Estadística y Procesos Estocásticos
Tema 3: Distribuciones Multivariantes

Grado en Ingeniería en Tecnologías de la Telecomunicación

3. Distribución conjunta de dos Variables Aleatorias Continuas.

Distribución conjunta de dos Variables Aleatorias Continuas.

Dadas dos variables aleatorias X e Y definidas sobre un mismo espacio de probabilidad $(\Omega, \mathcal{F}, P)$, la **función de distribución conjunta** de (X, Y) es:

$$F(s,t) = \Pr(X \leq s, Y \leq t) = \Pr(\{X \leq s\} \cap \{Y \leq t\})$$

En el caso de que X e Y sean continuas, si existe una función $f : \mathbb{R}^2 \to \mathbb{R}$, **integrable** y **no negativa**, tal que:

$$F(s,t) = \int_{-\infty}^{s} \int_{-\infty}^{t} f(x,y)\, dx dy$$

entonces la función de distribución $F(s,t)$ es **absolutamente continua** y $f(s,t)$ recibe el nombre de **función de densidad de probabilidad** del vector (X, Y).

Propiedades de la función de densidad bivariante

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

$$\Pr\left(a < X \leq b, c < Y \leq d\right) = \int_{a}^{b}\int_{c}^{d} f\left(x,y\right)dxdy$$

Si $f(x,y)$ es continua,

$$\frac{\partial^2}{\partial x \partial y} F\left(x,y\right) = f\left(x,y\right)$$

Ejemplo:

En estudios sobre la actividad neurológica del cerebro es habitual medir señales eléctricas en la corteza cerebral. En ausencia de estímulos externos, la señal medida en cierta región tiene una amplitud X y una fase Y aleatorias que se distribuyen de acuerdo a la función de densidad:

$$f(x,y) = \lambda xye^{-y^2},\ 0 \leq x \leq 1,\ \ 0 \leq y \leq \pi$$

1. Determinar el valor de λ para que la función anterior esté bien definida.

1. Determinar la probabilidad de que la señal registrada tenga una amplitud mayor que 0.5 y una fase entre $\pi/2$ y π.

Ejemplo:

1. El valor de la integral de la función de densidad sobre el rango en que están definidas las variables debe ser 1. Por tanto:

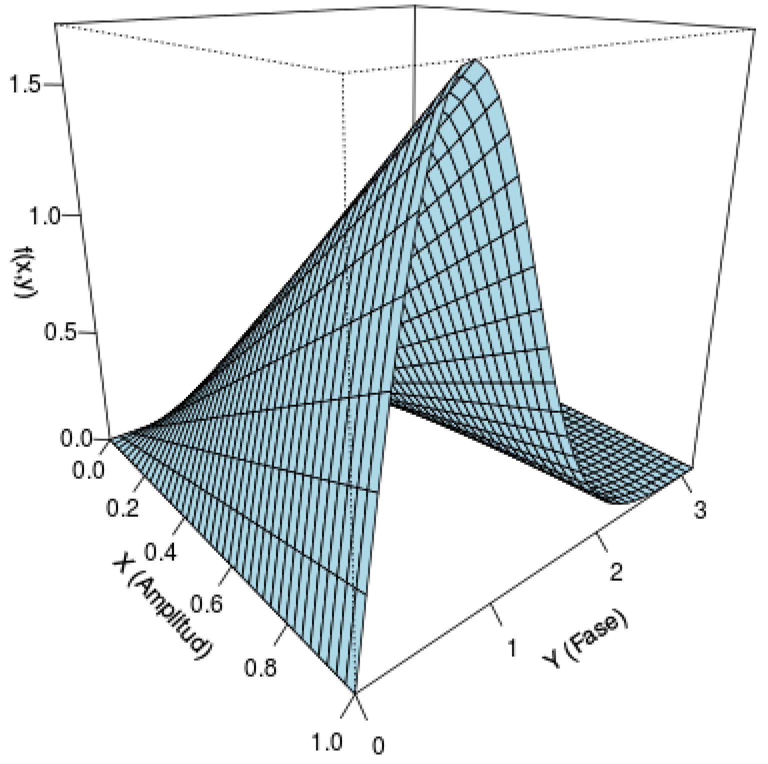
$$1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dxdy = \int_0^1 \int_0^\pi \lambda x y e^{-y^2} dxdy =$$

$$= \lambda \int_0^1 x \left\{ \int_0^\pi y e^{-y^2} dy \right\} dx = \lambda \int_0^1 x \left[-\frac{1}{2} e^{-y^2} \right]_0^\pi dx =$$

$$= -\frac{\lambda}{2} \left(e^{-\pi^2} - 1 \right) \left[\frac{x^2}{2} \right]_0^1 = \frac{\lambda \left(1 - e^{-\pi^2} \right)}{4} \implies \lambda = \frac{4}{1 - e^{-\pi^2}}$$

