Tema 3. Variables Aleatorias Conjuntas.

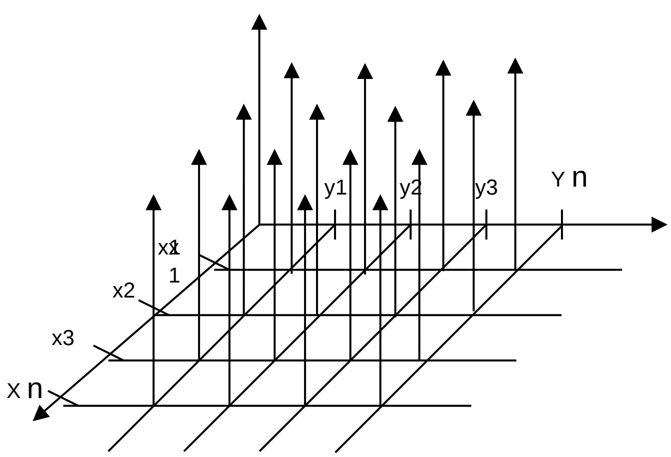
Objetivo: El alumno formulará funciones de probabilidad y densidad para variables aleatorias discretas y continuas, analizará su comportamiento utilizando los fundamentos de la teoría de la probabilidad conjunta e individualmente de las variables, e identificará las relaciones de dependencia entre dichas variables.

Variable aleatoria discreta conjunta (bivariantes).

Sean X y Y variables aleatorias discretas conjuntas. Para que X y Y definan una función de probabilidad conjunta, debe cumplirse que

1) $0 \leq P(x, y) \leq 1$

2) $\sum_x \sum_y P(x, y) = 1$



x / y	x_1	x_2	…	x_m	$h(y)$
y_1	$P(x_1,y_1)$	$P(x_2,y_1)$	…	$P(x_m,y1)$	$h(y_1)$
y_2	$P(x_1,y_2)$	$P(x_2,y_2)$	…	$P(x_m,y2)$	$h(y_2)$
…	…	…	…	…	…
y_n	$P(x_1,y_n)$	$P(x_2,y_n)$	…	$P(x_m,y_n)$	$h(y_n)$
$g(x)$	$g(x_1)$	$g(x_2)$	…	$g(x_m)$	1

Ejemplo:

De una caja que contiene 3 lápices azules, 2 rojos y 3 verdes. Se toman 2 lápices al azar. Si X es el número de lápices azules y Y el número de lápices rojos seleccionado, a) encuentre la función de la probabilidad conjunta f(x, y) y b) la probabilidad de (x, y);

$P[(x, y)] \ \varepsilon\ R]$; donde R es la región $R\{(x,y)|x+y \leq 1\}$

x=número de lápices azules = 3

y=número de lápices rojos = 2

Lápices verdes = 3

Se toman 2 lápices

X= {0, 1,2}

Y= {0, 1,2}

a)

x \ y	0	1	2	h(y)
0	3/28	9/28	3/28	15/28
1	6/28	6/28	0	12/28
2	1/28	0	0	1/28
g(x)	10/28	15/28	3/28	1

$$P(0,0)=\frac{\binom{3}{0}\binom{2}{0}\binom{3}{2}}{\binom{8}{2}}=\frac{3}{28}$$

$$P(1,0)=\frac{\binom{3}{1}\binom{2}{0}\binom{3}{1}}{\binom{8}{2}}=\frac{9}{28}$$

$$P(2,0)=\frac{\binom{3}{2}\binom{2}{0}\binom{3}{0}}{\binom{8}{2}}=\frac{3}{28}$$

$$P(0,1)=\frac{\binom{3}{0}\binom{2}{1}\binom{3}{1}}{\binom{8}{2}}=\frac{6}{28}$$

$$P(1,1)=\frac{\binom{3}{1}\binom{2}{1}\binom{3}{0}}{\binom{8}{2}}=\frac{6}{28}$$

$p(2,1)=0$ Ya que se requieren 3 y sólo se pueden sacar 2

$$P(0,2)=\frac{\binom{3}{0}\binom{2}{2}\binom{3}{0}}{\binom{8}{2}}=\frac{1}{28}$$

$$P(1,2)=0$$
$$P(2,2)=0$$

b) $P(x+y\leq 1)=P(0,0)+p(0,1)+P(1,0)=\frac{3}{28}+\frac{6}{28}+\frac{9}{28}=\frac{18}{28}=\frac{9}{14}$

Variables aleatorias conjuntas continuas.

