Material introductorio

Nombre del curso: Teoría Moderna de la Detección y Estimación

Autores: Vanessa Gómez Verdejo

Universidad
Carlos III de Madrid

Índice general

1

Variables aleatorias unidimensionales

1.1. Introducción

Cuando se realiza un experimento aleatorio, con frecuencia, nos interesa más conocer alguna función de dicho resultado, mejor que el resultado del experimento.

EJEMPLO 1 Se lanza una moneda tres veces consecutivas e interesa el número de caras obtenidas en el experimento aleatorio. Denotando con los símbolos H y T al hecho de obtener, respectivamente, una cara o una cruz, se tiene el siguiente conjunto de posibles sucesos del espacio muestral (Ω):

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

La variable aleatoria X es la función que cuenta el número de caras obtenidas en dicho experimento aleatorio:

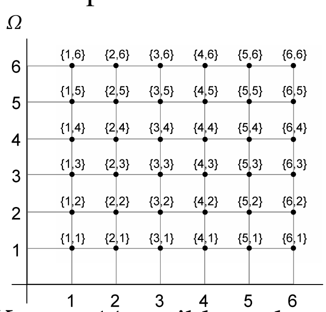
$$\begin{array}{cccc} X(\{HHH\}) = 3 & X(\{HHT\}) = 2 & X(\{HTH\}) = 2 & X(\{THH\}) = 2 \\ X(\{HTT\}) = 1 & X(\{THT\}) = 1 & X(\{TTH\}) = 1 & X(\{TTT\}) = 0 \end{array}$$

tal y como puede verse la v.a. X toma 4 posibles valores $\{0, 1, 2, 3\}$.

EJEMPLO 2 Se lanza un dado dos veces consecutivas y se pueden definir las v.a.:

- X: la suma de las caras obtenidas en el experimento aleatorio.
- Y: la diferencia (en valor absoluto) de las caras obtenidas en el experimento aleatorio.

Espacio muestral



X: v.a. suma $\quad Y$: v.a. diferencia

$$\begin{array}{ll} X(\{1,1\}) = 2 & Y(\{1,1\}) = 0 \\ X(\{1,2\}) = 3 & Y(\{1,2\}) = 1 \\ \vdots & \vdots \\ X(\{6,5\}) = 11 & Y(\{6,5\}) = 1 \\ X(\{6,6\}) = 12 & Y(\{6,6\}) = 0 \end{array}$$

La v.a. X toma 11 posibles valores $\{2, 3, 4, \ldots, 12\}$, mientras que Y toma 6 valores $\{0, 1, 2, 3, 4, 5\}$.

Formalmente llamaremos **variable aleatoria** a la función real que a cada suceso elemental del espacio muestral le asigna un valor numérico

$$\begin{aligned} X : \Omega &\longrightarrow \Re \\ \omega &\longrightarrow X(\omega) \end{aligned}$$

tal que $\forall a \in \Re \quad \{\omega : X(\omega) \leq a\}$, es un suceso del espacio muestral (subconjunto del espacio muestral Ω). En otras palabras, la imagen inversa de $(-\infty, a]$ es un suceso del espacio muestral Ω.

Distinguimos dos tipos de v.a.:

1. **Discretas**: toman un conjunto finito o infinito numerable de valores. Dicho conjunto de valores se denomina *recorrido* de la v.a.
   - Al contar el número de caras obtenidas al lanzar una moneda tres veces consecutivas, el recorrido de la función es el conjunto finito de valores $\{0, 1, 2, 3\}$.
   - Se lanza una moneda hasta obtener una cara y sea X la v.a. que cuenta el número de lanzamientos realizados. El recorrido de esta función es $\{1, 2, 3, \ldots\}$ un conjunto infinito, pero numerable, de valores.
2. **Continuas**: toman un conjunto infinito y no numerable de valores. En este caso, a este conjunto de valores se le denomina *soporte* de la función.
   - Un autobús pasa cada media hora y un viajero llega en un momento dado a la parada. X es la v.a. que cuenta el tiempo de espera desde que el viajero llega a la parada hasta que llega el autobús. El soporte de la función es el intervalo $(0, 30)$ minutos.

1.2. Distribución de probabilidad de una variable aleatoria

Ya que el valor de una v.a. está determinado por el resultado del experimento, podemos asignar probabilidades a los valores de la v.a. Cuando las v.a. son discretas, a la función que mide estas probabilidades se la llama **función de masa de probabilidad** (f.m.p) y la denotaremos $P_X(x)$ o, más breve, como $P(x)$. Esta función verifica que

$$P_X(x) = P\{X = x\}$$

Si la v.a. es continua se la designa como **función de densidad de probabilidad** (f.d.p.) y también se denotará como $p_X(x)$ ó, simplemente, $p(x)$.

Nota: En el caso continuo la probabilidad de un punto, $P\{X = x\}$, es nula (a pesar de que el suceso no es imposible); y la probabilidad de que $P\{X \neq x\} = 1$, a pesar de que no es seguro.

Se puede definir la f.d.p. de una v.a. aleatoria discreta haciendo uso de la función delta de dirac y se tiene que

$$p_X(x) = \sum_k P_X(x_k)\delta(x - x_k) = \sum_i P\{X = x_k\}\delta(x - x_k)$$

siendo $\{x_k\}$ el conjunto de puntos que forman el recorrido de la v.a. X.

Toda f.m.p o f.d.p. satisface las siguientes propiedades:

1.

$$\text{discreta} \quad P(x) \geq 0$$
$$\text{continua} \quad p(x) \geq 0$$

2. Se verifica

$$\text{discreta} \quad \sum P(x) = 1$$
$$\text{continua} \quad \int_{-\infty}^{+\infty} p(x)\, dx = 1$$

3. Se tiene que $\forall a, b \in \Re$ la probabilidad de que el valor de la v.a. X se encuentre entre a y b viene dada por:

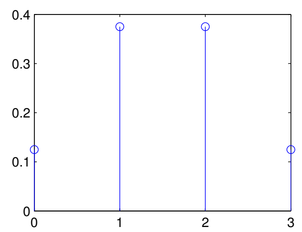
$$\text{discreta} \quad P\{a \leq X \leq b\} = \sum_{k=a}^{b} P(x = k) = \sum_{k=a}^{b} P\{X = k\}$$
$$\text{continua} \quad P\{a \leq X \leq b\} = \int_a^b p(x)\, dx$$

donde, eventualmente, a y b pueden ser, respectivamente, $-\infty$ y $+\infty$.

EJEMPLO 1 (cont.) Siendo X la v.a. resultante de contar el número de caras obtenidas al lanzar una moneda tres veces consecutivas, tenemos la siguiente f.m.p,

$$P\{X = 3\} = P_X(x = 3) = 1/8 = 0{,}125$$
$$P\{X = 2\} = P_X(x = 2) = 3/8 = 0{,}375$$
$$P\{X = 1\} = P_X(x = 1) = 3/8 = 0{,}375$$
$$P\{X = 0\} = P_X(x = 0) = 1/8 = 0{,}125$$

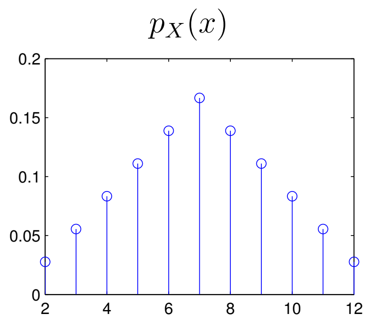
x	0	1	2	3
$p_X(x)$	1/8	3/8	3/8	1/8

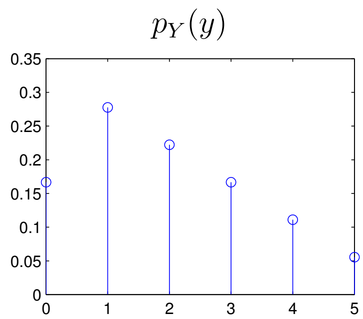


EJEMPLO 2 (cont.) Se lanza un dado dos veces consecutivas y son X e Y las v.a. que cuentan, respectivamente, la suma y la diferencia de las caras. Sus f.m.p son:

x	2	3	4	5	6	7	8	9	10	11	12
$p_X(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

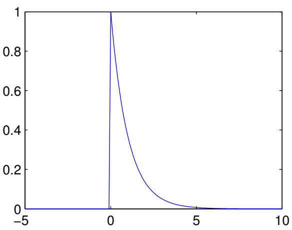
y	0	1	2	3	4	5
$p_Y(y)$	6/36	10/36	8/36	6/36	4/36	2/36





EJEMPLO 3 Consideremos que la duración de una bombilla viene dada por una v.a. con distribución exponencial,

$$p(x) = \begin{cases} \exp(-x) & x > 0 \\ 0, & \text{en otro caso} \end{cases}$$



Comprobemos que es una f.d.p.

$$\int_{-\infty}^{+\infty} p(x)\, dx = \int_{0}^{+\infty} \exp(-x)\ dx = -\exp(-x)\Big|_0^{+\infty} = 1$$

Calculemos la probabilidad de que la bombilla dure más de un año

$$P\{1 \leq X\} = \int_{1}^{+\infty} \exp(-x)\ dx = -\exp(-x)\Big|_1^{+\infty} = \exp(-1) - 0 = \frac{1}{e}$$

A partir de la f.d.p. o f.m.p. se puede definir formalmente el *soporte* o *recorrido* de una v.a. como el conjunto de valores de $x \in \Re$ donde la variable aleatoria X tiene una probabilidad mayor que cero, es decir, donde la distribución tiene masa. Se recuerda, que el término recorrido se emplea con v.a. discretas mientras que el término soporte es más adecuado para v.a. continuas.