La figura siguiente muestra la representación gráfica de esta función de densidad:

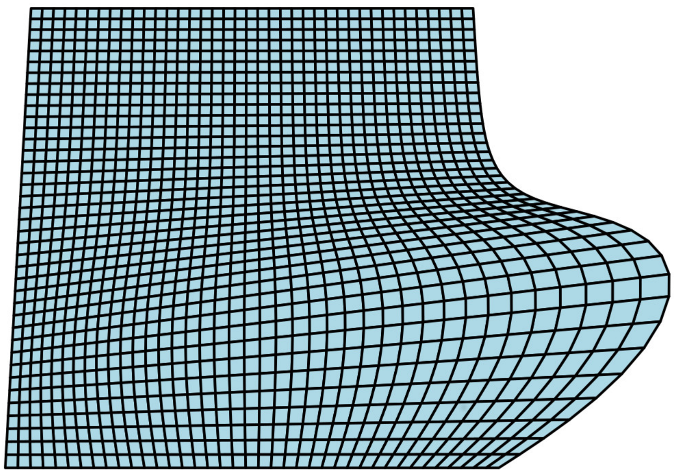
$$f(x,y) = \frac{4}{1-e^{-\pi^2}} xye^{-y^2}$$

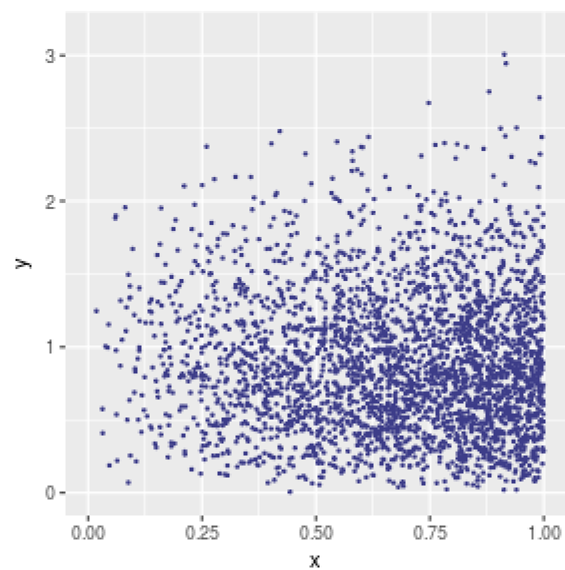


Pulsar aquí para rotar esta figura de forma interactiva

A continuación se muestra la función de densidad $f(x, y)$ vista "desde arriba", así como una muestra aleatoria de 5000 observaciones (puntos) de la variable aleatoria bidimensional (X, Y); puede verse que:

- la densidad (observada) de puntos es mayor donde la densidad de probabilidad (teórica) es más alta.
- donde la densidad de probabilidad es cero, no hay observaciones de (X, Y)





Ejemplo:

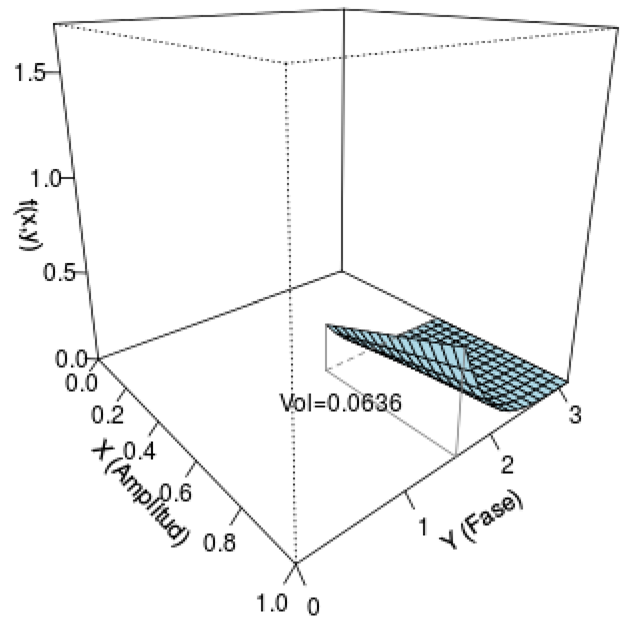
2. Determinar la probabilidad de que la señal registrada tenga una fase entre $\pi/2$ y π y una amplitud mayor que 0.5.

