
Variables aleatorias conjuntas

Distribución de probabilidad conjunta

Los resultados de un experimento pueden ser causa de múltiples variables como ocurre con el precio de un producto y sus ventas, el tiempo de preparación de un examen y su nota final, la cantidad de arena y de cemento en concreto, la cantidad de abono suministrado a una planta y su producción final. En estos casos se requiere una función de densidad que describa la variación de la probabilidad de ocurrencia de ambas variables, probabilidad que describe el comportamiento conjunto de las variables.

La función que tiene en cuenta efectos múltiples de variables aleatorias se denomina distribución de probabilidad conjunta. esta función puede ser una combinación de variables continua-continua, discreta-discreta o continua-discreta, dependiendo del experimento, en el caso bivariado.

En esta guía se presentan los casos : discreta-discreta y el caso continua-continua.

Distribución de probabilidad conjunta variables aleatorias discretas

Si se dispone de dos variables aleatorias se puede definir distribuciones bidimensionales de forma semejante al caso unidimensional. Para el caso discreto-discreto se define:

Función de distribución conjunta

$$f_{X,Y}(x, y) = P(X = x, Y = y)$$

la cual debe cumplir con las siguientes características:

$$\sum_{x=x_{(1)}}^{x_{(n)}} \sum_{y=y_{(1)}}^{y_{(n)}} f_{X,Y}(x, y) = 1$$

$$f(x, y) \geq 0$$

Ejemplo 1

Se considera el experimento aleatorio que consiste en el lanzamiento de dos dados distinguibles, legales donde se registran los resultados de los puntos obtenidos en las

caras superiores de los dados. Para este experimento el espacio muestral es S y se presenta a continuación:

$$S = \left\{ \begin{array}{l} (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6) \\ (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6) \\ (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6) \\ (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6) \\ (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6) \\ (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6) \end{array} \right\}$$

Cada pareja (i, j) representa un posible resultado al lanzar dos dados, donde el valor i corresponde al resultado del dado 1 y el valor j al resultado del dado 2.

Puede resultar interesante determinar la probabilidad que el primer dado caiga en 5 y el segundo en 6 o que el primero caiga en 2 y la suma del número de puntos que aparecen en los dos dados sea 5. De acuerdo con este interés se definen las variables y la función que las modela.

Para la segunda situación se puede definir las variables:

X_1 : número de puntos de la cara superior del primer dado.

X_2 : Suma de puntos de las caras superiores de los dos dados.

De acuerdo con la definición de las variables, los resultados posibles para cada una de ellas son los siguientes:

$$R_{X_1} : \{1, 2, 3, 4, 5, 6\}$$

$$R_{X_2} : \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

Para cada par ordenado posible (x_1, x_2) se determina el valor de la probabilidad conjunta $P(X_1 = x_1; X_2 = x_2)$ que se puede denotar por $f_{X_1, X_2}(x_1, x_2)$.

En particular para los pares $(1, 2)$, $(1, 3)$ y $(6, 12)$ las probabilidades conjuntas son $1/36$ y se muestran a continuación

$$\begin{aligned}
f_{X_1, X_2}(1, 2) &= P(X_1 = 1, X_2 = 2) = 1/36 \\
f_{X_1, X_2}(1, 3) &= P(X_1 = 1, X_2 = 3) = 1/36 \\
&\vdots \\
f_{X_1, X_2}(x_i, x_j) &= P(X_1 = x_i, X_2 = x_j) = 1/36 \\
&\vdots \\
f_{X_1, X_2}(6, 12) &= P(X_1 = 6, X_2 = 12) = 1/36
\end{aligned}$$

En la Tabla 1 se presentan todos los pares posibles (x_1, x_2) con sus respectivas probabilidades conjuntas, los casos imposibles tienen asignados probabilidad de cero.

Tabla 1. Función de probabilidad conjunta f_{X_1, X_2} para el Ejemplo 1

$f_{x_1x_2}(x_1, x_2)$	x_1						
	1	2	3	4	5	6	
x_2	2	1/36	0	0	0	0	0
	3	1/36	1/36	0	0	0	0
	4	1/36	1/36	1/36	0	0	0
	5	1/36	1/36	1/36	1/36	0	0
	6	1/36	1/36	1/36	1/36	1/36	0
	7	1/36	1/36	1/36	1/36	1/36	1/36
	8	0	1/36	1/36	1/36	1/36	1/36
	9	0	0	1/36	1/36	1/36	1/36
	10	0	0	0	1/36	1/36	1/36
	11	0	0	0	0	1/36	1/36
	12	0	0	0	0	0	1/36

El primer caso, $f_{X_1, X_2}(1, 2) = 1/36$ y se lee de la siguiente manera: la probabilidad de que la variable X_1 tome el valor de 1 y la variable X_2 el valor de 2, es igual a $1/36$. En contexto, la probabilidad de que al lanzar dos dados legales distinguibles, el número de puntos del dado1 sea 1 y la suma de puntos de los dos dados sea igual a 2, es de $1/36$. Es decir, cuando se lanzan un número grande de veces dos dados distinguibles y legales aproximadamente el 2.77 % de las veces el número de puntos del dado1 es uno y la suma de los puntos de los dos dados es dos.

En el caso bidimensional la altura representa la probabilidad del evento determinado por

el par ordenado (x_1, x_2) , es decir $f_{X_1, X_2}(x_1, x_2)$. También se pueden dibujar los puntos en lugar de segmentos de rectas. Sin embargo la lectura es menos fácil. El gráfico solamente se puede hacer para el caso de dos variables, dado que solo podemos representar gráficamente tres dimensiones. (R^3)

En la Figura 1 se representa gráficamente la función de probabilidad conjunta f_{X_1, X_2} del Ejemplo 1. En el eje y se encuentran los valores $f_{X_1, X_2}(x_1, x_2)$ y están representados por segmentos de rectas verticales.