1.3. Función de distribución

La función de distribución (f.d.) indica para cada valor de $x \in \Re$ la probabilidad de que la v.a. X sea menor que dicho valor x. Si denotamos con $F_X(x)$ o, directamente, $F(x)$ a la función de distribución de la v.a. X, ésta se define como:

$$F_X(x) = P\{X \leq x\}; \quad \forall x \in \Re$$

La f.d. es una función acumulativa, ya que calcula el área de la f.d.p $p_X(x)$ hasta el punto x,

$$\text{discreta } F_X(x) = P\{X \leq x\} = \sum_{k \leq x} P(x = k) = \sum_{k \leq x} P\{X = k\}$$

$$\text{continua } F_X(x) = P\{X \leq x\} = \int_{-\infty}^{x} p_X(t)\, dt$$

Propiedad: En v.a. discretas su f.d.p. es escalonada, mientras que en las v.a. continuas su f.d.p. es continua.

Dada la f.d $F_X(x)$ se puede calcular la f.d.p.

$$p_X(x) = \frac{dF_X(x)}{dx} = F'(x)$$

La función de distribución de una v.a. X debe verificar las siguientes propiedades:

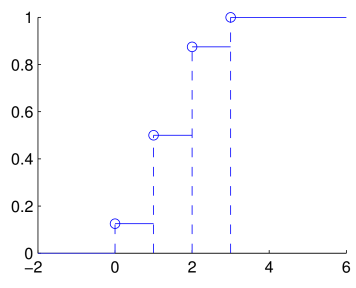
1. $\lim_{x \to -\infty} F(x) = F(-\infty) = 0$ y $\lim_{x \to +\infty} F(x) = F(+\infty) = 1$.
2. $F(x)$ es monótona no decreciente, es decir, si $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$.
3. Continua a la derecha de cada punto, es decir,

$$\lim_{x \to a^+} F(x) = F(a)$$

Y, recíprocamente, toda función que verifique las propiedades anteriores es una f. d.

EJEMPLO 1 (cont.) Obtenga la función de distribución de la v.a. que cuenta el número de caras obtenidas al lanzar una moneda 3 veces consecutivas.

$$\begin{aligned}
\text{Si } x < 0 &\Rightarrow F(x) = P\{X < x\} = 0 \\
\text{Si } 0 \leq x < 1 &\Rightarrow F(x) = P\{X < x\} = P(x = 0) = 1/8 = 0{,}375 \\
\text{Si } 1 \leq x < 2 &\Rightarrow F(x) = P\{X < x\} = P(x = 0) + P(x = 1) = 4/8 = 0{,}5 \\
\text{Si } 2 \leq x < 3 &\Rightarrow F(x) = P\{X < x\} = P(x = 0) + P(x = 1) + P(x = 2) = \\
& \qquad 7/8 = 0{,}625 \\
\text{Si } 3 \leq x &\Rightarrow F(x) = P\{X < x\} = P(x = 0) + P(x = 1) + P(x = 2) + \\
& \qquad +P(x = 3) = 8/8 = 1
\end{aligned}$$



EJEMPLO 3 (cont.) Se tiene una f.d.p de tipo exponencial,

$$p(x) = \begin{cases} \exp(-x) & x > 0 \\ 0, & \text{en otro caso} \end{cases}$$

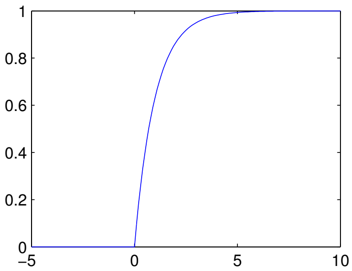
obtenga y represente la función de distribución.

$$\text{Si } x \leq 0 \Rightarrow F(x) = P\{X \leq x\} = \int_{-\infty}^{x} p_X(t)\, dt = \int_{-\infty}^{x} 0\, dt = 0$$

$$\text{Si } x > 0 \Rightarrow F(x) = P\{X \leq x\} = \int_{-\infty}^{x} p_X(t)\, dt = \int_{0}^{x} \exp(-t)\; dt =$$

$$= -\exp(-t)\Big|_0^x = 1 - \exp(-x)$$

$$F(x) = \begin{cases} 1 - \exp(-x) & x > 0 \\ 0, & x \leq 0 \end{cases}$$



1.4. Cuantiles

Llamaremos cuantiles a los valores de $x \in \Re$ que dividen la función de densidad de probabilidad en regiones con el mismo área. Los cuantiles más empleados son:

- La **mediana**: es el valor de $x \in \Re$ que divide la f.d.p. en dos intervalos de modo que el área de la f.d.p. en cada intervalo vale $1/2$, por lo que, la función de distribución en ese punto vale $1/2$.
- Los **cuartiles**: son tres cuantiles que dividen la f.d.p. en 4 regiones de modo que el área de la f.d.p. en cada intervalo vale $1/4$. En este caso la función de distribución en los cuartiles toma los valores $1/4$, $1/2$ y $3/4$.
- Los **deciles**: son 9 cuantiles que dividen la f.d.p. en 10 regiones de modo que el área de la f.d.p. en cada intervalo vale $1/10$. En este caso la función de distribución en los cuartiles toma los valores $1/10, 2/10, \ldots, 9/10$.
- Los **centiles**: son 99 cuantiles que dividen la f.d.p. en 100 regiones con área de $1/100$. En este caso la función de distribución toma los valores $1/100$, $2/100$, …, $99/100$.

EJEMPLO 3 (cont.) Obtenga la mediana y los cuartiles de la distribución exponencial.

A partir de la función de distribución:

$$F(x) = \begin{cases} 1 - \exp(-x) & x > 0 \\ 0, & x \leq 0 \end{cases}$$

$$
\begin{aligned}
F(x_{\text{MEDIANA}}) = \frac{1}{2} &\Rightarrow 1 - \exp\left(-x_{\text{MEDIANA}}\right) = \tfrac{1}{2} \Rightarrow x_{\text{MEDIANA}} = \log 2 \\
F(x_{\text{Q1}}) = \frac{1}{4} &\Rightarrow \quad 1 - \exp\left(-x_{\text{Q1}}\right) = \tfrac{1}{4} \Rightarrow \quad x_{\text{Q1}} = \log 4 \\
F(x_{\text{Q2}}) = \frac{1}{2} &\Rightarrow \quad 1 - \exp\left(-x_{\text{Q2}}\right) = \tfrac{1}{2} \Rightarrow \quad x_{\text{Q2}} = \log 2 \\
F(x_{\text{Q3}}) = \frac{3}{4} &\Rightarrow \quad 1 - \exp\left(-x_{\text{Q3}}\right) = \tfrac{3}{4} \Rightarrow \quad x_{\text{Q3}} = \log \frac{4}{3}
\end{aligned}
$$

1.5. Variables aleatorias idénticamente distribuidas

Dos v.a. X, Y están idénticamente distribuidas (i.d.) sii tienen la misma distribución.

EJEMPLO Se lanza un dado dos veces consecutivas:

$$
\text{Sea } X = \begin{cases} 1, & \text{el resultado del primer dado es } 1 \text{ ó } 3 \\ 2, & \text{si el resultado del primer dado es } 2, 4, 5 \text{ ó } 6 \end{cases}
$$

$$
\text{Sea } Y = \begin{cases} 1, & \text{el resultado del segundo dado es } 2 \text{ ó } 4 \\ 2, & \text{si el resultado del segundo dado es } 1, 3, 5 \text{ ó } 6 \end{cases}
$$

Claramente,

$$
\begin{aligned}
P\{X = 1\} = 2/6; \, P\{X = 2\} = 4/6 \\
P\{Y = 1\} = 2/6; \, P\{Y = 2\} = 4/6
\end{aligned}
$$

Además, son independientes, por lo que se dice que X e Y son i.i.d.

1.6. Momentos de una distribución

Sea X una v.a. discreta con f.m.p. p(x), se define el **valor esperado, la media o la esperanza matemática** de X como

$$
\mathbb{E}\left\{X\right\} = \sum_k k \, P_X(x = k) = \sum_k k \, P\{X = k\}
$$

Nota: es una media ponderada; si todos los valores son equiprobables, es la media aritmética.

Si X es una v.a. continua con f.d.p. p(x), entonces se define:

$$
\mathbb{E}\left\{X\right\} = \int_{-\infty}^{+\infty} x \, p(x) \, dx
$$

Nota: en este caso es el centro de masa de la distribución.

