduce el **concepto de distribución condicionada**.
- Más concretamente, para vectores aleatorios **discretos**, estudiaremos **la función masa de probabilidad condicionada**, y para vectores aleatorios **continuos** se estudiará **la función de densidad de probabilidad condicionada**.

Distribuciones condicionadas para el caso discreto

Notación

- Sea $\mathrm{X} = (X_1, \ldots, X_n)$ es un vector aleatorio discreto tal que $P(\mathrm{X} \in E_{\mathrm{X}}) = 1$, para un cierto subconjunto E_{X} numerable de $\mathbb{R}^n$, si y sólo si, para $i = 1, \ldots, n$, $\exists E_{X_i} \subset \mathbb{R}$, numerable tal que, $P(X_i \in E_{X_i}) = 1$.

- Por simplicidad, en la siguiente definición se notará, para cualesquiera dos subconjuntos de índices $\{j_1, \ldots, j_p\}$ e $\{i_1, \ldots, i_k\}$ del conjunto $\{1, \ldots, n\}$ tales que $i_m \neq j_l$, para $m = 1, \ldots, k$, y $l = 1, \ldots, p$, con $p + k = n$, las componentes del vector $\mathrm{x}_n = (x_1, \ldots, x_n)$, definidas a partir de las componentes de los vectores $\mathrm{x}_p = (x_{j_1}, \ldots, x_{j_p})$ y $\mathrm{x}_k = (x_{i_1}, \ldots, x_{i_k})$, como

$$\mathrm{x}_n = (x_1, \ldots, x_n) = (\mathrm{x}_p, \mathrm{x}_k),$$

donde se entiende que $(\mathrm{x}_p, \mathrm{x}_k)$ representa el vector resultante de reordenar las componentes de los vectores x_p y x_k, de forma adecuada, de acuerdo con los valores de los subíndices $j_1, \ldots, j_p$ e $i_1, \ldots, i_k$, dentro del conjunto $\{1, \ldots, n\}$.

Distribuciones condicionadas para el caso discreto

Definición

- Para cualesquiera dos subconjuntos de índices $\{j_1, \ldots, j_p\}$ e $\{i_1, \ldots, i_k\}$ del conjunto $\{1, \ldots, n\}$, definidos como antes, **la función masa de probabilidad de la variable aleatoria p-dimensional** $\mathrm{X}_p = (X_{j_1}, \ldots, X_{j_p})$ **condicionada** a los valores $\mathrm{y}_k = (y_{i_1}, \ldots, y_{i_k})$ de la variable aleatoria k- dimensional $(X_{i_1}, \ldots, X_{i_k})$, i.e., condicionada a los valores $(X_{i_1} = y_{i_1}, \ldots, X_{i_k} = y_{i_k})$, tal que $P(X_{i_1} = y_{i_1}, \ldots, X_{i_k} = y_{i_k}) > 0$, se define como sigue:

$$p_{X_{j_1},\ldots,X_{j_p}}(x_{j_1}, \ldots, x_{j_p} / X_{i_1} = y_{i_1}, \ldots, X_{i_k} = y_{i_k}) = \frac{p_{\mathrm{X}}(\mathrm{x}_p, \mathrm{y}_k)}{p_{X_{i_1},\ldots,X_{i_k}}(y_{i_1}, \ldots, y_{i_k})},$$

donde, siguiendo la notación anteriormente adoptada, p_{X} drepresenta la función masa de probabilidad conjunta del vector aleatorio discreto $\mathrm{X} = (X_1, \ldots, X_n)$ y $p_{X_{i_1},\ldots,X_{i_k}}$ denota la función masa de probabilidad marginal de las componentes aleatorias $X_{i_1}, \ldots, X_{i_k}$.