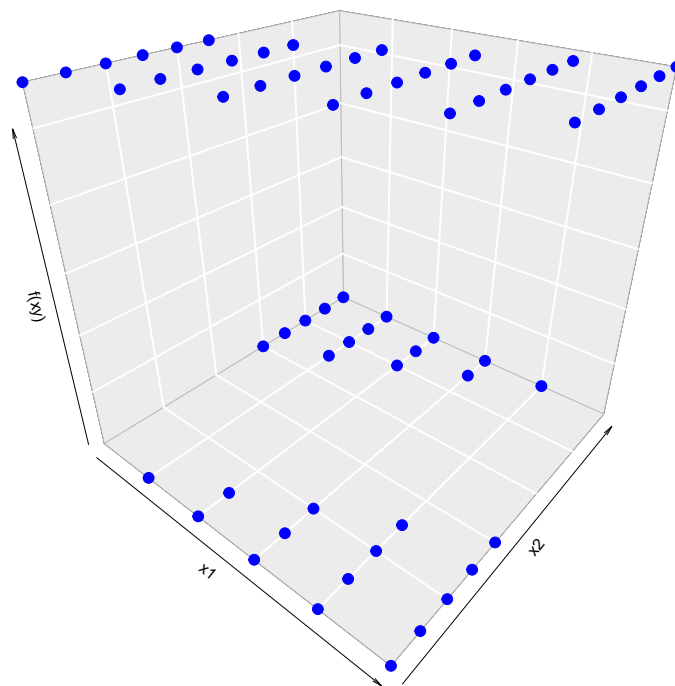


Figura 1. Función de probabilidad conjunta X_1, X_2 en el ejemplo

Cada una de las propiedades se puede ilustrar con la función $f_{X_1, X_2}(x_1, x_2)$ del Ejemplo 1 y los valores de la Tabla 1, pues en efecto cada valor $f_{X_1, X_2}(x_1, x_2)$ está entre 0 y 1 y la suma de todos ellos es exactamente 1.

Se puede encontrar las funciones de distribución de probabilidad para la variable X y también para la variable Y por separado a partir de la función de distribución de probabilidad conjunta. A estas funciones se les llama función de distribución marginal de X y función de distribución marginal de Y .

Funciones de probabilidad marginales

Si X, Y son dos variables aleatorias discretas, entonces se puede definir la función de probabilidad marginal de X (función de probabilidad de X al margen de Y) como

$$g(x) = f_X(x) = \sum_{y=y_{(1)}}^{y_{(n)}} f_{X,Y}(x, y)$$

La función de probabilidad marginal de Y (función de probabilidad de Y al margen de X) como

$$h(y) = f_Y(y) = \sum_{x=x_{(1)}}^{x_{(n)}} f_{X,Y}(x, y)$$

En la Tabla 2 se presenta además de la función de distribución de probabilidad conjunta de X, Y , las distribuciones marginales para la variable X y para la variable Y

Tabla 2. Función de probabilidad conjunta f_{X_1, X_2} y funciones de probabilidad marginal g y h del Ejemplo 1

$f_{XY}(x, y)$		x_1						
x_2		1	2	3	4	5	6	$h(x_2)$
	2	1/36	0	0	0	0	0	1/36
	3	1/36	1/36	0	0	0	0	2/36
	4	1/36	1/36	1/36	0	0	0	3/36
	5	1/36	1/36	1/36	1/36	0	0	4/36
	6	1/36	1/36	1/36	1/36	1/36	0	5/36
	7	1/36	1/36	1/36	1/36	1/36	1/36	6/36
	8	0	1/36	1/36	1/36	1/36	1/36	5/36
	9	0	0	1/36	1/36	1/36	1/36	4/36
	10	0	0	0	1/36	1/36	1/36	3/36
	11	0	0	0	0	1/36	1/36	2/36
	12	0	0	0	0	0	1/36	1/36
$g(x_1)$		6/36	6/36	6/36	6/36	6/36	6/36	1

Al igual que se presentó en el módulo 2, se puede requerir calcular probabilidades condicionales de $X|Y = y_o$, o de $Y|X = x_o$. Para ello es necesario construir las respectivas funciones de distribución de probabilidad condicionales

Función de probabilidad condicional

La función de probabilidad condicional de X dado que $Y = y_0$ está dada por:

$$f_{X|Y}(x|y_0) = \begin{cases} \frac{f_{X,Y}(x, y_0)}{h(y_0)} & ; h(y_0) > 0 \\ 0 & ; \text{en otro caso} \end{cases}$$

La función de probabilidad condicional de Y dado que $X = x_0$ está dada por:

$$f_{Y|X}(y|x_0) = \begin{cases} \frac{f_{X,Y}(x_0, y)}{g(x_0)} & ; g(x_0) > 0 \\ 0 & ; \text{en otro caso} \end{cases}$$

Por ejemplo en el caso del ejemplo 1 podríamos requerir determinar $P(X = 5|X_2 = 8)$. Para realizar el calculo, primero construiremos la función condicional $f_{X_1|X_2=8}(x_1, x_2)$. De acuerdo con lo visto en el modulo 2 tenemos que:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \text{ en este caso } P(X_1 = 5|X_2 = 8) = \frac{P(X_1 = 5; X_2 = 8)}{P(X_2 = 8)}$$

Tabla 3. Función de probabilidad condicional $f_{X_1|X_2=8}(x_1, x_2)$ del Ejemplo 1

$f_{X X_2=5}(x, x_2)$	x_1						
	1	2	3	4	5	6	
x_2	8	0	$\frac{1/36}{5/36}$	$\frac{1/36}{5/36}$	$\frac{1/36}{5/36}$	$\frac{1/36}{5/36}$	$\frac{5/36}{5/36}$

De donde se obtiene que :

$$P(X_1 = 5|X_2 = 8) = \frac{1}{5}$$

Variables aleatorias conjuntas continuas

En el caso de variables continuas se utilizan los mismo conceptos vistos en el caso discreto-discreto, haciendo el cambio de las sumatorias por integrales definidas.

En particular para $f_{X,Y}(x, y)$ definida en una región R , se cumple que la integral doble de $f_{X,Y}(x, y)$ en la región R proporciona la probabilidad de que las variables X y Y asuman los valores (x, y) en la región R . Esta integral puede interpretarse como el volumen bajo la superficie $f_{X,Y}(x, y)$ en la región R .

