Variable Aleatoria bidimensional

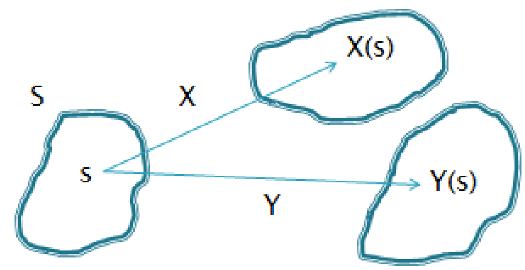
Hasta ahora el resultado de un experimento era un único valor; es decir, las v.a. eran unidimensionales. En ocasiones nos interesa observar simultáneamente 2 o más características

Ejemplos: Altura y peso de una persona, Dureza y resistencia de una pieza de acero.

Definición

Sea ε un experimento y S su espacio muestral. Sean $X=X(s)$ e $Y=Y(s)$ dos funciones que asignan un número real a cada $s \in S$.

Llamamos a (X,Y) ***v.a. bidimensional***



Clasificación

- (X, Y) *es v.a.* ***bidimensional discreta si X e Y son discretas***
- (X, Y) *es v.a.* ***bidimensional continua si X e Y son continuas***
- El caso X continua, Y discreta (o viceversa) no lo consideramos.

Recorrido de una v.a.bidimensional

Al conjunto de valores que toma la variable aleatoria bidimensional (X,Y) lo llamaremos ***recorrido de la*** *v.a. (X,Y)* y lo indicaremos R_{XY}

$$R_{XY} = \{(x,y) : x=X(s) \text{ e } y=Y(s) \text{ con } s \in S\}$$

Note que $R_{XY} \subset R^2$

Clasificación de recorridos R_{XY}

Como con cualquier espacio muestral, según el número de elementos que lo constituyen, podemos clasificar a los recorridos R_{XY} *en*

- *numerables (finitos o infinitos)*
- *no-numerables.*

Recorridos numerables

Los recorridos numerables son, en general, de la forma

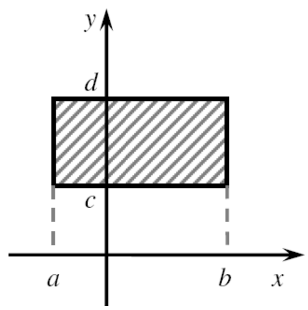
$R_{XY} = \{(x_i, y_j) \text{ con } i=1,2,\ldots n \text{ y } j=1,2, \ldots,m\}= \{(x_1,y_1), (x_1, y_2), \ldots, (x_n, y_m)\}$ **(FINITO)**

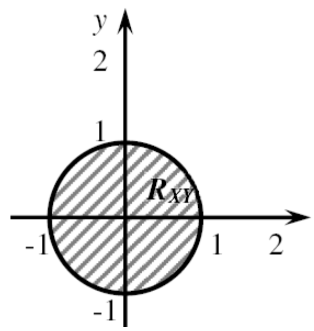
$R_{XY} = \{(x_i, y_j) \text{ con } i=1,2,\ldots \text{ y } j=1,2, \ldots\} = \{(x_1,y_1), (x_1, y_2), \ldots\}$ **(INFINITO NUMERABLE)**

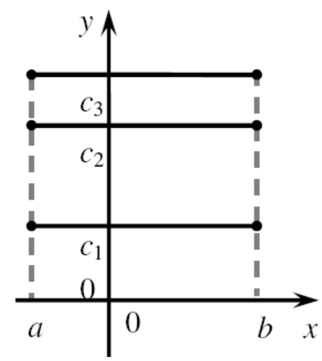
Recorridos no numerables

Los recorridos no numerables son regiones o subconjuntos no numerables del plano Euclidiano.

Ejemplos $R_{XY}=\{(x,y): x^2 + y^2 \leq 1\}$







$R_{XY}=\{(x,y): a \leq x \leq b;\ c \leq y \leq d\}$

$R_{XY}=\{(x,y_j): a \leq x \leq b;\ y_j=c_1,c_2,c_3\}$
(Este es un recorrido no numerable “mixto”)

Función de probabilidad puntual conjunta

Sea $(X\,,Y\,)$ una v.a. bidimensional discreta y sea R_{XY} su recorrido (numerable).

