TEMA 3

Definición de variable aleatoria bidimensional: Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad. Una v.a. bidimensional es una función en la forma:

$$\begin{array}{lll} \text{Z:} & \Omega \rightarrow & \mathbb{R}^2 \\ & \omega \rightarrow & Z(\omega) = (X(\omega), Y(\omega)) \end{array}$$

donde cada una de las componentes X e Y son variables aleatorias unidimensionales, es decir funciones en la forma:

$$X, Y : \Omega \rightarrow \mathbb{R}$$

verificando:

$X^{-1}\{(-\infty, x)\} = \{\omega \in \Omega \ / X(\omega) \leq x\} \subset \mathcal{A}$

$Y^{-1}\{(-\infty, y)\} = \{\omega \in \Omega \ / Y(\omega) \leq y\} \subset \mathcal{A}$

Veamos un ejemplo:

$\Omega =$"conjunto de los ciudadanos españoles"

$X =$"color del pelo"

$Y =$"comunidad autónoma de nacimiento"

Consideremos la variable aleatoria bidimensional $Z(\omega) = (X(\omega), Y(\omega))$. El campo de variación de las variables aleatorias es:

$$\text{Color de pelo: X=} \begin{cases} \text{Negro} \rightarrow 1 \\ \text{Rubio} \rightarrow 2 \\ \text{Pelirrojo} \rightarrow 3 \\ \text{Castaño} \rightarrow 4 \end{cases}$$

El campo de variación de X es 1, 2, 3, 4

$$\text{Comunidad Autónoma: Y=} \begin{cases} \text{Aragón} \rightarrow 1 \\ \text{Baleares} \rightarrow 2 \\ \text{Canarias} \rightarrow 3 \\ \text{Castilla León} \rightarrow 4 \\ . \\ . \\ . \\ . \end{cases}$$

El campo de variación de Y es 1, 2, 3, ..., 17

Si elegimos al ciudadano ω_1 = José Pi García, que tiene el pelo rubio y nació en Canarias, entonces $Z(\omega_1) = (X(\omega_1), Y(\omega_1)) = (2, 3)$

Dada la v.a. bidimensional $Z = (X, Y)$, a X se le llama primera característica y a Y se la llama segunda característica.

Tipos de variable aleatoria

a) Se dice que $Z = (X, Y)$ es discreta si son discretas X e Y

b) Se dice que $Z = (X, Y)$ es continua si son continuas X e Y

Recordemos lo que ocurría con variable aleatoria de dimensión uno. Para las discretas existía la función de probabilidad y función de distribución. Para las continuas existía función de densidad y función de distribución. Pues bien, a continuación vamos a definir un concepto que es común a la variable discreta y continua: la función de distribución.

Función de distribución de una variable aleatoria bidimensional

Consideremos la variable aleatoria bidimensional $Z(\omega) = (X(\omega), Y(\omega))$ en el espacio de probabilidad $(\Omega, \mathcal{A}, P)$. Se define la función de distribución como $F(x, y) = P[X \leq x, Y \leq y]$. La función de distribución en el punto (x_0, y_0) es la probabilidad acumulada por todos aquellos valores (x, y) tales que $x \leq x_0, y \leq y_0$ (en el caso discreto esta acumulación se haría mediante un sumatorio y en el caso continuo mediante una integral, ya veremos esto más adelante). Las propiedades que ha de cumplir la función de distribución de una v.a. bidimensional son:

1º $F(-\infty, y) = F(x, -\infty) = 0$, $F(+\infty, +\infty) = 1$ 2º F es monótona no decreciente en cada componente, esto quiere decir que:

Si $x_1 < x_2 \Rightarrow F(x_1, y) \leq F(x_2, y)\ \forall y \in \mathbb{R}$

Si $y_1 < y_2 \Rightarrow F(x, y_1) \leq F(x, y_2)\ \forall x \in \mathbb{R}$

3º F es continua a la derecha en cada componente, esto quiere decir que :

$$\lim_{\substack{\text{h} \to o \\ \text{h}>0}} F(x+h, y) = \lim_{h \to 0^+} F(x+h, y) = F(x, y) \quad \forall (x, y) \in \mathbb{R}^2$$

$$\lim_{\substack{\text{h} \to o \\ \text{h}>0}} F(x, y+h) = \lim_{h \to 0^+} F(x, y+h) = F(x, y) \quad \forall (x, y) \in \mathbb{R}^2$$

4º $\forall a, b, c, d \in \mathbb{R}$ tales que $a \leq b$ y $c \leq d$ se verifica:
$P(a < X \leq b, c < Y \leq d) = F(b, d) + F(a, c) - F(a, d) - F(b, c)$, además este valor ha de estar comprendido entre 0 y 1.

La expresión anterior nos da la forma de calcular la probabilidad a partir de la función de distribución. Si dibujamos el rectángulo obtenemos la figura 1

Consideremos el rectángulo $T = \{(x, y) \in \mathbb{R}^2 / a < x \leq b,\ c < y \leq d\}$.

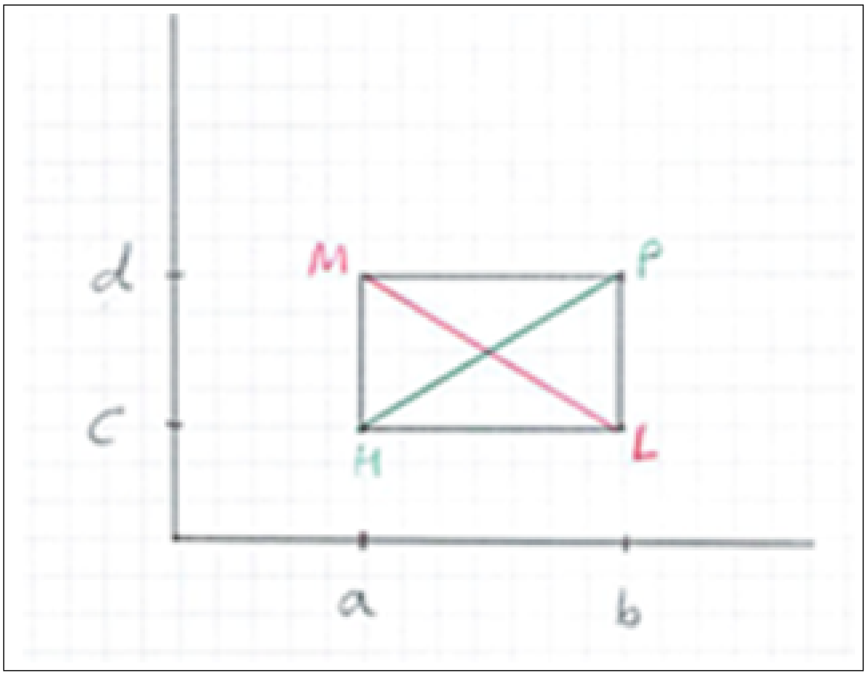


Figura 1: cálculo de la probabilidad a partir de la función de distribución

Una regla nemotécnica para la cuarta propiedad es la siguiente: *La probabilidad en el rectángulo T es la suma del valor de la función de distribución en los valores de la diagonal secundaria menos la suma del valor de la función de distribución en los vértices de la diagonal principal*, es decir $P(T) = F(H) + F(P) - [F(L) + F(M)]$

Variable aleatoria bidimensional discreta
Función de probabilidad conjunta

Si $Z = (X, Y)$ es una v.a. bidimensional discreta en el espacio de probabilidad $(\Omega, \mathcal{A}, P)$ se define su función de probabilidad como:
$P[X = x_i, Y = y_j] = P(x_i, y_j) = P_{i,j}$ y ha de verificar las siguientes propiedades:

i) $\forall i, j \Rightarrow 0 \leq P[X = x_i, Y = y_j] \leq 1$

ii) $\displaystyle\sum_i \sum_j P[X = x_i, Y = y_j] = 1$

Veamos un ejemplo: dada la función de probabilidad $P[X = x_i, Y = y_j] = = c(i + j),\ i = 1, 2, 3, 4\ \ j = 1, 2, 3$ obtenga el valor de c para que sea función de probabilidad conjunta de la variable (X, Y).

