Estadística

Tema 3. Probabilidad y variable aleatoria

María Dolores Frías Domínguez
Jesús Fernández Fernández
Carmen María Sordo

Departamento de Matemática Aplicada y Ciencias de la Computación

TEMA 3: Probabilidad y v. aleatoria

Probabilidad

- Experimento aleatorio y definiciones de probabilidad
- Probabilidad condicionada e independencia de sucesos
- Teorema de la probabilidad total y de Bayes

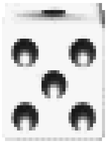
Variable aleatoria

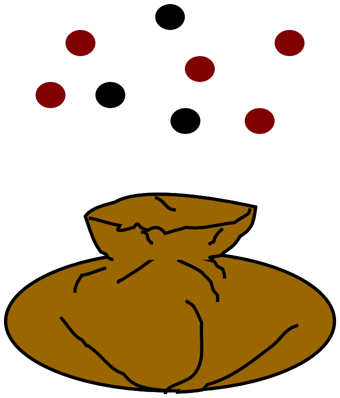
- Definición
- Funciónes de probabilidad, densidad y distribución
- Medidas de una variable aleatoria

Experimentos

Aleatorio: Cualquier situación que, realizada en las mismas condiciones, proporcione un resultado imposible de predecir, conociendo de antemano cuáles son todos los posibles resultados.

La probabilidad aparece cuando
se analizan experimentos aleatorios

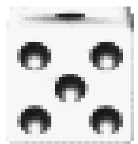


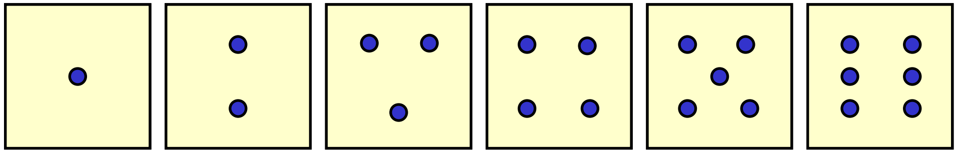


Determinista: Cualquier experimento que al realizarse bajo condiciones específicas conduce siempre al mismo resultado: (lo contrario de un experimento aleatorio).

Suceso Elemental y Espacio Muestral

Suceso elemental: cada uno de los posibles resultados del experimento

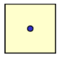
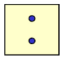
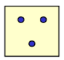
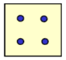
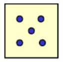
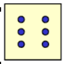




Se suele designar de forma genérica con la letra ω.

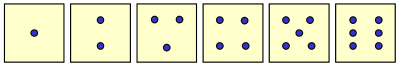
Siempre existe más de un posible resultado.

Pueden ser de cualquier naturaleza (no necesariamente números).

$\omega_1=\{$  $\}$ $\omega_2=\{$  $\}$ $\omega_3=\{$  $\}$ $\omega_4=\{$  $\}$ $\omega_5=\{$  $\}$ $\omega_6=\{$  $\}$

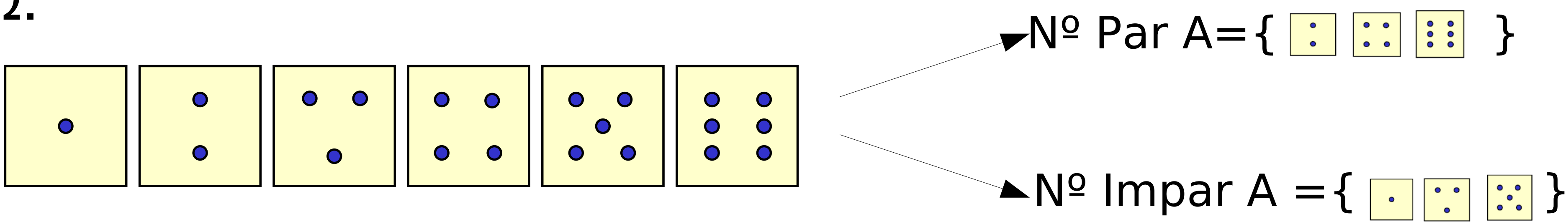
El conjunto de todos los sucesos elementales se denomina **espacio muestral**.

Se suele designar de forma genérica con la letra Ω

$\Omega=\{$  $\}$

Suceso Compuesto, Seguro e Imposible

Compuesto: Cualquier subconjunto del espacio muestral Ω.



Un suceso compuesto ocurre si y solo si ocurre alguno de los sucesos elementales que lo forman.