En esta ocasión empezaremos con un ejemplo para posteriormente definir los conceptos.

Ejemplo 2

Una mezcla se forma con la combinación de 2 ácidos en dos litros de agua. La cantidad de ácido (X) y la cantidad de ácido (Y), en litros, que se vierten en una mezcla se modela con la función $f_{X,Y}$ como se presenta a continuación

$$f_{X,Y}(x,y) = \begin{cases} k(x+y) & ; \quad 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & ; \quad \text{en otro caso} \end{cases}$$

- (a) Determinar el valor de k para que $f_{X,Y}(x,y)$ sea una función de densidad conjunta que modela la cantidad de ácidos vertidos en la mezcla.
- (b) ¿Qué porcentaje de las veces se vierte entre medio y un litro de cada ácido?
- (c) ¿Qué porcentaje de las veces que se vierten los ácidos se obtiene a lo sumo un litro entre ambos ácidos?

Solución:

En este experimento las variables X cantidad del primer ácido y Y la cantidad del segundo ácido son variables aleatorias continuas y su función de densidad de probabilidad depende de una constante k que es necesario identificar para resolver los puntos b) y c).

Las anteriores preguntas equivalen a responder lo siguiente:

- (a) Calcular el valor de k para el cual $f_{X,Y}$ es función de densidad.
- (b) $P(0.5 \leq X \leq 1; 0.5 \leq Y \leq 1)$
- (c) $P(X + Y \leq 1)$

Antes de dar respuesta a los interrogantes, es conveniente representar gráficamente la región donde está definida la función $f_{X,Y}$ que corresponde a un cuadrado con $0 \leq x \leq 1$ y $0 \leq y \leq 1$. Esto se presenta en la Figura 2

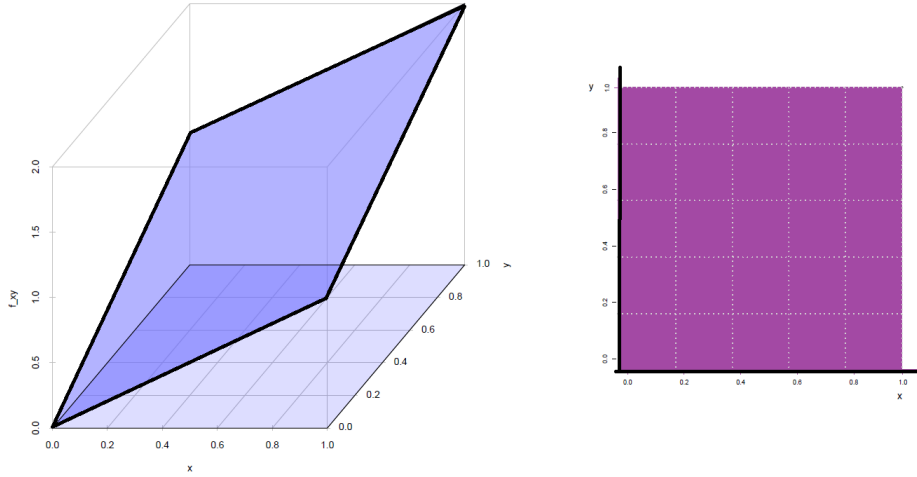


Figura 2 Dominio de $f_{X,Y}$ del Ejemplo 2

La figura 2 representa la función de densidad conjunta de $f_{X,Y}$ en su dominio debe determinar un sólido de volumen igual a 1, esto en concordancia con las propiedades de una función de densidad conjunta.

- (a) La segunda propiedad para el caso particular $n = 2$ (dos variables) que debe cumplir una función de densidad conjunta es:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$$

Por lo que

$$\begin{aligned} \Rightarrow 1 &= \int_0^1 \int_0^1 k(x+y) dx dy \\ &= k \int_0^1 \left[\frac{x^2}{2} + yx \right]_0^1 dy \\ &= k \int_0^1 \left[\frac{1^2}{2} + y \right] dy \\ &= k \left[\frac{1}{2}y + \frac{y^2}{2} \right]_0^1 \\ &= k \left(\frac{1}{2} + \frac{1}{2} \right) = k(1) \Rightarrow k = 1 \end{aligned}$$

La función de densidad conjunta que modela las cantidades de ácido X y Y en la mezcla es:

$$f_{XY}(x, y) = \begin{cases} x + y & ; 0 \leq x \leq 1; 0 \leq y \leq 1 \\ 0 & ; \text{en otro caso} \end{cases}$$

El gráfico de $f_{X,Y}$ se presenta en la Figura 2 y el volumen del sólido debajo del plano en la región $0 \leq x \leq 1, 0 \leq y \leq 1$ es igual a 1.

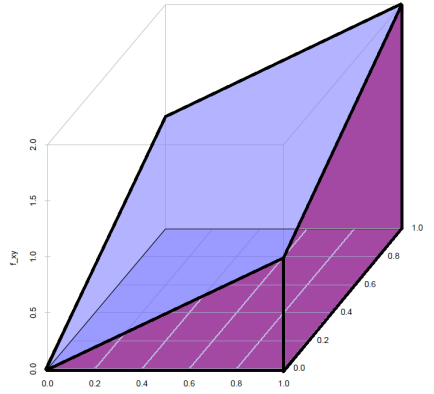


Figura 3 Superficie $f_{X,Y}(x,y)$ cuando $0 \leq x \leq 1$ y $0 \leq y \leq 1$.