Distribuciones condicionadas para el caso continuo

Definición

- Dado un vector aleatorio $\mathrm{X} = (X_1, \ldots, X_n)$ continuo con función de densidad de probabilidad conjunta f_{X}, adoptando la notación anterior, para cualesquiera dos subconjuntos de índices $\{j_1, \ldots, j_p\}$ e $\{i_1, \ldots, i_k\}$ del conjunto $\{1, \ldots, n\}$, definidos como antes, **la función de densidad de probabilidad de la variable aleatoria** p-**dimensional** $\mathrm{X}_p = (X_{j_1}, \ldots, X_{j_p})$, **condicionada** a los valores $\mathrm{y}_k = (y_{i_1}, \ldots, y_{i_k})$ de la variable aleatoria k- dimensional $(X_{i_1}, \ldots, X_{i_k})$, se define como

$$f_{X_{j_1},\ldots,X_{j_p}}(x_{j_1}, \ldots, x_{j_p} / y_{i_1}, \ldots, y_{i_k}) = \frac{f_{\mathrm{X}}(\mathrm{x}_p, \mathrm{y}_k)}{f_{X_{i_1},\ldots,X_{i_k}}(y_{i_1}, \ldots, y_{i_k})}$$

$$\forall \mathrm{x}_p = (x_{j_1}, \ldots, x_{j_p}),$$

donde y_k debe ser tal que $f_{X_{i_1},\ldots,X_{i_k}}(y_{i_1}, \ldots, y_{i_k}) > 0$.

Esquema de contenidos

1. Discreto a discreto

Sea (X_1, X_2) discreto con f.m.p.

$X_1 \backslash X_2$	-1	0	1
-2	$1/6$	$1/12$	$1/6$
0	$1/6$	0	$1/6$
1	$1/12$	$2/12$	0

y queremos $Y \equiv (|X_1|, X_2^2)$

Solución:

$E_{(X_1,X_2)} = \{(-2,-1),(-2,0),(-2,1),(0,-1),(0,0),(0,1),(1,-1),(1,0),(1,1)\}$

$E_Y = \{(0,0),(0,1),(1,0),(1,1),(2,0),(2,1)\}$

$\lvert X \rvert \backslash X^2$	0	1
0	0	$1/3$
1	$2/12$	$1/12$
2	$1/12$	$1/3$

$P_Y(0,0) = P_{(X_1,X_2)}(g^{-1}(0,0)) = P_{(X_1,X_2)}(0,0)$

$P_Y(0,1) = P_{(X_1,X_2)}(g^{-1}(0,1)) = P_{(X_1,X_2)}(0,-1) + P_{(X_1,X_2)}(0,1) = 1/3$

⋮

Y así sucesivamente

2. Continuo a discreto

Sea (X_1, X_2) v.a. continuo con $f(x_1,x_2) = \lambda\mu e^{-\lambda x_1 - \mu x_2}$ si $x_1, x_2 > 0$ (λ, μ ctes. > 0)

y queremos la distribución de: $Y = \begin{cases} 0 & \text{si } x_1 > x_2 \\ 1 & \text{si } x_1 \leq x_2 \end{cases}$ Como Y es discreto, necesito P_Y (f.m.p):

$$P_Y(0) = P_{(X_1,X_2)}(X_1 < X_2) = \int_{\{(x_1,x_2)/x_1<x_2\}} \lambda\mu e^{-\lambda x_1 - \mu x_2} = \int_0^{+\infty}\int_{x_1}^{+\infty} \lambda\mu e^{-\lambda x_1 - \mu x_2}\,dx_2\,dx_1 = \int_0^{+\infty} \lambda\mu e^{-\lambda x_1}\underbrace{\left(\int_{x_1}^{+\infty} e^{-\mu x_2}\,dx_2\right)}_{-\frac{1}{\mu}\left(e^{-\mu x_2}\right]_{x_1}^{+\infty} = -\frac{1}{\mu}(0 - e^{-\mu x_1})} dx_1 = \int_0^{+\infty} \lambda \cancel{\mu} \frac{1}{\cancel{\mu}} \cdot e^{-\lambda x_1} \cdot e^{-\mu x_1}\,dx_1 = \int_0^{+\infty} \lambda e^{-(\lambda+\mu)x_1}\,dx_1 =$$

$$= -\frac{\lambda}{\lambda+\mu}\left(e^{-(\lambda+\mu)x_1}\right]_0^{+\infty} = \frac{\lambda}{\lambda+\mu} \qquad P_Y(1) = \frac{\mu}{\lambda+\mu}$$