Sea $p : R_{XY} \rightarrow R$ una función que a cada $(x_i,\ y_j)$ le asigna un número real $p(x_i,y_j)$ tal que

$$P(X=x_i, Y=y_j)= p(x_i,y_j) \ \forall\ p(x_i,y_j) \in R_{XY}$$

y que verifica

$$a)\ p(x_i, y_j) \geq 0 \quad \forall (x_i, y_j) \in R_{XY}$$

$$b)\ \sum_{(x_i, y_j) \in R_{XY}} p(x_i, y_j) = \sum_i \sum_j p(x_i, y_j) = 1$$

A ***p*** la llamaremos ***función de probabilidad puntual conjunta*** *de la v.a. bidimensional* (*X ,Y).*

Ejemplo 4.1

Una agencia de seguros vende pólizas para casas y autos de manera conjunta. Para una póliza de auto los montos deducibles (a cargo del propietario) son 100$ y 250$ mientras que para una casa son 0$, 100$ y 200$. Cada cliente puede ser visto como una v.a. bidimensional (X,Y) donde X=“cantidad deducible de la póliza de auto” e Y = “cant.deducible de la póliza de la casa”. *Note que* S *tiene 6 elementos*

Suponga que la *función de probabilidad (fdp) conjunta* es la siguiente

p(x,y)	***0***	***100***	***200***
100	*0.20*	*0.10*	*0.20*
250	*0.05*	*0.5*	*0.30*

Es decir que

P(100,200) = P(X=100, Y=200)=0.20

P(Y≥100)=P(100,100) + P(100,200) + P(250,100) + P(250,200) =

= 0.1+0.2+0.15+0.3=0.75

Función de distribución marginal de una v.a. discreta

La función de distribución marginal de una v.a. discreta se obtiene sumando p(x,y) en los valores de la otra variable.

$$p_X(x_i) = P(X = x_i) = \sum_j p(x_i, y_j) \quad \forall i$$

$$p_Y(y_j) = P(Y = y_j) = \sum_i p(x_i, y_j) \quad \forall j$$

Las funciones $\boldsymbol{p_X}$ y $\boldsymbol{p_Y}$ coinciden con las distribuciones de probabilidad de X e Y respectivamente

Ejemplo : En el ejemplo de las pólizas de seguro

p(x,y)	***0***	***100***	***200***	$p_X(x)$
100	*0.20*	*0.10*	*0.20*	0.5
250	*0.05*	*0.15*	*0.30*	0.5
$p_Y(y)$	0.25	0.25	0.50	

Función de densidad de probabilidad conjunta

Sea (X ,Y) una v.a. bidimensional continua y sea R_{XY} su recorrido (no numerable).

Sea f: $R_{XY} \to R$ una función que a cada punto (x, y) de R_{XY} le asigna un número real *f(x,y)* tal que $$P(B) = \iint_B f(x,y)\,dx\,dy \quad \forall B \subseteq R_{XY}$$

y que verifica

$$a)\ f(x,y) \geq 0 \quad \forall (x,y) \in R^2$$

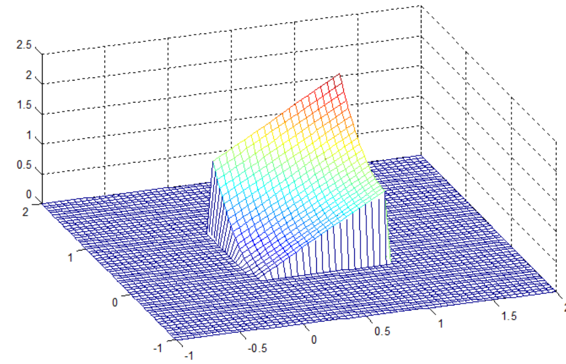
$$b) \iint_{R^2} f(x,y)\,dx\,dy = 1$$

A ***f*** la llamaremos ***función de densidad de probabilidad conjunta*** *de la v.a. bidimensional* (X ,Y).

Ejemplo 4.2

Un banco tiene dos ventanillas para atención al cliente. En un día elegido al azar se observa la parte del tiempo en que está ocupada cada ventanilla. Por ejemplo (0.25, 0.75) indica que la 1er. ventanilla sólo se ocupó la cuarta parte del tiempo mientras

que la 2da. estuvo activa un 75% del horario.
En este ejemplo $R_{XY}=\{(x,y): 0\leq x \leq 1,\ 0 \leq y \leq 1\}$
Suponga que la *fdp conjunta* de (X,Y) está dada por



$$f(x,y)=\begin{cases} \frac{6}{5}(x+y^2) & 0\leq x\leq 1, 0\leq y\leq 1 \\ 0 & \textit{en caso contrario} \end{cases}$$

a) Comprobar que es una *fdp* legítima.

Solución: Puede verse e la figura que f(x,y)≥0, falta ver si el volumen bajo la superficie f(x,y) y arriba del área determinada por (0,0) y (1,1) da 1

$$\int_0^1\int_0^1 \frac{6}{5}(x+y^2)\,dx\,dy \ =\int_0^1 \frac{6}{5}\left(\frac{1}{2}+y^2\right)dy \ =\frac{6}{5}\left(\frac{y}{2}+\frac{y^3}{3}\right)\Bigg|_0^1 =\frac{6}{5}\left(\frac{1}{2}+\frac{1}{3}\right)=1$$

b) ¿Cuál es la probabilidad de que las ventanillas estén ocupadas a lo sumo un cuarto del tiempo?

$$P\left(0\leq x\leq \frac{1}{4}, 0\leq y\leq \frac{1}{4}\right)=\int_0^{1/4}\int_0^{1/4}\frac{6}{5}(x+y^2)\,dx\,dy=0.0109$$