- (b) Para determinar $P(0.5 \leq X \leq 1, 0.5 \leq Y \leq 1)$ se debe calcular el volumen del sólido determinado por la superficie $f_{X,Y}(x,y)$ en la región $0.5 \leq X \leq 1, 0.5 \leq Y \leq 1$, esta región se presenta en la Figura 4. Si x varía entre 0 y 1 entonces y varía entre 0 y $-x + 1$, luego $P(X + Y)$ se calcula como sigue

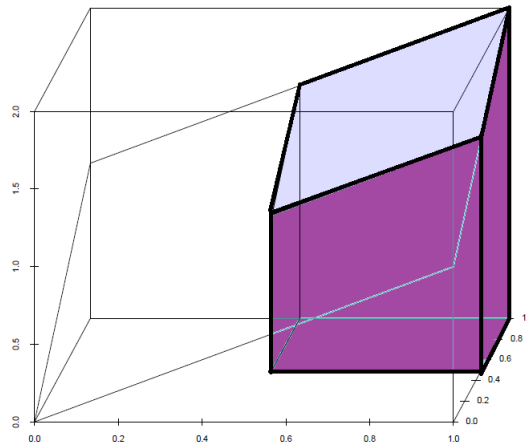


Figura 4 Región intersección entre el dominio de $f_{X,Y}(x,y)$, $0.5 \leq X \leq 1$ y $0.5 \leq Y \leq 1$

El cálculo del volumen lleva a:

$$\begin{aligned}
 P(0.5 \leq X \leq 1; 0.5 \leq Y \leq 1) &= \int_{0.5}^1 \int_{0.5}^1 (x+y) \, dx \, dy \\
 &= \int_{0.5}^1 \left[\frac{x^2}{2} + xy \right]_{0.5}^1 \, dy \\
 &= \int_{0.5}^1 \left[\left(\frac{1}{2} + y \right) - \left(\frac{1}{8} + \frac{y}{2} \right) \right] \, dy \\
 &= \int_{0.5}^1 \left(\frac{y}{2} + \frac{3}{8} \right) \, dy = \left[\frac{y^2}{4} + \frac{3}{8}y \right]_{0.5}^1 \\
 &= \frac{4}{16} + \frac{6}{16} - \frac{4}{16} = \frac{6}{16} = 0.375
 \end{aligned}$$

Aproximadamente el 37.5 % de las veces que se vierten los dos ácidos en la mezcla se vierte entre medio y un litro de cada ácido. Esto es cierto cuando se vierte un

número grande de veces ambos ácidos en la mezcla.

- (c) Para calcular $P(X + Y \leq 1)$ se debe determinar la región de integración determinada por $0 \leq x \leq 1$, $0 \leq y \leq 1$ y $x + y \leq 1$. Para esto en el rectángulo $0 \leq x \leq 1$ y $0 \leq y \leq 1$ se dibuja la recta $y = -x + 1$. Al reemplazar $(0, 0)$ en la desigualdad $x + y \leq 1$, el par $(0, 0)$ satisface la desigualdad, por tanto la región intersección es la sombreada y el volumen correspondiente a la probabilidad requerida se presentan en el Figura xx.

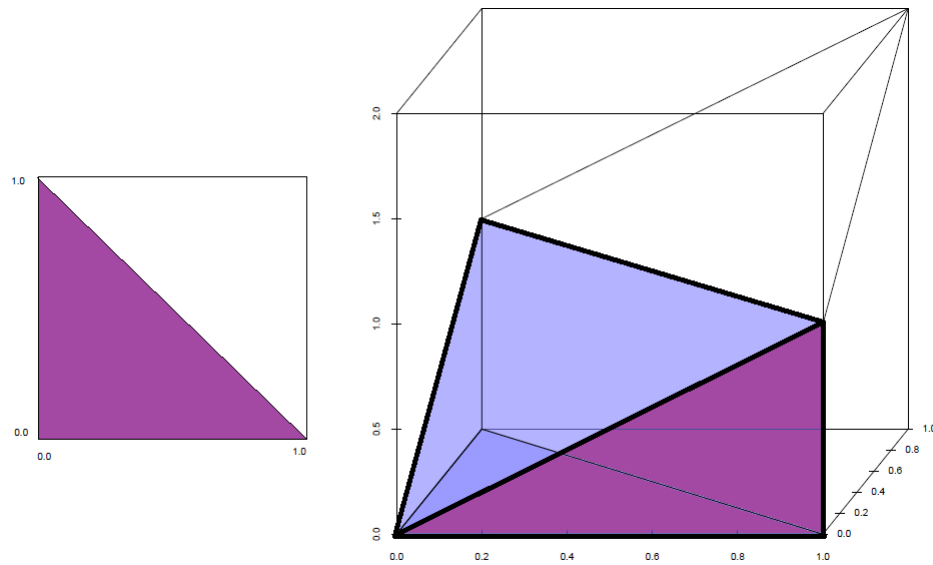


Figura 5 Región determinada por $0 \leq x \leq 1$, $0 \leq y \leq 1$ y $x + y \leq 1$.

En la Figura 5 se observa que mientras la variable X toma valores entre 0 y 1, la variable Y depende de los valores de X , con $Y = 1 - X$. Es así como Y varía entre 0 y $Y = 1 - X$. En ella se ilustra el volumen a calcular para determinar la probabilidad de $P(X + Y \leq 1)$. De acuerdo con esta región, X varía entre 0 y 1, luego $P(X + Y \leq 1)$

$$\begin{aligned}
P(X + Y \leq 1) &= \int_0^1 \int_0^{-x+1} (x + y) dy dx \\
&= \int_0^1 \left[xy + \frac{y^2}{2} \right]_{y=0}^{y=-x+1} dx \\
&= \int_0^1 \left[\frac{(-x+1)^2}{2} + x(-x+1) \right] dx \\
&= \int_0^1 \left(\frac{-x^2 + 1}{2} \right) dx \\
&= \left[\frac{-x^3}{6} + \frac{x}{2} \right]_0^1 \\
&= -\frac{1}{6} + \frac{1}{2} = \frac{1}{3} \approx 0.333
\end{aligned}$$

Aproximadamente el 33.3 % de las veces que se viertan dos ácidos se tiene a lo sumo un litro de ácidos entre los dos ácidos. Esto cuando se vierten ácidos bajo las mismas condiciones un número grande de veces.

Al igual que para las variables discretas, para las variables continuas también se definen las funciones de densidad marginal y condicional.