3. Continuo a continuo

Sea (X_1, X_2) con $f(x_1, x_2) = \lambda\mu e^{-\lambda x_1 - \mu x_2}$ si $x_1, x_2 > 0$ (λ, μ ctes. > 0)

y queremos la distribución de: $Y = (X_1 + X_2, X_1)$

Planteamiento

Sea $\mathrm{X} = (X_1, \ldots, X_n)$ un vector aleatorio sobre $(\Omega, \mathcal{A}, P)$, tal que

$$P(\mathrm{X} \in E_{\mathrm{X}}) = P_{\mathrm{X}}(E_{\mathrm{X}}) = 1,$$

para un cierto conjunto $E_{\mathrm{X}} \subseteq \mathbb{R}^n$. Se considera una función $g : E_{\mathrm{X}} \longrightarrow \mathbb{R}^m$ medible.

Entonces, $\mathrm{Y} = g(\mathrm{X})$ es una variable aleatoria m-dimensional sobre $(\Omega, \mathcal{A}, P)$, cuya distribución de probabilidad y función de distribución de probabilidad vienen respectivamente dadas por:

$$P_{\mathrm{Y}}(B) = P_{\mathrm{X}}(g^{-1}(B)) = P(X \in g^{-1}(B)), \quad \forall B \in \mathcal{B}^m$$
$$F_{\mathrm{Y}}(\mathrm{y}) = P_{\mathrm{X}}(g^{-1}((-\infty, y_1] \times \cdots \times (-\infty, y_m])), \quad \forall \mathrm{y} = (y_1, \ldots, y_m) \in \mathbb{R}^m. \qquad (5)$$

Función masa de probabilidad de $\mathrm{Y} = g(\mathrm{X})$, X discreto

Supongamos que $\mathrm{X} = (X_1, \ldots, X_n)$ en (5) es de tipo **discreto**, es decir, E_X es un subconjumto numerable de $\mathbb{R}^n$. Entonces,

$$P_\mathrm{Y}(g(E_\mathrm{X})) = P(\mathrm{Y} \in g(E_\mathrm{X})) = P(\mathrm{X} \in E_\mathrm{X}) = P_\mathrm{X}(E_\mathrm{X}) = 1.$$

Del mismo modo, entonces $\mathrm{Y} = g(\mathrm{X})$ **es también una variable *m*-dimensional discreta** que toma sus valores en el conjunto numerable $g(E_\mathrm{X})$, con **función masa de probabilidad** definida por:

Para cualquier $\mathrm{y} = (y_1, \ldots, y_m) \in g(E_\mathrm{X})$,

$$\begin{aligned} p_\mathrm{Y}(\mathrm{y}) &= P_\mathrm{X}(g^{-1}(\mathrm{y})) = P_\mathrm{X}(\{\mathrm{x} \in E_\mathrm{X};\ g(\mathrm{x}) = \mathrm{y}\}) \\ &= \sum_{\mathrm{x}=(x_1,\ldots,x_n)\in g^{-1}(\mathrm{y})} p_\mathrm{X}(\mathrm{x}) = \sum_{\mathrm{x}=(x_1,\ldots,x_n)\in g^{-1}(\mathrm{y})} p_\mathrm{X}(x_1, \ldots, x_n) \\ &= \sum_{\mathrm{x}=(x_1,\ldots,x_n)\in g^{-1}(\mathrm{y})} P(X_1 = x_1, \ldots, X_n = x_n), \quad \forall \mathrm{y} \in g(E_\mathrm{X}). \end{aligned}$$

Función masa de probabilidad de $\mathrm{Y} = g(\mathrm{X})$, X continuo e Y discreto

Supongamos ahora que $\mathrm{X} = (X_1, \ldots, X_n)$ en (5) es una variable aleatoria n-dimensional **continua**, con función de densidad de probabilidad f_{X}, y se considera una función medible $g : E_{\mathrm{X}} \longrightarrow \mathbb{R}^m$ tal que $\mathrm{Y} = g(\mathrm{X})$ es **discreta**.