Análogamente podemos definir los momentos de orden n (respecto al origen)

$$
\mathbb{E}\left\{X^n\right\}; \quad \forall n = 1, 2, \ldots
$$

y los momentos centrales (respecto a la media)

$$\mathbb{E}\left\{(X - \mathbb{E}\{X\})^n\right\}; \quad \forall n = 1, 2, \ldots$$

El momento central de orden 2

$$\mathbb{E}\left\{(X - \mathbb{E}\{X\})^2\right\} = Var\{X\}$$

se denomina **varianza** y es de especial importancia. Por la propia definición de varianza, $Var\{X\} \geq 0$; sólo en el caso en que X sea constante (degenerada) $Var\{X\} = 0$.

Habitualmente se denota con $\sigma^2 = Var\{X\}$ y se define la **desviación típica** de una v.a. como

$$\sigma = \sqrt{Var\{X\}}$$

En general, se puede definir el operador esperanza matemática, $\mathbb{E}\{\cdot\}$, sobre una función $g(\cdot)$ de una v.a. X como:

$$\text{discreta } \mathbb{E}\{g(X)\} = \sum_k g(k)\, P(x = k)$$

$$\text{continua } \mathbb{E}\{g(X)\} = \int_{-\infty}^{+\infty} g(x)\, p(x)\, dx$$

Es un operador lineal, por lo que tiene las siguientes propiedades:

1. $\mathbb{E}\{X + Y\} = \mathbb{E}\{X\} + \mathbb{E}\{Y\}$
2. $\mathbb{E}\{aX\} = a\mathbb{E}\{X\}$
3. $\mathbb{E}\{aX + Y\} = a\mathbb{E}\{X\} + \mathbb{E}\{Y\}$

Mientras que el operador varianza verifica:

1. $Var\{aX\} = a^2 Var\{X\}$
2. $Var\{aX + b\} = a^2 Var\{X\}$

El calculo de la varianza se puede realizar mediante la siguiente expresión:

$$Var\{X\} = \mathbb{E}\left\{(X - \mathbb{E}\{X\})^2\right\} = \mathbb{E}\left\{X^2\right\} - \mathbb{E}\{X\}^2$$

DEMOSTRACIÓN

$$\begin{aligned} Var\{X\} &= \mathbb{E}\left\{(X - \mathbb{E}\{X\})^2\right\} = \mathbb{E}\left\{X^2 - 2X\mathbb{E}\{X\} + \mathbb{E}\{X\}^2\right\} = \\ &= \mathbb{E}\left\{X^2\right\} - 2\mathbb{E}\{X\}\mathbb{E}\{X\} + (\mathbb{E}\{X\})^2 = \mathbb{E}\left\{X^2\right\} - \mathbb{E}\{X\}^2 \end{aligned}$$

EJEMPLO 1 (cont.) Se lanza una moneda 3 veces consecutivas, halle la esperanza matemática y la varianza de la v.a. que cuenta en número de caras obtenidas.

Teníamos que la f.m.p. es:

$$P_X(x = 3) = 1/8 \quad P_X(x = 2) = 3/8$$

$$P_X(x = 1) = 3/8 \quad P_X(x = 0) = 1/8$$

luego la esperanza matemática es

$$\mathbb{E}\{X\} = \sum_k kP(x=k) = 0\frac{1}{8} + 1\frac{3}{8} + 2\frac{3}{8} + 3\frac{1}{8} = \frac{12}{8} = \frac{3}{2}$$

Cuando f.m.p. es simétrica (como es el caso), la media coincide con el punto de simetría de la f.m.p. y, a su vez, con la mediana de X.
La varianza la calculamos con

$$\mathbb{E}\left\{X^2\right\} = \sum_k k^2P(x=k) = 0^2\frac{1}{8} + 1^2\frac{3}{8} + 2^2\frac{3}{8} + 3^2\frac{1}{8} = \frac{24}{8} = 3$$

$$Var\{X\} = \mathbb{E}\left\{X^2\right\} - \mathbb{E}\{X\}^2 = 3 - \left(\frac{3}{2}\right)^2 = \frac{3}{4}$$

EJEMPLO 3 (cont.) Obtenga la media y varianza de una v.a. exponencial,

$$p(x) = \begin{cases} \exp(-x) & x > 0 \\ 0, & \text{en otro caso} \end{cases}$$

$$\begin{aligned} \mathbb{E}\{X\} &= \int_{-\infty}^{+\infty} x\, p(x)\, dx = \int_0^{+\infty} x\ \exp(-x)\ dx = \\ & u = x \quad du = dx \\ & dv = \exp(-x)\ dx \quad v = -\exp(-x) \\ &= -x\exp(-x)\Big|_0^{+\infty} + \int_0^{+\infty} \exp(-x)\ dx = -\exp(-x)\Big|_0^{+\infty} = 1 \end{aligned}$$

Para calcular la varianza,

$$\begin{aligned} \mathbb{E}\left\{X^2\right\} &= \int_{-\infty}^{+\infty} x^2\, p(x)\, dx = \int_0^{+\infty} x^2\ \exp(-x)\ dx = \\ & u = x^2 \quad du = 2x\, dx \\ & dv = \exp(-x)\ dx \quad v = -\exp(-x) \\ &= -x^2\exp(-x)\Big|_0^{+\infty} + \int_0^{+\infty} 2x\ \exp(-x)\ dx = \\ &= 2\int_0^{+\infty} x\ \exp(-x)\ dx = 2 \\ Var\{X\} &= \mathbb{E}\left\{X^2\right\} - \mathbb{E}\{X\}^2 = 2 - 1^2 = 1 \end{aligned}$$

1.7. Cambio de variable aleatoria

1. **Suma de una constante**: consideremos que se tiene una v.a. X de la que se conoce su f.d.p, $p_X(x)$, y se desea calcular la f.d.p. de otra v.a. Y relacionada con la anterior mediante la expresión

$$Y = X + a; \quad a \in \Re$$

para calcular $p_Y(y)$ se recurre a obtener su f.d.

$$F_Y(y) = P\{Y \leq y\} = P\{X + a \leq y\} = P\{X \leq y - a\} =$$
$$= \int_{-\infty}^{y-a} p_X(x)\, dx = F_X(y-a)$$

como la f.d.p es la derivada de la función anterior, se tiene

$$p_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X(y-a)}{dy} = p_X(x = y - a)$$

2. **Multiplicación por una constante positiva**: consideremos que Y se ha generado a partir de X según

$$Y = aX; \quad a > 0$$

siguiendo el mismo proceso anterior, se tiene

$$F_Y(y) = P\{Y \leq y\} = P\{aX \leq y\} = P\{X \leq y/a\} =$$
$$= \int_{-\infty}^{y/a} p_X(x)\, dx = F_X(y/a)$$

luego

$$p_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X(y/a)}{dy} = \frac{1}{a}\, p_X\left(x = \frac{y}{a}\right)$$

Veamos mediante un ejemplo un caso más complejo:

EJEMPLO Sea X una v.a. con f.d.p. uniforme en el intervalo $(0, 1)$

$$p_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0, \text{ en el resto} \end{cases}$$

Halla la f.d.p. de la v.a. $Y = -2 \ln X$.
Comencemos analizando los soportes de ambas v.a.:

$$\text{Soporte } X: \quad (0,1)$$
$$\text{Soporte } Y: \quad (0,\infty) \quad \begin{cases} x = 0 \Rightarrow y = -2\ln 0 = -2\,(-\infty) = \infty \\ x = 1 \Rightarrow y = -2\ln 1 = 0 \end{cases}$$

Calculamos $F_X(x)$

$$x \leq 0 \;\; F_X(x) = 0$$
$$0 < x < 1 \;\; F_X(x) = \int_0^x p_X(t)\, dt = \int_0^x 1\, dt = x$$
$$x \geq 1 \;\; F_X(x) = 1$$

La f.d. de Y será

$$
\begin{aligned}
F_Y(y) &= P\{Y \leq y\} = P\{-2\ln X \leq y\} = P\{\ln X \geq -y/2\} = \\
&= P\left\{X \geq \exp\left(-y/2\right)\right\} = 1 - P\left\{X < \exp\left(-y/2\right)\right\} = \\
&= 1 - P\left\{X \leq \exp\left(-y/2\right)\right\} = 1 - F_X(\exp\left(-y/2\right))
\end{aligned}
$$

entonces

$$
F_Y(y) = \begin{cases} 1 - F_X\Big(\exp\left(-y/2\right)\Big) = 1 - \exp\left(-y/2\right), \; y > 0 \\ \\ 0, \quad \text{en el resto} \end{cases}
$$

y llegamos a

$$
p_Y(y) = \frac{dF_Y(y)}{dy} = \begin{cases} \dfrac{1}{2}\exp\left(-y/2\right), \; y > 0 \\ \\ 0, \qquad\qquad\qquad \text{en el resto} \end{cases}
$$

2

Variables aleatorias bidimensionales

2.1. Introducción

En muchos fenómenos estadísticos la observación da lugar a un par de v.a. o, en general, a n v.a.. Esto se puede representar con una v.a. bidimensional (X,Y) y sus pares de valores $(x_1, y_1), (x_2, y_2), \ldots$ o, bien, con una v.a. n-dimensional $\mathbf{X} = [X_1, X_2, \ldots, X_n]$ y sus valores $\mathbf{x}_1, \mathbf{x}_2, \ldots$. Por sencillez, en este tema nos centraremos en el caso bidimensional pero todo lo que se vea se puede extender al caso n-dimensional.