$$P\left(X > 0.5, \frac{\pi}{2} \leq Y \leq \pi\right) = \frac{4}{1-e^{-\pi^2}} \int_{0.5}^{1} \int_{\pi/2}^{\pi} xye^{-y^2} dxdy =$$

$$\frac{4}{1-e^{-\pi^2}} \int_{0.5}^{1} x \left[-\frac{1}{2} e^{-y^2}\right]_{\pi/2}^{\pi} dx = \frac{2}{1-e^{-\pi^2}} \int_{0.5}^{1} x \left(e^{-(\pi/2)^2} - e^{-\pi^2}\right) dx =$$

$$\frac{2}{1-e^{-\pi^2}} \left(e^{-(\pi/2)^2} - e^{-\pi^2}\right) \left[\frac{x^2}{2}\right]_{1/2}^{1} = \frac{2}{1-e^{-\pi^2}} \left(e^{-(\pi/2)^2} - e^{-\pi^2}\right) \left(\frac{1}{2} - \frac{1}{8}\right) = 0.0636$$

Esta probabilidad corresponde al volumen bajo la función de densidad sobre la región $0.5 \leq X \leq 1$, $\frac{\pi}{2} \leq Y \leq \pi$, tal como se muestra en la figura siguiente:



class: split-25

La proporción (observada) de puntos de la muestra anterior que caen en la región $0.5 \leq X \leq 1$, $\frac{\pi}{2} \leq Y \leq \pi$ es 0.06867, que se aproxima bastante al valor teórico 0.0636.

Función de densidad marginal

Si $f(x, y)$ es la función de densidad del vector aleatorio (X, Y), la **función de densidad marginal de X** se obtiene integrando (sumando) sobre todos los posibles valores de y:

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

Análogamente, la **densidad marginal de Y** se obtiene integrando (sumando) sobre todos los posibles valores de x:

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

Ejemplo:

Calculemos las densidades marginales de:

$$f(x,y) = \frac{4}{1-e^{-\pi^2}} xye^{-y^2}$$

1. Densidad marginal de X:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^{\pi} \frac{4}{1-e^{-\pi^2}} xye^{-y^2} dy = \frac{4x}{1-e^{-\pi^2}} \int_0^{\pi} ye^{-y^2} dy =$$

$$= \frac{4x}{1-e^{-\pi^2}} \left[-\frac{1}{2} e^{-y^2} \right]_0^{\pi} = \frac{4x}{1-e^{-\pi^2}} \frac{1}{2} \left(1 - e^{-\pi^2} \right) = 2x$$

2. Densidad marginal de Y:

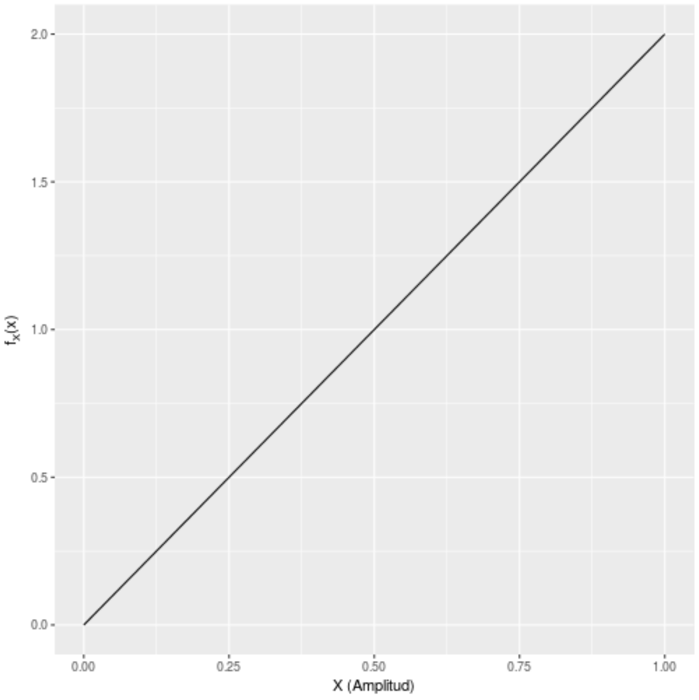
$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^{1} \frac{4}{1-e^{-\pi^2}} xye^{-y^2} dx = \frac{4ye^{-y^2}}{1-e^{-\pi^2}} \int_0^{1} xdx =$$

$$= \frac{4ye^{-y^2}}{1-e^{-\pi^2}} \left[\frac{x^2}{2} \right]_0^{1} = \frac{4ye^{-y^2}}{1-e^{-\pi^2}} \frac{1}{2} = \frac{2ye^{-y^2}}{1-e^{-\pi^2}}$$