Seguro: aquel que siempre ocurre. El suceso seguro es $\boldsymbol{\Omega}$

Imposible: aquel que nunca ocurre.El suceso imposible es el conjunto vacío (**Ø**)

Relaciones entre Sucesos

Dos sucesos son **compatibles** si pueden ocurrir en la misma realización del experimento.

$$A \cap B \neq \varnothing$$

En caso contrario se dice que son incompatibles

El suceso **opuesto** de uno dado A es el suceso que ocurre si y solo si no ocurre A.

$$A \cup A^c = \Omega \qquad A \cap A^c = \varnothing$$

Suceso complementario A^c

Ejemplo

$\Omega = \{1,2,3,4,5,6\}$

$A=\{2,4,6\}$ y $B=\{1,2\}$ compatibles, $A^c=\{1,3,5\}$ $B^c=\{3,4,5,6\}$

$A=\{2,4,6\}$ y $C=\{1,3,5\}$ incompatibles, $C^c=A$

María Dolores Frías, Jesús Fernández y Carmen María Sordo

Probabilidad

Clásica: Hay un número finito de resultados igualmente verosímiles que cubren todas las posibilidades

$$P(A) = \frac{resultados\ favorables}{resultados\ posibles}$$

Limitado a juegos de azar por la condición exigida de que todos los resultados deben ser igualmente verosímiles.

Probabilidad

Frecuentista: Situaciones prácticas que pueden repetirse en condiciones idénticas. Probabilidad verificable mediante experimentación.

Variabilidad aleatoria a corto plazo: Los resultados varían de una realización a otra de forma impredecible.

Regularidad a la larga: El cociente entre el número de veces que ocurre el suceso A (n_A) y el número total de repeticiones del experimento n parece converger a un límite cuando n crece.

Probabilidad P(A) del suceso A es el límite de la frecuencia relativa n_A/n del suceso A cuando el número de repeticiones n crece indefinidamente.

Probabilidad

Dado un espacio muestral Ω, una probabilidad P es una función que hace corresponder a todo suceso del espacio muestral un número real no negativo y que verifica los siguiente axiomas:

1. La probabilidad del suceso seguro es la unidad

$$P(\Omega) = 1$$

2. Si $A_1, A_2, \ldots A_n$ es una sucesión de sucesos disjuntos dos a dos ($A_i \cap A_j = \emptyset \quad i \neq j$), entonces la probabilidad de la unión de todos ellos es la suma de sus probabilidades:

$$P(A_1 \cup A_2 \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots P(A_n)$$

Para cualquier suceso A definiremos su probabilidad P(A) como una medida de A que tomará un valor entre 0 y 1.

Propiedades de la Probabilidad

1. La probabilidad de un suceso más la de su complementario suman 1

$$P(A^c)=1-P(A)$$

2. La probabilidad del suceso imposible es nula

$$P(\emptyset)=0$$

3. Si A es un suceso que incluye al suceso B ($B\subseteq A$), se tiene

$$P(B) \leq P(A) \qquad\qquad P(A-B)\equiv P(A\cap B^c)=P(A)-P(B)$$

4. La probabilidad de cualquier suceso no puede ser mayor de 1

$$P(A)\leq 1,$$

5. Si A y B son dos sucesos cualesquiera se cumple (regla de la adición):

$$P(A\cup B) = P(A)+P(B)-P(A\cap B)$$

Ejercicio

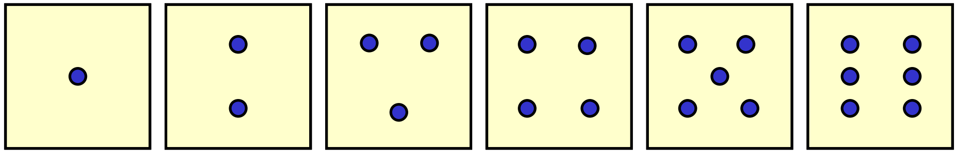
Sean A y B sucesos de un mismo espacio muestral tales que:

$P(B)=0.4 \quad P(A \cup B)=0.7 \quad P(A \cap B)=0.3$

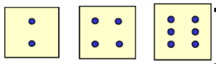
Calcular: $P(A^c)$

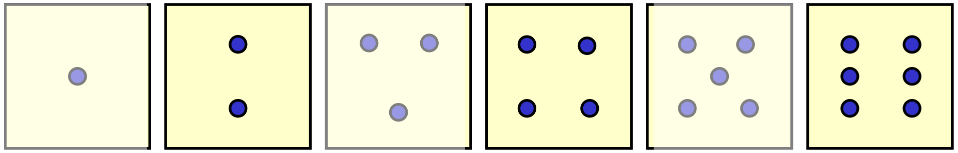
Probabilidad Condicionada

Al lanzar un dado



se sabe que ha salido un número par.

B={}



¿Cuál es la probabilidad de un suceso A conociendo está información?

Probabilidad Condicionada

La **probabilidad condicionada, P(A|B)**, "mide" la probabilidad relativa de A con respecto al espacio reducido B.

Sea un espacio muestral y un suceso B tal que P(B) > 0. Se define:

$$P(A|B) \equiv \frac{P(A \cap B)}{P(B)}$$

$P(\cdot|B)$ es una nueva medida de probabilidad

A|B **NO ES UN SUCESO**, | **NO ES UNA OPERACIÓN ENTRE SUCESOS!**

Regla del producto:

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

$$P(A_1 \cap \ldots \cap A_n) = P(A_1)P(A_2|A_1)\ldots P(A_n|A_1 \cap \ldots \cap A_{n-1})$$

Independencia de Sucesos

Sea un espacio muestal y los sucesos A, B tales que $P(B) > 0$, se dice que **A es independiente de B** si el conocimiento de que ha ocurrido B no altera la probabilidad de A:

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

Sea un espacio muestral y un conjunto de dos o más sucesos A_1 ,..., A_n. Se dice que estos n sucesos son independientes entre sí si se verifica:

$$P(A_1 \cap \ldots \cap A_n) = P(A_1)\, P(A_2) \ldots P(A_n)$$

y, además, todos los sucesos de cualquier selección formada por n-1 sucesos tomados de entre A_1 ,..., A_n son independientes entre sí.

Ejercicio

De una baraja de 40 cartas se extrae una al azar. Comprobar cuales de los siguientes pares de sucesos son independientes:

1. A={rey}, B={espadas}
2. A={figuras}, B={espadas}
3. A={rey}, B={figuras}

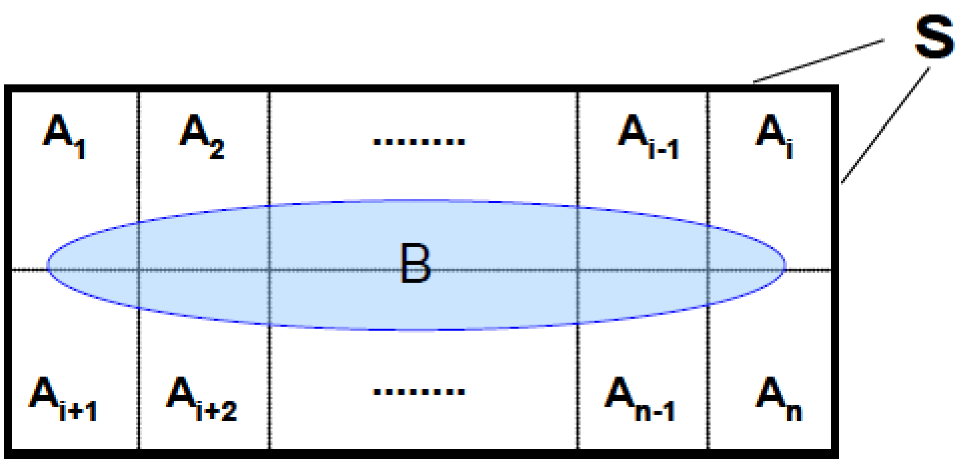
Ejercicio

Suponiendo que $P(A) = 1/3$, $P(B) = 1/4$ y $P(A \cap B) = 1/10$. Hallar: $P(A|B)$, $P(B|A)$, $P(A|A \cap B)$, $P(A \cap B|B)$ y $P(A|B^c)$.

Teorema de la Probabilidad Total

Supongamos que sobre el espacio muestral S tenemos una partición A_i, con i = 1, ..., n. Su unión es el total y los sucesos son incompatibles dos a dos.

Esto significa que cualquier suceso elemental de S necesariamente debe estar en uno y sólo uno de los eventos A_i



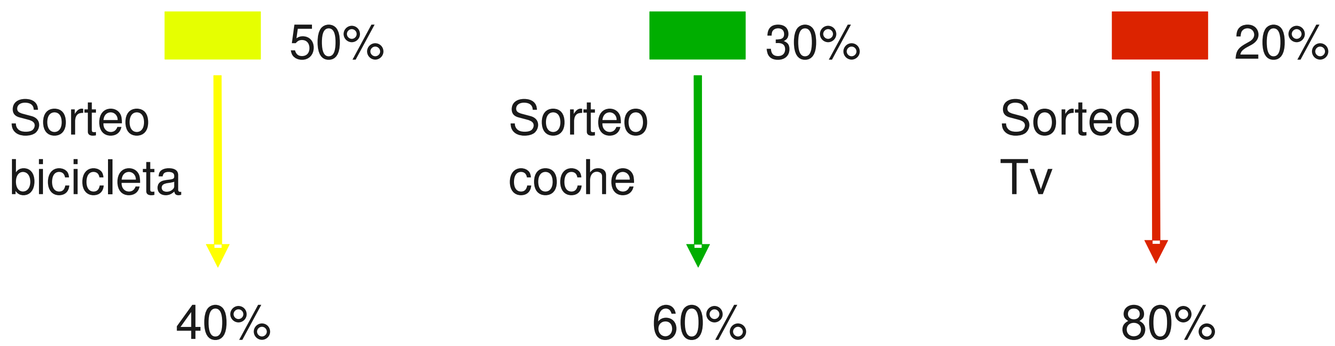
¿Cual es la probabilidad del suceso B, P(B)?