Función de Densidad de Probabilidad marginal

Las ***funciones de densidad de probabilidad marginal*** de X e Y, denotadas por $f_X(x)$ y $f_Y(y)$ respectivamente están dadas por

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy \qquad \textit{para} \quad -\infty < x < \infty$$

$$f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx \qquad \textit{para} \quad -\infty < y < \infty$$

f_X y f_Y coinciden con las fdp de X e Y respectivamente

Ejemplo 4.3

Retomando el ejemplo de las ventanillas de atención que tenían *fdp conjunta* de (X,Y) está dada por

$$f(x,y)=\begin{cases} \frac{6}{5}(x+y^2) & 0\leq x\leq 1, 0\leq y\leq 1 \\ 0 & \textit{en caso contrario} \end{cases}$$

¿ Cuál es la *fdp marginal* de la ventanilla representada por X ?

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 \frac{6}{5}(x+y^2)dy = \frac{6}{5}x + \frac{2}{5}$$

$$f_X(x) = \begin{cases} \dfrac{6}{5}x + \dfrac{2}{5} & 0 \le x \le 1 \\ 0 & en\ caso\ contrario \end{cases}$$

Ejercicio 4.1

Utilizando la siguiente fdp conjunta $f(x,y) = \begin{cases} \dfrac{6}{5}(x+y^2) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & en\ caso\ contrario \end{cases}$

calcular la fdp marginal de Y.

$$f_Y(x) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^1 \frac{6}{5}(x+y^2)dx = \frac{6}{5}\left(\frac{x^2}{2} + y^2 x\right)\Bigg|_0^1 = \frac{3}{5} + \frac{6}{5}y^2$$

$$f_Y(y) = \begin{cases} \dfrac{3}{5} + \dfrac{6}{5}y^2 & 0 \le y \le 1 \\ 0 & en\ caso\ contrario \end{cases}$$

Utilizar $f_Y(y)$ para calcular $P\left(\frac{1}{4} \le Y \le \frac{3}{4}\right)$

$$P\left(\frac{1}{4} \le Y \le \frac{3}{4}\right) = \int_{1/4}^{3/4}\left(\frac{6}{5}y^2 + \frac{3}{5}\right)dy = \left(\frac{2y^3}{5} + \frac{3y}{5}\right)\Bigg|_{1/4}^{3/4} = \frac{2}{5}\left(\frac{3}{4}\right)^3 + \frac{9}{20} - \frac{2}{5}\left(\frac{1}{4}\right)^3 - \frac{3}{20} = \frac{37}{80} = 0.4625$$

Ejercicio 4.2

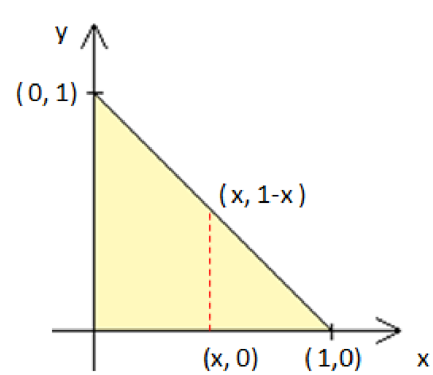
Una compañía vende latas de nueces mixtas de lujo que contienen almendras, castañas y maní.El peso neto de cada lata es de 1kg.pero la contribución en peso de cada tipo de nuez es aleatoria.Como los tres pesos suman 1, un modelo de probabilidad conjunta para dos tipos de fruto da la información necesaria a cerca del tercer tipo.

Se selecciona una lata al azar.

X="peso de las almendras en la lata"

Y= "peso de las castañas en la lata"

La región de densidad positiva es la de la figura



La *fdp conjunta* de X e Y es la siguiente

$$f(x,y)=\begin{cases}24\,x\,y & 0\le x\le 1,\ 0\le y\le 1,\ x+y\le 1\\ 0 & en\ caso\ contrario\end{cases}$$

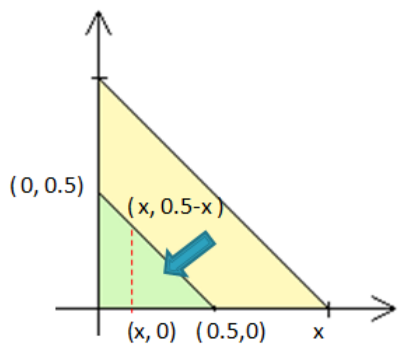
a) Comprobar que es una *fdp* legítima.

