ApLabAD

Apuntes de Laboratorio de Análisis de Datos

RICUIB

2025-02-01

Tabla de contenidos

Prefacio

Este libro en la web es una versión de las notas de clase de asignaturas introductorias al análisis de datos.

Ha sido elaborado con Quarto RStudio, PBC. (2022). Quarto (Version 1.0). Hemos utilizado el formato formato book.

Parte I

Parte 1: Probabilidad y variables aleatorias

En esta sección, veremos la teoría básica de la probabilidad y de las variables aleatorias. Resolveremos problemas prácticos para comprender mejor los conceptos y utilizaremos R para realizar cálculos y gráficos.

También exploraremos los modelos de probabilidad discretos y continuos más conocidos.

En ocasiones, trabajaremos con problemas de cálculo más complejos.

Además, introduciremos problemas sencillos de modelización con probabilidades. Estos consistirán en un enunciado, real o inventado, en el que se pedirá modelizar el problema mediante una variable aleatoria y responder a una serie de preguntas.

1 Preliminares: conjuntos y combinatoria

Para aprender cálculo de probabilidades son necesarios conocimientos de:

1. Cálculo: Derivadas, integrales, límites, sumas de series…
2. Geometría básica y álgebra lineal : rectas, hiperplanos, volúmenes… Matrices, valores propios…
3. Teoría de conjuntos y combinatoria…..

Por experiencia sabemos que la mayoría de estudiantes tienen más conocimientos de cálculo, geometría y matrices.

Pero muchos tienen una falta de conocimientos en teoría básica de conjuntos y combinatoria (matemática discreta).

1.1 Teoría de conjuntos

Definición de conjunto

La definición de conjunto es una idea o noción primitiva. Es decir es una idea básica del pensamiento humano: un conjunto es una colección de objetos: números, imágenes… cualquier cosa, jugadores de fútbol, palabras, colores ….

La definición de conjunto es una idea o noción primitiva. Es decir es una idea básica del pensamiento humano: un conjunto es una colección de objetos: números, imágenes… cualquier cosa, jugadores de fútbol, palabras, colores ….

La teoría de conjuntos básicas es simple y natural y es la que necesitamos para este curso.

La teoría de conjuntos matemática es más compleja y presenta varias paradojas como la paradoja de Russell.

La idea o noción práctica de conjunto es la de una colección de objetos de un cierto tipo.

Estas colecciones o conjuntos se pueden definir por:

- **Comprensión**: reuniendo los objetos que cumplen una propiedad p
- **Extensión**: dando una lista exhaustiva de los miembros del conjunto

1.1.1 Conjuntos básicos

Los conjuntos suelen tener un conjunto madre como por ejemplo

- $\mathbb{N} = \{0, 1, 2, ...\}$
- $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, ...\}$
- $\mathbb{Q} = \left\{ \frac{p}{q} \quad \middle| \quad p, q \in \mathbb{Z} \text{ y } q \neq 0. \right\}$
- $\mathbb{R} = \{\text{Todos los puntos de una recta.}\}$
- $\mathbb{C} = \{a + b \cdot i \quad | \quad a, b \in \mathbb{R}\}$ los números complejos $\quad a + b \cdot i$.
- Alfabeto $= \{a, b, c, \ldots, A, B, C, ...\}$.
- Palabras $= \{paz, guerra, amor, probabilidad, ...\}$.

Recordemos que i es la unidad imaginaria que cumple que $i = \sqrt{-1}$.

1.1.2 Características y propiedades básicas de los conjuntos

Si a cada objeto x de Ω le llamaremos **elemento del conjunto** Ω y diremos que x pertenece a Ω. Lo denotaremos por $x \in \Omega$.

Un **conjunto de un elemento**, por ejemplo $\{1\}$ recibe el nombre de **conjunto elemental** (o **singleton** del inglés).

Sea A otro conjunto diremos que A **es igual a** B si todos los elementos A están en B y todos los elementos de B están en A. Por ejemplo $A = \{1, 2, 3\}$ es igual a $B = \{3, 1, 2\}$.

Si B es otro conjunto, tal que si $x \in A$ entonces $x \in B$ diremos que A es un subconjunto de o que está contenido en B. Lo denotaremos por $A \subseteq B$.

El conjunto que no tiene elementos se denomina conjunto vacío y se denota por el símbolo $\emptyset$. Dado A un conjunto cualquiera obviamente $\emptyset \subseteq A$.