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) \Rightarrow P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Teorema de la Probabilidad Total

Ejemplo

En un saco hay papeletas de 3 colores con las siguientes probabilidades de ser elegidas:



Probabilidad de ganar en cada concurso

¿Qué probabilidad tienes de ganar en el sorteo?

$$\begin{aligned} P(G) = &\quad P(G \cap A) + P(G \cap V) + P(G \cap R) \\ = &\quad P(G|A)P(A) + P(G|V)P(V) + P(G|R)P(R) \\ = &\quad 0.54 \end{aligned}$$

Teorema de Bayes

Supongamos ahora que B ocurre, ¿cuál de los sucesos A_i ha ocurrido?

¿$P(A_i|B)$? $i = 1, ..., n$

Teorema de Bayes:
parte de una situación en la que es posible conocer las probabilidades de que ocurran una serie de sucesos A_i.

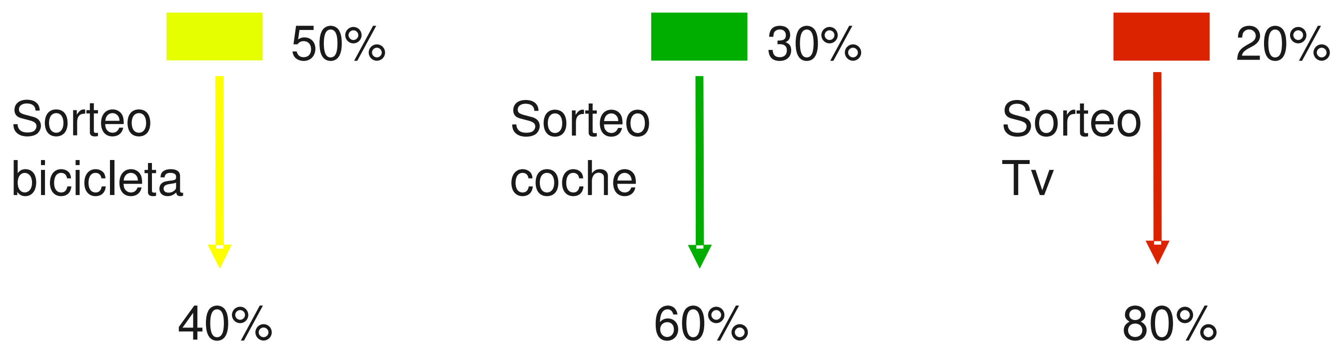
Conociendo que ha ocurrido el suceso B, la fórmula del teorema de Bayes nos indica como modifica esta información las probabilidades de los sucesos A_i.

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

Teorema de Bayes

Ejemplo

En un saco hay papeletas de 3 colores con las siguientes probabilidades de ser elegidas:



Probabilidad de ganar en cada concurso

¿Sabemos que Pablo ha ganado el sorteo, que probabilidad hay de que Pablo haya ganado una bicicleta?

$$P(A|G) = \frac{P(G \cap A)}{P(G)} = \frac{P(G|A)P(A)}{P(G)} = \frac{0.4 \cdot 0.5}{0.54} = 0.37$$

Ejercicio

La probabilidad de que una persona, seleccionada al azar, sea diabética es 0,03. Además en el 95% de las personas diabéticas los niveles de glucosa en sangre son superiores a 1000 mg/l, mientras que para las personas sanas ésto sólo ocurre el 2%.

Supongamos que al realizar un análisis de sangre los niveles de glucosa de una persona son superiores a 1.000 mg/l.

¿Cuál es la probabilidad de que esa persona sea diabética?

Probabilidad y variable aleatoria

Probabilidad

- Experimento aleatorio y definiciones de probabilidad
- Probabilidad condicionada e independencia de sucesos
- Teorema de la probabilidad total y de Bayes

Variable aleatoria

- Definición
- Funciónes de probabilidad, densidad y distribución
- Medidas de una variable aleatoria

Variable Aleatoria

Define de forma numérica los resultados de un experimento aleatorio. Es una función que asigna a cada suceso un número.

