Tema 2. Vectores Aleatorios
Parte I

Doble Grado en Ingeniería Informática y Matemáticas

Úrsula Torres Parejo

Departamento de Estadística e I.O. Universidad de Granada

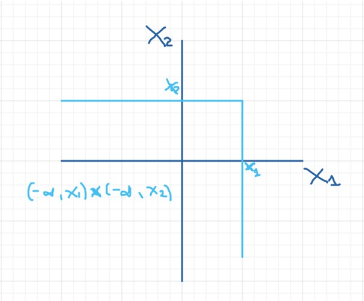
Contenido

- Sigma-álgebra de Borel sobre $\mathbb{R}^n$, $\mathcal{B}^n$
- Definición de vector aleatorio
- Distribución de probabilidad de un vector aleatorio
- Función de distribución de un vector aleatorio
- Propiedades de la función de distribución de un vector aleatorio
- Teorema de correspondencia
- Cálculo de probabilidades en intervalos bidimensionales
- Vectores aleatorios discretos
- Vectores aleatorios continuos

Sigma-álgebra de Borel sobre $\mathbb{R}^n$, $\mathcal{B}^n$

Es la mínima σ-álgebra que contiene todos los intervalos de $\mathbb{R}^n$.
En particular, se puede generar a partir de la colección de todos los intervalos de $\mathbb{R}^n$ de la forma:

$$(-\infty, x] = (-\infty, x_1] x (-\infty, x_2] x ... x (-\infty, x_n] \quad x = (x_1, ..., x_n) \in \mathbb{R}^n$$



Definición de vector aleatorio

Un vector aleatorio $X = (X_1, ..., X_n)$ sobre un espacio probabilístico base $(\Omega, \mathcal{A}, \mathcal{P})$ se define como una función:

$$X : (\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}^n, \mathcal{B}^n); \quad X^{-1}(B) \in \mathcal{A} \;\; \forall \;\; B \in \mathcal{B}^n$$

Caracterización I

$$\begin{aligned} X = (X_1, ..., X_n) \quad &: \quad (\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}^n, \mathcal{B}^n) \text{ es un vector aleatorio} \Leftrightarrow \\ X^{-1}((-\infty, x]) \quad &= \quad \{\omega \in \Omega / X(\omega) \leq x\} = \\ & \quad \{\omega \in \Omega / X_1(\omega) \leq x_1, ..., X_n(\omega) \leq x_n\} \in \mathcal{A} \\ & \quad \forall x = (x_1, ..., x_n) \in \mathbb{R}^n \end{aligned}$$

Caracterización II

$$\begin{aligned} X = (X_1, ..., X_n) \quad &: \quad (\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}^n, \mathcal{B}^n) \text{ es un vector aleatorio} \Leftrightarrow \\ X_i \quad &: \quad (\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}, \mathcal{B}) \text{ es una variable aleatoria} \end{aligned}$$

Distribución de Probabilidad de un vector aleatorio

Sea $X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}^n, \mathcal{B}^n)$ un vector aleatorio.
Se denomina distribución de probabilidad:

$$P_X : \mathcal{B}^n \to [0, 1]$$

satisfaciendo:

$$P_X(B) = P[\omega \in \Omega : X(\omega) \in B] = P[x \in B] \quad \forall B \in \mathcal{B}^n$$

X transforma el espacio de probabilidad original:

$$(\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}_X)$$

y el interés se centra en el estudio de este nuevo espacio, es decir, en el estudio de P_X.

Teorema: P_X es una medida de probabilidad sobre $(\mathbb{R}^n, \mathcal{B}^n)$

Función de Distribución de un vector aleatorio

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

$$F_X : \mathbb{R}^n \to [0, 1]$$

tal que

$$\begin{aligned} F_X(x) &= P_X[(-\infty, x]] = P[X \in (-\infty, x)] = P[X \leq x] \\ F_X(x_1, ..., x_n) &= P[X_1 \leq x_1, ..., X_n \leq x_n] \end{aligned}$$

Propiedades de la Función de Distribución de un vector aleatorio

1. Es monótona no decreciente en cada argumento

$$\begin{array}{rcl} \forall i = 1, ..., n & \forall & x_1, ..., x_{i-1}, x_{i+1}, ..., x_n \in \mathbb{R} \\ x_i < x_i^{'} & \Rightarrow & F_X(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n) \\ & \leq & F_X(x_1, ..., x_{i-1}, x_i^{'}, x_{i+1}, ..., x_n) \end{array}$$

2. Es continua a la derecha en cada argumento

$$\begin{array}{rcl} \forall i = 1, ..., n & \forall & x_1, ..., x_{i-1}, x_{i+1}, ..., x_n \in \mathbb{R} \\ & \lim_{x_i^{'} \to x_i^{+}} & F_X(x_1, ..., x_{i-1}, x_i^{'}, x_{i+1}, ..., x_n) = \\ & & F_X(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n) \end{array}$$

Propiedades de la Función de Distribución de un vector aleatorio

3.

$$
\begin{aligned}
&\forall i = 1, ..., n \quad \forall x_1, ..., x_n \in \mathbb{R} \\
&\exists \lim_{x_i \to -\infty} F_X(x_1, ..., x_i, ..., x_n) = F_X(x_1, ..., -\infty, ..., x_n) = 0 \\
&\exists \lim_{x_1, ..., x_n \to \infty} F_X(x_1, ..., x_i, ..., x_n) = F_X(\infty, ..., \infty, ..., \infty) = 1
\end{aligned}
$$

Propiedades de la Función de Distribución de un vector aleatorio

4.

$$
\begin{aligned}
&\forall(x_1, ..., x_n) \in \mathbb{R}^n, \quad \forall(\epsilon_1, ..., \epsilon_n) \in \mathbb{R}^n_+ \\
&F_X(x_1 + \epsilon_1, ..., x_n + \epsilon_n) - \\
&\sum_{i=1}^{n} F_X(x_1 + \epsilon_1, .., x_{i-1} + \epsilon_{i-1}, x_i, x_{i+1} + \epsilon_{i+1}, ..., x_n + \epsilon_n) + \\
&\sum_{i,j=1, i<j}^{n} F_X(X_1 + \epsilon_1, .., x_{i-1} + \epsilon_{i-1}, x_i, x_{i+1} + \epsilon_{i+1}, ..., x_{j-1} \\
&+\epsilon_{j-1}, x_j, x_{j+1} + \epsilon_{j+1}, ..., x_n + \epsilon_n) + ... + \\
&(-1)^n F_X(x_1, ..., x_n) \geq 0
\end{aligned}
$$

Propiedades de la Función de Distribución de un vector aleatorio

Caracterización

Las propiedades vistas caracterizan la función de distribución de los vectores aleatorios, i.e. toda función que las cumple es la función de distribución correspondiente a un vector aleatorio.