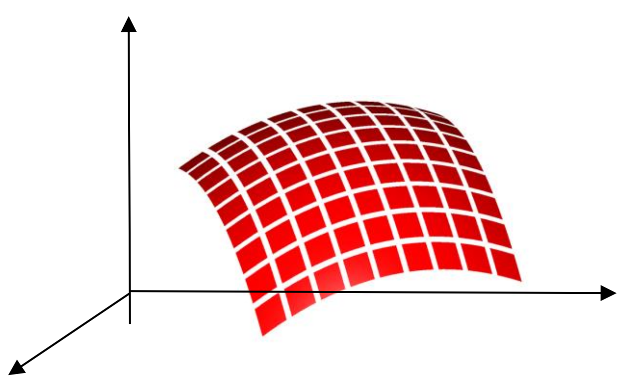
Sean X y Y dos variables aleatorias conjuntas continuas. Para que X y Y definan una función de densidad conjunta debe cumplirse:

1) f(x, y) > 0

2) $\int_c^d\int_a^b f(x,y)dxdy=1$

$$a\leq x\leq b$$
$$c\leq y\leq d$$

Gráfica



Ejemplo:

Considérese la función de densidad conjunta

$$f(x, y) = \begin{cases} \dfrac{x(1+3y^2)}{4} & ; 0 < x < 2 \quad 0 < y < 1 \\ 0 & \text{c.o.c} \end{cases}$$

a) Verifique que f(x, y) define una función densidad de probabilidad.
b) Encuentre la $P[(x, y)\varepsilon A]$ donde A es la región $\{(x, y)/0 \le x \le 1, \frac{1}{4} \le y \le \frac{1}{2}\}$

a)

$$\frac{1}{4}\int_0^1\int_0^2 (x+3xy^2)dxdy = \frac{1}{4}\int_0^1 (\frac{x^2}{2}+\frac{3}{2}x^2y^2)\Big]_0^2 dy = \frac{1}{4}\int_0^1 (2+6y^2)dy = \frac{1}{4}(2y+2y^3)\Big]_0^1 =$$

$$\frac{1}{4}(4) = 1$$

b)

$$\frac{1}{4}\int_{\frac{1}{4}}^{\frac{1}{2}}\int_0^1 x(1+3y^2)dxdy = \frac{1}{4}\int_{\frac{1}{4}}^{\frac{1}{2}} (\frac{x^2}{2}+\frac{3}{2}x^2y^2)dy\Big]_0^1 = \int_{\frac{1}{4}}^{\frac{1}{2}} (\frac{1}{2}+\frac{3}{2}y^2)dy = \frac{1}{2}y+\frac{y^3}{3}\Big]_{\frac{1}{4}}^{\frac{1}{2}} =$$

$$= (\frac{1}{4}+\frac{1}{24}) - (\frac{1}{8}+\frac{1}{192}) = \frac{48+8-3-1}{192} = \frac{52}{192}$$

DISTRIBUCIONES CONDICIONALES

Recordando:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

$$f(X=x\,|\,Y=y)=\frac{f(x,y)}{f(Y=y)}=\frac{f(x,y)}{h(y)}$$

$$f(Y=y\,|\,X=x)=\frac{f(x,y)}{f(X=x)}=\frac{f(x,y)}{g(x)}$$

$$P(a<x<b/Y=y)=\int_a^b f(x/y)dx$$

$$P(c<y<d/X=x)=\int_c^d f(y/x)dy$$

Ejemplo: Con referencia al ejemplo de los lápices de colores, encuentre la distribución condicional de X, dado que Y = 1 y utilícela para determinar: $P(X=0|Y=1)$.

