Distribuciones de Probabilidad Conjunta

V.a. bidimensional

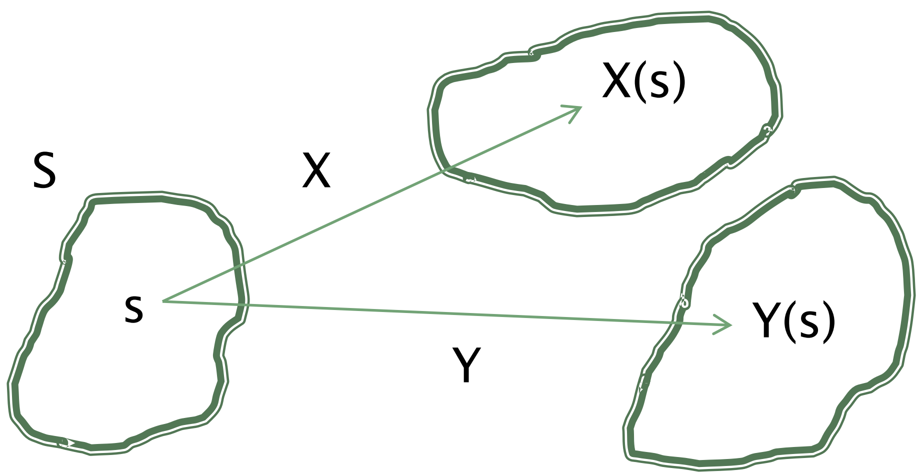
- Hasta ahora el resultado de un experimento era un único valor; es decir, las v.a. eran unidimensionales.

- 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.

V.a. bidimensional

- 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*

V.a. n-dimensional

- Sea ε un experimento y S su espacio muestral.
- Si $X_1=X_1(s)$, $X_2=X_2(s)$, …, $X_n=X_n(s)$ son *n* funciones, cada una de las cuales asigna un número real a cada resultado $s \in S$, entonces llamamos a

$$(X_1, X_2, \ldots, X_n)$$

una *variable aleatoria n-dimensional*.

V.a. bidimensionales

- 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

Finito

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

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

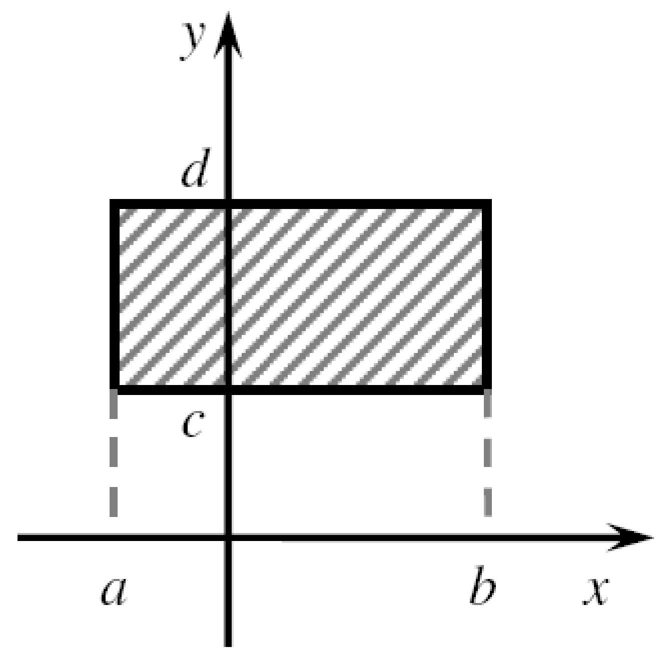
Infinito numerable

Recorridos no numerables

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

- **Ejemplo**

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

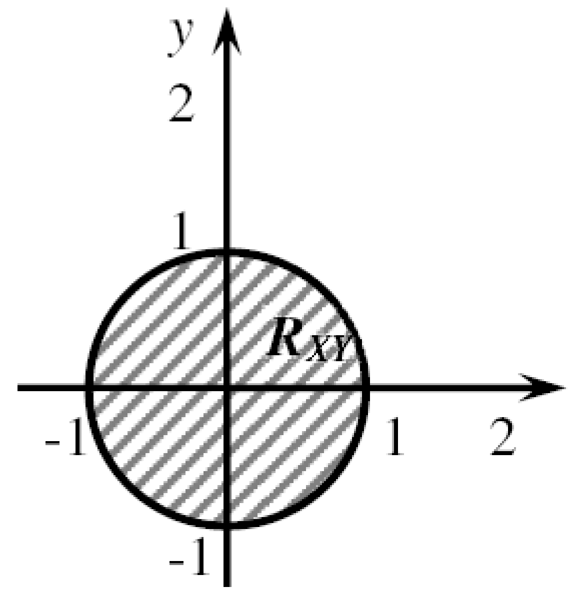


Recorridos no numerables

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

- **Ejemplo**

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

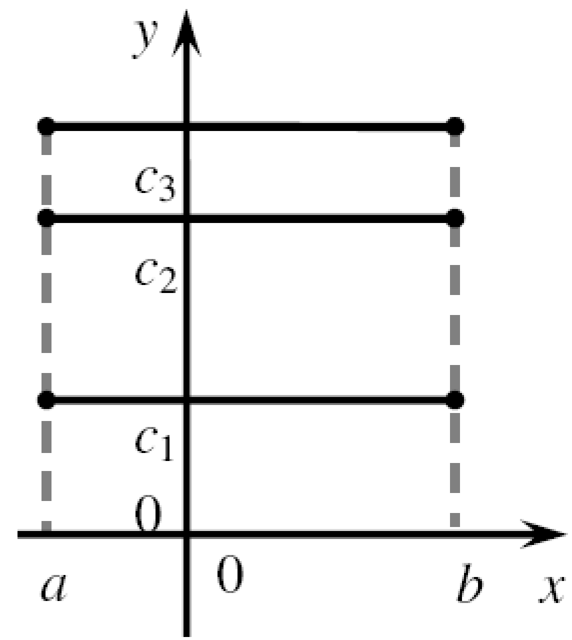


Recorridos no numerables

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

- **Ejemplo**

$R_{XY}=\{(x,y_j): a\leq x \leq b;\ y_j=c_1,c_2,c_3\}$



Este es un recorrido no numerable "mixto"