Para ver que es una fdp legítima debe cumplirse que f(x,y)≥0 (se cumple) y que el volumen bajo f(x,y) da 1.

$$\int_0^1\int_0^{1-x}24xy\,dy\,dx=\int_0^1 24x\frac{(1-x)^2}{2}dx=\int_0^1 12(x-2x^2+x^3)dx=6-8+3=1$$

b) Calcular la probabilidad de que a lo sumo el 50% de la lata contenga almendras y castañas.

$$\int_0^{0.5}\int_0^{0.5-x}24\,x\,y\,dy\,dx$$



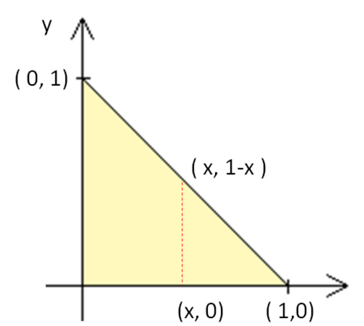
$$\int_0^{0.5}\int_0^{0.5-x}24\,x\,y\,dy\,dx=\int_0^{0.5}12x(0.5-x)^2\,dx=\int_0^{0.5}(3x-12x^2+12x^3)dx$$

$$=\left(3\frac{x^2}{2}-4x^3+3x^4\right)\Bigg|_0^{0.5}=\frac{3}{2}(0.5)^2-4(0.5)^3+3(0.5)^4=0.0625$$

c) Calcular la fdp para las almendras.

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)\,dy=\int_0^{1-x}24xy\,dy=12x(1-x)^2$$

$$f_X(x)=\begin{cases}12\,x\,(1-x)^2 & 0\le x\le 1\\ 0 & si\ no\end{cases}$$



Distribuciones Condicionales

Función de probabilidad puntual condicional

Sea (X,Y) una v.a. bidimensional discreta cuya *fdp conjunta* es p(x,y).
Sean $p_X(x)$ y $p_Y(y)$ las fdp marginales de X e Y respectivamente.
Definimos la ***función de probabilidad puntual de X condicional a Y*** como sigue

$$p_{X|Y}(x_i \mid y_j) = P(X = x_i \mid Y = y_j) = \frac{p(x_i, y_j)}{p_Y(y_j)}$$

Análogamente, definimos la función de probabilidad puntual de Y condicional a X como sigue

$$p_{Y|X}(y_j \mid x_i) = P(Y = y_j \mid X = x_i) = \frac{p(x_i, y_j)}{p_X(x_i)}$$

Ejemplo 4.4

Volviendo al ejemplo de las pólizas de seguro

p(x,y)	***0***	***100***	***200***	$p_X(x)$
100	*0.20*	*0.10*	*0.20*	0.5
250	*0.05*	*0.5*	*0.30*	0.5
$p_Y(y)$	0.25	0.25	0.50	

$$P(X = 250 \mid Y = 100) = \frac{P(X = 250 \mid Y = 100)}{P(Y = 100)} = \frac{p(250,100)}{p_Y(100)} = \frac{0.15}{0.25} = 0.6$$

Función de densidad de probabilidad condicional

Sea (X,Y) una v.a. bidimensional continua cuya *fdp conjunta* es *f(x,y)*.
Sean $f_X(x)$ y $f_Y(y)$ las fdp marginales de X e Y respectivamente.
Para cualquier valor y de Y para el cual $f_Y(y)>0$ definimos la ***función de densidad de probabilidad de X condicional a Y =y*** como sigue

$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} \qquad -\infty < x < \infty$$

Análogamente, para cualquier valor x de X para el cual $f_X(x)>0$ definimos la ***función de densidad de probabilidad de Y condicional a X=x*** como sigue

$$f_{Y|X}(y \mid x) = \frac{f(x,y)}{f_X(x)} \qquad -\infty < y < \infty$$

Ejemplo 4.5

Retomando el ejemplo de las ventanillas del banco donde X e Y representan la parte del tiempo en que están ocupadas las ventanillas 1 y 2 respectivamente.

$$f(x,y)=\begin{cases}\frac{6}{5}(x+y^2) & 0\le x\le 1, 0\le y\le 1\\ 0 & en\ caso\ contrario\end{cases}$$

¿Cuál es la probabilidad de que la ventanilla 2 esté ocupada a lo sumo la mitad del tiempo dado que X=0.8?

$$P(Y\le 0.5 \mid X=0.8)=\int_{-\infty}^{0.5} f_{Y|X}(y\mid 0.8)dy$$

Tenemos que calcular

$$f_{Y|X}(y\mid 0.8)=\frac{f(0.8,y)}{f_X(0.8)} \qquad -\infty<y<\infty$$

Ya habíamos calculado

$$f_X(x)=\begin{cases}\frac{6}{5}x+\frac{2}{5} & 0\le x\le 1\\ 0 & en\ caso\ contrario\end{cases}$$

De donde

$$f_X(0.8)=\frac{6}{5}*0.8+\frac{2}{5}=\frac{34}{25}$$

Luego

$$f_{Y|X}(y\mid 0.8)=\frac{f(0.8,y)}{\frac{34}{25}}=\frac{25f(0.8,y)}{34} \qquad -\infty<y<\infty$$

$$=\frac{25\left(\frac{6}{5}(0.8+y^2)\right)}{34}=\frac{30(0.8+y^2)}{34}$$

$$\therefore\ f_{Y|X}(y\mid 0.8)=\frac{24+30y^2}{34} \qquad 0<y<1$$

Por lo tanto, la probabilidad buscada se calcula como

$$P(Y\le 0.5\mid X=0.8)=\int_{-\infty}^{0.5} f_{Y|X}(y\mid 0.8)dy=\int_0^{0.5}\frac{1}{34}(24+30y^2)dy=$$

$$=\frac{1}{34}(24*0.5+10*(0.5)^3)=0.39$$

Variables Aleatorias Independientes

Sea (X ,Y) una v.a. bidimensional discreta.

