TEMA 5 - CARACTERISTICAS DE LA V.A. BIDIMENSIONAL

Esperanza de una función de varias variables
Sea (X,Y) una v.a. bidimensional y $g(X,Y)$ una función de la misma $g: \mathbb{R}^2 \longrightarrow \mathbb{R}$. Se define la esperanza de $g(X,Y)$ como:

a) Si (X,Y) es v.a. discreta con función de probabilidad conjunta $P[X=x_i, Y=y_j]$

$$E[g(X,Y)] = \sum_i g(x_i, y_j) \cdot P[X=x_i, Y=y_j]$$

b) Si (X,Y) es v.a. continua con función de densidad conjunta $f(x,y)$

$$E[g(X,Y)] = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} g(x,y) \cdot f(x,y)\, dy \right) dx$$

Veamos un ejemplo de esto (ejercicio 19)
Consideremos la v.a. bidimensional con la siguiente distribución de probabilidad conjunta:

X\Y	0	1
0	a	b
1	c	d

Obtenga $E[X \cdot Y]$

$$E[X\cdot Y] = \sum x_i \cdot y_i \cdot P[X=x_i, Y=y_j] = 0\cdot 0 \cdot P[X=0, Y=0] +$$

$$+ 0 \cdot 1 \cdot P[X=0, Y=1] + 1 \cdot 0 \cdot P[X=1, Y=0] + 1 \cdot 1 \cdot P[X=1, Y=1] = d$$

Ejemplo (ejercicio 18)
Sea (X,Y) una v.a. bidimensional con función de densidad conjunta

$$f(x,y) = \begin{cases} x+y & \text{si } 0<x<1,\ 0<y<1 \\ 0 & \text{en el resto} \end{cases}$$

Obtenga $E[X \cdot Y]$

$$E[X \cdot Y] = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} x \cdot y\, f(x,y)\, dy \right) dx = \int_0^1 \left(\int_0^1 x \cdot y \cdot (x+y)\, dy \right) dx$$

$$\int xy(x+y)\, dy = \int x^2 y + xy^2\, dy = \frac{x^2y^2}{2} + \frac{x\,y^3}{3}$$

$$\int_0^1 xy(x+y)\, dy = \left[\frac{x^2y^2}{2} + \frac{xy^3}{3} \right]_0^1 \quad \frac{x^2}{2} + \frac{x}{3}$$

$$E[X \cdot Y] = \int_0^1 \frac{x^2}{2} + \frac{x}{3}\, dx = \left[\frac{x^3}{6} + \frac{x^2}{6} \right]_0^1 = \frac{1}{3}$$

Las propiedades de la esperanza

Las propiedades de la esperanza en v.a. bidimensional son muy parecidas a las correspondientes propiedades en v.a. unidimensional. De nuevo te advierto que las demostraciones no te las tienes que estudiar, es un recurso que has de aprender a manejar.

1º Si $P[g(x,y) \geq 0] = 1 \Rightarrow E[g(x,y)] \geq 0$

Demostración:

Si $P[g(x,y) \geq 0] = 1 \Rightarrow P[g(x,y) < 0] = 0 \Rightarrow g(x,y) < 0$ es el suceso imposible $\Rightarrow g(x,y) \geq 0 \ \forall (x,y)$

a) Caso discreto

$E[g(x,y)] = \sum_{i,j} g(x_i, y_j) \cdot P[X = x_i, Y = y_j]$. Como $g(x_i, y_j) \geq 0$,

$P[X = x_i, Y = y_j] \geq 0 \Rightarrow g(x_i, y_j) \cdot P[X = x_i, Y = y_j] \geq 0 \Rightarrow$

$\Rightarrow \sum_{i,j} g(x_i, y_j) \cdot P[X = x_i, Y = y_j] \geq 0 \Rightarrow E[g(x,y)] \geq 0$

b) Caso continuo

Como $g(x,y) \geq 0$, $f(x,y) \geq 0 \Rightarrow g(x,y) \cdot f(x,y) \geq 0 \Rightarrow$

$\Rightarrow \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} g(x,y) \cdot f(x,y)\, dy \right) dx \geq 0 \Rightarrow E[g(x,y)] \geq 0.$

2ª Si $g(X,Y) \leq h(X,Y) \Rightarrow E[g(X,Y)] \leq E[h[X,Y]]$

Demostración

a) Caso discreto

Si $g(x,y) \leq h(x,y) \Rightarrow g(x_i, y_j) \leq h(x_i, y_j) \Rightarrow$

$\Rightarrow g(x_i, y_j) \cdot P[X=x_i, Y=y_j] \leq h(x_i, y_j) \cdot P[X=x_i, Y=y_j]$