Prof. Laura Lanzarini

Función de probabilidad

- 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).*

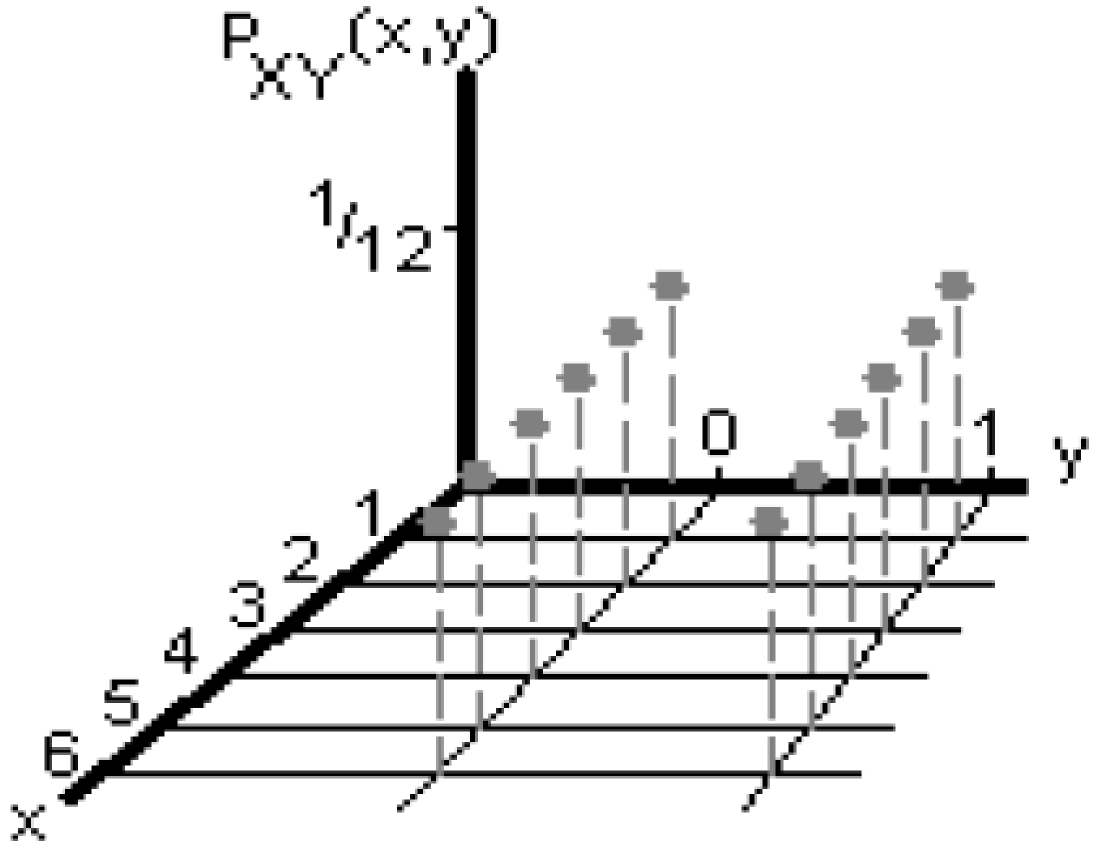
Prof. Laura Lanzarini

Ejemplo

- Cierto experimento consiste en arrojar un dado y una moneda honestos y observar qué número salió y si salió cara o no.
- Como resultado se obtendrá una v.a. bidimensional discreta (X,Y) donde
  - X: Nro. que sale al tirar un dado honesto.
  - Y: La cantidad de caras que salen al tirar una moneda.

Ejemplo

- X: Nro. que sale al tirar un dado honesto.
- Y: La cantidad de caras que salen al tirar una moneda.



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)*

Ejemplo 4.1

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

Y

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

- Es decir que

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

Ejemplo 4.1

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

Y

	p(x,y)	0	100	200
X	100	0.20	0.10	0.20
	250	0.05	0.15	0.30

$$P(Y \geq 100) = P(100,100) + P(100,200) + P(250,100) + P(250,200) =$$

$$0.1 + 0.2 + 0.15 + 0.3 = 0.75$$

Ejemplo 4.1

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

Y

p(x,y)	0	100	200
X 100	0.20	0.10	0.20
X 250	0.05	0.15	0.30

$$P(Y \geq X) = P(100,100) + P(100,200)$$

$$= 0.10 + 0.20 = 0.30$$

Ejercicio

- Dos líneas de producción fabrican cierto tipo de artículo.
- Supóngase que la capacidad (para un día cualquiera) es de 5 artículos para la línea I y 3 artículos para la línea II y que el número verdadero de artículos producidos por cada línea es una v.a.
- Sea (X,Y) la v.a. bidimensional que da el número de artículos producidos por la línea I y por la línea II respectivamente.