Teorema de correspondencia

Existe una correspondencia biunívoca entre las funciones de probabilidad P_X sobre $(\mathbb{R}^n, \mathcal{B}^n)$ y las funciones $F : \mathbb{R}^n \to [0,1]$ que satisfacen las propiedades vistas. Dicha correspondencia está determinada por la relación:

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

Más concretamente,

i) Si P_X es una medida de probabilidad sobre $(\mathbb{R}^n, \mathcal{B}^n)$, la función $F_X : \mathbb{R}^n \to [0,1]$ tal que $F(x) = P[(-\infty, x]]$ satisface las propiedades vistas.

ii) Recíprocamente, si $F_X : \mathbb{R}^n \to [0,1]$ satisface las propiedades $\Rightarrow$ Existe una única medida de probabilidad P_X sobre $(\mathbb{R}^n, \mathcal{B}^n)$ verificando:

$$P[(-\infty, x]] = F(x) \;\; \forall x \in \mathbb{R}^n$$

Corolario

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

b. Las propiedades vistas caracterizan a las funciones de distribución. Toda función que las cumpla es la función de distribución de un vector aleatorio.

Cálculo de probabilidades en intervalos bidimensionales

- $$P[a < X_1 \leq b, X_2 \in I] = P[X_1 \leq b, X_2 \in I] - P[X_1 \leq a, X_2 \in I]$$
- $$P[a < X_1 < b, X_2 \in I] = P[X_1 < b, X_2 \in I] - P[X_1 \leq a, X_2 \in I]$$
- $$\begin{aligned} P[X_1 \leq b, c < X_2 \leq d] &= P[X_1 \leq b, X_2 \leq d] \\ &\quad - P[X_1 \leq b, X_2 \leq c] \\ &= F(b, d) - F(b, c) \end{aligned}$$
- $$\begin{aligned} P[X_1 \leq b, c \leq X_2 < d] &= P[X_1 \leq b, X_2 < d] \\ &\quad - P[X_1 \leq b, X_2 < c] \\ &= F(b, d^-) - F(b, c^-) \end{aligned}$$

Vectores aleatorios discretos

Un vector aleatorio es discreto si su conjunto de valores es numerable.

$$\begin{aligned} \exists \quad E_X \subset \mathbb{R}^n & \text{ numerable, tal que} \\ P_X(E_X) = P[X \in E_X] & = 1 \end{aligned}$$

Función masa de probabilidad de un vector aleatorio discreto

$E_X \subset \mathbb{R}^n$ conjunto de valores del vector.

$$\begin{aligned} P_X &: E_X \to [0,1] \\ P_X &= P[X = x] = P[X_1 = x_1, ..., X_n = x_n] \;\; \forall x = (x_1, ..., x_n) \in E_X. \end{aligned}$$

satisfaciendo:

- $P[X = x] \geq 0 \;\; \forall x \in E_X$
- $\sum_{x \in E_X} P[X = x] = 1$

Caracterización. Toda colección numerable de números no negativos de suma 1 constituye la función masa de probabilidad de algún vector aleatorio n-dimensional de tipo discreto.

Función de distribución de un vector aleatorio discreto

$$F_X = P[X \leq x] = \sum_{x_i \in E_X, x_i \leq x} P[X = x_i] \;\; \forall x \in \mathbb{R}^n$$

Caracterización

$$X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}^n, \mathcal{B}^n, P_X) \text{ es discreto} \Leftrightarrow$$
$$X_i : (\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}, \mathcal{B}) \text{ es una componente discreta } \forall i = 1, ..., n$$

Vectores aleatorios continuos

Un vector aleatorio $X : (\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}^n, \mathcal{B}^n, P_X)$ es de tipo continuo si $\exists$ una función $f_X : \mathbb{R}^n \to \mathbb{R}$ no negativa e integrable, tal que:

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

i.e.

$$F_X(x_1, ..., x_n) = \int_{-\infty}^{x_n} \int_{-\infty}^{x_{n-1}} \int_{-\infty}^{x_1} f_X(t_1, ..., t_n) \qquad dt_1, ..., dt_n$$

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

Vectores aleatorios continuos

La función f_X recibe el nombre de función de densidad del vector aleatorio X y, como vemos en la definición, f_X determina $F(x)$ y, por tanto, P_X:

$$P_X(B) = P[x \in B] = \int_B f_X(t)dt$$

lo cual implica que si E es numerable $P[x \in E] = 0$.

Propiedades de la función de densidad

1. f_X es no negativa, integrable y

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_X(x_1, \dots, x_n)dx_1, \dots dx_n = 1$$

2.

$$lim_{x_i \to -\infty} f_X(x_1, \dots, x_n) = lim_{x_i \to \infty} f_X(x_1, \dots, x_n) = 0 \quad \forall i = 1, \dots, n$$

3. f_X es continua salvo en un conjunto de puntos con medida nula (es decir, el conjunto de discontinuidades de f_X es numerable) y F_X es derivable en los puntos de continuidad $x = (x_1, \dots, x_n)$ tal que:

$$\frac{\partial F_X(x_1, \dots, x_n)}{\partial x_1, \dots, \partial x_n} = f_X(x_1, \dots, x_n)$$

Propiedades de la función de densidad

4. f_X puede ser cambiada en puntos de medida nula sin afectar a la integral, o sea, a F_X.
5. Si E es un conjunto numerable $\Rightarrow P_X(E) = 0$

Caracterización de la función de densidad

$\rightarrow$ Si $f_X : \mathbb{R}^n \rightarrow \mathbb{R}$ es la función de densidad de un vector aleatorio $\Rightarrow$ es no negativa e integrable y su integral sobre $\mathbb{R}^n$ vale 1.

$\leftarrow$ Toda función $f_X : \mathbb{R}^n \rightarrow \mathbb{R}$ no negativa, integrable y tal que $\int_{\mathbb{R}^n} f_X(x)dx = 1$ es la función de densidad de algún vector aleatorio n-dimensional de tipo continuo.

Proposición: Si $X = (X_1, ..., X_n) : (\Omega, \mathcal{A}, \mathcal{P}) \rightarrow (\mathbb{R}^n, \mathcal{B}^n, P_X)$ es un vector aleatorio de tipo continuo
$\Rightarrow X_i : (\Omega, \mathcal{A}, \mathcal{P}) \rightarrow (\mathbb{R}, \mathcal{B}, P_{X_i})$ es de tipo continuo $\forall i = 1, ..., n$.