a) $f(x/1)=f(X=x/y=1)$

b) $f(x=0/y=1)=\dfrac{f(0,1)}{h(1)}$

Completando la tabla de distribución conjunta:

y \ x	0	1	2	h(y)
0	3/28	9/28	3/28	15/28
1	6/28	6/28	0	12/28
2	1/28	0	0	1/28
g(x)	10/28	15/28	3/28	1

$$h(1)=\sum_{x=0}^{2} f(x,1)=\frac{6}{28}+\frac{6}{28}+0=\frac{3}{7}$$

$f(x/1)=\dfrac{f(x,1)}{h(1)}=\dfrac{7}{3}f(x,1)$ x = 0, 1, 2

$$f(0/1)=\frac{7}{3}f(0,1)=(\frac{7}{3})(\frac{6}{28})=\frac{1}{2}$$
$$f(1/1)=\frac{7}{3}f(1,1)=(\frac{7}{3})(\frac{6}{28})=\frac{1}{2}$$
$$f(2/1)=\frac{7}{3}f(2,1)=(\frac{7}{3})(0)=0$$

a)

Distribución condicional

X	0	1	2
f(x/1)	½	½	0

b)

$$P(x=0/y=1)=\frac{1}{2}$$

DISTRIBUCIONES MARGINALES

$$\left.\begin{array}{l} g(x)=\sum_{\forall y} f(x,y) \\ h(y)=\sum_{\forall y} f(x,y) \end{array}\right\} \text{Distribuciones marginales caso discreto}$$

$$\left.\begin{array}{l} g(x)=\int_{c}^{d} f(x,y)dy \\ h(y)=\int_{a}^{b} f(x,y)dx \end{array}\right\} \text{Distribuciones marginales caso continuo}$$

Ejemplo:

Sea la función de densidad conjunta:

$$f(x,y)=\begin{cases} \frac{x(1+3y^2)}{4} & 0<x<2;0<y<1 \\ & 0 \end{cases}$$

Obtenga:

a) Las distribuciones marginales g(x) y h(y).
b) La distribución condicional f(x/y)
c) La probabilidad $P\left(\frac{1}{4} < x < \frac{1}{2} \middle| y = \frac{1}{3}\right)$

a) $g(x) = \int_0^1 \frac{x+3xy^2}{4} dy = \frac{1}{4}\int_0^1 xdy + \frac{3}{4}\int_0^1 xy^2 dy = \frac{x}{4}(y) + \frac{3x}{4}\left(\frac{y^3}{3}\right) = \frac{xy}{4} + \frac{xy^3}{4}\Big]_0^1$

$= \frac{x}{4} + \frac{x}{4} = \frac{x}{2}; 0 < x < 2$

$h(y) = \int_0^2 \frac{x+3xy^2}{4} dx = \frac{1}{4}\int_0^2 xdx + \frac{3y^2}{4}\int_0^2 xdx = \frac{x^2}{8} + \frac{3y^2x^2}{8}\Big]_0^2 = \frac{1}{2} + \frac{3y^2}{2} = \frac{1+3y^2}{2}$; 0<y<1

b) $f(x \mid y) = \frac{f(x,y)}{h(y)} = \frac{x(1+3y^2)}{4\left(\frac{1+3y^2}{2}\right)} = \frac{x}{2}$

c) $P(a < x < b / Y = y) = \int_a^b f(x/y)dx$

$P(\frac{1}{4} < x < \frac{1}{2} / y = \frac{1}{3}) = \int_{\frac{1}{4}}^{\frac{1}{2}} \frac{x}{2} dx = \frac{x^2}{4}\Big]_{\frac{1}{4}}^{\frac{1}{2}} = \frac{1}{16} - \frac{1}{64} = \frac{4-1}{64} = \frac{3}{64}$

Se puede observar en este caso que f(x/y) no depende de y, entonces es igual a g(x).