Ejercicio

- Esta es la distribución de probabilidad de conjunta de (X,Y)

Y \ X	0	1	2	3	4	5
0	0	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

Es una fdp conjunta válida?

Ejercicio

Y \ X	0	1	2	3	4	5
0	0	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

- Sea B={más artículos producidos por la línea I (v.a. X) que por la línea II (v.a. Y)}

$P(B) =$ Sumar $P(X=x_i,Y=y_j)$ con $y_j>x_i$

Ejercicio

Y \ X	0	1	2	3	4	5
0	0	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

- Sea B={más artículos producidos por la línea I (v.a. X) que por la línea II (v.a. Y)}

$$P(B) = \sum_{y_j > x_i} P(X = x_i; Y = y_j) = 0.75$$

Función de distribución marginal

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

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

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

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

Función de distribución marginal

- En el ejemplo de las pólizas de seguro

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

Ejercicio

- Calcular las distribuciones marginales de X y de Y

Y \ X	0	1	2	3	4	5	p(y)
0	0	0.01	0.03	0.05	0.07	0.09	0.25
1	0.01	0.02	0.04	0.05	0.06	0.08	0.26
2	0.01	0.03	0.05	0.05	0.05	0.06	0.25
3	0.01	0.02	0.04	0.06	0.06	0.05	0.24
p(x)	0.03	0.08	0.16	0.21	0.24	0.28	

Distribuciones Condicionales

- 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

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

Distribuciones Condicionales

- 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 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

Y

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

$$P(X = 250 \mid Y = 100) = \quad ?$$

Ejemplo 4.4

- Volviendo al ejemplo de las pólizas de seguro

		Y			
	p(x,y)	0	100	200	$p_X(x)$
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	

$$P(X=250 \mid Y=100) = \frac{P(X=250, Y=100)}{P(Y=100)}$$

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

Ejemplo 4.4

- Volviendo al ejemplo de las pólizas de seguro

Y

	p(x,y)	0	100	200	$p_X(x)$
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	

$$P(Y \geq 100 \mid X = 250) = \quad ?$$

Prof. Laura Lanzarini

Ejemplo 4.4

- Volviendo al ejemplo de las pólizas de seguro

		Y			
	$p(x,y)$	0	100	200	$p_X(x)$
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	

$$P(Y \geq 100 \mid X = 250) = \frac{P(Y \geq 100\,;\, X = 250)}{P(X = 250)} =$$

$$\frac{(0.15 + 0.30)}{0.5} = \frac{0.45}{0.5} = 0.9$$

Ejercicio

Y \ X	0	1	2	3	4	5	p(y)
0	0	0.01	0.03	0.05	0.07	0.09	0.25
1	0.01	0.02	0.04	0.05	0.06	0.08	0.26
2	0.01	0.03	0.05	0.05	0.05	0.06	0.25
3	0.01	0.02	0.04	0.06	0.06	0.05	0.24
p(x)	0.03	0.08	0.16	0.21	0.24	0.28	

$$P(Y=2 \mid X=5) = \frac{P(X=5\,;Y=2)}{P(X=5)} = \frac{0.06}{0.25} = 0.24$$

V.a. 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?

Ejemplo 4.6

X e Y son independientes

- Para demostrar que son independientes debe verificarse $p(x_i, y_j) = p_X(x_i) \,.\, p_Y(y_j) \ \forall (x_i, y_j) \in R_{XY}$

$$p(0,0)=0.1 = p_X(0).p_Y(0)=0.2*0.5$$
$$p(0,1)=0.04 = p_X(0).p_Y(1)=0.2*0.2$$
$$p(0,2)=0.06 = p_X(0).p_Y(2)=0.2*0.3$$
$$p(1,0)=0.2 = p_X(1).p_Y(0)=0.4*0.5$$
$$p(1,1)=0.08 = p_X(1).p_Y(1)=0.4*0.2$$
$$p(1,2)=0.12 = p_X(1).p_Y(2)=0.4*0.3$$
$$p(2,0)=0.2 = p_X(2).p_Y(0)=0.4*0.5$$
$$p(2,1)=0.08 = p_X(2).p_Y(1)=0.4*0.2$$
$$p(2,2)=0.12 = p_X(2).p_Y(2)=0.4*0.3$$

Función de una v.a. bidimensional

- Ejemplo: Sean X e Y v.a. aleatorias que representan las medidas de una pieza seleccionada al azar
  - X=ancho de la pieza
  - Y=longitud de la pieza

  - Z=2X+2Y es el perímetro de la pieza
  - W= X.Y es el área de la pieza

Z y W también son variables aleatorias

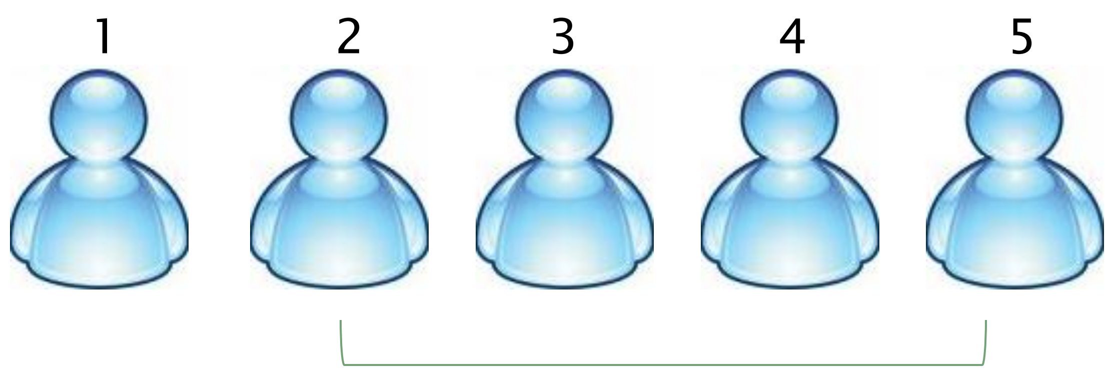
Esperanza de una función

- 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)

Prof. Laura Lanzarini

Ejemplo 4.8

- 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

$$p(x,y)=\begin{cases}\dfrac{1}{20} & x=1,\ldots,5;\ y=1,\ldots,5;\ x\neq y\\[2ex] 0 & en\ caso\ contrario\end{cases}$$

Ejemplo 4.8

- El nro. de asientos que separan a dos individuos es $h(X,Y)=|X-Y|-1$.
- En la siguiente 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	---

Ejemplo 4.8

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	---

$$E[h(X,Y)] = \sum_{x}\sum_{y} h(x,y)\,p(x,y)$$

$$= \sum_{x=1}^{5}\sum_{\substack{y=1 \\ x\neq y}}^{5} (|x-y|-1).\frac{1}{20} = 1$$

Esperanza de una suma de v.a.

- Sean X e Y dos variables aletorias arbitrarias entonces

$$E(X+Y)=E(X)+E(Y)$$

Generalizando

$$E(X_1+X_2+...+X_n)=E(X_1)+E(X_2)+...+E(X_n)$$

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

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

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

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

$$Cov(X,Y) = \begin{cases} \sum_x \sum_y (x - E(X))(y - E(Y))\, p(x,y) \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - E(X))(y - E(Y))\, f(x,y)\, dx\, dy \end{cases}$$

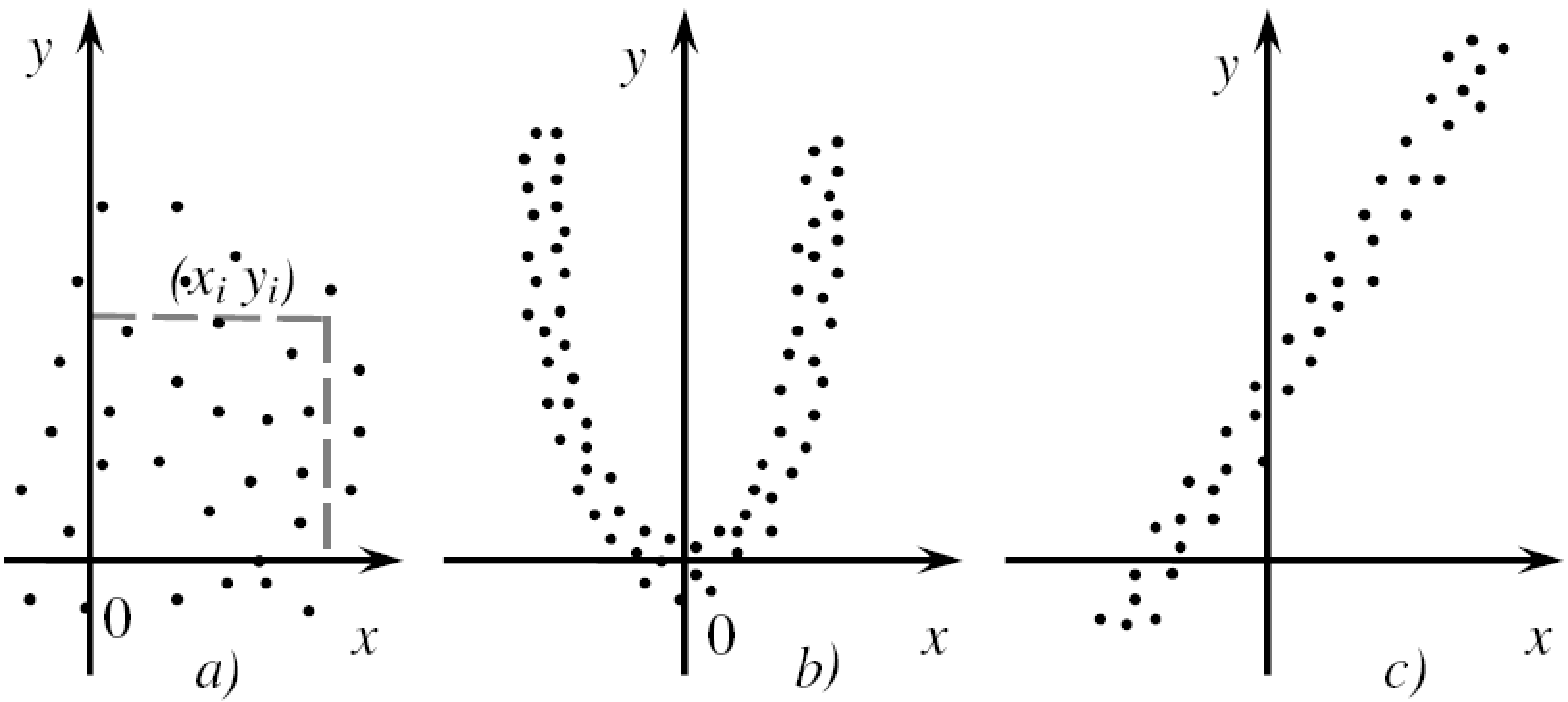
X e Y discretas

X e Y continuas

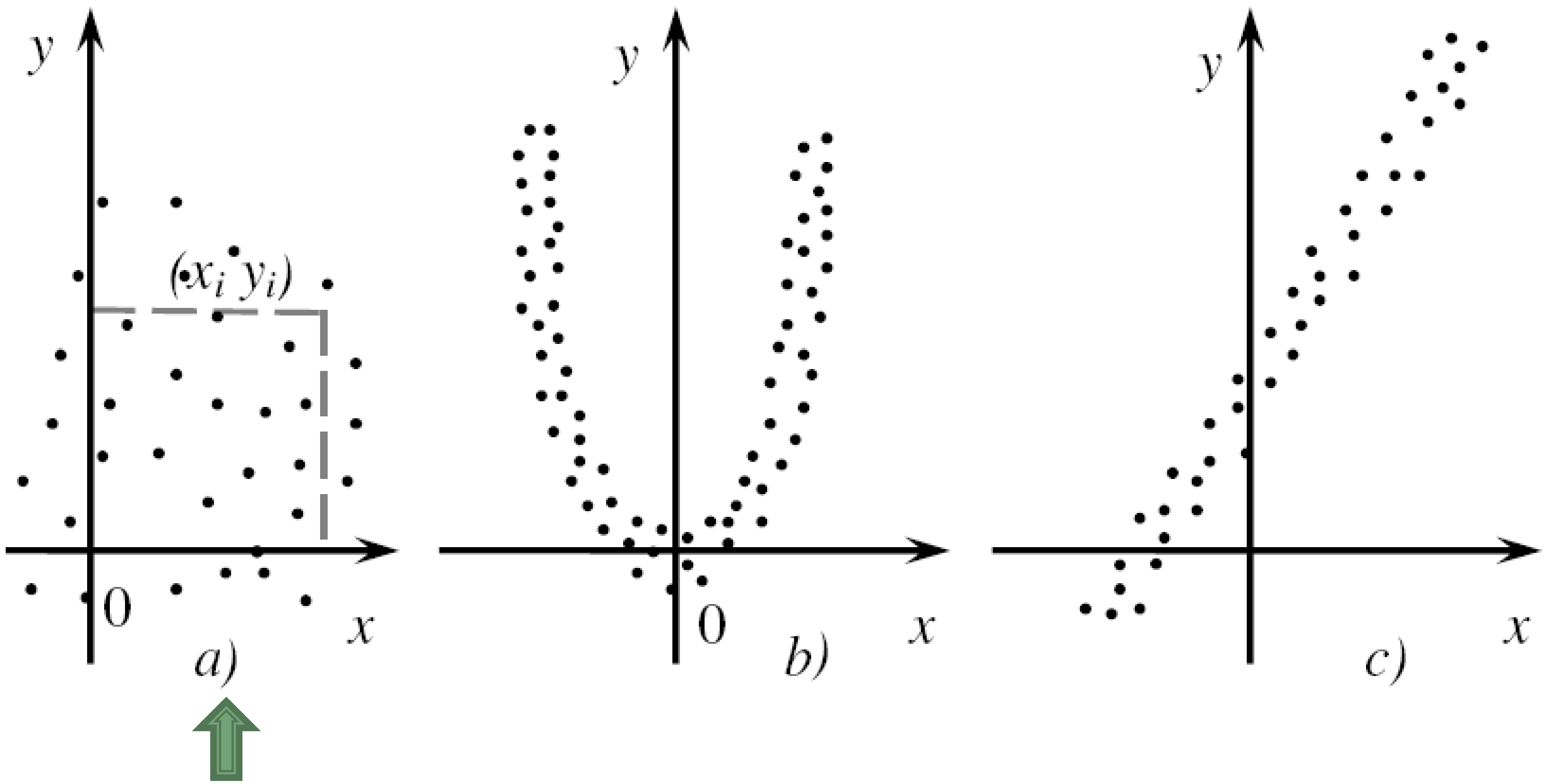
Prof. Laura Lanzarini

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

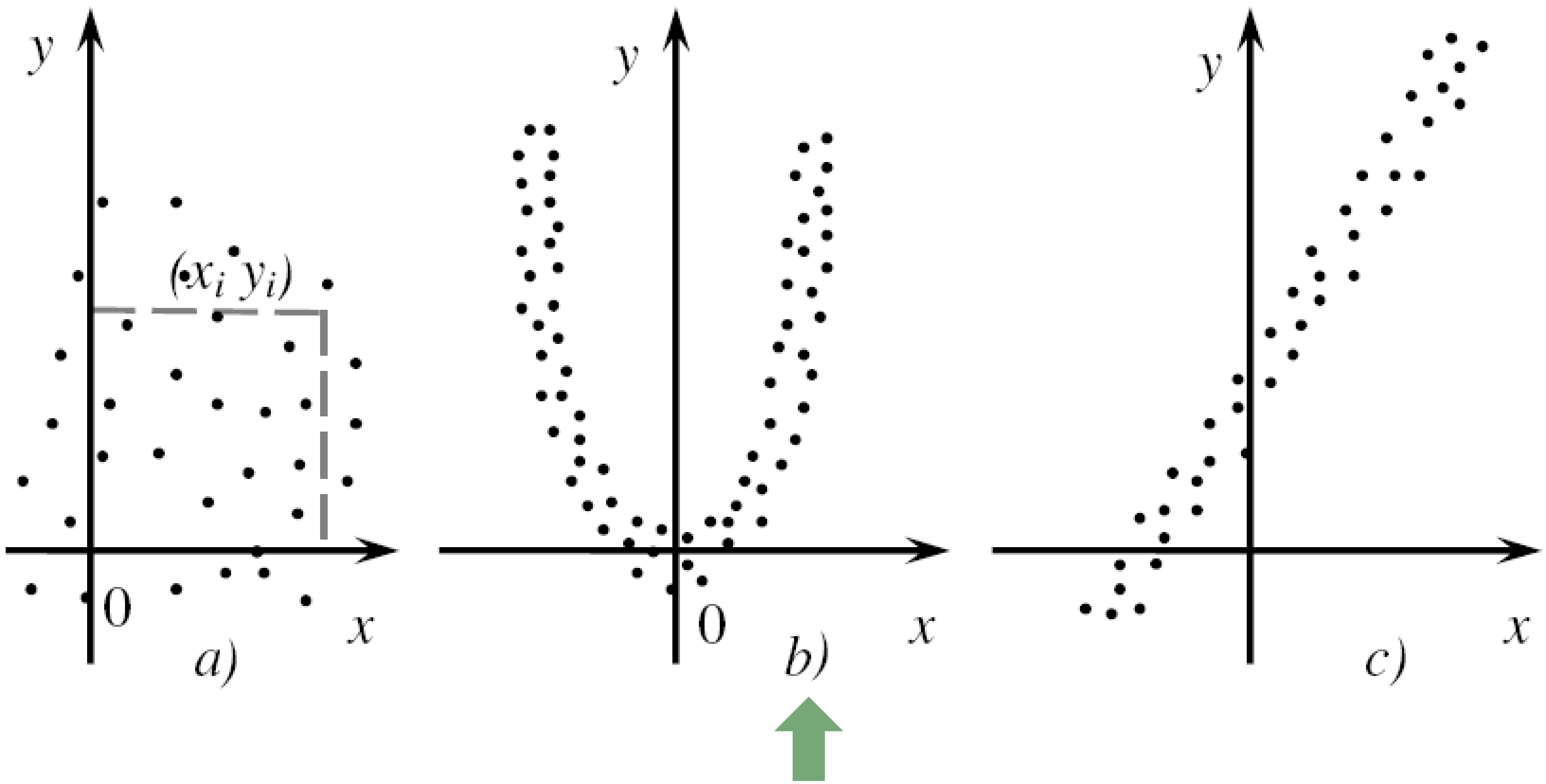


Diagramas de Dispersión



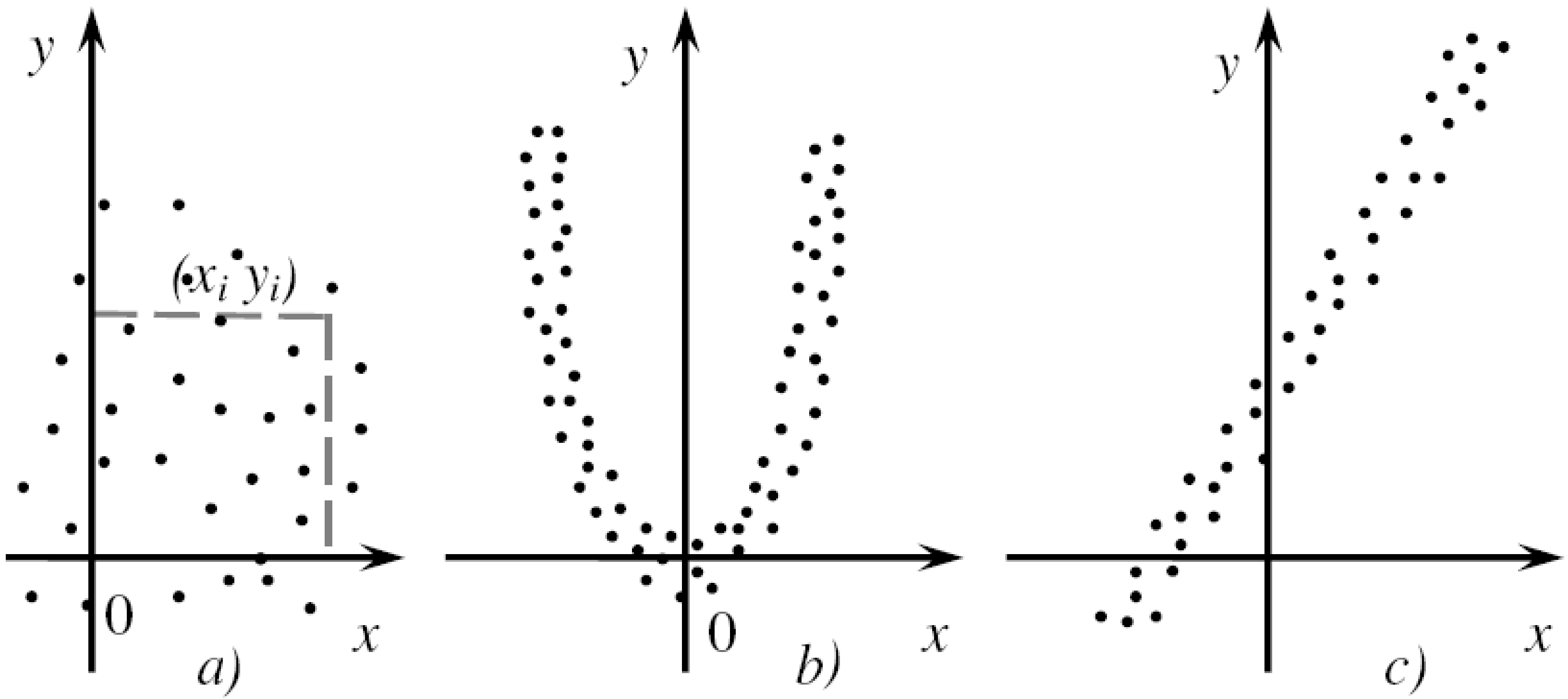
a) Entre X e Y no hay ninguna relación funcional