Tema 2. Vectores Aleatorios
Parte II

Doble Grado en Ingeniería Informática y Matemáticas

Úrsula Torres Parejo

Departamento de Estadística e I.O. Universidad de Granada

Contenido

- Distribuciones marginales
- Distribuciones condicionadas
- Funciones de vectores aleatorios. Cambio de variable
- Distribución del máximo y del mínimo
- Esperanza matemática
- Esperanza de la transformación $g : \mathbb{R}^n \to \mathbb{R}$
- Momentos centrados y no centrados
- Desigualdad de Cauchy-Schwarz
- Función generatriz de momentos

Distribuciones marginales

Al considerar un vector aleatorio como un conjunto de variables, a la distribución del vector se le llama **distribución conjunta** y a la distribución de cada variable **distribución marginal** de dicha componente.
Dichas distribuciones marginales pueden obtenerse a partir de la distribución conjunta:

Si $X = (X_1, ..., X_n)$ es un vector con función de distribución $F_X \Rightarrow$

$$\forall i = 1, ..., n \quad F_{X_i}(x_i) = F_X(+\infty, ..., +\infty, x_i, +\infty, ..., +\infty)$$

Análogamente:

$$F_{x_{i1},...,x_{ik}}(x_{i1}, ..., x_{ik}) = F_X(+\infty, ..., +\infty, x_{i1}, +\infty, ..., +\infty, x_{ik}, +\infty, ..., +\infty)$$

Distribuciones marginales. Caso discreto

Sea $X = (X_1, ..., X_n)$ un vector aleatorio discreto con $P[X \in E_X] = 1$ y función masa de probabilidad $P[X = x]$ $\forall x \in E_X$. Si X_i es una componente arbitraria y por tanto discreta con valores en E_{X_i}, entonces su función masa de probabilidad puede obtenerse a partir de la conjunta:

$$P[X_i = x_i] = \sum_{\substack{x_1,...,x_{i-1},x_{i+1},...,x_n,\\ x \in E_X}} P[X = x]$$

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

La función masa marginal de cualquier subvector se calcularía de igual forma.

Distribuciones marginales. Caso continuo

Sea $X = (X_1, ..., X_n)$ un vector aleatorio continuo con función de densidad f_X, entonces cada componente X_i es de tipo continuo y su función de distribución es:

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

con

$$f_{X_i}(t_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(t_1, ..., t_n) dt_1 \cdots dt_{i-1} \cdot dt_{i+1} \cdots dt_n \qquad \forall t_i \in \mathbb{R}$$

La función de densidad marginal de cualquier subvector se calcularía de igual forma.

Distribuciones condicionadas. Caso discreto

Distribución condicionada al valor de una variable
Sea $X = (X_1, ..., X_n)$ un vector aleatorio discreto y X_i una componente arbitraria y $x_i^* \in \mathbb{R}/P[X_i = x_i^*] > 0 \Rightarrow$ Se define la distribución condicionada de $(X_1, ..., X_{i-1}, X_{i+1}, ..., X_n)$ a $(X_i = x_i^*)$ como la determinada por la función masa de probabilidad:

$$P[X_1 = x_1, ..., X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}, ..., X_n = x_n / X_i = x_i^*] =$$

$$= \frac{P[X_1 = x_1, ..., X_{i-1} = x_{i-1}, X_i = x_i^*, X_{i+1} = x_{i+1}, ..., X_n = x_n]}{P[X_i = x_i^*]}$$

$$\forall (x_1, ..., x_{i-1}, x_{i+1}, ..., x_n)/(x_1, ..., x_{i-1}, x_i^*, x_{i+1}, ..., x_n) \in E_X$$

Distribuciones condicionadas. Caso discreto

Distribución condicionada a valores de varias variables
Sea $X = (X_1, ..., X_n)$ un vector aleatorio discreto y sea $X_{i1}, ..., x_{ik}$ un subvector arbitrario y
$(x^*_{i1}, ..., x^*_{ik}) \in \mathbb{R}^k / P[X_{i1} = x^*_{i1}, ..., X_{ik} = x^*_{ik}] > 0 \Rightarrow$ Se define la distribución condicionada de
$(X_1, ..., X_{i1-1}, X_{i1+1}, ..., X_{ik-1}, X_{ik+1}, ..., X_n)$ a
$(X_{i1} = x^*_{i1}, ..., X_{ik} = x^*_{ik})$ como la determinada por la función masa de probabilidad:

$$P[X_1 = x_1, ..., X_{i1-1} = x_{i1-1}, X_{i1+1} = x_{i1+1}, ..., X_{ik-1} = x_{ik-1},$$
$$X_{ik+1} = x_{ik+1}, ..., X_n = x_n / X_{i1} = x^*_{i1}, ..., X_{ik} = x^*_{ik}] =$$

$$= \frac{P[X_1 = x_1, ..., X_{i1} = x^*_{i1}, ..., X_{ik} = x^*_{ik}, ..., X_n = x_n]}{P[X_{i1} = x^*_{i1}, ..., X_{ik} = x^*_{ik}]}$$

$$\forall (x_1, ..., x_{i1-1}, x_{i1+1}, ..., x_{ik-1}, x_{ik+1}, ..., x_n) / (x_1, ..., x^*_{i1}, , x^*_{ik}, ..., x_n) \in E_X$$

Distribuciones condicionadas. Caso continuo

Distribución condicionada al valor de una variable
Sea $X = (X_1, ..., X_n)$ un vector aleatorio continuo con función de densidad f_X. Sea X_i una componente arbitraria y $x_i^* \in \mathbb{R} / f_{X_i}(x_i^*) > 0 \Rightarrow$ Se define la distribución condicionada de $(X_1, ..., X_{i-1}, X_{i+1}, ..., X_n)$ a $(X_i = x_i^*)$ como la determinada por la función de densidad:

$$f_{X_1,..,X_{i-1},X_{i+1},...,X_n/X_i=x_i^*}(x_1, ..., x_{i-1}, x_{i+1}, ..., x_n/x_i^*) =$$

$$= \frac{f_X(x_1, ..., x_i^*, ..., x_n)}{f_{X_i}(x_i^*)}$$

Distribuciones condicionadas. Caso continuo

Distribución condicionada a valorer de varias variables

Sea $X = (X_1, ..., X_n)$ un vector aleatorio continuo con función de densidad f_X. Sea $(X_{i1}, ..., X_{ik})$ un subvector arbitrario y $(x_{i1}^*, ..., x_{ik}^*) \in \mathbb{R}^k / f_{X_{i1},...X_{ik}}(x_{i1}^*, ..., x_{ik}^*) > 0 \Rightarrow$ Se define la distribución condicionada de $(X_1, ..., X_{i1-1}, X_{i1+1}, ..., X_{ik-1}, X_{ik+1}, ..., X_n)$ a $(X_{i1} = x_{i1}^*, ..., X_{ik} = x_{ik}^*)$ como la determinada por la función de densidad:

$$f_{X_1,..,X_{i1-1},X_{i1+1},...,X_{ik-1},X_{ik+1},...,X_n/X_{i1}=x_{i1}^*,...,X_{ik}=x_{ik}^*}(x_1, ..., x_{i1-1}, x_{i1+1}, ..., x_{ik-1}, x_{ik+1}, ..., x_n / x_{i1}^*, ..., x_{ik}^*) =$$

$$= \frac{f_X(x_1, ..., x_{i1}^*, ..., x_{ik}^*, ..., x_n)}{f_{X_{i1},...,X_{ik}}(x_{i1}^*, ..., x_{ik}^*)}$$

Funciones de vectores aleatorios. Cambio de variable

Si X es un vector aleatorio n-dimensional y
$g : (\mathbb{R}^n, \mathcal{B}^n) \to (\mathbb{R}^m, \mathcal{B}^m)$ es una transformación medible,
$Y = g(X)$ es un vector aleatorio m-dimensional cuya distribución puede hallarse a partir de la de X.