Solución:

Para que sea función de probabilidad conjunta ha de verificar que:

i)$0 \leq P[X = x_i, Y = y_j] \leq 1 \Rightarrow 0 \leq c(i + j) \leq 1 \Rightarrow c > 0$

ii)$\sum_i \sum_j P[X = x_i, Y = y_j] = 1$.Construimos la siguiente tabla:

X/Y	1	2	3	$\sum_i x_i$
1	2c	3c	4c	9c
2	3c	4c	5c	12c
3	4c	5c	6c	15C
4	5c	6c	7c	18c
$\sum_j y_j$	14c	18c	22c	54c

Tenemos que $\sum_i \sum_j P[X = x_i, Y = y_j] = 54c = 1 \Rightarrow c = \frac{1}{54}$

Función de distribución de una variable aleatoria bidimensional discreta

Sea $Z(\omega) = (X(\omega), Y(\omega))$ una v.a. bidimensional discreta en el espacio de probabilidad $(\Omega, \mathcal{A}, P)$. Se define su función de distribución como $F(x, y) = = P[X \leq x, Y \leq y] = \sum_{x_i \leq x} \sum_{y_j \leq y} P[X = x_i, Y = y_j]$

$F(x, y)$ verifica las siguientes propiedades vistas anteriormente pero que volvemos a repetir:

1º $F(-\infty, y) = F(x, -\infty) = 0,\ F(+\infty, +\infty) = 1$

2º F es monótona no decreciente en cada componente, esto quiere decir que:

Si $x_1 < x_2 \Rightarrow F(x_1, y) \leq F(x_2, y)\ \forall y \in \mathbb{R}$

Si $y_1 < y_2 \Rightarrow F(x, y_1) \leq F(x, y_2)\ \forall x \in \mathbb{R}$

3º F es continua a la derecha en cada componente, esto quiere decir que :

$$\lim_{\substack{h \to o \\ h>0}} F(x+h, y) = \lim_{h \to 0^+} F(x+h, y) = F(x, y) \quad \forall (x, y) \in \mathbb{R}^2$$

$$\lim_{\substack{h \to o \\ h>0}} F(x, y+h) = \lim_{h \to 0^+} F(x, y+h) = F(x, y) \quad \forall (x, y) \in \mathbb{R}^2$$

4º Sea T el rectángulo $a \leq b$ y $c \leq d$. Entonces la probabilidad en el mencionado rectángulo vale:
$P(a < X \leq b, c < Y \leq d) = F(b, d) + F(a, c) - F(a, d) - F(b, c)$.

Regionamientos en $\mathbb{R}^2$

Consideremos el rectángulo $T = \{(x, y) \in \mathbb{R}^2 / a \leq x \leq b,\ c \leq y \leq d\}$ Los regionamientos que se muestran en la figura 2 corresponden a las siguientes regiones:

Región I=$\{(x, y) \in \mathbb{R}^2\ /x < a\ o\ bien\ y < c\}$

Región II=$\{(x, y) \in \mathbb{R}^2\ /\ a \leq x \leq b,\ c \leq y \leq b\}$

Región III=$\{(x, y) \in \mathbb{R}^2\ /\ a \leq x \leq b,\ y > d\}$

Región IV=$\{(x, y) \in \mathbb{R}^2\ /\ x > b,\ c \leq y \leq b\}$

Región V=$\{(x, y) \in \mathbb{R}^2\ /\ x > a\ \ y > d\}$

Ejemplo:Obtenga la función de distribución conjunta para la variable bi-

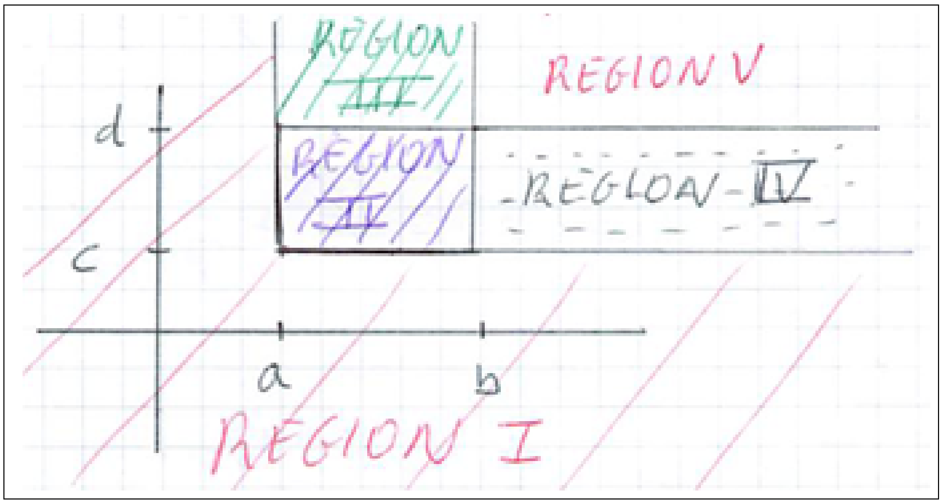


Figura 2: regionamientos

dimensional $Z = (X, Y)$ cuya función de probabilidad viene dada por la tabla:

X/Y	0	1
0	2/9	5/18
1	5/18	2/9

Solución:

Distribuir es sinónimo de repartir. Tenemos que dividir todo $\mathbb{R}^2$ en regiones para asignar probabilidad según esté localizado el punto en las distintas regiones.Dibujamos el campo de variación de X y el de Y (figura 3)

- Si $x < 0 \quad y < 0$ (región I) entonces:

$$F(x, y) = P[X \leq x, Y \leq y] = P[X = 0, Y = 0] = \frac{2}{9}$$

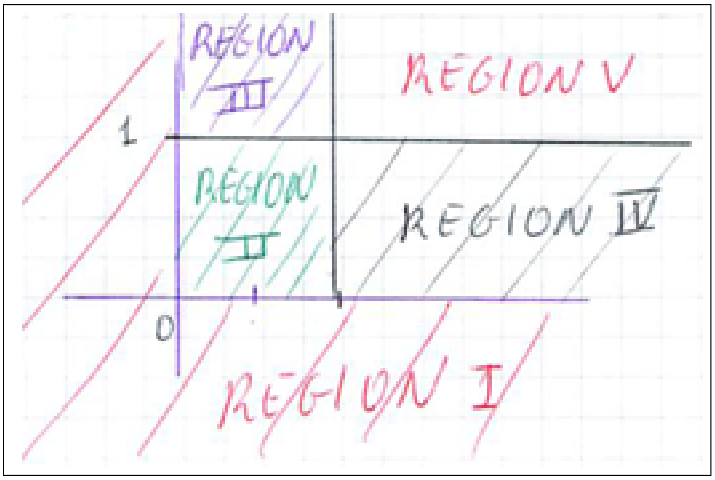


Figura 3:

- Si $0 \leq x < 1,\ \ 0 \leq y < 1$ (región II)

$$F(x,y) = P(X \leq x, Y \leq y) = P(x=0, y=0) = \frac{2}{9}$$

- Si $0 \leq x < 1,\ \ y \geq 1$ (región III)

$$F(x,y) = P(X=0, Y=1) = \frac{5}{18}$$

- Si $x \geq 1,\ \ 0 \leq y < 1$ (región IV)

$$F(x,y) = P(X=1, Y=0) = \tfrac{5}{18}$$

- Si $x \geq 1,\ \ y \geq 1$ (región V)

$$F(x,y) = P(X=1, Y=1) = \frac{2}{9}$$

Luego la función de distribución es:

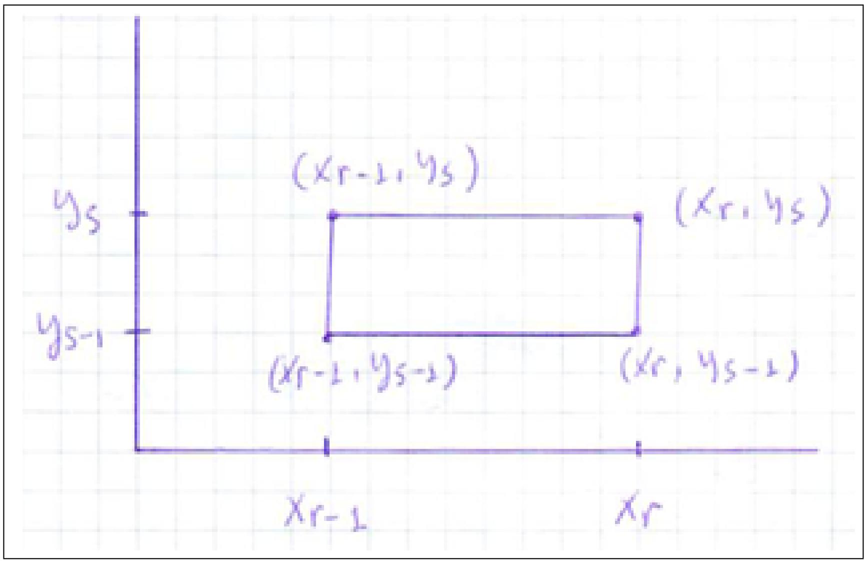


Figura 4:

$$F(x,y)=\begin{cases} 0 & \text{si } x<0 \ \ y<0 \\ \frac{2}{9} & \text{si } 0<1, \ 0\leq y<1 \\ \frac{5}{18} & \text{si } 0<1, \ y\geq 1 \\ \frac{5}{18} & \text{si } x\geq 10\leq y<1 \\ \frac{2}{9} & \text{si } x\geq 1, \ y\geq 1 \end{cases}$$

Una última cuestión referente a la v.a. bidimensional discreta. En variable unidimensional discreta, si disponíamos de la función de distribución $F(x)$, podíamos calcular la probabilidad en un punto x_i mediante la expresión $P[X=x_i]=F(x_i)-F(x_{i-1})$.¿Se puede hacer esto también en variable bidimensional?. Es decir, si conocemos la función de distribución de una variable bidimensional discreta $F(x,y)$, ¿cómo podemos calcular el valor de $P[X=x_r, Y=y_s]$?. Sean $x_{r-1}, \ y_{s-1}$ los valores inmediatamente anteriores a x_r *e* y_s respectivamente (figura 4). Entonces $P[X=x_r, Y=y_s]=$ $=F(x_r,y_s)+F(x_{r-1},y_{s-1})-F(x_r,y_{s-1})-F(x_{r-1},y_s)$ (regla nemotécnica del rectángulo).

Veamos un ejemplo. Sea la función de distribución de una v.a. bidimensional discreta dada por:

$$F(x,y)=\begin{cases} 0 & \text{si } x<0 & \text{ó } y<0 \\ \dfrac{x^2y+xy^2}{6} & \text{si } 0\le x<2, & 0\le y<1 \\ \dfrac{x^2+x}{6} & \text{si } 0\le x<2, & 1\le y \\ \dfrac{2y+y^2}{3} & \text{si } 2\le x, & 0\le y<1 \\ 1 & \text{si } 2\le x, & 1\le y \end{cases}$$

Se trata de calcular $P[X=2,Y=1]$. Los valores inmediatamente anteriores a $x=2$ e $y=1$ son $x=1$ e $y=0$. Si aplicamos a estos valores la fórmula anterior: $P[X=x_r,Y=y_s]==F(x_r,y_s)+F(x_{r-1},y_{s-1})-$ $-F(x_r,y_{s-1})-F(x_{r-1},y_s)$. Por tanto:
$P[X=2,Y=1]=F(2,1)+F(1,0)-F(2,0)-F(1,1)=1+0-0-\dfrac{2}{6}=\dfrac{2}{3}$

Variable aleatoria bidimensional: función de densidad

Sea $Z=(X,Y)$ una v.a. bidimensional continua en el espacio de probabilidad $(\Omega,\mathcal{A},P)$. Se dice que la función $f:\mathbb{R}^2\to\mathbb{R}$ es función de densidad de Z si verifica las siguientes propiedades:
i) $f(x,y)\ge 0 \quad \forall(x,y)\in\mathbb{R}^2$
ii) $\displaystyle\int_{-\infty}^{+\infty}(\int_{-\infty}^{+\infty} f(x,y)dy)dx=1$
iii) $P[a\le x\le b,\ c\le y\le d]=\displaystyle\int_a^b(\int_c^d f(x,y)dy)dx$

Veamos un ejemplo: se trata de obtener el valor de k para que sea función de densidad:

$$f(x,y)=\begin{cases} k(x+y) & \text{si } \le x\le 2 \quad \le y\le 1 \\ 0 & \text{en el resto} \end{cases}$$

<u>Solución</u>:

Para ser función de densidad ha de verificar dos condiciones:

* $f(x,y) \geq o \Rightarrow k > 0$

* $\displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dydx = 1 \Rightarrow \int_0^2 (\int_0^1 k(x+y)dy)dx = 1$. Se tiene que $\displaystyle\int_0^1 k(x+y)dy = kxy + k\frac{y^2}{2}|_0^1 = kx + \frac{k}{2}$. Por lo tanto $\int_0^2 kx + \frac{k}{2}dx = \frac{k}{2}x^2 + \frac{k}{2}x|_0^2 = 2k + k = 3k = 1 = \frac{1}{3}$. Para que f sea función de densidad ha de ser $k = \frac{1}{3}$

Función de distribución de una v.a bidimensional continua

Sea $Z = (X,Y)$ una v.a. bidimensional continua con función de densidad $f(x,y)$. Se define la función de distribución de Z como $F(x,y) = P[X \leq \leq x, Y \leq y] = P[-\infty < X \leq x, -\infty < Y \leq y] = \displaystyle\int_{-\infty}^{x}(\int_{-\infty}^{y} f(x,y)dy)dx$. Las propiedades que debe cumplir una función de distribución ya las hemos visto (pág. 3). Veamos un ejemplo
Se trata de obtener la función de distribución de una v.a. continua, sabiendo que su función de densidad es:

$$f(x,y) = \begin{cases} \frac{1}{3}(x+y) \text{ si} & 0 \leq x \leq 2, \quad 0 \leq y \leq 1 \\ 0 & \text{en el resto} \end{cases}$$

Para hallar la función de distribución hacemos la partición de $\mathbb{R}^2$ como se indica en la figura 5.
Si $x < 0 \;\; y < 0$ (región I) la función de densidad vale $f(x,y) = 0$ con lo que la función de distribución vale cero.

Si $0 \leq x < 2$, $0 \leq y < 1$ estamos en la región II. Aquí la función de distribución es $f(x,y) = \displaystyle\int_{-\infty}^{x}(\int_{-\infty}^{y}\frac{1}{3}(x+y)dy)dx = \int_0^x(\int_0^y \frac{1}{3}(x+y)dy)dx$. Se tiene

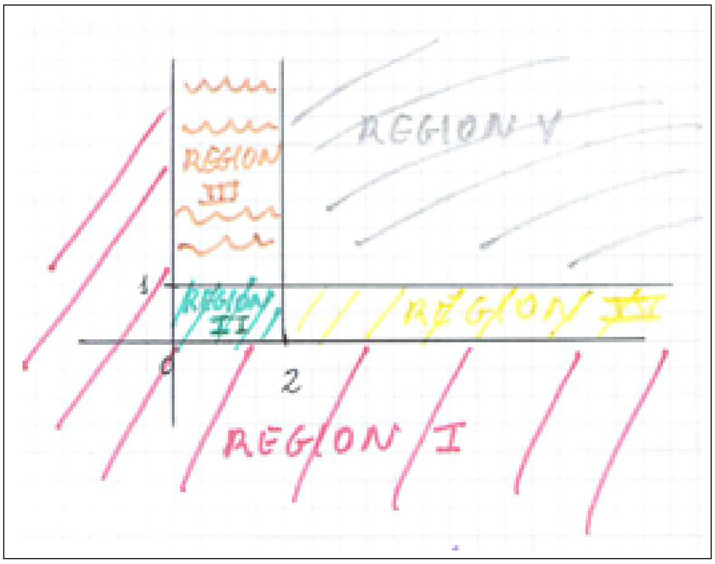


Figura 5:

que $\int_0^y \frac{1}{3}(x+y)dy = \frac{1}{3}x\int_0^y dy + \frac{1}{3}\int_0^y ydy = \left[\frac{1}{3}xy\right]_0^y + \frac{1}{3}\left[\frac{y^2}{2}\right]_0^y = \frac{1}{3}xy + \frac{1}{6}y^2$. Por lo tanto $F(x,y) = \int_0^x (\frac{1}{3}xy + \frac{1}{6}y^2)dx = \frac{1}{6}xy(x+y)$

Si $0 \leq x < 2,\ y \geq 1$ estamos en la región III. La función de distribución está dada por $F(x,y) = \int_0^x (\int_0^1 \frac{1}{3}(x+y)dy)dx$. Calculamos primero la integral entre paréntesis $\int_0^1 \frac{1}{3}(x+y)dy = \left[\frac{1}{3}xy + \frac{1}{6}y^2\right]_0^1 = \frac{1}{3}x + \frac{1}{6}$. Luego:
$F(x,y) = \int_0^x \frac{1}{3}x + \frac{1}{6}dx = \left[\frac{1}{6}x^2 + \frac{1}{6}x\right]_0^x = \frac{1}{6}x(x+1)$

Si $x \geq 2,\ 0 \leq y < 1$ estamos en la región IV y aquí la función de distribución vale $F(x,y) = \int_0^2 (\int_0^y \frac{1}{3}(x+y)dy)dx = \int_0^2 \frac{1}{3}xy + \frac{1}{6}y^2 dx = \left[\frac{1}{6}x^2y + \frac{1}{6}y^2x\right]_0^2 == \frac{2}{3}y + \frac{1}{3}y^2$

Si $x \geq 2,\ y \geq 1$ estamos en la región V y en esta zona la función de distribución vale $F(x,y) = \int_0^2 (\int_0^1 \frac{1}{3}(x+y)dy)dx = 1$. Luego la función de distribución vale:

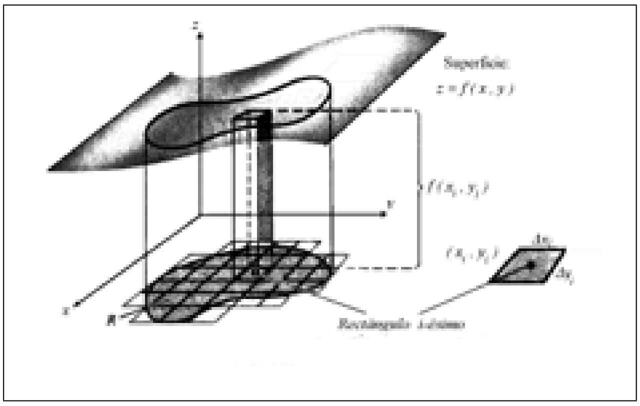


Figura 6: interpretación geométrica de la integral doble

$$F(x,y) = \begin{cases} 0 \text{ si} & x < 0 \text{ ó} \quad y < 0 \\ \dfrac{1}{6}xy(x+y) \text{ si} & 0 < 2 \ , \quad 0 \leq y < 1 \\ \dfrac{2}{3}y + \dfrac{1}{3}y^2 \text{ si} & x \geq 0 \ , \quad 0 \leq y < 1 \\ \dfrac{1}{6}x(x+1) \text{ si} & 0 < 2 \ , \quad y \geq 1 \\ 1 \text{ si} & x \geq 1 \ , \quad y \geq 1 \end{cases}$$

Obtención de la función de densidad a partir de la función de distribución

Sea $Z = X, Y)$ una v.a. bidimensional continua con función de distribución $F(x, y)$. Entonces la función de densidad viene dada por la expresión $f(x,y) = \dfrac{\partial^2 F(x,y)}{\partial x \partial y}$

Breve resumen de la integral doble

La integral doble $\int\int_D f(x,y)dxdy$ calcula el volumen del recinto que tiene como base el disco D y como altura f(x,y) (figura 6)

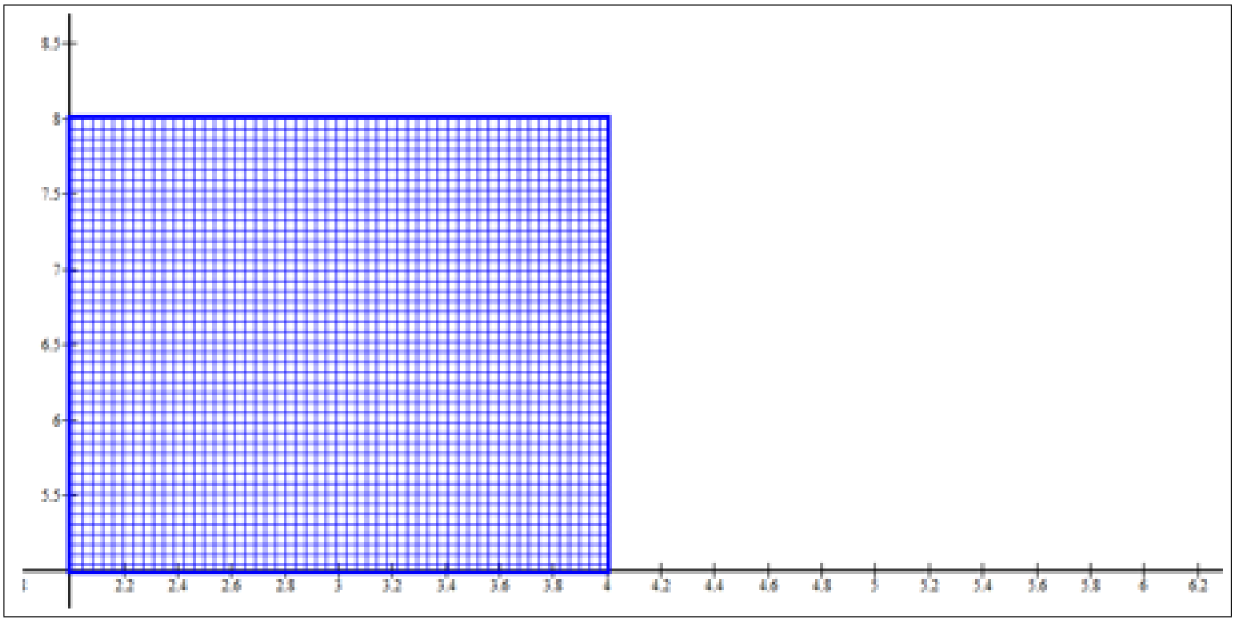


Figura 7: Recinto de integración del ejemplo Teorema de Fubini

Teorema de Fubini: expresa que en la integral doble no importa el orden de integración, es decir:

$$\int_a^b \left(\int_c^d f(x,y)dy\right)dx = \int_c^b \left(\int_a^b f(x,y)dx\right)dy$$

Veamos ejemplos:

Calcular $\displaystyle\int\int_D (2x^2+5xy)dxdy$, siendo $D = \{(x,y) \in \mathbb{R}^2 / 2 \leq x \leq 4,\ 5 \leq y \leq 8\}$

Este tipo de recintos son los más fáciles que se pueden presentar, se trata de un rectángulo. Es aconsejable dibujar el recinto.

En el recinto de variación de x tomamos un valor cualquiera x y miramos como varía y, o bien en el recinto de variación de y tomamos un valor y mirando como varía x, esto es lo que dice el Teorema de Fubini. En este caso la solución de la integral sería $\displaystyle\int_2^4 \left(\int_5^8 (2x^2+5xy)dy\right)dx \int_5^8 \left(\int_2^4 (2x^2+5xy)dx\right)dy$

Se tiene que $\displaystyle\int_5^8 (2x^2+5xy)dy = G(8) - G(5)$ (aquí x es constante). Como

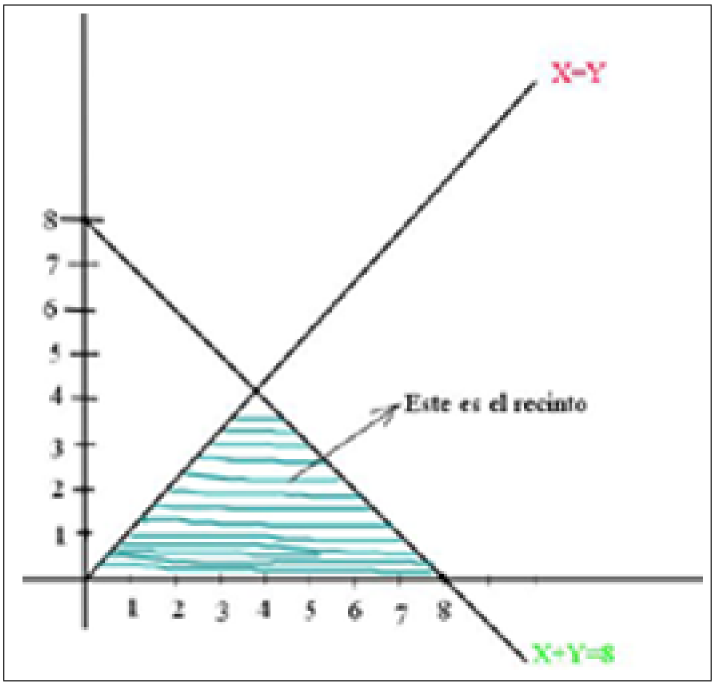


Figura 8: Representación gráfica del recinto D

$$G(y) = \int (2x^2 + 5xy)dy = 2x^2y + \frac{5}{2}xy^2, \text{ entonces: } \begin{cases} G(8) = 16x^2 + 160x \\ G(5) = 10x^2 + \dfrac{125}{2}x \end{cases}$$

Por tanto $\int_5^8 (2x^2 + 5xy)dy = 6x^2 + \frac{195}{2}x$. Para finalizar tenemos que:

$$\int_2^4 (\int_5^8 (2x^2+5xy)dy)dx = \int_2^4 (6x^2+\frac{195}{2}x)dx = \left| 2x^3 + \frac{195}{4}x^2 \right|_2^4 = 908 - 211 = 697$$

Ejemplo: Calcular $\iint_D xydxdy$, siendo:

$D = \{(x, y) \in \mathbb{R}^2 / 1 \leq x \leq 10,\ 0 \leq x - y,\ x + y \leq 8\}$
La solución de la integral viene dada por:

$$\int_0^4 (\int_0^x xy\ dy)dx + \int_4^8 (\int_0^{8-x} xy\ dy)dx$$

Distribuciones marginales

Cuando nos referimos a una v.a. bidimensional, de la que conocemos su comportamiento conjunto, nos preguntamos si es posible, a partir de esa información, conocer el comportamiento individual de alguna de sus variables, en definitiva, nos encontramos ante el problema de determinar como se distribuye cada una de las dos componentes sabiendo la distribución conjunta.

Función de probabilidad marginal en v.a. bidimensional discreta

Sea $Z = (X, Y)$ una v.a. bidimensional discreta con función de probabilidad $P(X = x_i, Y = y_i),\ i = 1, 2, ...r,\ j = 1, 2, ...s$. La función de probabilidad marginal de X es:

$$P[X = x_i] = P_{i\bullet} = \sum_{j=1}^{s} P[X = x_i, Y = y_j],\ i = 1, 2, ...r$$

De igual forma la probabilidad marginal de Y es:

$$P[Y = y_j] = P_{\bullet,j} = \sum_{i=1}^{r} P[X = x_i, Y = y_j],\ j = 1, 2, ...s$$

Veamos un ejemplo: **Obtenga las distribuciones marginales de probabilidad para la variable aleatoria bidimensional dada por la siguiente tabla:**

X $\backslash Y$	0	1
0	2/9	5/18
1	5/18	2/9

La función de probabilidad marginal de X es:

$$P[X = 0] = P[X = 0, Y = 0] + P[X = 0, Y = 1] = \frac{2}{9} + \frac{5}{18} = \frac{1}{2}$$

$$P[X = 1] = P[X = 1, Y = 0] + P[X = 1, Y = 1] = \frac{5}{18} + \frac{2}{9} = \frac{1}{2}$$

De igual manera la función de probabilidad marginal de Y es:

$$P[Y=0]=P[X=0,Y=0]+P[X=1,Y=0]=\frac{2}{9}+\frac{5}{18}=\frac{1}{2}$$

$$P[Y=1]=P[X=0,Y=1]+P[X=1,Y=1]=\frac{5}{18}+\frac{2}{9}=\frac{1}{2}$$

Esto lo podríamos haber obtenido de la tabla añadiendo una columna para la distribución marginal de X y una fila para la distribución marginal de Y de la siguiente forma:

X $\backslash Y$	0	1	Marginal X
0	$2/9$	$5/18$	$1/2$
1	$5/18$	$2/9$	$1/2$
Marginal Y	$1/2$	$1/2$	1

Veamos otro ejemplo. Sea la función de probabilidad conjunta: $P[X=i,\ Y=j]=\frac{1}{54}(i+j),\ i=1,2,3,....,\ j=1,2,3,.....$ Obtenga las funciones de probabilidad marginal de las variables X e Y.

La distribución marginal de X es:
$P[X=x_i]=\sum_{j=1}^{3}P[X=i,\ Y=j]=P[X=i,\ Y=1]+P[X=i,\ Y=2]++P[X=i,\ Y=3]=\frac{1}{54}(i+1+i+2+i+3)=\frac{3i+6}{54}=\frac{i+2}{18}$
La distribución marginal de Y es:
$P[Y=j]=\sum_{i=1}^{4}P[X=i,\ Y=j]=P[X=1,\ Y=j]+P[X=2,\ Y=j]+P[X=3,\ Y=j]+P[X=4,\ Y=j]=\frac{1}{54}(1+j+2+j+3+j+4+j)=$
$=\frac{1}{54}(10+4j)\Rightarrow P[Y=j]=\frac{5+2j}{27}\ \ j=1,2,3....$

Función de distribución marginal en v.a. bidimensional discreta

Sea $Z=(X,Y)$ una variable aleatoria bidimensional discreta con función de probabilidad $P[X=x_i,\ Y=y_j],\ i=1,2,....r,\ j=1,2,....s$. Entonces la función de distribución marginal de X es:

$F_1(x) = F(x, +\infty) = P[X \leq x,\ Y < +\infty] = \sum_{x_i \leq x} \sum_{j=1}^{s} P[X = x_i,\ Y = y_j]$

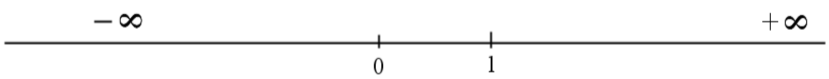
De manera análoga la distribución marginal de Y es:

$F_2(y) = F(+\infty, y) = P[X < +\infty,\ Y \leq y] = \sum_{i=1}^{r} \sum_{Y \leq y} P[X = x_i,\ Y = y_j]$

Veamos un ejemplo. Obtenga las funciones de distribución marginal para las variables aleatorias X e Y, sabiendo que la función de distribución conjunta es:

$$F(x,y) = \begin{cases} 0 \ \text{si} & x < 0 \ \ y < 0 \\ \frac{2}{9} \ si & 0 \leq x < 1,\ 0 \leq y < 1 \\ \frac{1}{2} \ si & 0 \leq x < 1,\ y \geqslant 1 \\ \frac{1}{2} \ si & \text{x} \geqslant 1,\ 0 \leq y < 1 \\ 1 \ \text{si} & x \geqslant 1,\ y \geqslant 1 \end{cases}$$

Veamos la distribución marginal de X. El campo de variación de X está dado por las siguientes regiones:

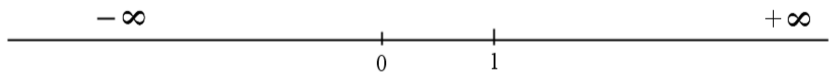


Como $Y \longrightarrow +\infty$ ya que calculamos $F_1(x) = F(x, +\infty)$ entonces:

$$F_1(x) = \begin{cases} 0 \ \text{si} & x < 0 \\ \frac{1}{2} \ \text{si} & 0 < 1 \\ 1 \ \text{si} & x \geqslant 1 \end{cases}$$

(Fijémonos en la definición de F(x,y) para obtener $\mathrm{F}_1(x,y)$)

Para hallar la distribución marginal de Y procedemos de forma similar, $F_2(y) = F(+\infty, y)$. El campo de variación de y viene dado por los mismos intervalos que x :



Entonces:

$$F_1(x) = \begin{cases} 0 \text{ si} & y < 0 \\ \dfrac{1}{2} \text{ si} & 0 < 1 \\ 1 \text{ si} & y \geqslant 1 \end{cases}$$

Función de densidad marginal en v.a. bidimensional continua

Sea $Z = (X, Y)$ una v.a. bidimensional contínua con función de densidad $f(x, y)$. Se define la función de densidad:

a) Marginal de X como $f_1(x) = \displaystyle\int_{-\infty}^{+\infty} f(x, y)dy, \ \forall x \in \mathbb{R}$

b) Marginal de Y como $f_2(y) = \displaystyle\int_{-\infty}^{+\infty} f(x, y)dx, \ \forall y \in \mathbb{R}$

Observemos que tanto f_1 como f_2 son funciones de densidad en una dimensión. Veamos un ejemplo

Obtenga las funciones de densidad marginal de x e y a partir de la función de densidad bidimensional:

$$f(x, y) = \begin{cases} \dfrac{1}{3}(x + y) & \text{si } \ 0 \leq x \leq 2 \quad 0 \leq y \leq 1 \\ & \\ 0 & \text{en el resto} \end{cases}$$

a) La función de densidad marginal de X es:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \int_0^1 \frac{1}{3}(x+y)dy = \frac{1}{3}\left[xy + \frac{1}{2}y^2\right]_0^1 = \frac{1}{3}x + \frac{1}{6}$$

Entonces la función marginal de X es $f_1(x) = \begin{cases} \frac{1}{3}x + \frac{1}{6} & \text{si } 0 \leq x \leq 2 \\ & \\ 0 & \text{en el resto} \end{cases}$

b) La función de densidad marginal de Y es:

$$f_2(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \int_0^2 \frac{1}{3}(x+y)dy = \frac{1}{3}\left[\frac{x^2}{2} + xy\right]_0^2 = \frac{2}{3} + \frac{2}{3}y$$

Entonces la función marginal de Y es $f_2(y) = \begin{cases} \frac{2}{3} + \frac{2}{3}y & \text{si } 0 \leq y \leq 1 \\ & \\ 0 & \text{en el resto} \end{cases}$

Función de distribución marginal en v.a. bidimensional continua

Sea $Z = (X,Y)$ una v.a. bidimensional continua con función de densidad conjunta f(x,y). Se define la función de distribución marginal de X como: $F_1(x) = F(x,+\infty) = \int_{-\infty}^{x} (\int_{-\infty}^{+\infty} f(x,y)dy)dx$. Al ser $f_1(x) = \int_{-\infty}^{x} f(x,y)dy$, podemos poner que $F_1(x) = \int_{-\infty}^{x} f_1(x)dx \ \forall x \in \mathbb{R}$. De modo análogo la función de distribución marginal de Y viene dada por:
$F_2(y) = F(-\infty,y) = \int_{-\infty}^{+\infty} (\int_{-\infty}^{y} f(x,y)dy)dx = \int_{-\infty}^{y} f_2(y)dy \ \forall y \in \mathbb{R}$

Ejemplo: Obtenga las funciones de distribución marginales de X e Y siendo la función de distribución:

$$F(x,y) = \begin{cases} 0 & \text{Si} \quad x < 0 \ o \ y < 0 \\ \dfrac{x^2 + xy^2}{6} & \text{Si} \quad 0 \leq x < 2, \ 0 \leq y < 1 \\ \dfrac{x^2 + x}{6} & \text{Si} \quad 0 \leq x < 2, \ y \geq 1 \\ \dfrac{2y + y^2}{3} & \text{Si} \quad x \geq 2, \ 0 \leq y < 1 \\ 1 & \text{Si} \quad x \geq 2, \ y \geq 1 \end{cases}$$

Solución:

Distribución marginal de X

- Si $x < 0 \Rightarrow F_1(x) = F(x, +\infty) = 0$

- Si $0 \leq x \leq 2 \Rightarrow F_1(x) = F(x, +\infty) = \dfrac{x^2 + x}{6}$ ya que ha de ser $y \geq 1$

- Si $x \geq 2 \Rightarrow F_1(x) = F(x, +\infty) = 1$ al ser $y \geq 1$

Luego la función de distribución marginal de x es:

$$F_1(x) = \begin{cases} 0 & \text{Si} \quad x < 0 \\ \dfrac{x^2 + x}{6} & \text{Si} \quad 0 \leq x < 2 \\ 1 & \text{Si} \quad x \geq 2 \end{cases}$$

Distribución marginal de Y

- Si $y < 0 \Rightarrow F_2(y) = F(-\infty, y) = 0$

- Si $0 \leq y < 1 \Rightarrow F_2(y) = F(-\infty, y) = \dfrac{2y + y^2}{3}$ al ser $x \geq 2$

- Si $y \geq 1$ al ser $x \geq 2 \Rightarrow F_2(y) = 1$

Luego la función de distribución marginal de Y es:

$$F_2(y) = \begin{cases} 0 & \text{Si} \quad y < 0 \\ \dfrac{2y + y^2}{3} & \text{Si} \quad 0 \le y < 1 \\ 1 & \text{Si} \quad y \ge 1 \end{cases}$$

Distribuciones condicionadas

Lo que pretende el estudio de las distribuciones condicionadas es ver como se opera con la variable aleatoria bidimensional cuando se impone una determinada condición a una de sus componentes.