Diagramas de Dispersión



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

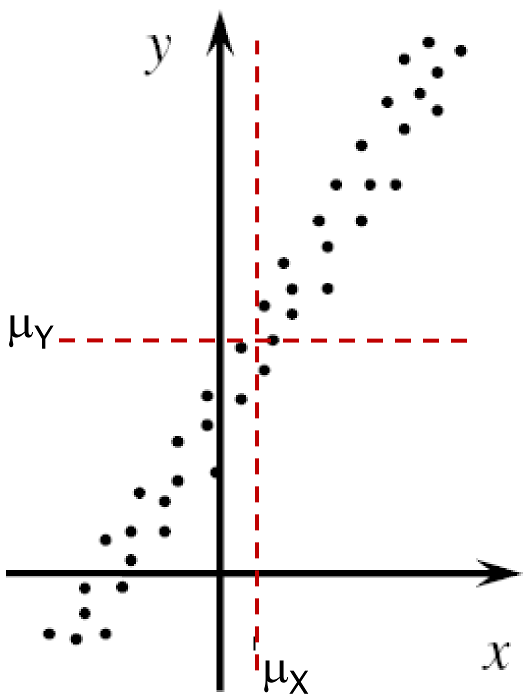
Diagramas de Dispersión



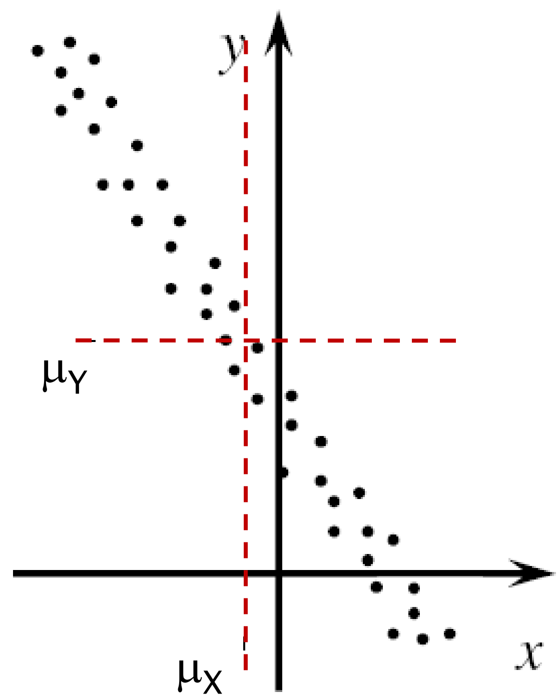
c) Entre X e Y existe una **relación lineal**.
Este es el tipo de relación que nos interesa

Covarianza

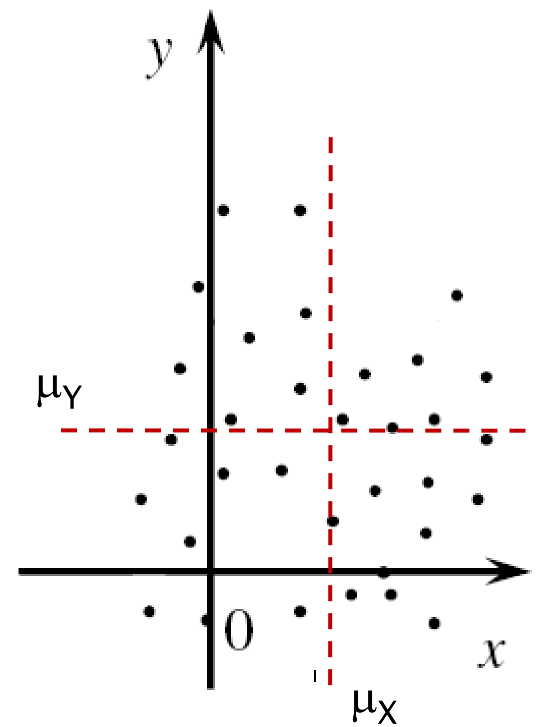
$$Cov(X,Y) = \sum_{x}\sum_{y}(x - E(X))(y - E(Y))\,p(x.y)$$



Covarianza Positiva



Covarianza Negativa



Covarianza cercana a cero

Ejemplo

- Dadas las variables X e Y con la sigte.distribución

Y

X

$p(x,y)$	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	

- Calcular **Cov(X,Y)**

Ejemplo

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	

(Rows: X; columns: Y)