Entonces la **función masa de probabilidad** de Y viene dada, $\forall \mathrm{y} = (y_1, \ldots, y_n) \in g(E_{\mathrm{X}})$, por la siguiente expresión:

$$p_{\mathrm{Y}}(\mathrm{y}) = p_{\mathrm{Y}}(y_1, \ldots, y_n) = \int_{\{\mathrm{x} \in E_{\mathbf{X}};\ g(\mathrm{x}) = \mathrm{y}\}} f_{\mathrm{X}}(\mathrm{v}) d\mathrm{v}.$$

Función de densidad de $Y = g(X)$, X continuo e Y continuo

Sea $X = (X_1, \ldots, X_n)$ en (5) es una variable aleatoria n-dimensional **continua**, con función de densidad de probabilidad f_X. Supongamos que $g : E_X \longrightarrow \mathbb{R}^n$ es una función medible satisfaciendo las siguientes condiciones:

(i) g es **derivable** en todos los argumentos

(ii) g es **inyectiva** ($g^{-1} : g(E_X) \longrightarrow E_X$, es una aplicación). Es decir, $\forall y \in g(E_X)$, existe un único $x = (x_1, \ldots, x_n) \in E_X$, tal que $g(x) = g(x_1, \ldots, x_n) = y$. Dado que x está unívocamente determinado por y, utilizaremos la notación

$$g^{-1}(y) = x(y) = (x_1(y), \ldots, x_n(y)).$$

(iii) El **jacobiano** de g^{-1} es **no nulo**, i.e., para cualquier $y \in g(E_X)$,

$$J(y) = \det\left[\begin{pmatrix} \frac{\partial x_1(y_1,\ldots,y_n)}{\partial y_1} & \ldots & \frac{\partial x_1(y_1,\ldots,y_n)}{\partial y_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial x_n(y_1,\ldots,y_n)}{\partial y_1} & \ldots & \frac{\partial x_n(y_1,\ldots,y_n)}{\partial y_n} \end{pmatrix}\right] = \begin{vmatrix} \frac{\partial x_1(y_1,\ldots,y_n)}{\partial y_1} & \ldots & \frac{\partial x_1(y_1,\ldots,y_n)}{\partial y_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial x_n(y_1,\ldots,y_n)}{\partial y_1} & \ldots & \frac{\partial x_n(y_1,\ldots,y_n)}{\partial y_n} \end{vmatrix} \neq 0$$

Bajo las condiciones (i)-(iii), $Y = g(X)$ es **una variable aleatoria n-dimensional continua**, cuya función de densidad f_Y viene dada por:

Valor absoluto del determinante

$$f_Y(y) = f_X(g^{-1}(y))|J(y)|, \quad \forall y \in g(E_X)$$

Observación 1

Una **propiedad interesante del jacobiano** es que cuando éste es **diferente de cero, en el entorno de un punto dado**, entonces el teorema de la función inversa garantiza que la función asociada **admite una función inversa alrededor de dicho punto**.

Siendo, por tanto, una **condición suficiente pero no necesaria**, es decir, el **jacobiano** de una función n-dimensional sobre un conjunto de $\mathbb{R}^n$ **se puede anular en un punto**, sin que ello implique la no existencia de la función inversa en un entorno de dicho punto.

Observación 2

Si g no cumple la condición (ii), pero para cada $\mathrm{y} \in g(E_{\mathbf{X}})$, existe un conjunto numerable de antiimágenes de y,

$$\{\mathrm{x}_k(\mathrm{y})\}_{k\in\mathbb{N}} = \{(x_{1k}(\mathrm{y}), \ldots, x_{nk}(\mathrm{y}))\}_{k\in\mathbb{N}},$$

tales que, para todo $k \in \mathbb{N}$, el jacobiano J_k de x_k satisface

$$J_k(\mathrm{y}) \neq 0, \quad \forall \mathrm{y} \in g(E_{\mathbf{X}}),$$

entonces $\mathbf{Y} = g(\mathbf{X})$ es una variable aleatoria n-dimensional continua, cuya función de densidad $f_{\mathbf{Y}}$ viene dada por:

$$f_{\mathbf{Y}}(\mathrm{y}) = \sum_{k\in\mathbb{N}} f_{\mathbf{X}}(\mathrm{x}_k(\mathrm{y}))|J_k(\mathrm{y})|, \quad \forall \mathrm{y} \in g(E_{\mathbf{X}}).$$

Esquema de contenidos

Distribuciones el máximo y del mínimo

Notación
Sea $\mathrm{X} = (X_1, \ldots, X_n) : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n, P_{\mathrm{X}})$ un vector aleatorio n-dimensional.
Se definen las variables aleatorias unidimensionales

$$\begin{aligned} \max(X_1, \ldots, X_n) &: (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}) \\ \min(X_1, \ldots, X_n) &: (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}) \end{aligned}$$

cuyas funciones de distribución se definen como sigue: para cualquier $y \in \mathbb{R}$,

$$\begin{aligned} F_{\max(X_1,\ldots,X_n)}(y) &= P(\max(X_1, \ldots, X_n) \leq y) \\ &= P(X_1 \leq y, \ldots, X_n \leq y) = F_{\mathrm{X}}(y, \ldots, y) \\ F_{\min(X_1,\ldots,X_n)}(y) &= P(\min(X_1, \ldots, X_n) \leq y) \\ &= 1 - P(\min(X_1, \ldots, X_n) > y) = 1 - P(X_1 > y, \ldots, X_n > y) \\ &= 1 - P_{\mathrm{X}}((y, \infty) \times \cdots \times (y, \infty)). \end{aligned}$$

Distribuciones el máximo y del mínimo

La **función de distribución conjunta** de la variable aleatoria bidimensional

$$(Z, T) = (\min(X_1, \ldots, X_n), \max(X_1, \ldots, X_n)),$$

viene dada por

$$F_{(Z,T)}(x, y) = \begin{cases} F_X(y, \ldots, y), & y \leq x \\ F_X(y, \ldots, y) - P(x < X_1 \leq y, \ldots, x < X_n \leq y), & y > x \end{cases}$$

Indicación

- Para deducir la expresión $F_{(Z,T)}(x, y)$ cuando $y > x$ se ha tenido en cuenta la igualdad de sucesos siguiente: $\{Z \leq x\} \cap \{T \leq y\} = \{Z > x\}^c \cap \{T \leq y\}$.

Ejercicio propuesto (voluntario)

- Obtener las funciones de distribución del máximo y mínimo, así como su conjunta en el supuesto de que X_i, $\forall i = 1, \ldots, n$ sean v.a. independientes e identicamente distribuidas con función de distribución $F_{X_i}(x) = F(x)$, $\forall i = 1, \ldots, n$, $x \in \mathbb{R}$.
- En el caso de que además todas v.a. sean continuas, obtener también las funciones de densidad asociadas.

(1) $F_{(Z,T)}(x,y) = P[Z\leq x, T\leq y] = P[\min\{X_1,\dots,X_n\}\leq x, \max\{X_1,\dots,X_n\}\leq y] =$

$= P[\min\{X_1,\dots,X_n\}\leq x]\cdot P[\max\{X_1,\dots,X_n\}\leq y] = F_X(y,\dots,y)$

Independencia de sucesos

(2) Lo hacemos nosotros

$F_{(Z,T)}(x,y) = P[Z\leq x, T\leq y] = P[\min(X_1,\dots,X_n)\leq x, \max(X_1,\dots,X_n)\leq y] = P[\min(X_1,\dots,X_n)\leq x]\cdot P[\max(X_1,\dots,X_n)\leq y] =$