Fórmula General del Cambio de Variable

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

Cambio de variable: Discreto a discreto

Si X es un vector aleatorio n-dimensional discreto con valores en $E_X \subset \mathbb{R}^n$ y $g : (\mathbb{R}^n, \mathcal{B}^n) \to (\mathbb{R}^m, \mathcal{B}^m)$ una transformación medible, $Y = g(X)$ es un vector aleatorio m-dimensional con valores en $g(E_X)$ cuya función masa de probabilidad puede hallarse a partir de la de X como:

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

Cambio de variable: Continuo a discreto

Si X es un vector aleatorio n-dimensional continuo con función de densidad $f_X(x)$ y 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 cuya función masa de probabilidad se puede obtener a partir de la función de densidad de X como:

$$P[Y = y] = P[X \in g^{-1}(y)] = \int_{g^{-1}(y)} f_X(x)dx, \quad y \in g(E_X), x \in \mathbb{R}^n$$

Cambio de variable: Continuo a continuo

Si X es un vector aleatorio n-dimensional continuo con función de densidad $f_X(x)$ y 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 cuya función de densidad se puede obtener derivando su función de distribución, que puede hallarse como:

$$
\begin{aligned}
F_Y(y) &= P[Y \leq y] = P[g(X) \leq y] = P[X \in g^{-1}((-\infty, y])] = \\
&= \int_{g^{-1}((-\infty,y])} f_X(x)dx, \quad x \in \mathbb{R}^n \quad y \in \mathbb{R}^m
\end{aligned}
$$

Teorema: Cambio de variable de continuo a continuo

Sea $X : (\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}_X)$ un vector aleatorio con función de densidad $f_X(x)$ y $g : (\mathbb{R}^n, \mathcal{B}^n) \to (\mathbb{R}^n, \mathcal{B}^n)$ una función medible tal que:

1. $g(g_1, ..., g_n)$ admite inversa $g^{-1}(g_1^*, ..., g_n^*)$
2. La inversa es derivable en todos los argumentos:
   $\forall \; i, j = 1, ..., n \; \exists \; \frac{\partial g_i^*(y_1,...,y_n)}{\partial y_j}$
3. El jacobiano de la inversa es no nulo:
   $J = \left| \left(\left(\frac{\partial g_i^*(y_1,...,y_n)}{\partial y_j} \right) \right)_{i,j} \right| \neq 0$

Bajo estas condiciones, el vector aleatorio $Y = g(X)$ es de tipo continuo y su función de densidad es:

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

Teorema: Cambio de variable de continuo a continuo

1. **Nota 1:** Si el vector transformado es de dimensión menor que el original, se condieran variables auxiliares y luego se calcula la marginal correspondiente.
2. **Nota 2:** Si g no admite inversa, pero cada $y \in \mathbb{R}^n$ tiene un número finito de antiimágenes, $g_1^{-1}(y), ..., g_k^{-1}(y)$, cada una satisfaciendo las hipótesis del teorema, entonces:

$$f_Y(y) = \sum_{i=1}^{k} |J_i| f_X(g_i^{-1}(y))$$

Distribución del máximo y del mínimo

Sea $X = (X_1, ..., X_n)$ un vector aleatorio n-dimensional definido sobre $(\Omega, \mathcal{A}, \mathcal{P})$.

1. Se define $Max = max(X_1, ..., X_n)$ como una variable aleatoria tal que:

$$Max(\omega) = Max\{X_1(\omega), X_2(\omega), ..., X_n(\omega)\} \quad \forall \omega \in \Omega$$

2. Se define $Min = min(X_1, ..., X_n)$ como una variable aleatoria tal que:

$$Min(\omega) = Min\{X_1(\omega), X_2(\omega), ..., X_n(\omega)\} \quad \forall \omega \in \Omega$$

Dado que estas transformaciones no satisfacen, en general, el teorema de cambio de variable, para obtener su distribución usamos la fórmula general.

Distribución del máximo y del mínimo

Sea $X = (X_1, ..., X_n)$ un vector aleatorio con función de distribución F_X, entonces:

Distribución del máximo: $Max = max(X_1, ..., X_n)$

$$\begin{aligned} F_{Max}(x) &= P[Max \leq x] = P[X_1 \leq x, ..., X_n \leq x] \\ &= F_X(x, ..., x) \;\; \forall x \in \mathbb{R} \end{aligned}$$

Distribución del mínimo: $Min = min(X_1, ..., X_n)$

$$\begin{aligned} F_{Min}(x) &= P[Min \leq x] = 1 - P[Min > x] = \\ &= 1 - P[X_1 > x, ..., X_n > x] \;\; \forall x \in \mathbb{R} \end{aligned}$$

Y a partir de ahí se obtiene la función masa o densidad de probabilidad, según corresponda.

Distribución del máximo y del mínimo

Distribución conjunta del máximo y del mínimo:

$$F_{Max,Min}(x,y) = P[Max \leq x, Min \leq y] =$$

$$= \begin{cases} P[Max \leq x] = F_X(x,...,x) & x \leq y \\ \\ P[Max \leq x] - P[Max \leq x, Min > y] = \\ = F_X(x,...,x) - P[y < X_1 \leq x, ..., y < X_n \leq x] & x > y \end{cases}$$

Esperanza matemática

Sea $X = (X_1, ..., X_n)$ un vector aleatorio n-dimensional. La esperanza matemática $E[X]$ de X, si existe, se define como:

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

también llamado vector de medias.

Consecuencia.

La esperanza matemática de un vector aleatorio $X = (X_1, ..., X_n)$ $\exists \Leftrightarrow \exists$ las esperanzas matemáticas de sus componentes aleatorias.
Equivalentemente, un vector aleatorio tiene media finita $\Leftrightarrow$ sus distribuciones marginales tienen media finita.

Esperanza matemática

Nota.

La esperanza matemática de cada componente X_i del vector, puede calcularse a partir de la distribución conjunta o a partir de la marginal.

Caso discreto.

$$E[X_i] = \sum_{x_i} x_i P[X_i = x_i] = \sum_{x_1,\ldots,x_n} x_i P[X_1 = x_1, \ldots, X_n = x_n]$$

Caso continuo

$$E[X_i] = \int_{\mathbb{R}} x_i f_{X_i}(x_i) dx_i = \int_{\mathbb{R}^n} x_i f_X(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

Esperanza matemática

Propiedades.

1. Linealidad.

$$\forall (a_1, ..., a_n), (b_1, ..., b_n) \in \mathbb{R}^n.$$
$$Si \ \ \exists\ E[X_i] \Rightarrow \exists\ E[a_i X_i + b_i] \ \ i = 1, ..., n \Rightarrow$$
$$\Rightarrow \exists\ E\left[\sum_{i=1}^{n} a_i X_i + b_i\right] = \sum_{i=1}^{n} a_i E[X_i] + b_i$$

2. Conservación del orden.

   Si X_1 y X_2 son variables aleatorias unidimensionales tales que $X_1 \leq X_2 \Rightarrow E[X_1] \leq E[X_2]$, siempre que $\exists\ E[X_1], E[X_2]$.

