EYP 1027 Modelos Probabilísticos

Clase 12

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística
Pontificia Universidad Católica de Chile

Segundo Semestre 2020

Contenido I

Vimos que un vector aleatorio $(X_1, \ldots, X_n)$ en $(\Omega,\, \mathcal{A},\, P)$ es una función,

$$(X_1, \ldots, X_n) : \Omega \longrightarrow \mathbb{R}^n,$$

la cual asigna a cada $\omega \in \Omega$ un vector de números reales,

$$(x_1, \ldots, x_n) = (X_1(\omega), \ldots, X_n(\omega)) \in \mathbb{R}^n.$$

Es decir, $(X_1, \ldots, X_n)$ un vector aleatorio en $(\Omega,\, \mathcal{A},\, P)$ si, y sólo si, todas sus coordenas $X_1, \ldots, X_n$ son variables aleatorias en $(\Omega,\, \mathcal{A},\, P)$.

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

También vimos que la distribución de probabilidad del vector aleatorio $(X_1, \ldots, X_n)$, o distribución de probabilidad conjunta de las variables aleatorias $X_1, \ldots, X_n$, esta dada por,

$$P_{X_1,\ldots,X_n}(B) = P\{(X_1, \ldots, X_n) \in B\}, \quad \forall\, B \subset \mathbb{R}^n,$$

tal que $\{(X_1, \ldots, X_n) \in B\} := \{\omega \in \Omega : (X_1(\omega), \ldots, X_n(\omega)) \in B\} \in \mathcal{A}$.

En particular, la función de distribución acumulada (fda) de $(X_1, \ldots, X_n)$ o fda conjunta de $X_1, \ldots, X_n$, esta dada por,

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n), \quad \forall\, (x_1, \ldots, x_n) \in \mathbb{R}^n,$$

la que a su vez determina la distribución de probabilidad de conjunta de $X_1, \ldots, X_n$.

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

Finalmente, concluimos que una distribución de probabilidad conjunta queda determinada por la fmp conjunta en el caso discreto (c.d.), y por la fdp conjunta en el caso continuo (c.c.). Es decir, para todo $B \subset \mathbb{R}^n$,

$$P_{X_1,\ldots,X_n}(B) = \begin{cases} \sum_{(x_1,\ldots,x_n)\in B} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) & \text{c.d.,} \\ \\ \int_{(x_1,\ldots,x_n)\in B} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)dx_n\cdots dx_1 & \text{c.c.,} \end{cases}$$

donde

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \begin{cases} P(X_1 = x_1,\ldots,X_n = x_n) & \text{fdp conjunta (c.d.),} \\ \\ \dfrac{\partial^n F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{\partial x_1\cdots\partial x_n} \text{ (cp1)} & \text{fdp conjunta (c.c.).} \end{cases}$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

Distribuciones marginales

Por otra parte, partir de la distribución conjunta de $X_1, \ldots, X_n$ se puede obtener la distribución de cualquier subvector de $(X_1, \ldots, X_n)$, y en particular las distribuciones marginales de cada una de sus coordenadas.

De hecho, si $(X_1, \ldots, X_k)$ es subvector formado por las primeras k $(1 \leq k < n)$ componentes del vector aleatorio $(X_1, \ldots, X_n)$, entonces la fda de $(X_1, \ldots, X_k)$ puede determinarse como,

$$F_{X_1,\ldots,X_k}(x_1, \ldots, x_k) = F_{X_1,\ldots,X_k,X_{k+1},\ldots,X_n}(x_1, \ldots, x_k, \infty, \ldots, \infty)$$
$$\forall\, (x_1, \ldots, x_k) \in \mathbb{R}^k.$$

donde $F_{X_1,\ldots,X_k,X_{k+1},\ldots,X_n}(x_1, \ldots, x_k, \infty, \ldots, \infty) := \lim_{x_{k+1},\ldots,x_n \to \infty} F_{X_1,\ldots,X_k,X_{k+1},\ldots,X_n}(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n)$.

De la misma forma se puede obtener la fda de cualquier subvector formado con diferentes X_i's.

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

En particular, para cada $i = 1, \ldots, n$, la fda marginal de X_i puede obtenerse como,

$$F_{X_i}(x_i) = F_{X_1,\ldots,X_{i-1},X_i,X_{i+1},\ldots,X_n}(\infty,\ldots,\infty,x_i,\infty,\ldots,\infty)$$
$$\forall\, x_i \in \mathbb{R}.$$

Es decir, la fda conjunta $F_{X_1,\ldots,X_n}$ determina completamente todas las dfa's marginales $F_{X_1},\ldots,F_{X_n}$.