Sea *$p(x_i, y_j)$* su *fdp conjunta* y *$p_X(x_i)$* y *$p_Y(y_j)$* las correspondientes fdp marginales de X e Y. Decimos que X e Y son **variables aleatorias independientes** si y sólo si

$$p(x_i, y_j) = p_X(x_i) \,.\, p_Y(y_j) \qquad \forall (x_i, y_j) \in R_{XY}$$

Note que debe verificarse **para todos** los pares de valores $(x_i, y_j) \in R_{XY}$

Ejemplo 4.6

Una máquina se usa para un trabajo a la mañana y para otro diferente en la tarde. Sean *X e Y el número de veces que la máquina falla en la mañana y en la tarde* respectivamente. Esta es la fdp conjunta de (X,Y)

Y/X	**0**	**1**	**2**	**$p_Y(y_j)$**
0	0.1	0.2	0.2	0.5
1	0.04	0.08	0.08	0.2
2	0.06	0.12	0.12	0.3
$p_X(x_i)$	0.2	0.4	0.4	1

¿ X e Y son independientes?

Para demostrar que son independientes debe verificarse

$$\mathbf{p(x_i, y_j) = p_X(x_i) \,.\, p_Y(y_j) \; \forall (x_i, y_j) \in R_{XY}}$$

$p(0,0)=0.1 = p_X(0).p_Y(0)=0.2*0.5$

$p(0,1)=0.04 = p_X(0).p_Y(1)=0.2*0.2$

$p(0,2)=0.06 = p_X(0).p_Y(2)=0.2*0.3$

$p(1,0)=0.2 = p_X(1).p_Y(0)=0.4*0.5$

$p(1,1)=0.08 = p_X(1).p_Y(1)=0.4*0.2$

$p(1,2)=0.12 = p_X(1).p_Y(2)=0.4*0.3$

$p(2,0)=0.2 = p_X(2).p_Y(0)=0.4*0.5$

$p(2,1)=0.08 = p_X(2).p_Y(1)=0.4*0.2$

$p(2,2)=0.12 = p_X(2).p_Y(2)=0.4*0.3$

Por lo tanto, *X e Y son independientes*

Variables Aleatorias Continuas Independientes

Sea (X ,Y) una v.a. bidimensional continua. Sea *$f(x,y)$* su *fdp conjunta* y sean *$f_X(x)$* y *$f_Y(y)$* las fdp marginales de X e Y respectivamente. Decimos que X e Y son **variables aleatorias independientes** si y sólo si

$$f(x,y) = f_X(x) \,.\, f_Y(y) \qquad \forall (x, y) \in R^2$$

Ejemplo 4.7

Sean X e Y v.a. continuas que representan el tiempo de vida de dos dispositivos electrónicos. La fdp conjunta de la v.a. continua (X,Y) es:

$$f(x,y)=\begin{cases} e^{-(x+y)} & 0\le x<\infty,\ 0\le y<\infty \\ 0 & si\ no \end{cases}$$

Decir si X e Y son v.a. independientes.

Solución: Calculamos las fdp marginales de X e Y

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy=\int_0^{\infty} e^{-(x+y)}dy=\int_0^{\infty} e^{-x}e^{-y}dy=e^{-x}$$

$$f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx=\int_0^{\infty} e^{-(x+y)}dx=\int_0^{\infty} e^{-x}e^{-y}dx=e^{-y}$$

Luego las fdp marginales son

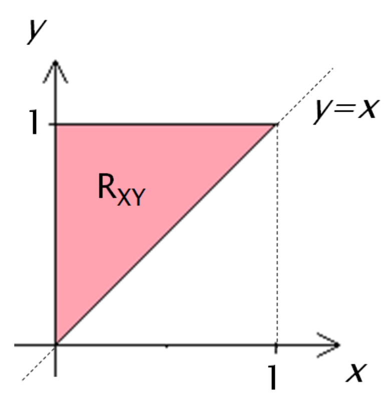
$$f_X(x)=\begin{cases} e^{-x} & 0\le x<\infty \\ 0 & si\ no \end{cases} \qquad f_Y(y)=\begin{cases} e^{-y} & 0\le y<\infty \\ 0 & si\ no \end{cases}$$

Vemos efectivamente que

$$f(x,y)=f_X(x).f_Y(y)=\begin{cases} e^{-(x+y)}=e^{-x}e^{-y} & 0\le x<\infty, 0\le y<\infty \\ 0 & si\ no \end{cases}$$

Es decir *X e Y son v.a. independientes*

Ejercicio



Sea una v.a. bidimensional continua con la siguiente *fdp* conjunta

$$f(x,y)=\begin{cases} 8xy & 0\le x\le y<1 \\ 0 & si\ no \end{cases}$$

Verificar que X e Y no son independientes.