Esperanza de la transformación $g : \mathbb{R}^n \to \mathbb{R}$

Sea $g : \mathbb{R}^n \to \mathbb{R}$ una función medible $\Rightarrow$ $Y = g(X)$ es una variable aleatoria. Entonces:

- Si X es de tipo discreto $\Rightarrow \exists\ E[g(X)] \Leftrightarrow$

$$\sum_{x_1,\dots,x_n} |g(X_1,\dots,X_n)| P[X_1 = x_1,\dots,X_n = x_n] < \infty$$

  Y en caso de existir:

$$E[g(X_1,..,X_n)] = \sum_{x_1,\dots,x_n} g(X_1,\dots,X_n) P[X_1 = x_1,\dots,X_n = x_n]$$

- Si X es de tipo continuo $\Rightarrow \exists E[g(X)] \Leftrightarrow$

$$\int_{\mathbb{R}^n} |g(X_1,\dots,X_n)| f_X(x_1,\dots,x_n) dx_1 \cdots dx_n < \infty$$

  Y en caso de existir:

$$E[g(X_1,..,X_n)] = \int_{\mathbb{R}^n} g(X_1,\dots,X_n) f_X(x_1,\dots,x_n) dx_1 \cdots dx_n$$

Momentos centrados y no centrados

Momentos no centrados

Sea $X = (X_1, ..., X_n)$ un vector aleatorio y $k_1, ..., k_n$ enteros no negativos. Se define el momento no centrado de orden $(k_1, .., k_n)$ como $E[X_1^{k_1} \cdots X_n^{k_n}]$, siempre que exista.

Caso discreto

$$m_{k_1,...,k_n} = E[X_1^{k_1} \cdots X_n^{k_n}] = \sum_{x_1,...,x_n} X_1^{k_1} \cdots X_n^{k_n} P[X_1 = x_1, ..., X_n = x_n]$$

Caso continuo

$$m_{k_1,...,k_n} = E[X_1^{k_1} \cdots X_n^{k_n}] = \int_{\mathbb{R}^n} X_1^{k_1} \cdots X_n^{k_n} f_X(x_1, ..., x_n) dx_1 \cdots dx_n$$

Nota. Si hacemos $k_1 = ... = k_{i-1} = k_{i+1} = ... = k_n = 0$, se obtienen los momentos no centrados de la variable $X_i \;\; \forall i = 1, ..., n$ o momentos marginales.

Momentos centrados y no centrados

Momentos centrados

Sea $X = (X_1, ..., X_n)$ un vector aleatorio y $k_1, ..., k_n$ enteros no negativos. Si $\exists\ E[X_i]\ \ \forall i = 1, ..., n$ se define el momento centrado de orden $(k_1, .., k_n)$ como $E[(X_1 - E[X_1])^{k_1} \cdots (X_n - E[X_n])^{k_n}]$, siempre que exista.

$$\mu_{k_1,...,k_n} = E[(X_1 - E[X_1])^{k_1} \cdots (X_n - E[X_n])^{k_n}]$$

Casos particulares

(X_1, X_2) vector aleatorio bidimensional.
$m_{10} = E[X_1]$, $m_{01} = E[X_2]$, $m_{20} = E[X_1^2]$, $m_{02} = E[X_2^2]$,
$m_{11} = E[X_1 X_2]$,
$\mu_{10} = E[X_1 - E[X_1]] = 0$, $\mu_{01} = 0$,
$\mu_{20} = Var[X_1] = E[(X_1 - E[X_1])^2] = E[X_1^2] - E[X_1]^2$,
$\mu_{02} = Var[X_2]$,
$\mu_{11} = Cov[X_1, X_2] = E[(X_1 - E[X_1])(X_2 - E[X_2])] =$
$E[X_1 X_2] - E[X_1]E[X_2]$

Momentos centrados y no centrados

Matriz de covarianzas

Para un vector de dimensión n se define la matriz de covarianzas:

$$Cov_X = ((Cov(X_i, X_j)))_{i,j=1,\ldots,n}$$

Ejemplo

(X_1, X_2) vector aleatorio bidimensional:

$$Cov_X = \begin{pmatrix} Var[X_1] & Cov[X_1, X_2] \\ Cov[X_2, X_1] & Var[X_2] \end{pmatrix}$$

Desigualdad de Cauchy-Schwarz

Sean X, Y variables aleatorias tales que $\exists\ E[X^2], E[Y^2]$, entonces:

- $E[XY]^2 \leq E[X^2]E[Y^2]$
- Si X o Y es degenerada en 0 o las dos los son, se da la igualdad.
- Si X e Y son no degeneradas, se da la igualdad si $\exists\ a, b \in \mathbb{R}$ no nulas tales que $P[aX + bY = 0] = 1$.

Corolario.

Sean X, Y variables aleatorias tales que $\exists\ E[X^2], E[Y^2]$, entonces:

- $Cov[X, Y]^2 \leq Var[X]Var[Y]$
- Si X o Y es degenerada en 0 o las dos los son, se da la igualdad.
- Si X e Y son no degeneradas, se da la igualdad si $\exists a \neq 0, b \neq 0 \quad y \quad c \in \mathbb{R}$ tales que $P[aX + bY = c] = 1$.

Función generatriz de momentos

Dado un vector $X = (X_1, ..., X_n)$, su función generatriz de momentos se define como:

$$M_X(t_1, ..., t_n) = E[e^{t_1 X_1 + t_2 X_2 + ... + t_n X_n}]$$
$$donde\ (t_i) \in (-a_i, b_i)\ a_i, b_i \in \mathbb{R}^+, i = 1, ..., n$$

Teorema de Unicidad

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

Función generatriz de momentos

Relación entre la función generatriz de momentos y los momentos

Si $\exists \; M_{X_1,\ldots,X_n} \Rightarrow \exists$ todos los momentos no centrados y se pueden calcular como:

$$m_{k_1,\ldots,k_n} = \frac{\partial^{k_1+\ldots+k_n} M_X(t_1,\ldots,t_n)}{\partial t_1^{K_1},\ldots,\partial t_n^{k_n}}\Big|_{t_1=0,\ldots,t_n=0}$$

Funciones generatrices de momentos marginales

Si $X=(X_1,\ldots,X_n)$ tiene función generatriz de momentos $M_X(t_1,\ldots,t_n) \Rightarrow \exists$ la función generatriz de momentos de cada subvector y, por tanto, de cada componente y puede calcularse a partir de la conjunta como:
$M_{(X_{i1},\ldots,X_{ik})}(t_{i1},\ldots,t_{ik}) = M_{X_1,\ldots,X_n}(0,\ldots,0,t_{i1},\ldots,t_{ik},0,\ldots,0)$, para $t_{i1},\ldots t_{ik}$ en los intervalos correspondientes.