Ejemplo

X: ”Número de puntos que aparece en la cara superior de un dado al lanzarlo” $x \in \{1,2,3,4,5,6\}$

Y: “Número de artículos defectuosos por lote”
$y \in \{0,1,2,3,4,5,6,\ldots n\}$

Z: “Número de personas que entran al día en una tienda”
$z \in \{0,1,2,3,4,5,6, \ldots\}$

W: “Altura de los alumnos de esta asignatura”
$w \in \mathbb{R}$

V: “Suma que aparece al lanzar dos dados”
$v \in \{2,3,4,5,6,7,8,9,10,11,12\}$

Variable Aleatoria

Una **variable aleatoria** unidimensional es una función definida para todos los sucesos elementales de un experimento aleatorio tal que se verifican las siguientes condiciones:

1) Los valores de la función son números reales

2) La probabilidad $P(X \leq x)$ del suceso $\{X \leq x\}$ está definida de forma única y consistente con el espacio de probabilidad.

Discretas: Entre dos valores próximos toma a lo sumo un número finito de valores. E.g. número de artículos defectuosos

Continuas: Entre dos valores próximos toma un número infinito de valores. E.g. peso de los individuos de una población.

Ejemplo

Experimento aleatorio:
lanzar dos dados (no trucados)

Variable aleatoria: función que hace corresponder a cada resultado del experimento la suma de ambos dados.

sucesos | Variable aleatoria (x_i)

1-1
1-2
1-3
….
3-1
….
6-6

2
3
4
…
12

V.A. Discreta: fc. Probabilidad

Asigna a cada posible valor de la variable discreta su probabilidad, $p_X(x)$. Cumple las siguientes propiedades

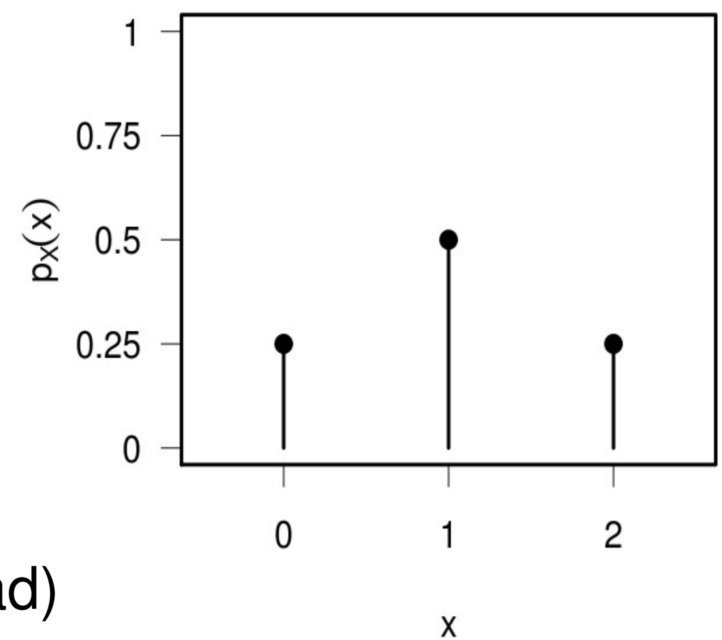
1. $p_X(x) = P(X=x)$

2. $p_X(x) \geq 0$

3. $\sum_{\forall i} p_X(x_i) = 1$

4. $P(a<X\leq b) = \sum_{a<x_i\leq b} p_X(x_i)$

5. Son valores adimensionales (probabilidad)



Ejemplo

Número de cruces al lanzar 2 monedas:

Suceso elemental $\xrightarrow{X}$ R $\xrightarrow{p_X(x)}$ [0 ,1]

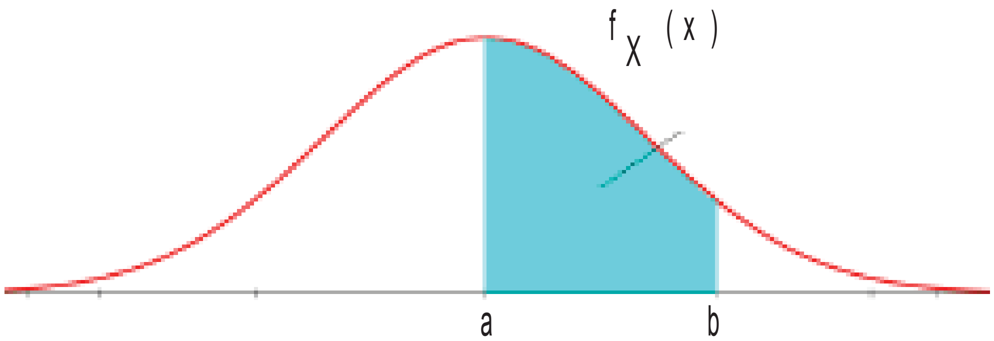
ω_i	x_i	$P(X=x_i)$
CC	0	1/4
CX, XC	1	1/2
XX	2	1/4

$$p_X(x) = \begin{cases} \frac{1}{4} & \text{si } x=0 \text{ o } x=2 \\ \frac{1}{2} & \text{si } x=1 \\ 0 & \text{en otro caso} \end{cases}$$

V.A. Continua: fc. Densidad

$f_X(x)$, es una función que cumple:

1. $f_X(x) \geq 0$
2. $\int_R f_X(x) = 1$
3. $P(a<X\leq b) = \int_a^b f_X(x)dx$



4. La función densidad **no** es la probabilidad de x.
5. Su valor tiene dimensiones del inverso de la variable

Identificamos la probabilidad de un intervalo como el área bajo la f. densidad.