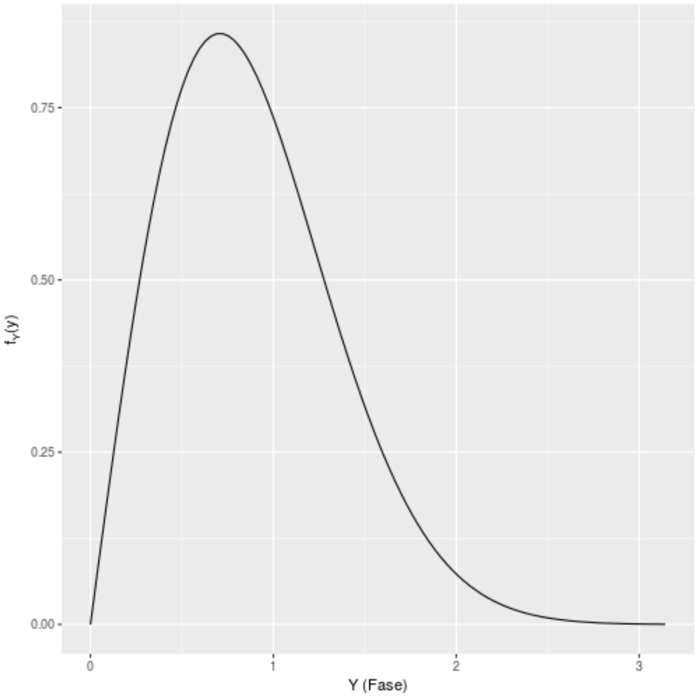
Ejemplo:

Gráficamente:

$f_X(x) = 2x$



$f_Y(y) = \frac{2ye^{-y^2}}{1-e^{-\pi^2}}$



Distribuciones de probabilidad condicionales.

La función de distribución acumulativa de Y condicionada por el valor de X se define como:

$$F_{Y|X=x}(y) = \Pr(Y \leq y \mid X = x)$$

En el caso discreto:

$$F_{Y|X=x}(t) = \Pr(Y \leq t \mid X = x) = \frac{P(Y \leq t, X = x)}{P(X = x)} = \sum_{y \leq t} \frac{P(Y = y, X = x)}{P(X = x)}$$

En el caso continuo sustituimos la suma por la integral (siempre que exista $f(x,y)$):

$$F_{Y|X=x}(t) = \Pr(Y \leq t \mid X = x) = \int_{-\infty}^{t} \frac{f(x,y)}{f_X(x)} dy$$

Distribuciones de probabilidad condicionales.

Derivando en la función de distribución condicionada, se obtiene la **función de densidad de Y condicionada por X**:

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

Análogamente, la **función de densidad de X condicionada por Y** es:

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

Ejemplo:

$$f\left(x,y\right)=\frac{4}{1-e^{-\pi^2}}xye^{-y^2}$$

1. Densidad condicional de *X* dado el valor de *Y*:

$$f_{X|Y}\left(x|y\right)=\frac{f\left(x,y\right)}{f_Y(y)}=\frac{\frac{4}{1-e^{-\pi^2}}xye^{-y^2}}{\frac{2ye^{-y^2}}{1-e^{-\pi^2}}}=2x$$

2. Densidad condicional de *Y* dado el valor de *X*:

$$f_{Y|X}\left(y|x\right)=\frac{f\left(x,y\right)}{f_X(x)}=\frac{\frac{4}{1-e^{-\pi^2}}xye^{-y^2}}{2x}=\frac{2ye^{-y^2}}{1-e^{-\pi^2}}$$

Nótese en este ejemplo que:

- Las densidades condicionales coinciden con las marginales (ésto no tiene por qué ocurrir siempre)
- La densidad de X condicionada por Y no depende de Y, y la densidad de Y condicionada por X no depende de X. (ésto tampoco ocurre siempre)

4. Independencia de Variables Aleatorias

Independencia de variables aleatorias.