Tema 3. Independencia de Variables Aleatorias

Doble Grado en Ingeniería Informática y Matemáticas

Úrsula Torres Parejo

Departamento de Estadística e I.O. Universidad de Granada

UNIVERSIDAD DE GRANADA

Contenido.

- Definición de independencia
- Caracterizaciones de independencia
- Propiedades de independencia
- Teorema de Multiplicación de Esperanzas
- Distribuciones reproductivas
- Independencia para familias de variables aleatorias
- Independencia de vectores aleatorios

Definición.

Sean $X_1, ..., X_n$ variables aleatorias definidas sobre un mismo espacio de probabilidad y $X = (X_1, ..., X_n)$ un vector aleatorio n-dimensional. Se dice que las variables $X_i, \quad i = 1, ..., n$ son independientes $\Leftrightarrow$:

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

Caracterizaciones.

Caracterización para variables aleatorias discretas
Sea $X = (X_1, ..., X_n)$ un vector aleatorio discreto $\Rightarrow X_1, ..., X_n$ son independientes $\Leftrightarrow$:

1. $P[X_1 = x_1, ..., X_n = x_n] = P[X_1 = x_1] \cdots P[X_n = x_n]$
2. $P[X_1 = x_1, ..., X_n = x_n] = h_1(x_1) \cdots h_n(x_n)$

$\forall x_1, ..., x_n \in \mathbb{R}$, con $h_i \quad i = 1, ..., n$ funciones arbitrarias.

Caracterización para variables aleatorias continuas
Sea $X = (X_1, ..., X_n)$ un vector aleatorio continuo $\Rightarrow X_1, ..., X_n$ son independientes $\Leftrightarrow$:

1. $f_X(x_1, ..., x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$
2. $f_X(x_1, ..., x_n) = h_1(x_1) \cdots h_n(x_n)$

$\forall x_1, ..., x_n \in \mathbb{R}$, con $h_i \quad i = 1, ..., n$ funciones arbitrarias.

Caracterizaciones.

Caracterización de independencia por conjuntos de Borel
Sean $X_1, ..., X_n$ variables aleatorias unidimensionales definidas sobre un mismo espacio de probabilidad $\Rightarrow X_1, ..., X_n$ son independientes $\Leftrightarrow$:

$$P[X_1 \in \mathcal{B}_1, ..., X_n \in \mathcal{B}_n] = P[X_1 \in \mathcal{B}_1] \cdots P[X_n \in \mathcal{B}_n] \;\; \forall\, \mathcal{B}_1, ..., \mathcal{B}_n \in \mathcal{B}.$$

Propiedades de independencia.

1. Una variable degenerada en un valor c, i.e. $P[X = c] = 1$ es independiente de cualquier otra.
2. Si $X_1, ..., X_n$ son variables aleatorias independientes $\Rightarrow$ Cualquier subconjunto de ellas está formado por variables aleatorias independientes.
3. Si $X_1, ..., X_n$ son variables aleatorias independientes $\Rightarrow$ Todas las distribuciones condicionadas coinciden con las marginales correspondientes.
4. Si $X_1, ..., X_n$ son variables aleatorias independientes y $\exists\ M_{X_i}(t)\ \forall t \in (-a_i, b_i)\ \forall\ a_i, b_i \in \mathbb{R}^+, i = 1, ..., n \Rightarrow$ $\exists\ M_{X_1,...,X_n}(t_1, ..., t_n)$ con $t_i \in (-a_i, b_i)$ $i = 1, ..., n$ y se tiene que:
$$M_{X_1,...,X_n}(t_1, ..., t_n) = M_{X_1}(t_1) \cdots M_{X_n}(t_n)$$
5. Si $X_1, ..., X_n$ son variables aleatorias independientes y $g_i : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ son funciones medibles $\forall\ i = 1, ..., n \Rightarrow$ Las variables aleatorias $g_1(x_1), ..., g_n(x_n)$ son independientes.

Teorema de Multiplicación de Esperanzas.

i) Si $X_1, \ldots, X_n$ son variables aleatorias independientes y $\exists\ E[X_i]\ \forall i = 1, \ldots, n \Rightarrow$

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

ii) Si $X_1, \ldots, X_n$ son variables aleatorias independientes y $g_1, \ldots, g_n : \mathbb{R} \to \mathbb{R}$ son funciones medibles $\Rightarrow g_1(x_1), \ldots, g_n(x_n)$ son variables aleatorias independientes y

$$E[g_1(x_1) \cdots g_n(x_n)] = E[g_1(x_1)] \cdots E[g_n(x_n)]$$

iii) Si X e Y son v.a. independientes $\Rightarrow Cov(X, Y) = 0$

iv) Si $X_1, \ldots, X_n$ son variables aleatorias independientes, tales que $\exists\ E[X_i^2]\ \forall i = 1, \ldots, n \Rightarrow$

$$Var[\sum_{i=1}^{n} a_i X_i] = \sum_{i=1}^{n} a_i^2 Var[X_i], \quad a_1, \ldots, a_n \in \mathbb{R}$$

Distribuciones reproductivas.

Son aquellas distribuciones que permanecen estables frente a la suma de variables aleatorias independientes. Específicamente:

- **Binomial**

$$X_i \sim B(n_i, p) \; indep., \;\; i = 1, ..., n \Rightarrow \sum_{i=1}^{n} X_i \sim B\left(\sum_{i=1}^{n} n_i, p\right)$$

- **Poisson**

$$X_i \sim P(\lambda_i) \; indep., \;\; i = 1, ..., n \Rightarrow \sum_{i=1}^{n} X_i \sim P\left(\sum_{i=1}^{n} \lambda_i\right)$$

- **Binomial Negativa**

$$X_i \sim BN(r_i, p) \; indep., \;\; i = 1, ..., n \Rightarrow \sum_{i=1}^{n} X_i \sim BN\left(\sum_{i=1}^{n} r_i, p\right)$$

Distribuciones reproductivas.

- **Geométrica**

$$X_i \sim G(p) \ indep., \ \ i = 1, ..., n \Rightarrow \sum_{i=1}^{n} X_i \sim BN(n, p)$$

- **Normal**

$$X_i \sim N(\mu_i, \sigma_i^2) \ indep., \ \ i = 1, ..., n \Rightarrow \sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

- **Gamma**

$$X_i \sim \Gamma(u_i, \lambda) \ indep., \ \ i = 1, ..., n \Rightarrow \sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} u_i, \lambda\right)$$

Distribuciones reproductivas.

- **Erlang**

$$X_i \sim \mathcal{E}(k_i, \lambda) \ indep., \ \ i = 1, ..., n \Rightarrow \sum_{i=1}^{n} X_i \sim \mathcal{E}\left(\sum_{i=1}^{n} k_i, \lambda\right)$$

- **Exponencial**

$$X_i \sim exp(\lambda) \ indep., \ \ i = 1, ..., n \Rightarrow \sum_{i=1}^{n} X_i \sim \mathcal{E}(n, \lambda)$$

Independencia para familias de variables aleatorias.