Sea $Z = (X, Y)$ una v.a. bidimensional y supongamos que se impone a Y una determinada condición dada por el conjunto A. El cálculo de la función de distribución viene dado por:

$$F(x/Y \in A) = \frac{P[X \le x,\ y \in A]}{P[Y \in A]}$$

Hay dos casos particularmente importantes:

a) Función de distribución de X condicionada a $Y \le y$

$$F(x/Y \le y) = \frac{P[X \le x,\ Y \le y]}{P[Y \le y]} = \frac{F(x, y)}{F_2(y)}$$

Ha de cumplirse que $P[Y \le y] = F_2(y) \ne 0$

b) Función de distribución de Y condicionada a $X \le x$ es:

$$\mathrm{F}(y/X \le x) = \frac{P(X \le x,\ Y \le y)}{P(X \le x)}$$

hemos de imponer la condición de que $P(X \le x) \ne 0$

Se pueden formular una gran cantidad de casos de distribuciones condicionadas pero todos se resuelven de la misma manera.

Distribuciones condicionadas discretas.- Sea $Z = (X, Y)$ una v.a. bidimensional discreta con función de probabilidad:

$$P_{ij} = P(X = x_{ij}, Y = y_{ij}),\ i = 1, 2, 3, ...,\ j = 1, 2, 3, ...$$

a) Las funciones de probabilidad condicionada son:

$$Pi/j = P(X = x_i/Y = y_j) = \frac{P(X = x_i,\ Y = y_j)}{P(Y = y_j)} = \frac{P_{ij}}{P_{\bullet j}}$$

$$Pj/i = P(X = x_i/Y = y_j) = \frac{P(X = x_i,\ Y = y_j)}{P(X = x_i)} = \frac{P_{ij}}{P_{i\bullet}}$$

Veamos un ejemplo: se trata de calcular las funciones de probabilidad condicionada siendo la función de probabilidad conjunta $P(X = i,\ Y = j) =$ $= \dfrac{1}{54}(i + j),\ i = 1, 2, 3, 4;\ j = 1, 2, 3$

<u>Distribución de probabilidad de X condicionada a Y=j</u>

$$P(X = i/Y = j) = \frac{P(X = i, Y = j)}{P(Y = j)} = \frac{\frac{1}{54}(i + j)}{\displaystyle\sum_{i=1}^{4} P(X = i, Y = j)} = \frac{\cancel{\frac{1}{54}}(i + j)}{\displaystyle\sum_{i=1}^{4} \cancel{\frac{1}{54}}(i + j)} =$$

$$= \frac{i + j}{10 + 4j},\ j = 1, 2, 3$$

<u>Distribución de probabilidad de Y condicionada a X=i</u>

$$P(Y = j/X = i) = \frac{P(X = i, Y = j)}{P(X = i)} = \frac{\cancel{\frac{1}{54}}(i + j)}{\displaystyle\sum_{j=1}^{3} \cancel{\frac{1}{54}}(i + j)} == \frac{i + j}{6 + 3i},\ i = 1, 2, 3, 4$$

b) Las funciones de distribución condicionada son:

$$F(x/y_j) = P(X \leq x/Y = y_j) = \frac{\displaystyle\sum_{x_i \leq x} P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{\displaystyle\sum_{x_i \leq x} P(X = x_i, Y = y_j)}{P_{\bullet j}} =$$

$= \displaystyle\sum_{x_i \leq x} \frac{P_{ij}}{P_{\bullet j}} = \sum_{x_i \leq x} P_{i/j}$. De la misma forma tenemos que:

$$F(y/x_i) = P(Y \leq y/X = x_i) = \frac{\sum_{y_j \leq y} P(X = x_i, Y = y_j)}{P(X = x_i)} = \frac{\sum_{y_j \leq y} P(X = x_i, Y = y_j)}{P_{i\bullet}} =$$

$$= \sum_{y_i \leq y} \frac{P_{ij}}{P_{i\bullet}} = \sum_{y_i \leq y} P_{j/i}$$

Distribuciones condicionadas en v.a bidimensional continua

Sea $Z = (X, Y)$ una v.a. bidimensional continua con función de densidad $f(x, y)$. Entonces la función de densidad marginal de x es $f_1(x) = = \int_{-\infty}^{+\infty} f(x, y)dy$. La función de densidad marginal de y será $f_2(y) = \int_{-\infty}^{+\infty} f(x, y)dx$. Se definen las funciones de densidad condicionada como:

$$f(x/y) = \frac{f(x, y)}{f_2(y)},\ f(y/x) = \frac{f(x, y)}{f_1(x)}$$

La razón de definir de esta manera la función de densidad condicionada es por la analogía entre la función de probabilidad en variable discreta y la función de densidad en variable continua.

Función de distribución condicionada en una v.a. bidimensional continua

Se definen las funciones de distribución condicionada como:

$$F(x/y) = \int_{-\infty}^{x} f(x/y)dx = \int_{-\infty}^{x} \frac{f(x, y)}{f_2(y)}dx$$
$$F(y/x) = \int_{-\infty}^{y} f(y/x)dy = \int_{-\infty}^{y} \frac{f(x, y)}{f_1(x)}dy$$

Veamos un ejemplo.- Sea $Z = (X, Y)$ una v.a. bidimensional continua con función de densidad conjunta :

$$f(x, y) = \begin{cases} \frac{1}{3}(x + y) & \text{Si} \quad 0 \leq x \leq 2 \quad 0 \leq y \leq 1 \\ 0 & \quad \text{en el resto} \end{cases}$$

Tenemos que obtener la función de densidad y la función de distribución de Y condicionada a $X = x$

Si $o \leq x \leq 2,\ 0 \leq y \leq 1$, tenemos que $f(y/x) = \dfrac{f(x,y)}{f_1(x)}$. Como se verifica que $f_1(x) = \displaystyle\int_{-\infty}^{+\infty} f(x,y)dy = \int_0^1 \frac{1}{3}(x+y)dy = \left[\frac{1}{3}xy + \frac{1}{6}y^2\right]_0^1 \Rightarrow f_1(x) =$ $= \dfrac{1}{3}x + \dfrac{1}{6} = \dfrac{1}{3}(x + \dfrac{1}{2})$. Luego la función de densidad condicionada es:

$f(y/x) = \dfrac{\frac{1}{3}(x+y)}{\frac{1}{3}(x+\frac{1}{2})} = \dfrac{x+y}{x+\frac{1}{2}} = \dfrac{2(x+y)}{2x+1}$. Nos piden la función de densidad de y dado un valor $X = x$, que será entonces:

$$f(y/x) = \begin{cases} \dfrac{2(x+y)}{2x+1} & \text{Si} \quad 0 \leq y \leq 1 \\ 0 & \text{en el resto} \end{cases}$$

Siendo $0 \leq x \leq 2$

Para calcular la función de distribución tenemos que tener en cuenta los intervalos de variación de y, que son los que indica la figura siguiente:

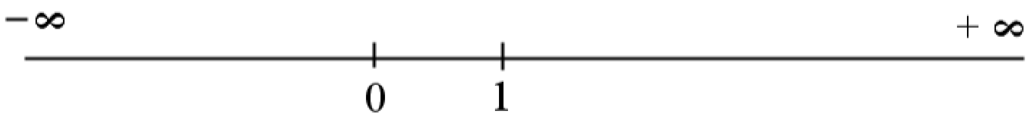


Figura 9: Variación de Y

También hemos de tener en cuenta que $F(y/x) = \displaystyle\int_{-\infty}^{y} f(y/x)dy$

- *Si* $y < 0 \Rightarrow F(y/x) = 0$

- *Si* $0 \leq y < 1 \Rightarrow F(y/x) = \displaystyle\int_{-\infty}^{+\infty} f(y/x)dy = \int_0^y \frac{2(x+y)}{2x+1} = \frac{2xy + y^2}{2x+1}$

ya que:

$\displaystyle\int \frac{2(x+y)}{2x+1}dy = \frac{2}{2x+1}\int (x+y)dy = \frac{2}{2x+1}\left(x+y\frac{y^2}{2}\right) = \frac{2xy}{2x+1} +$
$\displaystyle+\frac{y^2}{2x+1} = \frac{2xy+y^2}{2x+1}$

$\bullet$ *Si* $y \geq 1 \Rightarrow F(y/x) = \displaystyle\int_{-\infty}^{+\infty} f(y/x)dy = \int_0^1 \frac{2(x+y)}{2x+1}dy = \left[\frac{2xy+y^2}{2x+1}\right]_0^1 = 1$

Luego la función de distribución condicionada para $0 \leq x \leq 2$ es:

$$F(y/x) = \begin{cases} 0 & \text{Si } y < 0 \\ \dfrac{2xy+y^2}{2x+1} & \text{Si } 0 \leq y < 1 \\ 1 & \text{Si } y \geq 1 \end{cases}$$

INDEPENDENCIA DE VARIABLES ALEATORIAS

a) **Independencia de variables aleatorias discretas.-** Sea $Z = (X, Y)$ una v.a. discreta con función de probabilidad $P_{ij} = P(X = x_i, Y = y_j)$. Se dice que X e Y son independientes si:

$P(X = x_i / Y = y_j) = P(X = x_i)$

$P(Y = y_j / X = x_i) = P(Y = y_j)$

Proposición.- Si X e Y son independientes entonces se verifica que:

$$P_{ij} = P_{i\bullet} P_{\bullet j}$$

<u>Demostración:</u>

Según la definición dada, si X e Y son independientes $P(X = x_i / Y = y_j) = P(X = x_i) \Rightarrow$
$\Rightarrow \dfrac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = P(X = x_i)$ Por tanto tenemos:

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j) \Rightarrow P_{ij} = P_{i\bullet}P_{\bullet j}$$

Veamos un ejemplo. Sea la función de probabilidad conjunta de una v.a. bidimensional discreta $Z = (X, Y)$, $P(X = i, Y = j) = \dfrac{1}{54}(i + j),\ i = 1, 2, 3, 4\ j = 1, 2, 3$. Se trata de determinar si las variables X e Y son independientes.

Las variables son independientes si verifican $P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j) \Rightarrow P_{ij} = P_{i\bullet}P_{\bullet j}$. Vamos a obtener las distribuciones marginales

$$P_{i\bullet} = P(X = i) = \sum_{j=1}^{3} \frac{1}{54}(i + j) = \frac{i + 2}{18},\ i = 1, 2, 3, 4$$
$$P_{\bullet j} = P(Y = j) = \sum_{i=1}^{4} \frac{1}{54}(i + j) = \frac{2j + 5}{27},\ j = 1, 2, 3$$

La igualdad $P_{ij} = P_{i\bullet}P_{\bullet j}$ se debe verificar $\forall i, j$. Si tomamos $i = j = 1$ se tiene que $P_{11} = \dfrac{1}{54}(1 + 1) = \dfrac{1}{27};\ P_{i\bullet} = P_{1\bullet} = \dfrac{3}{18} = \dfrac{1}{6};\ P_{\bullet j} = \dfrac{7}{27}$. Como $P_{11} \neq P_{1\bullet}P_{\bullet 1}$ concluimos afirmando que X e Y no son independientes.

Proposición.- Sea $Z = (X, Y)$ una v.a. bidimensional discreta con función de probabilidad $P(X = x_i, Y = y_j)$. Entonces su función de distribución $F(X, Y)$ verifica la siguiente propiedad:

$$\text{X e Y son independientes} \Longleftrightarrow F(x, y) = F_1(x) \cdot F_2(y)$$

La razón de esto es que:

$$F(x, y) = \sum_{x_i \leq x} \sum_{y_j \leq y} P(X = x_i, Y = y_j) = \sum_{x_i \leq x} \sum_{y_j \leq y} P(X = x_i) \cdot P(Y = y_j) \Rightarrow$$
$$\Rightarrow F(x, y) = \sum_{x_i \leq x} P(X = x_i) \cdot \sum_{y_j \leq y} P(Y = y_j) = F_1(x) \cdot F_2(y)$$

b) **Independencia de variables aleatorias continuas.** Sea $Z = (X, Y)$ una v.a. bidimensional continua con función de densidad conjunta $f(x, y)$ y con funciones de densidad marginal de X $f_1(x)$, marginal de Y $f_2(y)$. Se dice que X e Y son independientes si se verifica que $f(x/y) = f_1(x);\ f(y/x) = f_2(y)$.

Como $f(x/y) = \dfrac{f(x,y)}{f_2(y)}$; $f(y/x) = \dfrac{f(x,y)}{f_1(x)}$, podemos decir que si X e Y son independientes:

$$f(x,y) = \frac{f(x,y)}{f_2(y)} = f_1(x) \Longleftrightarrow f(x,y) = f_1(x) \cdot f_2(y)$$

Inversamente, si $f(x,y) = f_1(x) \cdot f_2(y) \Longleftrightarrow f(x/y) = \dfrac{f(x,y)}{f_2(y)} = f_1(x)$ y de forma análoga también se cumple que $f(y/x) = f_2(y)$ con lo que X e Y son independientes. De todo esto se desprende la siguiente proposición

Proposición.- X e Y son independientes $\Longleftrightarrow f(x,y) = f_1(x) \cdot f_2(y)$

lo mismo se puede decir de la función de distribución para la v.a. continua

Proposición.- X e Y son independientes $\Longleftrightarrow F(x,y) = F_1(x) \cdot F_2(y)$

Veamos un ejemplo. Sea la v.a. bidimensional con función de densidad conjunta:

$$f(x,y) = \begin{cases} \frac{1}{3}(x+y) & \text{Si } 0 \leq x \leq 2 \quad 0 \leq y \leq 1 \\ 0 & \text{en el resto} \end{cases}$$

Tenemos que determinar si las variables X e Y son independientes. Anteriormente ya hemos visto que las funciones de densidad marginales eran:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)dy_1(x) = \frac{2x+1}{6} \; si \; 0 \leq x \leq 2$$
$$f_2(y) = \int_{-\infty}^{+\infty} f(x,y)dx_2(y) = \frac{2(1+y)}{3} \; si \; 0 \leq y \leq 1$$

Fuera del rectángulo $0 \leq x \leq 2$; $0 \leq y \leq 1$ se cumple $f(x,y) = 0$, $f_1(x) = f_2(y) = 0$, por tanto se cumple que $f(x,y) = f_1(x) \cdot f_2(y)$

Dentro del rectángulo $0 \leq x \leq 2$; $0 \leq y \leq 1$ se verifica $f(x,y) \neq$ $\neq f_1(x) \cdot f_2(y)$ ya que $\frac{1}{3}(x+y) \neq \dfrac{2x+1}{6} \cdot \dfrac{2(1+y)}{3}$, por tanto las variables X e Y no son independientes