Solución: Calculamos las *fdp* marginales de X e Y

$$f_X(x)=\begin{cases} \int_x^1 8\,x\,y\,dy=4x(1-x^2) & 0\le x\le 1 \\ 0 & si\ no \end{cases} \qquad f_Y(y)=\begin{cases} \int_0^y 8\,x\,y\,dx=4y^3 & 0\le y\le 1 \\ 0 & si\ no \end{cases}$$

Podemos ver que $f(x,y)=8xy\neq 16x(1-x^2)y^3 = f_X(x).f_Y(y)$

$\therefore$ ***X e Y son variables aleatorias dependientes***

Función de una v.a. bidimensional

Ejemplo: Sean X e Y v.a. aleatorias que representan las medidas de una pieza seleccionada al azar. Ej: X=ancho de la pieza e Y=longitud de la pieza, las siguientes variables también son variables aleatorias

Z=2X+2Y es el perímetro de la pieza

W= X.Y es el área de la pieza

Esperanza de una función

Sean X e Y dos v.a. aleatorias distribuidas de manera conjunta con función de probabilidad conjunta *p(x,y)* si son discretas o fdp conjunta *f(x,y)* si son continuas. Entonces el valor esperado de una función *h(X,Y)* está dado por

$$E[h(X,Y)] = \begin{cases} \sum_x \sum_y h(x,y).p(x,y) & \text{Si X e Y son discretas} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y).f(x,y)\,dx\,dy & \text{Si X e Y son continuas} \end{cases}$$

Ejemplo 4.8

Cinco amigos compraron entradas para un concierto. Los boletos son para los asientos 1 a 5 de una fila determinada. Si distribuyen al azar los boletos entre los cinco ¿cuál es el número esperado de asientos que separan a dos amigos cualesquiera de los cinco?

A cada par de amigos los separa un determinado número de asientos (ej:2)

Sean X e Y los números de asiento del 1er. y 2do. individuo respectivamente.

R_{XY} = {(1,2), (1,3),(1,4), (1,5), (2,1),(2,3),(2,4),(2,5), (3,1), (3,2), (3,4), (3,5), (4,1), (4,2), (4,3), (4,5), (5,1),(5,2),(5,3),(5,4)}

Todos los pares son igualmente probables, entonces

$$p(x,y) = \begin{cases} \dfrac{1}{20} & x = 1,\ldots,5;\ y = 1,\ldots,5;\ x \neq y \\ 0 & \textit{en caso contrario} \end{cases}$$

El nro. de asientos que separan a dos individuos es h(X,Y)=|X−Y|−1.

En la tabla se da h(x,y) para cada par posible (x,y)

h(x,y)	1	2	3	4	5
1	---	0	1	2	3
2	0	---	0	1	2
3	1	0	---	0	1
4	2	1	0	---	0
5	3	2	1	0	---

$$E[h(X,Y)] = \sum_x \sum_y h(x,y)\,p(x,y) \;\; = \sum_{x=1}^{5} \sum_{\substack{y=1 \\ x \neq y}}^{5} (|x-y|-1).\frac{1}{20} = 1$$

Ejemplo 4.9

En el ejemplo de la fdp conjunta de la cantidad X de almendras y la cantidad Y de castañas en una lata de 1 kg. de nueces

$$f(x,y) = \begin{cases} 24\,x\,y & 0 \le x \le 1,\; 0 \le y \le 1,\, x+y \le 1 \\ 0 & en\ caso\ contrario \end{cases}$$

Si un kilo de almendras le cuesta a la empresa 1$, 1 kg. de castañas cuesta $1.50, 1kg.de maní cuesta $0.50, indicar el valor esperado del contenido de la lata.
El costo del contenido de la lata puede expresarse como

h(X,Y)=(1)X+(1.5)Y+(0.50)(1−X−Y)= 0.5+0.5X+Y

El costo total esperado es

$$E[h(x,y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y).\,f(x.y)\,dx\,dy$$

$$= \int_0^1 \int_0^{1-x} (0.5 + 0.5x + y).24xy\,dx\,dy = \$1.10$$

Esperanza de una suma de v.a.

Sean X e Y dos variables aletorias arbitrarias entonces $E(X+Y) = E(X) + E(Y)$

Generalizando $E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$

o también $$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

Esperanza de un producto de v.a.