$=(1-P[\min(X_1,\dots,X_n)>x])\cdot P[\max(X_1,\dots,X_n)\leq y] = P[\max(X_1,\dots,X_n)\leq y] - P[\min(X_1,\dots,X_n)>x]\cdot P[\max(X_1,\dots,X_n)\leq y]$

$= \underbrace{P(X_1\leq y,\dots,X_n\leq y)}_{F_X(y,\dots,y)} - P[\min(X_1,\dots,X_n)>x, \max(X_1,\dots,X_n)\leq y] = F_X(y,\dots,y) - P_X(x<X_1\leq y,\dots,x<X_n\leq y)$

Ejemplo

Sea $X=(X_1,X_2)$ vector aleatorio discreto con f.m.p. $P[X_1=x, X_2=x_2] = p^2(1-p)^{x_1+x_2}$; $x_1,x_2=0,1,2,\dots$
$0<p<1$

Máximo

¿$P[Z=x]$?

$P[Z=x] \underset{\text{discreto}}{=} P[Z\leq x] - P[Z\leq x-1] = F_Z(x) - F_Z(x-1)$

$F_Z(x) = P[Z\leq x] = F_X(x,x) = \sum_{x_1,x_2=0}^{x} P[X_1=x, X_2=x] = \sum_{x_1,x_2=0}^{x} p^2(1-p)^{x_1+x_2} = p^2 \underbrace{\sum_{x_1=0}^{x}(1-p)^{x_1}\cdot\sum_{x_2=0}^{x}(1-p)^{x_2}}_{\text{Son la misma cantidad}} = p^2\Big[\underbrace{\sum_{x_1=0}^{x}(1-p)^{x_1}}_{\text{Serie geométrica}}\Big]^2 = p^2\left(\frac{1-(1-p)^{x+1}}{1-(1-p)}\right)^2 =$

$= (1-(1-p)^{x+1})^2$

$P[Z=x] = F_Z(x) - F_Z(x-1) = [1-(1-p)^{x+1}]^2 - [1-(1-p)^x]^2 \quad \forall x=0,1,\dots$

Mínimo

¿$P[T=x] = F_T(x) - F_T(x-1)$?

$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}$

$F_T(x) = 1 - P[X_1>x, X_2>x] = 1 - \sum_{x_1,x_2=x+1}^{\infty} P[X_1=x_1, X_2=x_2] =$

Esquema de contenidos

Definición

Sea $\mathrm{X} = (X_1, \ldots, X_n) : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n, P_{\mathrm{X}})$ un vector aleatorio n-dimensional. La **Esperanza Matemática**, $E[\mathrm{X}]$, de X, si existe, se define como un vector determinístico **cuyas componentes son las medias o esperanzas matemáticas de sus componentes aleatorias**, es decir

$$E[\mathrm{X}] = (E[X_1], \ldots, E[X_n]) = (\mu_1, \ldots, \mu_n) \in \mathbb{R}^n.$$

Consecuencia

La esperanza matemática de un vector aleatorio $\mathrm{X} = (X_1, \ldots, X_n)$ existe $\Leftrightarrow$ existen las esperanzas matemáticas de sus componentes aleatorias.

Equivalentemente, un vector aleatorio tiene media finita $\Leftrightarrow$ sus distribuciones marginales tienen media o momento de orden uno no centrado finito, i.e.,

$$\exists E[\mathrm{X}] \Leftrightarrow \exists E[X_i], \ \ i = 1, \ldots, n.$$

Propiedades de la esperanza matemática

- **Linealidad.** Para cualesquiera $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in \mathbb{R}^n$,

$$\exists E[X_i] \Rightarrow \exists E[a_i X_i + b_i], \quad i = 1, \ldots, n$$

$$\Downarrow$$

$$\exists E\left[\sum_{i=1}^{n} a_i X_i + b_i\right] = \sum_{i=1}^{n} a_i E[X_i] + b_i.$$

- **Monotonía.** Sean X_1 y X_2 variables aleatorias unidimensionales, tales que, $\exists E[X_1]$, $\exists E[X_2]$, entonces

$$X_1 \leq X_2 \;\Rightarrow\; E[X_1] \leq E[X_2].$$

Transformación $g : \mathbb{R}^n \longrightarrow \mathbb{R}$

Sea $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ una función medible, entonces, $Y = g(\mathrm{X})$ es una variable aleatoria. Si existe su media $E[g(\mathrm{X})]$, entonces se tienen las siguientes afirmaciones:

- X **es de tipo discreto,** $P(\mathrm{X} \in E_{\mathrm{X}}) = 1$, siendo $E_{\mathrm{X}} \subset \mathbb{R}$ numerable. Entonces:

$$\exists E[g(\mathrm{X})] \;\Leftrightarrow\; \sum_{(x_1,\dots,x_n)\in E_{\mathbf{X}}} |g(x_1,\dots,x_n)| p_{\mathrm{X}}(x_1,\dots,x_n) < \infty.$$

$$\exists E[g(\mathrm{X})] \;\Rightarrow\; E[g(\mathrm{X})] = \sum_{(x_1,\dots,x_n)\in E_{\mathbf{X}}} g(x_1,\dots,x_n) p_{\mathrm{X}}(x_1,\dots,x_n)$$

- X **es de tipo continuo con función de densidad de probabilidad** f_{X}. Entonces:

$$\exists E[g(\mathrm{X})] \;\Leftrightarrow\; \int_{\mathbb{R}^n} |g(x_1,\dots,x_n)| f_{\mathrm{X}}(x_1,\dots,x_n) dx_1 \dots dx_n < \infty.$$

$$\exists E[g(\mathrm{X})] \;\Rightarrow\; E[g(\mathrm{X})] = \int_{\mathbb{R}^n} g(x_1,\dots,x_n) f_{\mathrm{X}}(x_1,\dots,x_n) dx_1 \dots dx_n$$

Esquema de contenidos

Definición

Sea $\mathrm{X} = (X_1, \ldots, X_n) : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n, P_{\mathrm{X}})$ un vector aleatorio n-dimensional.

- **Momentos no centrados de orden k de** X. Se notarán $m_{k_1,\ldots,k_n}$, siendo $k = \sum_{i=1}^n k_i$. Corresponden al cálculo de la esperanza matemática de la función

$$g(\mathrm{X}) = g(X_1, \ldots, X_n) = X_1^{k_1} \cdots X_n^{k_n}, \quad k = \sum_{i=1}^n k_i.$$

  Por tanto, su existencia y cálculo se obtiene como en el apartado anterior.

- **Momentos centrados (respecto a la media) de orden k de** X. Se notarán $\mu_{k_1,\ldots,k_n}$, siendo $k = \sum_{i=1}^n k_i$. Su existencia y cálculo se derivan como en el apartado anterior, considerando la esperanza matemática de la función g dada por

$$g(\mathrm{X}) = g(X_1, \ldots, X_n) = (X_1 - E[X_1])^{k_1} \cdots (X_n - E[X_n])^{k_n}, \quad k = \sum_{i=1}^n k_i.$$

Caso bidimensional (X, Y). Momentos de segundo orden

Si existen $m_{1,0}$, $m_{0,1}$, y $m_{1,1}$ se define la covarianza de X e Y como

$$\mu_{1,1} = \text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])].$$

Equivalentemente,

$$\mu_{1,1} = \text{Cov}(X, Y) = E[XY] - E[X]E[Y].$$

Varianza de una combinación lineal n-dimensional

Para cualquier $(a_1, \ldots, a_n) \in \mathbb{R}^n$,

$$\text{Si } \exists E[X_i^2],\ \ i = 1, \ldots, n, \Rightarrow \exists \text{Var}\left[\sum_{i=1}^{n} a_i X_i\right]$$

$$\text{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 \text{Var}(X_i) + \sum_{i \neq j}^{n} a_i a_j \text{Cov}(X_i, X_j).$$

Desigualdad de Cauchy-Schwarz

Esta desigualdad se obtiene como consecuencia de la estructura de Hilbert definida a partir del conjunto $\mathcal{L}^2(\Omega, \mathcal{A}, P)$, constituido por las variables aleatorias unidimensionales sobre $(\Omega, \mathcal{A}, P)$, centradas, con momento de orden dos finito.
Específicamente, en dicho conjunto, se define el **producto escalar**

$$\langle X, Y \rangle_{\mathcal{L}^2(\Omega, \mathcal{A}, P)} = E[XY], \quad \forall X, Y \in \mathcal{L}^2(\Omega, \mathcal{A}, P).$$

La **norma asociada** viene entonces dada por

$$\|X\|^2_{\mathcal{L}^2(\Omega, \mathcal{A}, P)} = E[X^2], \quad \forall X \in \mathcal{L}^2(\Omega, \mathcal{A}, P).$$

- En este espacio, la **desigualdad de Cauchy-Schwarz** indica que:

$$[E[XY]]^2 \leq E[X^2]E[Y^2], \quad \forall X, Y \in \mathcal{L}^2(\Omega, \mathcal{A}, P).$$

- **Se da la igualdad** si y sólo existen a y b no nulos tales que $P(aX + bY = 0) = 1$.
- Si X e Y son **no degeneradas** se tiene también que $(\mathrm{Cov}(X, Y))^2 \leq \mathrm{Var}(X)\mathrm{Var}(Y)$. **Se da la igualdad** si y sólo existen a y b no nulos, y $c \in \mathbb{R}$, tales que

$$P(aX + bY = c) = 1.$$

Esquema de contenidos

Definición

Se define la **función generatriz de momentos n-dimensional** como

$$M_{\mathrm{X}}(t_1, \ldots, t_n) : (-a_1, b_1) \times \cdots \times (-a_n, b_n) \longrightarrow \mathbb{R}.$$

$$(t_1, \ldots, t_n) \longrightarrow E\left[\exp\left(\sum_{i=1}^{n} t_i X_i\right)\right],$$

donde $a_i, b_j \in \mathbb{R}^+$, $\forall i, j = 1, \ldots, n$.

Teorema de unicidad

Si existe la función generatriz de momentos de un vector aleatorio, **determina de forma unívoca** a su distribución de probabilidad.

Momentos no centrados

Si existe la función generatriz de momentos de un vector aleatorio n-dimensional, **los momentos no centrados** de dicho vector se calculan mediante **evaluación en el vector cero de la derivada cruzada** de M_X de orden correspondiente. Es decir,

$$m_{k_1,\ldots,k_n} = \left[\frac{\partial^{k_1+\cdots+k_n} M_X(t_1,\ldots,t_n)}{\partial t_1^{k_1}\ldots\partial t_n^{k_n}}\right]_{t_1=\cdots=t_n=0}.$$

Funciones generatrices de momentos marginales

Sea $\mathrm{X} = (X_1, \ldots, X_n) : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n, P_{\mathrm{X}})$ un vector aleatorio n-dimensional. Supongamos que existe la función generatriz de momentos M_{X} de X.

Entonces, la **función generatriz de momentos de cualquier subvector**

$$(X_{i_1}, \ldots, X_{i_k}), \quad k < n,$$

se calcula como sigue:

$$M_{X_{i_1}, \ldots, X_{i_k}}(t_{i_1}, \ldots, t_{i_k}) = M_{\mathrm{X}}((0, \mathrm{t}_k)),$$

$$\mathrm{t}_k = (t_{i_1}, \ldots, t_{i_k}) \in (-a_{i_1}, b_{i_1}) \times \cdots \times (-a_{i_k}, b_{i_k}),$$

donde el vector $(0, \mathrm{t}_k)$ se define recolocando las componentes del vector $\mathrm{t}_k = (t_{i_1}, \ldots, t_{i_k})$ en los lugares correspondientes, determinados por los valores $i_1, \ldots, i_k \in \{1, \ldots, n\}$, y asignando el valor cero a las restantes componentes, ubicadas en los lugares $\{1, \ldots, n\} - \{i_1, \ldots, i_k\}$.