EJEMPLO 1 Se lanza una moneda dos veces consecutivas. Se define X como la v.a. que cuenta el n^0 de caras de la 1^a tirada y la v.a. Y que cuenta el n^0 de caras de la 2^a tirada. El espacio muestral (Ω) es:

$$\Omega = \{HH, HT, TH, TT\}$$

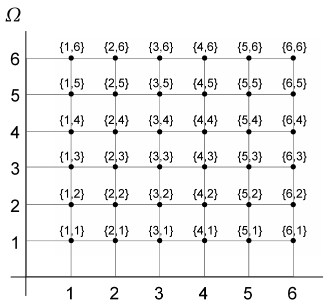
Las v.a. X e Y serán:

$$(X,Y) : \Omega \longrightarrow \Re^2$$

$$\begin{array}{ll} HH \to (1,1) & (X,Y)(\{HH\}) = (1,1) \\ HT \to (1,0) & (X,Y)(\{HT\}) = (1,0) \\ TH \to (0,1) & (X,Y)(\{TH\}) = (0,1) \\ TT \to (0,0) & (X,Y)(\{TT\}) = (0,0) \end{array}$$

EJEMPLO 2 Se lanza un dado dos veces consecutivas e interesa conocer la suma y la diferencia de las caras obtenidas en el experimento aleatorio. Podemos definir la v.a. X que nos da la suma de las caras y la v.a. Y que indica el valor absoluto de la diferencia de las caras.

Espacio muestral Pares de valores (X, Y)



$$
\begin{aligned}
(X, Y)(\{1,1\}) &= (2,0)\\
(X, Y)(\{1,2\}) &= (3,1)\\
&\vdots\\
(X, Y)(\{3,6\}) &= (9,3)\\
(X, Y)(\{4,1\}) &= (5,3)\\
&\vdots\\
(X, Y)(\{6,5\}) &= (11,1)\\
(X, Y)(\{6,6\}) &= (12,0)
\end{aligned}
$$

Llamaremos **variable aleatoria bidimensional** a la función

$$
\begin{aligned}
(X, Y) : \Omega &\longrightarrow \Re^2\\
\omega &\longrightarrow (X(\omega), Y(\omega))
\end{aligned}
$$

tal que $\forall (a, b) \in \Re^2 \quad \{\omega : X(\omega) \leq a \wedge Y(\omega) \leq b\}$, es un suceso del espacio muestral. En otras palabras, la imagen inversa de $(-\infty, a] \times (-\infty, b]$ es un suceso del espacio muestral Ω.

Del mismo modo que en el caso unidimensional, podemos definir dos tipos de v.a.:

1. **Discretas**: toman un conjunto finito o infinito numerable de valores.
2. **Continuas**: toman un conjunto infinito y no numerable de valores.

2.2. Distribución de probabilidad de una v.a. bidimensional

La función bidimensional que asigna probabilidades a los valores de una v.a. bidimensional es su distribución de probabilidad conjunta. Cuando las v.a. son discretas se la denomina como **función de masa de probabilidad** (f.m.p) conjunta y la denotaremos $P_{X,Y}(x, y)$ o, comúnmente, como $P(x, y)$. Verificando

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

Si la v.a. es continua se la designa como **función de densidad de probabilidad** (f.d.p.) conjunta y también se denotará como $p_{X,Y}(x, y)$ ó, más breve, $p(x, y)$.

Dicha función satisface las siguientes propiedades:

1. $p(x, y) \geq 0$
2. Se verifica

$$
\begin{aligned}
\text{discreta} \quad & \sum \sum P(x, y) = 1\\
\text{continua} \quad & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x, y)\, dxdy = 1
\end{aligned}
$$

3. Se tiene que $\forall (a,b) y (c,d) \in \Re^2$ la probabilidad de que el valor de la v.a. (X,Y) se encuentre en la región definida por los puntos (a,b) y (c,d) viene dada por:

$$\text{discreta} \quad P\{(a,b) \leq (X,Y) \leq (c,d)\} = \sum_{k=a}^{c} \sum_{k'=b}^{d} P(x=k, y=k') =$$

$$= \sum_{k=a}^{c} \sum_{k'=b}^{d} P\{X=k, Y=k'\}$$

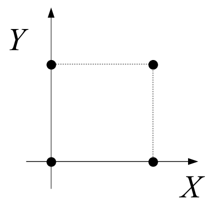
$$\text{continua} \quad P\{(a,b) \leq (X,Y) \leq (c,d)\} = \int_a^c \int_b^d p(x,y)\, dxdy,$$

EJEMPLO 1 (cont.) Se lanza una moneda dos veces consecutivas. Se define X como la v.a. que cuenta el n^0 de caras de la 1^a tirada y la v.a. Y que cuenta el n^0 de caras de la 2^a tirada. La f.m.p. conjunta es:

$$P\{(X,Y) = (1,1)\} = 1/4$$
$$P\{(X,Y) = (1,0)\} = 1/4$$
$$P\{(X,Y) = (0,1)\} = 1/4$$
$$P\{(X,Y) = (0,0)\} = 1/4$$

		X	
		0	1
Y	0	1/4	1/4
	1	1/4	1/4

El recorrido de esta f.m.p. lo podemos representar en el espacio de (X,Y)



EJEMPLO 2 (cont.) Se lanza un dado dos veces consecutivas y se define la v.a. X que nos da la suma de las caras y la v.a. Y que indica la diferencia de las caras.

		X										
		2	3	4	5	6	7	8	9	10	11	12
Y	0	1/36	0	1/36	0	1/36	0	1/36	0	1/36	0	1/36
	1	0	2/36	0	2/36	0	2/36	0	2/36	0	2/36	0
	2	0	0	2/36	0	2/36	0	2/36	0	2/36	0	0
	3	0	0	0	2/36	0	2/36	0	2/36	0	0	0
	4	0	0	0	0	2/36	0	2/36	0	0	0	0
	5	0	0	0	0	0	2/36	0	0	0	0	0

EJEMPLO 3 Sea (X,Y) una v.a. bidimensional con f.d.p. conjunta

$$p(x,y) = \begin{cases} x+y; \ 0 < x < 1 \quad 0 < y < 1 \\ 0, \qquad \text{en el resto} \end{cases}$$

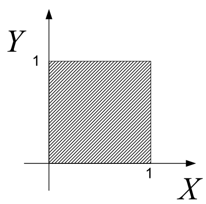
compruebe que esta función es una f.d.p.

Se debe verificar

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)\,dxdy = 1$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)\,dxdy = \int_0^1\int_0^1 (x+y)\ dxdy = \int_0^1 \left(\frac{x^2}{2}+yx\right)\Big|_0^1 dy =$$
$$\int_0^1 \left(\frac{1}{2}+y\right) dy = \left(\frac{y}{2}+\frac{y^2}{2}\right)\Big|_0^1 = \frac{1}{2}+\frac{1}{2} = 1$$

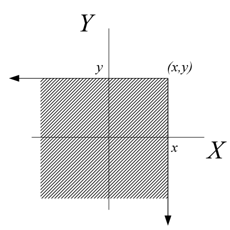
El soporte de esta función es la región $(0,1)\times(0,1)$



2.3. Función de distribución

Llamemos función de distribución de una v.a. (X,Y) a

$$F_{X,Y}(x,y) = P\{X \le x;\ Y \le y\} \quad \forall (x,y)\in \Re^2$$



discreta $F_{X,Y}(x,y) = P\{X \leq x;\ Y \leq y\} = \sum_{k \leq x} \sum_{k' \leq x} P(x = k, y = k') =$

$\sum_{k \leq x} \sum_{k' \leq x} P\{X = k, Y = k'\}$

continua $F_{X,Y}(x,y) = P\{X \leq x;\ Y \leq y\} = \int_{-\infty}^{x} \int_{-\infty}^{y} p_{X,Y}(t,t')\, dtdt'$

Dada la f.d $F_{X,Y}(x,y)$ se puede calcular la f.d.p.

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

La función de distribución de la v.a. (X,Y) debe verificar:

1. $F(-\infty, y) = 0$, $F(x, -\infty) = 0$ y $F(+\infty, +\infty) = 1$.
2. $F(x,y)$ es monótona no decreciente respecto a cada una de las variables.
3. Continua a la derecha de cada punto respecto a cada una de las variables.