Dados dos sucesos A y B de un espacio muestral Ω, ya habíamos visto que son **independientes** si B **no informa** sobre A, esto es:

$$P(A|B) = P(A)$$

De donde se seguía que si A y B son independientes, entonces:

$$P(A \cap B) = P(A) \cdot P(B)$$

De modo completamente análogo, dado un vector (X, Y) de variables aleatorias, ambas son **independientes** si:

- En el caso discreto:

$$P(X = x|Y = y) = P(X = x) \qquad P(X = x, Y = y) = P(X = x)P(Y = y)$$

- En el caso continuo:

$$f_{Y|X}(y/x) = f_Y(y) \qquad f(x, y) = f_X(x)f_Y(y)$$

Independencia de variables aleatorias.

- La independencia entre dos variables aleatorias X e Y significa que el conocimiento del valor de una de ellas no permite mejorar las predicciones que podamos hacer sobre el valor de la otra.

- En el ejemplo anterior, $f_{X|Y}(x|y) = 2x$, lo que significa que la densidad de X conociendo el valor de Y no depende de Y; por tanto Y no informa sobre X.

- Lo mismo puede decirse sobre la densidad de Y conocido el valor de X, en este caso $f_{Y|X}(y|x) = \frac{2ye^{-y^2}}{1-e^{-\pi^2}}$: no depende de X y por tanto el valor de X no informa sobre la probabilidad de Y.

- Podemos comprobar en este caso que se cumple que $f(x, y) = f_X(x)f_Y(y)$, pues $\frac{4xye^{-y^2}}{1-e^{-\pi^2}} = 2x \cdot \frac{2ye^{-y^2}}{1-e^{-\pi^2}}$

- Por tanto nuestras variables X e Y (amplitud y fase de una señal medida en un cerebro en reposo) **son independientes**.

Esperanza condicional.

Dado un vector aleatorio (X, Y), la **esperanza de Y condicionada por X** es el valor esperado de Y cuando la X toma un valor fijo x. Se calcula mediante:

- Cuando X es discreta:

$$E\left[Y \mid X = x\right] = \sum_y yP\left(Y = y \mid X = x\right)$$

- Cuando X es continua:

$$E\left[Y \mid X = x\right] = \int_{-\infty}^{\infty} yf\left(y \mid x\right) dy$$

La esperanza de X condicionada por Y se define de modo análogo.

Ejemplo 1:

$$f\left(x,y\right)=\frac{4}{1-e^{-\pi^2}}xye^{-y^2}$$

Esperanza de Y condicionada por X:

$$E\left[Y\mid X=x\right]=\int_0^{\pi} y f_{Y|X}(y|x)dy=\int_0^{\pi} y\frac{2ye^{-y^2}}{1-e^{-\pi^2}}dy=\frac{2}{1-e^{-\pi^2}}\int_0^{\pi} y^2e^{-y^2}dy$$

La función $y^2e^{-y^2}$ no puede integrarse analíticamente, pero puede calcularse la esperanza anterior utilizando matlab/octave:

```
function f=f(y)
  f=2*y*exp(-y^2)/(1-exp(-pi^2));
endfunction
mu=quad(@(y) y.*f(y),0,pi)
```

```
## mu =  0.88610
```

Ejemplo 1:

- Esperanza de X condicionada por Y:

$$E\left[X \mid Y=y\right]=\int_0^1 x f_{X|Y}(x|y)dx=\int_0^1 x2xdx=2\left[\frac{x^3}{3}\right]_0^1=\frac{2}{3}$$

- Nótese que en este ejemplo la esperanza de X condicionada por Y no depende del valor de Y, y la esperanza condicionada de Y por el valor de X no depende de X (cosa que ya podíamos sospechar ya que lo mismo ocurría con las densidades condicionales).

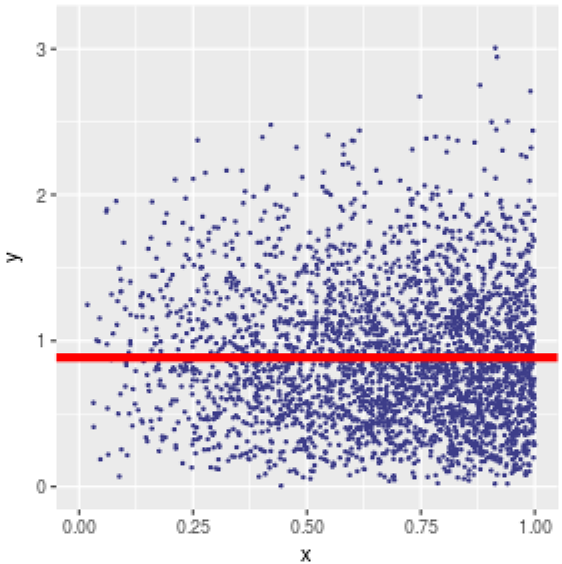
- Esto es otra consecuencia de la independencia entre X e Y.