En general ***la esperanza de un producto de variables aleatorias no es igual al producto de las esperanzas***.
Si (X,Y) *es una v.a. bidimensional tal que X e Y son variables aleatorias* ***independientes***, entonces

$$E(X.Y) = E(X).E(Y)$$

Varianza de una suma de v.a.

$$V(X+Y)=V(X)+V(Y)+2\sigma_{XY}$$

con $\quad \sigma_{XY}=E(X.Y)-E(X).E(Y)$

A $\boldsymbol{\sigma_{XY}}$ se la llama **covarianza** de X e Y

Covarianza

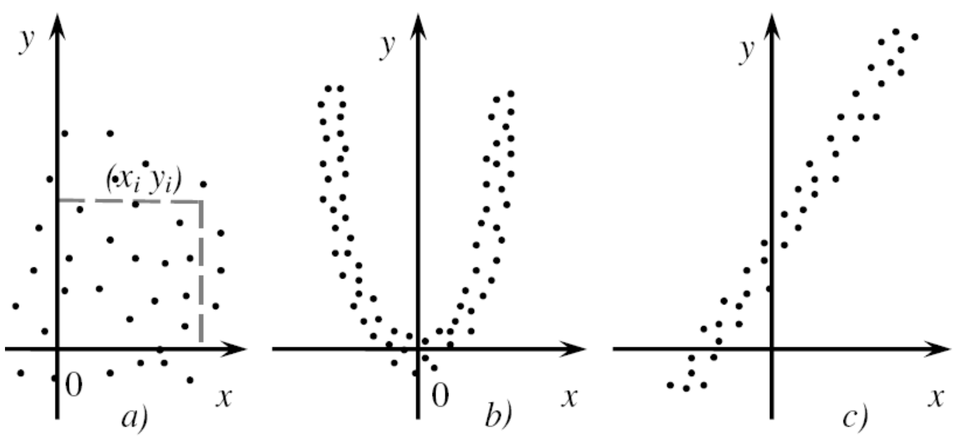
Cuando dos v.a.X e Y no son independientes suele ser interesante evaluar cuan estrecha es la relación entre si. La **covarianza** entre dos v.a. X e Y es

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

$$=\begin{cases} \sum_x\sum_y (x-E(X))(y-E(Y))\,p(x.y) & \text{(X e Y discretas)} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-E(X))(y-E(Y))\,f(x.y)dx\,dy & \text{(X e Y continuas)} \end{cases}$$

Diagramas de Dispersión

Consiste en dibujar pares de valores (x_i, y_j) medidos de la v.a. (X,Y) en un sistema de coordenadas

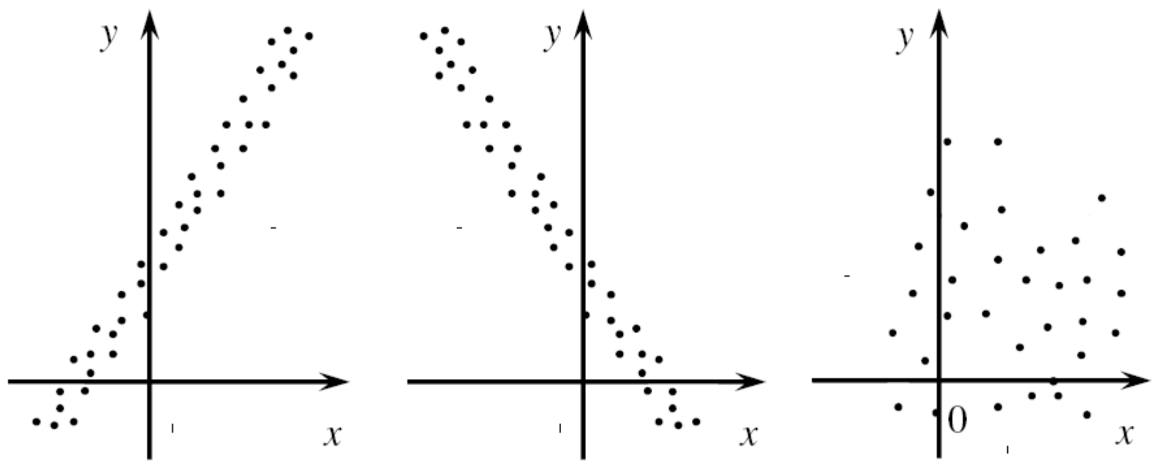


a) Entre X e Y no hay ninguna relación funcional

b) Entre X e Y podría existir un relación funcional que corresponde a una parábola

c) Entre X e Y existe una **relación lineal**. Este es el tipo de relación que nos interesa

Covarianza



Covarianza Positiva | Covarianza Negativa | Covarianza cercana a cero

Ejemplo

Dadas las variables X e Y con la sigte.distribución

p(x,y)	***0***	***100***	***200***	$p_X(x)$
100	*0.20*	*0.10*	*0.20*	0.5
250	*0.05*	*0.5*	*0.30*	0.5
$p_Y(y)$	0.25	0.25	0.50	

Calcular **Cov(X,Y)**

$$Cov(X,Y) = \sum_x \sum_y (x - E(X))(y - E(Y))\, p(x,y)$$

$$E(X) = \sum_x x.p_X(x) = 100*0.5 + 250*0.5 = 175$$

$$E(Y) = \sum_y y.p_Y(y) = 0 + 100*0.25 + 200*0.5 = 125$$

$$\begin{aligned} Cov(X,Y) = {} & (100-175)(0-125)0.2 + (100-175)(100-125)0.1 + \\ & (100-175)(200-125)0.2 + (250-175)(0-125)0.05 + \\ & (250-175)(100-125)0.15 + (250-175)(200-125)0.30 \\ & = 1875 \end{aligned}$$

Covarianza

La **covarianza** entre dos v.a. X e Y también puede expresarse así

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

Si X e Y son v.a. independientes entonces Cov(X,Y)=0

Ejemplo

Las fdp conjunta y marginales de X=cantidad de almendras e Y=cant.de castañas fueron

$$f(x,y) = \begin{cases} 24\,x\,y & 0 \le x \le 1,\ 0 \le y \le 1,\ x + y \le 1 \\ 0 & en\ caso\ contrario \end{cases}$$

$$f_X(x) = \begin{cases} 12\,x(1-x)^2 & 0 \le x \le 1 \\ 0 & si\ no \end{cases} \qquad f_Y(y) = \begin{cases} 12\,y(1-y)^2 & 0 \le y \le 1 \\ 0 & si\ no \end{cases}$$

Calcular Cov(X,Y) a partir de E(X.Y)–E(X).E(Y)

$$E(X) = \int_0^1 x.f_X(x)dx = \int_0^1 x.12.x(1-x)^2 = \frac{2}{5}$$

En este ejemplo E(Y) = E(X)

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x.y.f(x,y) = \int_0^1\int_0^{1-x} 24x^2y^2\,dx\,dy = 8\int_0^1 x^2(1-x)^3\,dx = \frac{2}{15}$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{2}{15} - \left(\frac{2}{5}\right)\left(\frac{2}{5}\right) = \frac{2}{15} - \frac{4}{25} = -\frac{2}{75}$$

Es razonable que la covarianza sea negativa porque más almendras en la lata significan menos castañas.

Propiedades de la Covarianza

Las siguientes propiedades pueden ser de utilidad

- $Cov(a + bX, c + dY) = bdCov(X,Y)$
- $Cov(X + Y, Z) = Cov(X,Z) + Cov(Y,Z)$
- $Cov(X,X) = V(X)$

Defecto de la Covarianza

Es importante notar que el valor calculado de la covarianza depende de las unidades de medición. En el ejemplo del seguro: Cov(X,Y)=1875. Pero, si la cantidad deducible se expresó en centavos y no en pesos, entonces 100X reemplazaría a X y 100Y reemplazaría a Y y por lo tanto Cov(100X,100Y)=100*100*Cov(X,Y)=18750000

La elección de las unidades no debe tener ningún efecto en la medida de la fuerza de la relación

Coeficiente de correlación lineal

Sea (X,Y) una v.a.bidimensional definimos el coeficiente de correlación lineal entre X e Y, representado por *Corr(X,Y)*, $\rho_{X,Y}$ o sólo ρ se define como

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}$$

Propiedad 1

Si a y c son positivas o negativas

Corr(a X + b, c Y + d) = Corr(X,Y)

Es decir que el coeficiente de correlación no resulta afectado por un cambio lineal en las unidades de medición. Esto corrige el defecto de la Covarianza.

Propiedad 2

Para dos v.a. aleatorias cualesquiera X e Y $\quad -1 \leq \rho_{X,Y} \leq 1$

En general se dice que la relación es

- <u>Fuerte</u> si $|\rho| \geq 0.8$
- <u>Moderada</u> si $0.5 < |\rho| < 0.8$
- <u>Débil</u> si $|\rho| \leq 0.5$

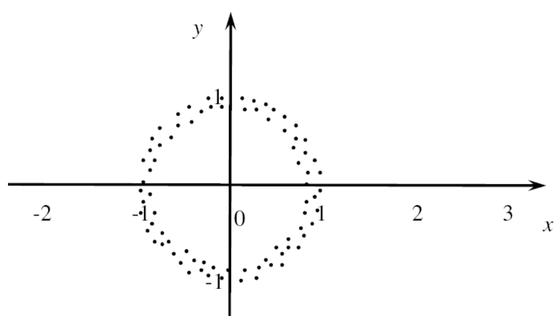
Propiedad 3

Si X e Y son independientes entonces $\rho=0$, pero $\rho=0$ **no implica independencia.**

ρ sólo es una medida del grado de relación **lineal** entre X e Y.

Que $\rho=0$ no significa que X e Y sean independientes sino que hay una ausencia completa de una relación lineal.

En este ejemplo $\rho_{X,Y}=0$ aunque puede verse que entre X e Y existe la siguiente relación $X^2+Y^2 = 1$



Propiedad 4

$\rho = 1$ o -1 si y sólo si **Y=aX + b** donde a y b son constantes y $a \neq 0$.

Sólo cuando X e Y tenga una relación lineal perfecta ρ será tan positiva o negativa como pueda ser.

Que $|\rho|<1$ sólo indica que la relación no es completamente lineal pero podría existir una relación no lineal muy fuerte.