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

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

Prof. Laura Lanzarini

Ejemplo

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	

(Filas: X; columnas: Y)

$$Cov(X,Y) = \sum_x \sum_y (x-175)(y-\boxed{E(Y)})\, p(x,y)$$

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

Ejemplo

		Y			
	p(x,y)	0	100	200	$p_X(x)$
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	

$$Cov(X,Y) = \sum_x \sum_y (x-175)(y-125)\,p(x,y)$$

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

Covarianza

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

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

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

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

$$Cov(X,Y) = E(X.Y) - E(X).E(Y)$$

Ejemplo

Y

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

X

$p_Y(y)$	0.25	0.25	0.50

E(X)=175

E(Y) = 125

$$E(X,Y) = \sum x_i * y_j * P(X = x_i ; Y = y_j) =$$

$$= (100*0*0.20 + 100*100*0.1 + 100*200*0.2 +$$

$$250*0*0.05 + 250*100*0.15 + 250*200*0.30) = 23750$$

$$Cov(X,Y) = E(X.Y) - E(X).E(Y) =$$

$$23750 - 175*125 = 1875$$

Prof. Laura Lanzarini

Propiedades de la Covarianza

- Las siguientes propiedades pueden ser de utilidad
  - $$Cov(a+bX,c+dY)=b\,d\,Cov(X,Y)$$
  - $$Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)$$
  - $$Cov(X,X)=V(X)$$

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

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

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

Ejemplo

$E(X) = 175 \; ; \; E(Y) = 125$

$Cov(X,Y) = 1875$

		Y			
	p(x,y)	0	100	200	$p_X(x)$
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	

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

Falta calcular los valores de las desviaciones de cada v.a.

Ejemplo

$E(X) = 175$; $E(Y) = 125$

$Cov(X,Y) = 1875$

		Y			
	p(x,y)	0	100	200	$p_X(x)$
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	

$$V(X) = (100 - 175)^2 * 0.5 + (200 - 175)^2 * 0.5 = 5625$$

$$\sigma_X = \sqrt{V(X)} = \sqrt{5625} = 75$$

Ejemplo

E(X) = 175 ; E(Y) = 125

Cov(X,Y) = 1875

Y

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

$$\sigma_X = 75$$

$$V(Y) = (0-125)^2 * 0.25 + (100-125)^2 * 0.25 + (200-125)^2 * 0.5 = 6875$$

$$\sigma_Y = \sqrt{V(Y)} = \sqrt{6875} = 82.9156$$

Ejemplo

E(X) = 175 ; E(Y) = 125

Cov(X,Y) = 1875

		Y			
	p(x,y)	0	100	200	$p_X(x)$
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	

$$\sigma_X = 75$$

$$\sigma_Y = 82.9156$$

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{1875}{75*82.9156}$$

$$\rho_{X,Y} = 0.3015$$

Prof. Laura Lanzarini

Coeficiente de correlación lineal

Propiedad 1

Si a y c son positivas o negativas

$$\mathbf{Corr(\ a\,X + b,\ c\,Y + d) = Corr(X,Y)}$$

- 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.

Coeficiente de correlación lineal

Propiedad 2

Para dos v.a. aleatorias cualesquiera X e Y

$$-1 \leq \rho_{X,Y} \leq 1$$

- En general se dice que la relación es
  - Fuerte si $|\rho| \geq 0.8$
  - Moderada si $0.5 < |\rho| < 0.8$
  - Débil si $|\rho| \leq 0.5$

Coeficiente de correlación lineal

Propiedad 3

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

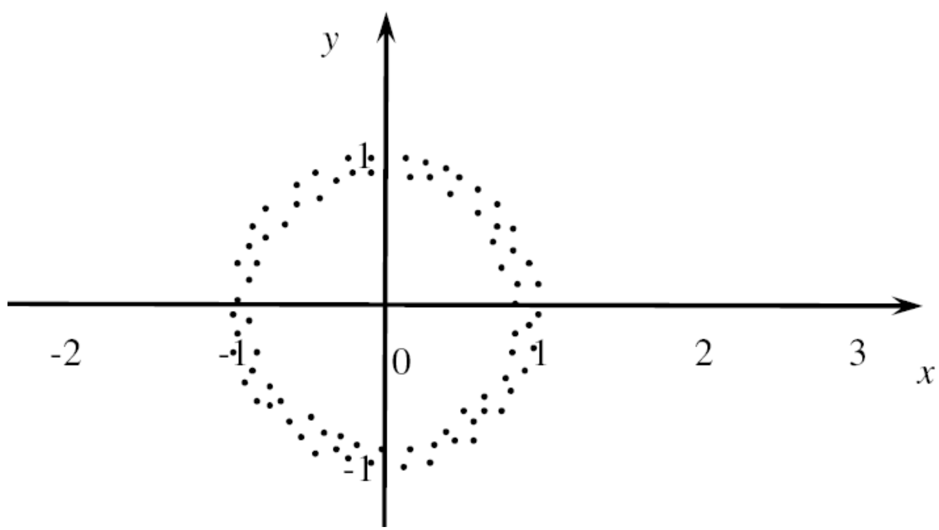
- ρ sólo es una medida del grado de relación **lineal** entre X e Y.
- Que $\rho=0$ no significa que X e Y sean independientes sino que hay una ausencia completa de una relación lineal.

Prof. Laura Lanzarini

Coeficiente de correlación lineal

Propiedad 3

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



Coeficiente de correlación lineal

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.

Resumen

- **V.A.bidimensional**
  - Recorrido
    - Numerable y no numerable
- **V.A.Bidimensional Discreta**
  - Func.de Prob.Conjunta
  - Funciones Marginales
  - Distrib.Condicionales
  - V.A.Independientes
- **Esperanza**
  - De una función
  - De una suma de v.a.
  - De un producto de v.a.
- **Varianza de una suma**
- **Covarianza**
  - Propiedades
  - Defecto
- **Coef.de correlación**
  - Propiedades