Ejemplo

$$f_X(x) = \begin{cases} ax^2/27 & \text{si } 0<x<a \\ 0 & \text{en otro caso} \end{cases}$$

Determinar el valor de la constante de normalización a.

María Dolores Frías, Jesús Fernández y Carmen María Sordo

Función de Distribución

$F_X(x)$, función que asigna a cada valor de X la probabilidad de que la variable aleatoria tome valores menores o iguales que x:

$$\mathbf{F_X(x) = P(X \leq x)}$$

v. a. discreta: $F_X(x) = P(X \leq x) = \sum_{x_i \leq x} p_X(x_i)$

v. a. continua: $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x')\,dx'$

Ejemplo

Número de caras al lanzar 2 monedas

sucesos		V.A. (x_i)		$p_X(x_i)$
CC	$\to$	0	$\to$	1/4
CX, XC	$\to$	1	$\to$	1/2
XX	$\to$	2	$\to$	1/4

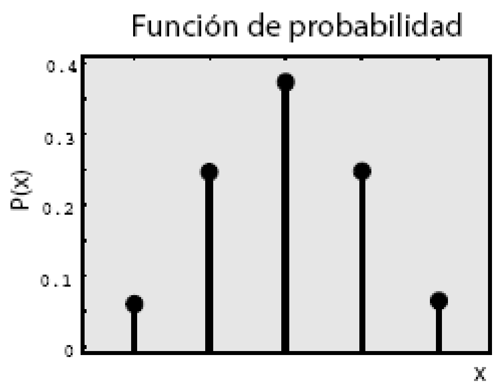
$$F_X(x) = \begin{cases} 0 & \text{si } x<0 \\ \frac{1}{4} & \text{si } 0 \leq x < 1 \\ \frac{3}{4} & \text{si } 1 \leq x < 2 \\ 1 & \text{si } 2 \leq x \end{cases}$$

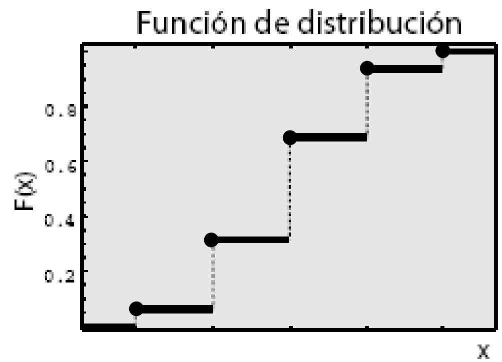
María Dolores Frías, Jesús Fernández y Carmen María Sordo

Propiedades de Fc. Distribución

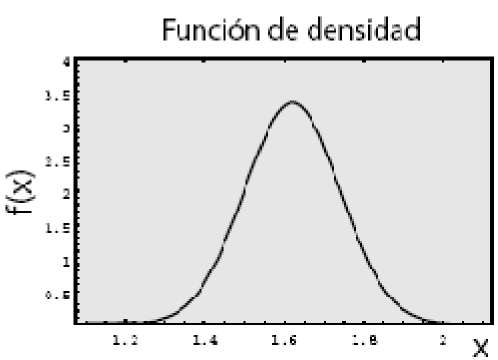
1. $F_X(-\infty) = 0$
2. $F_X(\infty) = 1$
3. Monótona no decreciente
4. Continua por la derecha
5. $P(a<X\leq b) = F_X(b) - F_X(a)$
6. $P(X=x) = F_X(x) - F_X(x^-)$. En particular, si X es una v.a. continua $P(X=x)=0 \ \ \forall x$