Ejemplo

Tomemos como conjunto base $\Omega = \{1, 2, 3\}$

- Ω es un conjunto de cardinal 3, se denota por $\#(\Omega) = 3$ o por $|\Omega| = 3$
- El conjunto Ω tiene $2^3 = 8$ subconjuntos.
  - el vacío $\emptyset$ y los elementales $\{1\}, \{2\}, \{3\}$
  - los subconjuntos de dos elementos: $\{1, 2\}, \{1, 3\}, \{2, 3\}$
  - el conjunto total de tres elementos $\Omega = \{1, 2, 3\}$.

Dado un conjunto Ω podemos construir el **conjunto de todas sus partes** (todos sus subconjuntos) al que denotamos por $\mathcal{P}(\Omega)$. También se denomina de forma directa partes de Ω.

Cardinal de las partes de un conjunto

Propiedad

Por ejemplo $\#\left(\mathcal{P}(\{1,2,3\})\right) = 2^{\#(\{1,2,3\})} = 2^3 = 8$.

Efectivamente

$$\mathcal{P}(\{1,2,3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

Dado un subconjunto A de Ω podemos construir la función característica de A

$$\chi_A : \Omega \to \{0,1\}$$

dado un $\omega \in \Omega$

$$\chi_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

1.1.3 Operaciones entre conjuntos

Intersección

Sea Ω un conjunto y A y B dos subconjuntos de Ω.

El conjunto **intersección** de A y B es el formado por todos los elementos que perteneces a A **Y** B, se denota por $A \cap B$.

Más formalmente

$$A \cap B = \left\{x \in \Omega \middle| x \in A \text{ y } x \in B\right\}.$$

Unión

El conjunto **unión** de A y B es el formado por todos los elementos que perteneces a A **O** pertenecen a B, se denota por $A \cup B$.

Más formalmente

$$A \cup B = \left\{x \in \Omega \middle| x \in A \text{ o } x \in B\right\}.$$

Diferencia.

El conjunto **diferencia** de A y B es el formado por todos los elementos que perteneces a A **Y NO** pertenecen a B, se denota por $A - B = A - (A \cap B)$.

Más formalmente

$$A - B = \left\{x \in \Omega \middle| x \in A \text{ y } x \notin B\right\}.$$

Complementario

El **complementario** de un subconjunto A de Ω es $\Omega - A$ y se denota por A^c o $\overline{A}$.

Más formalmente

$$A^c = \left\{x \in \Omega \middle| x \notin A\right\}.$$

1.1.4 Más propiedades

Sea Ω un conjunto y A, B, C tres subconjuntos de Ω

- Se dice que dos conjuntos A y B **son disjuntos** si $A \cap B = \emptyset$.
- $\Omega^c = \emptyset$.
- $\emptyset^c = \Omega$.
- $A \cup B = B \cup A$, $A \cap B = B \cap A$ conmutativas.
- $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$ asociativas.
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ distributivas.
- $\left(A^c\right)^c = A$ doble complementario.
- $\left(A \cup B\right)^c = A^c \cap B^c$, $\left(A \cap B\right)^c = A^c \cup B^c$ leyes de De Morgan.

1.1.5 Con R, ejemplos.

Con R los conjuntos de pueden definir como vectores

```
Omega=c(1,2,3,4,5,6,7,8,9,10)
A=c(1,2,3,4,5)
B=c(1,4,5)
C=c(4,6,7,8)
Omega
```

```
 [1]  1  2  3  4  5  6  7  8  9 10
```

```r
A
```

```
[1] 1 2 3 4 5
```

```r
B
```

```
[1] 1 4 5
```

```r
C
```

```
[1] 4 6 7 8
```

$A \cap B$

```r
A
```

```
[1] 1 2 3 4 5
```

```r
B
```

```
[1] 1 4 5
```

```r
intersect(A,B)
```

```
[1] 1 4 5
```

$A \cup B$

```r
A
```

```
[1] 1 2 3 4 5
```

```r
B
```

```
[1] 1 4 5
```

```r
union(A,B)
```

```
[1] 1 2 3 4 5
```

$B - C$

```r
B
```

```
[1] 1 4 5
```

```r
C
```

```
[1] 4 6 7 8
```

```r
setdiff(B,C)
```

```
[1] 1 5
```

$A^c = \Omega - A$

```r
Omega
```

```
 [1]  1  2  3  4  5  6  7  8  9 10
```

```r
A
```

```
[1] 1 2 3 4 5
```

```r
setdiff(Omega,A)
```

```
[1]  6  7  8  9 10
```