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

Similarmente, si $f_{X_1,\dots,X_n}$ es la fmp conjunta (c.d.) o fdp conjunta (c.c.) de $X_1,\dots,X_n$, entonces, para cada $1 \le k < n$, la fmp o fdp del subvector $(X_1,\dots,X_k)$ puede determinarse como,

$$f_{X_1,\dots,X_k}(x_1,\dots,x_k) = \begin{cases} \sum_{x_{k+1}\in\mathbb{R}} \cdots \sum_{x_n\in\mathbb{R}} f_{X_1,\dots,X_k,X_{k+1},\dots,X_n}(x_1,\dots,x_k,x_{k+1},\dots,x_n) & \text{c.d.} \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\dots,X_k,X_{k+1},\dots,X_n}(x_1,\dots,x_k,x_{k+1},\dots,x_n)\,dx_{k+1}\cdots dx_n & \text{c.c.} \end{cases}$$

De igual forma se puede obtener la fmp (c.d.) o fdp (c.c.) de cualquier subvector formado con diferentes X_i's.

En particular, para $k = 1$, se obtiene la fmp marginal (c.d.) o fdp marginal (c.c.) de X_1, es decir,

$$f_{X_1}(x_1) = \begin{cases} \sum_{x_2 \in \mathbb{R}} \cdots \sum_{x_n \in \mathbb{R}} f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) & \text{c.d.} \\ \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) dx_2 \cdots dx_n & \text{c.d.} \end{cases}$$

De forma análoga se obtienen las fmp's marginales (c.d.) o fdp's marginales (c.c.) de $X_2, \ldots, X_n$.

Una propiedad importante ocurre cuando la fda conjunta se puede factorizar como el producto de sus dfa's marginales.

Definición 1.1

Sea $(X_1, \ldots, X_n)$ un vector aleatorio (arbitrario) con fda conjunta $F_{X_1,\ldots,X_n}$ y fda's marginales $F_{X_1}, \ldots, F_{X_n}$, respectivamente. Entonces, las variables aleatorias $X_1, \ldots, X_n$ se dicen (mutuamente) independientes, ssi:

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i) \quad \forall\, (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

En el caso discreto o continuo, la definición de variables aleatorias independientes se puede etablecer de forma equivalente en términos de las fmp's conjuntas y marginales (c.d.) o las fdp's conjuntas y marginales (c.c.).

Definición 1.2

Si $(X_1, \ldots, X_n)$ es un vector aleatorio (discreto o continuo) con fmp conjunta (c.d.) o fdp conjunta (c.c.) $f_{X_1,\ldots,X_n}$, y fmp's marginales (c.d.) o fdp's marginales (c.c.) $f_{X_1}, \ldots, f_{X_n}$, respectivamente. Entonces, las variables aleatorias $X_1, \ldots, X_n$ se dicen (mutuamente) independientes, ssi:

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i) \quad \forall\, (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

Notas:

1) Sean $\mathcal{X}$ y $\mathcal{X}_1, \ldots, \mathcal{X}_n$ los recoridos conjunto y marginales de $X_1, \ldots, X_n$, respectivamente. Entonces, una condicion necesaria (pero no suficiente) para que $X_1, \ldots, X_n$ sean variables aleatorias (mutuamente) independientes es que $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$.

2) Si $X_1, \ldots, X_n$ son variables aleatorias (mutuamente) independientes, entonces todos los subvectores formados con distintas componentes también son independientes.

3) Si $X_1, \ldots, X_n$ son variables aleatorias (absolutamente) continuas (mutuamente) independientes, entonces el vector aleatorio $(X_1, \ldots, X_n)$ es (absolutamente) continuo.

4) Si $X_1, \ldots, X_n$ son variables aleatorias independientes e identicamente distribuidas (iid), es decir, todas ellas tienen la misma fda F (fmp/fdp f), escribimos $X_1, \ldots, X_n \overset{\text{iid}}{\sim} F$ o $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f$.

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

Caso bivariado

A continuación reforzamos algunos de los conceptos previos en el contexto bivariado, es decir, para $n = 2$, con $X_1 = X$ y $X_2 = Y$.

