Tema 1: Variables aleatorias multidimensionales

Cálculo de Probabilidades II

Grado en Estadistica

Universidad de Granada

Curso 2017/18

Índice

Planteamiento

Variables aleatorias: miden numéricamente una característica de un experimento aleatorio.

P. ej. elegir a una persona al azar y medir su peso.

Planteamiento

En muchas ocasiones, puede ser interesante medir varias características de ese experimento aleatorio.

P. ej. elegir a una persona al azar y medir su peso, su altura, su presión sanguínea, su color de ojos...

Concepto

- La formalización de estas medidas da lugar a las **variables aleatorias multidimensionales** o **vectores aleatorios**.
- Su definición es análoga a las vv. aa. unidimensionales, así como su tratamiento.

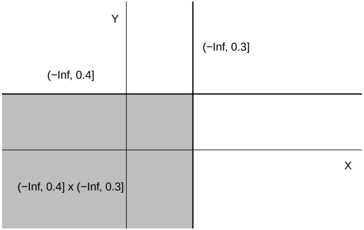
Experimento	x_1	x_2	x_3	x_4	...
Peso	78.5	76.6	75.7	77.3	...
Altura	176.3	184.9	181.4	171.2	...
Ojos azules	0	0	1	0	...

Espacios de Borel multidimensionales

Se define la σ-álgebra de Borel sobre $\mathbb{R}^n$, $\mathcal{B}^n$, como la mínima clase que contiene todos los intervalos de $\mathbb{R}^n$
Sea $\mathcal{Y}$ una clase que contiene todos los intervalos posibles en $\mathbb{R}$:

$$\mathcal{B}^n = \sigma(\mathcal{Y}^n)$$

$\mathcal{B}^n$ es la σ-álgebra generada por intervalos del todo tipo; en particular $(-\infty, x_1] \times (-\infty, x_2] \times \ldots \times (-\infty, x_n] = (-\infty, x], x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$



Definición

Se define una variable aleatoria multidimensional sobre un espacio probabilístico $(\Omega, \mathcal{A}, P)$ como la aplicación

$$X : \Omega \to \mathbb{R}^n$$

Esta aplicación es medible para todo conjunto B contenido en $\mathcal{B}^n$, esto es

$$X^{-1}(B) \in \mathcal{A}, \forall B \in \mathcal{B}^n$$

Caracterización

$X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^n, \mathcal{B}^n)$ es un vector aleatorio si y solo si:

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

Con $x = (x_1, ..., x_n) \in \mathbb{R}^n$

Caracterización

$X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, P) \rightarrow (\mathbb{R}^n, \mathcal{B}^n)$ es un vector aleatorio si y solo si:

$$\forall i = 1, ..., n \quad X_i : (\Omega, \mathcal{A}, P) \rightarrow (\mathbb{R}, \mathcal{B}) \text{ es variable aleatoria}$$

Ejemplo

Considérese el experimento de lanzar dos monedas al aire. Sobre dicho experimento, se cuentan por un lado el número de caras obtenidas y por otro lado la diferencia en valor absoluto entre el número de caras y el de cruces.
Los elementos del espacio de probabilidad, $(\Omega, \mathcal{A}, P)$, asociado al experimento aleatorio se definen de la siguiente forma:

$$\Omega = \{cc, cx, xc, xx\}$$

$$\mathcal{A} = \mathcal{P}(\Omega)$$

$$P(\omega) = 1/4, \omega \in \Omega$$

Ejemplo

Sea:
X_1: número de caras obtenidas.
X_2: diferencia en valor absoluto entre el número de caras y el de cruces.
El vector (X_1, X_2) está definido en el espacio medible $(\mathbb{R}^2, \mathcal{B}^2)$

Ejemplo

$$X^{-1}((-\infty,0],(-\infty,0]) = \{\omega \in \Omega / X_1(\omega) \le 0, X_2(\omega) \le 0\} = \emptyset \in \mathcal{A}$$

$$X^{-1}((-\infty,0],(-\infty,2]) = \{\omega \in \Omega / X_1(\omega) \le 0, X_2(\omega) \le 2\} = \{xx\} \in \mathcal{A}$$

$$X^{-1}((-\infty,1],(-\infty,0]) = \{\omega \in \Omega / X_1(\omega) \le 1, X_2(\omega) \le 0\} = \{cx, xc\} \in \mathcal{A}$$

$$X^{-1}((-\infty,2],(-\infty,0]) = \{\omega \in \Omega / X_1(\omega) \le 2, X_2(\omega) \le 0\} = \{cx, xc\} \in \mathcal{A}$$

$$X^{-1}((-\infty,1],(-\infty,2]) = \{\omega \in \Omega / X_1(\omega) \le 1, X_2(\omega) \le 2\} = \{cx, xc, xx\} \in \mathcal{A}$$

$$X^{-1}((-\infty,2],(-\infty,2]) = \{\omega \in \Omega / X_1(\omega) \le 2, X_2(\omega) \le 2\} = \\ = \{cc, cx, xc, xx\} = \Omega \in \mathcal{A}$$

etc. etc.

Además, si $x_1, x_2 < 0$, la imagen inversa será $\emptyset \in \mathcal{A}$ y si $x_1, x_2 > 2$ será $\Omega \in \mathcal{A}$. Por tanto, el vector (X_1, X_2) es un vector aleatorio.

Distribución de probabilidad

Sea $X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^n, \mathcal{B}^n)$, se denomina distribución de probabilidad de X a la función de conjunto

$$P_X = P \circ X^{-1} : \begin{array}{ccl} \mathcal{B}^n & \to & [0,1] \\ B & \mapsto & P_X(B) = P\{X^{-1}(B)\} = P\{X \in B\} \end{array}$$

X transforma el espacio probabilístico original $(\Omega, \mathcal{A}, P)$ en el nuevo espacio probabilístico

$$X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^n, \mathcal{B}^n, P_X)$$

Función de distribución

$X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^n, \mathcal{B}^n, P_X)$, se define la función de distribución de X tal que:

$$\begin{array}{rccl} F_X : & \mathbb{R}^n & \to & [0,1] \\ & x & \mapsto & F_X(x) = P_X((-\infty, x]) = P(X^{-1}((-\infty, x])) \end{array}$$

Con $x = (x_1, ..., x_n) \in \mathbb{R}^n$

Propiedades de la función de distribución

1. Monótona no decreciente en cada argumento ($\equiv$ para cada x_i)
2. Continua a la derecha en cada argumento ($\equiv$ para cada x_i)
3. $$\lim_{x_1,\ldots,x_n \to +\infty} F_X(x_1,...,x_n) = F_X(+\infty,...,+\infty) = 1$$
   $$\lim_{x_i \to -\infty} F_X(x_1,...,x_i,...,x_n) = F_X(x_1,...,x_{i-1},-\infty,x_{i+1},...,x_n) = 0$$

(Sus demostraciones son análogas a las del caso unidimensional, pero para vectores aleatorios se requiere otra propiedad más)

Propiedades de la función de distribución

4. Sean $(x_1, ..., x_n) \in \mathbb{R}^n$ y $(\epsilon_1, ..., \epsilon_n) \in \mathbb{R}^{n+}$

$$F_X(x_1 + \epsilon_1, ..., x_n + \epsilon_n) -$$

$$-\sum_{i=1}^{n} F_X(x_1 + \epsilon_1, ..., x_{i-1} + \epsilon_{i-1}, x_i, x_{i+1} + \epsilon_{i+1}, ..., x_n + \epsilon_n) +$$

$$+\sum_{i,j=1,\ i<j}^{n} F_X(x_1 + \epsilon_1, ..., x_{i-1} + \epsilon_{i-1}, x_i, x_{i+1} + \epsilon_{i+1}, ...,$$

$$x_{j-1} + \epsilon_{j-1}, x_j, x_{j+1} + \epsilon_{j+1}, ..., x_n + \epsilon_n) +$$

$$+ ... + (-1)^n F_X(x_1, ..., x_n) \geq 0$$

Ejemplo (continuación)

La función de distribución del vector (X_1, X_2) para cada parte se calcularía como sigue:

$$F_X(x_1, x_2) = P_X((-\infty, x_1], (-\infty, x_2]) = P(X^{-1}((-\infty, x_1], (-\infty, x_2]))$$

Por ejemplo, para la función de distribución en $X = (0, 0)$:

$$F_X(0,0) = P_X((-\infty, 0], (-\infty, 0]) = P(X^{-1}((-\infty, 0], (-\infty, 0])) = \\ = P(\emptyset) = 0$$

Otro ejemplo, este del punto $X = (1, 2)$:

$$F_X(1,2) = P_X((-\infty, 1], (-\infty, 2]) = P(X^{-1}((-\infty, 1], (-\infty, 2])) = \\ = P(\{cx, xc, xx\}) = 3/4$$

Ejemplo (continuación)

La función de distribución se podría expresar de la siguiente manera:

$$F_X(x_1, x_2) = \begin{cases} 0 & x_1 < 0 \text{ ó } x_2 < 0 \text{ ó } \{0 \leq x_1 < 1, 0 \leq x_2 < 2\} \\ 1/4 & 0 \leq x_1 < 1, x_2 \geq 2 \\ 1/2 & x_1 \geq 1, 0 \leq x_2 < 2 \\ 3/4 & 1 \leq x_1 < 2, x_2 \geq 2 \\ 1 & x_1 \geq 2, x_2 \geq 2 \end{cases}$$

Propiedades de la función de distribución (2)

- Para $(x_1, ..., x_n) \in \mathbb{R}^n$

$$P(X_1 \leq x_1, ..., X_{i-1} \leq x_{i-1}, X_i < x_i, X_{i+1} \leq x_{i+1}, ..., X_n \leq x_n) =$$

$$= F_X(x_1, ..., x_{i-1}, x_i^-, x_{i+1}, ..., x_n)$$

- Para $(x_1, ..., x_n) \in \mathbb{R}^n$

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

$$= F_X(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n) - F_X(x_1, ..., x_{i-1}, x_i^-, x_{i+1}, ..., x_n)$$

- F_X es continua en el argumento i-ésimo en el punto $x_i \in \mathbb{R}$ si y sólo si

$$P(X_1 \leq x_1, ..., X_{i-1} \leq x_{i-1}, X_i = x_i, X_{i+1} \leq x_{i+1}, ..., X_n \leq x_n) = 0$$