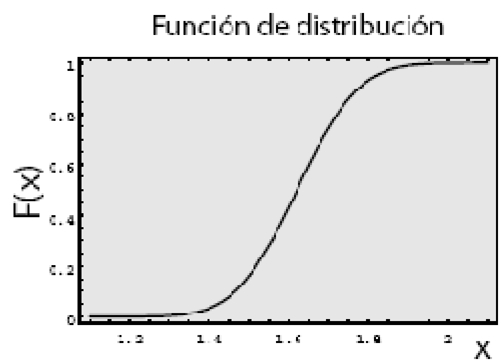
Fc. Distribución



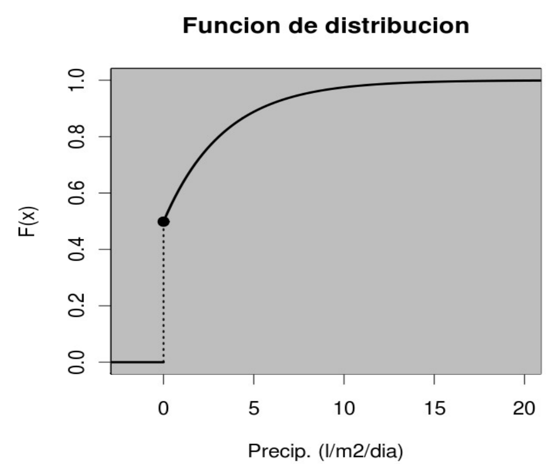


V.A. discreta: Número caras al lanzar 4 monedas





V.A. continua: Altura de los individuos de una población



V.A. mixta: lluvia recogida en un día en Santander

Ejercicio

Contestar razonadamente a las siguientes preguntas referidas a la figura de la derecha:

a) La función representada, ¿puede ser función de distribución de alguna variable aleatoria?. Si es así, ¿de qué tipo de variable aleatoria se trata?

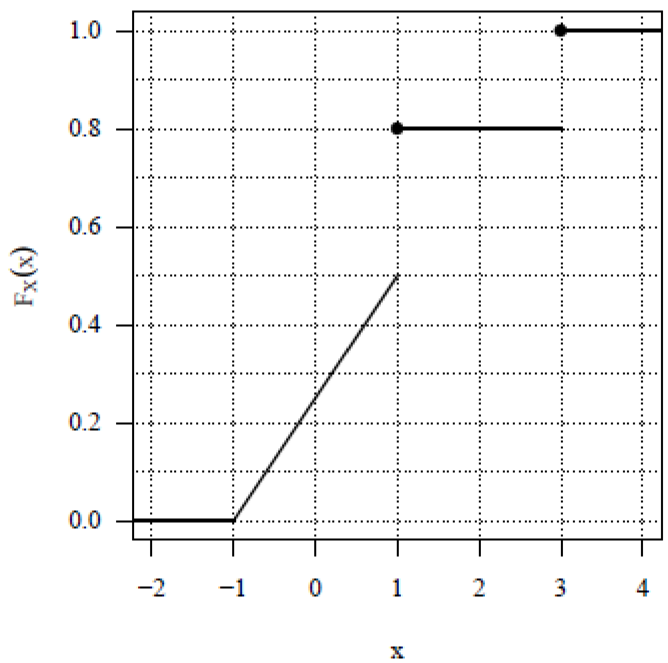
b) ¿Cuánto vale $P(X = 0)$?

c) ¿Cuánto vale $P(X = 1)$?

d) ¿Cuánto vale $P(X \leq 2)$?

e) ¿Cuánto vale la mediana de esta variable aleatoria?

f) ¿Y su rango intercuartílico?



Ejercicio

Dada la función $f(x) = a(1 + x^2)$ si $x \in [0, 3]$ (y 0 en otro caso)

a) Hallar el valor de a para que sea una función de densidad.

Medidas de una distribución

Valores numéricos que describen la distribución de probabilidad de una variable aleatoria.

Método más sencillo que el uso de funciones (distribución, probabilidad, densidad).

Descripción incompleta.

Medidas de posición: Cuantiles, deciles, cuartiles.

Medidas de localización: Informan sobre la localización de la distribución. Media y mediana.

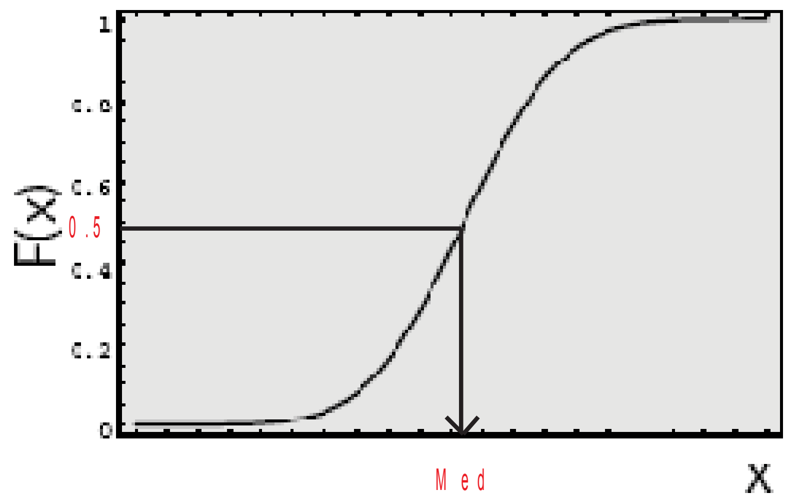
Medidas de dispersión: Miden el grado de variabilidad de la distribución. Varianza, desviación típica, rango intercuartílico.