Probabilidades conjuntas

$$P(a < X < b; c < Y < d) = \int_a^b \int_c^d f(x, y) dy dx$$

Funciones de densidad marginales

En el caso continuo- continuo las distribuciones de densidad marginales se pueden encontrar a partir de la función de densidad conjunta

Función de densidad marginal

Si X y Y son dos variables aleatorias continuas, entonces se define:

La función de densidad marginal de X como:

$$g(x) = f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$$

La función de densidad marginal de Y como:

$$h(y) = f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$$

Ejemplo 3

En el caso del Ejemplo 2 las funciones marginales se obtiene de la siguiente manera:

$$g(x) = f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy = \int_0^1 (x+y) dy = xy + \frac{y^2}{2} \Big|_0^1 = x + \frac{1}{2}$$
$$g(x) = \begin{cases} x + \frac{1}{2} & ; \quad 0 \leq x \leq 1 \\ 0 & ; \quad \text{en otro caso} \end{cases}$$

Para el caso de $h(y)$ se tiene:

$$h(y) = f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx = \int_0^1 (x+y) dx = \frac{x^2}{2} + yx \Big|_0^1 = \frac{1}{2} + y$$
$$h(y) = \begin{cases} y + \frac{1}{2} & ; \quad 0 \leq y \leq 1 \\ 0 & ; \quad \text{en otro caso} \end{cases}$$

Función de densidad condicional

Función de densidad condicional

La función de densidad condicional de X dado que $Y = y_0$ está dada por:

$$f_{X|Y}(x|y_0) = \begin{cases} \frac{f_{X,Y}(x, y_0)}{h(y_0)} & ; \quad h(y_0) > 0 \\ 0 & ; \quad \text{en otro caso} \end{cases}$$

La función de densidad condicional de Y dado que $X = x_0$ está dada por:

$$f_{Y|X}(y|x_0) = \begin{cases} \frac{f_{X,Y}(x_0, y)}{g(x_0)} & ; \quad g(x_0) > 0 \\ 0 & ; \quad \text{en otro caso} \end{cases}$$

Ejemplo 4

Si se desea construir la función de $X|Y = 0.5$, para la función de densidad conjunta del ejemplo 2:

Solución

$$f_{X|Y}(x, 0.5) = \frac{f_{XY}(x, 0.5)}{h(0.5)} = \frac{(x + 0.5)}{0.5 + 0.5} = x + 0.5$$

Función de densidad conjunta acumulada

Para $F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$ se tiene en el caso de variables aleatorias continuas

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(s, t) ds dt$$

Propiedades de la función de distribución conjunta

- $F_{X,Y}(x, y)$ es una función no decreciente.
- $F_{X,Y}(x, -\infty) = 0$
 $F_{X,Y}(-\infty, y) = 0$
 $F_{X,Y}(-\infty, x) = 0$
 $F_{X,Y}(\infty, \infty) = 1$
- $P(x_1 < X \leq x_2, y_1 < Y \leq y_2) =$
 $F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_1)$
- Para todo par de variables aleatorias continuas, si $F_{X,Y}$ tiene derivadas parciales de orden superior a dos, se cumple que:

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y}$$

Ejemplo 5

Para la función descrita en el ejemplo 2 determinar $F(1/2, 1/4)$

$$\begin{aligned} F(1/2, 1/4) &= P(X \leq 1/2; Y \leq 1/4) \\ &= \\ &= \int_0^{1/2} \int_0^{1/4} (x + y) dy dx = \int_0^{1/2} \left(xy + \frac{y^2}{2} \Big|_0^{1/4} \right) dx \\ &= \int_0^{1/2} \left(\frac{x}{4} + \frac{1}{32} \right) dx = \left(\frac{x^2}{8} + \frac{x}{32} \right) \Big|_0^{1/2} \\ &= \frac{1}{32} + \frac{1}{64} = \frac{3}{64} = 0.046875 \end{aligned}$$

Covarianza y correlación $Cov(XY) = \sigma_{XY}$

La covarianza es una medida de variación conjunta entre dos variables aleatorias con respecto a sus medias.

Definición : Covarianza

Sean X y Y variables aleatorias con distribución de probabilidad conjunta $f_{X,Y}(x, y)$. La covarianza de X y Y es:

$$Cov(XY) = \sigma_{xy} = E[(x - \mu_x)(y - \mu_y)] = E[XY] - E[X]E[Y]$$

$$\sigma_{xy} = \sum_x \sum_y (x - \mu_x)(y - \mu_y) f_{XY}(x, y) \quad \text{si } X \text{ y } Y \text{ son discretas}$$

$$\sigma_{XY} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)(y - \mu_Y) f_{XY}(x, y) dx dy \quad \text{si } X \text{ y } Y \text{ son continuas}$$

Si los puntos de la distribución de probabilidad conjunta de X y Y que reciben una probabilidad positiva tienden a caer en una recta con pendiente positiva (o negativa), entonces σ_{XY} es positiva o (negativa). Si los puntos tienden a caer en una recta con pendiente positiva, X tiende a ser mayor que μ_X cuando Y es mayor que μ_Y . Por lo tanto, el producto de los dos términos $x - \mu_X$ y $y - \mu_Y$ tiende a ser positivo. Si los puntos tienden a caer en una recta con pendiente negativa, entonces $x - \mu_X$ tiende a ser positiva cuando $y - \mu_Y$ es negativa, y viceversa. Por lo tanto, el producto de $x - \mu_X$ y $y - \mu_Y$ tiende a ser negativo. En este sentido, la covarianza entre X y Y describe la variación entre las dos variables aleatorias.

Tanto en el caso continuo como discreto, se cumple

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

A partir de la definición del concepto de covarianza, podemos enunciar la siguiente propiedad de las varianzas:

$$V(aX \pm bY) = a^2V(X) + b^2V(Y) \pm 2ab Cov(XY)$$

Propiedades de la covarianza

La covarianza es una medida de la variación común a dos variables y, es una medida del grado de relación lineal entre las dos variables.

- σ_{XY} es positiva si los valores altos de X están asociados a los valores altos de Y y viceversa.
- σ_{XY} es negativa si los valores altos de X están asociados a los valores bajos de Y y viceversa.
- Si X e Y son variables aleatorias independientes $\sigma_{XY} = 0$.
- La independencia es condición suficiente pero no necesaria para que la σ_{XY} sea nula.

