# 3.4: DISTRIBUCIONES DE PROBABILIDAD CONJUNTAS

Hasta ahora el estudio se ha referido a variables restringidas a espacios muestrales unidimensionales en los que se registran los resultados de un experimento como valores que toma una sola variable aleatoria. Sin embargo, en algunas situaciones será necesario registrar los resultados de dos (o más) variables aleatorias definidas en espacios muestrales bidimensionales (o superiores)

Si *X* e *Y* son dos variables aleatorias discretas, la función *f*(*x*,*y*) es la distribución de probabilidad para sus ocurrencias simultáneas para cualquier par de valores (*x*,*y*) dentro del rango de definición de las variables. Así a *f*(*x*,*y*) se la llama **distribución de probabilidad conjunta** de *X* e *Y*



## Definición de distribución de probabilidad conjunta discreta

Sean *X* e *Y* dos variables aleatorias discretas, ***f*(*x*,*y*) = *P*(*X* = *x* , *Y* = *y*)** se llamará **función distribución de probabilidad conjunta** (o función **masa** de probabilidad de las variables *X* e *Y*) a la probabilidad de que ocurran los resultados *x* e y al mismo tiempo si satisface las siguientes propiedades[1]:

1. $f(x,y) \geq 0$ para todo valor x de X y para todo y de Y
2. $\sum_{x}\sum_{y} f(x,y) = 1$

Como consecuencia, para cualquier región A en el plano *xy*,
$P[(X,Y) \in A] = \sum_{x}\sum_{y} f(x,y)$

Ejemplo 1

En una caja de bolas de billar se encuentran tres azules, dos rojas y tres verdes. Se definen las cantidades que se extraen *X* e *Y* como "Cantidad de bolas azules" y "Cantidad de bolas rojas", respectivamente. Si se seleccionan al azar dos bolas, determinar:

a) La función distribución de probabilidad conjunta

*X*: "Cantidad de bolas azules que se extraen" x: 0, 1, 2, 3

*Y*: "Cantidad de bolas rojas que se extraen" y: 0, 1, 2

E: Se extraen al azar dos bolas de la caja.

[1] Estas propiedades son, por supuesto, análogas al caso unidimensional.

Así, los pares ordenados (0,0), (0,1), (0,2), (1,0), (1,1) y (2,0) son los posibles valores que pueden tomar conjuntamente *X* e *Y*. Necesitamos encontrar una expresión para f(x) que permita calcular cada una de las probabilidades para los pares ordenados.

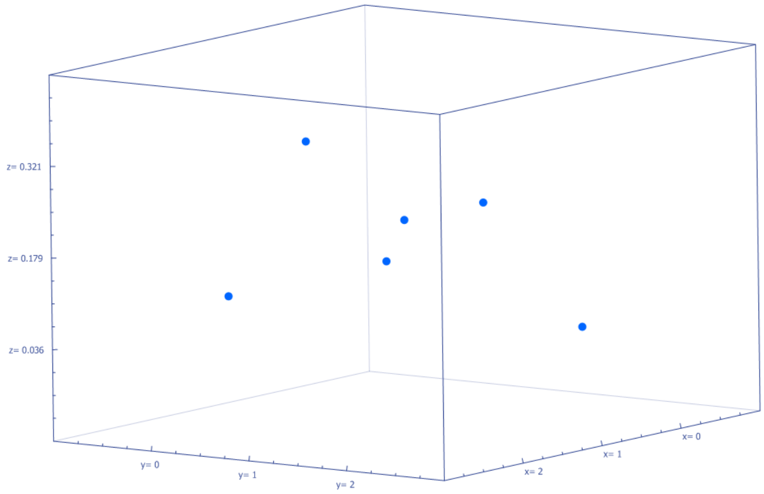
$$f(x,y)=\frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}}$$