Medidas de Posición

Cuantil de orden α, q_α

Cualquier valor que verifique: $P(X \leq q_\alpha) \geq \alpha \quad P(X \geq q_\alpha) \geq 1-\alpha$

Mediana, Med(X) : $q_{0,5}$



Cuartiles:

$Q_1 = q_{0.25}$

$Q_2 = q_{0.5}$

$Q_3 = q_{0.75}$

Ejemplo

$$f_X(x) = \begin{cases} 0.34x+0.83 & 0<x<1 \\ 0 & \text{en otro caso} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \text{si } x<0 \\ 0.17x^2+0.83x & 0\leq x<1 \\ 1 & \text{si } 1\leq x \end{cases}$$

$$P(X\leq q_{0,25})=0,17q_{0,25}^{\ 2}+0,83q_{0,25}=0,25 \Rightarrow q_{0,25}=0,2846$$

$$P(X\leq q_{0,5})=0,17q_{0,5}^{\ 2}+0,83q_{0,5}=0,5 \Rightarrow q_{0,5}=0,5422$$

Medidas de Localización

Esperanza matemática o media, E(X) o μ_X

V.A. Discreta: $$E(X) \equiv \mu_X = \sum_i x_i p_X(x_i)$$

V.A. Continua: $$E(X) \equiv \mu_X = \int_{-\infty}^{\infty} x f_X(x)\,dx$$

E(aX+bY+c) = aE(X)+bE(Y)+c donde X e Y son v.a. y a,b,c en R

Ejemplo

$$p_X(x) = \begin{cases} 0{,}2 \text{ para } x=0 \\ 0{,}8 \text{ para } x=1 \\ 0 \text{ en otro caso} \end{cases}$$

$$E(X) = 0*0{,}2+1*0{,}8 = 0{,}8$$

$$f_X(x) = \begin{cases} 1 \text{ para } 0\leq x\leq 1 \\ 0 \text{ en otro caso} \end{cases}$$

$$E(X) = \int_0^1 x*1dx = 0{,}5$$

Medidas de Dispersión

Varianza var(X) o σ^2

$$var(X) \equiv \sigma_X^2 = E\left(\left(X - \mu_X\right)^2\right)$$

V.A. Discreta:
$$var(X) = \sigma_X^2 = \sum_i (x_i - \mu_X)^2 p_X(x_i)$$

V.A. Continua:
$$var(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx$$

Desviación típica, dev(X) o σ: raiz cuadrada de la varianza

Rango Intercuantílico: $Q_3 - Q_1$

Ejemplo

$$p_X(x) = \begin{cases} 0{,}2 \text{ para } x=0 \\ 0{,}8 \text{ para } x=1 \\ 0 \text{ en otro caso} \end{cases}$$

$$Var(X) = (0-0{,}8)^2*0{,}2+(1-0{,}8)^2*0{,}8 = 0{,}16$$

$$f_X(x) = \begin{cases} 1 \text{ para } 0\leq x\leq 1 \\ 0 \text{ en otro caso} \end{cases}$$

$$Var(X) = \int_0^1 (x-0{,}5)^2*1dx = 0{,}08$$

María Dolores Frías, Jesús Fernández y Carmen María Sordo

Ejercicio

Dada la función $f(x) = a(1 + x^2)$ si $x \in [0, 3]$ (y 0 en otro caso)

a) Hallar el valor de a para que sea una función de densidad.

b) Hallar el valor esperado y la varianza de la distribución.

c) Hallar la probabilidad de que la variable aleatoria sea mayor que su valor esperado.

Media y Varianza

Sea Y=aX+b, donde a y b son contantes y X e Y variables aleatorias

$$\mu_Y = a\,\mu_X + b \qquad \sigma_Y^2 = a^2\sigma_X^2$$

Se puede pasar de una variable aleatoria X de media μ_X y varianza σ^2 a otra $Y=(X-\mu_X)/\sigma$ con media 0 y varianza 1.

Estandarización de la variable aleatoria X.

Ejercicio

La demanda, expresada en toneladas, de un determinado producto es una variable aleatoria cuya función de densidad es :

$$f_X(x) = \begin{cases} x/6 & 2 \leq x \leq 4 \\ 0 & \text{en otro caso} \end{cases}$$

¿Cuales son la media, la varianza y la mediana de esta demanda?

Suponiendo que los beneficios Y del producto pueden obtenerse a partir de la demanda mediante la fórmula $Y=c+dX$, se pide:

1. Calcular los beneficios esperados
2. Calcular la varianza de los beneficios

Ejercicio

Dada la función densidad:

$$f_x(x) = \begin{cases} ax^2/27 & 0<x<a \\ 0 & \text{en otro caso} \end{cases}$$

1. Determinar el valor de la constante de normalización a y representar gráficamente la función densidad.

2. Su función de distribución F(x). Representarla gráficamente.

3. Obtener P(1<x<2) y P(x>2)