Sea (X, Y) es un vector aleatorio con fmp conjunta (c.d.) o fdp conjunta (c.c.) $f_{X,Y}(x, y)$, entonces,

a) La distribución de probabilidad de (X, Y) es,

$$
\begin{aligned}
P_{X,Y}(B) &= P\{(X, Y) \in B\} \\
&= \begin{cases} \sum_{(x,y)\in B} f_{X,Y}(x, y) & \text{c.d.} \\ \\ \int_{(x,y)\in B} f_{X,Y}(x, y) dy dx & \text{c.c.} \end{cases} \qquad \forall\, B \subset \mathbb{R}^2.
\end{aligned}
$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

b1) En particular, la fda (da conjunta) de (X, Y) es,

$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$$
$$= \begin{cases} \sum_{u \leq x} \sum_{v \leq y} f_{X,Y}(u,v) & \text{c.d.} \\ \\ \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v) dv du & \text{c.c.} \end{cases} \qquad \forall (x,y) \in \mathbb{R}^2.$$

b1) Las fda's marginales son,

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x,y) := F_{X,Y}(x,\infty) \;\; \forall x \in \mathbb{R} \quad \text{(fda marginal de } X\text{)},$$
$$F_Y(x) = \lim_{x \to \infty} F_{X,Y}(x,y) := F_{X,Y}(\infty,y) \;\; \forall y \in \mathbb{R} \quad \text{(fda marginal de } Y\text{)}$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

c1) Si (X, Y) es discreto, las fmp marginales son,

$$\begin{aligned}
f_X(x) = P(X = x) &= P(X = x, Y \in \mathbb{R}) \\
&= P(\{X = x\} \cap [\cup_{y \in \mathbb{R}} \{Y = y\}]) \\
&= P(\cup_{y \in \mathbb{R}} [\{X = x\} \cap \{Y = y\}]) \\
&= P(\cup_{y \in \mathbb{R}} \{X = x, Y = y\}]) \\
&= \sum_{y \in \mathbb{R}} P(X = x, Y = y) \\
&= \sum_{y \in \mathbb{R}} f_{X,Y}(x, y) \;\; \forall\, x \in \mathbb{R} \;\; \text{(fmp marginal de } X\text{)},
\end{aligned}$$

y,

$$
\begin{aligned}
f_Y(y) = P(Y = y) &= P(Y \in \mathbb{R}, Y = y) \\
&= P([\cup_{x\in\mathbb{R}}\{Y = y\}] \cap \{X = x\}) \\
&= P(\cup_{x\in\mathbb{R}}[\{X = x\} \cap \{Y = y\}]) \\
&= P(\cup_{x\in\mathbb{R}}\{X = x, Y = y\}]) \\
&= \sum_{x\in\mathbb{R}} P(X = x, Y = y) \\
&= \sum_{x\in\mathbb{R}} f_{X,Y}(x, y) \;\; \forall\, y \in \mathbb{R} \;\; \text{(fmp marginal de } Y\text{)}.
\end{aligned}
$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

El resultado anterior, prueba el siguiente teorema.

Teorema 1.1

Sea (X, Y) un vector aleatorio bivariado discreto con fmp conjunta $f_{X,Y}(x, y)$. Entonces las fmp marginal de X e Y, $f_X(x) = P(X = x)$ y $f_Y(y) = P(Y = y)$, están dadas por

$$f_X(x) = \sum_{y \in \mathbb{R}} f_{X,Y}(x, y) \quad \text{y} \quad f_Y(y) = \sum_{x \in \mathbb{R}} f_{X,Y}(x, y)$$

c2) Recordando que si (X, Y) es un vector aleaotorio continuo, entonces X e Y también son variables aleatorias continuas, las fdp's marginales de X e Y se pueden definir como en el caso discreto, pero reemplazando las sumas por integrales. Específicamente, las fdp's marginales de X e Y están dados por,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy, \quad -\infty < x < \infty,$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx, \quad -\infty < y < \infty.$$

Alternativamente, como X e Y son variables aleatorias continuas, las fdp's marginales de X e Y también pueden obtenerse derivando las respectivas fda's marginales, es decir,

$$f_X(x) = \frac{dF_X(x)}{dx} \text{ (cp1) y } f_Y(y) = \frac{dF_Y(y)}{dy} \text{ (cp1)}$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

d) Finalmente, X e Y son variables aleatorias independientes si, y sólo si,

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) \quad \forall\,(x,y) \in \mathbb{R}^2.$$