Si ocurriera que f(x/y) no dependiera de x, entonces f(x/y)=h (y). En este caso se puede hacer la siguiente afirmación:

Si $f(x,y) = g(x) \bullet h(y)$

Se dice que las variables x e y son <u>independientes</u> (estadísticamente independientes)

Esperanza de una distribución conjunta

Recordando que:

$$E[h(x)] = \sum_{\forall x} h(x)\, p(x)$$ Para caso discreto

$$E[h(x)] = \int_a^b h(x)\, f(x) dx$$ Para caso continuo

$$E[h(x,y)] = \sum_x \sum_y h(x,y)\, f(x,y)$$ Para caso discreto

$$E[h(x,y)] = \int_c^d \int_a^b h(x,y)\, f(x,y) dxdy$$ Para caso continuo

Ejemplo:

Obtener el valor esperado de xy a partir de la tabla.

y\x	0	1	2	h(y)
0	3/28	9/28	3/28	15/28
1	6/28	6/28	0	12/28
2	1/28	0	0	1/28
g(x)	10/28	15/28	3/28	1

$$E[x,y] = \sum_{\forall x} \sum_{\forall y} x \bullet y f(x,y) = (0)(0)\frac{3}{28} + (1)(0)\frac{9}{28} + (2)(0)\frac{3}{28} + (0)(1)\frac{6}{28} +$$
$$(1)(1)\frac{6}{28} + (2)(1)0 + (0)(2)\frac{1}{28} + (1)(2)0 + (2)(2)0 = \frac{6}{28} = \frac{3}{14}$$

Obtener la E[y/x] para la distribución continua:

$$f(x,y)=\begin{cases}\dfrac{x+3xy^2}{4} & 0<x<2;0<y<1\\ 0 & \end{cases}$$

$$\mathrm{E}[y/x]=\int_0^1\int_0^2 \frac{y}{x}\left(\frac{x+3xy^2}{4}\right)dxdy=\int_0^1\int_0^2 \frac{xy+3xy^3}{4x}dxdy=\int_0^1\int_0^2 \frac{y+3y^3}{4}dxdy=\int_0^1 \frac{2(y+3y^3)}{4}dy=$$

$$\int_0^1 \frac{y+3y^3}{2}dy=\frac{1}{4}(y^2+\frac{3}{2}y)\Big]_0^1=\frac{1}{4}(1+\frac{3}{2})=\frac{5}{8}$$

Teorema:

Sean X e Y dos variables aleatorias independientes entonces:

$$\mathrm{E}[x\bullet y]=\mathrm{E}[x]\bullet\mathrm{E}[y]$$

$$\mathrm{E}[x\bullet y]=\int_0^1\int_0^2 xy\left[\frac{x+3xy^2}{4}\right]dxdy=$$

$$\frac{1}{4}\int_0^1\int_0^2 x^2y+3x^2y^3dydx=\frac{1}{4}\int_0^1 \frac{yx^3}{3}+x^3y^3\Big]_0^2 dy=\frac{1}{4}\int_0^1(\frac{8}{3}y+8y^3)dy=\frac{1}{4}\left[\frac{3}{4}y^2+2y^4\right]_0^1=\frac{1}{4}(\frac{4}{3}+2)=$$

$$\frac{1}{4}(\frac{16}{3})=\frac{16}{12}=\frac{5}{6}$$

Comprobación:

$$\mathrm{E}[x]=\frac{1}{4}\int_0^1\int_0^2(x^2+3x^2y^2)dxdy=\frac{1}{4}\int_0^1\left[\frac{x^3}{3}+x^3y^2\right]_0^2 dy=\frac{1}{4}\int_0^1\left[\frac{8}{3}+8y^2\right]dy=\frac{1}{4}\left(\frac{8}{3}y+\frac{8}{3}y^3\right)_0^1=\frac{1}{4}\left(\frac{8}{3}+\frac{8}{3}\right)=\frac{4}{3}$$

$$\mathrm{E}[y]=\frac{1}{4}\int_0^1\int_0^2\left(xy+3xy^3\right)dxdy=\frac{1}{4}\int_0^1\left[\frac{x^2y}{2}+\frac{3}{2}x^2y^2\right]_0^2 dy=\frac{1}{4}\int_0^1\left[2y+6y^3\right]dy=$$

$$\frac{1}{4}\left[y^2+\frac{3}{2}y^4\right]_0^1=\frac{1}{4}\left[1+\frac{3}{2}\right]=\frac{1}{4}\left(\frac{5}{2}\right)=\frac{5}{8}$$