La cantidad de formas en que dos bolas pueden extraerse de entre 8 se expresa mediante el número combinatorio $\binom{8}{2}$;así obtenemos la totalidad de resultados del espacio muestral, *#Ω*. La cantidad de formas en que *x* bolas azules pueden obtenerse de las 3 que hay en la caja es $\binom{3}{x}$, donde *x* sabemos que toma los valores de 0 a 3. La cantidad de formas en que *y* bolas rojas pueden obtenerse de las 2 que hay en la caja es $\binom{2}{y}$, donde sabemos que *y* toma los valores de 0 a 2. Dado que se extraen dos bolas al azar y que en la caja hay también bolas verdes, $\binom{3}{2-x-y}$expresa la cantidad de formas en que se pueden extraer bolas verdes cuando en alguna de las dos extracciones (o en ambas) no salió ninguna azul o roja. Ahora podemos armar la siguiente tabla:

| f(x,y) | | | x | | Total |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | |
| y | 0 | 3/28 | 9/28 | 3/28 | 15/28 |
| | 1 | 6/28 | 6/28 | | 12/28 |
| | 2 | 1/28 | | | 1/28 |
| Total | | 10/28 | 15/28 | 3/28 | 1 |

Puede verse que no hay valores para los pares ordenados (1,2), (2,1) y (2,2) debido a que a lo sumo pueden sacarse dos bolas. La suma total, o gran total, es igual a uno, cumpliendo así con la propiedad (2).

b) Si A={(x,y)/x+y<2}
P[(X,Y) $\in$ A] = P( X+Y<2 ) = f(0,0)+ f(0,1)+ f(1,0) =
= 3/28 + 6/28 + 9/28 = 18/28 = 9/14
La representación de f(x,y) es una gráfica de puntos de probabilidad en el espacio de tres dimensiones, cuyo dominio está en el plano xy como puede verse en la siguiente figura:




## Definición de distribución de probabilidad conjunta discreta

Sean *X* e *Y* dos variables aleatorias continuas, se llamará ***f(x,y)*** **función distribución de probabilidad conjunta** (o función **densidad** de probabilidad de las variables *X* e *Y*) a la probabilidad de que ocurran los resultados *x* e y al mismo tiempo si satisface las siguientes propiedades[2]:

1. $f(x,y) \geq 0$ para todo $(x,y)$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$

Como consecuencia, para cualquier región A en el plano xy,

$$P[(X,Y) \in A] = \iint_A f(x,y)dxdy$$

Cuando *X* e *Y* son variables aleatorias continuas, la función de densidad *f(x,y)* es una superficie sobre el plano xy y $P[(X,Y) \in A]$ es igual al volumen del cilindro recto delimitado por la base A y la superficie[3].

Ejemplo 2

Sea *X* el tiempo, en segundos, de reacción frente a la presencia de cierto catalizador e *Y* la temperatura (en °F) a la que la reacción comienza a suceder, y su función densidad conjunta es:

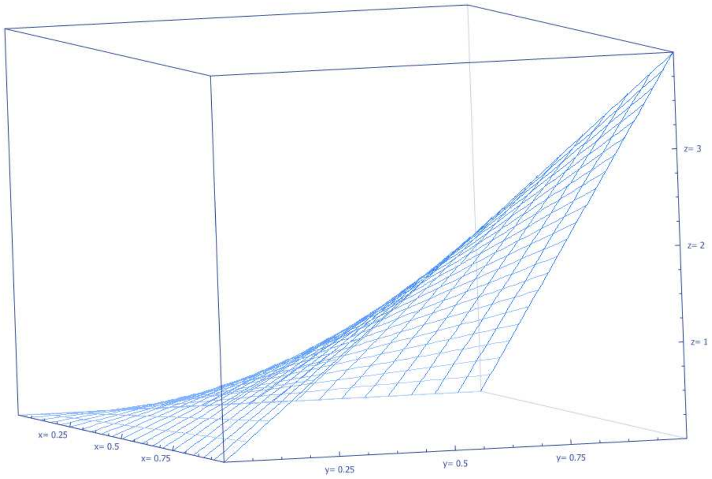
$$f(x,y) = \begin{cases} 4xy & 0 < x < 1 \quad 0 < y < 1 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

a) Compruebe la propiedad (2)

[2] Igual que en el caso unidimensional continuo, ahora f(x,y) no es una probabilidad

[3] Recordando que la integral doble se interpreta como el volumen del cuerpo acotado por el plano xy y por la superficie z = f(x,y), entonces la propiedad (2) es, en realidad, el volumen acotado por el plano xy y dicha superficie y es igual a 1 unidad de volumen.