Ejemplo 1:

Si volvemos a los 5000 valores observados del vector (X, Y) y representamos la función:

$$\mu_Y(x) = E[Y|X = x] = 0.8861$$

esta función es una recta horizontal que corresponde al valor medio de Y para cada x:



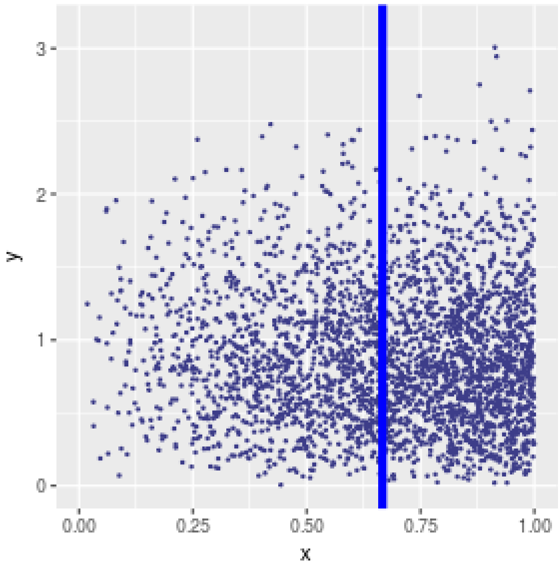
Por tanto, da igual lo que valga la x, el valor esperado de Y es siempre el mismo (0.8661) $\Longrightarrow$ el valor de x no informa sobre Y (son independientes).

Ejemplo 1:

Asimismo, si representamos la función:

$$\mu_X(y) = E[X|Y = y] = \frac{2}{3}$$

obtenemos una recta vertical que corresponde al valor medio de X para cada y:



Vemos que da igual lo que valga la y, el valor esperado de X es siempre el mismo (2/3) $\Longrightarrow$ el valor de y no informa sobre X (son independientes).

Ejemplo 2:

Consideremos ahora el vector aleatorio (X, Y) con función de densidad:

$$f(x,y) = \lambda(x+y^2), \quad 0 \leq x \leq 1;\ 0 \leq y \leq 1$$

Cálculo del valor de λ:

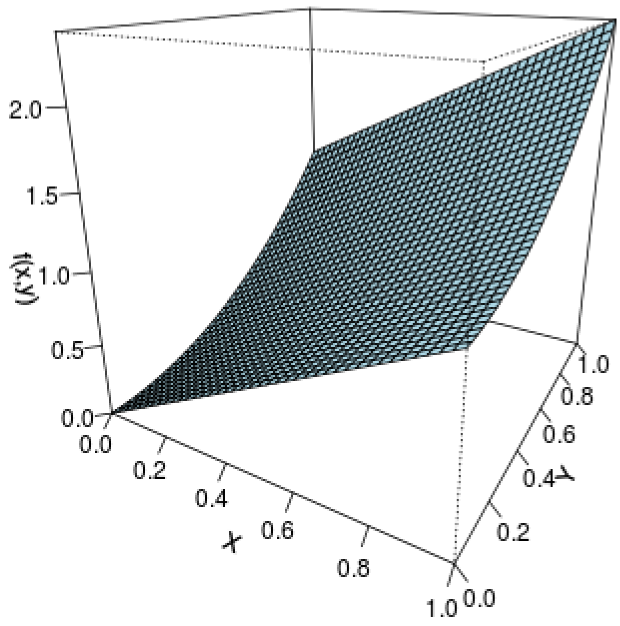
$$1 = \int_0^1 \int_0^1 f(x,y)dydx = \lambda \int_0^1 \int_0^1 (x+y^2)dydx = \lambda \left\{ \int_0^1 \int_0^1 xdydx + \int_0^1 \int_0^1 y^2 dydx \right\} =$$

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

$$= \lambda \left\{ \int_0^1 xdx + \int_0^1 \frac{1}{3} dx \right\} = \lambda \left\{ \left[\frac{x^2}{2} \right]_0^1 + \frac{1}{3}[x]_0^1 \right\} = \lambda \left(\frac{1}{2} + \frac{1}{3} \right) = \lambda \frac{5}{6} \implies \lambda = \frac{6}{5}$$