$\Rightarrow \sum_i \sum_j g(x_i, y_j) \cdot P[X=x_i, Y=y_j] \leq \sum_i \sum_j h(x_i, y_j) \cdot P[X=x_i, Y=y_j)$

$\Rightarrow E[g(x,y)] \leq E[h(x,y)]$

b) Caso continuo

Si $g(x,y) \leq h(x,y)$ al ser $f(x,y) \geq 0 \Rightarrow g(x,y) \cdot f(x,y) \leq h(x,y) \cdot f(x,y)$

$\Rightarrow \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} g(x,y) \cdot f(x,y)\, dy \right) dx \leq \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} h(x,y) \cdot f(x,y)\, dy \right) dx \Rightarrow$

$\Rightarrow E[g(x,y)] \leq E[h(x,y)]$

3ª $E[a+b\,g(x,y)] = a + b\,E[g(x,y)]$

Demostración:

a) Caso discreto

$E[a+b\,g(x,y)] = \sum_i \sum_j (a+b\,g(x_i, y_j)) \cdot P[X=x_i, Y=y_j] =$

$= \sum_i \sum_j a \cdot P[X=x_i, Y=y_j] + \sum_i \sum_j b \cdot g(x_i, y_j) \cdot P[X=x_i, Y=y_j]$

$= a \sum_i \sum_j P[X=x_i, Y=y_j] + b \sum_i \sum_j g(x_i, y_j) \cdot P[X=x_i, Y=y_j] =$

$= a \cdot 1 + b \cdot E[g(x,y)] = a + b \cdot E[g(x,y)]$

b) Caso continuo

$E[a+b\,g(x,y)] = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} [a+b\,g(x,y)] \cdot f(x,y)\, dy \right) dx =$

$= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} a \cdot f(x,y)\, dy \right) dx + \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} b\,g(x,y) \cdot f(x,y)\, dy \right) dx$

es decir:

$$E[a+b\,g(x,y)] = a\cdot\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} f(x,y)\,dy\right)dx + b\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} g(x,y)f(x,y)\,dy\right)dx =$$

$$= a\cdot 1 + b\cdot E[g(x,y)] = a + b\cdot E[g(x,y)]$$

4ª $E[g(x,y) + h[x,y)] = E[g(x,y)] + E[h(x,y)]$

Demostración

a) Caso discreto

$$E[g(x,y) + h(x,y)] = \sum_i \sum_j [g(x_i,y_j) + h(x_i,y_j)]\cdot P[X=x_i, Y=y_j] =$$

$$= \sum_i \sum_j g(x_i,y_j)\cdot P[X=x_i, Y=y_j] + h(x_i,y_j\cdot P[X=x_i, Y=y_j] =$$

$$= \sum_i \sum_j g(x_i,y_j)\cdot P[X=x_i, Y=y_j] + \sum_i \sum_j h(x_i,y_j)\cdot P[X=x_i, Y=y_j] =$$

$$= E[g(x,y)] + E[h(x,y)]$$

b) Caso continuo

$$E[g(x,y) + h(x,y)] = \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} [g(x,y) + h(x,y)]\cdot f(x,y)\,dy\right)dx =$$

$$= \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} g(x,y)\cdot f(x,y) + h(x,y)\cdot f(x,y)\,dy\right)dx =$$

$$= \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} g(x,y)\cdot f(x,y)\,dy\right)dx + \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} h(x,y)\cdot f(x,y)\,dy\right)dx =$$

$$= E[g(x,y)] + E[h(x,y)]$$

5ª $E[g(X) + h(Y)] = E[g(X)] + E[h(Y)]$

Demostración.

a) Caso discreto

$$E[g(X) + h[Y]] = \sum_i \sum_j (g(x_i) + h(y_j))\cdot P[X=x_i, Y=y_j] =$$

$$= \sum_i \sum_j g(x_i)\cdot P[X=x_i, Y=y_j] + \sum_i \sum_j h(y_j)\cdot P[X=x_i, Y=y_j]$$

$$E[g(X)+h(Y)] = \sum_i g(x_i)\cdot \sum_j P[X=x_i, Y=y_j] + \sum_j h(y_j) \sum_i P[X=x_i, Y=y_j]$$

$$= \sum_i g(x_i)\cdot P[X=x_i] + \sum_j h(y_j)\cdot P[Y=y_j] = E[g(X)] + E[h(Y)]$$

b) Caso continuo

$$E[g(X)+h(Y)] = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} [g(x)+h(Y)]\cdot f(x,y)\, dy \right) dx =$$

$$= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} g(x) f(x,y)\, dy \right) dx + \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} h[Y]\cdot f(x,y)\, dy \right) dx =$$

$$= \int_{-\infty}^{+\infty} g(x) \left(\int_{-\infty}^{+\infty} f(x,y)\, dy \right) dx + \int_{-\infty}^{+\infty} h(y)\cdot \left(\int_{-\infty}^{+\infty} f(x,y)\, dx \right) dy =$$

$$= \int_{-\infty}^{+\infty} g(x)\cdot f_1(x)\, dx + \int_{-\infty}^{+\infty} h(y)\cdot f_2(y)\, dy = E[g(x)] + E[h(y)]$$

