FUNDAMENTOS DA TEORÍA DA PROBABILIDADE
resumos numéricos e representacións gráficas

María Isabel Borrajo García, Mercedes Conde Amboage e Rosa María Crujeiras Casais (Departamento de Estatística, Análise Matemática e Optimización)

No mundo real pódense atopar dous tipos de fenómenos:

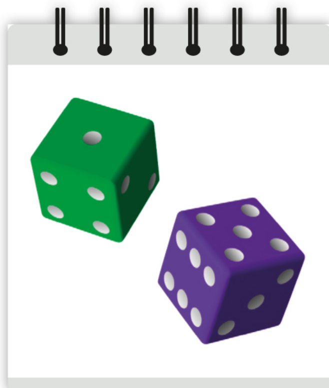
- **Experimento aleatorio**: é aquel experimento que, repetido sucesivas veces en idénticas condicións, pode dar lugar a resultados diferentes e imprevisibles. É dicir, estas probas teñen por característica fundamental a incerteza do seu resultado, aínda que todos os seus posibles resultados han de ser coñecidos de antemán. Exemplos clásicos son o lanzamento dun dado ou lanzamento dunha moeda.
- **Experimento determinista**: é aquel experimento que repetido sucesivamente en condicións idénticas produce sempre os mesmos resultados. Un exemplo é a hora de saída do sol. Por outra banda, un experimento determinista pode mudar a aleatorio se se introduce un erro asociado, por exemplo, un erro de medida.

A teoría da probabilidade ocúpase de entender e estudar esta incerteza que caracteriza aos fenómenos aleatorios, asignando a cada un dos seus posibles resultados un número que indicará o grao no que dito resultado pode ter lugar.

Para poder aplicar correctamente a teoría da probabilidade, precisamos coñecer algúns conceptos fundamentais:

- **Suceso elemental**: é cada un dos posibles resultados dun experimento aleatorio.
- **Espazo mostral**: é o conxunto de todos os sucesos elementais, xeralmente denótase por Ω.
- **Suceso**: é calquera subconxunto do espazo mostral; en particular Ω denomínase *suceso seguro*, e o conxunto baleiro, $\emptyset$, denomínase *suceso imposible*.

Exemplo 1: para ilustrar estes conceptos empregaremos un exemplo clásico e moi sinxelo de experimento aleatorio: o lanzamento dun dado. Nese caso os sucesos elementais son todos os posibles resultados, é dicir, 1, 2, 3, 4, 5 e 6; o espazo mostral é $\Omega = \{1,2,3,4,5,6\}$, e por exemplo $A =$ "saír un número par" $= \{2,4,6\}$ é un exemplo de suceso.



Sabías que...?

No ano 1713 publícase o *Ars Conjectandi* de Jacob Bernoulli, a primeira obra que asenta os conceptos básicos da probabilidade e da combinatoria, e que establece a lei feble dos grandes números: a probabilidade de ocorrencia dun evento podemos aproximala pola frecuencia relativa de ocorrencia do mesmo, nun número grande de repeticións. Este foi un primeiro paso para relacionar a probabilidade coa estatística!

Unha breve (e necesaria) revisión sobre teoría de conxuntos

A teoría de conxuntos é unha rama das matemáticas que se ocupa do estudo duns obxectos denominados conxuntos. Un conxunto defínese como unha agrupación, colección ou clase de obxectos coñecidos como elementos do conxunto.

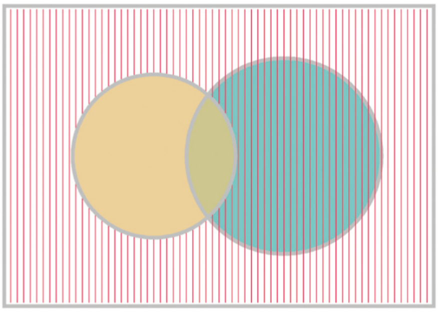
En probabilidade trabállase con sucesos, que son fundamentalmente conxuntos non ordenados. É por isto que moitos dos conceptos que se empregan en teoría da probabilidade están sustentados pola teoría de conxuntos. Para poder formalizar adecuadamente algunhas ideas da teoría da probabilidade, precísase ter presente o seguinte:

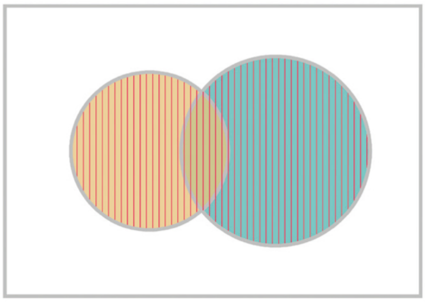
- **Partes dun conxunto:** son todos os subconxuntos que se poden formar a partir do conxunto dado, por exemplo, $\mathcal{P}(\Omega)$ denota as partes do conxunto Ω.
- **σ-álxebra:** dado $\mathcal{A} \subset \mathcal{P}(\Omega)$, dise que $\mathcal{A}$ é unha σ-álxebra de conxuntos sobre Ω se verifica estas condicións:
  ✓ o conxunto baleiro pertence á σ-álxebra, $\emptyset \in \mathcal{A}$,
  ✓ dado $A \subset \Omega$, se $A \in \mathcal{A}$, entón $A^c \in \mathcal{A}$, onde A^c denota o conxunto complementario de A, e
  ✓ dados $A_i \in \mathcal{A}$ para todo $i \in \mathbb{N}$ entón $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{A}$.
- **Conxuntos disxuntos:** dous conxuntos calquera chámanse disxuntos se non posúen ningún elemento en común.

OPERACIÓNS CON SUCESOS

A partir dos sucesos elementais dun experimento pódense definir outros sucesos derivados a través do complementario, a unión ou a intersección de sucesos, entre outras. Estas operacións teñen como base a teoría de conxuntos, propia da álxebra, e fundaméntanse en conceptos como os que se acaban de revisar na epígrafe anterior.

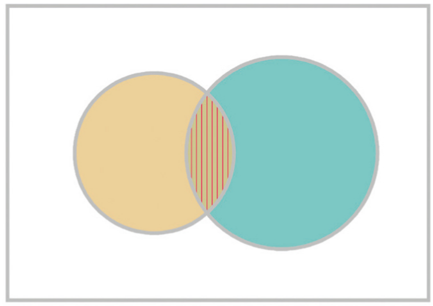
Dado un suceso A en Ω, defínese o seu **suceso complementario** como aquel que ocorre cando non ocorre A, e denótase por A^c.



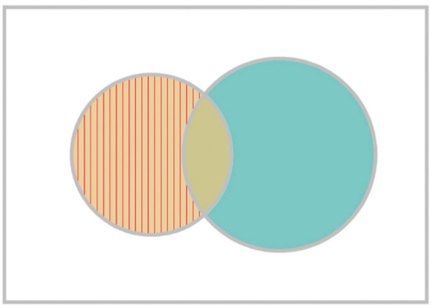


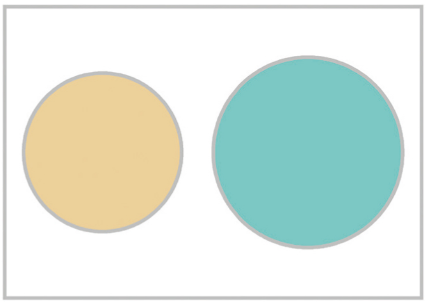
Dados dous sucesos A e B en Ω, defínese o **suceso unión** de A e B como aquel que ocorre cando ocorre A ou ocorre B, e denótase por $A \cup B$.

Dados dous sucesos A e B en Ω, defínese o **suceso intersección** de A e B como aquel que ocorre cando ocorre A e ocorre B simultaneamente, e denótase por $A \cap B$.



Dados dous sucesos A e B en Ω, defínese o **suceso diferenza** entre A e B como aquel que ocorre cando ocorre A e non ocorre B, e denótase por $A \setminus B$.

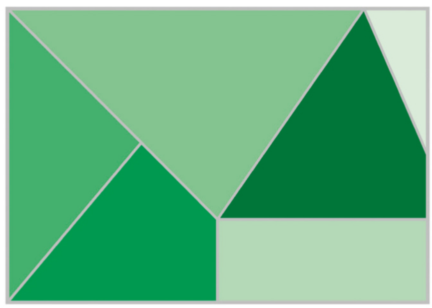




Dados dous sucesos A e B en Ω, dise que son **incompatibles** se a súa intersección é o conxunto baleiro, é dicir, verifícase que $A \cap B = \emptyset$.

Dada unha familia de sucesos $\{A_1, \dots, A_n\}$ en Ω, dise que conforma un **sistema completo de sucesos** se se verifica:

- $A_i \cap A_j = \emptyset \quad \forall i \neq j$
- $\cup_{i=1}^{n} A_i = \Omega$



Nótese que nas figuras precedentes denota o suceso A, denota o suceso B e o raiado denota o suceso resultante da operación realizada entre A e B.

Continuando co **Exemplo 1**, ademais do suceso $A =$ "saír un número par" $= \{2,4,6\}$, considérese o suceso $B =$ "saír un número maior que 3" $= \{4,5,6\}$. Entón, coas operacións con sucesos obteriamos o seguinte:

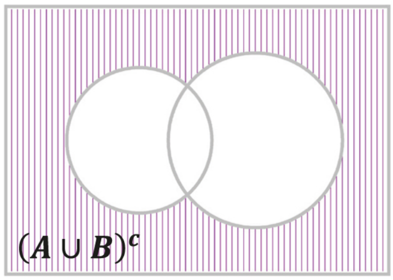
- Suceso complementario de A: $A^c =$ "non saír un número par" = "saír un número impar" $= \{1,3,5\}$.
- Suceso unión de A e B: $A \cup B =$ "saír un número par ou saír un número maior que 3" $= \{2,4,5,6\}$.
- Suceso intersección de A e B: $A \cap B =$ "saír un número par e saír un número maior que 3" $= \{4,6\}$.
- Suceso diferenza de A e B: $A \setminus B =$ "saír un número par, pero non saír un número maior que 3 " $= \{2\}$.
- A e B son sucesos compatibles, posto que a súa intersección non é o conxunto baleiro.
- A familia $\{A_1, A_2, A_3\}$, sendo $A_1 = \{1,2\}$, $A_2 = \{3,4\}$ e $A_3 = \{5,6\}$ conformarían un sistema completo de sucesos.

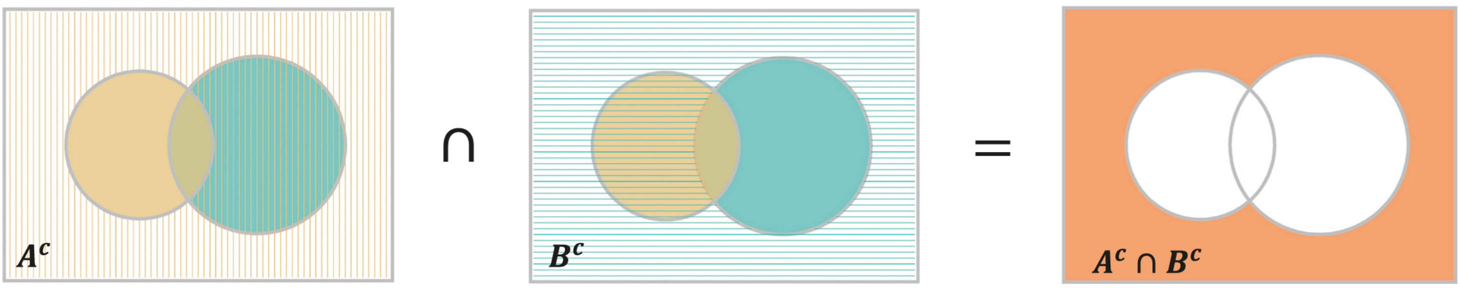
OPERACIÓNS CON SUCESOS (cont.)

As leis de De Morgan

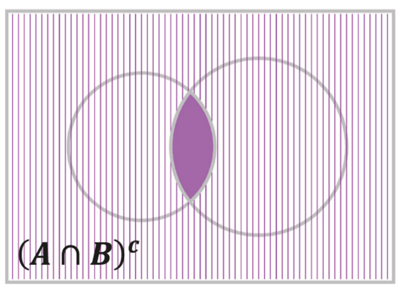
As leis de De Morgan, en honor a Augustus De Morgan (1806-1871), matemático e lóxico británico, profesor do University College London, e presidente da London Mathematical Society, permiten o intercambio de unións e intersección a través do complementario, e son útiles no cálculo de probabilidades de diferentes sucesos.

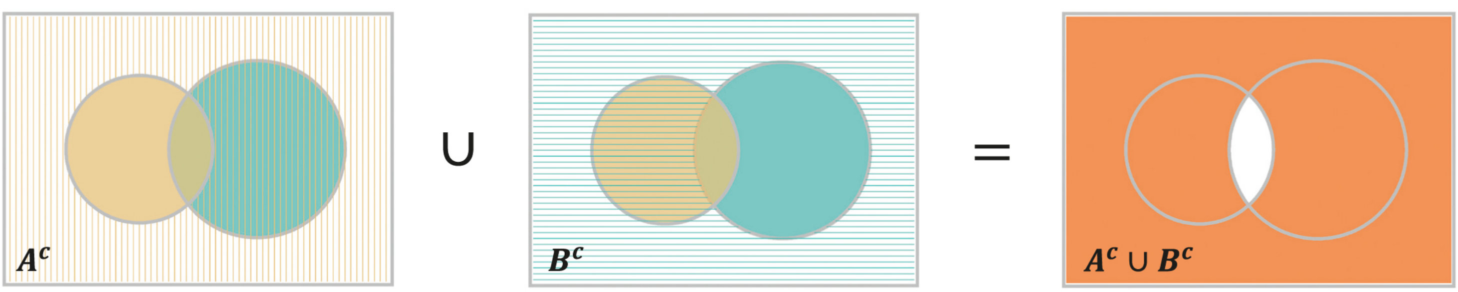
- $(\boldsymbol{A} \cup \boldsymbol{B})^{\boldsymbol{c}} = \boldsymbol{A}^{\boldsymbol{c}} \cap \boldsymbol{B}^{\boldsymbol{c}}$, é dicir, o complementario da unión de dous sucesos pode obterse como intersección dos respectivos complementarios.





- $(\boldsymbol{A} \cap \boldsymbol{B})^{\boldsymbol{c}} = \boldsymbol{A}^{\boldsymbol{c}} \cup \boldsymbol{B}^{\boldsymbol{c}}$, é dicir, o complementario da unión de dous sucesos pode obterse como intersección dos respectivos complementarios.





O CONCEPTO DE PROBABILIDADE

No estudo de experimentos aleatorios non só é importante coñecer os posibles resultados senón saber tamén con que probabilidade se produce cada un deles. Por iso, unha vez definidos os posibles sucesos débeselles asignar unha probabilidade.

Definición de Laplace

Supoñamos que Ω ten un número finito de sucesos elementais e que todos eles ocorren coa mesma probabilidade. Entón, a probabilidade dun suceso A en Ω, $\mathbb{P}(A)$, calcúlase como o cociente entre casos favorables e casos posibles. No **Exemplo 1** , a probabilidade do suceso $A =$ "saír un número par" $=\{2,4,6\}$ sería $\mathbb{P}(A) = \frac{3}{6} = \frac{1}{2}$.

Definición frecuentista

Se repetimos n veces un experimento aleatorio e n_A denota o número de veces que ocorre o suceso$A$, a frecuencia relativa deste suceso defínese como $fr(A)=\frac{n_A}{n}$, onde $0 \leq fr(A) \leq 1$. Ao facer $0 \leq fr(A) \leq 1$ grande, esta frecuencia estabilízase nun valor e defínese a probabilidade do suceso A como o límite das frecuencias cando $n \to \infty$.

Definición axiomática de Kolmogorov

Considérese un espazo mostral Ω e unha familia de subconxuntos del denotada por $\mathcal{A}$. Dise que $\mathbb{P}$ é unha probabilidade en $(\Omega, \mathcal{A})$ se verifica as seguintes condicións:

- a probabilidade do espazo mostral é a unidade: $\mathbb{P}(\Omega) = 1$,
- dada unha familia de sucesos $\{A_1, \dots, A_i, \dots\}$ en Ω, disxuntos dous a dous, $A_i \cap A_j = \emptyset$ se $i \neq j$ entón verifícase que:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i),$$

- dado calquera suceso $A \in \mathcal{A}$, entón $0 \leq \mathbb{P}(A) \leq 1$.

Probabilidade condicionada

Como xa se indicou, a teoría da probabilidade está ligada á incerteza, é dicir, ao noso descoñecemento sobre os resultados dun experimento. O feito de que ocorra ou non un determinado suceso (ter ese coñecemento) incorpora cambios no grao de certeza sobre os sucesos aleatorios, ao estar adquirindo nova información. Estes cambios na valoración da incerteza, poden producir cambios nas probabilidades dos sucesos aleatorios.

Dados dous sucesos A e B en Ω, con $\mathbb{P}(B) > 0$, defínese formalmente a probabilidade de A condicionada ao suceso B como:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

É importante notar que esta definición é consistente no sentido de que cumpre os axiomas da definición dunha probabilidade. Ademais, no caso particular de que $B = \Omega$, teríase que

$$\mathbb{P}(A|\Omega) = \frac{\mathbb{P}(A \cap \Omega)}{\mathbb{P}(\Omega)} = \mathbb{P}(A).$$

Exemplo2: consideremos o experimento aleatorio consistente en lanzar dous dados no que o espazo mostral vén dado por $\Omega=\{(1,1), (1,2), \dots, (1,6), (2,1), \dots, (2,6), \dots, (6,1), \dots, (6,6)\}$. Definamos os sucesos A = "saír polo menos un seis" e B = "a suma dos dous resultados sexa maior que 8". Entón

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{3/11}{11/36} = \frac{108}{121} \approx 0.8926.$$

Independencia de sucesos

Dados dous sucesos A e B en Ω , diremos que son independentes se se verifica que

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Esta condición resulta equivalente a dicir que $\mathbb{P}(A|B)=\mathbb{P}(A)$ se $\mathbb{P}(B) \neq 0$. Ademais, se A e B son independentes, é fácil ver que os seguintes pares de sucesos tamén o son: A e B^c, A^c e B, A^c e B^c.

RESULTADOS NOTABLES

Regra do produto

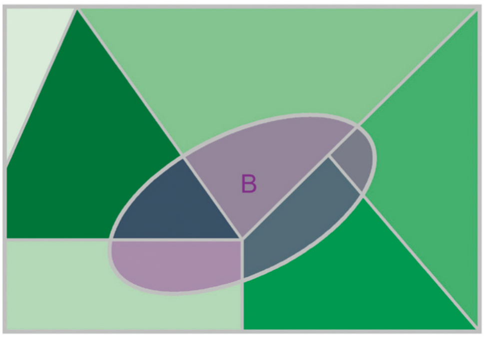
Tendo en conta a definición de probabilidade condicionada, dados dous sucesos A e B en Ω tales que $\mathbb{P}(B) > 0$, pódese calcular a probabilidade do suceso $A \cap B$ como:

$$\mathbb{P}\,(A \cap B) = \mathbb{P}(A) \cdot\ \mathbb{P}(B|A).$$

Do mesmo xeito, se consideramos $\{A_1, \dots, A_n\}$ un conxunto de sucesos de Ω, tales que a intersección de todos eles é non baleira, é dicir, $\mathbb{P}(\cap_{i=1}^{n}\ A_i) > 0$, entón, a intersección de todos os sucesos A_i (ilustrada na figura da dereita para $n = 4$) pode calcularse como segue:

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_1) \cdot\ \mathbb{P}(A_2|A_1) \cdot\ \mathbb{P}(A_3|A_1 \cap A_2) \cdot \dots \cdot\ \mathbb{P}\big(A_n|\cap_{i=1}^{n-1}\ A_i\big).$$

Teorema de probabilidades totais



Sexa $\{A_1, \dots, A_n\}$ un sistema completo de sucesos en Ω tal que $\mathbb{P}(A_i) > 0$ para todo $i = 1, \dots n$. Se consideramos un suceso B en Ω, tal e como se ilustra na figura da esquerda, verifícase que:

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i).$$

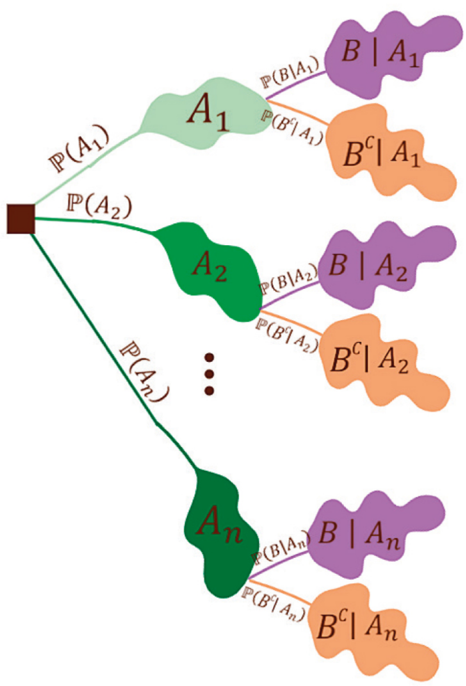
Como os A_i son disxuntos, a probabilidade de B pode escribirse como suma das probabilidades das interseccións $\mathbb{P}(B \cap A_i)$, e como $\mathbb{P}(A_i) > 0$ a probabilidade de cada intersección pode obterse a partir da probabilidade condicionada $\mathbb{P}(B|A_i)$.

Teorema de Bayes

Consideremos un experimento aleatorio que se realiza en dúas etapas: na primeira, tense un sistema completo de sucesos $\{A_1, \dots, A_n\}$, con probabilidades $\mathbb{P}(A_i) > 0$ que se denominan probabilidades *a priori*. Nunha segunda etapa, ocorre o suceso B, e coñécense as probabilidades condicionadas $\mathbb{P}(B|A_i)$ de obter nesta etapa o suceso B cando na primeira etapa se obtivo o suceso A_i. Nestas condicións o teorema de Bayes permite calcular as probabilidades:

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j)}{\sum_{i=1}^{n} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}.$$

Estas probabilidades reciben o nome de probabilidades *a posteriori*, pois calcúlanse despois de ter observado o suceso B.



VARIABLES ALEATORIAS

Unha **variable aleatoria** X é unha correspondencia que asocia a cada elemento do espazo mostral dun experimento aleatorio un número. Se X é unha variable aleatoria que toma unha cantidade finita ou infinita numerable de valores, diremos que X é unha **variable aleatoria discreta**. Pola contra, diremos que X é unha **variable aleatoria continua** se toma valores nun intervalo ou unión de intervalos da recta real.

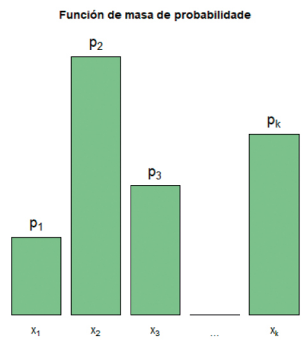
Variables discretas

Sexa X unha variable aleatoria discreta que toma os valores $\{x_1, x_2, \dots, x_k\}$, denomínase **función de masa de probabilidade** de X ao conxunto de probabilidades $\{p_1, p_2, \dots, p_k\}$, tales que

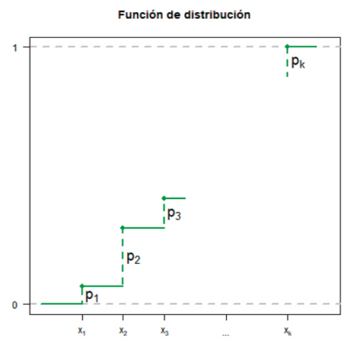
$$\mathbb{P}(X = x_i) = p_i > 0 \quad \text{para todo } \ i = 1, \dots, k \ \text{ e } \sum_{i=1}^{k} p_i = 1.$$

Ademais, o comportamento da variable X tamén se pode describir grazas á **función de distribución** (que denotaremos por F_X) que a cada valor $x \in \mathbb{R}$ lle asocia a probabilidade de que a variable X tome valores menores ou iguais a este, é dicir:

$$F_X(x) = \mathbb{P}_X(X \leq x) = \sum_{x_i \leq x} p_i.$$



A modo de ilustración, represéntase a función de masa de probabilidade e a función de distribución asociada a unha variable discreta que toma os valores $\{x_1, x_2, \dots, x_k\}$ con probabilidades $\{p_1, p_2, \dots, p_k\}$, respectivamente. Neste tipo de variables a función de distribución é unha función en esqueira.



Variables continuas

Dada unha variable aleatoria continua X, o seu comportamento estará caracterizado pola función de distribución ou pola función de densidade. De maneira análoga ao caso de variables discretas, pódese definir a función de distribución da variable X como a función que asocia a cada valor $x \in \mathbb{R}$ a probabilidade de que a variable tome valores menores ou iguais a este, é dicir:

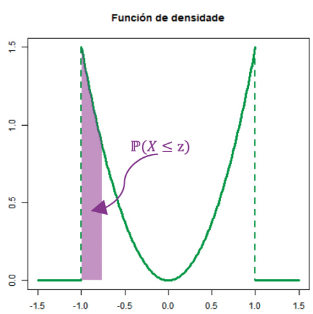
$$F_X(x) = \mathbb{P}(X \leq x).$$

Ademais, dada a variable continua X, pódese definir a función de densidade como:

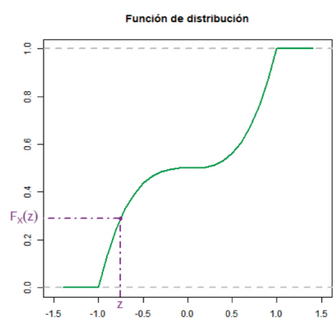
$$f_X(x) = \lim_{h \to 0} \frac{(x - h < X < x + h)}{2h}.$$

Polo tanto, temos a seguinte relación entre a función de densidade e a función de distribución:

$$f_X(x) = F'_X(x) \quad \text{ou ben} \quad F_X(x) = \int_{-\infty}^{x} f_X(t)dt.$$



Exemplo 3: móstrase a función de densidade e a función de distribución asociada a unha variable continua X con función de densidade $f_X(x) = \frac{3}{2}x^2$ no intervalo $[-1,1]$ e 0 noutro caso.



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA
XUNTA DE GALICIA
RAGC/ REAL ACADEMIA GALEGA DE CIENCIAS