Ejemplo 2:

Representación gráfica de $f(x,y) = \frac{6}{5}(x + y^2)$:



Pulsar aquí para rotar esta figura de forma interactiva

Ejemplo 2:

$$f(x,y) = \frac{6}{5}(x + y^2)$$

Densidad marginal de *X*:

$$f_X\left(x\right) = \int_0^1 f(x,y)dy = \frac{6}{5}\int_0^1 (x + y^2)dy = \frac{6}{5}\left[xy + \frac{y^3}{3}\right]_0^1 = \frac{6}{5}\left(x + \frac{1}{3}\right)$$

Densidad marginal de *Y*:

$$f_Y\left(y\right) = \int_0^1 f(x,y)dx = \frac{6}{5}\int_0^1 (x + y^2)dx = \frac{6}{5}\left[\frac{x^2}{2} + y^2x\right]_0^1 = \frac{6}{5}\left(\frac{1}{2} + y^2\right)$$

En este caso X e Y **no son independientes** pues $f(x,y) \neq f_X(x) \cdot f_Y(y)$

Ejemplo 2:

$$f(x,y) = \frac{6}{5}(x+y^2)$$

Densidad condicional de *Y* dado el valor de *X*:

$$f_{Y|X}\left(y|x\right) = \frac{f(x,y)}{f_X\left(x\right)} = \frac{\frac{6}{5}(x+y^2)}{\frac{6}{5}\left(x+\frac{1}{3}\right)} = \frac{(x+y^2)}{\left(x+\frac{1}{3}\right)}$$

Densidad condicional de *X* dado el valor de *Y*:

$$f_{X|Y}\left(x|y\right) = \frac{f(x,y)}{f_Y\left(y\right)} = \frac{\frac{6}{5}(x+y^2)}{\frac{6}{5}\left(\frac{1}{2}+y^2\right)} = \frac{(x+y^2)}{\left(\frac{1}{2}+y^2\right)}$$

Ejemplo 2:

$$f(x,y) = \frac{6}{5}(x+y^2)$$

Esperanza de *Y* condicionada por *X*:

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

$$= \frac{\frac{1}{2}\left(x+\frac{1}{2}\right)}{\left(x+\frac{1}{3}\right)} = \frac{3}{4}\left(\frac{2x+1}{3x+1}\right)$$

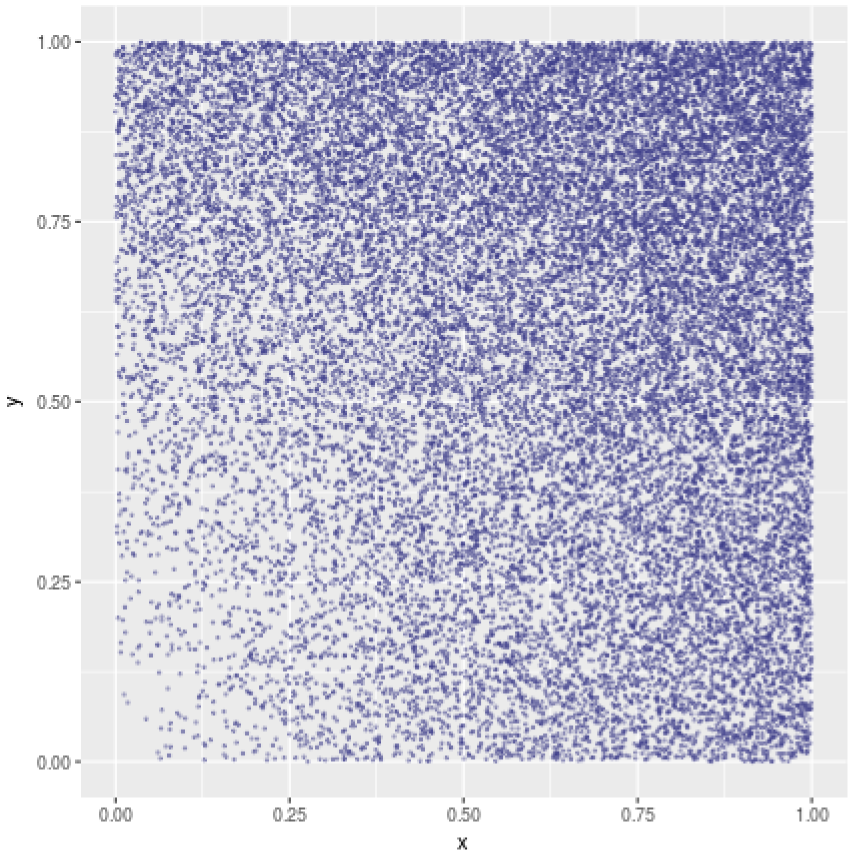
Esperanza de *X* condicionada por *Y*:

$$E\left[X|Y=y\right] = \int_0^1 x f\left(x|y\right)dx = \int_0^1 x\frac{(x+y^2)}{\left(\frac{1}{2}+y^2\right)}dx = \frac{1}{\left(\frac{1}{2}+y^2\right)}\left[\frac{x^3}{3}+y^2\frac{x^2}{2}\right]_0^1 =$$

$$\frac{\frac{1}{3}+\frac{1}{2}y^2}{\frac{1}{2}+y^2} = \frac{2+3y^2}{3+6y^2}$$

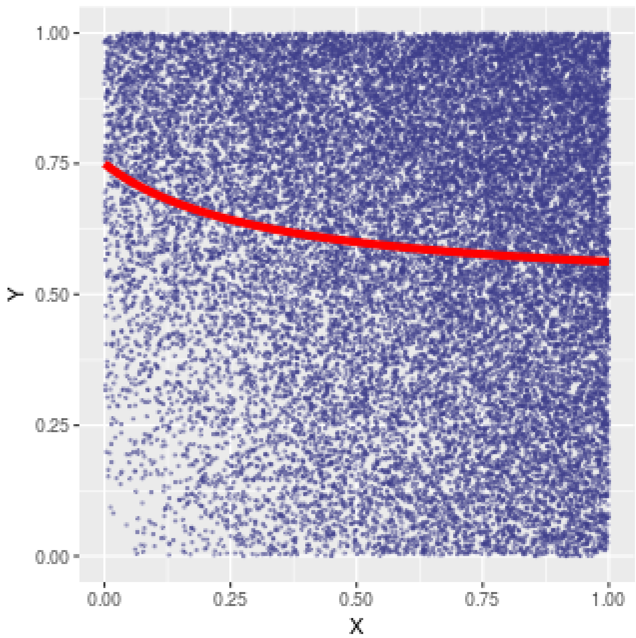
Ejemplo 2:

La gráfica siguiente muestra 25000 observaciones del vector (X, Y):



Ejemplo 2:

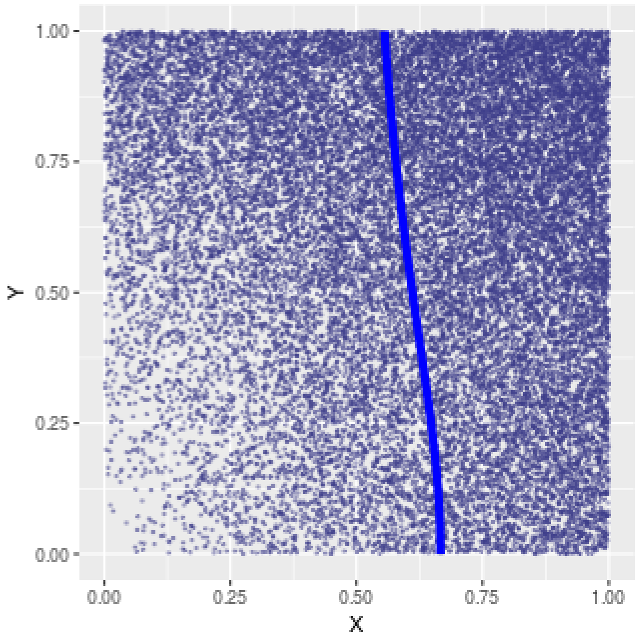
Si superponemos la función $\mu_{Y}(x) = E[y|X = x] = \frac{3}{4}\left(\frac{2x+1}{3x+1}\right)$:



La linea roja es el valor esperado de Y para cada X. Cuando el valor de X es pequeño la esperanza de Y es más alta; a medida que aumenta el valor de X disminuye el valor esperado de $Y \Longrightarrow$ El valor de X informa sobre Y

Ejemplo 2:

Asimismo, si superponemos la función $\mu_X(y) = E[X|Y=y] = \frac{2+3y^2}{3+6y^2}$ obtenemos:



La linea azul es el valor esperado de X para cada Y. A medida que Y aumenta, el valor esperado de X disminuye. $\Longrightarrow$ El valor de Y informa sobre X.