

# Probabilidad

FACULTAD  
DE  
CIENCIAS  
UNIVERSIDAD DE GRANADA



Los Del DGIIM, [losdeldgiim.github.io](https://losdeldgiim.github.io)

Doble Grado en Ingeniería Informática y Matemáticas  
Universidad de Granada



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional (CC BY-NC-ND 4.0).

Eres libre de compartir y redistribuir el contenido de esta obra en cualquier medio o formato, siempre y cuando des el crédito adecuado a los autores originales y no persigas fines comerciales.

# Probabilidad

Los Del DGIIM, [losdeldgiim.github.io](https://losdeldgiim.github.io)

Arturo Olivares Martos

Granada, 2024-2025



# Índice general

<b>1. Esperanza Condicionada</b>	<b>5</b>
1.1. Esperanza condicionada . . . . .	5
1.2. Momentos condicionados . . . . .	11
1.2.1. Momentos condicionados no centrados . . . . .	11
1.2.2. Momentos condicionados centrados . . . . .	11
1.3. Varianza condicionada . . . . .	12
1.4. Regresión Mínimo Cuadrática . . . . .	14
1.4.1. Búsqueda de la función de regresión óptima, $\varphi_{\text{opt}}$ . . . . .	14
1.4.2. Razones de Correlación . . . . .	18
1.5. Rectas de Regresión . . . . .	21
1.5.1. Coeficiente de determinación lineal . . . . .	23
1.5.2. Coeficiente de correlación lineal de Pearson . . . . .	24
<b>2. Relaciones de problemas</b>	<b>27</b>
2.1. Esperanza Condicionada . . . . .	28



# 1. Esperanza Condicionada

Cuando se considera un conjunto de variables aleatorias definidas en relación a un determinado experimento, es usual que existan relaciones entre ellas. El problema de regresión consiste en encontrar una función matemática que permita aproximar el valor de determinadas variables, conocidos los valores del resto.

*Observación.* Este concepto de regresión es el análogo al visto en la asignatura de EDIP. En este caso, en vez de establecer un carácter  $Y$  en función de un carácter  $X$  en función de sus frecuencias, buscamos establecer una relación entre dos variables aleatorias.

En el presente curso, por simplicidad, nos limitaremos a analizar el problema de regresión bidimensional; es decir, dadas dos variables aleatorias  $X$  e  $Y$ , queremos encontrar una función matemática que nos permita aproximar el valor de  $Y$  conocido el valor de  $X$ . Para encontrar esta función matemática usaremos el criterio de optimalidad denominado mínimos cuadrados, visto ya en EDIP.

## 1.1. Esperanza condicionada

Para encontrar esta función matemática, necesitamos introducir el concepto de esperanza condicionada.

**Definición 1.1** (Esperanza condicionada). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, se define la esperanza condicionada de  $X$  a  $Y$ , notada por  $E[X | Y]$ , como la variable aleatoria siguiente (en el caso de que exista):

$$E[X | Y](y) := E[X | Y = y]$$

Notemos que esta esperanza es una esperanza normal, solo que considerando la distribución condicionada de  $X$  a  $Y = y$ .

Veamos en función de si es discreta o continua:

- Si  $X$  e  $Y$  son discretas, entonces:

$$E[X | Y = y] = \sum_{x \in E_x} x \cdot P[X = x | Y = y] \quad \text{con } P[Y = y] > 0$$

- Si  $X$  e  $Y$  son continuas, entonces:

$$E[X | Y = y] = \int_{-\infty}^{\infty} x \cdot f_{X|Y=y}(x) dx \quad \text{con } f_Y(y) > 0$$

Notemos que la definición de  $E[X | Y = y]$  requiere que la distribución condicionada de  $X$  a  $Y = y$  esté bien definida. Por este motivo, es preciso exigir que  $P[Y = y] \neq 0$  en el caso discreto, y  $f_Y(y) \neq 0$  en el caso continuo. Así, la variable aleatoria  $E[X | Y]$  es una función de la variable  $Y$ , definida sobre el conjunto de sus valores,  $E_Y$ .

Veamos ahora qué hemos de imponer para que exista la esperanza condicionada de  $X$  a  $Y$ .

**Proposición 1.1.** *Sean dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad. Entonces:*

$$\exists E[X] \implies \exists E[X | Y]$$

*Demostración.* Distinguimos en función de si  $X$  e  $Y$  son discretas o continuas:

- Si  $X$  e  $Y$  son discretas, entonces:

$$\begin{aligned} E[X | Y = y] &= \sum_{x \in E_x} |x| \cdot P[X = x | Y = y] = \sum_{x \in E_x} |x| \cdot \frac{P[X = x, Y = y]}{P[Y = y]} = \\ &= \frac{\sum_{x \in E_x} |x| \cdot P[X = x, Y = y]}{P[Y = y]} \leq \frac{\sum_{x \in E_x} |x| \cdot P[X = x]}{P[Y = y]} = \frac{E[|X|]}{P[Y = y]} < \infty \end{aligned}$$

donde en la desigualdad hemos hecho uso de que

$$P[X = x, Y = y] \leq P[X = x] \forall x, y$$

y, al final, hemos usado que  $E[|X|] < \infty$  por tener  $E[X]$  finita. Por tanto,  $E[X | Y]$  existe.

- Si  $X$  e  $Y$  son continuas, entonces:

$$\begin{aligned} E[X | Y = y] &= \int_{-\infty}^{\infty} |x| \cdot f_{X|Y=y}(x) dx = \int_{-\infty}^{\infty} |x| \cdot \frac{f_{X,Y}(x, y)}{f_Y(y)} dx = \\ &= \frac{\int_{-\infty}^{\infty} |x| \cdot f_{X,Y}(x, y) dx}{f_Y(y)} \end{aligned}$$

Como  $E[|X|] < \infty$ , entonces:

$$\begin{aligned} E[|X|] &= \int_{-\infty}^{\infty} |x| \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x| \cdot f_{X,Y}(x, y) dx dy < \infty \implies \\ &\implies \int_{-\infty}^{\infty} |x| \cdot f_{X,Y}(x, y) dx < \infty \end{aligned}$$

Por tanto,  $E[X | Y]$  existe. □

**Ejemplo.** Sea  $(X, Y)$  un vector aleatorio bidimensional con función de densidad conjunta:

$$f_{X,Y}(x, y) = \begin{cases} 2 & \text{si } 0 < x < y < 1 \\ 0 & \text{en otro caso} \end{cases}$$



Calcular  $E[X | Y]$  y  $E[Y | X]$ .

Tenemos que:

$$E[X | Y = y] = \int_{-\infty}^{\infty} x \cdot f_{X|Y=y}(x) dx$$

$$E[Y | X = x] = \int_{-\infty}^{\infty} y \cdot f_{Y|X=x}(y) dy$$

Para calcular ambas funciones de densidad condicionada, necesitamos calcular las marginales de  $X$  e  $Y$ :

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy = \int_x^1 2 dy = 2(1 - x)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx = \int_0^y 2 dx = 2y$$

Así, las funciones de densidad condicionada son:

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{2}{2y} = \frac{1}{y} \quad \text{si } 0 < x < y < 1$$

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{2}{2(1-x)} = \frac{1}{1-x} \quad \text{si } 0 < x < y < 1$$

Por tanto, las esperanzas condicionadas son:

$$E[X | Y = y] = \int_{-\infty}^{\infty} x \cdot f_{X|Y=y}(x) dx = \int_0^y x \cdot \frac{1}{y} dx = \frac{1}{y} \int_0^y x dx = \frac{1}{y} \left[ \frac{x^2}{2} \right]_0^y = \frac{y}{2}$$

$$E[Y | X = x] = \int_{-\infty}^{\infty} y \cdot f_{Y|X=x}(y) dy = \int_x^1 y \cdot \frac{1}{1-x} dy = \frac{1}{1-x} \int_x^1 y dy = \frac{1}{1-x} \left[ \frac{y^2}{2} \right]_x^1 =$$

$$= \frac{1}{1-x} \left( \frac{1}{2} - \frac{x^2}{2} \right) = \frac{1}{2} \cdot \frac{1-x^2}{1-x} = \frac{1+x}{2}$$

Al igual que considerábamos  $E[g(X)]$ , consideramos ahora la esperanza condicionada de una función de  $X$  a  $Y$ .

**Definición 1.2** (Esperanza condicionada de una función). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, y una función  $g : \mathbb{R} \rightarrow \mathbb{R}$  medible, se define la esperanza condicionada de  $g(X)$  a  $Y$ , notada por  $E[g(X) | Y]$ , como la variable aleatoria siguiente (en el caso de que exista):

$$E[g(X) | Y](y) := E[g(X) | Y = y]$$

Veamos en función de si es discreta o continua:

- Si  $X$  e  $Y$  son discretas, entonces:

$$E[g(X) | Y = y] = \sum_{x \in E_x} g(x) \cdot P[X = x | Y = y] \quad \text{con } P[Y = y] > 0$$

- Si  $X$  e  $Y$  son continuas, entonces:

$$E[g(X) | Y = y] = \int_{-\infty}^{\infty} g(x) \cdot f_{X|Y=y}(x) dx \quad \text{con } f_Y(y) > 0$$

Al igual que ocurría con la esperanza condicionada de  $X$  a  $Y$ , para que exista basta con imponer que  $E[g(X)] < \infty$ . Es decir, la demostración de la siguiente proposición es análoga a la del proposición anterior.

**Proposición 1.2.** Sean dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, y una función  $g : \mathbb{R} \rightarrow \mathbb{R}$  medible. Entonces:

$$\exists E[g(X)] \implies \exists E[g(X) | Y]$$

**Ejemplo.** En el experimento aleatorio del lanzamiento de tres monedas se consideran las variables:

- $X$ : Número de caras.
- $Y$ : Diferencia, en valor absoluto, entre el número de caras y el número de cruces.

Calcular  $E[X^2 | Y]$ .

El espacio muestral es el siguiente:

$$\Omega = \{CCC, CC+, C+C, +CC, C++, +C+, ++C, +++\}$$

La función masa de probabilidad conjunta viene dada por:

$X \setminus Y$	1	3	$P[X = x_i]$
0	0	$1/2^3$	$1/2^3$
1	$3/2^3$	0	$3/2^3$
2	$3/2^3$	0	$3/2^3$
3	0	$1/2^3$	$1/2^3$
$P[Y = y_j]$	$3/2^2$	$1/2^2$	1

donde además hemos calculado las marginales de  $X$  e  $Y$ . Entonces, tenemos que:

$$\begin{aligned}
 E[X^2 | Y = 1] &= \sum_{x \in E_x} x^2 \cdot P[X = x | Y = 1] = \\
 &= 1^2 \cdot P[X = 1 | Y = 1] + 2^2 \cdot P[X = 2 | Y = 1] = \\
 &= 1 \cdot \frac{P[X = 1, Y = 1]}{P[Y = 1]} + 4 \cdot \frac{P[X = 2, Y = 1]}{P[Y = 1]} = \\
 &= 1 \cdot \frac{3/2^3}{3/2^2} + 4 \cdot \frac{3/2^3}{3/2^2} = \frac{5}{2} \\
 E[X^2 | Y = 3] &= \sum_{x \in E_x} x^2 \cdot P[X = x | Y = 3] = \\
 &= 0^2 \cdot P[X = 0 | Y = 3] + 3^2 \cdot P[X = 3 | Y = 3] = \\
 &= 0 \cdot \frac{P[X = 0, Y = 3]}{P[Y = 3]} + 9 \cdot \frac{P[X = 3, Y = 3]}{P[Y = 3]} = \\
 &= 9 \cdot \frac{1/2^3}{1/2^2} = \frac{9}{2}
 \end{aligned}$$

Por tanto, la esperanza condicionada de  $X^2$  a  $Y$  es:

$$E[X^2 | Y] = \begin{cases} 5/2 & \text{si } Y = 1 \\ 9/2 & \text{si } Y = 3 \end{cases}$$

### Propiedades de la esperanza condicionada

Introducimos las siguientes propiedades, que se deducen de forma directa en la mayoría de los casos haciendo uso de que la esperanza condicionada de  $X$  a  $Y$  es la esperanza de  $X$  con respecto a la distribución condicionada de  $X$  a  $Y$ .

**Proposición 1.3.** *Sea  $Y$  una variable aleatoria discreta sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$  y  $c \in \mathbb{R}$ . Entonces:*

$$E[c | Y] = c$$

*Demostración.* Como  $Y$  es discreta, entonces:

$$E[c | Y = y] = \sum_{x \in E_x} x \cdot P[X = x | Y = y] = c \cdot P[X = c | Y = y] = c$$

□

Notemos que no consideramos  $Y$  continua, ya que  $c$  es una variable aleatoria constante (discreta) y solo hemos definido la esperanza condicionada para variables aleatorias del mismo tipo.

**Proposición 1.4** (Linealidad). *Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Entonces:*

$$E[(aX + b) | Y] = aE[X | Y] + b \quad \forall a, b \in \mathbb{R}$$

**Proposición 1.5.** *Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Entonces:*

$$E[aX_1 + bX_2 | Y] = aE[X_1 | Y] + bE[X_2 | Y] \quad \forall a, b \in \mathbb{R}$$

**Proposición 1.6.** *Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Si  $X \geq 0$ , entonces:*

$$E[X | Y] \geq 0 \wedge E[X | Y] = 0 \iff P[X = 0] = 1$$

**Proposición 1.7** (Conservación del Orden). *Sean  $X_1, X_2$  e  $Y$  tres variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Si  $X_1 \leq X_2$ , entonces:*

$$E[X_1 | Y] \leq E[X_2 | Y]$$

**Proposición 1.8.** *Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ , y consideramos una función  $g : \mathbb{R} \rightarrow \mathbb{R}$  medible. Entonces:*

$$\text{Si } X, Y \text{ independientes} \implies E[g(X) | Y] = E[g(X)]$$

En particular, tomando  $g = Id$ , tenemos:

$$\text{Si } X, Y \text{ independientes} \implies E[X | Y] = E[X]$$

*Demostración.* Distinguimos en función de si  $X$  e  $Y$  son discretas o continuas:

- Si  $X$  e  $Y$  son discretas, como  $X$  e  $Y$  son independientes, entonces:

$$P[X = x \mid Y = y] = P[X = x] \quad \forall x, y$$

Por tanto, tenemos:

$$E[g(X) \mid Y = y] = \sum_{x \in E_X} g(x) \cdot P[X = x \mid Y = y] = \sum_{x \in E_X} g(x) \cdot P[X = x] = E[g(X)]$$

- Si  $X$  e  $Y$  son continuas, como  $X$  e  $Y$  son independientes, entonces:

$$f_{X|Y=y}(x) = f_X(x) \quad \forall x, y$$

Por tanto, tenemos:

$$E[g(X) \mid Y = y] = \int_{-\infty}^{\infty} g(x) \cdot f_{X|Y=y}(x) dx = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx = E[g(X)]$$

□

**Proposición 1.9.** Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ , y consideramos una función  $g : \mathbb{R} \rightarrow \mathbb{R}$  medible. Entonces:

$$E[E[g(X) \mid Y]] = E[g(X)]$$

En particular, tomando  $g = Id$ , tenemos:

$$E[E[X \mid Y]] = E[X]$$

*Demostración.* Distinguimos en función de si  $X$  e  $Y$  son discretas o continuas:

- Si  $X$  e  $Y$  son discretas, entonces, como  $E[g(X) \mid Y]$  es una variable aleatoria en función de  $Y$ , usando la esperanza de una función de una variable aleatoria, tenemos:

$$\begin{aligned} E[E[g(X) \mid Y]] &= \sum_{y \in E_Y} E[g(X) \mid Y = y] \cdot P[Y = y] = \\ &= \sum_{y \in E_Y} \sum_{x \in E_X} g(x) \cdot P[X = x \mid Y = y] \cdot P[Y = y] = \\ &= \sum_{x \in E_X} g(x) \cdot \sum_{y \in E_Y} P[X = x, Y = y] = \\ &= \sum_{x \in E_X} g(x) \cdot P[X = x] = E[g(X)] \end{aligned}$$

- Si  $X$  e  $Y$  son continuas, entonces, como  $E[g(X) \mid Y]$  es una variable aleatoria en función de  $Y$ , usando la esperanza de una función de una variable aleatoria,

tenemos:

$$\begin{aligned}
 E[E[g(X) | Y]] &= \int_{-\infty}^{\infty} E[g(X) | Y = y] \cdot f_Y(y) dy = \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x) \cdot f_{X|Y=y}(x) dx \cdot f_Y(y) dy = \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x) \cdot f_{X,Y}(x, y) dx dy = \\
 &= \int_{-\infty}^{\infty} g(x) \cdot \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy dx = \\
 &= \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx = E[g(X)]
 \end{aligned}$$

□

## 1.2. Momentos condicionados

Como una variable aleatoria que es, se pueden definir los momentos condicionados de una variable aleatoria a otra. Comenzamos con los momentos condicionados no centrados.

### 1.2.1. Momentos condicionados no centrados

**Definición 1.3** (Momento condicionado). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad y  $k \in \mathbb{N}$ , se define el momento condicionado de orden  $k$  de  $X$  a  $Y$  como la variable aleatoria siguiente (en el caso de que exista):

$$E[X^k | Y]$$

Usando la esperanza condicionada de una función, tenemos que  $\exists E[X^k]$  implica que exista  $E[X^k | Y]$ , y se calcula como:

- Si  $X$  e  $Y$  son discretas, entonces:

$$E[X^k | Y = y] = \sum_{x \in E_x} x^k \cdot P[X = x | Y = y] \quad \text{con } P[Y = y] > 0$$

- Si  $X$  e  $Y$  son continuas, entonces:

$$E[X^k | Y = y] = \int_{-\infty}^{\infty} x^k \cdot f_{X|Y=y}(x) dx \quad \text{con } f_Y(y) > 0$$

### 1.2.2. Momentos condicionados centrados

Respecto de los momentos condicionados centrados, tenemos la siguiente definición.

**Definición 1.4** (Momento condicionado centrado). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad y  $k \in \mathbb{N}$ , se define el momento condicionado de orden  $k$  centrado de  $X$  a  $Y$  como la variable aleatoria siguiente (en el caso de que exista):

$$E[(X - E[X | Y])^k | Y]$$

En caso de existencia, los valores de estas variables se calculan teniendo en cuenta que al considerar el condicionamiento a un valor arbitrario,  $Y = y$ , la variable  $E[X | Y]$  toma el valor  $E[X | Y = y]$  y, por tanto:

$$E[(X - E[X | Y])^k | Y = y] = E[(X - E[X | Y = y])^k | Y = y] \quad \forall y \in E_Y$$

Entonces, ya que dado  $Y = y$ ,  $E[X | Y = y]$  es una constante, la variable  $(X - E[X | Y = y])^k$  sólo depende de  $X$ , y aplicando de nuevo la expresión de la esperanza condicionada de una función de una variable aleatoria, se tiene:

- Si  $X$  e  $Y$  son discretas, entonces:

$$E[(X - E[X | Y])^k | Y = y] = \sum_{x \in E_X} (x - E[X | Y = y])^k \cdot P[X = x | Y = y] \quad \text{con } P[Y = y] > 0$$

- Si  $X$  e  $Y$  son continuas, entonces:

$$E[(X - E[X | Y])^k | Y = y] = \int_{-\infty}^{\infty} (x - E[X | Y = y])^k \cdot f_{X|Y=y}(x) dx \quad \text{con } f_Y(y) > 0$$

Las propiedades de los momentos condicionados son similares a las de los momentos sin condicionar. En particular, la existencia de los momentos condicionados no centrados equivale a la de los momentos condicionados centrados.

### 1.3. Varianza condicionada

La varianza condicionada es un caso particular de los momentos condicionados centrados, y es de especial relevancia en el estudio de la regresión, por lo que estudiaremos su definición y propiedades.

**Definición 1.5.** Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, se define la varianza condicionada de  $X$  a  $Y$  como el momento condicionado centrado de orden 2 de  $X$  a  $Y$ :

$$\text{Var}[X | Y] := E[(X - E[X | Y])^2 | Y]$$

Veamos ahora algunas de sus propiedades.

**Proposición 1.10.** Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Si  $\exists E[X^2]$ , entonces:

1.  $\exists \text{Var}[X | Y]$  y  $\text{Var}[X | Y] \geq 0$ .
2.  $\text{Var}[X | Y] = E[X^2 | Y] - (E[X | Y])^2$ .

*Demostración.* Demostramos cada uno de los apartados:

1. Dado que  $\exists E[X^2]$ , entonces  $\exists E[X^2 | Y]$ . Como la existencia de momentos no centrados implica la de momentos centrados, entonces  $\exists E[(X - E[X | Y])^2 | Y] = \text{Var}[X | Y]$ .

Por otro lado, como  $E[(X - E[X | Y])^2 | Y]$ , es la esperanza condicionada de una variable aleatoria no negativa, entonces  $\text{Var}[X | Y] \geq 0$ .

2. Partiendo de la definición de varianza condicionada, fijado  $Y = y$ , tenemos:

$$\begin{aligned} \text{Var}[X | Y = y] &= E[(X - E[X | Y = y])^2 | Y = y] = \\ &= E[X^2 - 2XE[X | Y = y] + (E[X | Y = y])^2 | Y = y] \stackrel{(*)}{=} \\ &\stackrel{(*)}{=} E[X^2 | Y = y] - 2E[E[X | Y = y] \cdot X | Y = y] + E[(E[X | Y = y])^2 | Y = y] \stackrel{(**)}{=} \\ &\stackrel{(**)}{=} E[X^2 | Y = y] - 2E[X | Y = y] \cdot E[X | Y = y] + (E[X | Y = y])^2 = \\ &= E[X^2 | Y = y] - 2(E[X | Y = y])^2 + (E[X | Y = y])^2 = \\ &= E[X^2 | Y = y] - (E[X | Y = y])^2 \end{aligned}$$

donde en  $(*)$  hemos usado la linealidad de la esperanza condicionada, y en  $(**)$  hemos usado que, como ya hemos condicionado a  $Y = y$ ,  $E[X | Y = y]$  es una constante.

□

Introducimos además esta última proposición, que cobrará gran importancia en la siguiente sección.

**Proposición 1.11** (Descomposición de la varianza). *Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Si  $\exists E[X^2]$ , entonces se tiene que  $\exists \text{Var}[E[X | Y]]$  y  $\exists E[\text{Var}[X | Y]]$ , y se cumple que:*

$$\text{Var}[X] = \text{Var}[E[X | Y]] + E[\text{Var}[X | Y]]$$

*Demostración.* En primer lugar, tenemos que  $\exists E[X^2]$  implica  $\exists E[E[X^2 | Y]]$ . Por tanto, del segundo apartado de la proposición anterior, tenemos que:

$$0 \leq \text{Var}[X | Y] = E[X^2 | Y] - (E[X | Y])^2 \implies 0 \leq (E[X | Y])^2 \leq E[X^2 | Y]$$

Por tanto, como  $(E[X | Y])^2$  es una variable aleatoria acotada por variables aleatorias con esperanza, entonces  $\exists E[(E[X | Y])^2]$ . Tenemos por tanto, que:

- $\exists E[(E[X | Y])^2] \implies \exists \text{Var}[E[X | Y]] = E[(E[X | Y])^2] - (E[E[X | Y]])^2$ .
- $\exists E[(E[X | Y])^2]$  y  $\exists E[E[X^2 | Y]]$ , lo que implica que:  

$$\exists E[E[X^2 | Y]] - E[(E[X | Y])^2] = E[E[X^2 | Y] - (E[X | Y])^2] = E[\text{Var}[X | Y]]$$

Por tanto, hemos demostrado las dos existencias. Para demostrar la igualdad, tenemos que:

$$\begin{aligned} \text{Var}[E[X | Y]] + E[\text{Var}[X | Y]] &= \\ &= E[(E[X | Y])^2] - (E[E[X | Y]])^2 + E[E[X^2 | Y] - E[(E[X | Y])^2]] = \\ &= - (E[X])^2 + E[X^2] = \\ &= \text{Var}[X] \end{aligned}$$

Como queríamos demostrar.

□

## 1.4. Regresión Mínimo Cuadrática

Explicamos de nuevo el problema de regresión, que ya se introdujo para Estadística Descriptiva. Dadas dos variables aleatorias,  $X$  e  $Y$ , definidas sobre el mismo espacio de probabilidad, el problema de regresión de  $Y$  sobre  $X$  consiste en determinar una función de  $X$  que proporcione una representación, lo más precisa posible, de la variable  $Y$ . Esto es, se trata de aproximar la variable  $Y$  por una función de  $X$ , de manera que la aproximación sea óptima en algún sentido preestablecido.

$$Y \approx \varphi(X)$$

donde:

- $Y$  es la variable dependiente, explicada o endógena.
- $X$  es la variable independiente, explicativa o exógena.
- $\varphi$  es la función de regresión.

El criterio de optimalidad más usual para abordar el problema de regresión es el basado en el principio de mínimos cuadrados, y consiste en encontrar la función que minimiza la media de las desviaciones cuadráticas de las aproximaciones respecto de los verdaderos valores de la variable aproximada. Esto es, se trata de encontrar una función,  $\varphi$ , que minimice  $E[(Y - \varphi(X))^2]$ .

**Definición 1.6** (Error cuadrático medio). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, el error cuadrático medio asociado a la función de regresión  $\varphi$  de forma que  $Y \approx \varphi(X)$  es:

$$\text{E.C.M.}(\varphi) = E[(Y - \varphi(X))^2]$$

Por tanto, el problema de regresión abordado bajo esta perspectiva consiste en encontrar una función,  $\varphi_{\text{opt}}$ , que minimice el E.C.M.

### 1.4.1. Búsqueda de la función de regresión óptima, $\varphi_{\text{opt}}$

En la presente subsección, suponemos que estamos buscando aproximar  $Y$  por una función de  $X$ ,  $\varphi(X)$ , y que hemos decidido que la función de regresión óptima es aquella que minimiza el E.C.M. asociado a la aproximación.

*Observación.* Se deja para el lector generalizar el caso para aproximar  $X$  por una función de  $Y$ .

Así, el problema de regresión se reduce a encontrar la función de regresión óptima,  $\varphi_{\text{opt}}$ , que minimiza el E.C.M. El siguiente teorema no se demostrará por ser un resultado que excede los conocimientos de la asignatura, pero nos proporciona la solución al problema de regresión bajo el criterio de mínimos cuadrados.

**Teorema 1.12.** Sea  $Y$  una variable aleatoria sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ , y  $X$  una variable aleatoria sobre el mismo espacio de probabilidad. Entonces,  $E[Y | X]$  minimiza el E.C.M. Es decir:

$$\varphi_{\text{opt}}(X) = E[Y | X]$$

Además, esta función de regresión óptima es única.



El E.C.M. asociado a la función de regresión óptima, que es el mínimo error cuadrático medio cometido al aproximar  $Y$  a partir de  $X$ , es:

$$\begin{aligned} \text{E.C.M.}(\varphi_{\text{opt}}) &= \text{E.C.M.}(E[Y | X]) = E[(Y - E[Y | X])^2] = \\ &= E[E[(Y - E[Y | X])^2 | X]] = E[\text{Var}[Y | X]] \end{aligned}$$

Considerando la descomposición de la varianza, tenemos:

$$\text{Var}[Y] = \text{Var}[E[Y | X]] + E[\text{Var}[Y | X]] = \text{Var}[\varphi_{\text{opt}}(X)] + \text{E.C.M.}(\varphi_{\text{opt}})$$

**Definición 1.7** (Curva de regresión mínimo cuadrática). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, la curva de regresión mínimo cuadrática de  $Y$  sobre  $X$  es la curva obtenida empleando  $\varphi_{\text{opt}}(X)$ . Es decir:

- Curva de Regresión Mínimo Cuadrática de  $Y$  sobre  $X$ :

$$\hat{Y}(x) = E[Y | X = x] \quad \forall x \in E_X$$

- Curva de Regresión Mínimo Cuadrática de  $X$  sobre  $Y$ :

$$\hat{X}(y) = E[X | Y = y] \quad \forall y \in E_Y$$

Veamos algunos casos particulares, para los que antes debemos introducir las siguientes definiciones.

**Definición 1.8** (Dependencia funcional). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, se dice que  $Y$  depende funcionalmente de  $X$  si existe una función  $f : E_X \rightarrow E_Y$  tal que:

$$Y = f(X) \quad \forall x \in E_X$$

Dicha función  $f$  se denomina curva de dependencia.

**Proposición 1.13.** Sean  $X, Y$  dos variables aleatorias sobre el mismo espacio de probabilidad.

$$\text{Var}[Y | X] = 0 \iff Y \text{ depende funcionalmente de } X$$

*Demostración.* Demostramos mediante la doble implicación:

$\implies$ ) Partimos de:

$$\text{Var}[Y | X = x] = 0 \quad \forall x \in E_X$$

Por tanto, fijado  $x \in E_X$ ,  $\exists!$   $c_x \in E_Y$  tal que:

$$[Y | X = x] = c_x$$

Por tanto, definimos la función  $f : E_X \rightarrow E_Y$  como:

$$f(x) = c_x \quad \forall x \in E_X$$

Por tanto,  $Y$  depende funcionalmente de  $X$ .

$\Leftarrow$ ) Partimos de que  $Y$  depende funcionalmente de  $X$ , es decir,  $\exists f : E_X \rightarrow E_Y$  tal que:

$$Y = f(X)$$

Entonces, tenemos que:

$$\begin{aligned} \text{Var}[Y \mid X = x] &= \text{Var}[f(X) \mid X = x] = \\ &= \text{Var}[f(x) \mid X = x] = 0 \quad \forall x \in E_X \end{aligned}$$

donde hemos usado que la varianza de una constante es 0.

□

**Definición 1.9** (Dependencia recíproca). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, se dice que hay dependencia recíproca entre  $X$  e  $Y$  si  $Y$  depende funcionalmente de  $X$  y  $X$  depende funcionalmente de  $Y$  con la misma función. Es decir, existe una función  $f : E_X \rightarrow E_Y$  tal que:

$$\begin{aligned} Y &= f(X) & \forall x \in E_X \\ X &= f^{-1}(Y) & \forall y \in E_Y \end{aligned}$$

Dicha función  $f$  se denomina curva de dependencia.

**Corolario 1.13.1.** Sean  $X, Y$  dos variables aleatorias sobre el mismo espacio de probabilidad.

$$\text{Var}[Y \mid X] = \text{Var}[X \mid Y] = 0 \iff Y \text{ y } X \text{ dependen recíprocamente}$$

*Demostración.* Demostramos mediante doble implicación:

$\Rightarrow$ ) Como corolario de la proposición anterior, sabemos que  $\exists f : E_X \rightarrow E_Y$ ,  $g : E_Y \rightarrow E_X$  tales que:

$$\begin{aligned} Y &= f(X) & \forall x \in E_X \\ X &= g(Y) & \forall y \in E_Y \end{aligned}$$

Sea  $c_x \in E_y$ , y consideramos su preimagen

$$f^{-1}(c_x) = \{x \in E_X \mid f(x) = c_x\} = \{x \in E_X \mid Y = c_x\}$$

No obstante, usando la demostración de la proposición anterior tenemos que dicho  $c_x$  es único. Por tanto,  $f^{-1} = g$  y, por tanto,  $Y$  y  $X$  dependen recíprocamente.

$\Leftarrow$ ) Se tiene de forma directa como corolario de la proposición anterior.

□

Algunos casos particulares son:

- Si  $Y$  depende funcionalmente de  $X$  con curva de dependencia  $Y = f(X)$ , entonces la curva de regresión mínimo cuadrática de  $Y$  sobre  $X$  es la propia curva de dependencia:

$$y = f(x) = E[Y | X = x] \quad \forall x \in E_X$$

*Demostración.* Dado que  $Y = f(X)$ , entonces:

$$E[Y | X = x] = E[f(X) | X = x] = E[f(x) | X = x] = f(x) \quad \forall x \in E_X$$

□

- Si hay dependencia recíproca entre  $X$  e  $Y$ , es decir,  $Y = f(X)$  y  $X = f^{-1}(Y)$ , entonces ambas curvas de regresión mínimo cuadrática coinciden con las curvas de dependencia:

$$\begin{aligned} y = f(x) &= E[Y | X = x] & \forall x \in E_X \\ x = f^{-1}(y) &= E[X | Y = y] & \forall y \in E_Y \end{aligned}$$

- Si  $X$  e  $Y$  son independientes, entonces la curva de regresión mínimo cuadrática de  $Y$  sobre  $X$  es la esperanza de  $Y$ , y la de  $X$  sobre  $Y$  es la esperanza de  $X$ :

$$\begin{aligned} \hat{Y}(x) &= E[Y] & \forall x \in E_X \\ \hat{X}(y) &= E[X] & \forall y \in E_Y \end{aligned}$$

Como vemos, estas curvas de regresión mínimo cuadrática son constantes y paralelas a los ejes, lo que muestra que no tiene sentido plantear un problema de regresión en este caso.

*Demostración.* Dado que  $X$  e  $Y$  son independientes, sabemos de forma directa que:

$$\begin{aligned} E[Y | X = x] &= E[Y] & \forall x \in E_X \\ E[X | Y = y] &= E[X] & \forall y \in E_Y \end{aligned}$$

□

Por tanto, a modo de resumen, en la Tabla 1.1 se recogen las predicciones de  $Y$  según lo que estemos observando.

	Sin observar $X$	Observando $X$	Para $X = x$
$\hat{Y}$	$E[Y]$	$E[Y   X]$	$E[Y   X = x]$
E.C.M.	$\text{Var}[Y]$	$\text{Var}[Y   X]$	$\text{Var}[Y   X = x]$

Tabla 1.1: Resumen de las predicciones de  $Y$  sobre en función de lo que se observe.

### 1.4.2. Razones de Correlación

Buscamos ahora estudiar el grado de bondad de la aproximación mínimo cuadrática de cada variable a partir de la otra. Partimos de la siguiente igualdad:

$$Y = E[Y | X] + (Y - E[Y | X])$$

Sabemos que  $\hat{Y} = E[Y | X]$  es la mejor aproximación que hemos obtenido. Vemos por tanto que el error es la variable aleatoria  $(Y - E[Y | X])$  unidades, que denominaremos *residuo*.

**Definición 1.10** (Residuo). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, el residuo asociado a la aproximación de  $Y$  a partir de  $X$  es la siguiente variable aleatoria:

$$R = Y - E[Y | X]$$

La aproximación será mejor cuanto menor sea  $|R|$ . Esta es una variable aleatoria, cuya comparación para estudiar la bondad no es sencilla. Usaremos por tanto valores numéricos que resuman la bondad de la aproximación.

Como primer intento, buscamos comparar mediante  $E[R]$ . Tenemos que:

$$E[R] = E[Y - E[Y | X]] = E[Y] - E[E[Y | X]] = E[Y] - E[Y] = 0$$

Por tanto, siempre será nulo, lo que no nos aporta información sobre la bondad de la aproximación. Buscamos por tanto comparar mediante la varianza de  $R$ , conocida como *varianza residual*.

**Definición 1.11** (Varianza residual). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, la varianza residual asociada a la aproximación de  $Y$  a partir de  $X$  es varianza del residuo, es decir:

$$\text{Var}[R] = \text{Var}[Y - E[Y | X]]$$

Calculemos el valor de la varianza residual.

$$\begin{aligned} \text{Var}[R] &= \text{Var}[Y - E[Y | X]] = E[(Y - E[Y | X])^2] - (E[Y - E[Y | X]])^2 = \\ &= E[(Y - E[Y | X])^2] - \cancel{E[R]^2} = E[(Y - E[Y | X])^2] = \\ &= E[E[(Y - E[Y | X])^2 | X]] = E[\text{Var}[Y | X]] = \text{E.C.M.}(\varphi_{\text{opt}}) \end{aligned}$$

Por tanto, usando la descomposición de la varianza, tenemos que:

$$\text{Var}[Y] = \text{Var}[E[Y | X]] + E[\text{Var}[Y | X]] = \text{Var}[\hat{Y}] + \text{Var}[R] = \text{Var}[\hat{Y}] + \text{E.C.M.}(\varphi_{\text{opt}}) \quad (1.1)$$

Vemos que el  $\text{E.C.M.}(\varphi_{\text{opt}})$  será menor conforme mayor sea  $\text{Var}[\hat{Y}]$ , por lo que la aproximación será mejor conforme mayor sea  $\text{Var}[\hat{Y}] = \text{Var}[E[Y | X]]$ . No obstante, usar este valor para comparar la bondad de la aproximación introduce los siguientes problemas:

- $\text{Var}[E[Y | X]]$  no es adimensional, por lo que no es un valor que podamos comparar.
- No es invariante frente a cambios de escala, lo cual puede llevar a conclusiones engañosas. Por ejemplo, sean:
  - $Y$  una variable aleatoria que mide la altura de una persona en metros.
  - $Y'$  una variable aleatoria que mide la altura de una persona en centímetros.

Entonces, es razonable pensar que cualquier variable  $X$  debe aproximar igual de bien a  $Y$  que a  $Y' = 100Y$ . No obstante, tenemos que:

$$\text{Var}[E[Y' | X]] = \text{Var}[E[100Y | X]] = \text{Var}[100E[Y | X]] = 100^2 \text{Var}[E[Y | X]]$$

De esta forma, midiendo la bondad de la aproximación por su varianza sin tener en cuenta la unidad de medida de las variables aproximadas, podríamos concluir que la variable  $X$  aproxima mucho mejor a  $Y'$  que a  $Y$ .

Estos inconvenientes se salvan normalizando la varianza de la función de regresión y usando el siguiente coeficiente, que es el que buscábamos presentar en esta sección.

**Definición 1.12** (Razón de correlación). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, se define la razón de correlación de  $Y$  sobre  $X$ :

$$\eta_{Y/X}^2 = \frac{\text{Var}[E[Y | X]]}{\text{Var}[Y]}$$

De forma análoga se define la razón de correlación de  $X$  sobre  $Y$ :

$$\eta_{X/Y}^2 = \frac{\text{Var}[E[X | Y]]}{\text{Var}[X]}$$

Usando la Ecuación 1.1, tenemos:

$$\eta_{Y/X}^2 = \frac{\text{Var}[E[Y | X]]}{\text{Var}[Y]} = \frac{\text{Var}[Y] - E[\text{Var}[Y | X]]}{\text{Var}[Y]} = 1 - \frac{E[\text{Var}[Y | X]]}{\text{Var}[Y]}$$

Veamos que  $\eta_{Y/X}^2$  solventa los dos problemas que presentábamos al usar  $\text{Var}[E[Y | X]]$  para medir la bondad de la aproximación.

- $\eta_{Y/X}^2$  es adimensional, ya que las unidades de medida de  $\text{Var}[E[Y | X]]$  y  $\text{Var}[Y]$  son las mismas.
- Veamos que es invariante frente a cambios de escala. Sean  $Y$  e  $Y'$  dos variables aleatorias, y  $Y' = aY$  con  $a \in \mathbb{R}$ . Entonces:

$$\eta_{Y'/X}^2 = \eta_{aY/X}^2 = \frac{\text{Var}[E[aY | X]]}{\text{Var}[aY]} = \frac{\text{Var}[aE[Y | X]]}{a^2 \text{Var}[Y]} = \frac{a^2 \text{Var}[E[Y | X]]}{a^2 \text{Var}[Y]} = \eta_{Y/X}^2$$

Por tanto, la razón de correlación es un valor adimensional e invariante frente a cambios de escala que nos permite comparar la bondad de la aproximación de una variable a partir de la otra. En particular,  $\eta_{Y/X}^2$  mide la proporción de la varianza de  $Y$  que queda explicada por la función de regresión  $\hat{Y} = E[Y | X]$ . En este sentido, se puede interpretar como una medida de la bondad del ajuste de la distribución a la curva de regresión correspondiente, de forma que *a mayor valor del coeficiente, mejor será la aproximación*.

**Proposición 1.14.** Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Entonces:

1.  $0 \leq \eta_{Y/X}^2, \eta_{X/Y}^2 \leq 1$ .
2.  $\eta_{Y/X}^2 = 0 \iff$  La curva de regresión de  $Y$  sobre  $X$  es  $Y = E[Y]$ .  
 $\eta_{X/Y}^2 = 0 \iff$  La curva de regresión de  $X$  sobre  $Y$  es  $X = E[X]$ .
3.  $\eta_{Y/X}^2 = 1 \iff Y$  depende funcionalmente de  $X$ .  
 $\eta_{X/Y}^2 = 1 \iff X$  depende funcionalmente de  $Y$ .
4.  $\eta_{Y/X}^2 = \eta_{X/Y}^2 = 1 \iff X$  e  $Y$  tienen dependencia recíproca.

*Demostración.* Demostraremos tan solo los resultados para  $\eta_{Y/X}^2$ , ya que los resultados para  $\eta_{X/Y}^2$  son análogos.

1. Tenemos en primer lugar  $\text{Var}[E[Y | X]] \geq 0$  y  $\text{Var}[Y] \geq 0$  por ser varianzas. Por tanto lado, usamos que:

$$\eta_{Y/X}^2 = 1 - \frac{E[\text{Var}[Y | X]]}{\text{Var}[Y]}$$

Tenemos que  $\text{Var}[Y | X], \text{Var}[Y] \geq 0$  por ser varianzas. Además, como la esperanza de una variable aleatoria no negativa es no negativa, tenemos que  $E[\text{Var}[Y | X]] \geq 0$ . Por tanto,  $\eta_{Y/X}^2 \leq 1$  y tenemos lo buscado.

2. Tenemos que:

$$\eta_{Y/X}^2 = 0 \iff \text{Var}[E[Y | X]] = 0 \iff \exists c \in \mathbb{R} \text{ tal que } E[Y | X] = c \quad \forall x \in E_X$$

donde la última implicación se debe a las propiedades de la varianza. Por tanto, tenemos que:

$$c = E[c] = E[E[Y | X]] = E[Y]$$

Por tanto, tenemos que:

$$\eta_{Y/X}^2 = 0 \iff E[Y | X] = E[Y]$$

3. Tenemos que:

$$\begin{aligned} \eta_{Y/X}^2 = 1 &\iff E[\text{Var}[Y | X]] = 0 \iff P[\text{Var}[Y | X] = 0] = 1 \iff \\ &\iff \text{Var}[Y | X] = 0 \end{aligned}$$

Por tanto, tenemos que  $Y$  depende funcionalmente de  $X$ .

4. Por lo razonado anteriormente, tenemos que:

$$\eta_{Y/X}^2 = \eta_{X/Y}^2 = 1 \iff \text{Var}[Y | X] = 0 = \text{Var}[X | Y]$$

Por tanto,  $X$  e  $Y$  tienen dependencia recíproca.

□

En esta última proposición, vemos que nos falta estudiar el caso en el que se da  $\eta_{Y/X}^2 = \eta_{X/Y}^2 = 0$ . Este lo veremos más adelante, porque se trata de ejemplos de rectas de regresión.

## 1.5. Rectas de Regresión

Cuando el cálculo de la esperanza condicionada es complicado y se tiene relativa seguridad de que los puntos de la distribución teórica se ajustan a una determinada forma funcional (exponencial, parabólica, etc.) puede ser útil restringir la búsqueda de la función de regresión óptima a la clase de funciones de dicha forma.

Un caso particularmente importante es el de la regresión lineal. Las rectas que mejor se ajustan a los puntos de la distribución en el sentido de mínimos cuadrados (para aproximar cada una de las variables en términos de la otra) se denominan rectas de regresión.

*Observación.* Supongamos de nuevo que queremos obtener la recta de regresión de  $Y$  sobre  $X$ , es decir,  $Y = \varphi(X)$ .

En estos casos, tenemos que  $\varphi$  ha de ser una recta, por lo que:

$$\varphi(X) = aX + b \quad a, b \in \mathbb{R}$$

Por tanto, razonando de forma idéntica a como hicimos en la sección anterior, tenemos que el problema de regresión se reduce a encontrar los valores de  $a$  y  $b$  que minimizan el E.C.M. asociado a la aproximación. Es decir, la recta de regresión de  $Y$  sobre  $X$  es:

$$\varphi_{\text{opt}}^L(X) = \min_{a, b \in \mathbb{R}} E[(Y - aX - b)^2]$$

El problema de regresión lineal se reduce por tanto a minimizar la siguiente función de dos variables:

$$L(a, b) = E[(Y - aX - b)^2] = E[Y^2] + a^2 E[X^2] + b^2 - 2aE[XY] - 2bE[Y] + 2abE[X]$$

Para calcular el mínimo de  $L$ , derivamos parcialmente e igualamos a 0:

$$\begin{aligned} \frac{\partial L}{\partial a} &= 2aE[X^2] - 2E[XY] + 2bE[X] = 0 \\ \frac{\partial L}{\partial b} &= 2b - 2E[Y] + 2aE[X] = 0 \implies 2b = 2E[Y] - 2aE[X] \end{aligned}$$

Por tanto, suponiendo  $\text{Var}[X] > 0$  (caso que estudiaremos más adelante), tenemos que:

$$\begin{aligned}
 2aE[X^2] - 2E[XY] + (2E[Y] - 2aE[X])E[X] &= 0 \implies \\
 \implies a(E[X^2] - E[X]^2) &= E[XY] - E[X]E[Y] \implies \\
 \implies a &= \frac{\text{Cov}(X, Y)}{\text{Var}[X]} \implies \\
 \implies b &= E[Y] - \frac{\text{Cov}(X, Y)}{\text{Var}[X]} \cdot E[X]
 \end{aligned}$$

que, como se comprueba calculando la matriz de derivadas segundas<sup>1</sup>, proporciona el mínimo de la función  $L$ :

$$\begin{vmatrix} \frac{\partial^2 L}{\partial a^2} & \frac{\partial^2 L}{\partial a \partial b} \\ \frac{\partial^2 L}{\partial b \partial a} & \frac{\partial^2 L}{\partial b^2} \end{vmatrix} = \begin{vmatrix} 2E[X^2] & 2E[X] \\ 2E[X] & 2 \end{vmatrix} = 4(E[X^2] - E[X]^2) = 4 \text{Var}[X] > 0$$

Por tanto, tenemos que la recta de regresión de  $Y$  sobre  $X$  es:

$$\hat{Y} = \varphi_{\text{opt}}^L(X) = E[Y] + \frac{\text{Cov}(X, Y)}{\text{Var}[X]}(X - E[X])$$

Este resultado refuerza el resultado que ya vimos del signo de la covarianza:

- Si  $\text{Cov}(X, Y) > 0$ , entonces las rectas de regresión son crecientes.
- Si  $\text{Cov}(X, Y) < 0$ , entonces las rectas de regresión son decrecientes.
- Si  $\text{Cov}(X, Y) = 0$ , entonces las rectas de regresión son constantes e iguales a la esperanza de la variable explicada.

Calculemos ahora su error cuadrático medio:

$$\begin{aligned}
 \text{E.C.M.}(\varphi_{\text{opt}}^L) &= E[(Y - \varphi_{\text{opt}}^L(X))^2] = E \left[ \left( Y - E[Y] - \frac{\text{Cov}(X, Y)}{\text{Var}[X]}(X - E[X]) \right)^2 \right] = \\
 &= E \left[ (Y - E[Y])^2 - 2 \frac{\text{Cov}(X, Y)}{\text{Var}[X]}(Y - E[Y])(X - E[X]) + \frac{\text{Cov}^2(X, Y)}{\text{Var}^2[X]}(X - E[X])^2 \right] = \\
 &= E[(Y - E[Y])^2] - 2 \frac{\text{Cov}(X, Y)}{\text{Var}[X]} E[(Y - E[Y])(X - E[X])] + \frac{\text{Cov}^2(X, Y)}{\text{Var}^2[X]} E[(X - E[X])^2] = \\
 &= \text{Var}[Y] - 2 \frac{\text{Cov}(X, Y)}{\text{Var}[X]} \cdot \text{Cov}(X, Y) + \frac{\text{Cov}^2(X, Y)}{\text{Var}^2[X]} \cdot \text{Var}[X] = \\
 &= \text{Var}[Y] - \frac{\text{Cov}^2(X, Y)}{\text{Var}[X]}
 \end{aligned}$$

Veamos ahora el caso particular en el que  $\text{Var}[X] = 0$ . Entonces,  $X = k \in \mathbb{R}$  es una variable degenerada (constante). Por tanto, cualquier función lineal de  $X$

<sup>1</sup>Minimización de funciones en varias variables es materia de Análisis Matemático II.



es constante, por lo que el problema de regresión lineal se reduce a encontrar la constante que minimiza  $E[(Y - c)^2]$ . Dicha constante es la esperanza de  $Y$ :

$$\varphi_{\text{opt}}^L(X) = E[Y]$$

De aquí en adelante, supondremos que  $\text{Var}[X] > 0$  cuando estemos trabajando con la regresión lineal de  $Y$  sobre  $X$ , y análogamente para la regresión lineal de  $X$  sobre  $Y$ .

### 1.5.1. Coeficiente de determinación lineal

Veamos ahora el equivalente a las razones de correlación en el caso de la regresión lineal. Para calcularlo, seguimos el mismo razonamiento que en la sección anterior, y este será:

$$\frac{\text{Var}[\varphi_{\text{opt}}^L(X)]}{\text{Var}[Y]}$$

Calculemos dicha varianza:

$$\begin{aligned} \text{Var}[\varphi_{\text{opt}}^L(X)] &= \text{Var}\left[E[Y] + \frac{\text{Cov}(X, Y)}{\text{Var}[X]}(X - E[X])\right] = \\ &= \frac{\text{Cov}^2(X, Y)}{\text{Var}^2[X]} \cdot \text{Var}[X] = \frac{\text{Cov}^2(X, Y)}{\text{Var}[X]} \end{aligned}$$

Por tanto, tenemos que:

$$\frac{\text{Var}[\varphi_{\text{opt}}^L(X)]}{\text{Var}[Y]} = \frac{\text{Cov}^2(X, Y)}{\text{Var}[X] \cdot \text{Var}[Y]}$$

Introducimos entonces la siguiente definición:

**Definición 1.13** (Coeficiente de determinación lineal). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, se define el coeficiente de determinación lineal de  $Y$  sobre  $X$ , notado como  $\rho_{X,Y}^2$ , como:

$$\rho_{X,Y}^2 = \frac{\text{Cov}^2(X, Y)}{\text{Var}[X] \cdot \text{Var}[Y]}$$

*Observación.* Notemos que este coeficiente coincide con el producto de las pendientes de las rectas de regresión de  $Y$  sobre  $X$  y de  $X$  sobre  $Y$ . Esto nos será muy útil en la práctica.

Como primer resultado directo que se deduce de la definición, tenemos que:

$$\rho_{X,Y}^2 = \rho_{Y,X}^2$$

Por tanto, tenemos que ambos coeficientes de correlación son iguales, lo que nos permite referirnos simplemente a  $\rho_{X,Y}^2$  como coeficiente de determinación lineal.

Al igual que en el caso de las razones de correlación, el coeficiente de determinación lineal es un valor adimensional e invariante frente a cambios de escala que nos permite comparar la bondad de la aproximación de una variable a partir de la otra.

En este sentido, se puede interpretar como una medida de la bondad del ajuste de la distribución a la recta de regresión correspondiente, de forma que *a mayor valor del coeficiente, mejor será la aproximación*.

Como resultados, incluimos las siguientes propiedades del coeficiente de determinación lineal:

**Proposición 1.15.** *Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Entonces:*

1.  $\rho_{aX+b, cY+d}^2 = \rho_{X,Y}^2$  para  $a, b, c, d \in \mathbb{R}$ .

2.  $0 \leq \rho_{X,Y}^2 \leq 1$ .

3. *Tenemos que:*

$$\begin{aligned} 0 = \rho_{X,Y}^2 &\iff \text{La recta de regresión de } Y \text{ sobre } X \text{ es } Y = E[Y] \\ &\iff \text{La recta de regresión de } X \text{ sobre } Y \text{ es } X = E[X] \end{aligned}$$

4.  $1 = \rho_{X,Y}^2 \iff X$  y  $Y$  tienen dependencia funcional lineal recíproca.

5.  $\rho_{X,Y}^2 \leq \eta_{Y/X}^2, \eta_{X/Y}^2$ .

6.  $\rho_{X,Y}^2 = \eta_{Y/X}^2 \iff$  La curva de regresión de  $Y$  sobre  $X$  coincide con la recta de regresión de  $Y$  sobre  $X$ .

$$\rho_{X,Y}^2 = \eta_{X/Y}^2 \iff \text{La curva de regresión de } X \text{ sobre } Y \text{ coincide con la recta de regresión de } X \text{ sobre } Y.$$

### 1.5.2. Coeficiente de correlación lineal de Pearson

El coeficiente de determinación lineal es una medida de la bondad de la aproximación de una variable a partir de la otra, pero no nos da información sobre la dirección de la relación entre las variables. Para ello, introducimos el siguiente concepto:

**Definición 1.14** (Coeficiente de correlación lineal de Pearson). Dadas dos variables aleatorias  $X$  e  $Y$  sobre el mismo espacio de probabilidad, se define el coeficiente de correlación lineal de Pearson entre  $Y$  y  $X$ , notado como  $\rho_{X,Y}$ , como:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X] \cdot \text{Var}[Y]}}$$

Como primer resultado, de la definición deducimos que  $\rho_{X,Y} = \rho_{Y,X}$ , por lo que nos referiremos a  $\rho_{X,Y}$  como coeficiente de correlación lineal de Pearson. Notemos que  $\rho_{X,Y}$  es la raíz cuadrada del coeficiente de determinación lineal empleado el signo de la covarianza. Debido a que el signo de  $\rho_{X,Y}$  coincide con el signo de la covarianza, y este nos da la dirección de la relación entre las variables, el coeficiente de correlación lineal de Pearson nos da información sobre la dirección de la relación entre las variables.

- Si  $\rho_{X,Y} > 0$ , entonces las variables están positivamente correladas; es decir, si una de las variables aumenta, la otra también lo hace.

- Si  $\rho_{X,Y} < 0$ , entonces las variables están negativamente correladas; es decir, si una de las variables aumenta, la otra disminuye.
- Si  $\rho_{X,Y} = 0$ , entonces las rectas de regresión son constantes e iguales a las esperanzas de las variables dependientes. En este caso, se dice que son *inco-reladas*.

Como resultado directo de esta última definición, tenemos la siguiente caracterización:

**Proposición 1.16.** *Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Entonces:*

$$X, Y \text{ son incorreladas} \iff \text{Cov}[X, Y] = 0$$

Llegados a este punto, podemos estudiar el caso que nos faltaba de las curvas de regresión, que es el caso en el que  $\eta_{Y/X}^2 = \eta_{X/Y}^2 = 0$ .

**Proposición 1.17.** *Sea  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Entonces:*

$$\eta_{Y/X}^2 = \eta_{X/Y}^2 = 0 \implies X, Y \text{ son incorreladas.}$$

*Demostración.* Al ser ambas razones de correlación nulas, tenemos que las curvas de regresión son las esperanzas (constantes), luego son rectas y tenemos que:

$$\rho_{X,Y}^2 = \eta_{Y/X}^2 = \eta_{X/Y}^2 = 0$$

Por tanto, tenemos que  $X$  e  $Y$  son incorreladas. □

El siguiente resultado se deduce de forma directa de la definición del coeficiente de correlación lineal de Pearson y de las propiedades del coeficiente de determinación lineal:

**Proposición 1.18.** *Sean  $X$  e  $Y$  dos variables aleatorias sobre el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ . Entonces:*

1.  $-1 \leq \rho_{X,Y} \leq 1$ .
2.  $\rho_{X,Y}$  es invariante frente a cambios de escala (salvo signo). Esto se suele notar como:

$$\rho_{aX+b, cY+d} = \pm \rho_{X,Y} \quad a, b, c, d \in \mathbb{R}$$



## 2. Relaciones de problemas

## 2.1. Esperanza Condicionada

**Ejercicio 2.1.1.** Sea  $X$  una variable aleatoria que se distribuye uniformemente en el intervalo  $]0, 1[$ . Comprobar si las variables aleatorias  $X$  y  $|1/2 - X|$  son incorreladas.

Como  $X$  es uniforme en  $]0, 1[$ , tenemos que:

$$f_X(x) = \begin{cases} 1, & x \in ]0, 1[ \\ 0, & \text{en otro caso} \end{cases}$$

Calculamos las siguientes esperanzas:

$$E[X] = \int_0^1 x f_X(x) dx = \int_0^1 x dx = \left[ \frac{x^2}{2} \right]_0^1 = \frac{1}{2}$$

$$E[|1/2 - X|] = \int_0^{1/2} (1/2 - x) dx + \int_{1/2}^1 (x - 1/2) dx = \left[ \frac{x}{2} - \frac{x^2}{2} \right]_0^{1/2} + \left[ \frac{x^2}{2} - \frac{x}{2} \right]_{1/2}^1 = \frac{1}{4}$$

$$E[X \cdot |1/2 - X|] = \int_0^{1/2} x(1/2 - x) dx + \int_{1/2}^1 x(x - 1/2) dx = \left[ \frac{x^2}{4} - \frac{x^3}{3} \right]_0^{1/2} + \left[ \frac{x^3}{3} - \frac{x^2}{4} \right]_{1/2}^1 = \frac{1}{8}$$

Calculamos ahora la covarianza:

$$\text{Cov}[X, |1/2 - X|] = E[X \cdot |1/2 - X|] - E[X]E[|1/2 - X|] = \frac{1}{8} - \frac{1}{2} \cdot \frac{1}{4} = 0$$

Por tanto, tenemos que  $X$  y  $|1/2 - X|$  son incorreladas.

**Ejercicio 2.1.2.** Calcular las curvas de regresión y las razones de correlación para las siguientes distribuciones, comentando los resultados.

1. Considerar la distribución conjunta  $(X, Y)$  con función de masa de probabilidad dada por:

$X \setminus Y$	10	15	20	
1	0	$2/6$	0	$2/6$
2	$1/6$	0	0	$1/6$
3	0	0	$3/6$	$3/6$
	$1/6$	$2/6$	$3/6$	1

Tras haber calculado las distribuciones marginales, calculamos ahora las distribuciones condicionadas. La siguiente tabla muestra la distribución condicionada de  $Y$  dado  $X$ ,  $P[Y = y | X = x]$ :

$X \setminus Y$	10	15	20
1	0	1	0
2	1	0	0
3	0	0	1

La distribución condicionada de  $X$  dado  $Y$ ,  $P[X = x | Y = y]$  viene dada por la misma tabla, ya que en este caso tenemos que:

$$P[X = x | Y = y] = P[Y = y | X = x] \quad \forall x, y$$

Calculemos ahora las curvas de regresión y las razones de correlación.

- Curva de regresión de  $Y$  sobre  $X$ :

$$\hat{Y}(x) = E[Y | X = x] = \sum_y yP[Y = y | X = x] \quad \forall x \in E_x$$

Por tanto, la curva de regresión de  $Y$  sobre  $X$  es:

$$\hat{Y}(1) = 15 \quad \hat{Y}(2) = 10 \quad \hat{Y}(3) = 20$$

Para calcular la razón de correlación de  $Y$  sobre  $X$ , hay dos opciones:

**Opción 1.** Método rutinario.

Usamos la fórmula:

$$\eta_{Y/X}^2 = \frac{\text{Var}(E[Y | X])}{\text{Var}(Y)}$$

- Calculemos  $E[Y]$ :

$$E[Y] = \sum_y yP[Y = y] = 10 \cdot 1/6 + 15 \cdot 2/6 + 20 \cdot 3/6 = \frac{50}{3}$$

- Calculemos ahora  $E[Y^2]$ :

$$E[Y^2] = \sum_y y^2 P[Y = y] = 10^2 \cdot 1/6 + 15^2 \cdot 2/6 + 20^2 \cdot 3/6 = \frac{875}{3}$$

- Calculemos ahora  $E[(E[Y | X])^2]$ :

$$\begin{aligned} E[(E[Y | X])^2] &= \sum_x (E[Y | X = x])^2 P[X = x] = \\ &= 10^2 \cdot 1/6 + 15^2 \cdot 2/6 + 20^2 \cdot 3/6 = \frac{875}{3} = E[Y^2] \end{aligned}$$

- Calculemos ahora  $E[E[Y | X]]$ .

$$E[E[Y | X]] = E[Y]$$

Por tanto, usando lo anterior, tenemos:

$$\begin{aligned} \eta_{Y/X}^2 &= \frac{\text{Var}(E[Y | X])}{\text{Var}(Y)} = \frac{E[(E[Y | X])^2] - E[E[Y | X]]^2}{E[Y^2] - E[Y]^2} \\ &= \frac{E[Y^2] - E[Y]^2}{E[Y^2] - E[Y]^2} = 1 \end{aligned}$$

**Opción 2.** Razonando por dependencia funcional.

En este caso, vemos que  $Y$  es función de  $X$ . Por tanto:

$$\eta_{Y/X}^2 = 1$$

- Curva de regresión de  $X$  sobre  $Y$ :

$$\hat{X}(y) = E[X | Y = y] = \sum_x xP[X = x | Y = y] \quad \forall y \in E_y$$

Por tanto, la curva de regresión de  $X$  sobre  $Y$  es:

$$\hat{X}(10) = 2 \quad \hat{X}(15) = 1 \quad \hat{X}(20) = 3$$

De nuevo, razonando ahora por dependencia funcional, tenemos que:

$$\eta_{X/Y}^2 = 1$$

Como vemos, en este caso, hay dependencia recíproca entre  $X$  e  $Y$ . Por tanto, el ajuste es el ideal, ya que  $Y = f(X)$  y  $X = g(Y)$ . Cada una explica la totalidad de la variabilidad de la otra.

2. Considerar la distribución conjunta  $(X, Y)$  con función de masa de probabilidad dada por:

$X \setminus Y$	10	15	20	25	
1	0	$3/7$	0	$1/7$	$4/7$
2	0	0	$1/7$	0	$1/7$
3	$2/7$	0	0	0	$2/7$
	$2/7$	$3/7$	$1/7$	$1/7$	1

Tras haber calculado las distribuciones marginales, calculamos ahora las distribuciones condicionadas. La siguiente tabla muestra la distribución condicionada de  $Y$  dado  $X$ ,  $P[Y = y | X = x]$ :

$X \setminus Y$	10	15	20	25
1	0	$3/4$	0	$1/4$
2	0	0	1	0
3	1	0	0	0

La distribución condicionada de  $X$  dado  $Y$ ,  $P[X = x | Y = y]$  viene dada por la siguiente tabla:

$X \setminus Y$	10	15	20	25
1	0	1	0	1
2	0	0	1	0
3	1	0	0	0

Calculemos ahora las curvas de regresión y las razones de correlación.

- Curva de regresión de  $Y$  sobre  $X$ :

$$\hat{Y}(x) = E[Y | X = x] = \sum_y yP[Y = y | X = x] \quad \forall x \in E_x$$

Por tanto, la curva de regresión de  $Y$  sobre  $X$  es:

$$\hat{Y}(1) = 15 \cdot 3/4 + 25 \cdot 1/4 = 17,5$$

$$\hat{Y}(2) = 20$$

$$\hat{Y}(3) = 10$$



Para calcular la razón de correlación de  $Y$  sobre  $X$ , tenemos que:

$$\eta_{Y/X}^2 = \frac{\text{Var}(E[Y | X])}{\text{Var}(Y)}$$

- Calculemos  $E[Y]$ :

$$E[Y] = \sum_y yP[Y = y] = 10 \cdot 2/7 + 15 \cdot 3/7 + 20 \cdot 1/7 + 25 \cdot 1/7 = \frac{110}{7}$$

- Calculemos ahora  $E[Y^2]$ :

$$\begin{aligned} E[Y^2] &= \sum_y y^2 P[Y = y] \\ &= 10^2 \cdot 2/7 + 15^2 \cdot 3/7 + 20^2 \cdot 1/7 + 25^2 \cdot 1/7 = \frac{1900}{7} \end{aligned}$$

- Calculemos ahora  $E[(E[Y | X])^2]$ :

$$\begin{aligned} E[(E[Y | X])^2] &= \sum_x (E[Y | X = x])^2 P[X = x] \\ &= 17,5^2 \cdot 4/7 + 20^2 \cdot 1/7 + 10^2 \cdot 2/7 = \frac{1825}{7} \end{aligned}$$

- Calculemos ahora  $E[E[Y | X]]$ .

$$E[E[Y | X]] = E[Y]$$

Por tanto, usando lo anterior, tenemos:

$$\begin{aligned} \eta_{Y/X}^2 &= \frac{\text{Var}(E[Y | X])}{\text{Var}(Y)} = \frac{E[(E[Y | X])^2] - E[E[Y | X]]^2}{E[Y^2] - E[Y]^2} \\ &= \frac{E[(E[Y | X])^2] - E[Y]^2}{E[Y^2] - E[Y]^2} \\ &= \frac{\frac{1825}{7} - \left(\frac{110}{7}\right)^2}{\frac{1900}{7} - \left(\frac{110}{7}\right)^2} = \frac{9}{16} \approx 0,5625 \end{aligned}$$

Por tanto, tenemos que  $X$  explica el 56,25 % de la variabilidad de  $Y$ . Tenemos entonces que no es un ajuste ideal.

- Curva de regresión de  $X$  sobre  $Y$ :

En este caso, tenemos que  $X = f(Y)$ , por lo que:

$$\hat{X}(10) = 3, \quad \hat{X}(15) = 1, \quad \hat{X}(20) = 2, \quad \hat{X}(25) = 1$$

Por tanto, como  $X$  es función de  $Y$ , tenemos que:

$$\eta_{X/Y}^2 = 1$$

Tenemos que  $Y$  explica la totalidad de la variabilidad de  $X$ , por lo que el ajuste es el ideal.

**Ejercicio 2.1.3.** Sea  $X$  el número de balanzas e  $Y$  el número de dependientes en los puntos de venta de un mercado. Determinar las rectas de regresión y el grado de ajuste a la distribución, si la función masa de probabilidad de  $(X, Y)$  viene dada por:

$X \setminus Y$	1	2	3	4	
1	$1/24$	$2/24$	0	0	$3/24$
2	$1/24$	$2/24$	$3/24$	$1/24$	$7/24$
3	0	$1/24$	$2/24$	$6/24$	$9/24$
4	0	0	$2/24$	$3/24$	$5/24$
	$2/24$	$5/24$	$7/24$	$10/24$	1

Tras calcular las distribuciones marginales, calculamos ahora las esperanzas necesarias para que el resto de cálculos posteriormente sean directos.

$$E[X] = 1 \cdot 3/24 + 2 \cdot 7/24 + 3 \cdot 9/24 + 4 \cdot 5/24 = 8/3$$

$$E[Y] = 1 \cdot 2/24 + 2 \cdot 5/24 + 3 \cdot 7/24 + 4 \cdot 10/24 = 73/24$$

$$E[X^2] = 1^2 \cdot 3/24 + 2^2 \cdot 7/24 + 3^2 \cdot 9/24 + 4^2 \cdot 5/24 = 8$$

$$E[Y^2] = 1^2 \cdot 2/24 + 2^2 \cdot 5/24 + 3^2 \cdot 7/24 + 4^2 \cdot 10/24 = 245/24$$

$$E[XY] = \sum_{\substack{x \in E_X \\ y \in E_Y}} xyP[X = x, Y = y] = 209/24$$

$$\text{Var}[X] = E[X^2] - E[X]^2 = 8/9$$

$$\text{Var}[Y] = E[Y^2] - E[Y]^2 = 551/576$$

$$\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = 43/72$$

Calculamos ahora las rectas de regresión.

- Recta de regresión de  $Y$  sobre  $X$ :

$$\begin{aligned}\hat{Y}(x) &= E[Y] + \frac{\text{Cov}[X, Y]}{\text{Var}[X]}(X - E[X]) = \frac{73}{24} + \frac{43/72}{8/9}(X - 8/3) = \\ &= 73/24 + \frac{43}{64}(X - 8/3)\end{aligned}$$

- Recta de regresión de  $X$  sobre  $Y$ :

$$\begin{aligned}\hat{X}(y) &= E[X] + \frac{\text{Cov}[X, Y]}{\text{Var}[Y]}(Y - E[Y]) = 8/3 + \frac{43/72}{551/576}(Y - 73/24) = \\ &= 8/3 + \frac{344}{551}(Y - 73/24)\end{aligned}$$

Para determinar el grado de ajuste a la distribución, calculamos ahora el coeficiente de determinación lineal:

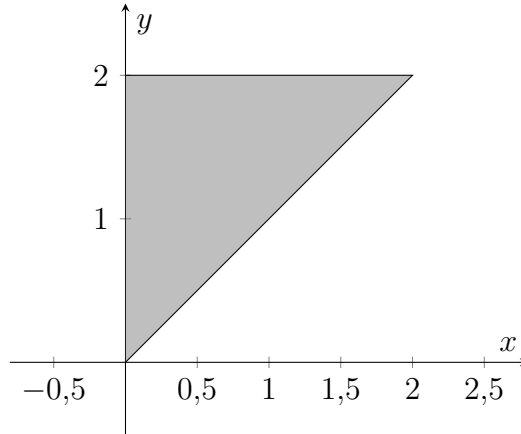
$$\rho_{X,Y}^2 = \frac{\text{Cov}[X, Y]^2}{\text{Var}[X] \text{Var}[Y]} = \frac{1849}{4408} \approx 0,419$$

Por tanto, tenemos que el 41,9% de la variabilidad de  $Y$  queda explicada por la regresión lineal de  $Y$  sobre  $X$ . Vemos por tanto que el ajuste no es bueno.

**Ejercicio 2.1.4.** Sea  $(X, Y)$  un vector aleatorio con valores en el conjunto dado por  $T = \{(x, y) \in \mathbb{R}^2 / 0 < x < y < 2\}$  y función de densidad constante. Calcular:

1. Función de densidad de probabilidad conjunta.

Veamos en primer lugar el conjunto en el que se distribuye el vector aleatorio  $(X, Y)$ :



Como la función de densidad es constante, tenemos que:

$$f_{(X,Y)}(x, y) = \begin{cases} k, & (x, y) \in T \\ 0, & \text{en otro caso} \end{cases}$$

Para calcular  $k$ , tenemos que:

$$1 = \int_T f_{(X,Y)} = \int_T k = k\lambda(T) = k \cdot 2 \implies k = 1/2$$

2. Curvas y rectas de regresión de  $X$  sobre  $Y$  y de  $Y$  sobre  $X$ .

Comenzamos calculando las curvas de regresión, ya que si estas son funciones lineales, las rectas de regresión coincidirán con las curvas de regresión. Para ello, calculamos las funciones de densidad marginales y condicionadas.

- Función de densidad de  $X$ . Para  $x \in ]0, 2[$ :

$$f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x, y) dy = \int_x^2 1/2 dy = 1/2(2 - x) = 1 - \frac{x}{2}$$

- Función de densidad de  $Y$ . Para  $y \in ]0, 2[$ :

$$f_Y(y) = \int_{-\infty}^{\infty} f_{(X,Y)}(x, y) dx = \int_0^y 1/2 dx = \frac{y}{2}$$

- Función de densidad condicionada de  $X$  dado  $Y = y^* \in ]0, 2[$ . Para  $x \in ]0, y^*[$ :

$$f_{X|Y=y^*}(x) = \frac{f_{(X,Y)}(x, y^*)}{f_Y(y^*)} = \frac{1/2}{y^*/2} = \frac{1}{y^*}$$

- Función de densidad condicionada de  $Y$  dado  $X = x^* \in ]0, 2[$ . Para  $y \in ]x^*, 2[$ :

$$f_{Y|X=x^*}(y) = \frac{f_{(X,Y)}(x^*, y)}{f_X(x^*)} = \frac{1/2}{1 - x^*/2} = \frac{1}{2 - x^*}$$

Calculamos ahora las curvas de regresión.

- Curva de regresión de  $X$  sobre  $Y$ :

$$\begin{aligned}\hat{X}(y) &= E[X | Y = y] = \int_0^y x f_{X|Y=y}(x) dx = \int_0^y x \cdot \frac{1}{y} dx = \frac{1}{y} \int_0^y x dx = \\ &= \frac{1}{y} \left[ \frac{x^2}{2} \right]_0^y = \frac{y}{2} \quad \forall y \in ]0, 2[ \end{aligned}$$

- Curva de regresión de  $Y$  sobre  $X$ :

$$\begin{aligned}\hat{Y}(x) &= E[Y | X = x] = \int_x^2 y f_{Y|X=x}(y) dy = \int_x^2 y \cdot \frac{1}{2-x} dy = \frac{1}{2-x} \int_x^2 y dy = \\ &= \frac{1}{2-x} \left[ \frac{y^2}{2} \right]_x^2 = \frac{2^2 - x^2}{2(2-x)} = \frac{4 - x^2}{2(2-x)} = \frac{2+x}{2} = 1 + \frac{x}{2} \quad \forall x \in ]0, 2[ \end{aligned}$$

Tenemos además que las curvas de regresión son funciones lineales, por lo que las rectas de regresión coinciden con las curvas de regresión.

### 3. Razones de correlación y coeficiente de correlación lineal.

Como las curvas de regresión son funciones lineales, tenemos que las razones de correlación coinciden y son iguales al coeficiente de determinación lineal.

$$\eta_{Y/X}^2 = \eta_{X/Y}^2 = \rho_{X,Y}^2 = \frac{\text{Cov}[X, Y]^2}{\text{Var}[X] \text{Var}[Y]}$$

Como este es el producto de las pendientes de las rectas de regresión, tenemos que:

$$\rho_{X,Y}^2 = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Como la correlación es positiva por ser ambas rectas crecientes, tenemos que:

$$\rho_{X,Y} = +\sqrt{\rho_{X,Y}^2} = 1/2$$

### 4. Error cuadrático medio asociado a cada una de las funciones de regresión.

Por lo visto en teoría, sabemos que:

$$\begin{aligned}\text{ECM}(\hat{X}) &= \text{Var}[X] - \frac{\text{Cov}[X, Y]^2}{\text{Var}[Y]} \\ \text{ECM}(\hat{Y}) &= \text{Var}[Y] - \frac{\text{Cov}[X, Y]^2}{\text{Var}[X]} \end{aligned}$$

Calculamos por tanto los valores necesarios:

$$E[X] = \int_0^2 x f_X(x) dx = \int_0^2 x(1 - x/2) dx = \left[ \frac{x^2}{2} - \frac{x^3}{6} \right]_0^2 = \frac{4}{2} - \frac{8}{6} = \frac{2}{3}$$

$$E[Y] = \int_0^2 y f_Y(y) dy = \int_0^2 y \cdot y/2 dy = \left[ \frac{y^3}{6} \right]_0^2 = \frac{8}{6} = \frac{4}{3}$$

$$E[X^2] = \int_0^2 x^2 f_X(x) dx = \int_0^2 x^2(1 - x/2) dx = \left[ \frac{x^3}{3} - \frac{x^4}{8} \right]_0^2 = \frac{8}{3} - \frac{16}{8} = \frac{2}{3}$$

$$E[Y^2] = \int_0^2 y^2 f_Y(y) dy = \int_0^2 y^2 \cdot y/2 dy = \left[ \frac{y^4}{8} \right]_0^2 = \frac{16}{8} = 2$$

$$\begin{aligned} E[XY] &= \int_0^2 \int_x^2 xy f_{(X,Y)}(x, y) dy dx = \int_0^2 \int_x^2 xy \cdot 1/2 dy dx = \frac{1}{2} \int_0^2 x \left[ \frac{y^2}{2} \right]_x^2 dx = \\ &= \frac{1}{4} \int_0^2 x (4 - x^2) dx = \frac{1}{4} \left[ 2x^2 - \frac{x^4}{4} \right]_0^2 = \frac{1}{4} (8 - 4) = 1 \end{aligned}$$

$$\text{Var}[X] = E[X^2] - E[X]^2 = \frac{2}{3} - \left( \frac{2}{3} \right)^2 = \frac{2}{9}$$

$$\text{Var}[Y] = E[Y^2] - E[Y]^2 = 2 - \left( \frac{4}{3} \right)^2 = \frac{2}{9}$$

$$\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = 1 - \frac{2}{3} \cdot \frac{4}{3} = \frac{1}{9}$$

Por tanto, tenemos que:

$$\text{ECM}(\hat{X}) = \text{Var}[X] - \frac{\text{Cov}[X, Y]^2}{\text{Var}[Y]} = \frac{2}{9} - \frac{1/9^2}{2/9} = \frac{2}{9} - \frac{1}{18} = \frac{3}{18} = \frac{1}{6}$$

$$\text{ECM}(\hat{Y}) = \text{Var}[Y] - \frac{\text{Cov}[X, Y]^2}{\text{Var}[X]} = \frac{2}{9} - \frac{1/9^2}{2/9} = \frac{2}{9} - \frac{1}{18} = \frac{3}{18} = \frac{1}{6}$$

**Ejercicio 2.1.5.** Dada la función masa de probabilidad del vector aleatorio  $(X, Y)$

$X \setminus Y$	0	1	2	3	
0	0,2	0,2	0,05	0	0,45
1	0,1	0,1	0,1	0,05	0,35
2	0	0,05	0,05	0,1	0,2
	0,3	0,35	0,2	0,15	1

1. Determinar la aproximación lineal mínimo cuadrática de  $Y$  para  $X = 1$ .

Nos piden la recta de regresión de  $Y$  sobre  $X$ , y evaluarla en el punto  $X = 1$ .

Calculamos las esperanzas necesarias para el cálculo de la recta de regresión.

$$E[X] = 0 \cdot 0,3 + 1 \cdot 0,35 + 2 \cdot 0,2 = 0,75$$

$$E[Y] = 0 \cdot 0,3 + 1 \cdot 0,35 + 2 \cdot 0,2 + 3 \cdot 0,15 = 1,2$$

$$E[X^2] = 0^2 \cdot 0,3 + 1^2 \cdot 0,35 + 2^2 \cdot 0,2 = 1,15$$

$$E[XY] = \sum_{\substack{x \in E_X \\ y \in E_Y}} xyP[X = x, Y = y] = 1,35$$

$$\text{Var}[X] = E[X^2] - E[X]^2 = 1,15 - 0,75^2 = 0,5875$$

$$\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = 1,35 - 0,75 \cdot 1,2 = 0,45$$

Calculamos ahora la recta de regresión de  $Y$  sobre  $X$ .

$$\hat{Y}(x) = E[Y] + \frac{\text{Cov}[X, Y]}{\text{Var}[X]}(x - E[X]) = 1,2 + \frac{0,45}{0,5875}(x - 0,75) = 1,2 + \frac{36}{47}(x - 0,75)$$

Evaluando en  $X = 1$ , tenemos que:

$$\hat{Y}(1) = 1,2 + \frac{36}{47}(1 - 0,75) = 1,2 + \frac{36}{47} \cdot 0,25 = 1,2 + \frac{9}{47} \approx 1,391489$$

2. Determinar la aproximación mínimo cuadrática de  $Y$  para  $X = 1$ .

En este caso, piden  $E[Y | X = 1]$ . Para ello, hemos de calcular las distribución condicionada de  $Y$  dado  $X = 1$ .

$$P[Y = 0 | X = 1] = \frac{P[X = 1, Y = 0]}{P[X = 1]} = \frac{0,1}{0,35} = \frac{2}{7}$$

$$P[Y = 1 | X = 1] = \frac{P[X = 1, Y = 1]}{P[X = 1]} = \frac{0,1}{0,35} = \frac{2}{7}$$

$$P[Y = 2 | X = 1] = \frac{P[X = 1, Y = 2]}{P[X = 1]} = \frac{0,1}{0,35} = \frac{2}{7}$$

$$P[Y = 3 | X = 1] = \frac{P[X = 1, Y = 3]}{P[X = 1]} = \frac{0,05}{0,35} = \frac{1}{7}$$

Por tanto, tenemos que:

$$\begin{aligned} E[Y | X = 1] &= \sum_{y \in E_Y} yP[Y = y | X = 1] = \\ &= 0 + 1 \cdot P[Y = 1 | X = 1] + 2 \cdot P[Y = 2 | X = 1] + 3 \cdot P[Y = 3 | X = 1] = \\ &= \frac{2}{7} + 2 \cdot \frac{2}{7} + 3 \cdot \frac{1}{7} = \frac{9}{7} \end{aligned}$$

**Ejercicio 2.1.6.** Dadas las siguientes distribuciones, determinar qué variable,  $X$  ó  $X'$ , aproxima mejor a la variable  $Y$ :

$X \setminus Y$	0	1	2	
0	$\frac{1}{5}$	0	0	$\frac{1}{5}$
2	0	$\frac{1}{5}$	0	$\frac{1}{5}$
3	$\frac{1}{5}$	0	$\frac{1}{5}$	$\frac{2}{5}$
4	0	0	$\frac{1}{5}$	$\frac{1}{5}$
	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$	1

$X' \setminus Y$	0	1	2	
0	$\frac{1}{5}$	0	$\frac{1}{5}$	$\frac{2}{5}$
2	0	$\frac{1}{5}$	0	$\frac{1}{5}$
3	$\frac{1}{5}$	0	0	$\frac{1}{5}$
4	0	0	$\frac{1}{5}$	$\frac{1}{5}$
	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$	1

Hemos de obtener  $\eta_{Y/X}^2$  y  $\eta_{Y/X'}^2$  para compararlas.

$$\eta_{Y/X}^2 = \frac{\text{Var}[E[Y | X]]}{\text{Var}[Y]}, \quad \eta_{Y/X'}^2 = \frac{\text{Var}[E[Y | X']]}{\text{Var}[Y]}$$

Calculamos las esperanzas necesarias para el cálculo de las varianzas.

$$\begin{aligned} E[Y] &= 0 \cdot 2/5 + 1 \cdot 1/5 + 2 \cdot 2/5 = 1 \\ E[Y^2] &= 0^2 \cdot 2/5 + 1^2 \cdot 1/5 + 2^2 \cdot 2/5 = 9/5 \\ E[E[Y | X]] &= E[Y] = 1 \\ E[E[Y | X']] &= E[Y] = 1 \\ E[(E[Y | X])^2] &= \sum_{x \in E_X} (E[Y | X = x])^2 P[X = x] \\ E[(E[Y | X'])^2] &= \sum_{x \in E_{X'}} (E[Y | X' = x])^2 P[X' = x] \end{aligned}$$

Calculamos por tanto las esperanzas condicionadas.

$$\begin{aligned} E[Y | X = 0] &= 0 \cdot 1 + 1 \cdot 0 + 2 \cdot 0 = 0 \\ E[Y | X = 2] &= 0 \cdot 0 + 1 \cdot 1 + 2 \cdot 0 = 1 \\ E[Y | X = 3] &= 0 \cdot 1/2 + 1 \cdot 0 + 2 \cdot 1/2 = 1 \\ E[Y | X = 4] &= 0 \cdot 0 + 1 \cdot 0 + 2 \cdot 1 = 2 \\ E[Y | X' = 0] &= 0 \cdot 1/2 + 1 \cdot 0 + 2 \cdot 1/2 = 1 \\ E[Y | X' = 2] &= 0 \cdot 0 + 1 \cdot 1 + 2 \cdot 0 = 1 \\ E[Y | X' = 3] &= 0 \cdot 1/2 + 1 \cdot 0 + 2 \cdot 1/2 = 1 \\ E[Y | X' = 4] &= 0 \cdot 0 + 1 \cdot 0 + 2 \cdot 1 = 2 \end{aligned}$$

Por tanto, hemos comprobado que  $E[Y | X] = E[Y | X']$ , por lo que:

$$E[Y | X] = E[Y | X'] \implies \text{Var}[E[Y | X]] = \text{Var}[E[Y | X']] \implies \eta_{Y/X}^2 = \eta_{Y/X'}^2$$

Por tanto, ambas variables aproximan igual de bien a  $Y$ . Veamos cómo es la aproximación calculando aun así los coeficientes. Tenemos que:

$$\begin{aligned} E[(E[Y | X])^2] &= \sum_{x \in E_X} (E[Y | X = x])^2 P[X = x] = \\ &= 0^2 \cdot 1/5 + 1^2 \cdot 1/5 + 1^2 \cdot 2/5 + 2^2 \cdot 1/5 = 7/5 \end{aligned}$$

Por tanto, tenemos que:

$$\begin{aligned} \text{Var}[Y] &= E[Y^2] - E[Y]^2 = 9/5 - 1 = 4/5 \\ \text{Var}[E[Y | X]] &= E[(E[Y | X])^2] - E[E[Y | X]]^2 = 7/5 - 1 = 2/5 \end{aligned}$$

Por tanto, tenemos que:

$$\eta_{Y/X}^2 = \frac{\text{Var}[E[Y | X]]}{\text{Var}[Y]} = \frac{2/5}{4/5} = 1/2 = \eta_{Y/X'}^2$$

Por tanto, tenemos que ninguno de los ajustes es bueno, ya que solo explican la mitad de la variabilidad de  $Y$ .

**Ejercicio 2.1.7.** Probar que las variables  $X = U + V$  e  $Y = U - V$  son incorreladas, pero no independientes, si  $U$  y  $V$  son variables aleatorias con función de densidad conjunta:

$$f_{U,V}(u, v) = \exp(-u - v), \quad u, v > 0.$$

Para estudiar si son incorreladas, calcularemos su covarianza. Previamente calculamos las marginales.

$$\begin{aligned} f_U(u) &= \int_0^\infty f_{U,V}(u, v) dv = \int_0^\infty \exp(-u - v) dv = e^{-u} \int_0^\infty e^{-v} dv = e^{-u} [-e^{-v}]_0^\infty = e^{-u} \\ f_V(v) &= \int_0^\infty f_{U,V}(u, v) du = \int_0^\infty \exp(-u - v) du = e^{-v} \int_0^\infty e^{-u} du = e^{-v} [-e^{-u}]_0^\infty = e^{-v} \end{aligned}$$

Por tanto, tenemos que:

$$U, V \sim \exp(1) \implies E[U^k] = E[V^k] = 1 \quad \forall k \in \mathbb{N}$$

Por tanto, tenemos:

$$\begin{aligned} E[X] &= E[U + V] = E[U] + E[V] = 1 + 1 = 2 \\ E[Y] &= E[U - V] = E[U] - E[V] = 1 - 1 = 0 \\ E[XY] &= E[(U + V)(U - V)] = E[U^2 - V^2] = E[U^2] - E[V^2] = 1 - 1 = 0 \end{aligned}$$

Tenemos por tanto que:

$$\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = 0 - 2 \cdot 0 = 0$$

Por tanto, tenemos que  $X$  e  $Y$  son incorreladas.

Veamos que no son independientes. Para ello, calculamos en primer lugar  $f_{X,Y}(x, y)$ . Para ello, definimos la transformación:

$$\begin{aligned} g : \quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (U, V) &\longmapsto (X, Y) = (U + V, U - V) \end{aligned}$$

Para obtener  $g^{-1}$ , buscamos obtener  $U, V$  en función de  $X, Y$ :

$$\begin{cases} X = U + V \\ Y = U - V \end{cases} \implies \begin{cases} U = \frac{X+Y}{2} \\ V = \frac{X-Y}{2} \end{cases}$$

Por tanto, tenemos que  $\exists g^{-1}$ , con:

$$\begin{aligned} g^{-1} : \quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (X, Y) &\longmapsto (U, V) = \left(\frac{X+Y}{2}, \frac{X-Y}{2}\right) \end{aligned}$$

Tenemos que todas las componentes de  $g^{-1}$  son derivables:

$$\frac{\partial U}{\partial X} = \frac{1}{2}, \quad \frac{\partial U}{\partial Y} = \frac{1}{2}, \quad \frac{\partial V}{\partial X} = \frac{1}{2}, \quad \frac{\partial V}{\partial Y} = -\frac{1}{2}$$

Además, veamos que el jacobiano no se anula:

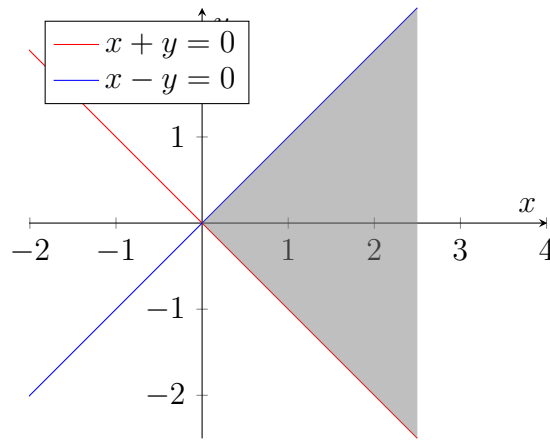
$$\det Jg^{-1}(x, y) = \begin{vmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{vmatrix} = \frac{1}{2^2} \cdot \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -\frac{1}{2} \neq 0 \quad \forall (x, y) \in \mathbb{R}^2.$$



Por tanto, podemos aplicar el teorema del cambio de variable para obtener  $f_{X,Y}(x, y)$ :

$$\begin{aligned} f_{X,Y}(x, y) &= f_{U,V}(g^{-1}(x, y)) \cdot |Jg^{-1}(x, y)| = \frac{1}{2} \cdot f_{U,V}\left(\frac{x+y}{2}, \frac{x-y}{2}\right) \\ &= \frac{1}{2} \cdot \exp\left(-\frac{x+y}{2} - \frac{x-y}{2}\right) = \frac{1}{2} \cdot \exp\left(\frac{-x-y-x+y}{2}\right) = \\ &= \frac{1}{2} \cdot \exp(-x) \quad \forall (x, y) \in \mathbb{R}^2 \text{ tal que } x+y, x-y > 0 \end{aligned}$$

Veamos el conjunto gráficamente:



Tenemos que el conjunto donde  $f_{X,Y}(x, y)$  no se puede expresar como producto cartesiano, por lo que intuimos que  $X$  e  $Y$  no son independientes. Para comprobarlo, calculamos la función de densidad marginal de  $X$  y  $Y$ :

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy = \int_{-x}^x \frac{1}{2} \exp(-x) dy = \frac{1}{2} \exp(-x) [y]_{-x}^x = \\ &= \frac{1}{2} \exp(-x)(x+x) = x e^{-x} \quad \forall x > 0 \end{aligned}$$

Para calcular la marginal de  $Y$ , distinguimos en función de  $y$ :

- Si  $y > 0$ :

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx = \int_y^{\infty} \frac{1}{2} \exp(-x) dx = \frac{1}{2} [-\exp(-x)]_y^{\infty} = \\ &= \frac{1}{2} [0 - (-\exp(-y))] = \frac{e^{-y}}{2} \quad \forall y > 0 \end{aligned}$$

- Si  $y \leq 0$ :

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx = \int_{-y}^{\infty} \frac{1}{2} \exp(-x) dx = \frac{1}{2} [-\exp(-x)]_{-y}^{\infty} = \\ &= \frac{1}{2} [0 - (-\exp(y))] = \frac{e^y}{2} \quad \forall y \leq 0 \end{aligned}$$

Por tanto, tenemos que:

$$f_Y(y) = \begin{cases} \frac{e^{-y}}{2} & y > 0 \\ \frac{e^y}{2} & y \leq 0 \end{cases}$$

Para el  $(1, 1)$ , tenemos que:

$$\begin{aligned} f_X(1) &= e^{-1}, & f_{X,Y}(1, 1) &= \frac{1}{2}e^{-1} \\ f_Y(1) &= \frac{1}{2}e^{-1}, & f_X(1)f_Y(1) &= \frac{1}{2}e^{-2} \neq f_{X,Y}(1, 1) \end{aligned}$$

Por tanto, no se cumple que  $f_{X,Y}(x, y) = f_X(x)f_Y(y)$  para todo  $(x, y) \in \mathbb{R}^2$ , por lo que  $X$  e  $Y$  no son independientes.

**Ejercicio 2.1.8.** Sea  $X$  una variable aleatoria con distribución uniforme en el intervalo  $[0, 1]$ , y sea  $Y$  una variable aleatoria continua tal que

$$f_{Y|X=x}(y) = \begin{cases} 1/x^2 & y \in [0, x^2] \\ 0 & \text{en caso contrario} \end{cases}$$

1. Calcular la función de densidad de probabilidad conjunta de  $X$  e  $Y$ . Calcular la función de densidad de probabilidad marginal de  $Y$ .

Para calcular  $f_{X,Y}(x, y)$ , usamos el siguiente resultado:

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

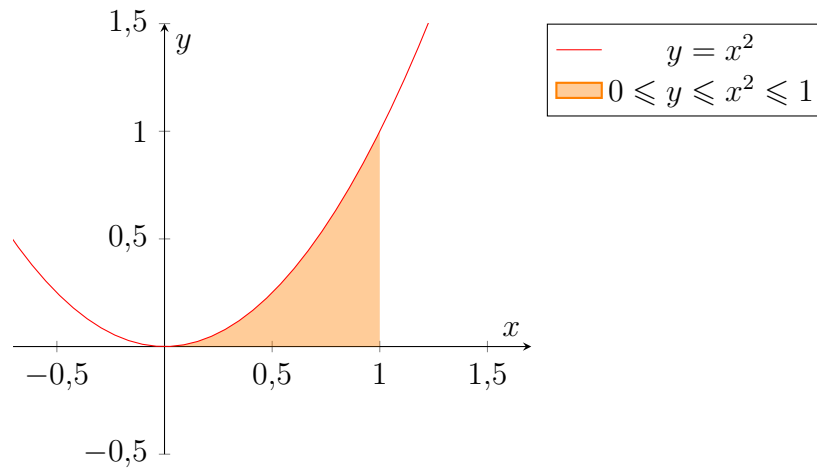
Por tanto, calculamos la marginal de  $X$ . Como esta es uniforme en  $[0, 1]$ , tenemos que:

$$1 = \int_0^1 f_X(x) dx = \int_0^1 k dx = k \implies f_X(x) = 1 \quad \forall x \in [0, 1]$$

Por tanto, tenemos que:

$$f_{X,Y}(x, y) = f_{Y|X=x}(y) \cdot f_X(x) = \begin{cases} 1/x^2 & 0 \leq y \leq x^2 \leq 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Veamos gráficamente el conjunto donde  $f_{X,Y}(x, y)$  no se anula:



Para calcular la marginal de  $Y$ , para cada  $y \in [0, 1]$ , tenemos que:

$$f_Y(y) = \int_{\sqrt{y}}^1 f_{X,Y}(x, y) \, dx = \int_{\sqrt{y}}^1 1/x^2 \, dx = \left[ -\frac{1}{x} \right]_{\sqrt{y}}^1 = -1 + \frac{1}{\sqrt{y}}$$

Por tanto, tenemos que:

$$f_Y(y) = \begin{cases} -1 + 1/\sqrt{y} & 0 \leq y \leq 1 \\ 0 & \text{en caso contrario} \end{cases}$$

2. Calcular  $E[X | Y = y]$  y  $E[Y | X = x]$ .

Calculamos en primer lugar la distribución condicionada de  $X$  dado un valor  $Y = y \in [0, 1]$ .

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{1/x^2}{-1 + 1/\sqrt{y}} \quad \forall x \in [0, 1] \mid \sqrt{y} \leq x$$

Calculamos ahora por tanto las esperanzas condicionadas:

$$\begin{aligned} E[X | Y = y] &= \int_{\sqrt{y}}^1 x \cdot f_{X|Y=y}(x) \, dx = \int_{\sqrt{y}}^1 x \cdot \frac{1/x^2}{-1 + 1/\sqrt{y}} \, dx = \frac{1}{-1 + 1/\sqrt{y}} \int_{\sqrt{y}}^1 1/x \, dx = \\ &= \frac{1}{-1 + 1/\sqrt{y}} [\ln(x)]_{\sqrt{y}}^1 = \frac{1}{-1 + 1/\sqrt{y}} [\ln(1) - \ln(\sqrt{y})] = \\ &= -\frac{\ln(\sqrt{y})}{-1 + 1/\sqrt{y}} \quad \forall y \in [0, 1] \end{aligned}$$

$$\begin{aligned} E[Y | X = x] &= \int_0^{x^2} y \cdot f_{Y|X=x}(y) \, dy = \int_0^{x^2} y \cdot \frac{1}{x^2} \, dy = \frac{1}{x^2} \int_0^{x^2} y \, dy = \\ &= \frac{1}{x^2} \left[ \frac{y^2}{2} \right]_0^{x^2} = \frac{1}{x^2} \cdot \frac{x^4}{2} = \frac{x^2}{2} \quad \forall x \in [0, 1] \end{aligned}$$

3. Para la misma densidad de probabilidad condicionada del apartado 1, considerando ahora que  $X$  es una variable aleatoria continua con función de densidad

de probabilidad:

$$f_X(x) = \begin{cases} 3x^2 & x \in [0, 1] \\ 0 & \text{en caso contrario} \end{cases}$$

Calcular de nuevo la función de densidad de probabilidad conjunta de  $X$  e  $Y$ , y la función de densidad de probabilidad marginal de  $Y$ , así como  $E[X | Y = y]$  y  $E[Y | X = x]$ .

Repetimos todo el proceso anterior, pero con la nueva función de densidad de  $X$ .

$$f_{X,Y}(x, y) = f_{Y|X=x}(y) \cdot f_X(x) = \begin{cases} 3 & 0 \leq y \leq x^2 \leq 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Para calcular la marginal de  $Y$ , para cada  $y \in [0, 1]$ , tenemos que:

$$f_Y(y) = \int_{\sqrt{y}}^1 f_{X,Y}(x, y) dx = 3 \int_{\sqrt{y}}^1 1 dx = 3 [x]_{\sqrt{y}}^1 = 3 - 3\sqrt{y} \quad \forall y \in [0, 1]$$

Calculamos ahora la distribución condicionada de  $X$  dado un valor  $Y = y \in [0, 1]$ .

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{3}{3 - 3\sqrt{y}} = \frac{1}{1 - \sqrt{y}} \quad \forall x \in [0, 1] \mid \sqrt{y} \leq x$$

Calculamos ahora las esperanzas condicionadas:

$$\begin{aligned} E[X | Y = y] &= \int_{\sqrt{y}}^1 x \cdot f_{X|Y=y}(x) dx = \int_{\sqrt{y}}^1 x \cdot \frac{1}{1 - \sqrt{y}} dx = \frac{1}{1 - \sqrt{y}} \int_{\sqrt{y}}^1 x dx = \\ &= \frac{1}{1 - \sqrt{y}} \left[ \frac{x^2}{2} \right]_{\sqrt{y}}^1 = \frac{1}{1 - \sqrt{y}} \left[ \frac{1}{2} - \frac{y}{2} \right] = \frac{1}{2(1 - \sqrt{y})} (1 - y) = \\ &= \frac{1 + \sqrt{y}}{2} \quad \forall y \in [0, 1] \end{aligned}$$

$$\begin{aligned} E[Y | X = x] &= \int_0^{x^2} y \cdot f_{Y|X=x}(y) dy = \int_0^{x^2} y \cdot \frac{1}{x^2} dy = \frac{1}{x^2} \int_0^{x^2} y dy = \\ &= \frac{1}{x^2} \left[ \frac{y^2}{2} \right]_0^{x^2} = \frac{1}{x^2} \cdot \frac{x^4}{2} = \frac{x^2}{2} \quad \forall x \in [0, 1] \end{aligned}$$

**Ejercicio 2.1.9.** Sean  $X$  e  $Y$  variables aleatorias con función de densidad conjunta:

$$f_{(X,Y)}(x, y) = \begin{cases} x + y & (x, y) \in [0, 1] \times [0, 1] \\ 0 & \text{en caso contrario} \end{cases}$$

1. Calcular la predicción mínimo cuadrática de  $Y$  a partir de  $X$  y el error cuadrático medio asociado.

Para calcular la predicción mínimo cuadrática de  $Y$  a partir de  $X$ , hemos de calcular  $E[Y | X]$ . Para ello, hemos de calcular la distribución marginal de  $X$  y la distribución condicionada de  $Y$  dado  $X = x \in [0, 1]$ . condicionada de  $Y$  dado  $X = x \in [0, 1]$ .

$$f_X(x) = \int_0^1 f_{(X,Y)}(x, y) dy = \int_0^1 x + y dy = \left[ xy + \frac{y^2}{2} \right]_0^1 = x + \frac{1}{2} \quad \forall x \in [0, 1]$$

$$f_{Y|X=x}(y) = \frac{f_{(X,Y)}(x, y)}{f_X(x)} = \frac{x + y}{x + 1/2} \quad \forall y \in [0, 1]$$

Calculamos ahora la esperanza condicionada de  $Y$  dado  $X = x$ :

$$\begin{aligned} E[Y | X = x] &= \int_0^1 y \cdot f_{Y|X=x}(y) dy = \int_0^1 y \cdot \frac{x + y}{x + 1/2} dy = \frac{1}{x + 1/2} \int_0^1 y \cdot (x + y) dy = \\ &= \frac{1}{x + 1/2} \int_0^1 xy + y^2 dy = \frac{1}{x + 1/2} \left[ \frac{xy^2}{2} + \frac{y^3}{3} \right]_0^1 = \frac{1}{x + 1/2} \left( \frac{x}{2} + \frac{1}{3} \right) = \\ &= \frac{2}{2x + 1} \left( \frac{3x + 2}{6} \right) = \frac{3x + 2}{6x + 3} \quad \forall x \in [0, 1] \end{aligned}$$

Calculamos su error cuadrático medio:

$$\text{ECM}(E[Y | X]) = E[\text{Var}[Y | X]]$$

Calculamos por tanto en primer lugar  $\text{Var}[Y | X] = E[Y^2 | X] - E[Y | X]^2$ :

$$\begin{aligned} E[Y^2 | X = x] &= \int_0^1 y^2 \cdot f_{Y|X=x}(y) dy = \int_0^1 y^2 \cdot \frac{x + y}{x + 1/2} dy = \frac{1}{x + 1/2} \int_0^1 y^2 \cdot (x + y) dy = \\ &= \frac{1}{x + 1/2} \int_0^1 xy^2 + y^3 dy = \frac{1}{x + 1/2} \left[ \frac{xy^3}{3} + \frac{y^4}{4} \right]_0^1 = \frac{1}{x + 1/2} \left( \frac{x}{3} + \frac{1}{4} \right) = \\ &= \frac{4x + 3}{12x + 6} \quad \forall x \in [0, 1] \end{aligned}$$

$$\begin{aligned} \text{Var}[Y | X = x] &= E[Y^2 | X = x] - E[Y | X = x]^2 = \frac{4x + 3}{12x + 6} - \left( \frac{3x + 2}{6x + 3} \right)^2 = \\ &= \frac{4x + 3}{6(2x + 1)} - \frac{(3x + 2)^2}{9(2x + 1)^2} = \frac{3(4x + 3)(2x + 1) - 2(3x + 2)^2}{18(2x + 1)^2} = \\ &= \frac{3(8x^2 + 10x + 3) - 2(9x^2 + 4 + 12x)}{18(2x + 1)^2} = \\ &= \frac{24x^2 + 30x + 9 - 18x^2 - 8 - 24x}{18(2x + 1)^2} = \\ &= \frac{6x^2 + 6x + 1}{18(2x + 1)^2} \quad \forall x \in [0, 1] \end{aligned}$$

Por tanto, tenemos que:

$$\begin{aligned} \text{ECM}(E[Y | X]) &= E[\text{Var}[Y | X]] = \int_0^1 \frac{6x^2 + 6x + 1}{18(2x + 1)^2} \cdot f_X(x) dx = \\ &= \int_0^1 \frac{6x^2 + 6x + 1}{18(2x + 1)^2} \cdot (x + 1/2) dx = \int_0^1 \frac{6x^3 + 6x^2 + x}{36(2x + 1)} dx \end{aligned}$$

Para resolver dicha integral, realizamos la división de polinomios:

$$\begin{array}{r} (6x^3 + 6x^2 + x) : (2x + 1) = 3x^2 + \frac{3}{2}x - \frac{1}{4} + \frac{\frac{1}{4}}{2x + 1} \\ \underline{-6x^3 - 3x^2} \phantom{+ x} \\ 3x^2 + x \\ \underline{-3x^2 - \frac{3}{2}x} \\ -\frac{1}{2}x \\ \underline{-\frac{1}{2}x + \frac{1}{4}} \\ \frac{1}{4} \end{array}$$

Por tanto, tenemos que:

$$\begin{aligned} \text{ECM}(E[Y | X]) &= \frac{1}{36} \int_0^1 3x^2 + \frac{3}{2}x - \frac{1}{4} + \frac{1/4}{2x + 1} dx = \\ &= \frac{1}{36} \left[ x^3 + \frac{3}{4}x^2 - \frac{x}{4} + \frac{1}{8} \ln(2x + 1) \right]_0^1 = \frac{1}{36} \left( 1 + \frac{3}{4} - \frac{1}{4} + \frac{1}{8} \ln(3) \right) = \\ &= \frac{1}{36} \left( \frac{3}{2} + \frac{1}{8} \ln(3) \right) \approx 0,045 \end{aligned}$$

2. Si se observa  $X = 1/2$ , ¿qué predicción de  $Y$  tiene menor error cuadrático medio? Calcular dicho error.

En este caso, tenemos que la predicción de  $Y$  a partir de  $X = 1/2$  es:

$$E[Y | X = 1/2] = \frac{3 \cdot 1/2 + 2}{6 \cdot 1/2 + 3} = \frac{7}{12}$$

Calculamos su error cuadrático medio. Para esto, hay dos opciones:

**Usando el apartado anterior** En este caso, tenemos que:

$$\begin{aligned} \text{ECM}(E[Y | X = 1/2]) &= E[\text{Var}[Y | X = 1/2]] = \\ &= E \left[ \frac{6 \cdot 1/4 + 6 \cdot 1/2 + 1}{18(2 \cdot 1/2 + 1)^2} \right] = E \left[ \frac{11}{144} \right] = \frac{11}{144} \approx 0,07639 \end{aligned}$$

**Usando la definición de ECM** En este caso, tenemos que:

$$\begin{aligned} \text{ECM}(E[Y | X = 1/2]) &= E[(Y - E[Y | X = 1/2])^2] = E \left[ \left( Y - \frac{7}{12} \right)^2 \right] = \\ &= E[Y^2] - \frac{7}{6}E[Y] + \frac{49}{144} \end{aligned}$$

Para calcular las esperanzas, calculamos en primer lugar la marginal de  $Y$ . Para  $y \in [0, 1]$ :

$$f_Y(y) = \int_0^1 f_{(X,Y)}(x, y) dx = \int_0^1 x + y dx = \left[ \frac{x^2}{2} + xy \right]_0^1 = \frac{1}{2} + y \quad \forall y \in [0, 1]$$

Calculamos por tanto las esperanzas necesarias:

$$\begin{aligned} E[Y] &= \int_0^1 y \cdot f_Y(y) dy = \int_0^1 y \cdot \left(\frac{1}{2} + y\right) dy = \int_0^1 \frac{y}{2} + y^2 dy = \\ &= \left[\frac{y^2}{4} + \frac{y^3}{3}\right]_0^1 = \frac{1}{4} + \frac{1}{3} = \frac{7}{12} \\ E[Y^2] &= \int_0^1 y^2 \cdot f_Y(y) dy = \int_0^1 y^2 \cdot \left(\frac{1}{2} + y\right) dy = \int_0^1 \frac{y^2}{2} + y^3 dy = \\ &= \left[\frac{y^3}{6} + \frac{y^4}{4}\right]_0^1 = \frac{1}{6} + \frac{1}{4} = \frac{5}{12} \end{aligned}$$

Por tanto, tenemos que:

$$\text{ECM}(E[Y | X = 1/2]) = E[Y^2] - \frac{7}{6}E[Y] + \frac{49}{144} = \frac{5}{12} - \frac{7}{6} \cdot \frac{7}{12} + \frac{49}{144} = \frac{11}{144}$$

3. Supóngase que una persona debe pagar una cantidad  $C$  por la oportunidad de observar el valor de  $X$  antes de predecir el valor de  $Y$ , o puede simplemente predecir el valor de  $Y$  sin observar  $X$ . Si la persona considera que su pérdida total es la suma de  $C$  y el error cuadrático medio de su predicción, qué valor máximo de  $C$  estaría dispuesta a pagar?

Estudiemos cada una de las opciones por separado.

**Observar  $X$  antes de predecir  $Y$ :** En este caso, la pérdida total es:

$$C + \text{ECM}(E[Y | X]) = C + \frac{1}{36} \left( \frac{3}{2} + \frac{1}{8} \ln(3) \right)$$

**Predecir  $Y$  sin observar  $X$ :** En este caso, la pérdida total es (ya que no pagamos  $C$ ):

$$\text{ECM}(E[Y]) = E[(Y - E[Y])^2] = \text{Var}[Y] = E[Y^2] - E[Y]^2 = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144}$$

Por tanto, la persona estaría dispuesta a pagar un máximo de:

$$C = \frac{11}{144} - \frac{1}{36} \left( \frac{3}{2} + \frac{1}{8} \ln(3) \right) \approx 0,03090$$

**Ejercicio 2.1.10.** Sea  $(X, Y)$  un vector aleatorio con función de densidad:

$$f_{(X,Y)}(x, y) = e^{-y}, \quad 0 < x < y$$

Obtener y representar las rectas y curvas de regresión. Calcular el coeficiente de correlación lineal, las razones de correlación y el error cuadrático medio cometido al predecir cada variable según cada una de las funciones de regresión. Interpretar los resultados.

Obtenemos en primer lugar las marginales:

$$\begin{aligned} f_X(x) &= \int_x^\infty e^{-y} dy = [-e^{-y}]_x^\infty = e^{-x} \quad \forall x > 0 \\ f_Y(y) &= \int_0^y e^{-y} dx = ye^{-y} \quad \forall y > 0 \end{aligned}$$

Obtenemos ahora las condicionadas. Dados  $x^*, y^* \in \mathbb{R}^+$ , tenemos que:

$$\begin{aligned} f_{Y|X=x^*}(y) &= \frac{f_{(X,Y)}(x^*, y)}{f_X(x^*)} = \frac{e^{-y}}{e^{-x^*}} = e^{x^*-y} \quad \forall y > x^* \\ f_{X|Y=y^*}(x) &= \frac{f_{(X,Y)}(x, y^*)}{f_Y(y^*)} = \frac{e^{-y^*}}{y^* e^{-y^*}} = \frac{1}{y^*} \quad \forall x \in ]0, y^*[ \end{aligned}$$

Calculamos ahora las curvas de regresión. Para la de  $Y$  sobre  $X$ , tenemos que:

$$\hat{Y}(x) = E[Y | X = x] = \int_x^\infty y \cdot e^{x-y} dy = e^x \int_x^\infty y \cdot e^{-y} dy$$

Para calcular la integral, realizamos integración por partes:

$$\left[ \begin{array}{ll} u(y) = y & u'(y) = 1 \\ v'(y) = e^{-y} & v(y) = -e^{-y} \end{array} \right] \implies \int_x^\infty y \cdot e^{-y} dy = [-ye^{-y}]_x^\infty - \int_x^\infty -e^{-y} dy =$$

$$= [-ye^{-y} - e^{-y}]_x^\infty = e^{-x}(x + 1)$$

Por tanto, tenemos que la curva de regresión de  $Y$  sobre  $X$  es:

$$\hat{Y}(x) = e^x e^{-x}(x + 1) = x + 1 \quad \forall x > 0$$

Como vemos, la curva de regresión de  $Y$  sobre  $X$  es  $y = x + 1$ , que es una recta y por tanto coincide con la recta de regresión.

Calculamos ahora la curva de regresión de  $X$  sobre  $Y$ . Para ello, tenemos que:

$$\hat{X}(y) = E[X | Y = y] = \int_0^y x \cdot \frac{1}{y} dx = \frac{1}{y} \int_0^y x dx = \frac{1}{y} \left[ \frac{x^2}{2} \right]_0^y = \frac{y}{2} \quad \forall y > 0$$

Como vemos, la curva de regresión de  $X$  sobre  $Y$  es  $x = y/2$ , que es una recta y por tanto coincide con la recta de regresión.

Calculamos ahora el coeficiente de determinación lineal. Para ello, calculamos



las varianzas y covarianzas:

$$\begin{aligned}
 E[X] &= \int_0^{\infty} x \cdot e^{-x} dx = [-xe^{-x} - e^{-x}]_0^{\infty} = 1 \\
 E[Y] &= \int_0^{\infty} y^2 \cdot e^{-y} dy = [-y^2 e^{-y}]_0^{\infty} + 2 \int_0^{\infty} y \cdot e^{-y} dy = 0 + 2 \cdot 1 = 2 \\
 E[X^2] &= \int_0^{\infty} x^2 \cdot e^{-x} dx = 2 \\
 E[Y^2] &= \int_0^{\infty} y^3 \cdot e^{-y} dy = [-y^3 e^{-y}]_0^{\infty} + 3 \int_0^{\infty} y^2 \cdot e^{-y} dy = 0 + 3 \cdot 2 = 6 \\
 E[XY] &= \int_0^{+\infty} \int_0^y xy f_{(X,Y)}(x, y) dx dy = \int_0^{+\infty} \int_0^y xye^{-y} dx dy = \\
 &= \int_0^{+\infty} ye^{-y} \cdot \frac{y^2}{2} dy = \frac{1}{2} \int_0^{+\infty} y^3 e^{-y} dy = \frac{1}{2} \cdot 6 = 3 \\
 \text{Var}[X] &= E[X^2] - E[X]^2 = 2 - 1 = 1 \\
 \text{Var}[Y] &= E[Y^2] - E[Y]^2 = 6 - 4 = 2 \\
 \text{Cov}[X, Y] &= E[XY] - E[X]E[Y] = 3 - 2 \cdot 1 = 1
 \end{aligned}$$

Por tanto, podemos calcular el coeficiente de determinación lineal. Además, como las rectas de regresión coinciden con las curvas de regresión, el coeficiente de determinación lineal coincide con las razones de correlación:

$$\eta_{X/Y}^2 = \eta_{Y/X}^2 = \rho_{X,Y}^2 = \frac{\text{Cov}[X, Y]^2}{\text{Var}[X] \text{Var}[Y]} = \frac{1^2}{1 \cdot 2} = \frac{1}{2}$$

Como podemos ver, tenemos que la mitad de la varianza de  $X$  y  $Y$  se explica por la regresión, lo que nos indica que los ajustes no son muy buenos.

Calculamos ahora el coeficiente de correlación lineal. Como la covarianza es positiva, el coeficiente de correlación lineal también lo será:

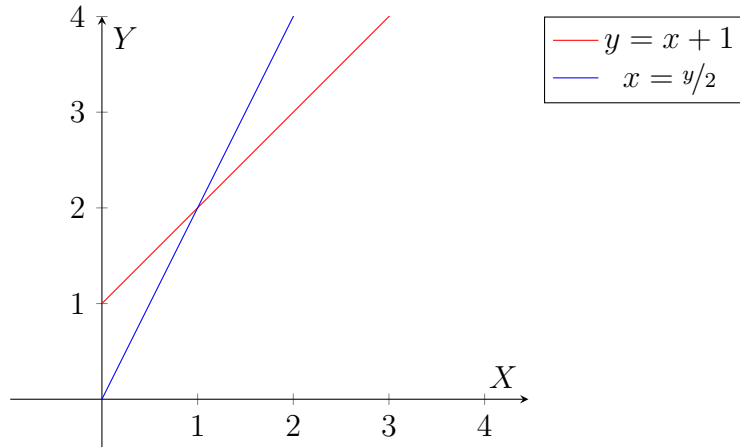
$$\rho_{X,Y} = +\sqrt{\rho_{X,Y}^2} = +\sqrt{\frac{1}{2}} = +\frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$$

Como podemos ver, el coeficiente de correlación lineal es positivo, lo que nos indica que las variables están correlacionadas positivamente.

Calculamos ahora el error cuadrático medio cometido al predecir cada variable según cada una de las funciones de regresión.

$$\begin{aligned}
 \text{ECM}(\hat{Y}(X)) &= E[(Y - \hat{Y}(X))^2] = E[(Y - X - 1)^2] = E[Y^2 + X^2 + 1 - 2Y + 2X - 2YX] = \\
 &= E[Y^2] + E[X^2] + 1 - 2E[Y] + 2E[X] - 2E[XY] = 6 + 2 + 1 - 4 + 2 - 6 = 1 \\
 \text{ECM}(\hat{X}(Y)) &= E[(X - \hat{X}(Y))^2] = E[(X - Y/2)^2] = E[X^2 + Y^2/4 - YX] = \\
 &= E[X^2] + \frac{1}{4}E[Y^2] - E[XY] = 2 + \frac{1}{4} \cdot 6 - 3 = \frac{1}{2}
 \end{aligned}$$

Por último, representamos las rectas de regresión:



**Ejercicio 2.1.11.** Sea  $(X, Y)$  un vector aleatorio con función de densidad uniforme sobre el cuadrado unidad. Obtener y representar las rectas y curvas de regresión. Calcular el coeficiente de correlación lineal, las razones de correlación y el error cuadrático medio cometido al predecir cada variable según cada una de las funciones de regresión. Interpretar los resultados.

Dado que la función de densidad es uniforme en el cuadrado unidad, tenemos que (compruébese):

$$f_{(X,Y)}(x, y) = \begin{cases} 1, & (x, y) \in [0, 1] \times [0, 1] \\ 0, & \text{en caso contrario} \end{cases}$$

Definimos la siguiente función auxiliar:

$$\begin{aligned} g: \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto \begin{cases} 1, & t \in [0, 1] \\ 0, & \text{en caso contrario} \end{cases} \end{aligned}$$

Tenemos que  $f_{(X,Y)}(x, y) = g(x)g(y)$  para todo  $(x, y) \in \mathbb{R}^2$ , por lo que  $X$  e  $Y$  son independientes. Por tanto, tenemos que las curvas de regresión son:

$$\begin{aligned} \hat{Y}(x) &= E[Y \mid X = x] = E[Y] = \int_0^1 y \, dy = \left[ \frac{y^2}{2} \right]_0^1 = \frac{1}{2} \quad \forall x \in [0, 1] \\ \hat{X}(y) &= E[X \mid Y = y] = E[X] = \int_0^1 x \, dx = \left[ \frac{x^2}{2} \right]_0^1 = \frac{1}{2} \quad \forall y \in [0, 1] \end{aligned}$$

Al ser independientes, en particular son incorreladas, por lo que:

$$\rho_{X,Y}^2 = \eta_{X/Y}^2 = \eta_{Y/X}^2 = 0$$

Calculamos ahora el error cuadrático medio cometido al predecir cada variable según cada una de las funciones de regresión.

$$\begin{aligned} \text{ECM}(\hat{Y}(X)) &= E[(Y - \hat{Y}(X))^2] = E[(Y - E[Y])^2] = \text{Var}[Y] \\ \text{ECM}(\hat{X}(Y)) &= E[(X - \hat{X}(Y))^2] = E[(X - E[X])^2] = \text{Var}[X] \end{aligned}$$

Calculamos por tanto ambas varianzas:

$$\begin{aligned} E[X^2] &= \int_0^1 x^2 dx = \left[ \frac{x^3}{3} \right]_0^1 = \frac{1}{3} \\ E[Y^2] &= \int_0^1 y^2 dy = \left[ \frac{y^3}{3} \right]_0^1 = \frac{1}{3} \\ \text{Var}[X] &= E[X^2] - E[X]^2 = \frac{1}{3} - \left( \frac{1}{2} \right)^2 = \frac{1}{12} \\ \text{Var}[Y] &= E[Y^2] - E[Y]^2 = \frac{1}{3} - \left( \frac{1}{2} \right)^2 = \frac{1}{12} \end{aligned}$$

Por tanto, tenemos que:

$$\begin{aligned} \text{ECM}(\hat{Y}(X)) &= \text{Var}[Y] = \frac{1}{12} \\ \text{ECM}(\hat{X}(Y)) &= \text{Var}[X] = \frac{1}{12} \end{aligned}$$

**Ejercicio 2.1.12.** Supongamos que  $(X, Y)$  tiene función de densidad de probabilidad conjunta dada por:

$$f_{(X,Y)}(x, y) = \begin{cases} 1, & |y| < x, x \in ]0, 1[ \\ 0, & \text{en otro caso} \end{cases}$$

Obtener y representar las rectas y curvas de regresión. Calcular el coeficiente de correlación lineal, las razones de correlación y el error cuadrático medio cometido al predecir cada variable según cada una de las funciones de regresión. Interpretar los resultados.

Calculamos las marginales:

$$\begin{aligned} f_X(x) &= \int_{-x}^x 1 dy = 2x \quad \forall x \in ]0, 1[ \\ f_Y(y) &= \int_{|y|}^1 1 dx = 1 - |y| \quad \forall y \in ]-1, 1[ \end{aligned}$$

Calculamos ahora las condicionadas. Dados  $x^* \in ]0, 1[$  e  $y^* \in ]-1, 1[$ , tenemos que:

$$\begin{aligned} f_{Y|X=x^*}(y) &= \frac{f_{(X,Y)}(x^*, y)}{f_X(x^*)} = \frac{1}{2x^*} \quad \forall y \in ]-x^*, x^*[ \\ f_{X|Y=y^*}(x) &= \frac{f_{(X,Y)}(x, y^*)}{f_Y(y^*)} = \frac{1}{1 - |y^*|} \quad \forall x \in ]|y^*|, 1[ \end{aligned}$$

Calculamos ahora las curvas de regresión. Tenemos que:

$$\begin{aligned} \hat{Y}(x) &= E[Y | X = x] = \int_{-x}^x y \cdot \frac{1}{2x} dy = \frac{1}{2x} \int_{-x}^x y dy = \frac{1}{2x} \left( \frac{x^2}{2} - \frac{x^2}{2} \right) = 0 \quad \forall x \in ]0, 1[ \\ \hat{X}(y) &= E[X | Y = y] = \int_{|y|}^1 x \cdot \frac{1}{1 - |y|} dx = \frac{1}{1 - |y|} \int_{|y|}^1 x dx = \frac{1}{1 - |y|} \left( \frac{1}{2} - \frac{y^2}{2} \right) = \\ &= \frac{1 - y^2}{2(1 - |y|)} = \frac{(1 - y^2)(1 + |y|)}{2(1 - |y|)(1 + |y|)} = \frac{1 + |y|}{2} \quad \forall y \in ]-1, 1[ \end{aligned}$$

Como la curva de regresión de  $Y$  sobre  $X$  es constante, la recta de regresión de  $Y$  sobre  $X$  es la misma que la curva de regresión. Calculamos ahora la recta de regresión  $X$  sobre  $Y$ . Para ello, previamente calculamos:

$$E[X] = \int_0^1 x \cdot 2x \, dx = 2 \int_0^1 x^2 \, dx = 2 \left[ \frac{x^3}{3} \right]_0^1 = \frac{2}{3}$$

$$\begin{aligned} E[Y] &= \int_{-1}^1 y \cdot (1 - |y|) \, dy = \int_{-1}^0 y(1 + y) \, dy + \int_0^1 y(1 - y) \, dy = \\ &= \left[ \frac{y^2}{2} + \frac{y^3}{3} \right]_{-1}^0 + \left[ \frac{y^2}{2} - \frac{y^3}{3} \right]_0^1 = -\frac{1}{2} + \frac{1}{3} + \frac{1}{2} - \frac{1}{3} = 0 \end{aligned}$$

$$E[X^2] = \int_0^1 x^2 \cdot 2x \, dx = 2 \int_0^1 x^3 \, dx = 2 \left[ \frac{x^4}{4} \right]_0^1 = \frac{1}{2}$$

$$\begin{aligned} E[Y^2] &= \int_{-1}^1 y^2 \cdot (1 - |y|) \, dy = \int_{-1}^0 y^2(1 + y) \, dy + \int_0^1 y^2(1 - y) \, dy = \\ &= \left[ \frac{y^3}{3} + \frac{y^4}{4} \right]_{-1}^0 + \left[ \frac{y^3}{3} - \frac{y^4}{4} \right]_0^1 = \frac{1}{3} - \frac{1}{4} + \frac{1}{3} - \frac{1}{4} = \frac{1}{6} \end{aligned}$$

$$E[XY] = \int_0^1 \int_{-x}^x xy \, dy \, dx = \int_0^1 x \left[ \frac{y^2}{2} \right]_{-x}^x \, dx = \int_0^1 x \left( \frac{x^2}{2} - \frac{x^2}{2} \right) \, dx = 0$$

$$\text{Var}[X] = E[X^2] - E[X]^2 = \frac{1}{2} - \left( \frac{2}{3} \right)^2 = \frac{1}{18}$$

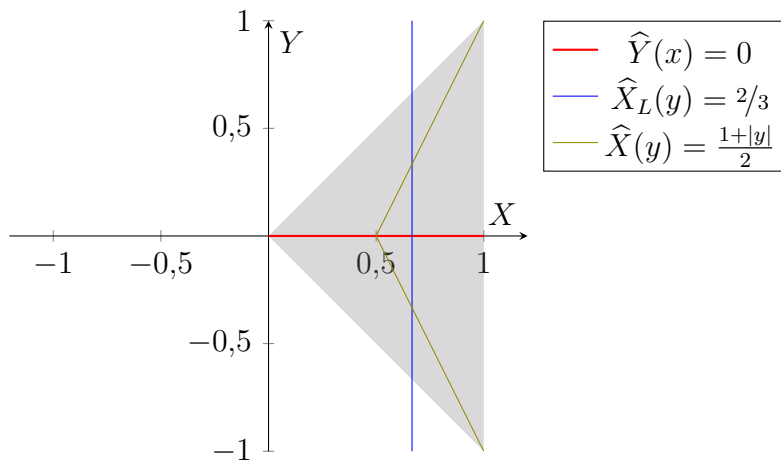
$$\text{Var}[Y] = E[Y^2] - E[Y]^2 = \frac{1}{6} - 0 = \frac{1}{6}$$

$$\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = 0$$

Por tanto, tenemos que la recta de regresión de  $X$  sobre  $Y$  es:

$$\hat{X}_L(y) = E[X] + \frac{\text{Cov}[X, Y]}{\text{Var}[Y]} (y - E[Y]) = \frac{2}{3} \quad \forall y \in ]-1, 1[$$

La representación gráfica de las rectas y curvas de regresión es:



El coeficiente de determinación lineal es:

$$\rho_{X,Y}^2 = \frac{\text{Cov}[X, Y]^2}{\text{Var}[X] \text{Var}[Y]} = 0$$

Como vemos, el ajuste lineal es nulo, ya que no explica nada de la variabilidad de  $X$  ni de  $Y$ .

El coeficiente de correlación lineal es:

$$\rho_{X,Y} = 0$$

Como vemos, las variables son linealmente incorreladas, como era de esperar tras el cálculo del coeficiente de determinación lineal.

Calculamos las razones de correlación. Como la curva de regresión de  $Y$  sobre  $X$  coincide con la recta de regresión, tenemos que:

$$\eta_{Y/X}^2 = \rho_{X,Y}^2 = 0$$

Respecto de la curva de regresión de  $X$  sobre  $Y$ , tenemos que:

$$\eta_{X/Y}^2 = \frac{\text{Var}[\hat{X}(Y)]}{\text{Var}[X]} = \frac{\text{Var}\left[\frac{1+|Y|}{2}\right]}{\text{Var}[X]}$$

Calculemos la varianza del numerador:

$$\begin{aligned} E[|Y|] &= \int_{-1}^1 |y| \cdot (1 - |y|) dy = \int_{-1}^0 -y(1 + y) dy + \int_0^1 y(1 - y) dy = \\ &= -\left[\frac{y^2}{2} + \frac{y^3}{3}\right]_{-1}^0 + \left[\frac{y^2}{2} - \frac{y^3}{3}\right]_0^1 = \frac{1}{2} - \frac{1}{3} + \frac{1}{2} - \frac{1}{3} = \frac{1}{3} \\ E\left[\frac{1+|Y|}{2}\right] &= \frac{1}{2} + \frac{1}{2}E[|Y|] = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2} + \frac{1}{6} = \frac{2}{3} \\ E\left[\left(\frac{1+|Y|}{2}\right)^2\right] &= \frac{1}{4} [1 + 2E[|Y|] + E[Y^2]] = \frac{1}{4} \left[1 + 2 \cdot \frac{1}{3} + \frac{1}{6}\right] = \frac{11}{24} \\ \text{Var}\left[\frac{1+|Y|}{2}\right] &= E\left[\left(\frac{1+|Y|}{2}\right)^2\right] - E\left[\frac{1+|Y|}{2}\right]^2 = \frac{11}{24} - \left(\frac{2}{3}\right)^2 = \frac{1}{72} \end{aligned}$$

Por tanto, tenemos que:

$$\eta_{X/Y}^2 = \frac{\text{Var}\left[\frac{1+|Y|}{2}\right]}{\text{Var}[X]} = \frac{1/72}{1/18} = \frac{1}{4}$$

En este caso, tan solo se explica un cuarto de la variabilidad de  $X$  por la regresión, por lo que el ajuste sigue siendo malo.

Calculamos ahora el error cuadrático medio de cada una de las funciones de

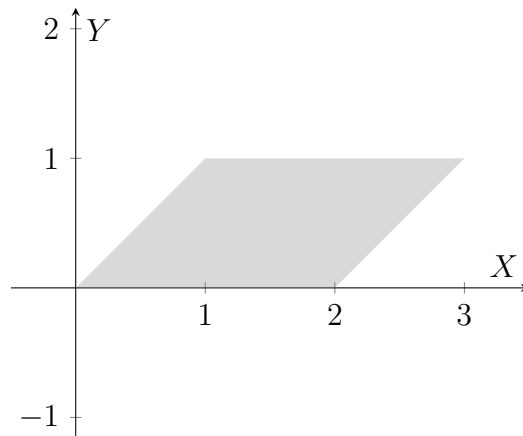
regresión:

$$\begin{aligned}
 \text{ECM}(\hat{Y}(X)) &= E[(Y - \hat{Y}(X))^2] = E[Y^2] = \frac{1}{6} \\
 \text{ECM}(\hat{X}(Y)) &= E[(X - \hat{X}(Y))^2] = E\left[\left(X - \frac{1+|Y|}{2}\right)^2\right] = \\
 &= E[X^2] + E\left[\left(\frac{1+|Y|}{2}\right)^2\right] - E[X(1+|Y|)] = \\
 &= E[X^2] + E\left[\left(\frac{1+|Y|}{2}\right)^2\right] - E[X] - E[|Y| \cdot X] = \\
 &= \frac{1}{2} + \frac{11}{24} - \frac{2}{3} - \frac{1}{4} = \frac{1}{24} \\
 E[X \cdot |Y|] &= \int_0^1 \int_{-x}^x x|y| \cdot 1 \, dy \, dx = \int_0^1 2x \int_0^x y \, dy \, dx = \\
 &= \int_0^1 2x \cdot \frac{x^2}{2} \, dx = \int_0^1 x^3 \, dx = \left[\frac{x^4}{4}\right]_0^1 = \frac{1}{4} \\
 \text{ECM}(\hat{X}_L(Y)) &= E[(X - E[X])^2] = \text{Var}[X] = \frac{1}{18}
 \end{aligned}$$

Aquí podemos ver que el error cuadrático medio cometido con la regresión lineal es mayor que el cometido con la curva de regresión, lo que indica que la regresión lineal es peor que la curva de regresión (como es de esperar).

**Ejercicio 2.1.13.** Sea  $(X, Y)$  un vector aleatorio distribuido uniformemente en el paralelogramo de vértices  $(0, 0)$ ;  $(2, 0)$ ;  $(3, 1)$  y  $(1, 1)$ . Calcular el error cuadrático medio asociado a la predicción de  $X$  a partir de la variable  $Y$  y a la predicción de  $Y$  a partir de la variable aleatoria  $X$ . Determinar la predicción más fiable a la vista de los resultados obtenidos.

Veamos gráficamente el paralelogramo:



Notemos por  $P$  al paralelogramo. Tenemos que la función de densidad conjunta es:

$$f_{(X,Y)}(x,y) = \begin{cases} k, & (x,y) \in P \\ 0, & \text{en otro caso} \end{cases}$$

Para calcular  $k$ , tenemos que:

$$1 = \int_P f_{(X,Y)} = \int_P k = k\lambda(P) = k \cdot 2 \cdot 1 = 2k \implies k = \frac{1}{2}$$

Calculamos en primer lugar las marginales. Para  $x \in [0, 3]$ , tenemos que:

- Si  $x \in [0, 1]$ , tenemos que:

$$f_X(x) = \int_0^x \frac{1}{2} dy = \frac{x}{2}$$

- Si  $x \in [1, 2]$ , tenemos que:

$$f_X(x) = \int_0^1 \frac{1}{2} dy = \frac{1}{2}$$

- Si  $x \in [2, 3]$ , tenemos que:

$$f_X(x) = \int_{x-2}^1 \frac{1}{2} dy = \frac{1}{2}(1 - x + 2) = \frac{3-x}{2}$$

Por tanto, tenemos que:

$$f_X(x) = \begin{cases} x/2, & x \in [0, 1] \\ 1/2, & x \in [1, 2] \\ \frac{3-x}{2}, & x \in [2, 3] \\ 0, & \text{en otro caso} \end{cases}$$

Para  $y \in [0, 1]$ , tenemos que:

$$f_Y(y) = \int_y^{y+2} \frac{1}{2} dx = \frac{1}{2}(y + 2 - y) = 1$$

Calculamos ahora las condicionadas. Dado  $y^* \in [0, 1]$ , tenemos que:

$$f_{X|Y=y^*}(x) = \frac{f_{(X,Y)}(x, y^*)}{f_Y(y^*)} = \frac{1/2}{1} = \frac{1}{2} \quad \forall x \in [0, 3]$$

Respecto a la condicionada de  $Y$  dado  $X = x^*$ , distinguimos en función de  $x^*$ :

$$f_{Y|X=x^*}(y) = \frac{f_{(X,Y)}(x^*, y)}{f_X(x^*)} = \begin{cases} 1/x^*, & x^* \in [0, 1], y \in [0, x^*] \\ 1/2, & x^* \in [1, 2], y \in [0, 1] \\ \frac{1}{3-x^*}, & x^* \in [2, 3], y \in [x^* - 2, 1] \end{cases}$$

Tenemos que calcular los siguientes errores cuadráticos medios:

$$\text{ECM}(\hat{X}(Y)) = E[\text{Var}[X | Y]]$$

$$\text{ECM}(\hat{Y}(X)) = E[\text{Var}[Y | X]]$$

Para calcular  $E[\text{Var}[X | Y]]$ , tenemos que:

$$\begin{aligned} E[X | Y = y] &= \int_y^{y+2} x \cdot \frac{1}{2} dx = \frac{1}{2} \left[ \frac{x^2}{2} \right]_y^{y+2} = \frac{1}{2} \left( \frac{(y+2)^2}{2} - \frac{y^2}{2} \right) = \frac{4y+4}{4} = y+1 \\ E[X^2 | Y = y] &= \int_y^{y+2} x^2 \cdot \frac{1}{2} dx = \frac{1}{2} \left[ \frac{x^3}{3} \right]_y^{y+2} = \frac{1}{2} \left( \frac{(y+2)^3}{3} - \frac{y^3}{3} \right) = \frac{6y^2+12y+8}{6} = \\ &= y^2 + 2y + 4/3 \\ \text{Var}[X | Y = y] &= E[X^2 | Y = y] - E[X | Y = y]^2 = y^2 + 2y + \frac{4}{3} - (y+1)^2 = \\ &= y^2 + 2y + 4/3 - y^2 - 2y - 1 = 1/3 \\ E[\text{Var}[X | Y]] &= E \left[ \frac{1}{3} \right] = \frac{1}{3} \end{aligned}$$

Para calcular  $E[\text{Var}[Y | X]]$ , tenemos que:

$$\begin{aligned} E[\text{Var}[Y | X]] &= E[E[Y^2 | X] - E[Y | X]^2] = E[E[Y^2 | X]] - E[E[Y | X]^2] = \\ &= E[Y^2] - E[E[Y | X]^2] \end{aligned}$$

Calculamos lo necesario:

**Ejercicio 2.1.14.** Sea  $(X, Y)$  un vector aleatorio con rectas de regresión

$$x + 4y = 1 \quad x + 5y = 2$$

1. ¿Cuál es la recta de regresión de  $Y$  sobre  $X$ ?

Suponemos que la recta de regresión de  $Y$  sobre  $X$  es  $x + 4y = 1$ . Por tanto, tenemos que:

$$y = \hat{Y}(x) = \frac{1}{4} - \frac{x}{4} = E[Y] + \frac{\text{Cov}[X, Y]}{\text{Var}[X]} (x - E[X]) \implies \frac{\text{Cov}[X, Y]}{\text{Var}[X]} = -\frac{1}{4}$$

Por otro lado, tenemos que la recta de regresión de  $X$  sobre  $Y$  es  $x + 5y = 2$ . Por tanto, tenemos que:

$$x = \hat{X}(y) = 2 - 5y = E[X] + \frac{\text{Cov}[X, Y]}{\text{Var}[Y]} (y - E[Y]) \implies \frac{\text{Cov}[X, Y]}{\text{Var}[Y]} = -5$$

Por tanto, el coeficiente de determinación lineal es:

$$\rho_{X,Y}^2 = \frac{\text{Cov}[X, Y]^2}{\text{Var}[X] \text{Var}[Y]} = -\frac{1}{4} \cdot (-5) = \frac{5}{4} > 1$$

Esto es un absurdo, luego nuestra suposición era errónea. La recta de regresión de  $Y$  sobre  $X$  es  $x + 5y = 2$  y la de  $X$  sobre  $Y$  es  $x + 4y = 1$ .

2. Calcular el coeficiente de correlación lineal y la proporción de varianza de cada variable que queda explicada por la regresión lineal.



Como la recta de regresión de  $Y$  sobre  $X$  es  $x + 5y = 2$ , tenemos que:

$$y = \hat{Y}(x) = \frac{2-x}{5} = E[Y] + \frac{\text{Cov}[X, Y]}{\text{Var}[X]} (x - E[X]) \implies \frac{\text{Cov}[X, Y]}{\text{Var}[X]} = -\frac{1}{5}$$

Por otro lado, como la recta de regresión de  $X$  sobre  $Y$  es  $x + 4y = 1$ , tenemos que:

$$x = \hat{X}(y) = 1 - 4y = E[X] + \frac{\text{Cov}[X, Y]}{\text{Var}[Y]} (y - E[Y]) \implies \frac{\text{Cov}[X, Y]}{\text{Var}[Y]} = -4$$

Por tanto, el coeficiente de determinación lineal es:

$$\rho_{X,Y}^2 = \frac{\text{Cov}[X, Y]^2}{\text{Var}[X] \text{Var}[Y]} = -\frac{1}{5} \cdot (-4) = \frac{4}{5} = 0,8$$

Por tanto, la proporción de varianza de cada variable que queda explicada por la regresión lineal es un 80 %. Además, como la covarianza es negativa, el coeficiente de correlación lineal es:

$$\rho_{X,Y} = -\sqrt{\rho_{X,Y}^2} = -\sqrt{\frac{4}{5}} = -\frac{2\sqrt{5}}{5}$$

### 3. Calcular las medias de ambas variables.

De la forma de las rectas de regresión, planteamos el siguiente sistema:

$$\begin{cases} E[Y] + 1/5 E[X] = 2/5 \\ E[X] + 4E[Y] = 1 \end{cases}$$

Resolviendo el sistema, llegamos a que:

$$E[X] = -3 \quad E[Y] = 1$$