1.1.6 Con python

```python
Omega=set([1,2,3,4,5,6,7,8,9,10])
A=set([1,2,3,4,5])
B=set([1,4,5])
C=set([4,6,7,8])
Omega
```

```
{1, 2, 3, 4, 5, 6, 7, 8, 9, 10}
```

```python
A
```

```
{1, 2, 3, 4, 5}
```

```python
B
```

```
{1, 4, 5}
```

```python
C
```

```
{8, 4, 6, 7}
```

```python
A & B   # intersección (&: and/y)
```

```
{1, 4, 5}
```

```python
A | B   # unión (|: or/o)
```

```
{1, 2, 3, 4, 5}
```

```python
A - C   # diferencia
```

```
{1, 2, 3, 5}
```

```python
Omega-C # complementario.
```

```
{1, 2, 3, 5, 9, 10}
```

1.2 Combinatoria

La combinatoria es una rama de la matemática discreta que entre otras cosas cuenta distintas configuraciones de objetos de un conjunto.

Por ejemplo si tenemos un equipo de baloncesto con 7 jugadores ¿cuántos equipos de 5 jugadores distintos podemos formar?

1.2.1 Número Binonial

Número combinatorio o número binomial

Nos da el número de subconjuntos de tamaño k de un conjunto de tamaño n. Este número es

$$C_n^k = \binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}.$$

Recordemos que

$$n! = 1 \cdot 2 \cdot 3 \cdots n.$$

Ejemplo

En nuestro caso con 7 jugadores $n = 7$ el número de equipos distintos de $k = 5$ es

$$\begin{aligned} C_7^5 &= \tbinom{7}{5} = \tfrac{7!}{5! \cdot (7-5)!} = \tfrac{7!}{5! \cdot 2!} \\ &= \tfrac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 1 \cdot 2} = \tfrac{6 \cdot 7}{2} = \tfrac{42}{2} = 21. \end{aligned}$$

Puedo formar 21 equipos distintos.

Ejercicio: el paquete `gtools`

Carga el paquete `gtools` de R y investiga la función `combinations(n, r, v, set, repeats.allowed)` para calcular todas las combinaciones anteriores.

1.2.2 Combinaciones con repetición

En combinatoria, las combinaciones con repetición de un conjunto son las distintas formas en que se puede hacer una selección de elementos de un conjunto dado, permitiendo que las selecciones puedan repetirse.

Combinaciones con repetición

El número CR_n^k de multiconjuntos con k elementos escogidos de un conjunto con n elementos satisface:

- Es igual al número de combinaciones con repetición de k elementos escogidos de un conjunto con n elementos.
- Es igual al número de formas de repartir k objetos en n grupos.

$$CR_n^k = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}.$$

Ejemplo: caramelos

Vamos a imaginar que vamos a repartir 12 caramelos entre Antonio, Beatriz, Carlos y Dionisio (que representaremos como A, B, C, D). Una posible forma de repartir los caramelos sería: dar 4 caramelos a Antonio, 3 a Beatriz, 2 a Carlos y 3 a Dionisio. Dado que no importa el orden en que se reparten, podemos representar esta selección como AAAABBBCCDDD.

Otra forma posible de repartir los caramelos podría ser: dar 1 caramelo a Antonio, ninguno a Beatriz y Carlos, los 11 restantes se los damos a Dionisio. Esta repartición la representamos como ADDDDDDDDDDD

Recíprocamente, cualquier serie de 12 letras A, B, C, D se corresponde a una forma de repartir los caramelos. Por ejemplo, la serie AAAABBBBBDDD corresponde a: Dar 4 caramelos a Antonio, 5 caramelos a Beatriz, ninguno a Carlos y 3 a Dionisio.

De esta forma, el número de formas de repartir los caramelos es:

$$CR_{n=4}^{k=12} = \binom{4+12-1}{12} = 455.$$

1.2.3 Variaciones.

Con los número $\{1,2,3\}$ ¿cuántos números de dos cifras distintas podemos formar sin repetir ninguna cifra?

La podemos escribir

$$12, 13, 21, 23, 31, 32$$

Luego hay seis casos, estas son las variaciones de orden $k = 2$ de un conjunto de $n = 3$ elementos.

Variaciones

Denotaremos las variaciones (sin repetición) de k elementos (de orden k) de un conjunto de n elementos por V_n^k su valor es

$$V_n^k = \frac{n!}{(n-k)!} = (n-k+1)\cdot(n-k+2)\cdots n.$$

Ejemplo

En nuestro ejemplo con $n = 3$ dígitos podemos escribir las siguientes variaciones de orden $k = 2$

$$V_{n=3}^{k=2} = \frac{3!}{(3-2)!} = \frac{1\cdot 2\cdot 3}{1} = 6.$$

Ejercicio

Carga el paquete `gtools` de R y investiga la función `permutations(n, r, v, set, repeats.allowed)` para calcular todas las variaciones anteriores.

1.2.4 Variaciones con repetición.

¿Y si en el caso anterior permitimos que se repita algún dígito?

Variaciones on repetición

Las variaciones de orden k de un conjunto de n elementos permitiendo que se repitan los elementos. Las denotamos y valen:

$$VR_n^k = n^k$$

Ejemplo

Efectivamente en nuestro caso

$$11, 12, 13, 21, 22, 23, 31, 32, 33$$

$$VR_{n=3}^{k=2} = n^k = 3^2 = 9.$$

1.2.5 Permutaciones

Permutaciones

Las permutaciones de un conjunto de cardinal n son todas las variaciones de orden máximo n. Las denotamos y valen:

$$P_n = V_n^n = n!$$

Variaciones on repetición

Por ejemplo todos los números que se pueden escribir ordenando todos los dígitos $\{1, 2, 3\}$ sin repetir ninguno

```
library(combinat)
for(permutacion in permn(3)) print(permutacion)
```

```
[1] 1 2 3
[1] 1 3 2
[1] 3 1 2
[1] 3 2 1
[1] 2 3 1
[1] 2 1 3
```

Efectivamente $P_3 = 3! = 1 \cdot 2 \cdot 3$.

Ejercicio

Carga el paquete `combinat` de R e investiga la función `permn` para calcular todas las permutaciones anteriores.

Investiga también el paquete `itertools` y la función `comb` de `scipy.misc` de Python e investiga sus funciones para todas las formas de contar que hemos visto en este tema.

Ejercicio

La función gamma de Euler, cobrará mucha importancia en el curso de estadística. Comprueba que la función `gamma(x+1)` da el mismo valor que la función `factorial(x)` en `R` para todo $x = \{1, 2, 3, \ldots, 10\}$.

1.2.6 Números multinomiales. Permutaciones con repetición.

Ejercicio

Consideremos un conjunto de elementos $\{a_1, a_2, \ldots, a_k\}$.

Entonces, si cada uno de los objetos a_i de un conjunto, aparece repetido n_i veces para cada i desde 1 hasta k, entonces el número de permutaciones con elementos repetidos es:

$$PR_n^{n_1,n_2,\ldots,n_k} = \binom{n}{n_1 \quad n_2 \quad \ldots \quad n_k} = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!},$$

donde $n = n_1 + n_2 + \cdots + n_k$.

Ejercicio

¿Cuantas palabras diferentes se pueden formar con las letras de la palabra `PROBABILIDAD`?

El conjunto de letras de la palabra considerada es el siguiente: $\{A, B, D, I, L, O, P, R\}$ con las repeticiones siguientes: las letras A, B, D, e I, aparecen 2 veces cada una; y las letras L, O, P, R una vez cada una de ellas.

Por tanto, utilizando la fórmula anterior, tenemos que el número de palabras (permutaciones con elementos repetidos) que podemos formar es

$$PR_{12}^{2,2,2,2,1,1,1,1} = \frac{12!}{(2!)^4(1!)^4} = 29937600.$$

1.3 Para acabar

1.3.1 Principios básicos para contar cardinales de conjuntos

El principio de la suma

Sean $A_1, A_2, \ldots, A_n$ conjuntos disjuntos dos a dos, es decir $A_i \cap A_j = \emptyset$ para todo $i \neq j$, $i, j = 1, 2, \ldots n$. Entonces

$$\#(\cup_{i=1}^n A_i) = \sum_{i=1}^n \#(A_i).$$

Principio de unión exclusión

Consideremos dos conjuntos cualesquiera A_1, A_2 entonces el cardinal de su unión es

$$\#(A_1 \cup A_2) = \#(A_1) + \#(A_2) - \#(A_1 \cap A_2).$$

El principio del producto

Sean $A_1, A_2, \ldots A_n$

$$\begin{aligned}\#(A_1 \times A_2 \times \cdots A_n) = & \ \# \left(\{(a_1, a_2, \ldots a_n) | a_i \in A_i, i = 1, 2, \ldots n\}\right) \\ & = \prod_{i=1}^{n} \#(A_i).\end{aligned}$$

1.3.2 Otros aspectos a tener en cuenta

Evidentemente nos hemos dejado muchas otras propiedades básicas de teoría de conjuntos y de combinatoria como:

- Propiedades de los números combinatorios.
- Binomio de Newton.
- Multinomio de Newton.

Si nos son necesarias las volveremos a repetir a lo largo del curso o bien daremos enlaces para que las podáis estudiar en paralelo.

2 Teoría elemental de la probabilidad

2.1 Definiciones básicas

Definición experimento aleatorio

Un experimento que repetido en las mismas condiciones puede dar resultados diferentes, pero que a largo plazo son predecibles recibe el nombre de **experimento aleatorio**.

Daremos nombres a distintos tipos de sucesos:

Espacio muestral y tipos de sucesos

- Llamaremos **suceso elemental** a cada uno de los posibles resultados del experimento aleatorio.
- Llamaremos **espacio muestral** (Ω,E) al conjunto formado por todos los sucesos elementales del experimento aleatorio.
- Llamaremos **suceso** a cualquier subconjunto del espacio muestral.
- **Suceso seguro o cierto** $A \subseteq \Omega$
- **Suceso imposible o vacio**: $\emptyset$
- **Partes de un conjunto**: $\mathcal{P}(\Omega)$: conjunto de todos los sucesos del experimento aleatorio (es decir, el conjunto de todos los subconjuntos de Ω)

A continuación describimos el clásico experimento del lanzamiento de un dado.

Ejemplo

Consideremos el experimento aleatorio que consiste en lanzar un dado. El espacio muestral de este experimento es $\Omega = \{1, 2, 3, 4, 5, 6\}$ o las figuras de las caras del dado.

Si lo representamos gráficamente, tendríamos:

Por comodidad y conveniencia se opta por representar el espacio muestral por

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

Recordemos la notación $\mathcal{P}(\Omega)$ que usamos para referirnos al conjunto de todos los subconjuntos de Ω. Este conjunto se llama **conjunto de partes** de Ω.

Ejercicio

¿Cuantos elementos contiene el conjunto de partes de Ω del experimento anterior?

Veamos algún ejemplo menos clásico. Podemos considerar el experimento aleatorio que consiste en calcular los n gramas de una palabra escogida al azar.

Ejemplo n-gramas

Se define un n-grama de una palabra como el conjunto de n letras consecutivas de la misma (contando los blancos de inicio y final de palabra que marcamos como "_").

Consideremos el experimento aleatorio que consiste en escoger al azar un 3-grama de la palabra "_Baleares_". Vamos a escribir el espacio muestral y algunos sucesos elementales del mismo.

En este caso, si consideramos la palabra "_Baleares_", el espacio muestral del experimento sería:

$$\Omega = \{\_Ba, Bal, ale, lea, ear, are, res, es\_\}$$

Algunos sucesos serían:

- 3-gramas que empiezan por a: $\{ale, are\}$.
- 3-gramas de inicio y final de palabra: $\{\_Ba, es\_\}$.
- 3-gramas que contengan una l: $\{Bal, ale, lea\}$.

Esiten bases de datos que estudián la frecuencias de n-grams de caracteres en textos en diferentes idiomas; generalmente de palabras. Por ejemplo, en español, los bigramas de sílabas más frecuentes son "EN" (3.01%) y "DE" (2.77%) y los trigramas de sílabas son "QUE" (1.66%) y "ENT" (1.38%). Podéis consultar más estadísticas en por ejemplo en Stefan Trost Media frecuencias de sílabas en español.

2.1.1 Operaciones con sucesos

Si tenemos dos sucesos $A, B \subseteq \Omega$, podemos definir:

- Ω: *suceso* total o *seguro*.
- $\emptyset$: suceso *vacío* o *imposible*.
- $A \cup B$: suceso *unión*; el que ocurre si sucede A o B.
- $A \cap B$: suceso *intersección*; el que ocurre si sucede A y B.
- A^c: suceso *complementario* el que sucede si NO sucede A.
- $A - B = A \cap B^c$: suceso *diferencia*, que acontece si sucede A y NO sucede B.

Sucesos incompatibles

Dos sucesos cualesquiera A y B son *incompatibles* (o *disjuntos*) cuando $A \cap B = \emptyset$.

Otro ejemplo se observa el sexo y la lateralidad de los estudiantes de una clase.

Ejemplo

Supongamos que el sexo se divide entre Mujeres y Hombres y la lateralirad en diestros y zurdos. Vamos a definir el espacio muestral, los sucesos elementales y a realizar algunas operaciones entre ellos.

Estudiantes de esta clase: Ω. - Mujeres de esta clase: A. - Estudiantes que son zurdos B.

Algunas operaciones entre los sucesos anteriores serían:

- $A \cup B$: Est. que son mujeres o que son zurdos.
- $A \cap B$: Mujeres de esta clase que son zurdas.
- A^c: Hombres de esta clase.
- $A - B$: Mujeres de la clases que NO son zurdas.
- $B - A$: Hombres de la clase que son zurdos.
- ¡Cuidado! No son incompatibles.

2.1.2 Propiedades

Propiedades

Conmutativas:

$$A \cup B = B \cup A, \quad A \cap B = B \cap A$$

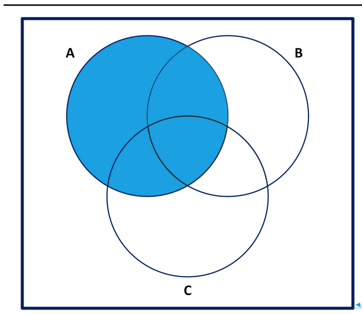
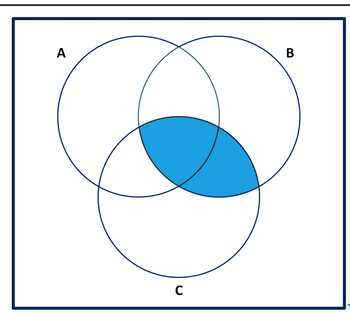
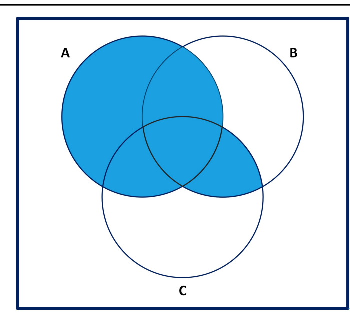
Asociativas:

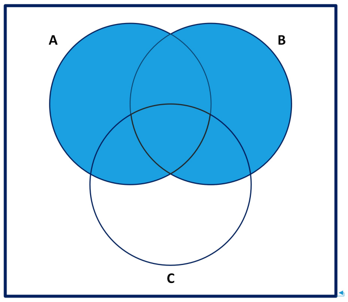
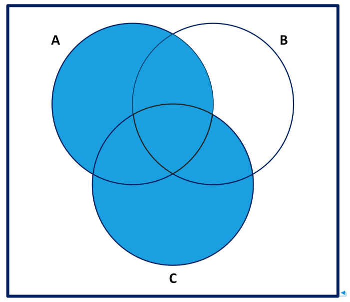
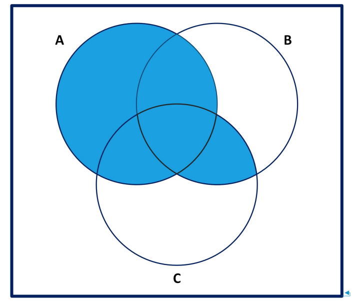
$$\begin{aligned} A \cup (B \cup C) &= (A \cup B) \cup C, \\ A \cap (B \cap C) &= (A \cap B) \cap C. \end{aligned}$$

Distributivas

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned}$$

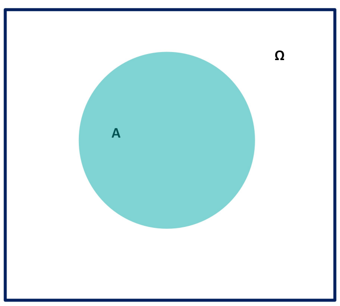
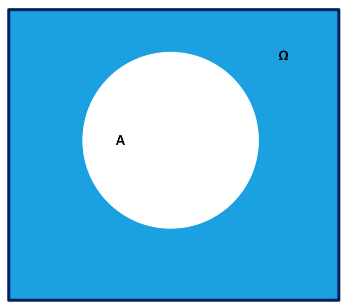
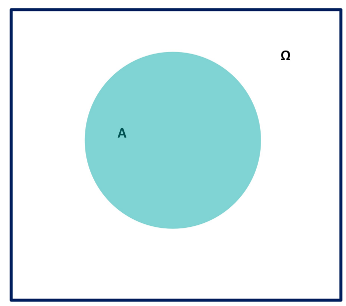
Veamos algunos diagramas que nos ayuda a demostrar las propiedades anteriores.

A	$B \cap C$	$A \cup (B \cap C)$
		

$A \cup B$	$A \cup C$	$(A \cup B) \cap (A \cup C)$
		

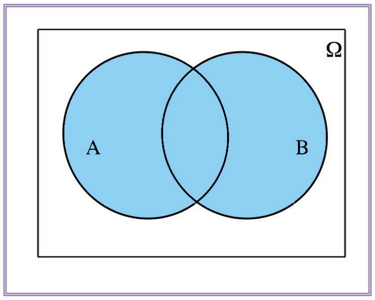
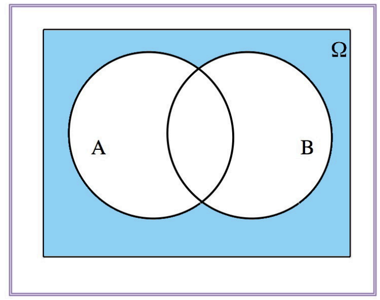
Complementario del complementario

$$(A^c)^c = A$$

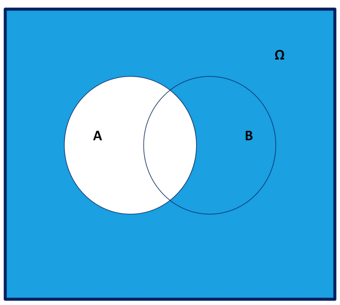
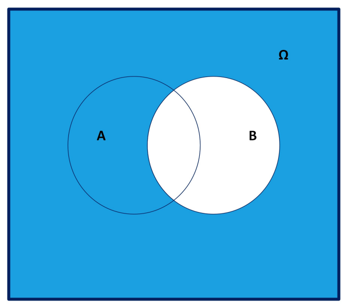
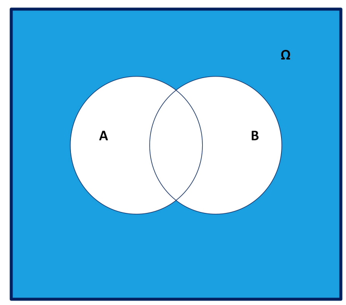
A	A^c	$(A^c)^c$
		

Leyes de De Morgan

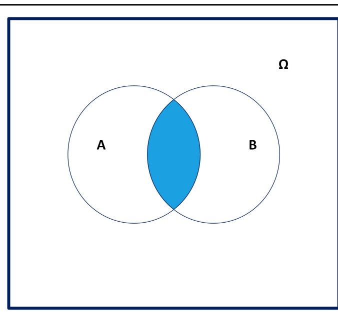
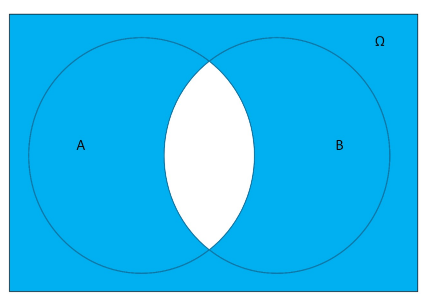
$$(A \cup B)^c = A^c \cap B^c$$

$A \cup B$	$(A \cup B)^c$
	

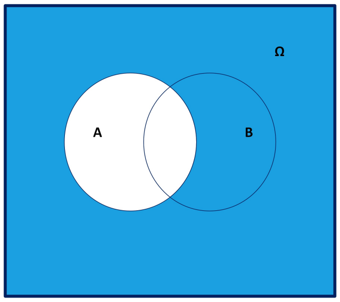
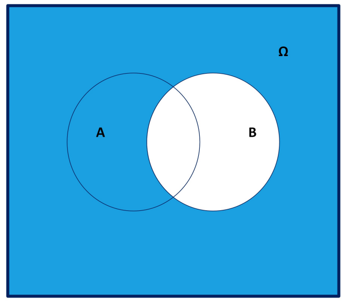
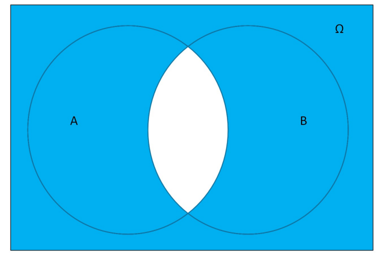
$$(A \cup B)^c = A^c \cap B^c$$

A^c	B^c	$A^c \cap B^c$
		

$$(A \cap B)^c = A^c \cup B^c$$

$A \cap B$	$(A \cap B)^c$
	

$$(A \cap B)^c = A^c \cup B^c$$

A^c	B^c	$A^c \cup B^c$
		

2.2 Definición de probabilidad

La probabilidad de un suceso es una puntuación (*score*) numérico entre 0 y 1 que mide la verosimilitud de que este evento se produzca.

Esta verosimilitud puede estar justificada por:

- Estimación personal
- Estimación de expertos
- La frecuencia con la que se da
- Cálculo formal

Definición formal de probabilidad

Sea Ω el espacio muestral de un experimento aleatorio. Supongamos que el número de posibles resultados, por el momento, es finito.

Una probabilidad sobre Ω es una aplicación $P : \mathcal{P}(\Omega) \to [0, 1]$ con las siguientes propiedades:

1. $0 \leq P(A) \leq 1$, para todo suceso A.
2. $P(\Omega) = 1$.
3. Si $\{A_1, A_2, \dots, A_n\}$ son sucesos disjuntos dos a dos, entonces

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

Si $a \in \Omega$ es un suceso elemental cometeremos el abuso de notación de poner $P(a)$ en lugar de $P(\{a\})$.

Veamos un ejemplo real de cómo se calcula la probabilidad de un suceso.

Ejemplo

En la página de la Fundación Banco de Sangre y Tejidos de las Islas Baleares (17-08-2023) podemos encontrar información sobre los porcentajes de tipos de sangre de los donantes de las Islas Baleares:

$$A : 46\%;\ B : 7.5\%;\ AB : 3.5\%;\ O : 43\%.$$

¿Cuál es la probabilidad de que un balear donante de sangre no sea del tipo O?

Experimento aleatorio: tipo de sangre de un paciente humano:

$$\Omega = \{\text{A,B,AB,O}\}$$

Probabilidad de un suceso: se asimila al porcentaje observado de individuos.

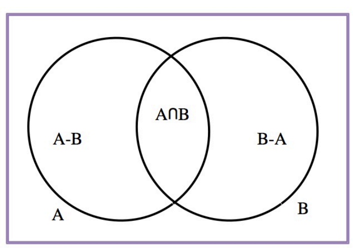
Suceso: $\{\text{O}\}^c = \{\text{A,B,AB}\}$.

$$P(\{\text{O}\}^c) = P(\{\text{A,B,AB}\}) = P(\text{A}) + P(\text{B}) + P(\text{AB}) = 0.57.$$

Necesitaremos tener propiedades y fórmulas prácticas para poder calcular probabilidades de sucesos más complejos. Veamos algunas de ellas.

Propiedades básicas de la probabilidad

- $P(\emptyset) = 0$.
- $P(A - B) = P(A) - P(A \cap B)$ porque $P(A) = P(A - B) + P(A \cap B)$.

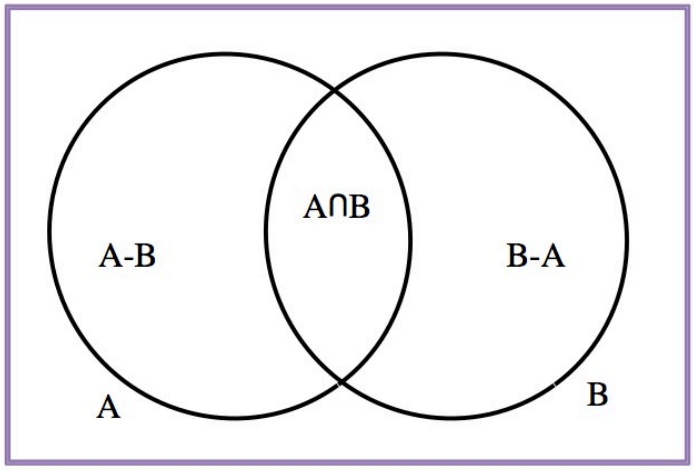


- Si $B \subseteq A$, entonces $0 \leq P(B) \leq P(A)$.
- $P(A^c) = 1 - P(A)$.

Una identidad muy utilizada es la de la probabilidad de la unión de dos sucesos cualesquiera.

La Probabilidad de la unión de dos sucesos

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



La demostración analítica de esta propiedad es la siguiente:

$$
\begin{aligned}
P(A) + P(B) - P(A \cap B) &= P(A - B) + P(A \cap B) \\
&\quad + P(B - A) + P(A \cap B) - P(A \cap B) \\
&= P(A - B) + P(A \cap B) + P(B - A) \\
&= P(A \cup B).
\end{aligned}
$$

Probabilidad de la unión de n conjuntos

Sean $A_1, A_2, \ldots A_n$ sucesos. Entonces:

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n).$$

La demostración es sncilla por inducción, ya tenemos el caso de dos sucesos, suponemos que es cierta para n y la extendemos a $n+1$ sucesos.