Una familia arbitraria es una colección de variables aleatorias que no tiene por qué ser finita, sino que puede ser infinita numerable.

- **Independencia mutua** Las variables aleatorias $\{X_t,\ t \in T\}$ donde T es un conjunto de índices infinito numerable, son mutuamente independientes si $\forall k \in \mathbb{N}$ las variables aleatorias $X_{t_1}, ..., X_{t_k}$ son independientes, $t_1, ..., t_k \in T$.
- **Independencia dos a dos** Las variables aleatorias $\{X_t,\ t \in T\}$ donde T es un conjunto de índices infinito numerable, son independientes dos a dos si $\forall t_i, t_j \in T$ X_{t_i} y X_{t_j} son independientes, con $t_i \neq t_j, i \neq j$.

$$\textbf{Independencia mutua} \quad \Rightarrow \quad \textbf{Independencia dos a dos}$$

$$\not\Leftarrow$$

Definición.

La definición de independencia de variables aleatorias se extiende de forma inmediata a vectores aleatorios.

Si $X^1, ..., X^m$ son vectores aleatorios definidos sobre un espacio de probabilidad, con dimensiones $n_1, ..., n_m$, respectivamente, $X^1, ..., X^m$ son indendientes $\Leftrightarrow$

$$F(X^1, ..., X^m) = F_{X^1}(x^1) \cdots F_{X^m}(x^m) \quad \forall x^1 \in \mathbb{R}^{n_1}, ..., x^m \in \mathbb{R}^{n_m}$$

Tema 4. Parte I. Esperanza Condicionada: Regresión y Correlación

Doble Grado en Ingeniería Informática y Matemáticas

Úrsula Torres Parejo

Departamento de Estadística e I.O. Universidad de Granada

Contenido.

- Esperanza condicionada de una variable aleatoria
- Esperanza condicionada de una función medible
- Propiedades y teoremas de la esperanza condicionada
- Momentos condicionados bidimensionales
- Teorema de descomposición de la varianza

Esperanza condicionada.

Sean X e Y variables aleatorias definidas sobre un mismo espacio de probabilidad, tales que $\exists\ E[X]$. Se define la esperanza de X dada Y (o esperanza de X condicionada a Y) y se denota por $E[X|Y]$ como la variable aleatoria que toma el valor $E[X|Y=y_0]$ cuando $Y = y_0$ y se calcula como sigue:

- **Caso discreto**
  $E[X|Y=y_0] = \sum_x x \cdot P[X=x|Y=y_0]$ con $P[Y=y_0] > 0$
- **Caso continuo**
  $E[X|Y=y_0] = \int_{-\infty}^{\infty} x \cdot f_{X|Y=y_0}(x)dx$ con $f_Y(y) > 0$

Notas:

- $E[X|Y]$ es una función de Y
- $E[X|Y=y_0]$ no es más que la media de X considerando como distribución la condicionada de X dado $Y = y_0$
- De forma similar se puede definir la esperanza de Y dada X, supuesta la existencia de E[Y].

Esperanza condicionada de una función medible.

Sean X e Y variables aleatorias definidas sobre un mismo espacio de probabilidad y sea $g : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ una función medible tal que $\exists\ E[g(X)]$. Se define la esperanza condicionada de g(X) dado Y y se denota por $E[g(X)|Y]$ como la variable aleatoria que toma el valor $E[g(X)|Y = y_0]$ cuando $Y = y_0$ y se calcula como sigue:

- **Caso discreto**
  $E[g(X)|Y = y_0] = \sum_x g(x) \cdot P[X = x|Y = y_0]; P[Y = y_0] > 0$
- **Caso continuo**
  $E[g(X)|Y = y_0] = \int_{-\infty}^{\infty} g(x) \cdot f_{X|Y=y_0}(x)dx \quad \text{con } f_Y(y) > 0$

Nota: De forma similar se puede definir $E[g(Y)|X]$

Propiedades de la esperanza condicionada.

1. $E[c/Y] = c$
2. **Linealidad:** Sean X e Y variables aleatorias definidas en el mismo espacio de probabilidad tales que $\exists\ E[X]$ y $a, b \in \mathbb{R} \Rightarrow \exists\ E[(aX + b)|Y] = aE[X|Y] + b$
3. Sean $X_1, ..., X_n$ variables aleatorias tales que $\exists\ E[X_i], i = 1, 2, ..., n \Rightarrow \forall a_1, ..., a_n \in \mathbb{R}$ $\exists\ E[(a_1X_1 + ... + a_nX_n)|Y] = a_1E[X_1|Y] + ... + a_nE[X_n|Y]$
4. Si $X \geq 0$ y $\exists\ E[X] \Rightarrow E[X|Y] \geq 0$ y $E[X|Y] = 0 \Leftrightarrow P[X = 0] = 1$
5. **Conservación del orden**: Si X_1 y X_2 son variables aleatorias tales que $\exists\ E[X_1], E[X_2]$ y $X_1 \leq X_2 \Rightarrow E[X_1|Y] \leq E[X_2|Y]$

Teoremas.

Teorema 1
Si X e Y son variables aleatorias independientes y $\exists\ E[g(X)]$ siendo g una función medible $\Rightarrow$

$$E[g(X)|Y] = E[g(X)]$$

En particular, $E[X|Y] = E[X]$

Teorema 2
Si X e Y son variables aleatorias definidas en el mismo espacio de probabilidad tales que $\exists\ E[g(X)]$ siendo g una función medible $\Rightarrow$

$$\exists\ E[E[g(X)|Y]] = E[g(X)]$$

En particular, $E[E[X|Y]] = E[X]$

Momentos condicionados bidimensionales.

Sean X e Y variables aleatorias definidas sobre el mismo espacio de probabilidad:

- Si $\exists\ E[X^n],\ n \in \mathbb{N} \Rightarrow$ a $E[X^n|Y]$ se le llama momento condicionado no centrado de orden n de X dada Y.
- Si $\exists\ E[X^n],\ n \in \mathbb{N} \Rightarrow$ a $E[(X - E[X|Y])^n|Y]$ se le llama momento condicionado centrado de orden n de X dada Y.

Ejemplo de cálculo:

$E[X^n|Y = y_0] = \sum_x X^n \cdot P[X = x|Y = y_0]$

$E[X^n|Y = y_0] = \int_{-\infty}^{\infty} X^n \cdot f_{X|Y=y_0}(x)dx$

Casos particulares:

- Momento condicionado no centrado de orden uno: Esperanza condicionada
- Momento condicionado centrado de orden dos: Varianza condicionada:
  $Var[X|Y] = E[(X - E[X|Y])^2|Y] = E[X^2|Y] - E[X|Y]^2$

Teorema de descomposición de la varianza.

Si $\exists\ E[X^2] \Rightarrow \exists\ Var[E[X|Y]]$ y $E[Var[X|Y]]$ y además:

$$Var[X] = Var[E[X|Y]] + E[Var[X|Y]]$$

Corolario

En las condiciones del teorema:

- $Var[X] \geq Var[E[X|Y]]$
- $Var[X] \geq E[Var[X|Y]]$

Tema 5. Algunos Modelos de Distribución de Probabilidad Multidimensionales

Úrsula Torres Parejo

Departamento de Estadística e I.O. Universidad de Granada

Contenido.