En el caso de la covarianza como en el de la varianza, ambas se expresan en términos del producto de las unidades de medida de ambas variables, lo cual no siempre es fácilmente interpretable. Por otra parte también es difícil comparar situaciones diferentes entre sí. En este caso, ambos problemas se solucionan de una vez mediante la definición del **coeficiente de correlación** ρ , que se define como el cociente entre la covarianza y el producto de las desviaciones típicas de las dos variables.

Coeficiente de correlación

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}}$$

$$-1 < \rho_{XY} < 1$$

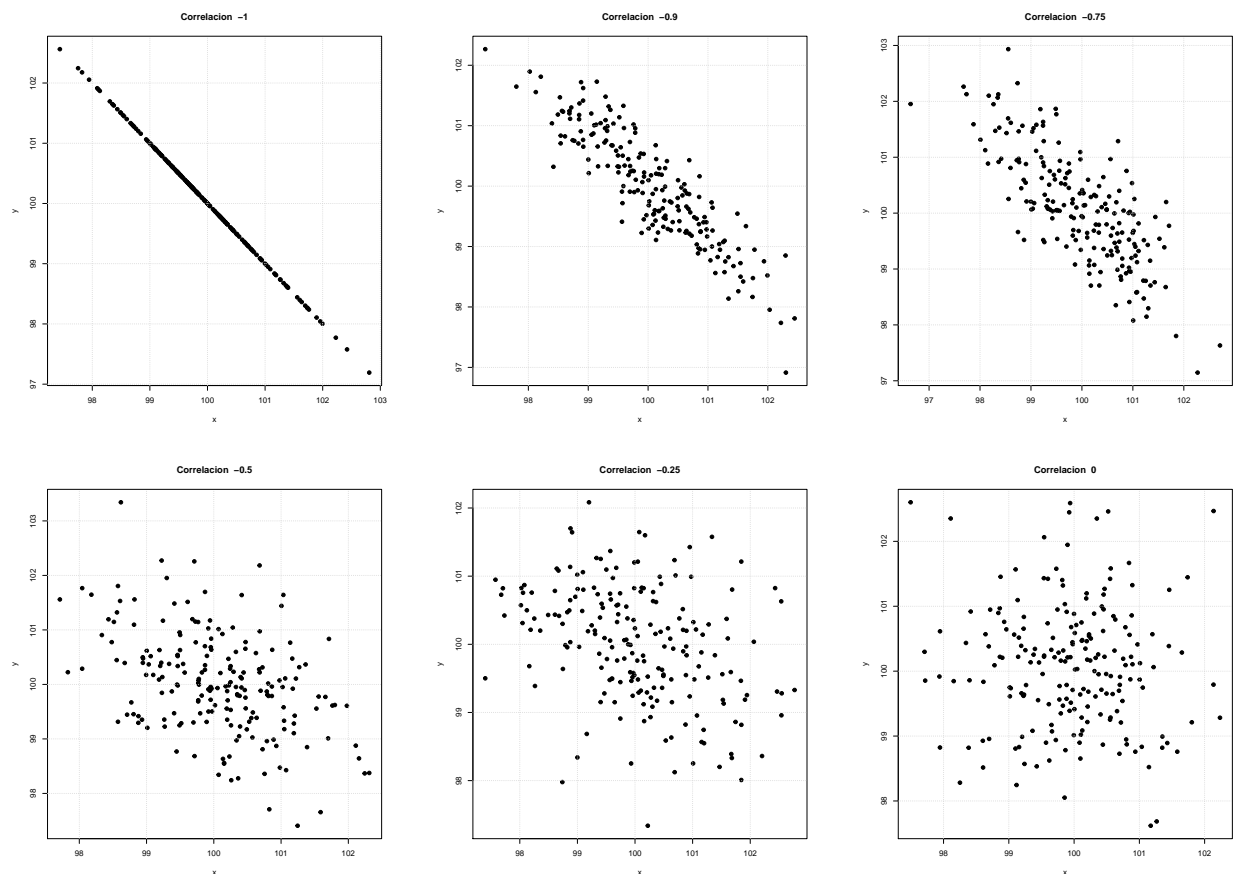
La correlación simplemente escala la covarianza por la desviación estándar de cada variable. Por consiguiente, la correlación es una cantidad adimensional que puede usarse para comparar las relaciones lineales entre pares de variables con unidades diferentes. Si los puntos de la distribución de probabilidad conjunta de X y Y que reciben una probabilidad positiva tienden a caer en una recta con pendiente positiva (o negativa), entonces ρ_{XY} está cerca de $+1$ (o de -1). Si ρ_{XY} es igual a $+1$ o -1 , entonces puede demostrarse que los puntos de la distribución de probabilidad conjunta que reciben una probabilidad positiva caen exactamente en una línea recta. Se dice que dos variables aleatorias con correlación diferente de cero están correlacionadas. En forma similar a la covarianza, la correlación es una medida de relación lineal entre variables aleatorias.