Equivalentemente, si las variables aleatorias X e Y son ambas discretas o ambas continuas, ellas son independientes si, y sólo si,

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad \forall\,(x,y) \in \mathbb{R}^2.$$

Ejemplo 1.1

Caso discreto: Sean A y B dos eventos de Ω. Defina las variables aleatorias,

$$X = I_A := \begin{cases} 1 & \text{si ocurre } A, \\ 0 & \text{si no.} \end{cases} \quad \text{e} \quad Y := I_B = \begin{cases} 1 & \text{si ocurre } B, \\ 0 & \text{si no.} \end{cases}$$

Note que:

$$\begin{array}{ccccccccc} \Omega & = & A^cB^c & + & A^cB & + & AB^c & + & AB \\ (X,Y) \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \mathcal{X} & = & \{(0,0), & & (0,1), & & (1,0), & & (1,1)\} \\ & = & \{0,1\} & \times & \{0,1\} & & & & \\ & = & \mathcal{X}_1 & \times & \mathcal{X}_2 & & & & \end{array}$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

Es decir, la fmp conjunta de X e Y es,

$$
\begin{aligned}
f_{X,Y}(x,y) &= P(X=x, Y=y) \\
&= \begin{cases} P(A^cB^c) & \text{si } x=0,\ y=0, \\ P(A^cB) & \text{si } x=0,\ y=1, \\ P(AB^c) & \text{si } x=1,\ y=0, \\ P(AB) & \text{si } x=1,\ y=1, \\ 0 & \text{eoc} \end{cases}
\end{aligned}
$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

La fmp conjunta $f_{X,Y}(x,y) = P(X = x, Y = y)$ anterior se puede reescribir como sigue,

$x \backslash y$	0	1	$P(X = x)$
0	$P(A^c B^c)$	$P(A^c B)$	$P(A^c) = 1 - P(A)$
1	$P(AB^c)$	P(AB)	$P(A)$
$P(Y = y)$	$P(B^c) = 1 - P(B)$	$P(B)$	1

donde en los margenes de la tabla se obtienen la distribuciones marginales $f_X(x) = P(X = x)$ y $f_Y(y) = P(Y = y)$ de X e Y, respectivamente.

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

Es decir,

$$\begin{aligned} f_X(x) &= P(X = x) \\ &= \sum_{y \in \{0,1\}} P(X = x, Y = y) = \begin{cases} 1 - P(A) & \text{si } x = 0, \\ P(A) & \text{si } x = 1, \\ 0 & \text{eoc.} \end{cases} \end{aligned}$$

$$\begin{aligned} f_Y(y) &= P(Y = y) \\ &= \sum_{x \in \{0,1\}} P(X = x, Y = y) = \begin{cases} 1 - P(B) & \text{si } y = 0, \\ P(B) & \text{si } y = 1, \\ 0 & \text{eoc.} \end{cases} \end{aligned}$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

En particular, si A y B son eventos independientes, entonces

$$\begin{aligned}
P(A^cB^c) &= P(A^c)P(B^c) = (1-P(A))(1-P(B)),\\
P(A^cB) &= P(A^c)P(B) = (1-P(A))P(B),\\
P(AB^c) &= P(A)P(B^c) = P(A)(1-P(B)),\\
P(AB) &= P(A)P(B).
\end{aligned}$$

En tal caso, se tiene que,

$$f_{X,Y}(x,y) = P(X=x, Y=y) = P(X=x)P(Y=y) = f_X(x)f_Y(y) \qquad \forall\,(x,y).$$

lo cual ocurre si, y sólo si, X e Y son variables aleatorias independientes.

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

Aplicaciones:

1) Lanzar una moneda al aire dos veces consecutivas. Sean,

$$A = \text{ sale cara en el primer lanzamiento}$$
$$B = \text{ sale cara en el segundo lanzamiento}$$
$$\Longrightarrow A \text{ y } B \text{ son eventos independientes.}$$

En este caso,

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$
$$= \begin{cases} p^{x+y}(1-p)^{2-(x+y)} & \text{si } x, y = 0, 1, \\ 0 & \text{eoc,} \end{cases}$$

donde $p =$ probabilidad de obtener cara al lanzar la moneda. Aquí, se tiene que $X, Y \overset{iid}{\sim} Ber(p)$, es decir, con fmp $f(z) = p^z(1-p)^{1-z}I_{\{0,1\}}(z)$.