EJEMPLO 1 (cont.) Se lanza una moneda dos veces consecutivas. Se define X como la v.a. que cuenta el n^0 de caras de la 1^a tirada y la v.a. Y que cuenta el n^0 de caras de la 2^a tirada. Obtenga la f.d conjunta:

$$F(x,y) = \begin{cases} 0; & x < 0 \text{ ó } y < 0 \\ 1/4; & 0 \leq x < 1 \text{ y } 0 \leq y < 1 \\ 1/2; & 1 \leq x \text{ y } 0 \leq y < 1 \\ 1/2; & 0 \leq x < 1 \text{ y } 1 \leq y \\ 1; & x \geq 1y \text{ y } \geq 1 \end{cases}$$

EJEMPLO 3 (cont.) Sea (X,Y) una v.a. bidimensional con f.d.p. conjunta

$$p(x,y) = \begin{cases} x + y; & 0 < x < 1 \quad 0 < y < 1 \\ 0, & \text{en el resto} \end{cases}$$

calcule su f.d.
En la región $0 \leq x \leq 1; 0 \leq y \leq 1$ se tiene

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p(u,v)\, dudv = \int_{0}^{x} \int_{0}^{y} (u+v)\ dudv = \int_{0}^{x} \left(uv + \frac{v^2}{2} \right) \Big|_0^y du =$$
$$= \int_0^x \left(uy + \frac{y^2}{2} \right) du = \left(\frac{u^2}{2} y + \frac{y^2}{2} u \right) \Big|_0^x = \frac{x^2 y}{2} + \frac{y^2 x}{2} = \frac{xy\,(x+y)}{2}$$

Luego

$$F(x,y)=\begin{cases} 0; & x<0 \text{ ó } y<0 \\ xy\,(x+y)/2; & 0\le x<1,\ 0\le y<1 \\ (y^2+y)\,/2; & 1\le x,\ 0\le y<1 \\ (x^2+x)\,/2; & 0\le x<1,\ 1\le y \\ 1; & x\ge 1,\ y\ge 1 \end{cases}$$

para obtener la función en la región $x \geq 1; 0 \leq y \leq 1$ e $y \geq 1; 0 \leq x \leq 1$ directamente se particulariza la expresión obtenida para $0 \leq x \leq 1; 0 \leq y \leq 1$ en $x = 1$ e $y = 1$, respectivamente.

2.4. Distribuciones marginales

Sea $F_{X,Y}(x,y)$ la f.d. de la v.a. (X,Y). Se llama función de distribución marginal de la v.a. X e Y, respectivamente, a las funciones:

$$F_X(x) = F_{X,Y}(x,\infty) = P\{X \leq x, Y < \infty\}$$

$$F_Y(y) = F_{X,Y}(\infty,y) = P\{X < \infty, Y \leq y\}$$

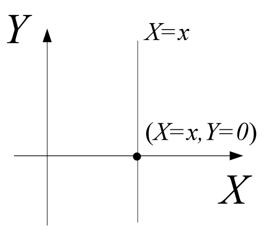
Las f.d. marginales cumplen las propiedades de una f.d.

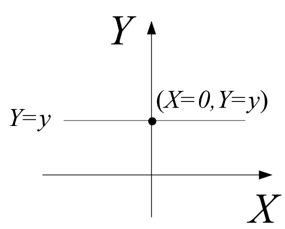
Se pueden definir, a partir de la f.d.p. conjunta $p_{X,Y}(x,y)$, **las funciones de densidad de probabilidad marginales** de las v.a. X e Y como:

$$\text{continua}\quad \begin{aligned} p_X(x) &= \textstyle\int_{-\infty}^{+\infty} p_{X,Y}(x,y)\,dy \\ p_Y(y) &= \textstyle\int_{-\infty}^{+\infty} p_{X,Y}(x,y)\,dx \end{aligned}$$

$$\text{discreta}\quad \begin{aligned} P_X(x) &= \textstyle\sum_k P_{X,Y}(x,y=k) = \sum_k P\{X=x, Y=k\} \\ P_Y(y) &= \textstyle\sum_k P_{X,Y}(x=k,y) = \sum_k P\{X=k, Y=y\} \end{aligned}$$

Se puede ver así, que la distribución marginal X en el punto x se obtiene sumando todas las masas situadas en la recta $X = x$ y asignado esta suma al punto $(X = x, Y = 0)$. Y, del mismo modo, la f.d.p. marginal de Y en el punto y se obtiene sumando las masas en la recta $Y = y$ y asignándoselas al punto $(X = 0, Y = y)$.





EJEMPLO 1 (cont.) Se lanza una moneda dos veces consecutivas. A partir de la f.d.p. conjunta de las v.a. X e Y que cuentan el n^0 de caras de la 1^a y 2^a tirada, respectivamente, obtenga las f.d.p. marginales.

$P_{X,Y}(x,y)$

		X	
		0	1
Y	0	$1/4$	$1/4$
	1	$1/4$	$1/4$

$P_X(x)$

X	$P\{X=x\}$
0	$2/4$
1	$2/4$

$P_Y(y)$

Y	$P\{Y=y\}$
0	$2/4$
1	$2/4$

EJEMPLO 2 (cont.) A partir de f.d.p. conjunta de las v.a. X e Y que resultan de sumar las caras y calcular la diferencia de las caras, respectivamente, cuando se lanza un dado dos veces consecutivas, obtenga las f.d.p. marginales.
Las marginales se calculan:

$$P_X(x=2) = \sum_{k=0}^{5} P(x=2, y=k) = 1/36$$
$$P_X(x=3) = \sum_{k=0}^{5} P(x=3, y=k) = 2/36$$
$$\vdots$$
$$P_X(x=12) = \sum_{k=0}^{5} P(x=12, y=k) = 1/36$$

$$P_Y(x=0) = \sum_{k=2}^{12} P(x=k, y=0) = 6/36$$
$$P_Y(x=1) = \sum_{k=2}^{12} P(x=k, y=1) = 10/36$$
$$\vdots$$
$$P_Y(x=5) = \sum_{k=2}^{12} P(x=k, y=5) = 2/36$$

x	2	3	4	5	6	7	8	9	10	11	12
$P_X(x)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

y	0	1	2	3	4	5
$P_Y(y)$	$6/36$	$10/36$	$8/36$	$6/36$	$4/36$	$2/36$

EJEMPLO 3 (cont.) Sea (X,Y) una v.a. bidimensional con f.d.p. conjunta

$$p(x,y) = \begin{cases} x+y; \ 0<x<1 \quad 0<y<1 \\ 0, \quad \text{en el resto} \end{cases}$$

obtenga las f.d.p. marginales.
En la región $0 \le x \le 1; 0 \le y \le 1$ se tiene

$$p_X(x) = \int_{-\infty}^{\infty} p(x,y)\,dy = \int_0^1 (x+y)\ dy = \left(xy + \frac{y^2}{2} \right)\Bigg|_0^1 = x + \frac{1}{2} \quad 0<x<1$$

$$p_Y(y) = \int_{-\infty}^{\infty} p(x,y)\,dx = \int_0^1 (x+y)\ dx = \left(xy + \frac{x^2}{2} \right)\Bigg|_0^1 = y + \frac{1}{2} \quad 0<y<1$$

$$p_X(x) = \begin{cases} x + \frac{1}{2}; \ 0<x<1 \\ 0, \quad \text{en el resto} \end{cases} \qquad p_Y(x) = \begin{cases} y + \frac{1}{2}; \ 0<y<1 \\ 0, \quad \text{en el resto} \end{cases}$$

2.5. Distribuciones condicionadas

Sea $p_{X,Y}(x,y)$ la f.d.p conjunta de las v.a. X,Y, se define la f.d.p de Y condicionada a $X = x$ y la f.d.p. de X condicionada a $Y = y$ como:

$$p_{Y|X}(y|x) = p_{X,Y}(x,y)/p_X(x)$$
$$p_{X|Y}(x|y) = p_{X,Y}(x,y)/p_Y(y)$$

En el caso de v.a. discretas, se tiene

$$P_{Y|X}(y|x) = P_{X,Y}(x,y)/P_X(x)$$
$$P_{X|Y}(x|y) = P_{X,Y}(x,y)/P_Y(y)$$

EJEMPLO 4 Se lanza una moneda tres veces consecutivas. Se consideran las v.a.:

- X: cuenta el n^0 de caras en las tres tiradas.
- Y: es el valor absoluto de la diferencia entre el n^0 de caras y de cruces en las tres tiradas.

Se pide:

a) Distribución de probabilidad de la v.a. (X,Y)

Espacio muestral
$\Omega = \{HHH, HHT, HTH, HTT,$
$THH, THT, TTH, TTT\}$

$P_{X,Y}(x,y)$		x			
		0	1	2	3
y	1	0	3/8	3/8	0
	3	1/8	0	0	1/8

b) Distribuciones marginales de X e Y.

x	$P_X(x)$
0	1/8
1	3/8
2	3/8
3	1/8

y	$P_Y(y)$
1	6/8
3	1/8

c) Distribución de X condicionada a $Y = 3$.

$$P(x=0|y=3) = \frac{P(x=0,y=3)}{P(y=3)} = \frac{1/8}{2/8} = 1/2$$
$$P(x=1|y=3) = \frac{P(x=1,y=3)}{P(y=3)} = 0$$
$$P(x=2|y=3) = \frac{P(x=2,y=3)}{P(y=3)} = 0$$
$$P(x=3|y=3) = \frac{P(x=3,y=3)}{P(y=3)} = \frac{1/8}{2/8} = 1/2$$

x	0	1	2	3
$P(x\|y=3)$	1/2	0	0	1/2

c) Distribución de Y condicionada a $X = 2$.