La siguiente figura muestra la superficie z = f(x,y) que es 4xy. De acuerdo a la propiedad (2), el volumen encerrado por ella y el plano xy para 0<x<1 e 0<y<1 debe ser igual a 1.




$$\int_0^1 \int_0^1 4xydxdy = 4\int_0^1 \left.\frac{x^2}{2}\right|_0^1 dy = 4\frac{1}{2}\int_0^1 dy = 4\frac{1}{2}\left.\frac{y^2}{2}\right|_0^1 = 4\frac{1}{2}\frac{1}{2} = 1$$

b) Encuentre e interprete $P\left(0 \le X \le \frac{1}{2}, \frac{1}{4} \le Y \le \frac{1}{2}\right)$

$$P\left(0 < X < \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2}\right) = \int_{1/4}^{1/2}\int_0^{1/2} 4xydxdy = 4\int_0^{1/2} \left.\frac{x^2}{2}\right|_0^{1/2} dy = 4\frac{1}{8}\int_{1/4}^{1/2} dy =$$

$$= 4\frac{1}{8}\left.\frac{y^2}{2}\right|_{1/4}^{1/2} = 4\frac{1}{8}\left(\frac{1}{8} - \frac{1}{32}\right) = \frac{3}{64}$$

La interpretación de este resultado es: "La probabilidad de que el tiempo de reacción con el catalizador esté entre 0 y la mitad del tiempo total (en segundos) y la temperatura esté entre un cuarto y la mitad de la temperatura total es 3/64"

## Distribuciones marginales de probabilidad

Aun cuando estemos analizando simultáneamente X e Y, es posible seguir considerando las distribuciones de probabilidad de cada una de ellas individualmente. Así, dada la distribución de probabilidad conjunta f(x,y) , la distribución de probabilidad de X considerada individualmente se obtiene al sumar f(x,y) sobre los valores de Y. Igualmente, la distribución de probabilidad de Y considerada individualmente se obtiene al sumar f(x,y) sobre los valores de X (como ilustración, ver del ejemplo 1). Puede observarse que las sumas así obtenidas son las probabilidades para los distintos valores de x e y, respectivamente. Estas probabilidades sumadas dan 1, como corresponde a toda distribución de probabilidad de cualquier variable aleatoria; y por encontrarse estos valores en los márgenes de la tabla (de distribución de probabilidades conjuntas para dos variables discretas) reciben el nombre de distribuciones de probabilidad marginales.

Sean *X* e *Y* dos variables aleatorias, sus **distribuciones de probabilidad marginales** son:

$g(x) = \sum_{y} f(x,y)$ y $h(y) = \sum_{x} f(x,y)$ para el caso discreto y

$g(x) = \int_{-\infty}^{\infty} f(x,y)dy$ y $h(y) = \int_{-\infty}^{\infty} f(x,y)dx$ para el caso continuo[4].

Ejemplo 3

Compruebe que los totales de las columnas y filas de la tabla *f*(*x*,*y*) del ejemplo 1 son las distribuciones marginales de *X* e *Y*.

Para *X*, de acuerdo con la definición

$$g(0) = P(X=0) = \sum_{y=0}^{2} f(0,y) = f(0,0) + f(0,1) + f(0,2) = \frac{3}{28} + \frac{6}{28} + \frac{1}{28} = \frac{10}{28}$$

$$g(1) = P(X=1) = \sum_{y=0}^{2} f(1,y) = f(1,0) + f(1,1) + f(1,2) = \frac{9}{28} + \frac{6}{28} = \frac{15}{28}$$

$$g(2) = P(X=2) = \sum_{y=0}^{2} f(2,y) = f(2,0) + f(2,1) + f(2,2) = \frac{3}{28}$$

Y estos son los valores de las filas en el margen de la tabla de *f*(*x*,*y*). Para *Y*, de acuerdo con la definición

$$h(0) = P(Y=0) = \sum_{x=0}^{2} f(x,0) = f(0,0) + f(1,0) + f(2,0) = \frac{3}{28} + \frac{9}{28} + \frac{3}{28} = \frac{15}{28}$$

$$h(1) = P(Y=1) = \sum_{x=0}^{2} f(x,1) = f(0,1) + f(1,1) + f(2,1) = \frac{6}{28} + \frac{6}{28} = \frac{12}{28}$$

$$h(2) = P(Y=2) = \sum_{x=0}^{2} f(x,2) = f(0,2) + f(1,2) + f(2,2) = \frac{1}{28}$$

Y estos son los valores de las columnas en el margen de la tabla f(x,y). En forma tabular, las distribuciones marginales se presentan:

| *x* | 0 | 1 | 2 |
|---|---|---|---|
| g(*x*) | 10/28 | 15/28 | 3/28 |
| *y* | 0 | 1 | 2 |
| *h*(*x*) | 15/28 | 12/28 | 1/28 |

Se observa que, siendo g(x) y h(y) distribuciones de probabilidad, cumplen con $\sum g(x)=1$ y $\sum h(y)=1$.