$\therefore \; \mathrm{E}[x \bullet y] = \mathrm{E}[x] \bullet \mathrm{E}[y]$

$\frac{5}{6} = \frac{4}{3} \bullet \frac{5}{8}$

$\frac{5}{6} = \frac{5}{6}$

Teorema:

La covarianza de dos variables aleatorias X e Y con medias μ_x y μ_y respectivamente está dada por

$$\sigma_{xy} = \mathrm{E}\left[\left(x - \mu_x\right)\left(y - \mu_y\right)\right]$$

O sea:

$$\sigma_{xy} = \mathrm{E}[xy] - \mathrm{E}[x] \bullet \mathrm{E}[y]$$ Covarianza

Corolario:

Si X e Y son estadísticamente independientes:

$$\sigma_{xy} = 0$$

Ejemplo:

Basándose en la tabla de distribución conjunta del ejercicio sobre los lápices de colores. Calcule la covarianza COV [X, Y] y concluya si las variables aleatorias X y Y son estadísticamente independientes o no.

Caso discreto

y\x	0	1	2	h(y)
0	3/28	9/28	3/28	15/28
1	6/28	6/28	0	12/28
2	1/28	0	0	1/28
g(x)	10/28	15/28	3/28	1

$$\sigma_{xy} = \mathrm{cov}[x,y] = \mathrm{E}[xy] - \mathrm{E}[x] \bullet \mathrm{E}[y]$$
$$\mathrm{E}[xy] = \sum_{\forall x}\sum_{\forall y} xyP(x,y)$$

$$\mathrm{E}[xy] = \sum_{\forall x}\sum_{\forall y} xyP(x,y) =$$
$$(0)(0)\frac{3}{28} + (1)(0)\frac{9}{28} + (2)(0)\frac{3}{28} + (0)(1)\frac{6}{28} + (1)(1)\frac{6}{28} + (2)(1)(0) + (0)(2)\frac{1}{28}$$
$$+(1)(2)(0) + (2)(2)(0) = \frac{6}{28}$$

$$E[xy] = \frac{6}{28}$$

$$E[x] = \sum_{x} xg(x) = 0\left(\frac{10}{28}\right) + (1)\left(\frac{15}{28}\right) + (2)\left(\frac{3}{28}\right) = \frac{15}{28} + \frac{6}{28} = \frac{21}{28}$$

$$\therefore\ E[x] = \frac{21}{28}$$

$$E[y] = \sum_{y} yh(y) = (0)\left(\frac{15}{28}\right) + (1)\left(\frac{12}{28}\right) + (2)\left(\frac{1}{28}\right) = \frac{12}{28} + \frac{2}{28} = \frac{14}{28}$$

$$\therefore\ E[y] = \frac{14}{28}$$

$$\mathrm{cov}[xy] = \frac{6}{28} - \left(\frac{21}{28}\right)\left(\frac{14}{28}\right) = \frac{6}{28} - \frac{294}{784} = \frac{168-249}{784} = \frac{-126}{784}$$

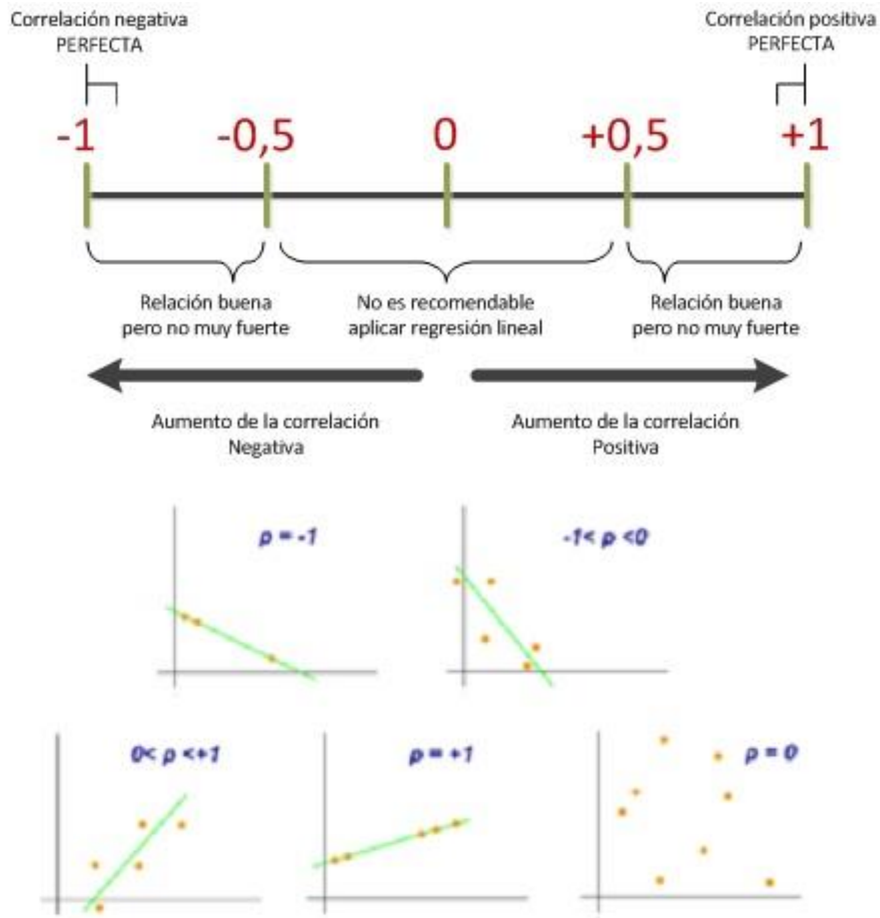
$\therefore\quad \sigma_{xy} = \mathrm{cov}[xy] = -\frac{126}{784} = -\frac{63}{392} = -0.1607 \neq 0$, por lo que las variables aleatorias

X y Y son estadísticamente dependientes.

COEFICIENTE DE CORRELACIÓN

El coeficiente de correlación es una medida que determina el grado al que se asocian los movimientos de dos variables.

Así, el coeficiente de correlación es un número que cuantifica algún tipo de relación y/o dependencia, es decir, relaciones estadísticas entre dos o más variables aleatorias o valores de datos observados.



El coeficiente de correlación para dos variables, puede obtenerse mediante la siguiente expresion matemática:

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \qquad -1 \le \rho \le 1$$

Donde

$$\sigma_x^2 = E[x^2] - E^2[x]$$
$$\sigma_y^2 = E[y^2] - E^2[y]$$

$$\sigma_x = \sqrt{\sigma_x^2}$$
$$\sigma_y = \sqrt{\sigma_y^2}$$

$$\sigma_{xy} = \mathrm{cov}[x, y] = \mathrm{E}[xy] - \mathrm{E}[x] \bullet \mathrm{E}[y]$$
$$\mathrm{E}[xy] = \sum_{\forall x} \sum_{\forall y} xyP(x, y)$$

Ejemplo 1:

Obtener el coeficiente de correlación para la siguiente distribución de probabilidad:

y \ x	0	1	2	h(y)
0	3/28	9/28	3/28	15/28
1	6/28	6/28	0	12/28
2	1/28	0	0	1/28
g(x)	10/28	15/28	3/28	1

$$E[x^2]=\sum_y x^2 g(x)=0^2\left(\frac{10}{28}\right)+1^2\left(\frac{15}{28}\right)+2^2\left(\frac{3}{28}\right)=\frac{15}{28}+\frac{12}{28}=\frac{37}{28}$$

$$E[y^2]=\sum_x y^2 h(y)=0^2\left(\frac{15}{28}\right)+1^2\left(\frac{12}{28}\right)+2^2\left(\frac{1}{28}\right)=\frac{12}{28}+\frac{4}{28}=\frac{16}{28}$$

$$\sigma_x^2=\frac{37}{28}-\left[\frac{21}{28}\right]^2=0.7589$$

$$\sigma_y^2=\frac{16}{28}-\left[\frac{14}{28}\right]^2=0.3214$$

$$\sigma_x=0.8712$$

$$\sigma_y=0.5669$$

$$\sigma_{xy}=\text{cov}[xy]=-\frac{126}{784}=-\frac{63}{392}=-0.1607$$

$$\rho=\frac{-0.1607}{(0.8712)(0.7071)}=-0.2609$$

Ejemplo 2.