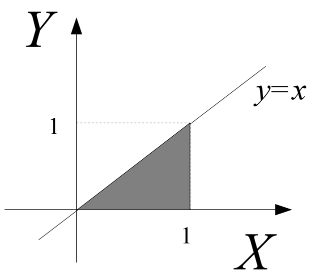
$$P(y=1|x=2) = \frac{P(x=2, y=1)}{P(x=2)} = \frac{3/8}{3/8} = 1$$
$$P(y=3|x=2) = \frac{P(x=2, y=3)}{P(x=2)} = 0$$

y	1	3
$P(y\|x=2)$	1	0

EJEMPLO 5 Sea (X, Y) una v.a. bidimensional con f.d.p. conjunta

$$p(x,y) = \begin{cases} 10x^2y; \ 0 \le y \le x \le 1 \\ 0, \quad \text{en el resto} \end{cases}$$

Halla la distribución de X condicionada a $Y = 1/2$ y la distribución de Y condicionada a $X = 1/2$.
El soporte de la función es:



Calculamos las distribuciones marginales

$$p_X(x) = \int_{-\infty}^{+\infty} p(x,y)\,dy = \int_0^x 10x^2y\,dy = 10x^2 \left.\frac{y^2}{2}\right|_0^x = 5x^4 \quad 0 \le x \le 1$$
$$p_Y(y) = \int_{-\infty}^{+\infty} p(x,y)\,dx = \int_y^1 10x^2y\,dx = 10y \left.\frac{x^3}{3}\right|_y^1 = \frac{10}{3}y\left(1-y^3\right) \quad 0 \le y \le 1$$

Las distribuciones condicionas serán

$$p_{X|Y}(x|y=1/2) = \frac{p(x,1/2)}{p(1/2)} = \frac{10x^2 - \frac{1}{2}}{10\frac{1}{2}\left(1-\frac{1}{8}\right)} = \frac{24}{7}x^2 \quad \frac{1}{2} \le x \le 1$$
$$p_{Y|X}(y|x=1/2) = \frac{p(1/2,y)}{p(1/2)} = \frac{10\frac{1}{4}y}{5\frac{1}{16}} = 8y \quad 0 \le y \le \frac{1}{2}$$

2.6. Variables aleatorias independientes

Una v.a. (X,Y) de f.d. $F(x,y)$ con f.d. marginales $F_X(x)$ y $F_Y(y)$ son independientes si

$$F(x,y) = F_X(x) \cdot F_Y(y)$$

Para v.a. continuas se tiene q las f.d.p. deben verificar

$$p(x,y) = \frac{\partial F(x,y)}{\partial x \partial y} = \frac{\partial F_X(x) \cdot F_Y(y)}{\partial x \partial y} = \frac{dF_X(x)}{dx} \cdot \frac{dF_Y(y)}{dy} = p_X(x) \cdot p_Y(y)$$

Nota: Para que 2 v.a. continuas sean independientes es necesario, pero no suficiente, que el soporte de su f.d.p. conjunta sea un rectángulo.

Ya que

$$p_{X,Y}(x,y) = p_{Y|X}(y|x) \cdot p_X(x) = p_{X|Y}(x|y) \cdot p_Y(y)$$

se tiene que las v.a. X o Y son independientes sii

$$p_{Y|X}(y|x) = p_Y(y) \quad , \quad p_{X|Y}(x|y) = p_X(x)$$

Para distribuciones discretas se debe verificar

$$P_{X,Y}(x,y) = P\{X = x, Y = y\} = P\{X = x\} \cdot P\{Y = y\} = P_X(x) \cdot P_Y(y)$$

o bien

$$P_{Y|X}(y|x) = P\{Y = y | X = x\} = P\{Y = y\} = P_Y(y)$$
$$P_{X|Y}(x|y) = P\{X = x | Y = y\} = P\{X = x\} = P_X(x)$$

EJEMPLO 6 Sea la v.a. discreta (X,Y) con f.m.p. conjunta

$$P(x,y) = \frac{1}{2^{x+y}} \qquad \forall x = 1,2,\ldots \quad \forall y = 1,2,\ldots$$

Halla las distribuciones marginales, condicionadas y analiza si son independientes.
Las distribuciones marginales son:

$$P_X(x) = \sum_{x=1}^{\infty} \frac{1}{2^{x+y}} = \frac{1}{2^y} \sum_{x=1}^{\infty} \frac{1}{2^x} = \frac{1}{2^y} \frac{1/2}{1-1/2} = \frac{1}{2^y} \quad \forall y = 1,2,\ldots$$
$$P_Y(y) = \frac{1}{2^x} \quad \forall x = 1,2,\ldots$$

Se tiene que

$$P_{X,Y}(x,y) = \frac{1}{2^{x+y}} = \frac{1}{2^y} \cdot \frac{1}{2^x} = P_X(x) \cdot P_Y(y) = P_X(x)$$

luego, X e Y son independientes y las distribuciones condicionadas son:

$$P_{X|Y}(x|y) = P(x) = \frac{1}{2^y} \frac{1/2}{1-1/2} = \frac{1}{2^y} \quad \forall y = 1,2,\ldots$$
$$P_{Y|X}(y|x) = P(y) = \frac{1}{2^x} \quad \forall x = 1,2,\ldots$$

EJEMPLO 3 (cont.) Sea (X, Y) una v.a. bidimensional con f.d.p. conjunta

$$p(x,y) = \begin{cases} x+y; \; 0<x<1 \quad 0<y<1 \\ 0, \qquad \text{en el resto} \end{cases}$$

¿son X e Y independientes?

Teníamos que $p_X(x) = \begin{cases} x + \frac{1}{2}; \; 0<x<1 \\ 0, \qquad \text{en el resto} \end{cases}$ $p_Y(x) = \begin{cases} y + \frac{1}{2}; \; 0<y<1 \\ 0, \qquad \text{en el resto} \end{cases}$

si son independientes $p(x,y) = p_X(x) \cdot p_Y(y)$; en nuestro caso:

$$p_X(x) \cdot p_Y(y) = \begin{cases} \left(x+\frac{1}{2}\right)\left(y+\frac{1}{2}\right) = xy + \frac{1}{2}(x+y) + \frac{1}{4}; \; 0<x<1, \; 0<y<1 \\ 0, \qquad \text{en el resto} \end{cases}$$

como $p(x,y) \neq p_X(x) \cdot p_Y(y)$, X e Y son dependientes.

2.7. Teorema de Probabilidad Total

El **Teorema de probabilidad Total** nos permite calcular la f.m.p. ó f.d.p. de una v.a. X a partir de sus probabilidades condicionadas respecto a otra variable Y:

$$\text{discreta} \;\; P_X(x) = \sum_k P_{X|Y}(x|y=k) \cdot P_Y(y=k)$$

$$\text{continua} \;\; p_X(x) = \int p_{X|Y}(x|y) \cdot p_Y(y)\, dy$$

2.8. Teorema de Bayes

Consideremos que se tiene una v.a. X y un determinado fenómeno aleatorio genera a partir de ella otra v.a. Y. Conociendo la probabilidad de Y dado el valor de X, es decir, $p_{Y|X}(y|x)$ y conocida la probabilidad marginal de X, el **Teorema de Bayes** proporciona la distribución de probabilidad condicionada de X dada Y, $p_{X|Y}(x|y)$, mediante

$$p_{X|Y}(x|y) = \frac{p_{Y|X}(y|x) \cdot p_X(x)}{\int p_{Y|X}(y|x) \cdot p_X(x)\, dx}$$

Normalmente, a la distribución marginal de X, $p_X(X)$, se la llama probabilidad "a priori", a las probabilidades $p_{X|Y}(X|Y)$ se las denomina probabilidades "a posteriori" y la probabilidad $p_{Y|X}(y|x)$ se denomina "verosimilitud de X".