[4] Cuando las variables son continuas, las sumatorias se reemplazan por integrales. En este caso, no es posible determinar las probabilidades para los valores individuales de las variables y esto conlleva no poder construir una tabla. Sin embargo, se mantiene el nombre de "marginal".

<u>Ejemplo 4</u>

Encuentre las distribuciones marginales de la función densidad conjunta del ejemplo 2.

De acuerdo con la definición

$$g(x) = \int_0^1 4xy\,dy = 4x\int_0^1 y\,dy = 4x\left.\frac{y^2}{2}\right|_0^1 = 2x \qquad 0 < x < 1$$

$$h(y) = \int_0^1 4xy\,dx = 4y\int_0^1 x\,dx = 4y\left.\frac{x^2}{2}\right|_0^1 = 2y \qquad 0 < y < 1$$

Expresadas como funciones por intervalos

$$g(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{en cualquier otro caso} \end{cases} \qquad h(y) = \begin{cases} 2y, & 0 < y < 1 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Se verifica que $g(x)$ y $h(y)$ son funciones de densidad de probabilidad.

$$\int_0^1 g(x)dx = \int_0^1 2x\,dx = 2\int_0^1 x\,dx = 2\left.\frac{x^2}{2}\right|_0^1 = 1 \quad \int_0^1 h(y)dy = \int_0^1 2y\,dy = 2\int_0^1 y\,dy = 2\left.\frac{y^2}{2}\right|_0^1 = 1$$

En general

$$\int_{-\infty}^{\infty} g(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$$

$$P(a < X < b) = P(a < X < b, -\infty < Y < \infty) = \int_a^b \int_{-\infty}^{\infty} f(x,y)dydx = \int_a^b g(x)dx$$

## Distribuciones condicionales de probabilidad

Recordemos que en la UT–2 se definió a la probabilidad para eventos condicionados de la siguiente manera:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \qquad P(A) > 0$$

Como el valor $x$ de $X$ es un evento que representa un subconjunto de □, entonces $A$ y $B$ son ahora eventos definidos por $X = x$ e $Y = y$, respectivamente. Entonces

$$P(X = x/Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f(x,y)}{h(y)}, \qquad h(y) > 0$$

$$P(Y = y/X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{f(x,y)}{g(x)}, \qquad g(x) > 0$$

donde $X$ e $Y$ son variables discretas. Debe notarse que $f(x,y)/h(y)$ es una función de $x$ con $y$ fija. Igualmente, $f(x,y)/g(x)$ es una función de $y$ con $x$ fija. Ambas satisfacen todas las propiedades de una distribución de probabilidad. Esto también es válido para cuando $f(x,y)$, $g(x)$ y $h(y)$ son las distribuciones conjunta y marginales de variables continuas. Diremos entonces que:

Siendo *X* e *Y* dos variables aleatorias, discretas o continuas, la **distribución condicional de probabilidad de la variable *X*, dado que *Y* toma un valor *y* (o un intervalo c<Y<d)** es:

$$f(x/y)=\frac{f(x,y)}{h(y)},\quad h(y)>0$$

De manera análoga, la **distribución condicional de probabilidad de la variable *Y*, dado que X toma un valor x (o un intervalo a<X<b)** es:

$$f(y/x)=\frac{f(x,y)}{g(x)},\quad g(x)>0$$

<u>Ejemplo 5</u>

Determine la distribución condicional de *Y* del ejemplo 2 y luego utilícela para determinar *P*(*Y* = 1/*X* = 1) e interprete el resultado.

De acuerdo a la definición:

$$f(y/x)=\frac{f(x,y)}{g(x)}, \text{entonces } f(y/1)=\frac{f(1,y)}{g(1)}$$

Del ejemplo 3

$$g(1)=P(X=1)=\sum_{y=0}^{2} f(1,y)=f(1,0)+f(1,1)+f(1,2)=\frac{9}{28}+\frac{6}{28}=\frac{15}{28}$$

Por lo tanto,

$$f(y/1)=\frac{f(1,y)}{15/28}=\frac{28}{15}f(1,y)\qquad y=0,1,2$$

Entonces:

$$f(0/1)=\frac{28}{15}f(1,0)=\frac{28}{15}\frac{9}{28}=\frac{9}{15}=\frac{3}{5}$$

$$f(1/1)=\frac{28}{15}f(1,1)=\frac{28}{15}\frac{6}{28}=\frac{6}{15}=\frac{2}{5}$$

$$f(2/1)=\frac{28}{15}f(2,1)=\frac{28}{15}0=0$$

y la distribución condicional de *Y* dado que *X* = 1 es

| y | 0 | 1 | 2 |
|---|---|---|---|
| f(y/1) | 3/5 | 2/5 | 0 |

Por último P(Y = 1/X = 1) = f(1/1)=2/5. Este resultado se entiende como "la probabilidad de extraer al azar una bola roja si ya se ha extraído una azul es de 2/5"

<u>Ejemplo 6</u>

Determine P(1/2<Y/X=3/4) para el ejemplo 3 e interprete:

Distribuciones conjuntas

$$g(x) = \int_0^1 4xydy = 2x, \quad 0 < x < 1 \quad \text{(como ya se vio)}$$

$$P(1/2 < Y/X = 3/4) = \int_{1/2}^1 \frac{4xy}{2x}dy, \quad 0 < x < 1, 0 < y < 1$$

$$P(1/2 < Y/X = 3/4) = \int_{1/2}^1 2ydy = 2\frac{y2}{2} = y2\Big|_{1/2}^1 = (1 - \frac{1}{4}) = 0{,}75$$

"La probabilidad de que la temperatura de reacción supere el 50% de la temperatura total que puede alcanzar cuando el tiempo alcanza el 75% del tiempo total es de 0,75".

Note que en este cálculo el valor x = 3/4 fue irrelevante y la probabilidad para Y **no dependió** de X. El procedimiento se redujo a integrar h(y) [recordemos que h(y) = 2y] entre los extremos correspondientes. Este resultado está relacionado con el siguiente punto.

## Independencia estadística

Si *f*(*y*/*x*) no depende de *x*, entonces *f*(*y*/*x*) = *h*(*y*). Análogamente, si *f*(*x*/*y*) no depende de *y*, entonces *f*(*y*/*x*) = *g*(*x*)

Sean *X* e *Y* dos variables aleatorias, discretas o continuas, con distribución de probabilidad conjunta *f*(*x*,*y*) y distribuciones marginales *g*(*x*) y *h*(*y*), respectivamente. Se dice que las variables aleatorias *X* e *Y* son **estadísticamente independientes** *sí y sólo sí f*(*x*,*y*) = *g*(*x*) . *h*(*y*)

Esto es, si la distribución de probabilidad conjunta es igual al producto de las distribuciones marginales, entonces *X* e *Y* son estadísticamente independientes. Dado que es una doble implicación, también ocurre que si las variables *X* e *Y* son estadísticamente independientes, entonces la distribución de probabilidad conjunta es igual al producto de las distribuciones marginales.

Se demostrará solamente la implicación a izquierda, es decir, si las variables *X* e *Y* son estadísticamente independientes, entonces la distribución de probabilidad conjunta es igual al producto de las distribuciones marginales.

Hipótesis:

$$f(y/x) = \frac{f(x,y)}{g(x)}, \quad g(x) > 0$$

*X* e *Y* son independientes.

Tesis

*f*(*x*,*y*) = *g*(*x*)*h*(*y*)

Demostración

$$f(y/x)g(x) = f(x,y) \qquad \text{[1], por Hipótesis}$$

$$h(y) = \int_{-\infty}^{\infty} f(x,y)dx \qquad \text{remplazando } f(x,y) \text{ por [1]}$$

$h(y) = \int_{-\infty}^{\infty} f(x/y)g(x)dx$

$h(y) = f(x/y)\int_{-\infty}^{\infty} g(x)dx$ dado que $f(x/y)$ es independiente de $y$ por Hipótesis

$h(y) = f(x/y)$ porque $\int_{-\infty}^{\infty} g(x)dx = 1$
por ser distribución de probabilidad

Finalmente, reemplazando en [1]
$h(y)g(x) = f(x,y)$

Análogamente
<u>Hipótesis</u>:

$f(x/y) = \dfrac{f(x,y)}{h(y)}, \quad h(y) > 0$

*X* e *Y* son independientes.
<u>Tesis</u>
f(x,y) = g(x)h(y)
<u>Demostración</u>

$f(x/y)h(y) = f(x,y)$ [1], por Hipótesis

$g(x) = \int_{-\infty}^{\infty} f(x,y)dy,$ reemplazando $f(x,y)$ por [1]

$g(x) = \int_{-\infty}^{\infty} f(y/x)h(y)dy,$

$g(x) = f(y/x)\int_{-\infty}^{\infty} h(y)dy$ dado que $f(y/x)$ es independiente de $x$ por Hipótesis

$g(x) = f(y/x)$ porque $\int_{-\infty}^{\infty} h(y)dy = 1$
por ser distribución de probabilidad

Finalmente, reemplazando en [1]
$f(x,y) = g(x)h(y)$

<u>Ejemplo 7</u>
Demuestre que las variables *X* e *Y* del ejemplo 3.4 son estadísticamente independientes.
De acuerdo a la definición *X* e *Y* son independientes $\Leftrightarrow f(x,y) = g(x)h(y)$
Entonces, recordando que:

$$g(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{en cualquier otro caso} \end{cases} \qquad h(y) = \begin{cases} 2y, & 0 < y < 1 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Luego $g(x)h(y) = 2x2y = 4xy = f(x,y) \Rightarrow$ X e Y son independientes
(Este resultado ya era esperado al analizar las conclusiones del ejemplo 6).

## Valor esperado de una distribución conjunta de probabilidad

Se extiende el concepto de **valor esperado** de una función de una variable aleatoria al caso de dos variables *X* e *Y* con distribución de probabilidad conjunta *f*(*x*,*y*)

Sean *X* e *Y* variables aleatorias con función distribución de probabilidad conjunta *f*(*x*,*y*). El **valor esperado de una función g(X,Y)** es:

$$\mu_{g(X,Y)} = E[g(X,Y)] = \sum_x \sum_y g(x,y)f(x,y), \qquad \text{si X e Y son discretas, y}$$

$$\mu_{g(X,Y)} = E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy, \qquad \text{si X e Y son continuas}$$

En particular, si g(X,Y) = X, entonces

$$\mu_{g(X,Y)} = E(X) = \sum_x \sum_y xf(x,y) = \sum_x xg(x), \qquad \text{si X e Y son discretas, y}$$

$$\mu_{g(X,Y)} = E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dydx = \int_{-\infty}^{\infty} xg(x)dx, \qquad \text{si X e Y son continuas}$$

donde g(x) es la distribución marginal de x. Un análisis análogo puede hacerse si g(X,Y) = Y.

$$\mu_{g(X,Y)} = E(Y) = \sum_x \sum_y yf(x,y) = \sum_x yh(y), \qquad \text{si X e Y son discretas, y}$$

$$\mu_{g(X,Y)} = E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy = \int_{-\infty}^{\infty} yh(y)dy, \qquad \text{si X e Y son continuas}$$

donde *h*(*y*) es la distribución marginal de *y*.

Sean *X* e *Y* variables aleatorias con distribución de probabilidad conjunta *f*(*x*,*y*). La **covarianza $\sigma_{XY}$ de X e Y** es:

$$\sigma_{(X,Y)} = E[(X-\mu_X)(Y-\mu_Y)] = \sum_x \sum_y (x-\mu_X)(y-\mu_Y)f(x,y), \qquad \text{si X e Y son discretas, y}$$

$$\sigma_{(X,Y)} = E[(X-\mu_X)(Y-\mu_Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y)f(x,y)dxdy, \qquad \text{si X e Y son continuas}$$

<u>Fórmula alternativa</u>

Igual que en el caso unidimensional, es preferible en algunos casos trabajar con una expresión más cómoda. Ésta es $\sigma_{(X,Y)} = E(XY) - \mu_X\mu_Y$

Demostración.

$$\sigma_{(X,Y)} = E[(X-\mu_X)(Y-\mu_Y)] =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y)f(x,y)dxdy = \qquad \text{(por definición)}$$

Distribuciones conjuntas

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(xy - x\mu_y - y\mu_x + \mu_x\mu_y)f(x,y)dxdy =$$ (aplicando propiedad distributiva)

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy - \mu_y\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy -$$
$$- \mu_x\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy + \mu_x\mu_y\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy =$$

$$= E(XY) - \mu_y\mu_x - \mu_x\mu_y + \mu_x\mu_y =$$ (por definición y recordando que $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy=1$ )
$$= E(XY) - \mu_y\mu_x$$

La covarianza es la medida de la fuerza de la asociación entre dos variables aleatorias.
Analizando su signo se interpreta que:

- Si $\rho_{XY} > 0$, la asociación es directa (a medida que crece *X* también lo hace *Y*)
- Si $\rho_{XY} < 0$, la asociación es indirecta (a medida que crece *X*, *Y* decrece)
- Si $\rho_{XY} = 0$, no existe asociación entre las variables y son estadísticamente independientes.

Recordemos que la covarianza es un número que no está acotado y puede ser positivo pero muy grande (o negativo pero muy grande, en valor absoluto), por lo que no proporciona una medida eficaz del grado de asociación entre las variables[5].

## Propiedades del valor esperado de dos (o más) funciones de variables aleatorias

I. El valor esperado de la suma o diferencia de dos (o más) funciones de variables aleatorias *X* e *Y* es la suma o diferencia de los valores esperados de las funciones.

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

<u>Demostración</u>

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[g(x,y) \pm h(x,y)]f(x,y)dxdy =$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy \pm \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f(x,y)dxdy =$$
$$= E[g(X,Y) \pm h(X,Y)]$$

**Corolario 1**. Si $g(X,Y) = g(X)$ y $h(X,Y) = h(Y)$
$$= E[g(X,Y) \pm h(X,Y)] = E[g(X)] \pm E[h(Y)]$$

**Corolario 2**. Si $g(X,Y) = X$ y $h(X,Y) = Y$
$$= E[g(X,Y) \pm h(X,Y)] = E(X) \pm E(Y)$$

[5] El coeficiente de correlación ρ (o de Pearson) está acotado ( $0 \le \rho \le 1$ ) y por lo tanto proporciona una medida eficaz del grado de asociación entre las variables. Su expresión es $\rho_{xy} = \dfrac{\sigma_{xy}}{\sigma_x \sigma_y}$

II. El valor esperado del producto de dos variables aleatorias $X$ e $Y$ independientes es igual al producto de los valores esperados

$$E(XY) = E(X)E(Y)$$

Demostración

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyg(x)h(y)dxdy$$ dado que $f(x,y) = g(x)h(y)$ por hipótesis de independencia.

$$= \int_{-\infty}^{\infty} xg(x)dx \int_{-\infty}^{\infty} yh(y)dy = E(X)E(Y)$$

III. La varianza de una función de $X$ e $Y$ tal como $aX+bY$ es

$$\sigma^2{}_{(aX+bY)} = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY}$$

Demostración

$$\sigma^2{}_{(aX+bY)} = E\{[(aX+bY) - \mu_{(aX+bY)}]^2\}$$

Donde $\mu_{(aX+bY)} = E(aX+bY) = aE(X) + bE(Y) = a\mu_X + b\mu_Y$, entonces

$$E\{[(aX+bY) - \mu_{(aX\pm bY)}]^2\} = E\{[(aX - a\mu_X + bY + b\mu_Y)]^2\}$$

$$= E\{[a(X-\mu_X) + b(Y + b\mu_Y)]^2\}$$

$$= a^2E[(X-\mu_X)]^2 + 2abE[(X-\mu_X)(Y-\mu_Y)] + b^2E[(Y-\mu_Y)]^2$$

$$= a^2\sigma_X^2 + 2ab\sigma_{XY} + b^2\sigma_Y^2$$

**Corolario 1**. Si $X$ e $Y$ son independientes, $\sigma^2{}_{(aX+bY)} = a^2\sigma_X^2 + b^2\sigma_Y^2$ dado que $\sigma_{XY} = E(XY) - \mu_X\mu_Y = E(X)E(Y) - \mu_X\mu_Y = 0$ (por Propiedad II)

**Corolario 2**. Si $X$ e $Y$ son independientes, $\sigma^2{}_{(X\pm Y)} = a^2\sigma_X^2 + b^2\sigma_Y^2$, ya que b se reemplaza por -b en el Corolario 1.

**Corolario 3**. Si $X_1$, $X_2$,…,$X_n$ son n variables aleatorias independientes

$$\sigma^2{}_{(a_1X_1+a_2X_2+\cdots+a_nX_n)} = a_1^2\sigma_{X_1}^2 + a_2^2\sigma_{X_2}^2 + \cdots + a_n^2\sigma_{X_n}^2$$

Distribuciones conjuntas