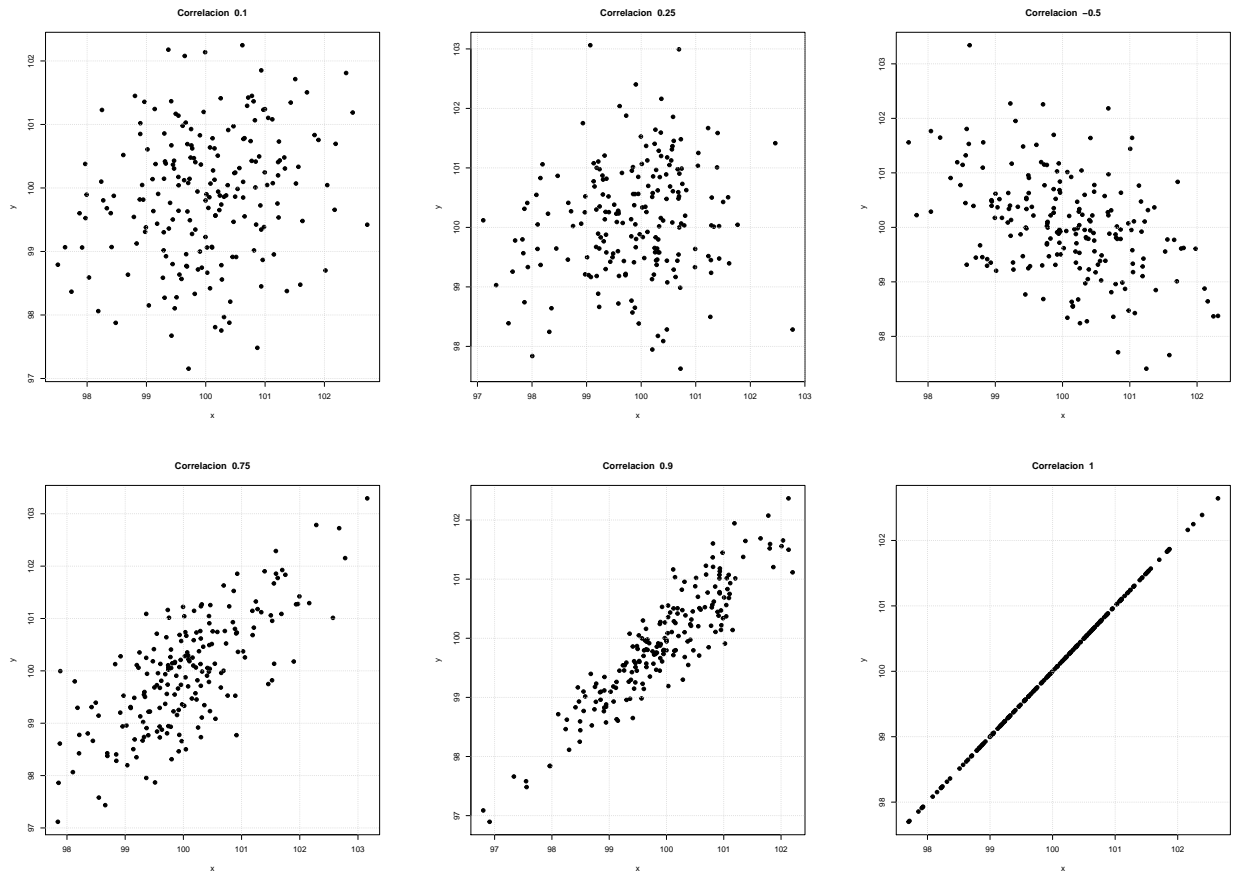
La representación gráfica para difernetes valores de correlación nos ayuda a entender el mismo concepto de relación entre X y Y .

Valores cercanos cero, indican que no existe una relación lineal entre las variables. Valores positivos indican una relación lineal entre las variables, es decir que cuando una variable aumenta la otra también aumenta. Una manera de interpretar los valores obtenidos para el coeficiente de correlación ρ se presentan en la Figura 6

Tabla 3 . Interpretación de valores del coeficiente de correlación

-1	-0.90	-0.75	-0.50	-0.25	-0.10	0	0.10	0.25	0.50	0.75	0.90	1.0
Negativa perfecta	Negativa muy fuerte	Negativa considerable	Negativa media	Negativa debil	Negativa muy debil	No existe correlación	Positiva muy debil	Positiva debil	Positiva media	Positiva considerable	Positiva muy fuerte	Positiva perfecta





Ejemplo 6

Sean X el número de veces en que una máquina presenta una falla y Y el número de veces en que se debe llamar a un ingeniero para que la repare, variables aleatorias con distribución de probabilidad conjunta como se indica en la Tabla XX, encuentre la covarianza entre X y Y $Cov(XY)$ y la correlación entre X y Y $Cor(XY)$

$f_{XY}(x, y)$		x			
		0	1	2	$h(y)$
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

Solución:

Para calcular la covarianza se utiliza la forma $Cov(XY) = E[XY] - E[X]E[Y]$

Inicialmente se calcula $E[XY]$

$$\begin{aligned}
E(XY) &= \sum_{x=0}^2 \sum_{y=0}^2 xy f(x, y) \\
&= (0)(0)f(0, 0) + (0)(1)f(0, 1) + (0)(2)f(0, 2) + (1)(0)f(1, 0) + (1)(1)f(1, 1) + \\
&= (1)(2)f(1, 2) + (2)(0)f(2, 0) + (2)(1)f(2, 1) + (2)(2)f(2, 2) \\
&= 3/14
\end{aligned}$$

Los valores de $E[X]$ y $E[Y]$, se calculan a partir de las respectivas distribuciones marginales $g(X)$ y $h(Y)$.

$$\begin{aligned}
E(X) &= \sum_{x=0}^2 x f(x) \\
&= (0)f(0) + (1)f(1) + (2)f(2) = (0)5/14 + (1)15/28 + (2)3/28 \\
&= 21/28
\end{aligned}$$

$$\begin{aligned}
E(X^2) &= \sum_{x=0}^2 x^2 f(x) \\
&= (0^2)f(0) + (1^2)f(1) + (2^2)f(2) = (0)5/14 + (1)15/28 + (4)3/28 \\
&= 27/28
\end{aligned}$$

$$\begin{aligned}
E(Y) &= \sum_{y=0}^2 y f(y) \\
&= (0)f(0) + (1)f(1) + (2)f(2) = (0)15/28 + (1)3/7 + (2)1/28 \\
&= 14/28
\end{aligned}$$

$$\begin{aligned}
E(Y^2) &= \sum_{y=0}^2 y^2 f(y) \\
&= (0^2)f(0) + (1^2)f(1) + (2^2)f(2) = (0)15/28 + (1)3/7 + (4)1/28 \\
&= 16/28
\end{aligned}$$

$$V[X] = E[X^2] - E[X]^2 = 27/28 - (21/28)^2 = 0.4017857$$

$$V[Y] = E[Y^2] - E[Y]^2 = 16/28 - (14/28)^2 = 0.3214286$$

Por tanto $Cov(XY)$ será:

$$\begin{aligned} Cov(XY) &= E(XY) - E(X)E(Y) = 3/14 - (31/28)(14/28) \\ &= 3/14 - (21/28)(14/28) = -9/56 \end{aligned}$$

Ahora la correlación $Cor(XY)$

$$\begin{aligned} \rho_{XY} &= \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}} \\ &= \frac{-9/56}{\sqrt{0.4017857 \times 0.3214286}} = -0.4472136 \end{aligned}$$

Ejemplo 7

Para el caso continuo-continuo se retoma el Ejemplo 2 con función de distribución conjunta $f(x, y) = x + y$ con $0 \leq x \leq 1$ y $0 \leq y \leq 1$.

Solución

Primero se calculan para cada una de las variables los valores esperados y sus varianzas a partir de las funciones de densidad marginales $g(x)$ y $h(y)$, halladas con anterioridad.

$$E[X] = \int_0^1 xg(x) dx = \int_0^1 x(x+y) dx = \int_0^1 x^2 + \frac{x}{2} dx = \frac{x^3}{3} + \frac{x^2}{4} \Big|_0^1 = \frac{7}{12}$$

$$E[Y] = \int_0^1 yh(y) dx = \int_0^1 y(x+y) dx = \int_0^1 y^2 + \frac{y}{2} dx = \frac{y^3}{3} + \frac{y^2}{4} \Big|_0^1 = \frac{7}{12}$$

$$E[X^2] = \int_0^1 x^2g(x) dx = \int_0^1 x^2(x+y) dx = \int_0^1 x^3 + \frac{x^2}{2} dx = \frac{x^4}{3} + \frac{x^3}{4} \Big|_0^1 = \frac{9}{18} = \frac{1}{2}$$

$$E[Y^2] = \int_0^1 y^2h(y) dy = \int_0^1 y^2(x+y) dy = \int_0^1 y^3 + \frac{y^2}{2} dx = \frac{y^4}{3} + \frac{y^3}{4} \Big|_0^1 = \frac{9}{18} = \frac{1}{2}$$

$$V[X] = E[X^2] - (E[X])^2 = \frac{1}{2} - \left(\frac{7}{12}\right)^2 = \frac{1}{2} - \frac{49}{144} = \frac{23}{144}$$

$$V[Y] = E[Y^2] - (E[Y])^2 = \frac{1}{2} - \left(\frac{7}{12}\right)^2 = \frac{1}{2} - \frac{49}{144} = \frac{23}{144}$$

$$\begin{aligned}
E[XY] &= \int_0^1 \int_0^1 xyf(x,y) \, dx \, dy \\
&= \int_0^1 \int_0^1 xy(x+y) \, dx \, dy \\
&= \int_0^1 \int_0^1 x^2y + y^2x \, dx \, dy \\
&= \int_0^1 \left(\frac{x^3y}{3} + \frac{y^2x^2}{2} \Big|_0^1 \right) dy \\
&= \int_0^1 \left(\frac{y}{3} + \frac{y^2}{2} \right) dy \\
&= \frac{1}{6} + \frac{1}{6} \Big|_0^1 = \frac{1}{3}
\end{aligned}$$

$$Cov[XY] = E[XY] - E[X]E[Y] = \frac{1}{3} - \frac{7}{12} \times \frac{7}{12} = \frac{1}{3} - \frac{49}{144} = \frac{1}{144}$$

$$\rho_{xy} = \frac{Cov[XY]}{\sqrt{V[X]V[Y]}} = \frac{1/144}{\sqrt{23/144 \times 23/144}} = \frac{1}{23} = 0.04$$

De acuerdo con la escala representada en la Tabla 3 este valor se puede interpretar como una relación muy débil de asociación lineal entre las dos variables .

Independencia

La independencia de dos variables se puede determinar mediante el coeficiente de correlación ($\rho = 0$), pero también si se cumple la siguiente condición.

Independencia de variables

Sean X y Y dos variables aleatorias discretas o continuas con función de probabilidad conjunta $f(x, y)$ y funciones de probabilidad marginales $g(x)$ y $h(y)$, respectivas, entonces se dice que las variables X y Y son estadísticamente independientes si:

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

Ejemplo 8

Determinar si las variables X y Y del ejemplo 2 son independientes estadísticamente.

Solución

Debemos determinar si se cumple la condición de independencia

$$\begin{aligned}g(x)h(y) &= \left(x + \frac{1}{2}\right)\left(y + \frac{1}{2}\right) \\&= xy + \frac{x}{2} + \frac{y}{2} + \frac{1}{4}\end{aligned}$$

Como el resultado es diferente a $f(xy)$, se concluye que las variables X y Y no son independientes estadísticamente.

Código R gráfico 1

```
x=c(1,1,1,1,1,1,1,1,1,1,1)
x1=c(1*x,2*x,3*x,4*x,5*x,6*x)
y=2:12
x2=c(y,y,y,y,y,y)
fx1x2=c(1,1,1,1,1,1,0,0,0,0,0,
0,1,1,1,1,1,1,0,0,0,0,
0,0,1,1,1,1,1,1,0,0,0,
0,0,0,1,1,1,1,1,1,0,0,
0,0,0,0,1,1,1,1,1,1,0,
0,0,0,0,0,1,1,1,1,1,1,1)/36

library(plot3D)
scatter3D(x1, x2, fx1x2, colvar = NULL, col = "blue",
pch = 19, cex = 1.5,
phi = 30, theta = 40,
zlab="f(xy)", xlab="x1", ylab="x2",
bty = "g",
col.panel = "steelblue",
col.grid = "darkblue",
add_lines=TRUE)
```

Código R gráfico 2

```
library("scatterplot3d")
x <- c(0,1,1,0,0)
y <- c(0,0,1,1,0)
z <- c(0,1,2,1,0)
s <- scatterplot3d(x,y,z, type='l',
xlim=c(0,1),ylim=c(0,1),zlim=c(0,2), angle=60,
xlab="x", ylab="y", zlab="f(x,y)",
scale.y=0.4,cex.axis=1, cex.names = 1,
grid=TRUE, box=TRUE,
label.tick.marks=TRUE)

x0=c(0,1,1,0)
y0=c(0,0,1,1)
z0=c(0,0,0,0)
polygon(s$xyz.convert(x0,y0,z0),col="#0000ff22") # morado

x1=c(0,1,1,0)
y1=c(0,0,1,1)
z1=c(0,1,2,1)
polygon(s$xyz.convert(x,y,z),col="#80FF8099") # verde
```