DEMOSTRACIÓN A partir de la definición de probabilidad condicionada, se tiene

$$p_{X,Y}(x,y) = p_{X|Y}(x|y) \cdot p_Y(y) = p_{Y|X}(y|x) \cdot p_X(x)$$

Despejando $p_{X|Y}(x|y)$ y utilizando el teorema de probabilidad total

$$p_{X|Y}(x|y) == \frac{p_{Y|X}(y|x) \cdot p_X(x)}{p_Y(y)} = \frac{p_{Y|X}(y|x) \cdot p_X(x)}{\int p_{Y|X}(y|x) \cdot p_X(x)\, dx}$$

Para el caso de v.a. discretas la formulación del Teorema de Bayes es equivalente:

$$P_{X|Y}(x|y) = \frac{P_{Y|X}(y|x) \cdot P_X(x)}{\sum_k^n P_{Y|X}(y|x=k) \cdot P_X(x=k)}$$

2.9. Momentos de variables aleatorias bidimensionales

El operador esperanza matemática sobre una v.a. bidimensional con f.d.p. conjunta $p(x, y)$ se define como:

discreta $\mathbb{E}\{g(X,Y)\} = \sum_k \sum_{k'} g(k,k')\, P_{X,Y}(x=k, y=k') = \sum_k \sum_{k'} g(k,k')\, P\{X=k, Y=k'\}$

continua $\mathbb{E}\{g(X,Y)\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y)\, p(x,y)\, dxdy$

que particularizando para $g(x, y) = x^n y^m$ nos permite definir los momentos centrados (respecto al origen) de la v.a. (X, Y). Así tenemos:

$n=1, m=0 \quad \mathbb{E}\{X\} \quad n=0, m=1 \quad \mathbb{E}\{Y\}$, **Medias marginales**

$n=1, m=1 \quad \mathbb{E}\{XY\}$

$n=2, m=0 \quad \mathbb{E}\{X^2\} \quad n=0, m=2 \quad \mathbb{E}\{Y^2\}$,

Cuando las v.a. X e Y son *independientes*, se verifica

$$\mathbb{E}\{XY\} = \mathbb{E}\{X\} \cdot \mathbb{E}\{Y\}$$

DEMOSTRACIÓN: para el caso de una v.a. continua

$$\mathbb{E}\{XY\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy\, p(x,y)\, dxdy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy\, p(x)p(y)\, dxdy =$$
$$= \int_{-\infty}^{+\infty} x\, p(x)\, dx \int_{-\infty}^{+\infty} y\, p(y) dy = \mathbb{E}\{X\} \cdot \mathbb{E}\{Y\}$$

Si $\mathbb{E}\{XY\} = 0$ se dice que las v.a. X e Y son *ortogonales*.

Si $g(x, y) = (x - \mathbb{E}\{X\})^n (y - \mathbb{E}\{Y\})^m$, se definen los momentos centrales (respecto a las medias marginales) de la v.a. (X, Y),

$n=2, m=0 \quad Var\{X\} = \mathbb{E}\left\{(X - \mathbb{E}\{X\})^2\right\}$

$n=0, m=2 \quad Var\{Y\} = \mathbb{E}\left\{(Y - \mathbb{E}\{Y\})^2\right\}$, **Varianzas marginales**

$n=1, m=1 \quad Covar\{X,Y\} = \mathbb{E}\{(X - \mathbb{E}\{X\}) \cdot (Y - \mathbb{E}\{Y\})\}$ **Covarianza**

El calculo de las varianzas y la covarianza se puede realizar según:

$$
\begin{aligned}
Var\{X\} &= \mathbb{E}\left\{(X-\mathbb{E}\{X\})^2\right\} = \mathbb{E}\left\{X^2\right\} - \mathbb{E}\{X\}^2 \\
Var\{Y\} &= \mathbb{E}\left\{(Y-\mathbb{E}\{Y\})^2\right\} = \mathbb{E}\left\{Y^2\right\} - \mathbb{E}\{Y\}^2 \\
Covar\{X,Y\} &= \mathbb{E}\{(X-\mathbb{E}\{X\})\cdot(Y-\mathbb{E}\{Y\})\} = \mathbb{E}\{XY\} - \mathbb{E}\{X\}\mathbb{E}\{Y\}
\end{aligned}
$$

También se pueden definir los momentos condicionados de X dado Y o de Y dado X como:

$$
\begin{aligned}
\text{discreta}\;\; \mathbb{E}\{g(X)|Y=y\} &= \sum_k g(k)\, P_{X|Y}(x=k|y) \\
\mathbb{E}\{g(Y)|X=x\} &= \sum_k g(k)\, P_{Y|X}(y=k|x) \\
\text{continua}\;\; \mathbb{E}\{g(X)|Y=y\} &= \int_{-\infty}^{+\infty} g(x)\, p_{X|Y}(x|y)\, dx \\
\mathbb{E}\{g(Y)|X=x\} &= \int_{-\infty}^{+\infty} g(y)\, p_{Y|X}(y|x)\, dy
\end{aligned}
$$

A partir de los momentos condicionados, se define el **Teorema de la Esperanza Matemática Total** como:

$$
\begin{aligned}
X\ \text{discreta}\;\; \mathbb{E}\{Y\} &= \sum_k \mathbb{E}\{Y|X=x\}\cdot P_X(x=k) \\
X\ \text{continua}\;\; \mathbb{E}\{Y\} &= \int \mathbb{E}\{Y|X=x\}\cdot p_X(x)\, dx
\end{aligned}
$$

Nota: En ambos casos Y puede ser una v.a discreta o una v.a. continua; es decir, siendo X discreta Y puede ser continua o discreta; e, igualmente, cuando X es continua, Y puede ser tanto continua como discreta.

Se define **matriz de varianzas-covarianzas**, o comúnmente matriz de covarianzas, como:

$$
\begin{pmatrix} Var\{X\} & Covar\{X,Y\} \\ Covar\{X,Y\} & Var\{Y\} \end{pmatrix}
$$

y esta matriz es una forma cuadrática definida positiva, es decir,

$$
\begin{vmatrix} Var\{X\} & Covar\{X,Y\} \\ Covar\{X,Y\} & Var\{Y\} \end{vmatrix} \geq 0
$$

Y, del mismo modo que en el caso unidimensional, se define la **desviación típica** de una v.a. como

$$
\sigma_X = \sqrt{Var\{X\}} \quad \sigma_Y = \sqrt{Var\{Y\}} \quad \sigma_{X,Y} = \sqrt{Covar\{X,Y\}}
$$

Además, para el caso bidimensional, se define el **coeficiente de correlación** como:

$$\rho = \frac{Covar\{X,Y\}}{\sigma_X \cdot \sigma_Y}$$

este coeficiente mide la interdependencia lineal de 2 v.a. Del hecho de que la matriz de varianzas-covarianzas sea definida positiva, se tiene:

$$Var\{Y\} \cdot Var\{X\} - Covar\{X,Y\}^2 \geq 0$$

$$Var\{Y\} \cdot Var\{X\} \geq Covar\{X,Y\}^2$$

$$\frac{Covar\{X,Y\}^2}{Var\{Y\} \cdot Var\{X\}} = \rho^2 \leq 1$$

luego $0 \leq \rho^2 \leq 1 \Rightarrow -1 \leq \rho \leq 1$.

Cuando $\rho = 0$ se dice que X e Y están *incorreladas*. Se verifica que si 2 v.a. son independientes, van a estar incorreladas, ya que:

$$Covar\{X,Y\} = \mathbb{E}\{XY\} - \mathbb{E}\{X\}\mathbb{E}\{Y\} = \mathbb{E}\{X\}\mathbb{E}\{Y\} - \mathbb{E}\{X\}\mathbb{E}\{Y\} = 0$$

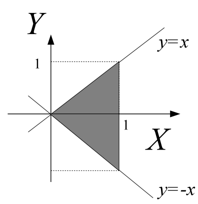
Nota: El hecho de que 2 v.a. estén incorreladas, no implica que sean independientes; es decir:

$$\textit{Independientes} \underset{\not\Leftarrow}{\Rightarrow} \textit{Incorreladas}$$

EJEMPLO 7 Sea (X,Y) una v.a. bidimensional con f.d.p. conjunta

$$p(x,y) = \begin{cases} 1;\ |y| < x \quad 0 < x < 1 \\ 0, \text{ en el resto} \end{cases}$$

Halla:
a) Distribuciones marginales



Calculamos las distribuciones marginales

$$p_X(x) = \int_{-\infty}^{+\infty} p(x,y)\,dy = \int_{-x}^{x} 1\,dy = y\Big|_{-x}^{x} = 2x \quad 0 \leq x \leq 1$$

$$p_Y(y) = \int_{-\infty}^{+\infty} p(x,y)\,dx = \int_{|y|}^{1} 1\,dx = x\Big|_{|y|}^{1} = 1 - |y| \quad -1 \leq y \leq 1$$

b) Distribuciones condicionadas
Las distribuciones condicionas serán

$$p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)} = \frac{1}{1-|y|} \quad |y| \le x \le 1$$
$$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)} = \frac{1}{2x} \quad |y| \le x$$

Vemos que las variables son dependientes: $p_{X|Y}(x|y) \neq p_X(x)$ y $p_{Y|X}(y|x) \neq p_Y(y)$.

c) Coeficiente de correlación

$$\rho = \frac{Covar\{X,Y\}}{\sigma_X \cdot \sigma_Y}$$

Calculamos los diferentes términos:

$$\mathbb{E}\{X\} = \int_{-\infty}^{+\infty} x\, p(x)\, dx = \int_0^1 2x^2\, dx = \frac{2}{3}x^3\Big|_0^1 = \frac{2}{3}$$
$$\mathbb{E}\{Y\} = \int_{-\infty}^{+\infty} y\, p(y)\, dy = \int_{-1}^{0} y\,(1+y)\ dy + \int_0^1 y\,(1-y)\ dy =$$
$$= \frac{y^2}{2} + \frac{y^3}{3}\Big|_{-1}^{0} + \frac{y^2}{2} - \frac{y^3}{3}\Big|_0^1 = -\left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) = 0$$
$$\mathbb{E}\{XY\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\, p(x,y)\, dxdy = \int_0^1\int_{-x}^{x} xy\, dxdy = \int_0^1 x\frac{y^2}{2}\Big|_{-x}^{x} dx = 0$$
$$Covar\{X,Y\} = \mathbb{E}\{XY\} - \mathbb{E}\{X\}\,\mathbb{E}\{Y\} = 0 - \frac{2}{3}0 = 0$$

Luego,

$$\rho = \frac{Covar\{X,Y\}}{\sigma_X \cdot \sigma_Y} = 0$$

Como puede verse $\rho = 0 \Rightarrow X$ e Y son incorreladas, sin embargo, son dependientes.