Cálculo de probabilidades en intervalos bidimensionales

- $P(X_1 \leq b, c < X_2 \leq d) = F(b, d) - F(b, c)$
- $P(X_1 \leq b, c < X_2 < d) = F(b, d^-) - F(b, c)$
- $P(X_1 \leq b, c \leq X_2 \leq d) = F(b, d) - F(b, c^-)$
- $P(X_1 \leq b, c \leq X_2 < d) = F(b, d^-) - F(b, c^-)$
- $P(X_1 < b, c < X_2 \leq d) = F(b^-, d) - F(b^-, c)$
- $P(X_1 < b, c < X_2 < d) = F(b^-, d^-) - F(b^-, c)$
- $P(X_1 < b, c \leq X_2 \leq d) = F(b^-, d) - F(b^-, c^-)$
- $P(X_1 < b, c \leq X_2 < d) = F(b^-, d^-) - F(b^-, c^-)$

Ejemplo (continuación)

- $P(X_1 \leq 1, 1 < X_2 \leq 2) = F(1,2) - F(1,1) = 3/4 - 1/2 = 1/4$
- $P(X_1 \leq 1, 1 < X_2 < 2) = F(1,2^-) - F(1,1) = 1/2 - 1/2 = 0$
- $P(X_1 \leq 1, 1 \leq X_2 \leq 2) = F(1,2) - F(1,1^-) = 3/4 - 1/2 = 1/4$
- $P(X_1 \leq 1, 1 \leq X_2 < 1) = F(1,2^-) - F(1,1^-) = 1/2 - 1/2 = 0$
- $P(X_1 < 1, 1 < X_2 \leq 2) = F(1^-,2) - F(1^-,1) = 1/4 - 0 = 1/4$
- $P(X_1 < 1, 1 < X_2 < 2) = F(1^-,2^-) - F(1^-,1) = 0 - 0 = 0$
- $P(X_1 < 1, 1 \leq X_2 \leq 2) = F(1^-,2) - F(1^-,1^-) = 1/4 - 0 = 1/4$
- $P(X_1 < 1, 1 \leq X_2 < 2) = F(1^-,2^-) - F(1^-,1^-) = 0 - 0 = 0$

Definición

Un vector aleatorio $X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^n, \mathcal{B}^n)$ se denomina de tipo discreto si existe un conjunto numerable, $E_X \subseteq \mathbb{R}^n$, tal que $P(X \in E_X) = 1$.
Sea $x = (x_1, ..., x_n) \in \mathbb{R}^n$, la distribución de probabilidad P_X se especifica por:

$$P_X(x) = P(\omega \in \Omega / X_1(\omega) = x_1, ..., X_n(\omega) = x_n)$$

Para un subconjunto $A \subseteq \mathbb{R}^n$ se tiene

$$P_X(A) = \sum_{x \in A} P_X(x) = \sum_{x \in A} P(X_1 = x_1, ..., X_n = x_n)$$

Definición

La distribución de probabilidad se maneja a través de la función masa de probabilidad, p, formada por los puntos de salto de la función de distribución.

$$
\begin{array}{rccc}
p: & E_X & \rightarrow & [0,1] \\
 & x_n & \mapsto & P(X = x_n) = p_n
\end{array}
$$

- $p_n \geq 0$
- $\sum_{x \in E_X} p_n = 1$

Definición

El cálculo de la función de distribución se realiza teniendo en cuenta el cálculo de probabilidades en subconjuntos:

- Sea un conjunto $B \in \mathcal{B}^n$

$$P_X(B) = P(X \in B) = \sum_{x \in B \cap E_X} P(X = x)$$

- Sea un vector $x \in \mathbb{R}^n$

$$F_X(x) = \sum_{x_j \in E_X, x_j \leq x} P(X = x_j)$$

El Teorema de Correspondencia aplicable a las vv. aa. discretas unidimensionales también es aplicable en esta situación.

Caracterización

- $X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^n, \mathcal{B}^n, P_X)$ es discreto si y sólo si $\forall i = 1, ..., n$ las componentes $X_i : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B})$ son discretas.
- Por tanto, un vector aleatorio n-dimensional discreto es un conjunto de n vv. aa. unidimensionales discretas.

Ejemplo

Un jurado popular de 10 personas está formado por los siguientes perfiles profesionales:

- 3 funcionarios
- 3 personas con contratos de formación
- 2 personas en situación de desempleo
- 2 estudiantes

Se eligen 3 personas al azar y se cuenta cuántas personas de esas 3 son funcionarias y cuántas tienen contratos de formación.

Ejemplo

El experimento aleatorio es el de seleccionar tres personas al azar del jurado popular. El espacio probabilístico asociado, $(\Omega, \mathcal{A}, P)$ es de la siguiente forma:

$$\Omega = \{(F_1, F_2, F_3), (F_1, F_2, CF_1), (F_1, D_1, CF_1), \ldots\}$$

$$\text{(habría } \binom{10}{3} = 120 \text{ sucesos elementales)}$$

$$\mathcal{A} = \mathcal{P}(\Omega)$$

$P \equiv$ uniforme (todos los sucesos elementales tienen la misma probabilidad)

Ejemplo

Sobre este espacio probabilístico se definen las variables aleatorias:
X_1: número de funcionarios en la muestra seleccionada. X_1 toma valores dentro de un conjunto numerable: $E_1 = (0, 1, 2, 3) \subset \mathbb{R}$
X_2: número de personas con contratos de formación en la muestra seleccionada. X_2 toma valores dentro de un conjunto numerable: $E_1 = (0, 1, 2, 3) \subset \mathbb{R}$

Ejemplo

Como X_1 y X_2 son vv. aa. discretas, el vector (X_1, X_2) sería un vector aleatorio discreto bidimensional. Dicho vector toma valores dentro del conjunto numerable $\{(0,0),(0,1),(0,2),(0,3),(1,0),(1,1),(1,2),(1,3),(2,0),(2,1),$ $(2,2),(2,3),(3,0),(3,1),(3,2),(3,3)\} \subset \mathbb{R}^2$. Este conjunto es el resultado de $E_1 \times E_2$

Ejemplo

La probabilidad de cada valor se calcula por la fórmula vista anteriormente:

$$P_X(x) = P(\omega \in \Omega / X_1(\omega) = x_1, ..., X_n(\omega) = x_n)$$

Por ejemplo, la probabilidad del valor $(0, 0)$ es:

$$P_X((0,0)) = P(\omega \in \Omega / X_1(\omega) = 0, X_2(\omega) = 0) =$$

$$= P(\{(D_1, D_2, E_1), (D_1, D_2, E_2), (D_1, E_1, E_2), (D_2, E_1, E_2)\}) =$$

$$= \frac{1}{120} + \frac{1}{120} + \frac{1}{120} + \frac{1}{120} = \frac{4}{120}$$

Ejemplo

Para el caso particular de este ejemplo, la probabilidad de cada valor $(i, j), i, j = 0, 1, 2, 3$ se podría calcular a través de la siguiente fórmula:

$$P(X_1 = i, X_2 = j) = \frac{\begin{pmatrix} 3 \\ i \end{pmatrix} \begin{pmatrix} 3 \\ j \end{pmatrix} \begin{pmatrix} 4 \\ 3 - i - j \end{pmatrix}}{120}$$

La función masa de probabilidad queda recogida en la siguiente tabla:

X_1 / X_2	0	1	2	3
0	4/120	18/120	12/120	1/120
1	18/120	36/120	9/120	0
2	12/120	9/120	0	0
3	1/120	0	0	0

Ejemplo

Para terminar, la función de distribución de (X_1, X_2) sería:

$$F_X(x_1, x_2) = \begin{cases} 0 & x_1 < 0 \text{ ó } x_2 < 0 \\ 4/120 & 0 \leq x_1 < 1, 0 \leq x_2 < 1 \\ 22/120 & 0 \leq x_1 < 1, 1 \leq x_2 < 2 \\ 34/120 & 0 \leq x_1 < 1, 2 \leq x_2 < 3 \\ 35/120 & 0 \leq x_1 < 1, x_2 \geq 3 \\ 22/120 & 1 \leq x_1 < 2, 0 \leq x_2 < 1 \\ 76/120 & 1 \leq x_1 < 2, 1 \leq x_2 < 2 \\ 97/120 & 1 \leq x_1 < 2, 2 \leq x_2 < 3 \\ 98/120 & 1 \leq x_1 < 2, x_2 \geq 3 \\ 34/120 & 2 \leq x_1 < 3, 0 \leq x_2 < 1 \\ 97/120 & 2 \leq x_1 < 3, 1 \leq x_2 < 2 \\ 118/120 & 2 \leq x_1 < 3, 2 \leq x_2 < 3 \\ 119/120 & 2 \leq x_1 < 3, x_2 \geq 3 \\ 35/120 & x_1 \geq 3, 0 \leq x_2 < 1 \\ 98/120 & x_1 \geq 3, 1 \leq x_2 < 2 \\ 119/120 & x_1 \geq 3, 2 \leq x_2 < 3 \\ 1 & x_1 \geq 3, x_2 \geq 3 \end{cases}$$

Ejemplo 2

Se lanzan 2 dados y se miden el mayor y el menor número obtenido en la tirada. Calcular la función masa de probabilidad y la función de distribución del vector aleatorio correspondiente.

Ejemplo 2

Comenzamos definiendo el experimento aleatorio de lanzar 2 dados. El espacio probabilístico, $(\Omega, \mathcal{A}, P)$ está compuesto de los siguientes elementos:

$$\Omega = \{11, 12, 13, ...\}$$

(Habría un total de $VR_{6,2} = 6^2 = 36$ sucesos elementales)

$$\mathcal{A} = \mathcal{P}(\Omega)$$

$$P(\omega) = \frac{1}{6^2}, \forall \omega \in \Omega$$

Ejemplo 2

Sean:

- X_1: valor más alto obtenido en el lanzamiento de los dados.
- X_2: valor más bajo obtenido en el lanzamiento de los dados.

Cada $X_i, i = 1, 2$ cumple la propiedad $P(X_i \in E_X^i) = 1$, con $E_X^i = 1, 2, 3, 4, 5, 6$.