2) Sacar una carta al azar de una baraja de 52 cartas. Sean,

$$\begin{aligned} A =& \text{ sale rey} \\ B =& \text{ sale corazón} \\ \Longrightarrow & A \text{ y } B \text{ son eventos independientes.} \end{aligned}$$

En este caso,

$$\begin{aligned} P(X = x, Y = y) &= P(X = x)P(Y = y) \\ &= \begin{cases} \left(\frac{1}{13}\right)^x \left(1 - \frac{1}{13}\right)^{1-x} \left(\frac{1}{4}\right)^y \left(1 - \frac{1}{4}\right)^{1-y} & \text{si } x, y = 0, 1, \\ 0 & \text{eoc.} \end{cases} \end{aligned}$$

$\Longrightarrow$ X e Y son variables aleatorias independientes, con $X \sim Ber(1/13)$ e $Y \sim Ber(1/4)$.

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

3) Lanzar un dado honesto dos veces consecutivas. Sean,

$$A = \text{ la suma es impar } \implies P(A) = 1/2,$$
$$B = \text{ la suma es par } \implies P(B) = 1/2$$
$$\implies A \text{ y } B \text{ no son eventos independientes}$$
$$(AB = \emptyset, \quad A + B = \Omega, \text{ es decir, } B = A^c)$$
$$\implies X + Y = 1 \text{ y } X \overset{d}{=} Y \sim Ber(1/2),$$
$$\text{pero } X \text{ e } Y \text{ no son independientes.}$$

En este caso,

$$P(X = x, Y = y) = \begin{cases} 1/2 & \text{si } x, = 0, 1, \ x + y = 1, \\ 0 & \text{eoc.} \end{cases}$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

Ejemplo 1.2

Caso continuo:

1) Sean X e Y variables aleatorias continuas con fda conjunta dada por,

$$F_{X,Y}(x,y) = \begin{cases} (1-e^{-x})(1-e^{-y}) & \text{si } x, y > 0, \\ 0 & \text{eoc.} \end{cases}$$

Sigue que,

$$F_X(x) = F_{X,Y}(x,\infty) = \begin{cases} 1-e^{-x} & \text{si } x > 0, \\ 0 & \text{eoc.} \end{cases}$$

$$F_Y(y) = F_{X,Y}(\infty,y) = \begin{cases} 1-e^{-y} & \text{si } y > 0, \\ 0 & \text{eoc.} \end{cases}$$

$$\Longrightarrow F_{X,Y}(x,y) = F_X(x)F_Y(y) \;\; \forall\ (x,y).$$

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

Alternativamente, como

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \begin{cases} e^{-x}e^{-y} & \text{si } x, y > 0, \\ 0 & \text{eoc.} \end{cases}$$

y

$$F_X(x) = \frac{\partial F_X(x)}{\partial x} = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{eoc.} \end{cases}$$

$$F_Y(y) = \frac{\partial F_Y(y)}{\partial y} = \begin{cases} e^{-y} & \text{si } y > 0, \\ 0 & \text{eoc.} \end{cases}$$

$$\Longrightarrow f_{X,Y}(x,y) = f_X(x) f_Y(y) \ \ \forall \ (x,y).$$

En este caso, se tiene que $X, Y \overset{iid}{\sim} \exp(1)$. Note, además, que $\mathcal{X} = \mathbb{R}_+^2 = \mathbb{R}_+ \times \mathbb{R}_+ = \mathcal{X}_1 \times \mathcal{X}_2$.

Vectores Aleatorios: Distribuciones conjuntas, marginales e independencia

2) Sea

$$f(x,y) = \begin{cases} c(|x|+|y|) & \text{si } |x|+|y| \leq 1, \\ 0 & \text{eoc.} \end{cases}$$

a) Haga un gráfico de $f(x,y)$
b) Para qué valores de la constante c, $f(x,y)$ es una fdp bivariada? (Resp. $c = 3/4$)
c) Calcule $P((X;Y) \in \{(x,y) : x, y \geq 0, x+y \leq 1\})$. (Resp. $1/4$)
d) Encuentre las fdp's marginales f_X y f_Y.
e) Son X e Y variables aleatorias continuas?. Justifique.