Sea la función

$$f(x,y) = \begin{cases} x+y; & 0<x<1,\ 0<y<1 \\ 0 & , otros\ valores \end{cases}$$

Calcular A) σ_{xy} , $B)\ \rho_{xy}$

A) $E[xy] = \int_0^1 \int_0^1 (x+y)xydxdy = \int_0^1 \int_0^1 (x^2y + xy^2)\, dxdy$

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

$$= \left[\left(\frac{y^2}{6}\right) + \left(\frac{y^3}{6}\right)\right]_0^1 = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

$$E[x] = \int_0^1 \int_0^1 x(x+y)dxdy = \int_0^1 \int_0^1 (x^2 + xy)\, dxdy$$

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

$$= \frac{y}{3} + \frac{y^2}{4}\Bigg|_0^1 = \frac{7}{12}$$

$$E[y] = \int_0^1 \int_0^1 y(x+y)dxdy = \int_0^1 \int_0^1 (xy + y^2)\, dxdy$$

$$= \int_0^1 \left[\left(\frac{x^2y}{2}\right) + (xy^2)\right]\Bigg|_0^1 dy \ = \int_0^1 \left[\left(\frac{y}{2}\right) + (y^2)\right] dy = \frac{y^2}{4} + \frac{y^3}{3}\Bigg|_0^1 = \frac{7}{12}$$

$$\sigma_{xy} = E[xy] - E[x]\cdot E[y] = \left(\frac{1}{3}\right) - \left(\frac{7}{12}\right)\left(\frac{7}{12}\right) = -\frac{1}{144}$$

B) $\rho_{xy} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

$$\sigma_x^2 = E[x^2] - E^2[x]$$

$$E[x^2] = \int_0^1 \int_0^1 x^2(x+y)dxdy = \int_0^1 \left(\frac{x^4}{4} + \frac{x^3 y}{3}\right)\Bigg|_0^1 dy$$

$$= \int_0^1 \left(\frac{1}{4} + \frac{y}{3}\right) dy = \left(\frac{y}{4} + \frac{y^2}{6}\right)\Bigg|_0^1 = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}$$

$$E[x] = \int_0^1 \int_0^1 x(x+y)dxdy = \int_0^1 \left(\frac{x^3}{3} + \frac{x^2 y}{2}\right)\Bigg|_0^1 dy$$

$$= \int_0^1 \left(\frac{1}{3} + \frac{y}{2}\right) dy = \left(\frac{y}{3} + \frac{y^2}{4}\right)\Bigg|_0^1 = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$$

$$\sigma_x^2 = \left(\frac{5}{12}\right) - \left(\frac{7}{12}\right)^2 = \frac{11}{144}$$

Por lo que

$$\sigma_x = \sqrt{\frac{11}{144}}$$

De manera semejante se puede calcular

$$\sigma_y^2 = E[y^2] - E^2[y]$$

$$E[y^2] = \int_0^1 \int_0^1 y^2(x+y)dxdy = \frac{5}{12}$$

$$E[y] = \int_0^1 \int_0^1 y(x+y)dxdy = \frac{7}{12}$$

Cuyo resultado es también:

$$\sigma_y^2 = \left(\frac{5}{12}\right) - \left(\frac{7}{12}\right)^2 = \frac{11}{144}$$

$$\sigma_y^2 = \frac{11}{144}$$

Entonces:

$$\sigma_y = \sqrt{\frac{11}{144}}$$

Finalmente, el coeficiente de correlación queda:

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{-\frac{1}{144}}{\sqrt{\frac{11}{144}} . \sqrt{\frac{11}{144}}} = -\frac{1}{11}$$

¿Cuál es la interpretación de este resultado?