2.10. Cambios de variables aleatorias

Consideremos que se conocen las f.d.p. de 2 v.a. X e Y y mediante una función g de ellas obtenemos otra v.a. Z de la que se desea calcular su f.d.p. Es decir, sea

$$Z = g(X,Y)$$

Para obtener $p_Z(z)$, calcularemos la f.d.p. conjunta de Z y X mediante,

$$p_{Z,X}(z,x) = p_{Z|X}(z|x) \cdot p_X(x)$$

dado que $p_X(x)$ es conocido, la única dificultad es obtener $p_{Z|X}(z|x)$; sin embargo, como se está condicionando a X, el cambio de variable será de una v.a. a otra, por lo que puede calcularse mediante lo visto en la Sección 1.7. Una vez obtenida $p_{Z|X}(z|x)$, $p_Z(z)$ será:

$$p_Z(z) = \int_{-\infty}^{+\infty} p_{Z|X}(z|x) \cdot p_X(x)\, dx$$

Veamos con un par de ejemplos como aplicar este procedimiento.

1. **Suma de dos v.a. independientes**: sea

$$Z = X + Y$$

Dadas 2 v.a. X e Y, uniformes en $(0,1)$, e independientes entre sí, obtenga $p_Z(z)$. Como se ha visto, $p_Z(z)$ viene dado por:

$$p_Z(z) = \int_{-\infty}^{+\infty} p_{Z|X}(z|x) \cdot p_X(x)\, dx$$

donde $p_{Z|X}(z|x)$ se calcula aplicando el cambio de v.a. (unidimensional) $Z = x+Y$, por lo que,

$$p_{Z|X}(z|x) = p_Y(y = z - x)$$

Entonces,

$$p_Z(z) = \int_{-\infty}^{+\infty} p_Y(y = z-x) \cdot p_X(x)\, dx = p_Y(y) * p_X(x)$$

que como puede verse, en este caso, $p_Z(z)$ puede obtenerse como la convolución de las f.d.p. de X e Y. Finalmente, considerando que X e Y son distribuciones uniformes, se tiene

$$\begin{aligned} z \leq 0 \quad & p_Z(z) = 0 \\ 0 < z \leq 1 \quad & p_Z(z) = \int_0^z dx = z \\ 1 < z < 2 \quad & p_Z(z) = \int_{z-1}^1 dx = 1 - (z-1) = 2 - z \\ z \geq 2 \quad & p_Z(z) = 0 \end{aligned}$$

2. **Producto de dos v.a. independientes**: sea

$$Z = X \cdot Y$$

Dadas 2 v.a. X e Y, uniformes en $(0,1)$, e independientes entre sí, obtenga $p_Z(z)$.

$$p_Z(z) = \int_{-\infty}^{+\infty} p_{Z|X}(z|x) \cdot p_X(x)\, dx$$

donde $p_{Z|X}(z|x)$ se obtiene el cambio de v.a. (unidimensional) $Z = x \cdot Y$, por lo que,

$$p_{Z|X}(z|x) = \frac{1}{x}\, p_Y(y = \frac{z}{x})$$

Entonces,

$$p_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{x}\, p_Y(y = \frac{z}{x}) \cdot p_X(x)\, dx$$

Sustituyendo,

$$\begin{aligned} z \leq 0 &\quad p_Z(z) = 0 \\ 0 < z < 1 &\quad p_Z(z) = \int_z^1 \frac{1}{x}\, dx = \ln x \Big|_z^1 = \ln 1 - \ln z = -\ln z \\ z \geq 1 &\quad p_Z(z) = 0 \end{aligned}$$

2.11. Variable gaussiana multidimensional

Una v.a. de especial interés es la v.a. gaussiana multidimensional, que tiene la siguiente f.d.p.

$$p_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}\, |V|^{1/2}} \exp\left(\frac{1}{2}\, (\mathbf{x}-\mathbf{m})^T\, V^{-1}\, (\mathbf{x}-\mathbf{m})\right)$$

donde **m** es el vector de medias de X, $\mathbf{m} = \mathbb{E}\{\mathbf{X}\}$, y V es la matriz de covarianzas de las componentes de **X**; es decir, siendo $\mathbf{X} = [X_1, X_2, \ldots, X_N]$

$$V = \begin{pmatrix} Var\{X_1\} & \ldots\, Covar\{X_1, X_N\} \\ \vdots & \ddots \quad\quad \vdots \\ Covar\{X_N, X_1\} & \ldots \quad Var\{X_N\} \end{pmatrix}$$

Una característica bastante importante de esta distribución es que conociendo los momentos de primer y segundo orden (medias, varianzas y covarianzas) se tiene definida completamente a la v.a., ya que se información suficiente para obtener la f.d.p. conjunta de **X**.

La matriz de covarianzas directamente indica si las componentes de **X** están relacionadas entre sí, ya que si la matriz de covarianzas resulta ser diagonal se tiene que la f.d.p. conjunta se puede expresar como el producto de las distribuciones marginales,

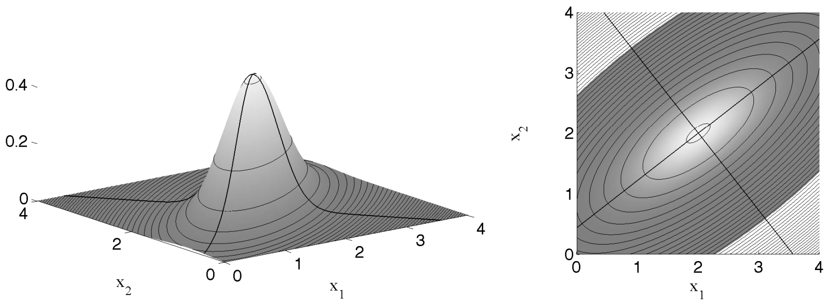
$$\begin{aligned} p_{\mathbf{X}}(\mathbf{x}) &= \frac{1}{(2\pi)^{N/2}\, |V|^{1/2}} \exp\left(\frac{1}{2}\, (\mathbf{x}-\mathbf{m})^T\, V^{-1}\, (\mathbf{x}-\mathbf{m})\right) \\ &= \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi v_n}} \exp\left(\frac{(x_n - m_n)^2}{2 v_n}\right) = \prod_{n=1}^{N} p_{X_n}(x_n) \end{aligned}$$

lo que demuestra que las componentes de **X** son independientes entre sí.

Mientras que el vector de medias indica donde se encuentra el centro de la distribución, la matriz de covarianza indica dispersión en cada componente. Para analizar este aspecto en detalle, consideremos una v.a. aleatoria gaussiana bidimensional de media $\mathbf{m} = [2, 2]^T$ y matriz de covarianza

$$V = \begin{pmatrix} 1 & 0{,}6 \\ 0{,}6 & 0{,}7 \end{pmatrix}$$

en este caso se tiene que f.d.p. tiene la siguiente representación en el espacio (dibujo de la izquierda) y en el plano $x_1 - x_2$ (dibujo de la derecha)

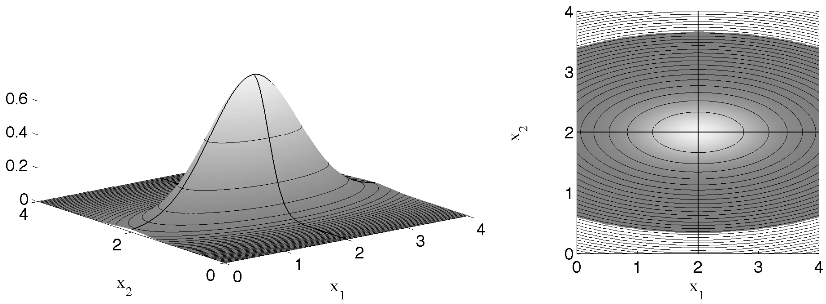


Como puede verse la distribución tiene curvas de nivel elípticas; puede demostrarse que los ejes principales de las elipses se corresponden con los vectores autovectores de la matriz de covarianza, que en este caso son: $q_1 = [0{,}615, -0{,}788]^T$ y $q_2 = [0{,}788, 0{,}615]^T$.

Cuando la matriz de covarianza es diagonal (los términos de covarianza son nulos), los ejes principales coinciden con las direcciones de x_1 y x_2; así, por ejemplo, si consideramos

$$V = \begin{pmatrix} 1 & 0 \\ 0 & 0{,}2 \end{pmatrix}$$

se tiene:



Y si además las varianzas de las componentes son iguales, las curvas nivel pasan de tener forma elíptica a ser círculos; si ahora

$$V = \begin{pmatrix} 0{,}5 & 0 \\ 0 & 0{,}5 \end{pmatrix}$$

la f.d.p es: