

Variables Aleatorias. Distribuciones de Probabilidad

Variables Aleatorias. Distribuciones de Probabilidad

Departamento Matemática Aplicada

Universidad de Málaga

Curso 2015-2016

Introducción

Dado un experimento aleatorio, al conjunto de sucesos elementales le hemos llamado espacio muestral (E). Supongamos que es discreto.

A cada suceso elemental podemos asociarle un número real de muchas formas diferentes. Cada una será una variable aleatoria.

Por ejemplo, al lanzar una moneda 4 veces, el espacio muestral es:

$E = \{CCCC, CCCF, CCFC, CCFF, CFCC, CFCF, CFFC, CFFF, FCCC, FCCF, FCFC, FCFF, FFCC, FFCF, FFFC, FFFF\}$ donde $C = \text{'Cara'}$ y $F = \text{'Cruz'}$.

- 'Número de caras obtenidas'. Al suceso $CFFF$ le corresponde un 1 y a $FCFC$ un 2.
- 'Número de caras antes de la primera aparición de cruz'. Al suceso $CFFF$ le corresponde un 1 y a $FCFC$ un 0.
- 'Cada cara se valora multiplicada por el lugar de aparición'. Al suceso $CFFF$ le corresponde un 1 ($1+0+0+0$) y a $FCFC$ un 6 ($0+2+0+4$).



Ejemplo

Ejemplo

Dado el experimento aleatorio consistente en lanzar un dado y una moneda y la variable aleatoria ξ : 'valor obtenido en el dado, que se duplica si resulta cara al lanzar la moneda'.

- *Determinar su espacio muestral.*
- *Hallar el aplicado de cada uno de los elementos de E .*
- *Determinar el soporte de X .*

a) $E = \{'1C', '2C', '3C', '4C', '5C', '6C', '1F', '2F', '3F', '4F', '5F', '6F'\}$

b) $X('1C')=2, X('2C')=4, X('3C')=6, X('4C')=8, X('5C')=10,$
 $X('6C')=12, X('1F')=1, X('2F')=2, X('3F')=3, X('4F')=4,$
 $X('5F')=5, X('6F')=6.$

c) El soporte $S_X = \{1, 2, 3, 4, 5, 6, 8, 10, 12\}$



Ejemplo

Ejemplo

Determinar el soporte para:

- *La variable aleatoria 'Número de llamadas recibidas'.*
- *Obtener un número aleatorio x en $[0,1)$. ($x=\text{rand}(1,1)$)*
- *Multiplicar 6 por x y sumarle 1. ($y=6*x+1$)*
- *Quedarte con la parte entera de y . ($z=\text{floor}(y)$)*

Ejemplo

Un componente es sustituido cuando se avería o al cabo de 10 años.

- *Determinar el soporte de la v.a. 'Duración del componente'.*
- *Si disponemos de 5 componentes iguales y cada uno sustituye al anterior. Lo mismo para la v.a. 'Duración de los 5 componentes'.*
- *Repetir los anteriores apartados si solo se sustituye por avería.*

Función de distribución

Definición

*Dado un espacio probabilístico (E, \mathcal{A}, P) y una variable aleatoria X . Definimos la **función de distribución de la variable aleatoria X** como una función que verifica:*

- 1 $F : \mathbb{R} \rightarrow [0, 1]$
- 2 $F(x) = P(\omega \in \mathcal{A} / X(\omega) \leq x)$

Propiedad

La función de distribución asociada a una variable aleatoria es única y caracteriza a la misma.

Esto es, si conocemos la función de distribución podemos calcular la probabilidad de que la variable aleatoria tome un valor.

Propiedades de la función de distribución

- 1 $P(X \leq x) = F(x)$
- 2 $P(X > x) = 1 - P(X \leq x) = 1 - F(x)$
- 3 $P(a < X \leq b) = F(b) - F(a)$
- 4 $P(X = a) = F(a) - \lim_{h \rightarrow 0^+} F(a - h)$
- 5 $0 \leq F(x) \leq 1$
- 6 $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$
- 7 $F(\infty) = \lim_{x \rightarrow \infty} F(x) = 1$
- 8 **$F(x)$ es monótona no decreciente**, es decir:
 $a < b \Rightarrow F(a) \leq F(b)$
- 9 **$F(x)$ es continua por la derecha**, es decir,
 $\lim_{h \rightarrow 0^+} F(x + h) = F(x)$

Ejemplo

Ejemplo

La función de distribución $F(x)$ de una v.a. discreta ξ viene dada por:

x	$(-\infty, 1)$	$[1, 2)$	$[2, 3)$	$[3, 4)$	$[4, 5)$	$[5, 6)$	$[6, 8)$	$[8, 10)$	$[10, 12)$	$[12, \infty)$
$F(x)$	0	$\frac{1}{12}$	$\frac{3}{12}$	$\frac{4}{12}$	$\frac{6}{12}$	$\frac{7}{12}$	$\frac{9}{12}$	$\frac{10}{12}$	$\frac{11}{12}$	1

Hallar:

1) $P(\xi \leq 3.5)$, 2) $P(\xi \leq 7)$, 3) $P(\xi > 2)$, 4) $P(\xi \geq 2)$, 5) $P(3 < \xi \leq 8)$.

$$1) P(\xi \leq 3.5) = F(3.5) = \frac{4}{12} = \frac{1}{3}.$$

$$2) P(\xi \leq 7) = F(7) = \frac{9}{12} = \frac{3}{4}.$$

$$3) P(\xi > 2) = 1 - F(2) = \frac{3}{4}.$$

$$4) P(\xi \geq 2) = P(\xi = 2) + P(\xi > 2) = (F(2) - \lim_{h \rightarrow 0} F(2 - h)) + \frac{3}{4} = \left(\frac{3}{12} - \frac{1}{12}\right) + \frac{3}{4} = \frac{11}{12}$$

$$5) P(3 < \xi \leq 8) = F(8) - F(3) = \frac{10}{12} - \frac{4}{12} = 0.5.$$

NOTA: La variable ξ es la de un ejemplo anterior con lanzamiento de moneda y dado.



Tipos de variables aleatorias

Dependiendo de como sea el conjunto soporte podemos distinguir varios tipos de variables aleatorias.

$$\text{Tipos de variables aleatorias.} \left\{ \begin{array}{l} \textit{Discretas} \\ \textit{Continuas} \\ \textit{Mixtas} \end{array} \right\} \left\{ \begin{array}{l} \text{Soporte finito} \\ \text{Soporte numerable} \end{array} \right.$$

Finita: Lanzamiento de un dado. Puntos obtenidos por un equipo de fútbol en una jornada ($S_P = \{0, 1, 3\}$).

Numerable: Días transcurridos hasta realizar un cambio de tarifa de móvil. Número de averías recibidas por un servicio técnico.

Continua: Alcance de una antena. Distancia recorrida por un vehículo en 1 hora. Consumo en litros de combustible en 100Km.

Mixta: a) Distancia alcanzada por un lanzador, si no lo realiza se codifica como -1.

b) Premio obtenido. Se realiza un sorteo entre 10 concursantes, si es seleccionado puede obtener un premio entre 5 y 20.



Variable aleatoria discreta

Definición

Sea (E, \mathcal{A}, P) un espacio probabilístico asociado a un experimento aleatorio y sea X una variable aleatoria. Diremos que se trata de una **variable aleatoria discreta** si su soporte es un conjunto discreto. Es decir, la variable X toma solo un conjunto finito o infinito numerable de valores en \mathbb{R} .

Para este tipo de variable, la forma más simple de definirlas es dar la probabilidad p_i de que tome cada uno de los posibles valores x_i de su soporte.

Definición

Definimos la **función de probabilidad** de una variable aleatoria discreta, mediante:

$$p(x_i) = P(\omega \in \mathcal{A} / X(\omega) = x_i) = P(X = x_i)$$

Ejemplos

Ejemplo

La función de probabilidad de la variable aleatoria: 'Suma de los valores obtenidos al lanzar 2 dados' es:

x_i	$p(x_i)$		x_i	$p(x_i)$
2	$\frac{1}{36}$		8	$\frac{5}{36}$
3	$\frac{2}{36}$		9	$\frac{4}{36}$
4	$\frac{3}{36}$		10	$\frac{3}{36}$
5	$\frac{4}{36}$		11	$\frac{2}{36}$
6	$\frac{5}{36}$		12	$\frac{1}{36}$
7	$\frac{6}{36}$			

Ejemplo

La función de probabilidad de 'Número de caras' al lanzar 3 monedas es: $P(0)=1/8$, $P(1)=3/8$, $P(2)=3/8$, $P(3)=1/8$.

Ejemplo para v.a. numerable

Ejemplo

Sea la variable aleatoria 'Lanzar un dado hasta la primera aparición de un 5'. Hallar la función de probabilidad.

$$P(X=1)=P(\text{sacar } 5 \text{ a la } 1^{\text{a}})=1/6.$$

$$P(X=2)=P(\text{fallar la primera y sacar } 5 \text{ a la } 2^{\text{a}})=(5/6)(1/6).$$

$$P(X=3)=P(\text{fallar 2 veces y sacar } 5 \text{ a la } 3^{\text{a}})=(5/6)(5/6)(1/6)=\\=(5/6)^2(1/6),$$

...

$$P(X = k) = (5/6)^{k-1}(1/6), \text{ para } k = 1, 2, \dots$$

que es la función de probabilidad.

Propiedades

- **La función de probabilidad caracteriza a la variable aleatoria discreta.**
- Si P es una función de \mathbb{R} en $[0, 1]$ que verifica:
 $\sum_i P(X = x_i) = 1$, entonces existe una v.a. que la tiene como función de probabilidad.

Ejemplo

¿Determina la función: $P(X = k) = (2/3)^k(1/3)$ con $k = 0, 1, 2, \dots$, una función de probabilidad?

Para todo k , $0 \leq P(X = k) \leq 1$. Veamos si $\sum_{k=0}^{\infty} P(X = k) = 1$.
 $\sum_k P(X = k) = P(X = 0) + P(X = 1) + \dots = (2/3)^0(1/3) + (2/3)^1(1/3) + (2/3)^2(1/3) + (2/3)^3(1/3) + \dots = 1/3 + (2/3)(1/3) + (2/3)^2(1/3) + \dots = \frac{1/3}{1-2/3} = 1$ luego define una función de probabilidad.

Representación gráfica. Ejemplo.

Tanto la función de probabilidad, como la de distribución de una variable aleatoria pueden ser representadas gráficamente.

Ejemplo

Sea ξ la variable aleatoria 'Número de caras' al lanzar 4 monedas.
Hallar:

- Función de probabilidad.
- Función de distribución.
- Representar ambas gráficamente.

$P(0) = P('FFFF') = 1/16$, $P(1) = P('CFFF' \cup 'FCFF' \cup 'FFCF' \cup 'FFFC') = 4/16$,
 $P(2) = P('CCFF' \cup 'CFCF' \cup 'CFFC' \cup 'FCCF' \cup 'FCFC' \cup 'FFCC') = 6/16$,
 $\dots P(3) = 4/16$, $P(4) = 1/16$.



Ejemplo-cont

La función de probabilidad expresada como tabla y la La función de distribución $F(x) = P(\xi \leq x)$ quedan:

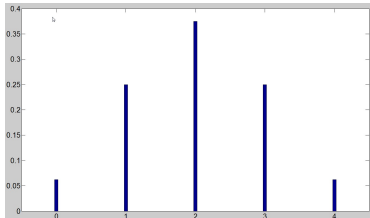
x_i	p_i
0	$\frac{1}{16}$
1	$\frac{4}{16}$
2	$\frac{6}{16}$
3	$\frac{4}{16}$
4	$\frac{1}{16}$

$$P(X) = \begin{cases} 0 & \frac{1}{16} \\ 1 & \frac{4}{16} \\ 2 & \frac{6}{16} \\ 3 & \frac{4}{16} \\ 4 & \frac{1}{16} \\ \text{Otro} & 0 \end{cases}$$

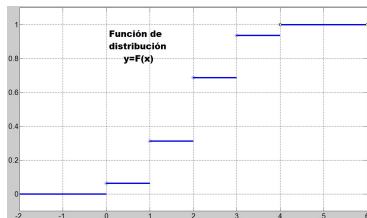
x	$F(x)$
$(-\infty, 0)$	0
$[0, 1)$	$\frac{1}{16}$
$[1, 2)$	$\frac{5}{16}$
$[2, 3)$	$\frac{11}{16}$
$[3, 4)$	$\frac{15}{16}$
$[4, \infty)$	1

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{16} & 0 \leq x < 1 \\ \frac{5}{16} & 1 \leq x < 2 \\ \frac{11}{16} & 2 \leq x < 3 \\ \frac{15}{16} & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

Función de probabilidad



Función de distribución



Esperanza matemática. Caso discreto

Definición

Se llama **esperanza matemática** de la variable aleatoria discreta X a:

$$E(X) = \sum_{x_i \in S_X} x_i P(X = x_i)$$

En el caso de que el soporte S_X sea un conjunto infinito numerable, será necesario que la serie sea absolutamente convergente, esto es

$$\sum_{x_i \in S_X} |x_i| P(X = x_i) < \infty$$

Ejemplo

Hallar la esperanza matemática del 'Número de caras' al lanzar 4 monedas.

$$E(X) = 0P(0) + 1P(1) + 2P(2) + 3P(3) + 4P(4) = 0 \frac{1}{16} + 1 \frac{4}{16} + 2 \frac{6}{16} + 3 \frac{4}{16} + 4 \frac{1}{16} = 2$$

Generalización del concepto

Definición

Sea g una función real y X una variable aleatoria llamamos **esperanza de $g(x)$** a:

$$E(g(X)) = \sum_{x_i \in S_X} g(x_i) P(X = x_i)$$

con la condición de que la serie sea absolutamente convergente:
 $\sum_{x_i \in S_X} |g(x_i)| P(X = x_i)$

Ejemplo

Si X es la v.a. 'Número de caras' al lanzar 4 monedas. Hallar las esperanzas de: 1) X^2 , 2) X^3 , 3) $\sin(\frac{\pi X}{2})$:

$$1) E(X^2) = 0^2 P(0) + 1^2 P(1) + 2^2 P(2) + 3^2 P(3) + 4^2 P(4) = \frac{1 \cdot 4}{16} + \frac{4 \cdot 6}{16} + \frac{9 \cdot 4}{16} + \frac{16 \cdot 1}{16} = 5$$

Ejemplo-cont

$$2) E(X^3) = 0^3P(0) + 1^3P(1) + 2^3P(2) + 3^3P(3) + 4^3P(4) = \\ = \frac{1 \cdot 4}{16} + \frac{8 \cdot 6}{16} + \frac{27 \cdot 4}{16} + \frac{64 \cdot 1}{16} = 14$$

$$3) E\left(\sin\left(\frac{\pi X}{2}\right)\right) = \sin(0)P(0) + \sin\left(\frac{\pi}{2}\right)P(1) + \sin(\pi)P(2) + \\ + \sin\left(\frac{3\pi}{2}\right)P(3) + \sin(2\pi)P(4) = \frac{4}{16} - \frac{4}{16} = 0$$

Ejemplo

Sea ξ la variable aleatoria 'Número de lanzamientos' de un dado hasta la primera aparición de la cara '5'. Hallar $E(\xi)$.

Conocemos que $P(k) = (5/6)^{k-1}(1/6)$ para $k=1,2,\dots$, luego:

$$E(X) = 1(1/6) + 2(5/6)(1/6) + 3(5/6)^2(1/6) + 4(5/6)^3(1/6) + \dots = \\ = \frac{1}{6}[1 + 2(5/6) + 3(5/6)^2 + 4(5/6)^3 + \dots] = \frac{S}{6}$$

$$\begin{array}{rcl} S & = & 1 + 2(5/6) + 3(5/6)^2 + 4(5/6)^3 + \dots \\ \frac{5}{6}S & = & +1(5/6) + 2(5/6)^2 + 3(5/6)^3 + \dots \\ \hline S - \frac{5}{6}S & = & 1 + (5/6) + (5/6)^2 + (5/6)^3 + \dots \Rightarrow \end{array}$$

$$\frac{S}{6} = \frac{1}{1 - 5/6} = 6 \Rightarrow S = 36 \Rightarrow E(X) = 6$$

Variable aleatoria continua

Definición

Se dice que una **variable aleatoria es continua** si su soporte es un intervalo real (finito o infinito) o unión de ellos.

Son variables aleatorias continuas: Temperatura, peso, duración de un componente,

Definimos la probabilidad de un intervalo:

$$P(a < \xi \leq b) = P(\xi^{-1}(a, b]) = P(\omega \in \mathcal{A} / a < \xi(\omega) \leq b)$$

En las variables aleatorias continuas (v.a.c.) se da la circunstancia de que **la probabilidad de un punto es cero**, aunque pueda ocurrir. Por ello:

1) $P(\xi < b) = P(\xi \leq b) = F(b),$

2) $P(\xi > a) = P(\xi \geq a) = 1 - F(a)$

3) $P(a < \xi < b) = P(a < \xi \leq b) = P(a \leq \xi < b) = P(a \leq \xi \leq b) = F(b) - F(a)$



Función de densidad

Definición

*Dada una variable aleatoria continua ξ , decimos que la función $y = f(t)$ real y no negativa es una **función de densidad** asociada a ξ , si verifica: $F(x) = \int_{-\infty}^x f(t)dt$, $\forall x \in \mathbb{R}$*

Propiedades:

- $\int_{-\infty}^{\infty} f(t)dt = 1$ (Consecuencia de $F(+\infty) = 1$)
- $P(a < \xi < b) = \int_a^b f(t)dt$ (Consecuencia de que $P(a < \xi < b) = F(b) - F(a)$)
- **La función de densidad es la derivada de la función de distribución: $f(x) = \frac{dF(x)}{dx}$**
- $P(\xi = a) = \int_a^a f(t)dt = 0$ (La probabilidad de un punto es 0)
- $P(\xi > b) = \int_b^{\infty} f(t)dt = 1 - F(b)$

Interpretación de la función de densidad

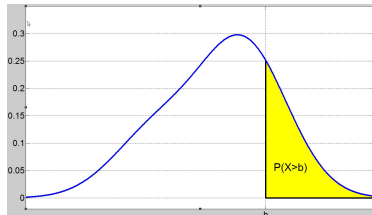
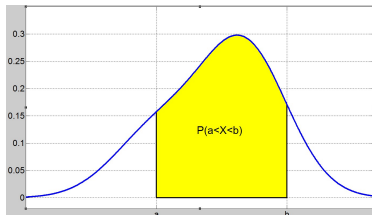
La función de densidad además de caracterizar a la variable aleatoria sirve para calcular probabilidades. Así, las probabilidades:

$$P(a < X < b) = \int_a^b f(t)dt = F(b) - F(a)$$

$$P(X > b) = \int_b^{\infty} f(t)dt = 1 - F(b)$$

$$P(X < a) = \int_{-\infty}^a f(t)dt$$

pueden ser interpretadas como áreas bajo la función de densidad.



Ejemplo

Ejemplo

Sea ξ la v.a.c. determinada por la función de densidad:

$$f(x) = \max\{0, a - |2 - x|\}.$$

- Determinar el valor de a para que $f(t)$ sea una función de densidad.
- Hallar la función de distribución.
- Hallar las probabilidades de los sucesos a) $P(\xi \leq 1.5)$, b) $P(\xi > 2.3)$, c) $P(1.1 \leq \xi \leq 1.7)$ d) $P(1.5 \leq \xi \leq 2.5)$

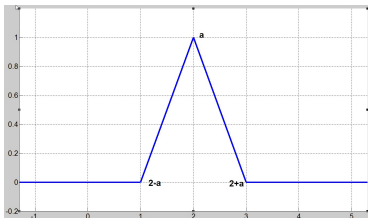
a) La forma de la función $y = f(x)$ es triangular con máximo en $x=2$ y altura a , y base $(2 - a, 2 + a)$. Para que el área valga 1:

$$S = Bh/2 = (2a)a/2 = a^2 = 1 \Rightarrow a = 1 \text{ y resulta la figura.}$$

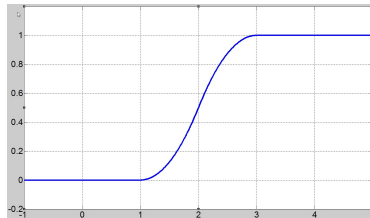
$$b) f(x) = \max\{0, 1 - |2 - x|\} = \begin{cases} x - 1 & x \in (1, 2] \\ 3 - x & x \in (2, 3) \\ 0 & x \notin (1, 3) \end{cases}$$



Ejemplo-cont



Función de densidad de ξ



Función de distribución de ξ

Integramos para hallar la función de distribución: (la constante de integración se ajusta para que sea continua, así $F(1)=0$, $F(2)=0.5$, $F(3)=1$).

$$F(x) = \begin{cases} 0 & x \leq 1 \\ \frac{x^2}{2} - x + \frac{1}{2} & 1 < x \leq 2 \\ 3x - \frac{x^2}{2} - 3.5 & 2 < x < 3 \\ 1 & x \leq 3 \end{cases}$$

Ejemplo-cont2

- a) $P(\xi \leq 1.5) = \int_{-\infty}^{1.5} f(t)dt = \int_1^{1.5} (t-1)dt = \{\text{Más fácil}\} = F(1.5) = 0.125$
- b) $P(\xi > 2.3) = \int_{2.3}^{\infty} f(t)dt = \int_{2.3}^{\infty} (3-t)dt = \{\text{Más fácil}\} = 1 - F(2.3) = 1 - \left[3(2.3) - \frac{2.3^2}{2} - 3.5 \right] = 0.245$
- c) $P(1.1 \leq \xi \leq 1.7) = \int_{1.1}^{1.7} f(t)dt = \int_{1.1}^{1.7} (t-1)dt = \{\text{Más fácil}\} = F(1.7) - F(1.1) = 0.245 - 0.005 = 0.24$
- d) $P(1.5 \leq \xi \leq 2.5) = \int_{1.5}^{2.5} f(t)dt = \int_{1.5}^2 (t-1)dt + \int_2^{2.5} (3-t)dt = \{\text{Más fácil}\} = F(2.5) - F(1.5) = 0.875 - 0.125 = 0.75$

Esperanza matemática. Caso Continuo

Definición

Se llama **esperanza matemática** de la v.a.c. X a:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

Solo quedará definida cuando la integral sea absolutamente convergente, es decir: $\int_{-\infty}^{\infty} |x|f(x)dx < \infty$

Ejemplo

Hallar la esperanza matemática de la función de la variable aleatoria ξ del ejemplo anterior.

La función de densidad era: $f(x) = \begin{cases} x-1 & x \in (1, 2) \\ 3-x & x \in (2, 3) \\ 0 & x \notin (1, 3) \end{cases}$, entonces:

$$E(\xi) = \int_{-\infty}^{\infty} xf(x)dx = \int_1^2 x(x-1)dx + \int_2^3 x(3-x)dx = \left[\frac{x^3}{3} - \frac{x^2}{2} \right]_1^2 + \left[3\frac{x^2}{2} - \frac{x^3}{3} \right]_2^3 = 2$$

Generalización del concepto

Definición

Sea g una función real y X una variable aleatoria continua llamamos **esperanza de $g(x)$** a:

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

con la condición de que la integral sea absolutamente convergente:

$$\int_{-\infty}^{\infty} |g(x)|f(x)dx$$

Ejemplo

Hallar $E(\sin(\xi))$

$$E[\sin(\xi)] = \int_{-\infty}^{\infty} \sin(x)f(x)dx = \int_1^2 \sin(x)(x-1)dx + \int_2^3 \sin(x)(3-x)dx \approx 0.836$$

Propiedades de la esperanza matemática

Las siguientes propiedades son válidas para cualquier variable aleatoria:

- Es la media de la variable aleatoria: $\bar{X} = E(X)$.
- $E(k)=k$. (La esperanza matemática de una constante k es k .)
- $E(aX+b)=aE(X)+b$ (Transformación afín)
- $E(aX + bY) = aE(X) + bE(Y)$ (Linealidad)
- Si son independientes se verifica: $E(XY) = E(X)E(Y)$

Ejemplo

Sea X la v.a. 'Número de caras' al lanzar 4 monedas e Y 'Puntos obtenidos' al lanzar un dado. Hallar $E(Y)$, $E(X+Y)$ y $E(3Y-2X+7)$.

$$E(Y) = 1\frac{1}{6} + 2\frac{1}{6} + \dots + 6\frac{1}{6} = \frac{21}{6} = 3.5$$

$$E(X + Y) = E(X) + E(Y) = 2 + 3.5 = 5.5,$$

$$E(3Y - 2X + 7) = 3E(Y) - 2E(X) + 7 = 3(3.5) - 2(2) + 7 = 13.5$$



Ejemplo

Ejemplo

Sea ξ la variable aleatoria continua dada por la función de

$$\text{distribución: } F(x) = \begin{cases} 0 & x \leq 0 \\ x^4 & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}.$$

Hallar: a) $E(\xi)$, b) $E(3\xi - 2)$, c) $E(\sin(\xi))$

a) Lo primero será hallar la función de densidad, $f(x) = \frac{dF(x)}{dx} = 4x^3$ para

$x \in [0, 1]$. Luego: $E(\xi) = \int_0^1 x(4x^3)dx = \left[4\frac{x^5}{5}\right]_0^1 = \frac{4}{5} = 0.8$

$$E(3\xi - 2) = 3E(\xi) - 2 = 3(0.8) - 2 = 0.4$$

$$E(\sin(\xi)) = \int_0^1 \sin(x)(4x^3)dx = 0.7084.$$

Cálculo realizado con la instrucción Matlab:

```
>> h=inline('sin(x).*4.*x.^3'),quad(h,0,1)
```

aunque podría haberse calculado mediante integración por partes.



Momentos

Definición

Llamamos **momento de orden r respecto al punto ' a ' de la variable aleatoria X** , a

$$M_r(a) = E((X - a)^r)$$

Cuando $a=0$ se denominan **momentos ordinarios de orden r** :

$$m_r = E(X^r).$$

Cuando $a = \bar{X} = E(X)$ se denominan **momentos centrales de orden r** :

$$\mu_r = E((X - E(X))^r)$$

Definición

Llamamos **Varianza de una variable aleatoria X a su momento central de orden 2**:

$$\mu_2 = E((X - E(X))^2) = E(X^2) - [E(X)]^2 = m_2 - \bar{X}^2$$

Relaciones entre momentos

Definición

A la raíz cuadrada de la varianza la llamamos **desviación típica** de la variable aleatoria: $\sigma_x = +\sqrt{V(X)}$

Propiedades de los momentos:

- $m_0 = 1, m_1 = E(X) = \bar{X}$
- $m_2 = V(x) + \bar{X}^2.$
- $\mu_0 = 1, \mu_1 = 0,$
- $\mu_2 = V(x) = m_2 - \bar{X}^2$
- $\mu_3 = m_3 - 3m_2\bar{X} + 2\bar{X}^3$
- $\mu_4 = m_4 - 4m_3\bar{X} + 6m_2\bar{X}^2 - 3\bar{X}^4$

Parámetros

Parámetros de tendencia central: Son la media, moda y mediana.

- **Media:** Es la esperanza de X .
- **Moda:** Valor máximo de la función de probabilidad (discretas) o de la de densidad (continuas).
- **Mediana:** Valor x tal que $F(x) = 0.5$

Al igual que en descriptiva podemos hablar de cuantiles, cuartiles, El cuantil $c \in (0, 1)$ es el valor x tal que $F(x) = c$.

Medidas de dispersión: Podemos hablar de rango, desviación típica, varianza.

Coeficiente de variación: $CV = \frac{\sigma_x}{|\bar{X}|}$

Medidas de forma: (Coeficientes de Fisher)

Sesgo: $\gamma_1 = \frac{\mu_3}{\sigma^3}$

Curtosis: $\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$

Tipos de variables aleatorias bidimensionales

Cada una de las variables X e Y que forman la variable bidimensional podrá ser: discreta, continua o mixta. Sin embargo vamos a estudiar solamente:

- **Variables aleatorias bidimensionales discretas:** Cuando ambas son discretas (soporte finito o infinito numerable).
- **Variables aleatorias bidimensionales continuas:** Si ambas son continuas.



Variables aleatorias bidimensionales discretas

Definición

*Podemos caracterizar la variable aleatoria bidimensional mediante la **Función de probabilidad conjunta**, donde*

$$P(X = x_i, Y = y_j) = p_{ij}$$

La función de probabilidad viene usualmente expresada en forma tabular.

$x \backslash y$	y_1	y_2	\cdots	y_j	\cdots	y_L
x_1	p_{11}	p_{12}	\cdots	p_{1j}	\cdots	p_{1L}
x_2	p_{21}	p_{22}	\cdots	p_{2j}	\cdots	p_{2L}
\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
x_i	p_{i1}	p_{i2}	\cdots	p_{ij}	\cdots	p_{iL}
\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
x_k	p_{k1}	p_{k2}	\cdots	p_{kj}	\cdots	p_{kL}

Variables discretas - cont.

Podremos definir todo lo que se ha indicado para tablas de frecuencias bidimensionales:

- Variables aleatorias marginales: (X e Y)
- Variables aleatorias condicionadas: ($X/Y = y_j$) e ($Y/X = x_i$)
- Conceptos de dependencia, independencia y dependencia funcional.
- Momentos bidimensionales respecto a (a, b):

$$M_{rs}(a, b) = E[(X - a)^r (Y - b)^s] = \sum_i \sum_j (x_i - a)^r (y_j - b)^s p_{ij}$$

Si (a, b) = (0,0) se llaman momentos ordinarios:

$$m_{rs} = E(X^r Y^s)$$

Si (a, b) = (\bar{X}, \bar{Y}) se llaman momentos centrales:

$$\mu_{rs} = E[(X - \bar{X})^r (Y - \bar{Y})^s]$$

- Conceptos de media, varianza, covarianza,

Ejemplo bidimensional discreta

1) Veamos como se calcula alguno de los términos de la tabla, por ejemplo para $X=1, Y=2$: Los casos que lo producen deben tener 1 cara en las dos primeros lanzamientos y otra cara en los dos últimos. Así, $\{X = 1, Y = 2\} = \{CFCF, CFFC, FCCF, FCFC\}$ y su probabilidad será $4/16$.

$X \backslash Y$	0	1	2	3	4	
0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	0	0	$\frac{4}{16}$
1	0	$\frac{2}{16}$	$\frac{4}{16}$	$\frac{2}{16}$	0	$\frac{8}{16}$
2	0	0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{4}{16}$
	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$	1

2) Tendría que verificarse que p_{ij} fuese el producto de las marginales para todo i, j , y falla el primero: $P(X = 0, Y = 0) \neq P(X = 0)P(Y = 0)$, luego **no son independientes** existiendo dependencia estadística.



Ejemplo bidimensional discreta-2

3) La función de distribución expresada como tabla resulta:

$X \backslash Y$	$(-\infty, 0)$	$[0, 1)$	$[1, 2)$	$[2, 3)$	$[3, 4)$	$[4, \infty)$
$(-\infty, 0)$	0	0	0	0	0	0
$[0, 1)$	0	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{4}{16}$	$\frac{4}{16}$
$[1, 2)$	0	$\frac{1}{16}$	$\frac{5}{16}$	$\frac{10}{16}$	$\frac{12}{16}$	$\frac{12}{16}$
$[2, \infty)$	0	$\frac{1}{16}$	$\frac{5}{16}$	$\frac{11}{16}$	$\frac{15}{16}$	1

Para calcular el valor de una casilla, por ejemplo $x \in [1, 2)$, $y \in [2, 3)$ miramos en la tabla de la función de probabilidad y sumamos todas las probabilidades de las casillas con $X \leq 1$ e $Y \leq 2$, es decir:

$$P(0,0)+P(0,1)+P(0,2)+P(1,1)+P(1,2)=10/16.$$

Ejemplo bidimensional discreta-2

4) Para las medias y varianzas podemos usar las marginales (función de probabilidad)

Media de X: $E(X) = 0 \cdot \frac{4}{16} + 1 \cdot \frac{8}{16} + 2 \cdot \frac{4}{16} = 1$

Media de Y: $E(Y) = 0 \cdot \frac{1}{16} + 1 \cdot \frac{4}{16} + 2 \cdot \frac{6}{16} + 3 \cdot \frac{4}{16} + 4 \cdot \frac{1}{16} = 2$

Varianza de X: $m_{2,0} = E(X^2) = 0^2 \cdot \frac{4}{16} + 1^2 \cdot \frac{8}{16} + 2^2 \cdot \frac{4}{16} = \frac{24}{16} \Rightarrow$

$V(X) = \frac{24}{16} - 1^2 = 0.5$

Varianza de Y:

$m_{0,2} = E(Y^2) = 0^2 \cdot \frac{1}{16} + 1^2 \cdot \frac{4}{16} + 2^2 \cdot \frac{6}{16} + 3^2 \cdot \frac{4}{16} + 4^2 \cdot \frac{1}{16} = \frac{80}{16} \Rightarrow$

$V(Y) = \frac{80}{16} - 2^2 = 1$

Covarianza:

$m_{1,1} = E(XY) = 0 \cdot 0 \cdot \frac{1}{16} + 0 \cdot 1 \cdot \frac{2}{16} + 0 \cdot 2 \cdot \frac{1}{16} + 1 \cdot 1 \cdot \frac{2}{16} + 1 \cdot 2 \cdot \frac{4}{16} + 1 \cdot 3 \cdot \frac{2}{16} + 2 \cdot 2 \cdot \frac{1}{16} + 2 \cdot 3 \cdot \frac{2}{16} + 2 \cdot 4 \cdot \frac{1}{16} = \frac{40}{16} \Rightarrow$
Cov(X, Y) = $\frac{40}{16} - 1 \cdot 2 = 0.5$



Ejemplo bidimensional discreta-3

5) $P(X^2 + Y \leq 3)$: Las casillas de la tabla de la función de probabilidad que verifican la condición, (con $p_{ij} > 0$), son: $(0,0), (0,1), (0,2), (1,1)$ y $(1,2)$. $P(\{(0,0) \cup (0,1) \cup (0,2) \cup (1,1) \cup (1,2)\}) = \frac{10}{16}$

$P(2X + Y > 5)$: Las casillas de la tabla de la función de probabilidad que verifican la condición, (con $p_{ij} > 0$), son: $(2,2), (2,3)$ y $(2,4)$.
 $P(\{(2,2) \cup (2,3) \cup (2,4)\}) = \frac{4}{16} = 0.25$

6) Podemos usar la fórmula: $Y - \bar{Y} = b(X - \bar{X})$, donde sabemos que: $\bar{X} = 1, \bar{Y} = 2$. $b = \frac{Cov}{V(x)} = \frac{0.5}{0.5} = 1 \Rightarrow m_{Y/X} = 1$ $b' = \frac{Cov}{V(y)} = \frac{0.5}{1} = 0.5 \Rightarrow m_{X/Y} = 2$ y las rectas son:

$$Y/X : \quad Y - 2 = 1(X - 1) \Rightarrow Y = X + 1$$

$$X/Y : \quad Y - 2 = 2(X - 1) \Rightarrow X = 0.5Y$$

7) $\rho = \sqrt{bb'} = \sqrt{1 \cdot 0.5} \approx 0.7071$



Variables bidimensionales continuas

Definición

Llamamos **función de densidad bidimensional** de la variable aleatoria (X, Y) a una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, tal que sea:

- Integrable y no negativa.

-

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

- $\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$ en todos los puntos en donde esté definida esa derivada segunda.

Dada la función de densidad de la variable aleatoria, la probabilidad de $(a \leq X \leq b) \cap (c \leq Y \leq d)$ queda definida por:

$$P(\{(a \leq X \leq b) \cap (c \leq Y \leq d)\}) = \int_c^d \int_a^b f(x, y) dx dy$$

Variables continuas-cont

- Variables aleatorias marginales: (X e Y)**

$$F_1(x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x, y) dy dx, \quad F_2(y) = \int_{-\infty}^{\infty} \int_{-\infty}^y f(x, y) dx dy$$

- Variables aleatorias condicionadas: (X/Y ≤ b) e (Y/X ≤ a)**

$$F_{X/Y \leq b}(x) = \frac{P(X \leq x, Y \leq b)}{P(Y \leq b)} = \frac{\int_{-\infty}^x \int_{-\infty}^b f(x, y) dy dx}{\int_{-\infty}^{\infty} \int_{-\infty}^b f(x, y) dy dx} = \frac{F(x, b)}{F(\infty, b)} = \frac{F(x, b)}{F_2(b)}$$

$$F_{Y/X \leq a}(y) = \frac{P(X \leq a, Y \leq y)}{P(X \leq a)} = \frac{\int_{-\infty}^y \int_{-\infty}^a f(x, y) dx dy}{\int_{-\infty}^{\infty} \int_{-\infty}^a f(x, y) dx dy} = \frac{F(a, y)}{F(a, \infty)} = \frac{F(a, y)}{F_1(a)}$$

- Variables aleatorias condicionadas: (X/Y = b) e (Y/X = a)**

Ahora los sucesos $Y = b$ y $X = a$ tienen probabilidad 0 (no está definida la condicionada). Tomamos un intervalo de amplitud ϵ y límites cuando $\epsilon \rightarrow 0$:

$$F_{X/Y=b}(x) = \lim_{\epsilon \rightarrow 0} \frac{P(X \leq x, b \leq Y \leq b+\epsilon)}{P(b \leq Y \leq b+\epsilon)} = \lim_{\epsilon \rightarrow 0} \frac{\int_{-\infty}^x \int_b^{b+\epsilon} f(x, y) dy dx}{\int_{-\infty}^{\infty} \int_b^{b+\epsilon} f(x, y) dy dx} =$$

$$= \frac{\int_{-\infty}^x f(x, b) dx}{\int_{-\infty}^{\infty} f(x, b) dx}$$

$$F_{Y/X \leq a}(y) = \lim_{\epsilon \rightarrow 0} \frac{P(a \leq X \leq a+\epsilon, Y \leq y)}{P(a \leq X \leq a+\epsilon)} = \lim_{\epsilon \rightarrow 0} \frac{\int_{-\infty}^y \int_a^{a+\epsilon} f(x, y) dx dy}{\int_{-\infty}^{\infty} \int_a^{a+\epsilon} f(x, y) dx dy} =$$

$$\frac{\int_{-\infty}^y f(a, y) dy}{\int_{-\infty}^{\infty} f(a, y) dy}$$



Variables continuas-cont2

- **Conceptos de dependencia, independencia y dependencia funcional.**
- **Momentos bidimensionales respecto a (a,b):**

$$M_{rs}(a, b) = E[(X - a)^r (Y - b)^s] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - a)^r (y - b)^s f(x, y) dx dy$$

Si (a,b)=(0,0) se llaman momentos ordinarios:

$$m_{rs} = E(X^r Y^s)$$

Si (a, b) = (\bar{X} , \bar{Y}) se llaman momentos centrales:

$$\mu_{rs} = E[(X - \bar{X})^r (Y - \bar{Y})^s]$$

- **Conceptos de media, varianza, covarianza, ...**



Ejemplo variable bidimensional continua

Ejemplo

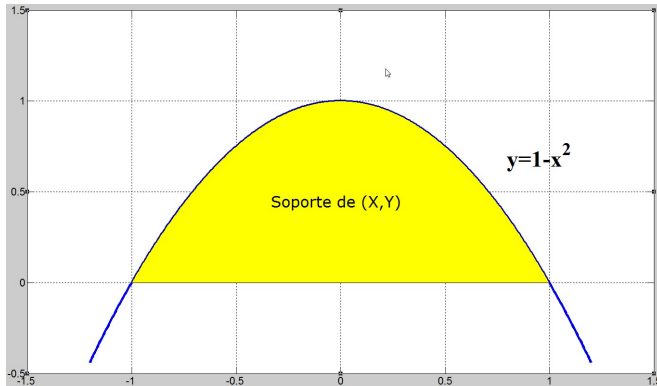
Sea la variable aleatoria (X, Y) con función de densidad:

$$f(x, y) = \begin{cases} c(x^2 + y) & 0 \leq y \leq 1 - x^2 \\ 0 & \text{en el resto} \end{cases}$$

- 1 Dibuje la región de \mathbb{R}^2 que representa el soporte de la variable aleatoria.
- 2 Determine el valor de c .
- 3 Hallar $P(X \geq 0)$.
- 4 Hallar la probabilidad de que la variable aleatoria tome valor en el cuadrado de lado 1 y centro el origen de coordenadas.
- 5 Determinar la distribución marginal de X .
- 6 ¿Son X e Y variables independientes?
- 7 Calcular la recta de regresión de Y/X y determinar la bondad del ajuste.

Ejemplo bidimensional continua

1) El soporte será la región donde $f(x, y) > 0$ y está representado en la figura:



Ejemplo bidimensional continua-2

2) Las funciones de densidad deben verificar:

$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$ que aquí queda:

$$\int_{-1}^1 \int_0^{1-x^2} c(x^2 + y) dy dx = 1$$

$$\begin{aligned} \int_{-1}^1 c \left[x^2 y + \frac{y^2}{2} \right]_0^{1-x^2} dx &= \int_{-1}^1 c \left[x^2(1-x^2) + \frac{(1-x^2)^2}{2} \right] dx = \\ &= \frac{c}{2} \int_{-1}^1 1 - x^4 dx = \frac{c}{2} \left[x - \frac{x^5}{5} \right]_{-1}^1 = \frac{4c}{5} = 1 \Rightarrow c = 1.25 \end{aligned}$$

$$3) P(X \geq 0) = 1.25 \int_0^1 \int_0^{1-x^2} (x^2 + y) dy dx =$$

$$= 1.25 \int_0^1 \left[x^2 y + \frac{y^2}{2} \right]_0^{1-x^2} dx = \frac{1.25}{2} \int_0^1 1 - x^4 dx = \frac{1}{2}$$

Ejemplo bidimensional continua-3

4) Será el cuadrado C de extremos $(-0.5, -0.5)$, $(0.5, -0.5)$, $(0.5, 0.5)$ y $(-0.5, 0.5)$. Su intersección con el soporte es el rectángulo con $y > 0$, es decir: $(-0.5, 0)$, $(0.5, 0)$, $(0.5, 0.5)$ y $(-0.5, 0.5)$.

$$\begin{aligned} P(C) &= 1.25 \int_{-0.5}^{0.5} \int_{-0.5}^{0.5} [x^2 + y] dy dx = 1.25 \int_{-0.5}^{0.5} \int_0^{0.5} \left[x^2 y + \frac{y^2}{2} \right]_0^{0.5} dx = \\ &= \frac{1.25}{2} \int_{-0.5}^{0.5} (x^2 + 0.25) dx = \frac{1.25}{2} \left[\frac{x^3}{3} + 0.25x \right]_{-0.5}^{0.5} \approx 0.2083 \end{aligned}$$

5) Marginal de X: Para $-1 \leq x \leq 1$ se obtiene:

$$\begin{aligned} F_1(x) &= 1.25 \int_{-1}^x \int_0^{1-x^2} (x^2 + y) dy dx = 1.25 \int_{-1}^x \left[x^2 y + \frac{y^2}{2} \right]_0^{1-x^2} dx = \\ &= \frac{1.25}{2} \int_{-1}^x (1 - x^4) dx = \frac{1.25}{2} \left[x - \frac{x^5}{5} \right]_{-1}^x = \frac{1}{8} (5x - x^5 + 4) \end{aligned}$$

Si $x \leq -1$, $F_1(x) = 0$, y si $x \geq 1$, $F_1(x) = 1$

Ejemplo bidimensional continua-4

6) Existen varias formas de verlo, por ejemplo, calculando la marginal de y y viendo que la función de densidad conjunta es el producto de las marginales. Pero es evidente que no son independientes, pues viendo la forma del soporte las condicionadas de $Y/X = 0.99$ e $Y/X = 0$ son diferentes. En la primera $f(0.99, 0.5) = 0$ y en la segunda $F(0.99, 0.5) > 0$.

7) Sin embargo vamos a comprobar que verifica $E(XY) = E(X)E(Y)$ y por tanto son incorreladas ($\text{Cov}=0$).

$$\begin{aligned} E(X) &= \frac{5}{4} \int_{-1}^1 \int_0^{1-x^2} x(x^2 + y) dy dx = \frac{5}{4} \int_{-1}^1 \left[x(x^2 y + \frac{y^2}{2}) \right]_0^{1-x^2} dx = \\ &= \frac{5}{8} \int_{-1}^1 (-x^5 + x) dx = \frac{5}{8} \left[-\frac{x^6}{6} + \frac{x^2}{2} \right]_{-1}^1 = 0 \end{aligned}$$

Ejemplo bidimensional continua-5

$$\begin{aligned} E(Y) &= \frac{5}{4} \int_{-1}^1 \int_0^{1-x^2} y(x^2 + y) dy dx = \frac{5}{4} \int_{-1}^1 \left[(x^2 \frac{y^2}{2} + \frac{y^3}{3}) \right]_0^{1-x^2} dx = \\ &= \frac{5}{24} \int_{-1}^1 (x^6 - 3x^2 + 2) dx = \frac{10}{21} \end{aligned}$$

$$\begin{aligned} E(XY) &= \frac{5}{4} \int_{-1}^1 \int_0^{1-x^2} xy(x^2 + y) dy dx = \frac{5}{4} \int_{-1}^1 \left[(x^3 \frac{y^2}{2} + x \frac{y^3}{3}) \right]_0^{1-x^2} dx = \\ &= \frac{5}{24} \int_{-1}^1 (x^7 - 3x^3 + 2x) dx = 0 \end{aligned}$$

Cov=E(XY)-E(X)E(Y)=0 $\Rightarrow r = 0, b = 0, m_{Y/X} = 0$ y la recta de regresión de Y/X queda: $Y - \frac{10}{21} = 0(X - 0)$ y es la recta **Y** = $\frac{10}{21}$ paralela al eje OX.

La recta X/Y será la recta vertical **X** = **0** pues $m_{Y/X} = \frac{1}{r} \frac{\sigma_y}{\sigma_x}$ con $r = 0$.



Ejemplo bidimensional continua-6

La recta X/Y será la recta vertical $\mathbf{X} = \mathbf{0}$ pues $m_{X/Y} = \frac{1}{r} \frac{\sigma_y}{\sigma_x} = \infty$ con $r = 0$.

$$Y - \frac{10}{21} = \infty(X - 0) \Rightarrow \mathbf{X} = \mathbf{0}$$

Es la recta de pendiente ∞ que pasa por $(0, \frac{10}{21})$.

Bondad del ajuste: El ajuste es muy malo pues $r = 0$ (incorreladas), esto quiere decir que mediante la regresión lineal, el conocer una variable no aporta conocimiento sobre la otra. $V_r = (1 - r^2)V(y) = V(y)$