Ejemplo 2

$$X_1(\omega) = \begin{cases} 1 & \omega \in \{11\} \\ 2 & \omega \in \{21, 12, 22\} \\ \ldots & \\ 6 & \omega \in \{61, 62, 63, 64, 65, 16, 26, 36, 46, 56, 66\} \end{cases}$$

$$X_2(\omega) = \begin{cases} 1 & \omega \in \{11, 21, 31, 41, 51, 61, 12, 13, 14, 15, 16\} \\ 2 & \omega \in \{22, 23, 24, 25, 26, 32, 42, 52, 62\} \\ \ldots & \\ 6 & \omega \in \{66\} \end{cases}$$

Se puede comprobar que $X_1^{-1}(B) \in \mathcal{A}, B \in \mathcal{B}$ y $X_2^{-1}(B) \in \mathcal{A}, B \in \mathcal{B}$, por lo que se demuestra que son variables aleatorias discretas.

Ejemplo 2

El vector aleatorio $X = (X_1, X_2)$ es un vector aleatorio discreto. El conjunto en el que toma valores, E_X, se define como:

$$E_X = E_X^1 \times E_X^2 = \{(1,1),(1,2),(1,3),...\}$$

Y la descripción matemática de la aplicación X sería:

$$X : \Omega \to \mathbb{R}^2$$

Ejemplo 2

Para cada intervalo en $\mathbb{R}^2$ de la forma $(-\infty, x_1] \times (-\infty, x_2], x_1, x_2 \in \mathbb{R}$, y que por tanto pertenezca a la σ-álgebra de Borel $\mathcal{B}^2$, se puede obtener una imagen inversa de X en ese intervalo que equivalga a un suceso aleatorio de $(\Omega, \mathcal{A}, P)$

$$X^{-1}((-\infty, x_1]x(-\infty, x_2]) = \begin{cases} \emptyset & x_1 < 1 \text{ ó } x_2 < 1 \\ \{11\} & 1 \leq x_1 < 2, x_2 \geq 1 \\ \{11, 21, 12\} & 2 \leq x_1 < 3, 1 \leq x_2 < 2 \\ \{11, 21, 12, 22\} & 2 \leq x_1 < 3, x_2 \geq 2 \\ ... & ... \\ \Omega & x_1 \geq 6, x_2 \geq 6 \end{cases}$$

Ejemplo 2

La función de distribución $F(x_1, x_2), x_1, x_2 \in \mathbb{R}$ equivale a $P_X((-\infty, x_1]x(-\infty, x_2]) = P(X^{-1}((-\infty, x_1]x(-\infty, x_2]))$. Por tanto, basta con calcular las probabilidades de los sucesos aleatorios de la diapositiva anterior.

$$P(X^{-1}((-\infty, x_1]x(-\infty, x_2])) = \begin{cases} P(\emptyset) = 0 & x_1 < 1 \text{ ó } x_2 < 1 \\ P(\{11\}) = \dfrac{1}{36} & 1 \leq x_1 < 2, x_2 \geq 1 \\ P(\{11, 21, 12\}) = \dfrac{3}{36} & 2 \leq x_1 < 3, 1 \leq x_2 < 2 \\ P(\{11, 21, 12, 22\}) = \dfrac{4}{36} & 2 \leq x_1 < 3, x_2 \geq 2 \\ ... & ... \\ P(\Omega) = 1 & x_1 \geq 6, x_2 \geq 6 \end{cases}$$

Ejemplo 2

$$F(x,y) = \begin{cases} 0 & x_1 < 1 \text{ ó } x_2 < 1 & 9/36 & 5 \leq x_1 < 6, 1 \leq x_2 < 2 \\ 1/36 & 1 \leq x_1 < 2, x_2 \geq 1 & 16/36 & 5 \leq x_1 < 6, 2 \leq x_2 < 3 \\ 3/36 & 2 \leq x_1 < 3, 1 \leq x_2 < 2 & 21/36 & 5 \leq x_1 < 6, 3 \leq x_2 < 4 \\ 4/36 & 2 \leq x_1 < 3, x_2 \geq 2 & 24/36 & 5 \leq x_1 < 6, 4 \leq x_2 < 5 \\ 5/36 & 3 \leq x_1 < 4, 1 \leq x_2 < 2 & 25/36 & 5 \leq x_1 < 6, x_2 \geq 5 \\ 8/36 & 3 \leq x_1 < 4, 2 \leq x_2 < 3 & 11/36 & x_1 \geq 6, 1 \leq x_2 < 2 \\ 9/36 & 3 \leq x_1 < 4, x_2 \geq 3 & 20/36 & x_1 \geq 6, 2 \leq x_2 < 3 \\ 7/36 & 4 \leq x_1 < 5, 1 \leq x_2 < 2 & 27/36 & x_1 \geq 6, 3 \leq x_2 < 4 \\ 12/36 & 4 \leq x_1 < 5, 2 \leq x_2 < 3 & 32/36 & x_1 \geq 6, 4 \leq x_2 < 5 \\ 15/36 & 4 \leq x_1 < 5, 3 \leq x_2 < 4 & 35/36 & x_1 \geq 6, 5 \leq x_2 < 6 \\ 16/36 & 4 \leq x_1 < 5, x_2 \geq 4 & 1 & x_1 \geq 6, x_2 < 6 \end{cases}$$

Ejemplo 2

La función masa de probabilidad se puede obtener directamente a partir de la expresión

$$P_X(x) = P(\omega \in \Omega / X_1(\omega) = x_1, ..., X_n(\omega) = x_n)$$

Por ejemplo:

$$P_X(1,1) = P(\omega \in \Omega / X_1(\omega) = 1, X_2(\omega) = 1) = P(\{11\}) = 1/36$$

También podría obtenerse hallando los puntos de salto en la función de distribución:

$$P_X(1,1) = F(1,1) - F(1^-,1) - F(1,1^-) + F(1^-,1^-) =$$
$$= 1/36 - 0 - 0 + 0 = 1/36$$

Ejemplo 2

La función masa de probabilidad se resume en la siguiente tabla

$$P(X_1 = i, X_2 = j) =$$

i/j	1	2	3	4	5	6
1	1/36	0	0	0	0	0
2	2/36	1/36	0	0	0	0
3	2/36	2/36	1/36	0	0	0
4	2/36	2/36	2/36	1/36	0	0
5	2/36	2/36	2/36	2/36	1/36	0
6	2/36	2/36	2/36	2/36	2/36	1/36

Ejemplo 3

La Encuesta de Condiciones de Vida (ECV), llevada a cabo por el Instituto Nacional de Estadística, recoge una cantidad de preguntas acerca del estado en el que se encuentran las viviendas donde residen las familias. Supóngase que en una aldea de 5 viviendas se elige una al azar para que participe en la encuesta. Cada una de las viviendas tiene las siguientes características:

Vivienda	H_1	H_2	H_3	H_4	H_5
¿Goteras?	Sí	Sí	No	No	No
¿Calefacción?	No	Sí	No	Sí	No

Sobre la vivienda elegida se mide, a través de variables binarias (0/1) si tiene problemas de goteras, humedades, etc. y si tiene problemas para mantener la temperatura adecuada en invierno. Calcular la función de distribución y la función masa de probabilidad del vector aleatorio resultante.

Ejemplo 3

El experimento aleatorio es el de extraer un hogar al azar.
El espacio probabilístico asociado $(\Omega, \mathcal{A}, P)$ cuenta con los siguientes elementos:

$$\Omega = \{H_1, H_2, H_3, H_4, H_5\}$$

$$\mathcal{A} = \mathcal{P}(\Omega)$$

$$P(\omega) = \frac{1}{5}, \forall \omega \in \Omega$$

Ejemplo 3

Se definen las variables

$$X_1 = \begin{cases} 1 & \text{si hay goteras en el hogar} \\ 0 & \text{si no hay goteras en el hogar} \end{cases}$$

$$X_2 = \begin{cases} 1 & \text{si } \textbf{no} \text{ hay calefacción en el hogar} \\ 0 & \text{si hay calefacción en el hogar} \end{cases}$$

Así pues

$$X_1(\omega) = \begin{cases} 1 & \omega \in \{H_1, H_2\} \\ 0 & \omega \in \{H_3, H_4, H_5\} \end{cases}$$

$$X_2(\omega) = \begin{cases} 1 & \omega \in \{H_1, H_3, H_5\} \\ 0 & \omega \in \{H_2, H_4\} \end{cases}$$

Ejemplo 3

Se demuestra que X_1 y X_2 son variables aleatorias discretas, y que por tanto el vector (X_1, X_2) es un vector aleatorio discreto. El conjunto de puntos donde toma valores, E_X, se obtiene a partir de

$$\left.\begin{array}{l} P(X_1 \in E_X^1) = 1, E_X^1 = \{0, 1\} \\ P(X_2 \in E_X^2) = 1, E_X^2 = \{0, 1\} \end{array}\right\} P((X_1, X_2) \in E_X) = 1,$$

$$E_X = E_X^1 \times E_X^2 = \{(0,0), (0,1), (1,0), (1,1)\}$$

Ejemplo 3

Al igual que en el ejemplo 2, la función de distribución se puede obtener mediante la expresión

$$F(x_1, x_2) = P_X((-\infty, x_1] \times (-\infty, x_2]) = P(X^{-1}((-\infty, x_1] \times (-\infty, x_2]))$$

Así:

$$F(x_1, x_2) = \begin{cases} P(\emptyset) = 0 & x_1 < 0 \text{ ó } x_2 < 0 \\ P(\{H_4\}) = \dfrac{1}{5} & 0 \leq x_1 < 1, 0 \leq x_2 < 1 \\ P(\{H_2, H_4\}) = \dfrac{2}{5} & x_1 \geq 1, 0 \leq x_2 < 1 \\ P(\{H_3, H_4, H_5\}) = \dfrac{3}{5} & 0 \leq x_1 < 1, x_2 \geq 1 \\ P(\Omega) = 1 & x_1 \geq 1, x_2 \geq 1 \end{cases}$$

Ejemplo 3

La función masa de probabilidad se puede obtener directamente a partir de la expresión

$$P_X(x) = P(\omega \in \Omega / X_1(\omega) = x_1, ..., X_n(\omega) = x_n)$$

También podría obtenerse hallando los puntos de salto en la función de distribución:

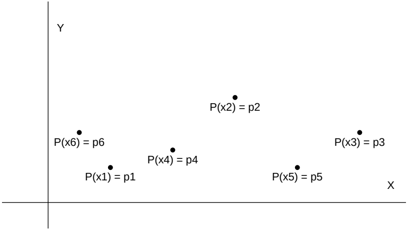
$$P_X(x_1, x_2) = F(x_1, x_2) - F(x_1^-, x_2) - F(x_1, x_2^-) + F(x_1^-, x_2^-)$$

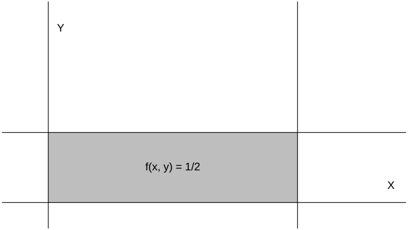
La función masa de probabilidad de este ejemplo se define como sigue:

$$P_X(0,0) = P(X_1 = 0, X_2 = 0) = P(\{H_4\}) = 1/5$$
$$P_X(0,1) = P(X_1 = 0, X_2 = 1) = P(\{H_3, H_5\}) = 2/5$$
$$P_X(1,0) = P(X_1 = 1, X_2 = 0) = P(\{H_2\}) = 1/5$$
$$P_X(1,1) = P(X_1 = 1, X_2 = 1) = P(\{H_1\}) = 1/5$$

Introducción

Los valores de interés no siempre han de ser puntos numerables. En ocasiones, puede definirse la probabilidad de un vector mediante funciones multidimensionales.





Definición

Sea $X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^n, \mathcal{B}^n, P_X)$ un vector aleatorio, se dice que éste es de tipo continuo si existe una función

$$f_X : \mathbb{R}^n \to \mathbb{R}$$

no negativa e integrable tal que

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt, \quad \forall x \in \mathbb{R}^n$$

Equivalentemente, puede decirse que un vector aleatorio es continuo si, para todo conjunto $A \subseteq \mathbb{R}^n$, su función de probabilidad está dada por

$$P_X(A) = P(X \in A) = \int_A f_X(x)dx$$

Propiedades

La función f_X es la función de densidad del vector aleatorio X, y sus propiedades son análogas al caso unidimensional

- $\int_{-\infty}^{+\infty} f_X(t)dt = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} f_X(t_1, ..., t_n)dt_1...dt_n = 1$
- f_X es totalmente continua salvo en un conjunto con medida de Lebesgue nula, y en los puntos de continuidad de f_X, F_X es derivable y
$$\frac{\partial^n F_X(x_1, ..., x_n)}{\partial x_1 ... \partial x_n} = f_X(x_1, ..., x_n)$$
- f_X no es única: puede ser cambiada en conjuntos de medida nula sin afectar a F_X.

El Teorema de Correspondencia es aplicable para vectores aleatorios continuos.

Caracterización

Sea $X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^n, \mathcal{B}^n, P_X)$ un vector aleatorio continuo, cada una de sus componentes $X_i : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}, P_{X_i}), \forall i = 1, ..., n$ es una variable aleatoria continua.

El recíproco no siempre es cierto

(*n* vv. aa. continuas no siempre componen un vector aleatorio continuo)

Ejemplo

Supóngase la función:

$$f(x, y) = k, \quad 0 \leq x \leq 1, 0 \leq y \leq 1$$

¿Qué valor ha de tomar k para que $f(x, y)$ sea función de densidad de un vector aleatorio continuo?

Ejemplo

$f(x,y)$ ha de ser no negativa (por tanto, $k \geq 0$), integrable, y debe cumplir que

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = \int_0^1\int_0^1 k\ dxdy = k\int_0^1\int_0^1 1dxdy =$$

$$= k\int_0^1 [x]_0^1\, dy = k\int_0^1 1dy = k\,[y]_0^1 = k = 1 \Rightarrow \mathbf{k = 1}$$

Ejemplo 2

Supóngase la función:

$$f(x, y) = kx, \quad 0 \leq x \leq 1, 1 \leq y \leq 2$$

¿Qué valor ha de tomar k para que $f(x, y)$ sea función de densidad de un vector aleatorio continuo?

Ejemplo 2

$f(x, y)$ ha de ser no negativa (por tanto, $kx \geq 0 \Rightarrow k \geq 0$), integrable, y debe cumplir que

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dxdy = 1$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dxdy = \int_{1}^{2} \int_{0}^{1} kx \ dxdy = k \int_{1}^{2} \left(\int_{0}^{1} x \ dx \right) dy =$$

$$= k \int_{1}^{2} \left[\frac{x^2}{2} \right]_0^1 dy = \frac{k}{2} \int_{1}^{2} 1 dy = \frac{k}{2} [y]_1^2 = \frac{k}{2} = 1 \Rightarrow \mathbf{k = 2}$$

Ejemplo 2

¿Probabilidad de que X esté entre 0 y 1/2 y que Y esté a su vez por encima de 5/3?

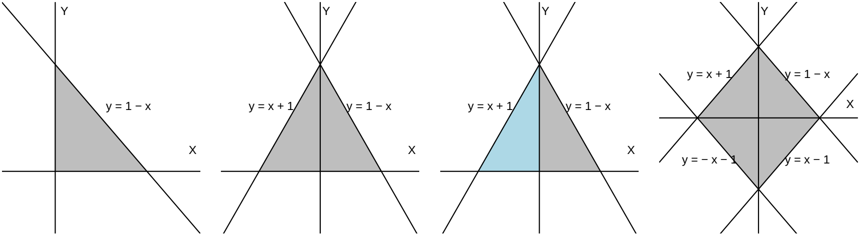
$$P(0 < X < 1/2, Y > 5/3) = \int_{5/3}^{+\infty} \int_0^{1/2} 2x \, dxdy =$$

$$= \int_{5/3}^{2} \left(\int_0^{1/2} 2x \, dx \right) dy = \int_{5/3}^{2} \left[x^2 \right]_0^{1/2} dy = \int_{5/3}^{2} \frac{1}{4} dy =$$

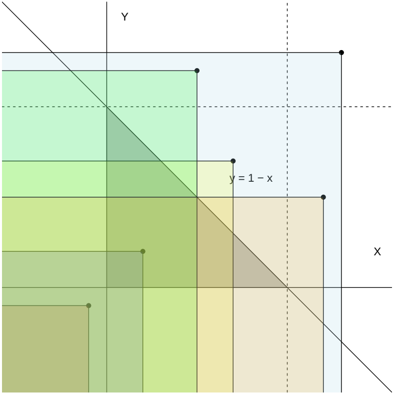
$$= \frac{1}{4} [y]_{5/3}^{2} = \frac{1}{12}$$

Cálculo de la función de distribución

De forma similar al caso discreto, para el cálculo de la función de distribución hay que tener en cuenta las diferentes áreas que pueden aparecer al tomar intervalos de la forma $(-\infty, x_1] \times (-\infty, x_2]$

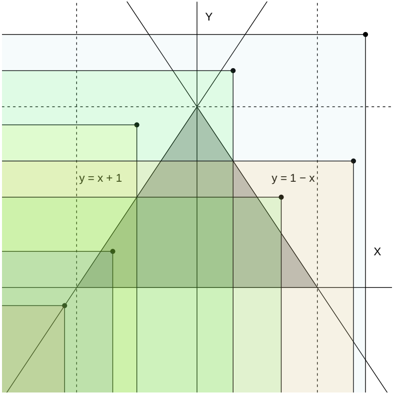


Cálculo de la función de distribución



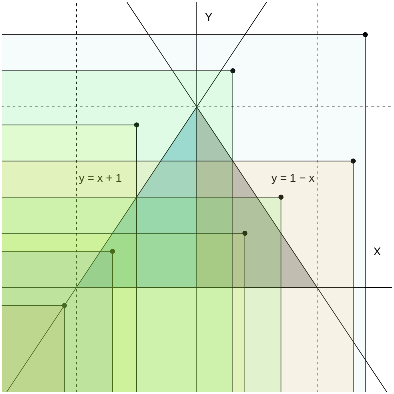
La forma funcional de $F(x, y)$ es diferente en cada uno de los puntos.

Cálculo de la función de distribución



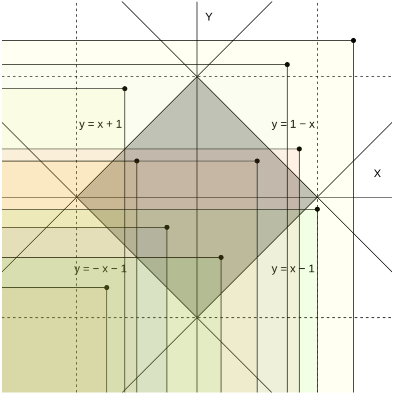
La forma funcional de $F(x, y)$ es diferente en cada uno de los puntos.

Cálculo de la función de distribución



La forma funcional de $F(x, y)$ es diferente en cada uno de los puntos.

Cálculo de la función de distribución



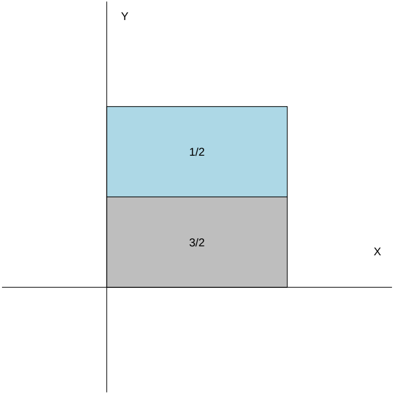
La forma funcional de $F(x, y)$ es diferente en cada uno de los puntos.

Ejemplo

Calcular la función de distribución del vector aleatorio continuo bidimensional con función de densidad:

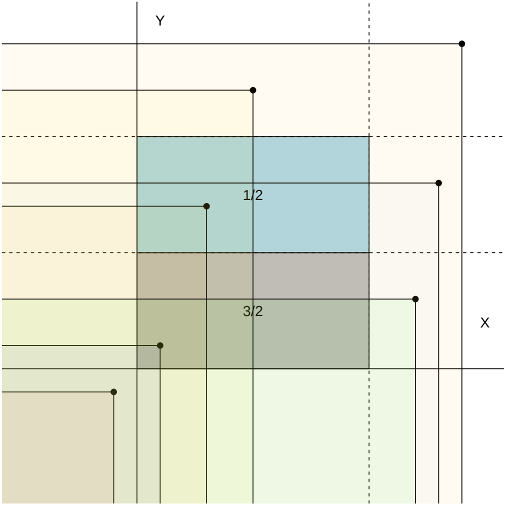
$$f(x,y) = \begin{cases} \dfrac{3}{2} & 0 < x < 1,\ 0 < y < \dfrac{1}{2} \\[2ex] \dfrac{1}{2} & 0 < x < 1, \dfrac{1}{2} < y < 1 \end{cases}$$

Ejemplo



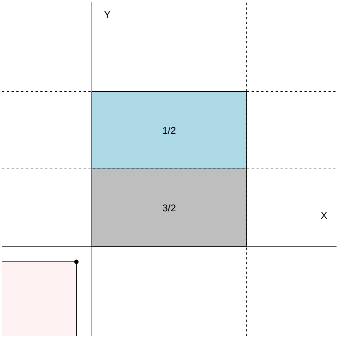
Hay hasta 7 zonas para las cuales la función de distribución tendrá una forma distinta.

Ejemplo



Ejemplo

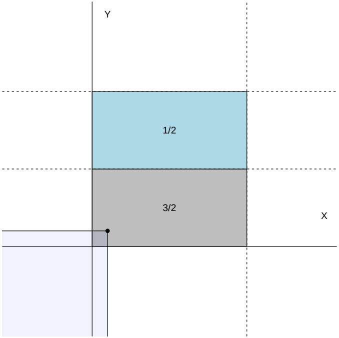
Zona 1: $x < 0$ ó $y < 0$



$$F(x, y) = 0, x < 0 \text{ ó } y < 0$$

Ejemplo

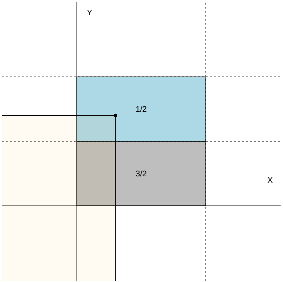
Zona 2: $0 < x < 1, 0 < y < 1/2$



$$F(x,y) = \int_0^y \int_0^x \frac{3}{2} dt_1 \, dt_2 = \int_0^y \frac{3}{2} x \, dt_2 = \frac{3}{2} xy, \; 0 < x < 1, 0 < y < 1/2$$

Ejemplo

Zona 3: $0 < x < 1, 1/2 < y < 1$

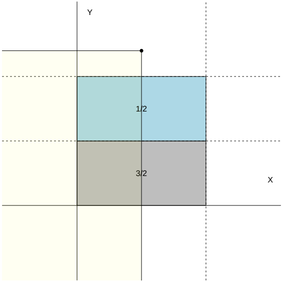


$$F(x,y) = \int_0^{1/2}\int_0^x \frac{3}{2} dt_1\ dt_2 + \int_{1/2}^y \int_0^x \frac{1}{2} dt_1\ dt_2 = \int_0^{1/2} \frac{3}{2} x\ dt_2 + \int_{1/2}^y \frac{1}{2} x\ dt_2 =$$

$$= \frac{3}{4}x + \frac{1}{2}xy - \frac{1}{4}x = \frac{1}{2}(x + xy),\ 0 < x < 1, 1/2 < y < 1$$

Ejemplo

Zona 4: $0 < x < 1, y \geq 1$

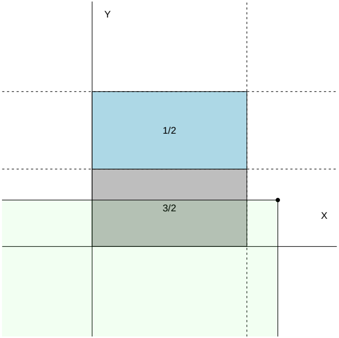


$$F(x,y) = \int_0^{1/2}\int_0^x \frac{3}{2} dt_1\ dt_2 + \int_{1/2}^1 \int_0^x \frac{1}{2} dt_1\ dt_2 = \int_0^{1/2} \frac{3}{2} x\ dt_2 + \int_{1/2}^1 \frac{1}{2} x\ dt_2 =$$

$$= \frac{3}{4}x + \frac{1}{4}x = x,\ 0 < x < 1, y \geq 1$$

Ejemplo

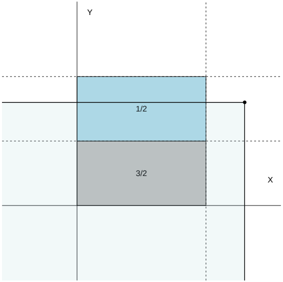
Zona 5: $x \geq 1, 0 < y < 1/2$



$$F(x,y) = \int_0^y \int_0^1 \frac{3}{2} dt_1 \; dt_2 = \int_0^y \frac{3}{2} dt_2 = \frac{3}{2} y, \;\; x \geq 1, 0 < y < 1/2$$

Ejemplo

Zona 6: $x \geq 1, 1/2 < y < 1$

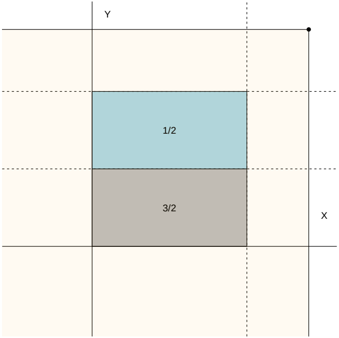


$$F(x,y) = \int_0^{1/2}\int_0^1 \frac{3}{2} dt_1\ dt_2 + \int_{1/2}^{y}\int_0^1 \frac{1}{2} dt_1\ dt_2 = \int_0^{1/2} \frac{3}{2}\ dt_2 + \int_{1/2}^{y} \frac{1}{2}\ dt_2 =$$

$$= \frac{3}{4} + \frac{1}{2}y - \frac{1}{4} = \frac{1}{2}(1+y),\ x \geq 1, 1/2 < y < 1$$

Ejemplo

Zona 7: $x \geq 1, y \geq 1$



$$F(x,y) = \int_0^{1/2} \int_0^1 \frac{3}{2} dt_1 \ dt_2 + \int_{1/2}^1 \int_0^1 \frac{1}{2} dt_1 \ dt_2 = 1, \ x \geq 1, y \geq 1$$

Ejemplo

$$F(x,y) = \begin{cases} 0 & x < 0 \text{ ó } y < 0 \\ \frac{3}{2}xy & 0 < x < 1, 0 < y < 1/2 \\ \frac{1}{2}(x + xy) & 0 < x < 1, 1/2 < y < 1 \\ x & 0 < x < 1, y \geq 1 \\ \frac{3}{2}y & x \geq 1, 0 < y < 1 \\ \frac{1}{2}(1 + y) & x \geq 1, 1/2 < y < 1 \\ 1 & x \geq 1, y \geq 1 \end{cases}$$

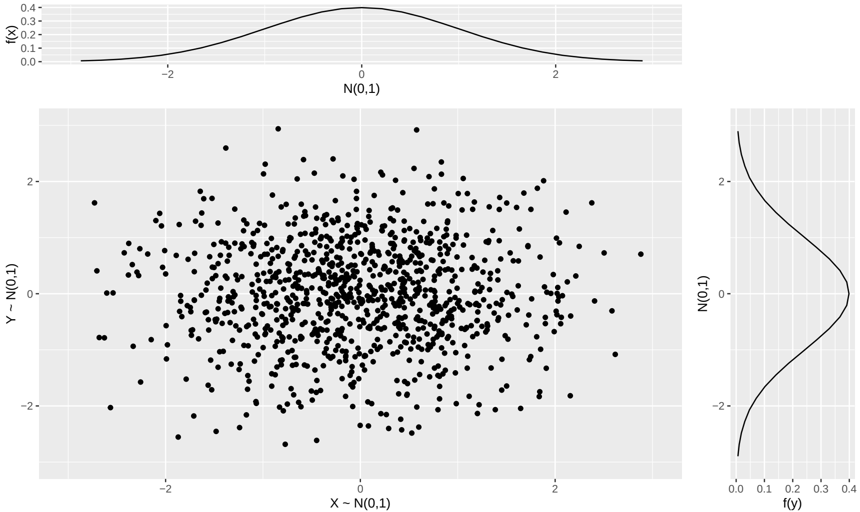
Introducción

La distribución de un vector aleatorio:

$$X = (X_1, X_2, ..., X_n)$$

se denomina **distribución conjunta.**
La distribución de cada componente del vector aleatorio: X_1, X_2, ... se denomina **distribución marginal** de dicha componente

Introducción



Función de distribución

$X = (X_1, ..., X_n)$ vector aleatorio con función de distribución
$F_X(x), x = (x_1, ..., x_n) \in \mathbb{R}$
Función de distribución de una componente:

$$F_{X_i}(x_i) = F_X(+\infty, ..., +\infty, x_i, +\infty, ..., +\infty),\ \forall i = 1, ..., n/x_i \in \mathbb{R}$$

Función de distribución de k componentes:

$$F_{X_{i_1},...,X_{i_k}}(x_{i_1}, ..., x_{i_k}) = F_X(+\infty, ..., +\infty, x_{i_1}, +\infty, ..., +\infty, x_{i_k}, +\infty, ..., +\infty)$$

Planteamiento

$X = (X_1, ..., X_n)$ vector aleatorio discreto con función masa de prob. $P(X = x) \forall x \in E_X$
La función masa de prob. de una componente se obtiene tal que:

$$P(X_i = x_i) = \sum_{\substack{x \in E_X \\ (x)_i = x_i}} P(X = x) =$$

$$= \sum_{\substack{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n \\ (x_1,\ldots,x_i,\ldots,x_n) \in E_X}} P(X_1 = x_1, ..., X_{i-1} = x_{i-1}, X_i = x_i, X_{i+1} = x_{i+1}, ..., X_n = x_n)$$

Se puede generalizar a k componentes; la fórmula quedaría igual pero quedarían fijos los valores $x_{i_1}, ..., x_{i_k}$

Ejemplo

Se extrae una muestra sin reemplazamiento de dos piezas de una caja de piezas de ajedrez de ambos colores (blancas y negras), y sobre esas dos se mide:

- X: número de piezas blancas en la muestra
- Y: numero de caballos en la muestra

A partir de la distribución conjunta de (X, Y), obtener las distribuciones marginales.

Ejemplo

Experimento aleatorio: la extracción sin reemplazamiento de dos piezas de la caja.
Sean (P, T, C, A, R, RR) los tipos de piezas que pueden obtenerse y (B, N) su color:

$$\Omega = \{(P_1B, P_2B), \ldots, (T_1B, T_2B), (C_1B, RB), \ldots, (P_1N, P_2N), \ldots\}$$

(Habría $\binom{32}{2} = 496$ sucesos elementales)

$$\mathcal{A} = \mathcal{P}(\Omega)$$

$$P(\omega) = \frac{1}{496}, \omega \in \Omega$$

Ejemplo

La función masa de probabilidad de (X, Y) se refleja en la siguiente tabla:

X/Y	0	1	2
0	91/496	28/496	1/496
1	196/496	56/496	4/496
2	91/496	28/496	1/496

Ejemplo

X/Y	0	1	2	p_{X_i}
0	91/496	28/496	1/496	120/496
1	196/496	56/496	4/496	256/496
2	91/496	28/496	1/496	120/496

X es un vector aleatorio discreto con $P(X \in E_X) = 1, E_X = \{0, 1, 2\}$
Función masa de probabilidad de X:

$$P(X=0) = P(X=0, Y=0) + P(X=0, Y=1) + P(X=0, Y=2) = \frac{120}{496}$$

$$P(X=1) = P(X=1, Y=0) + P(X=1, Y=1) + P(X=1, Y=2) = \frac{256}{496}$$

$$P(X=2) = P(X=2, Y=0) + P(X=2, Y=1) + P(X=2, Y=2) = \frac{120}{496}$$

Ejemplo

X/Y	0	1	2
0	91/496	28/496	1/496
1	196/496	56/496	4/496
2	91/496	28/496	1/496
p_{Y_j}	378/496	112/496	6/496

Y es un vector aleatorio discreto con $P(Y \in E_Y) = 1, E_Y = \{0, 1, 2\}$
Función masa de probabilidad de Y:

$$P(Y=0) = P(X=0, Y=0) + P(X=1, Y=0) + P(X=2, Y=0) = \frac{378}{496}$$

$$P(Y=1) = P(X=0, Y=1) + P(X=1, Y=1) + P(X=2, Y=1) = \frac{112}{496}$$

$$P(Y=2) = P(X=0, Y=2) + P(X=1, Y=2) + P(X=2, Y=2) = \frac{6}{496}$$

Planteamiento

$X = (X_1, ..., X_n)$ vector aleatorio continuo con función de densidad f_X
Cada componente X_i es una variable aleatoria continua con función de distribución

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

siendo $f_i(t_i)$ el resultado de

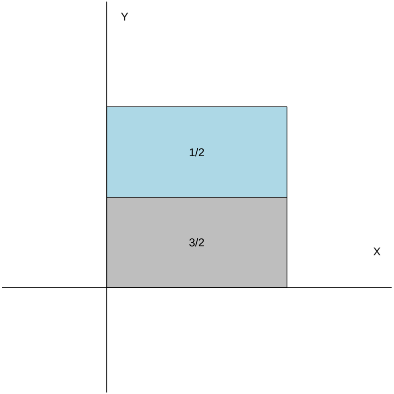
$$f_i(t_i) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} f_x(t_1, ..., t_i, ..., t_n)dt_1 ... dt_{i-1}dt_{i+1} ... dt_n \quad \forall t_i \in \mathbb{R}$$

(se integra la función de densidad conjunta respecto al resto de componentes)

Ejemplo

Calcular las distribuciones marginales del vector (X, Y) con función de densidad:

$$f(x,y) = \begin{cases} \dfrac{3}{2} & 0 < x < 1,\ 0 < y < \dfrac{1}{2} \\[2ex] \dfrac{1}{2} & 0 < x < 1, \dfrac{1}{2} < y < 1 \end{cases}$$



Ejemplo

Función de densidad de X, $f_1(x)$:

$$f_1(x) = \int_0^{1/2} \frac{3}{2} dy + \int_{1/2}^{1} \frac{1}{2} dy = \frac{3}{2} \left[y\right]_0^{1/2} + \frac{1}{2} \left[y\right]_{1/2}^{1} = 1$$

$$f_1(x) = 1, \; 0 < x < 1$$

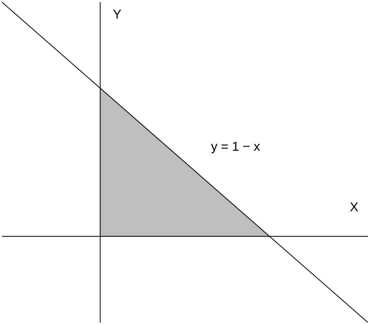
Función de densidad de Y, $f_2(y)$:

$$f_2(y) = \begin{cases} \displaystyle\int_0^1 \frac{3}{2} dx = \frac{3}{2} \left[x\right]_0^1 = \frac{3}{2}, & 0 < y < \dfrac{1}{2} \\[2ex] \displaystyle\int_0^1 \frac{1}{2} dx = \frac{1}{2} \left[x\right]_0^1 = \frac{1}{2}, & \dfrac{1}{2} < y < 1 \end{cases}$$

Ejemplo 2

Calcular las distribuciones marginales del vector (X, Y) con función de densidad:

$$f(x, y) = 2, \quad 0 < y < 1 - x < 1$$



Ejemplo 2

Función de densidad de X, $f_1(x)$:

$$f_1(x) = \int_0^{1-x} 2dy = [2y]_0^{1-x} = 2 - 2x, \quad 0 < x < 1$$

Función de densidad de Y, $f_2(y)$:

$$f_2(y) = \int_0^{1-y} 2dx = [2x]_0^{1-y} = 2 - 2y, \quad 0 < y < 1$$

Introducción

- En vectores aleatorios, la información sobre un subconjunto de los componentes puede ser importante.
- Esta información será a la que estén condicionados los cálculos de probabilidad posteriores, siguiendo el razonamiento de la probabilidad condicionada:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Definición

$X = (X_1, ..., X_n)$ vector aleatorio discreto con función masa de prob. $P(X = x)$ $\forall x \in E_X$
Sea $(X_{i1}, ..., X_{i_k})$ un subconjunto de componentes de dicho vector aleatorio y $(x^*_{i_1}, ..., x^*_{i_k}) \in \mathbb{R}^k$ los valores a los que se condicionan dichos componentes, de forma que $P(X_{i_1} = x^*_{i_1}, ..., X_{i_k} = x^*_{i_k}) > 0$.
Se define la distribución condicionada como:

$$P(X_1 = x_1, ..., X_{i_1-1} = x_{i_1-1}, X_{i_1+1} = x_{i_1+1}, ..., X_{i_k-1} = x_{i_k-1}, X_{i_k+1} = x_{i_k+1}/$$

$$/X_{i_1} = x^*_{i_1}, ..., X_{i_k} = x^*_{i_k}) =$$

$$= \frac{P(X_1 = x_1, ..., X_{i_1} = x^*_{i_1}, ..., X_{i_k} = x^*_{i_k}, ..., X_n = x_n)}{P(X_{i_1} = x^*_{i_1}, ..., X_{i_k} = x^*_{i_k})}$$

Distribución bidimensional condicionada a una variable

$X = (X_1, X_2)$ vector aleatorio discreto con función masa de prob. $P(X = x)\ \forall x \in E_X$
Tras el experimento, se sabe que X_1 ha tomado un determinado valor. ¿Cómo afectará a las probabilidades de cada x_2?

$$P(X_2 = x_2 | X_1 = x_1^*) = \begin{cases} \dfrac{P(X_1 = x_1^*, X_2 = x_2)}{P(X_1 = x_1^*)} & x_2 \in E_{X_2} \\ 0 & x_2 \notin E_{X_2} \end{cases}$$

Ejemplo

Se considera el experimento aleatorio de seleccionar al azar a una persona en este aula, sobre la cual se miden dos variables:
X_1: número de hermanos de la persona
X_2: número de hermanas de la persona
Considérese que $E_{X_1} = E_{X_2} = \{0, 1, 2\}$. La función masa de probabilidad del vector aleatorio $X = (X_1, X_2)$ se observa en la siguiente tabla:

X_1/X_2	0	1	2
0	0.45	0.2	0.02
1	0.2	0.05	0.02
2	0.02	0.02	0.02

Ejemplo

Supóngase que se elige a alguien al azar y se verifica que no tiene hermanos.
Probabilidad de que no tenga hermanas:

$$P(X_2 = 0|X_1 = 0) = \frac{P(X_1 = 0, X_2 = 0)}{P(X_1 = 0)} =$$

$$= \frac{P(X_1 = 0, X_2 = 0)}{P(X_1 = 0, X_2 = 0) + P(X_1 = 0, X_2 = 1) + P(X_1 = 0, X_2 = 2)} =$$

$$= \frac{0.45}{0.45 + 0.2 + 0.02} = 0.67$$

X_1/X_2	0	1	2
0	**0.45**	**0.2**	**0.02**
1	0.2	0.05	0.02
2	0.02	0.02	0.02

Ejemplo

Supóngase que se elige a alguien al azar y se verifica que no tiene hermanos.
Probabilidad de que tenga alguna hermana:

$$P(X_2 > 0 | X_1 = 0) = \frac{P(X_1 = 0, X_2 > 0)}{P(X_1 = 0)} =$$

$$= \frac{P(X_1 = 0, X_2 = 1) + P(X_1 = 0, X_2 = 2)}{P(X_1 = 0, X_2 = 0) + P(X_1 = 0, X_2 = 1) + P(X_1 = 0, X_2 = 2)} =$$

$$= \frac{0.2 + 0.02}{0.45 + 0.2 + 0.02} = 0.33$$

X_1/X_2	0	1	2
0	**0.45**	**0.2**	**0.02**
1	0.2	0.05	0.02
2	0.02	0.02	0.02

Ejemplo

Supóngase que se elige a alguien al azar y se verifica que tiene menos de dos hermanos.
Probabilidad de que no tenga ninguna hermana:

$$P(X_2 = 0 | X_1 < 2) = \frac{P(X_1 < 2, X_2 = 0)}{P(X_1 < 2)} =$$

$$= \frac{P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = 0)}{P(X_1 = 0, X_2 = 0) + ... + P(X_1 = 1, X_2 = 2)} =$$

$$= \frac{0.45 + 0.2}{0.45 + 0.2 + 0.02 + 0.2 + 0.05 + 0.02} = 0.69$$

X_1/X_2	0	1	2
0	**0.45**	**0.2**	**0.02**
1	**0.2**	**0.05**	**0.02**
2	0.02	0.02	0.02

Definición

$X = (X_1, ..., X_n)$ vector aleatorio continuo con función de densidad conjunta $f_X(x_1, ..., x_n), \forall\ (x_1, ..., x_n) \in \mathbb{R}$
Sea $(X_{i1}, ..., X_{i_k})$ un subconjunto de componentes de dicho vector aleatorio y $(x^*_{i_1}, ..., x^*_{i_k}) \in \mathbb{R}^k$ los valores a los que se condicionan dichos componentes, de forma que $f_{X_{i1},...,X_{ik}}(x^*_{i1}, ..., x^*_{ik}) > 0$.
Se define la distribución condicionada como:

$$f_{X_1,\ldots,X_{i_1-1},X_{i_1+1},\ldots,X_{i_k-1},X_{i_k+1},\ldots,X_n|X_{i_1}=x^*_{i_1},\ldots,X_{i_k}=x^*_{i_k}}(x_1, \ldots, x_{i_1-1}, x_{i_1+1}, \ldots, x_{i_k-1}, \ldots, x_n/x^*_{i_1}, \ldots, x^*_{i_k}) =$$

$$= \frac{f_X(x_1, ..., x^*_{i_1}, ..., x^*_{i_k}, ..., x_n)}{f_{X_{i_1},...,X_{i_k}}(x^*_{i_1}, ..., x^*_{i_k})}$$

Distribución bidimensional condicionada a una variable

$X = (X_1, X_2)$ vector aleatorio continuo con función de densidad conjunta $f_X(x_1, x_2)$
Tras el experimento, se sabe que X_1 ha tomado un determinado valor x_1^* tal que $f_{X_1}(x_1) > 0$. ¿Cómo afectará a la densidad de x_2?

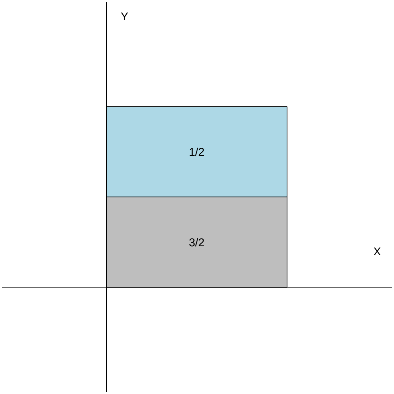
$$f_{X_2|X_1=x_1^*}(x_2/x_1^*) = \frac{f_X(x_1^*, x_2)}{f_{X_1}(x_1^*)}$$

(Cociente entre la densidad conjunta y la densidad marginal de x_1, evaluadas en los puntos x_1^* y x_2)

Ejemplo

Sea $X = (X, Y)$ el vector aleatorio continuo con función de densidad:

$$f(x,y) = \begin{cases} \dfrac{3}{2} & 0 < x < 1,\ 0 < y < \dfrac{1}{2} \\ \dfrac{1}{2} & 0 < x < 1, \dfrac{1}{2} < y < 1 \end{cases}$$



Se pide calcular $P(Y > 0.5|X = 0.25)$ y $P(X > 0.5|Y = 0.25)$

Ejemplo

Las densidades marginales se obtuvieron en la diapositiva 85:

$$f_X(x) = 1,\ 0 < x < 1$$

$$f_Y(y) = \begin{cases} \dfrac{3}{2}, & 0 < y < \dfrac{1}{2} \\ \dfrac{1}{2}, & \dfrac{1}{2} < y < 1 \end{cases}$$

Por tanto:

$$f_{Y|X=0.25}(y) = \frac{f(0.25, y)}{f_X(0.25)} = \begin{cases} \dfrac{3/2}{1} = \dfrac{3}{2}, & 0 < y < \dfrac{1}{2} \\ \dfrac{1/2}{1} = \dfrac{1}{2}, & \dfrac{1}{2} < y < 1 \end{cases}$$

$$f_{X|Y=0.25}(x) = \frac{f(x, 0.25)}{f_Y(0.25)} = \frac{3/2}{3/2} = 1, 0 < x < 1$$

Ejemplo

Así pues, las probabilidades buscadas se calculan utilizando las funciones de densidad condicionadas:

$$P(Y > 0.5|X = 0.25) = \int_{0.5}^{+\infty} f_{Y|X=0.25}(y)dy = \int_{0.5}^{1} \frac{1}{2}dy = 0.25$$

$$P(X > 0.5|Y = 0.25) = \int_{0.5}^{+\infty} f_{X|Y=0.25}(x)dx = \int_{0.5}^{1} 1dx = 0.5$$

Definición

Sea X un vector aleatorio n-dimensional e $Y = g(X)$ con

$$g : (\mathbb{R}^n, \mathcal{B}^n) \to (\mathbb{R}^m, \mathcal{B}^m)$$

Y es un vector aleatorio m-dimensional cuya distribución puede hallarse a partir de X

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \in g^{-1}((-\infty, y]))$$

Definición

X vector aleatorio n-dimensional discreto definido en E_X y $g : (\mathbb{R}^n, \mathcal{B}^n) \to (\mathbb{R}^m, \mathcal{B}^m)$. El vector $Y = g(X)$ es un vector aleatorio m-dimensional discreto definido en $g(E_X)$ con función masa de probabilidad:

$$P(Y = y) = P(X \in g^{-1}(y)) = \sum_{x \in g^{-1}(y)} P(X = x), \quad y \in g(E_X)$$

Ejemplo

Se define el experimento aleatorio de observar el resultado de un partido del Granada Club de Fútbol de la presente temporada, sobre el que se miden dos variables:
X_1: número de goles a favor
X_2: número de goles en contra
Considérese que la función masa de probabilidad del vector aleatorio $X = (X_1, X_2)$ se define según la siguiente tabla:

X_1/X_2	0	1	2	3
0	0.09	0.09	0.03	0.03
1	0.06	0.09	0.14	0
2	0.11	0.15	0.06	0.03
3	0	0.06	0.03	0
4	0	0.03	0	0

Ejemplo

Sobre ese experimento, se desea obtener el vector $Y = (Y_1, Y_2)$ donde:
Y_1: número de goles totales en el partido observado
Y_2: puntos conseguidos por el Granada en el partido observado (3 si gana, 1 si empata, 0 si pierde)
Por lo tanto:

$$Y_1 = X_1 + X_2$$

$$P(Y_1 \in E_{Y_1}) = 1, E_{Y_1} = \{0, 1, 2, 3, 4, 5\}$$

$$Y_2 = \begin{cases} 3 & X_1 > X_2 \\ 1 & X_1 = X_2 \\ 0 & X_1 < X_2 \end{cases}$$

Ejemplo

X_1/X_2	0	1	2	3
0	0.09	0.09	0.03	0.03
1	0.06	0.09	0.14	0
2	0.11	0.15	0.06	0.03
3	0	0.06	0.03	0
4	0	0.03	0	0

$$P(Y_1 = 0, Y_2 = 0) = P(X_1 + X_2 = 0, X_1 < X_2) = 0$$

$$P(Y_1 = 0, Y_2 = 1) = P(X_1 + X_2 = 0, X_1 = X_2) = P(X_1 = 0, X_2 = 0) = 0.09$$

$$P(Y_1 = 0, Y_2 = 3) = P(X_1 + X_2 = 0, X_1 > X_2) = 0$$

$$P(Y_1 = 1, Y_2 = 0) = P(X_1 + X_2 = 1, X_1 < X_2) = P(X_1 = 0, X_2 = 1) = 0.09$$

$$P(Y_1 = 1, Y_2 = 1) = P(X_1 + X_2 = 1, X_1 = X_2) = 0$$

$$P(Y_1 = 1, Y_2 = 3) = P(X_1 + X_2 = 1, X_1 > X_2) = P(X_1 = 1, X_2 = 0) = 0.06$$

...

Ejemplo

La tabla final quedaría como sigue:

Y_1 / Y_2	0	1	3
0	0	0.09	0
1	0.09	0	0.06
2	0.03	0.09	0.11
3	0.18	0	0.14
4	0	0.06	0.06
5	0.03	0	0.06

Definición

X vector aleatorio n-dimensional continuo con función de densidad $f_X(x)$ y $g : (\mathbb{R}^n, \mathcal{B}^n) \to (\mathbb{R}^m, \mathcal{B}^m)$. El vector $Y = g(X)$ es un vector aleatorio m-dimensional discreto definido en $g(E_X)$ con función masa de probabilidad:

$$P(Y = y) = P(X \in g^{-1}(y)) = \int_{g^{-1}(y)} f_X(x)dx$$

Ejemplo

Considérese el vector aleatorio continuo (X, Y) con función de densidad conjunta:

$$f(x,y) = e^{-x-y}, \quad x > 0, y > 0$$

Se desea obtener la distribución del vector aleatorio $Z = (Z_1, Z_2)$, donde:

$$Z_1 = \begin{cases} 0 & x > y \\ 1 & x < y \end{cases} \qquad\qquad Z_2 = \begin{cases} 0 & y < 1 - x \\ 1 & y > 1 - x \end{cases}$$

Ejemplo

$$P(Z_1 = 0, Z_2 = 0) = P(X > Y, Y < 1 - X) = P(X > Y, X < 1 - Y) =$$

$$= P(Y < X < 1 - Y) = \int_0^{1/2} \int_y^{1-y} e^{-x-y} dx\, dy =$$

$$= \int_0^{1/2} (-1) \left[e^{-x-y}\right]_y^{1-y} dy = ... = \frac{1}{2} - \frac{1}{e} \approx 0.132$$

$$P(Z_1 = 1, Z_2 = 0) = P(X < Y, Y < 1 - X) = P(X < Y < 1 - X) =$$

$$= \int_0^{1/2} \int_x^{1-x} e^{-x-y} dy\, dx = ... \approx 0.132$$

Ejemplo

$$P(Z_1 = 0, Z_2 = 1) = P(X > Y, Y > 1 - X) = P(0 < 1 - X < Y < X) =$$

$$= \int_{1/2}^{1} \int_{1-x}^{x} e^{-x-y} dy\, dx + \int_{1}^{+\infty} \int_{0}^{x} e^{-x-y} dy\, dx = \dots =$$

$$= \frac{1}{2e^2} + \frac{2e-1}{2e^2} \approx 0.368$$

$$P(Z_1 = 1, Z_2 = 1) = P(X < Y, Y > 1 - X) = P(X < Y, X > 1 - Y) =$$

$$= P(0 < 1-Y < X < Y) = \int_{1/2}^{1} \int_{1-y}^{y} e^{-x-y} dx\, dy + \int_{1}^{+\infty} \int_{0}^{y} e^{-x-y} dx\, dy =$$

$$= \dots = \frac{1}{2e^2} + \frac{2e-1}{2e^2} \approx 0.368$$

Definición

X vector aleatorio n-dimensional continuo con función de densidad $f_X(x)$ y $g : (\mathbb{R}^n, \mathcal{B}^n) \to (\mathbb{R}^m, \mathcal{B}^m)$ una transformación medible tal que $Y = g(X)$ es un vector aleatorio m-dimensional continuo.
La función de distribución de Y se puede obtener como:

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \in g^{-1}((-\infty, y])) =$$

$$= \int_{g^{-1}((-\infty,y])} f_X(x)dx$$

Ejemplo

Sea (X, Y) un vector aleatorio continuo con función de densidad conjunta:

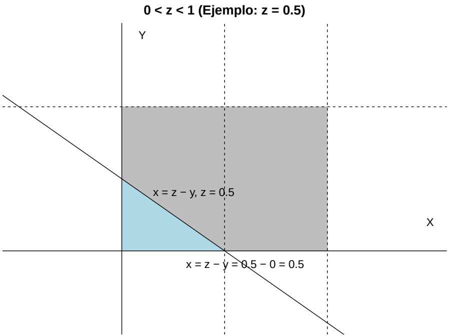
$$f(x, y) = 1, 0 < x < 1, 0 < y < 1$$

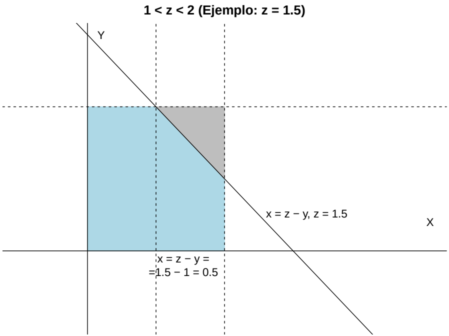
Se pide obtener la distribución de probabilidad de $Z = X + Y$.

Ejemplo

$$F_Z(z) = P(X + Y \leq z) = P(X \leq z - Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{z-y} 1 \, dx \, dy$$

La función de distribución resultante se dividirá en dos trozos:





Ejemplo

$$F_Z(z) = \begin{cases} \int_0^z \int_0^{z-y} 1 \, dx \, dy = \frac{z^2}{2} & 0 \leq z < 1 \\ \int_0^1 \int_0^{z-1} 1 \, dx \, dy + \int_{z-1}^1 \int_0^{z-x} 1 \, dy \, dx = \\ = -\frac{z^2}{2} + 2z - 1 & 1 \leq z < 2 \end{cases}$$

La función de densidad se obtiene derivando la función de distribución:

$$f_Z(z) = \begin{cases} z & 0 \leq z < 1 \\ 2 - z & 1 \leq z < 2 \end{cases}$$

Teorema del cambio de variable multidimensional

La función de densidad de $Y = g(X)$, en el caso de que $g : (\mathbb{R}^n, \mathcal{B}^n) \to (\mathbb{R}^{\mathbf{n}}, \mathcal{B}^{\mathbf{n}})$ sea derivable e invertible (salvo en casos especiales), se puede obtener mediante la aplicación del teorema de cambio de variable en el caso multidimensional:

$$f_Y(y) = f_X(g^{-1}(y))|J|, \ \forall y \in \mathbb{R}^n$$

donde $J \neq 0$ es el jacobiano (determinante de la matriz jacobiana) de g.

$$J = \begin{vmatrix} \frac{\partial g_1^{-1}(y_1,\ldots,y_n)}{\partial y_1} & \ldots & \frac{\partial g_1^{-1}(y_1,\ldots,y_n)}{\partial y_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial g_n^{-1}(y_1,\ldots,y_n)}{\partial y_1} & \ldots & \frac{\partial g_n^{-1}(y_1,\ldots,y_n)}{\partial y_n} \end{vmatrix}$$

Ejemplo (teorema del cambio de variable)

Retomamos el ejemplo anterior:

$$f(x, y) = 1, 0 < x < 1, 0 < y < 1$$

Se pide obtener la distribución de probabilidad de $Z = X + Y$, ahora mediante el teorema del cambio de variable.
Para poder aplicarlo, se obtiene una nueva componente auxiliar: $U = X - Y$.

Ejemplo (teorema del cambio de variable)

$$
\begin{array}{c} Z = X + Y \\ \\ U = X - Y \end{array} \quad \Rightarrow \quad \begin{array}{c} X = \dfrac{Z+U}{2} = g_1^{-1}(z,u) \\ \\ Y = \dfrac{Z-U}{2} = g_2^{-1}(z,u) \end{array}
$$

$$
J = \begin{vmatrix} \frac{\partial g_1^{-1}(z,u)}{\partial z} & \frac{\partial g_2^{-1}(z,u)}{\partial z} \\ \\ \frac{\partial g_1^{-1}(z,u)}{\partial u} & \frac{\partial g_2^{-1}(z,u)}{\partial u} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}
$$

$$
f_{Z,U}(z,u) = f_{X,Y}\left(\frac{z+u}{2}, \frac{z-u}{2}\right)\frac{1}{2} = \frac{1}{2},\ 0 < \frac{z+u}{2} < 1, 0 < \frac{z-u}{2} < 1
$$

Ejemplo (teorema del cambio de variable)

Para sacar la función de densidad de Z (la variable que realmente nos interesa), calculamos la marginal de (Z, U).

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{Z,U}(z,u)\, du = \begin{cases} \int_{-z}^{z} 1/2\, du = z & 0 < z < 1 \\ \int_{z-2}^{2-z} 1/2\, du = 2 - z & 1 < z < 2 \end{cases}$$

Introducción

En apartados anteriores, se ha estudiado la equivalencia:

$$\text{Distribución conjunta} \Rightarrow \text{Distribuciones marginales}$$

Su recíproco:

$$\text{Distribución conjunta} \Leftarrow \text{Distribuciones marginales}$$

en general **no** es cierto... salvo cuando las vv. aa. que componen el vector son **independientes** entre sí.

Definición

Sean $X_1, ..., X_n$ vv. aa. definidas sobre un mismo espacio probabilístico $(\Omega, \mathcal{A}, P)$. Se dice que las variables $X_1, ..., X_n$ son mutuamente independientes si

$$F_X(x_1, ..., x_n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot ... \cdot F_{X_n}(x_n), \ \forall x_1, ..., x_n \in \mathbb{R}$$

(la función de distribución conjunta equivale al producto de las funciones de distribución marginales)

Caso discreto

Sean $X_1, \ldots, X_n$ vv. aa. unidimensionales **discretas** definidas sobre un mismo espacio probabilístico $(\Omega, \mathcal{A}, P)$

$$X_1, \ldots, X_n \text{ son independientes}$$

$$\Updownarrow$$

$$P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \ldots \cdot P(X_n = x_n)$$

$$\forall (x_1, \ldots, x_n) \in \mathbb{R}^n$$

Caso continuo

Sean $X_1, ..., X_n$ vv. aa. unidimensionales **continuas** definidas sobre un mismo espacio probabilístico $(\Omega, \mathcal{A}, P)$

$$X_1, ..., X_n \text{ indep.} \Leftrightarrow \begin{cases} X = (X_1, ..., X_n) \text{ es un vector aleatorio continuo} \\ \\ f_X(x_1, ..., x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot ... \cdot f_{X_n}(x_n), \\ \forall (x_1, ..., x_n) \in \mathbb{R}^n \end{cases}$$

Caracterización por conjuntos de Borel

Sean $X_1, ..., X_n$ vv. aa. unidimensionales definidas sobre un mismo espacio probabilístico $(\Omega, \mathcal{A}, P)$

$$X_1, ..., X_n \text{ son independientes}$$

$$\Updownarrow$$

$$P(X_1 \in B_1, ..., X_n \in B_n) = P(X_1 \in B_1) \cdot P(X_2 \in B_2) \cdot ... \cdot P(X_n \in B_n)$$

$$\forall B_1, ..., B_n \in \mathcal{B}$$

Caracterización por factorización

Sean $X_1, ..., X_n$ vv. aa. unidimensionales **discretas** definidas sobre un mismo espacio probabilístico $(\Omega, \mathcal{A}, P)$

$$X_1, ..., X_n \text{ son independientes}$$

$$\Updownarrow$$

$$P(X_1 = x_1, ..., X_n = x_n) = h_1(x_1) \cdot h_2(x_2) \cdot ... \cdot h_n(x_n)$$

$$\forall (x_1, ..., x_n) \in \mathbb{R}^n$$

Caracterización por factorización

Sean $X_1, ..., X_n$ vv. aa. unidimensionales **continuas** definidas sobre un mismo espacio probabilístico $(\Omega, \mathcal{A}, P)$

$$X_1, ..., X_n \text{ indep.} \Leftrightarrow \begin{cases} X = (X_1, ..., X_n) \text{ es un vector aleatorio continuo} \\ \\ f_X(x_1, ..., x_n) = h_1(x_1) \cdot h_2(x_2) \cdot ... \cdot h_n(x_n), \\ \forall (x_1, ..., x_n) \in \mathbb{R}^n \end{cases}$$

Teoremas

Si $X_1, ..., X_n$ son independientes:

1. Si $g_i : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ son funciones medibles, entonces $g_1(X_1), ..., g_n(X_n)$ son independientes.
2. Cualquier subcolección $X_{i_1}, ..., X_{i_k}$ también lo son.
3. Las distribuciones condicionadas coinciden con las marginales.

Definición

Sean $X_1, ..., X_n$ vv. aa. unidimensionales definidas sobre un mismo espacio probabilístico $(\Omega, \mathcal{A}, P)$, se define la independencia dos a dos como:

$$\forall i, j = 1, ..., n, i \neq j, X_i \perp X_j \text{ } (X_i \text{ y } X_j \text{ son independientes})$$

Si $X_1, ..., X_n$ son independientes, entonces hay independencia dos a dos en todas ellas, pero el recíproco **no** es cierto.

Definición

Sean $X^1, ..., X^m$ vectores aleatrios definidos sobre un mismo espacio probabilístico $(\Omega, \mathcal{A}, P)$, con dimensiones $n_1, ..., n_m$ respectivamente.

$$X^1, ..., X^m \text{ indep. } \Leftrightarrow F_X(x^1, ..., x^m) = F_{X^1}(x^1) \cdot F_{X^2}(x^2) \cdot ... \cdot F_{X^m}(x^m),$$

$$\forall x^1 \in \mathbb{R}^{n_1}, ..., x^m \in \mathbb{R}^{n_m}$$