6ª Si X e Y son v.a. independientes, entonces se verifica:

$$E[X\cdot Y] = E[X]\cdot E[Y]$$

Demostración

a) Caso discreto

$$E[X\cdot Y] = \sum_i \sum_j x_i\cdot y_j\cdot P[X=x_i, Y=y_j]$$

Al ser X e Y independientes $P[X=x_i, Y=y_j] = P[X=x_i]\cdot P[Y=y_j]$

Por tanto:

$$E[X\cdot Y] = \sum_i \sum_j x_i\cdot y_j\cdot P[X=x_i]\cdot P[Y=y_j] = \sum_i x_i\cdot P[X=x_i] \sum_j y_j\cdot P[Y=y_j]$$

$$E[X\cdot Y] = E[X]\cdot E[Y].$$

b) Caso continuo

$E[X\cdot Y] = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} x\cdot y\, f(x,y)\, dy \right) dx$. Si X e Y son independientes entonces se verifica que $f(x,y) = f_1(x)\cdot f_2(y)$. Luego:

$$E[X\cdot Y] = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} x\cdot y\cdot f_1(x)\cdot f_2(y)\, dy \right) dx =$$

$$E[X\cdot Y] = \int_{-\infty}^{+\infty} x\cdot f_1(x) \left(\int_{-\infty}^{+\infty} y\cdot f_2(y)\, dy \right) dx = E[X]\cdot E[Y].$$

La propiedad 6ª se puede generalizar del siguiente modo:
Si X e Y son independientes entonces:

$$E[\,g(X)\cdot h(Y)\,] = E[\,g(X)\,]\cdot E[\,h(Y)\,]$$

Covarianza.

Dada la v.a. bidimensional (X,Y) la covarianza $Cov(X,Y)$ es la esperanza del producto de las desviaciones de X respecto a su media por las desviaciones de Y respecto a su media, es decir:

$$Cov(X,Y) = E[\,(X-\mu_X)\cdot(Y-\mu_Y)\,]$$

Propiedades de la covarianza

1ª $Cov(X,Y) = E[X\cdot Y] - E[X]\cdot E[Y]$

Demostración

$$Cov(X,Y) = E[\,(X-\mu_X)(Y-\mu_Y)\,] = E[XY - \mu_Y X - \mu_X Y + \mu_X\mu_Y] =$$

$$= E[XY] - \mu_Y E[X] - \mu_X E[Y] + \mu_X\mu_Y = E[XY] - \mu_Y\mu_X - \cancel{\mu_X\mu_Y} + \cancel{\mu_X\mu_Y}$$

luego $Cov(X,Y) = E[XY] - E[X]\cdot E[Y]$

2ª Si X e Y son independientes entonces $Cov(X,Y)=0$ pero si $Cov(X,Y)=0$ no tienen porqué ser X e Y independientes

Demostración

Si X e Y son independientes entonces $E[X\cdot Y] = E[X]\cdot E[Y]$

Como:

$$Cov(X,Y) = E[X\cdot Y] - E[X]\cdot E[Y] = E[X]\cdot E[Y] - E[X]\cdot E[Y] = 0$$

$\Rightarrow Cov(X,Y) = 0.$

3ª Si efectuamos las transformaciones lineales

$Z = a + bX$

$W = c + dY$, entonces:

$Cov(Z, W) = bd\, Cov(X, Y)$

Demostración:

$Cov(Z, W) = E[(Z - \mu_Z) \cdot (W - \mu_W)]$

Como $\mu_Z = a + b\mu_X \qquad \mu_W = c + d\mu_Y$

$Cov(Z, W) = E[(\cancel{a} + bx - \cancel{a} - b\mu_x)(\cancel{c} + dy - \cancel{c} - d\mu_y)] =$

$= E[b(x - \mu_x) \cdot d(y - \mu_y)] = E[bd(x - \mu_x)(Y - \mu_y)] =$

$= bd\, E[(x - \mu_x)(y - \mu_y)] = bd\, Cov(X, Y)$

4ª Sean $a, b \in \mathbb{R}$ dos constantes. Entonces:

$Var[aX + bY] = a^2 Var[X] + b^2 Var[Y] + 2ab\, Cov(X, Y)$

Demostración

$Var[ax + bY] = E[(ax + bY - \mu_{ax+bY})^2] = E[(ax + bY - a\mu_x - b\mu_y)^2]$

ya que $\mu_{ax+by} = E[ax + bY] = aE[X] + bE[Y] = a\mu_x + b\mu_y$

Luego:

$Var[ax + bY] = E[(a(x - \mu_x) + b(y - \mu_y))^2] =$

$= E[a^2(x - \mu_x)^2 + b^2(Y - \mu_y)^2 + 2ab(X - \mu_x)(Y - \mu_y)] =$

$= E[a^2(x - \mu_x)^2] + E[b^2(y - \mu_y)^2 + E[2ab(x - \mu_x)(y - \mu_y)] =$

$= a^2 E[(x - \mu_x)^2] + b^2 E[(y - \mu_y)^2] + 2ab E[(x - \mu_x)(y - \mu_y)] =$

$= a^2 \cdot Var[X] + b^2 Var[Y] + 2ab\, Cov(X, Y)$

5ª Sean X e Y dos v.a. independientes (o simplemente que $Cov(X, Y) = 0$) y sean $a, b \in \mathbb{R}$ constantes. Entonces:

$Var[aX + bY] = a^2 Var[X] + b^2 Var[Y]$

La demostración de la propiedad 5ª es trivial.

Observación: Casos particulares a tener en cuenta

- $Var[X+Y] = 1^2 Var[X] + 1^2 Var[Y] + 2\cdot 1\cdot 1\cdot Cov(X,Y)$ es decir:

$$Var[X+Y] = Var[X] + Var[Y] + 2\,Cov(X,Y)$$

- $Var[X-Y] = 1^2 Var[X] + (-1)^2 Var[Y] + 2\cdot 1(-1)\cdot Cov(X,Y)$ es decir:

$$Var[X-Y] = Var[X] + Var[Y] - 2\,Cov(X,Y)$$

Por lo tanto:

$$Var[X+Y] + Var[X-Y] = 2\,Var[X] + 2\,Var[Y]$$

Veamos un ejemplo

Ejemplo (ejercicio 19)

Consideremos dos v.a X, Y con la siguiente distribución de probabilidad conjunta:

X\Y	0	1
0	a	c
1	b	d

Obtenga $Cov(X,Y)$

La covarianza la podemos calcular mediante una de las dos expresiones siguientes:

a) $$Cov(X,Y) = \sum_i \sum_j (x_i - \mu_x)(y_j - \mu_y)\cdot P[X = x_i, Y = y_j]$$

b) $$Cov(X,Y) = E[X\cdot Y] - E[X]\cdot E[Y]$$

En cualquiera de las dos expresiones que utilicemos necesitamos los valores $E[X]$, $E[Y]$

La distribución marginal de X es:

x_i	$P[X=x_i]$
0	$a+c$
1	$b+d$

$\Rightarrow E[X] = 0(a+c)+1\cdot(b+d) = b+d$

La distribución marginal de Y es:

y_i	$P[X=x_i]$
0	$a+b$
1	$c+d$

$\Rightarrow E[Y] = 0\cdot(a+b)+1\cdot(c+d) = c+d$

Si empleamos para el cálculo de $Cov(X,Y)$ la primera de las expresiones:

$$Cov(X,Y) = \sum_i \sum_j (x_i-(b+d))(y_j-(c+d))\, P[X=x_i, Y=y_j] =$$

$$Cov(X,Y) = (0-(b+d))(0-(c+d))\,a + (0-(b+d))(1-(c+d))\,c +$$
$$+ (1-(b+d))\cdot(0-(c+d))\cdot c + (1-(b+d))(1-(c+d))\cdot d =$$
$$= d-(b+d)(c+d)$$

Si empleamos la segunda de las expresiones

$Cov(X,Y) = E[X\cdot Y] - E[X]\cdot E[Y]$ necesitamos el valor de $E[XY]$

$$E[XY] = \sum_i \sum_j x_i\cdot y_j\cdot P[X=x_i, Y=y_j] =$$

$$= 0\cdot 0\cdot P[X=0, Y=0] + 0\cdot 1\cdot P[X=0,Y=1] + 1\cdot 0\cdot P[X=1,Y=0] +$$
$$+ 1\cdot 1\cdot P[X=1,Y=1] = P[X=1,Y=1] = d.$$ Luego:

$Cov(X,Y) = d-(b+d)(c+d)$, esta segunda forma de calcular $Cov(X,Y)$ es mucho más facil.

Veamos otro ejemplo en variable continua

Ejemplo (ejercicio 18)

Sea (X, Y) una v.a. bidimensional con función de densidad conjunta

$$f(x,y) = \begin{cases} x+y & \text{si } 0<x<1,\ 0<y<1 \\ 0 & \text{en el resto} \end{cases}$$

Obtenga $Cov(X,Y)$

De nuevo, podemos calcular $Cov(X,Y)$ mediante una de las dos expresiones siguientes:

a) $Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} (x-\mu_X)(y-\mu_Y)\, dy \right) dx$

b) $Cov(X,Y) = E[XY] - E[X]E[Y]$

En cualquiera de las formas que elijamos para el cálculo de $Cov(X,Y)$ necesitamos los valores $E[X]$, $E[Y]$.

$E[X] = \int_{-\infty}^{+\infty} x \cdot f_1(x)\, dx$ siendo $f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\, dy = \int_0^1 x+y\, dy$

$$f_1(x) = \left[xy + \frac{y^2}{2}\right]_0^1 = x + \frac{1}{2}$$

Luego:

$$E[X] = \int_0^1 x \cdot \left(x+\frac{1}{2}\right) dx = \left[\frac{1}{3}x^3 + \frac{1}{4}x^2\right]_0^1 = \frac{1}{3} + \frac{1}{4} = \boxed{\frac{7}{12} = E[X]}$$

$E[Y] = \int_{-\infty}^{\infty} y\, f_2(y)\, dy$ siendo $f_2(y) = \int_{-\infty}^{\infty} f(x,y)\, dx = \int_0^1 (x+y)\, dx$

$$f_2(y) = \left[\frac{x^2}{2} + xy\right]_0^1 = \frac{1}{2} + y$$

Luego:

$$E[Y] = \int_0^1 y \cdot \left(\frac{1}{2} + y\right) dy = \left[\frac{1}{4}y^2 + \frac{1}{3}y^3\right]_0^1 = \frac{1}{4} + \frac{1}{3} = \boxed{\frac{7}{12} = E[Y]}$$

Si elegimos la primera forma:

$$Cov(X,Y) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} \left(x - \frac{7}{12}\right)\left(y - \frac{7}{12}\right) \cdot f(x,y)\, dy \right) dx =$$

$$= \int_0^1 \left(\int_0^1 \left(x-\frac{7}{12}\right)\left(y-\frac{7}{12}\right)(x+y)\, dy \right) dx = \int_0^1 \left(x-\frac{7}{12}\right)\left(\int_0^1 \left(y-\frac{7}{12}\right)(x+y)\, dy \right) dx$$

$$\int_0^1 (y-\frac{7}{12})(x+y)\,dy = \int_0^1 xy + y^2 - \frac{7}{12}x - \frac{7}{12}y\,dy =$$

$$= \left[x\frac{y^2}{2} + \frac{y^3}{3} - \frac{7}{12}xy - \frac{7}{12}\frac{y^2}{2} \right]_0^1 = \frac{x}{2} + \frac{1}{3} - \frac{7}{12}x - \frac{7}{24} =$$

$$= -\frac{1}{12}x + \frac{1}{24}$$. Entonces:

$$\mathrm{Cov}(X,Y) = \int_0^1 (x-\frac{7}{12})(-\frac{1}{12}x+\frac{1}{24})\,dx = \int_0^1 -\frac{1}{12}x^2 + \frac{1}{24}x + \frac{7}{144}x - \frac{7}{288}dx$$

$$\mathrm{Cov}(X,Y) = \left[-\frac{1}{36}x + \frac{1}{48}x^2 + \frac{7}{288}x^2 - \frac{7}{288}x \right]_0^1 =$$

$$\mathrm{Cov}(X,Y) = -\frac{1}{36} + \frac{1}{48} + \frac{7}{288} - \frac{7}{288} = -\frac{1}{144} \;;\; \mathrm{Cov}(X,Y) = -\frac{1}{144}$$

Si queremos calcular $\mathrm{Cov}(X,Y)$ mediante la segunda expresión:

$\mathrm{Cov}(X,Y) = E[XY] - E[X]\cdot E[Y]$ necesitamos el valor de $E[XY]$.

$$E[XY] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\cdot y\, f(x,y)\,dx\,dy = \int_0^1 \left(\int_0^1 xy(x+y)\,dy \right) dx$$

$$E[XY] = \int_0^1 \frac{1}{2}x^2 + \frac{1}{3}x\,dx = \left[\frac{1}{2}\frac{x^3}{3} + \frac{1}{3}\frac{x^2}{2} \right]_0^1 = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$$\int_0^1 xy(x+y)\,dy = \int_0^1 x^2y + xy^2\,dy = \left. \frac{x^2\cdot y^2}{2} + \frac{xy^3}{3} \right]_0^1 = \frac{1}{2}x^2 + \frac{1}{3}x$$

Entonces: $\mathrm{Cov}(X,Y) = \frac{1}{3} - (\frac{7}{12}\cdot\frac{7}{12}) = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}$

Es más facil calcular $\mathrm{Cov}(X,Y)$ mediante esta segunda forma.

Coeficiente de correlación lineal

Dada una v.a. bidimensional (X,Y) se nos plantea el problema de estudiar la influencia que tiene la variable X en la variable Y (o al contrario). De este modo podríamos explicar el comportamiento de la variable Y a partir del conocimiento de la variable X. Esto es lo que se pretende con el coeficiente de correlación. Por tanto el coeficiente de correlación mide el grado de asociación entre dos variables. Se define el coeficiente de correlación lineal ρ_{XY} como:

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

- Si $\rho_{XY} > 0$ se dice que hay correlación directa entre X e Y, cuando una toma valores grandes (pequeños) la otra toma también valores grandes (pequeños)
- Si $\rho_{XY} < 0$ se dice que hay correlación inversa entre X e Y, cuando una toma valores grandes (pequeños) la otra toma valores pequeños (grandes)
- Si $|\rho_{XY}| = 1$ hay una relación lineal perfecta, es decir que $Y = aX + b$
- Si $\rho_{XY} = 0$ las variables se dicen incorreladas, son independientes ya que $Cov(X,Y) = 0$.

Veamos un ejemplo de esto.

Ejemplo: Sea (X,Y) una v.a. bidimensional con función de densidad conjunta

$$f(x,y) = \begin{cases} x+y & \text{si } 0 < x < 1 \quad 0 < y < 1 \\ 0 & \text{en el resto} \end{cases}$$

Obtenga el coeficiente de correlación lineal de (X,Y)

En el ejemplo anterior hemos calculado la covarianza $Cov(X,Y) = -\frac{1}{144}$

Entonces ahora tenemos que calcular σ_X y σ_Y para ello

$Var[X] = E[X^2] - E[X]^2$

$Var[Y] = E[Y^2] - E[Y]^2$

En el ejemplo de la página 5 (vuelta) hemos visto que:

$E[X] = \frac{7}{12}$, $E[Y] = \frac{7}{12}$. Tan sólo nos queda calcular $E[X^2]$ y $E[Y^2]$

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 \cdot f_1(x)\, dx$$

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\, dy = \int_0^1 (x+y)\, dy = \left[xy + \frac{y^2}{2} \right] = x + \frac{1}{2}$$

$$f_1(x) = \begin{cases} x + \frac{1}{2} & \text{si } 0 < x < 1 \\ 0 & \text{en el resto} \end{cases}$$

$$E[X^2] = \int_0^1 x^2 \left(x + \frac{1}{2}\right) dx = \frac{x^4}{4} + \frac{x^3}{6} \Big]_0^1 = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}$$

Del mismo modo

$$E[Y^2] = \int_{-\infty}^{+\infty} Y^2 \cdot f_2(Y)\, dy$$

$$f_2(y) = \int_{-\infty}^{+\infty} f(x,y)\, dx = \int_0^1 (x+y)\, dx = \frac{x^2}{2} + xy \Big]_0^1 = \frac{1}{2} + y$$

$$E[Y^2] = \int_0^1 y^2 \left(\frac{1}{2} + y\right) dy = \frac{y^3}{6} + \frac{y^4}{4} \Big]_0^1 = \frac{5}{12}$$

Entonces:

$$Var[X] = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144} \Rightarrow \sigma_X = \sqrt{\frac{11}{144}}$$

$$Var[Y] = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144} \Rightarrow \sigma_Y = \sqrt{\frac{11}{144}}$$

Entonces:

$$\rho_{XY} = \frac{-\frac{1}{144}}{\sqrt{\frac{11}{144}} \cdot \sqrt{\frac{11}{144}}} = -\frac{1}{11}$$

Momentos respecto al origen en v. a. bidimensional
Si (X,Y) es v. a. bidimensional se define el momento de orden r y s respecto al origen como:

$$m_{rs} = E[X^r Y^s]$$

Momentos respecto a la media en v. a. bidimensional
Si (X,Y) es v. a. bidimensional se define el momento de orden r y s respecto a la media como:

$$\mu_{rs} = E[(X-\mu_x)^r (Y-\mu_y)^s]$$

Podemos observar que $\mu_{10} = \mu_{01} = 0$, $\mu_{11} = Cov(X,Y)$, etc.

Relacionar los momentos respecto a la media con los momentos respecto al origen no es tarea fácil.
¿Cómo está relacionado el momento de orden 2 y 1 respecto a la media con los momentos respecto al origen?

$$\mu_{21} = E[(X-\mu_x)^2 (Y-\mu_y)] = E[(X^2 - 2\mu_x X + \mu_x^2)(Y-\mu_y)]$$

$$= E[X^2 Y - \mu_y X^2 - 2\mu_x XY + 2\mu_x \mu_y X + \mu_x^2 Y - \mu_x^2 \mu_y]$$

$$\mu_{21} = E[X^2 Y] - \mu_y E[X^2] - 2\mu_x \cdot E[XY] + 2\mu_x^2 \mu_y + \cancel{\mu_x^2 \mu_y} - \cancel{\mu_x^2 \mu_y}$$

Luego $\mu_{21} = m_{21} - \mu_y \cdot m_{20} - 2\mu_x m_{11} + 2\mu_x^2 \mu_y$. Lo que sí podemos afirmar es que conocidos los momentos respecto al origen podemos hallar los momentos respecto a la media.
Como ocurría en una variable, el cálculo de los momentos respecto al origen puede complicarse mucho. Para ello tenemos el recurso de hallar la función generatriz de momentos.

Función generatriz de momentos de v.a. bidimensional
La función generatriz de momentos de la v.a. bidimensional (X,Y) se representa por $\phi_{(X,Y)}(s,t)$ y se define:

como:

$$\phi_{(X,Y)}(s,t) = E[e^{sX+tY}]$$

- En el caso discreto

$$\phi_{(X,Y)}(s,t) = \sum_i \sum_j e^{sx_i+ty_j} \cdot P[X=x_i, Y=y_j]$$

- En el caso continuo

$$\phi_{(X,Y)}(s,t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{sx+ty} f(x,y)\, dx\, dy$$

Funciones generatrices de momentos marginales.

La función generatriz de momentos marginal de una variable se obtiene poniendo un cero a la otra variable anulandola. Es decir, si la función generatriz de momentos es $\phi_{(X,Y)}(s,t)$, la función generatriz de momentos marginal respecto a X es:

$$\phi_{(X,Y)}(s,0) = E[e^{sX+0Y}] = E[e^{sX}] = \phi_X(s)$$

lo mismo para la distribución marginal de momentos respecto a Y:

$$\phi_{(X,Y)}(0,t) = E[e^{0X+tY}] = E[e^{tY}] = \phi_Y(t)$$

Calculo de momentos ordinarios a partir de la función generatriz de momentos

$$E[X^h Y^k] = \left. \frac{\partial^{h+k} \phi_{(X,Y)}(s,t)}{\partial s^h \, \partial t^k} \right|_{\substack{s=0 \\ t=0}}$$

$$E[X^h] = \left. \frac{\partial^h \phi_{(X,Y)}(s,0)}{\partial s^h} \right|_{s=0}$$

$$E[Y^k] = \left. \frac{\partial^k \phi_{(X,Y)}(0,t)}{\partial t^k} \right|_{t=0}$$

Ejemplos (Ejercicio 31)

Sea la v.a. bidimensional (X,Y) con función de densidad conjunta $f(x,y)=1$ si $0<x<1$, $0<y<1$ (0 en el resto).

a) Obtenga su función generatriz de momentos

b) Obtenga la función generatriz de $Z=X+Y$

La función generatriz de momentos de (X,Y) es:

$$\phi_{(X,Y)}(s,t)=E[e^{sX+tY}]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}e^{sx+ty}\cdot f(x,y)\,dx\,dy=$$

$$\phi_{(X,Y)}(s,t)=\int_0^1\left(\int_0^1 e^{sx+ty}\cdot 1\,dy\right)dx$$

$$\int e^{sx+ty}\,dy\ \left\{\begin{array}{l} w=sx+ty\\ dw=t\,dy\\ dy=\frac{1}{t}dw\end{array}\right\}=\int e^{w}\cdot\frac{1}{t}dw=\frac{1}{t}\cdot e^{w}=\frac{1}{t}\cdot e^{sx+ty}$$

$$\int_0^1 e^{sx+ty}\,dy=\left[\frac{1}{t}\cdot e^{sx+ty}\right]_0^1=\frac{1}{t}\cdot e^{t+sx}-\frac{1}{t}e^{sx}$$

$$\int\frac{1}{t}\cdot e^{t+sx}-\frac{1}{t}\cdot e^{sx}\,dx=\frac{1}{t}\int e^{t+sx}\,dx-\frac{1}{t}\int e^{sx}\,dx$$

$$\int e^{t+sx}\,dx\ \left\{\begin{array}{l} w=t+sx\\ dw=s\,dx\\ dx=\frac{1}{s}dt\end{array}\right\}=\frac{1}{s}e^{w}=\frac{1}{s}e^{t+sx}$$

$$\int e^{sx}\,dx\ \left\{\begin{array}{l} w=sx\\ dw=\ dx\Rightarrow dx=\frac{1}{s}dw\end{array}\right\}=\frac{1}{s}\cdot e^{sx}$$

$$\int\frac{1}{t}\cdot e^{t+sx}-\frac{1}{t}\,e^{sx}\,dx=\frac{1}{ts}e^{t+sx}-\frac{1}{ts}e^{sx}$$

$$\phi_{(X,Y)}(s,t)=\left[\frac{1}{t\cdot s}\left(e^{t+sx}-e^{sx}\right)\right]_0^1=\frac{1}{t\cdot s}\left[\left(e^{t+s}-e^{s}\right)-\left(e^{t}-1\right)\right]$$

$$\phi_{(X,Y)}(s,t)=\frac{e^{t+s}-e^{s}-e^{t}+1}{t\cdot s}\quad \text{si } t\neq 0,\ s\neq 0$$

b) La función generatriz de $Z = X+Y$ es:

$$\phi_Z(t) = E[e^{tZ}] = E[e^{t(X+Y)}] = E[e^{tX+tY}] = \phi_{(X,Y)}(t,t)$$

$$\phi_Z(t) = \frac{e^{2t} - e^t - e^t + 1}{t^2} \quad \text{si } t \neq 0.$$

Ejemplo: Sea la v. a. bidimensional (X,Y) con función generatriz $\phi_{(X,Y)}(s,t) = \dfrac{1}{(1-s)(1-t)}$ si $s<1$, $t<1$.

Demuestre que la covarianza de (X,Y) es igual a cero.

Se tiene que:

$$Cov(X,Y) = E[X\cdot Y] - E[X]\cdot E[Y]$$

- $$E[X\cdot Y] = \left.\frac{\partial^2 \phi_{(X,Y)}(s,t)}{\partial s\, \partial t}\right|_{\substack{s=0\\t=0}} = \frac{\partial}{\partial s}\left.\left(\frac{\partial \phi_{(X,Y)}(s,t)}{\partial t}\right)\right|_{\substack{s=0\\t=0}} =$$

$$= \frac{\partial}{\partial s}\left.\left(+\frac{(1-s)}{(1-s)^2(1-t)^2}\right)\right|_{\substack{s=0\\t=0}} = \frac{\partial}{\partial s}\left.\left(\frac{1}{(1-s)(1-t)^2}\right)\right|_{\substack{s=0\\t=0}} =$$

$$= \left.\frac{(1-t)^2}{(1-s)^2(1-t)^4}\right|_{\substack{s=0\\t=0}} = \left.\frac{1}{(1-s)^2(1-t)^2}\right|_{\substack{s=0\\t=0}} = 1$$

- $$E[X] = \left.\frac{\partial \phi_{(X,Y)}(s,0)}{\partial s}\right|_{s=0} = \frac{\partial}{\partial s}\left.\left(\frac{1}{1-s}\right)\right|_{s=0} = \left.\frac{1}{(1-s)^2}\right|_{s=0} = 1$$

- $$E[Y] = \left.\frac{\partial \phi_{(X,Y)}(0,t)}{\partial t}\right|_{t=0} = \frac{\partial}{\partial t}\left.\left(\frac{1}{1-t}\right)\right|_{t=0} = \left.\frac{1}{(1-t)^2}\right|_{t=0} = 1$$

Luego $Cov(X,Y) = 1 - 1\cdot 1 = 0$.

Función generatriz de momentos de la suma de n-variables independientes

Sean $(x_i)_{1\le i\le n}$ v. a. con función generatriz de momentos para cada una $\phi_{x_i}(t)$. Si $(x_i)_{1\le i\le n}$ son independientes entonces la función generatriz de momentos de la suma $\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$ es:

$$\phi_{x_1+x_2+\cdots+x_n}(t) = \prod_{i=1}^{n} \phi_{x_i}(t) = \phi_{x_1}(t)\cdot\phi_{x_2}(t)\cdots\phi_{x_n}(t)$$

Demostración:

$$\phi_{x_1+x_2+\cdots+x_n}(t) = E\left[e^{(x_1+x_2+\cdots+x_n)t}\right] = E\left[e^{x_1 t}\cdot e^{x_2 t}\cdots e^{x_n t}\right]$$

$$= E[e^{x_1 t}]\cdot E[e^{x_2 t}]\cdots E[e^{x_n t}] = \phi_{x_1}(t)\cdot\phi_{x_2}(t)\cdots\phi_{x_n}(t)$$

$\downarrow$ por ser independientes