- Distribución Multinomial
  - Definición
  - Función generatriz de momentos. Distribuciones marginales y condicionadas
  - Regresión y correlación. Caso bidimensional
  - Reproductividad
- Distribución Normal Bivariante
  - Definición
  - Distribuciones marginales y condicionadas
  - Regresión y correlación. Caso bidimensional
  - Normalidad de las combinaciones lineales de las componentes
  - Función generatriz de momentos

Distribución Multinomial.

Es una generalización de la distribución binomial cuando el experimento aleatorio considerado no tiene sólo dos posibles resultados (éxito o fracaso), sino 3 o más.
Indica el número de veces que aparecen k sucesos excluyentes y no necesariamente exhaustivos con probabilidades $p_1, ..., p_k$ en repeticiones independientes de un experimento.

$$\text{Si } X = (X_1, ..., X_k) \sim M_k(n; p_1, ..., p_k) \Rightarrow$$

$$P[X_1 = x_1, ..., X_k = x_k] =$$

$$= \frac{n!}{x_1! \cdots x_k!(n - \sum_{i=1}^k x_i)!} p_1^{x_1} \cdots p_k^{x_k} \left(1 - \sum_{i=1}^k p_i\right)^{n - \sum_{i=1}^k x_i}$$

$$n \in \mathbb{N}, 0 < p_i < 1, \sum_{i=1}^k p_i \leq 1, \sum_{i=1}^k x_i \leq n,\ x_i \in \{0, ..., n\}$$

$$i = 1, ..., k$$

F.G.M. Distribución marginales y condicionadas.

Función generatriz de momentos

$$M_X(t_1, ..., t_k) = \left(p_1 e^{t_1} + ... + p_k e^{t_k} + \left(1 - \sum_{i=1}^{k} p_i\right)\right)^n \forall t_1, ..., t_k \in \mathbb{R}$$

Distribuciones marginales

$M_{X_i}(t_i) = M_X(0, ..., 0, t_i, 0, ..., 0) = (p_i e^{t_i} + (1 - p_i))^n \, \forall t_i \in \mathbb{R}$
de donde se deduce que $X_i \sim B(n, p_i)$, $i = 1, .., k$.
De forma análoga, la distribución de cualquier subvector $X_{i1}, ..., X_{il}$, $l < k$ es $M_l(n; p_1, ..., p_l)$.

Definición.

Sea $(X_1, ..., X_k) \sim M_k(n; p_1, ..., p_k)$. Para cualquier subvector $X_{i1}, ..., X_{il}$ con $i_1, ..., i_l \in \{1, ..., k\}$, $X_{i1}, ..., X_{il} \sim M_l(n; p_1, ..., p_l)$.
En particular, $X_i \sim B(n, p_i), i = 1, ..., k$.

F.G.M. Distribución marginales y condicionadas.

Distribuciones condicionadas

Las distribuciones condicionadas de una distribución Multinomial, también se distribuyen según una Multinomial.

$$(X_1, ..., X_l | X_{l+1} = x_{l+1}, ..., X_k = x_k) \sim$$
$$M_l \left(n - \sum_{i=l+1}^{k} x_i ; \frac{p_1}{1 - \sum_{i=l+1}^{k} p_i}, ..., \frac{p_l}{1 - \sum_{i=l+1}^{k} p_i} \right)$$

Y en particular:

$$X_i | X_j = x_j \sim B\left(n - x_j ; \frac{pi}{1 - p_j} \right), i, j = 1, ..., k \quad i \neq j$$

Regresión y correlación. Caso bidimensional.

La curva de regresión de X_i sobre X_j es:

$$X_i = E[X_i | X_j = x_j] = \frac{np_i}{1 - p_j} - \frac{p_i}{1 - p_j} X_j$$

que coincide con la recta de regresión.

Las razones de correlación coinciden con el coeficiente de determinación.

$$\eta^2_{X_i|X_j} = \eta^2_{X_j|X_i} = \rho^2_{X_i X_j} = \frac{p_i p_j}{(1 - p_i)(1 - p_j)}$$

$i, j = 1, ..., k \quad i \neq j$

Reproductividad

La distribución Multinomial es reproductiva respecto al parámetro n:

$X_i \sim M(n_i; p_1, ..., p_k)$ $i = 1, ..., p$ v.a. independientes $\Rightarrow$

$$\sum_{i=1}^{p} X_i \sim M\left(\sum_{i=1}^{p} n_i; p_1, ..., p_k\right)$$

Distribución Normal Bidimensional

Definición

Se dice que (X_1, X_2) se distribuye según una Normal Bidimensional con vector de medias (μ_1, μ_2) y matriz de varianzas-covarianzas $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}$ si su función de densidad viene dada por:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]}$$

con $x_i \in \mathbb{R}$, $\mu_i \in \mathbb{R}, \sigma_i > 0. -1 < \rho < 1, i = 1, 2$

Expresión matricial: $f(x_1, x_2) = \frac{1}{2\pi\sqrt{|\Sigma|}} e^{\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)}$

Distribuciones marginales y condicionadas

Distribuciones marginales

$$X_1 \sim N(\mu_1, \sigma_1^2), \qquad X_2 \sim N(\mu_2, \sigma_2^2)$$

Distribuciones condicionadas

$$X_1|X_2 = x_2 \sim N\left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), \sigma_1^2(1-\rho^2)\right)$$

$$X_2|X_1 = x_1 \sim N\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), \sigma_2^2(1-\rho^2)\right)$$

Regresión y correlación

Las curvas de regresión coinciden con las rectas de regresión y vienen dadas por la media de las distribuciones normales condicionadas:

$$X_1 = E[X_1|X_2 = x_2] = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2)$$

$$X_2 = E[X_2|X_1 = x_1] = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1)$$

Los E.C.M asociados son:

$$ECM[X_1|X_2] = E[Var[X_1|X_2]] = \sigma_1^2(1-\rho^2)$$

$$ECM[X_2|X_1] = E[Var[X_2|X_1]] = \sigma_2^2(1-\rho^2)$$

Además:

$$\eta^2_{X_1|X_2} = \eta^2_{X_2|X_1} = \rho^2_{X_1,X_2}$$

Normalidad de las Combinaciones Lineales y Función Generatriz de Momentos

Normalidad de las Combinaciones Lineales

$$X \sim N_2(\mu, \Sigma) \Rightarrow Y = X \cdot A_{2xq} \sim N_q(\mu A, A^T \Sigma A), \quad q = 1, 2$$

Función Generatriz de Momentos

$$M_{X_1,X_2}(t_1, t_2) = e^{t_1\mu_1 + t_2\mu_2 + \frac{t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2t_1t_2\rho\sigma_1\sigma_2}{2}} \quad \forall t_1, t_2 \in \mathbb{R}^2$$