

Tema 2. Fundamentos de probabilidad

Estadística
Grado en Ingeniería Informática

Índice

1. Experimento aleatorio. Sucesos y espacio muestral	1
2. Operaciones con sucesos	2
2.1. Suceso complementario	2
2.2. Unión de sucesos	2
2.3. Intersección de sucesos	3
2.4. Suceso diferencia	4
2.5. Leyes de De Morgan	4
2.6. Sistema exhaustivo. Sistema completo	5
3. Asignación y definición de probabilidad	5
3.1. Definición axiomática de la probabilidad	7
3.2. Probabilidad condicionada	7
3.3. Sucesos independientes	8
4. Resultados importantes	8
4.1. Regla del producto	8
4.2. Teorema de las Probabilidades Totales	9
4.3. Teorema de Bayes	10

En este tema estudiaremos los conceptos fundamentales de la Teoría de la Probabilidad, ilustrándolos desde la perspectiva de la Teoría de Conjuntos. Comenzaremos describiendo datos obtenidos de experimentos aleatorios, que son aquellos en los que interviene el azar, y analizando los sucesos que pueden surgir, a través de operaciones básicas como la unión o la intersección. Pasaremos posteriormente a definir cómo asignar probabilidad a dichos sucesos, centrándonos en la perspectiva frecuentista y la axiomática de Kolmogorov. Finalmente proporcionaremos una serie de resultados notables de extensa aplicación para la resolución de problemas probabilísticos.

1. Experimento aleatorio. Sucesos y espacio muestral

Debemos distinguir entre experimentos deterministas, en los que existe una relación de causa-efecto, y experimentos aleatorios.

- **Experimento aleatorio:** es aquel experimento que, repetido sucesivas veces en condiciones idénticas, produce resultados diferentes e imprevisibles.
- **Experimento determinista:** es aquel experimento que repetido sucesivamente en condiciones idénticas siempre produce los mismos resultados. Un experimento determinista puede volverse aleatorio si se introduce un error asociado, por ejemplo, un error de medida.

Para caracterizar un experimento aleatorio, necesitamos, por un lado, conocer cuáles son los posibles resultados del mismo, y por otro, con qué probabilidad puede ocurrir cada uno de ellos. Los resultados de un experimento aleatorio se denominan sucesos.

- **Suceso elemental:** es cada uno de los posibles resultados de un experimento aleatorio.
- **Espacio muestral:** es el conjunto de todos los sucesos elementales. Se denotará por Ω .
- **Suceso:** es cualquier subconjunto del espacio muestral Ω . En particular, Ω se denomina suceso seguro, mientras que \emptyset (el conjunto vacío) es el suceso imposible. Al conjunto formado por todos los sucesos asociados a un experimento aleatorio lo denotamos por \mathcal{A} .

Ejemplo. Ilustraremos estos conceptos a través de un ejemplo de experimento aleatorio muy sencillo: el lanzamiento de un dado. Son sucesos elementales los posibles resultados (es decir: 1, 2, 3, 4, 5 y 6). El espacio muestral será $\Omega = \{1, 2, 3, 4, 5, 6\}$. Pero estos no son los únicos sucesos. Así, podemos considerar $A = \text{par} = \{2, 4, 6\}$ o $B = \text{mayor que 3} = \{4, 5, 6\}$.

2. Operaciones con sucesos

A partir de los sucesos elementales de un experimento, se pueden definir otros sucesos derivados a través del complementario, la unión o la intersección de sucesos. Dichas operaciones se pueden ilustrar a través de teoría de conjuntos, manejando conceptos básicos que ya se han estudiado en otras materias del curso.

2.1. Suceso complementario

Sea un suceso $A \in \mathcal{A}$. El suceso complementario de A es el que ocurre cuando no ocurre A y se denota por A^c o \bar{A} . En la Figura 1, el suceso A está señalado en color verde, mientras que $A^c = \Omega \setminus A$ es el conjunto de rayas.

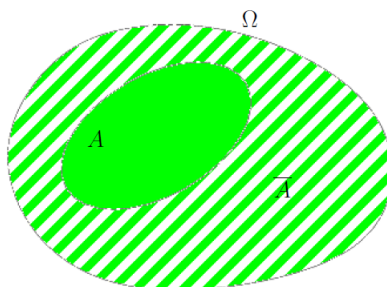


Figura 1. Suceso complementario.

Ejemplo. En el ejemplo del dado, tenemos que: $A^c = \{1, 3, 5\}$ y $B^c = \{1, 2, 3\}$ (es decir, sacar impar es el complementario de A y sacar menor o igual que 3 es el complementario de B).

2.2. Unión de sucesos

Consideremos $A, B \in \mathcal{A}$ dos sucesos. El suceso unión se denota por $A \cup B$ y es lo que ocurre cuando sucede A o sucede B .

La unión de sucesos es conmutativa ($A \cup B = B \cup A$). Además, si consideramos \emptyset , el suceso imposible, tenemos que $A \cup \emptyset = A$ y por otra parte $A \cup A^c = \Omega$. La unión de los sucesos A y B está representada en la Figura 2, parte (a).

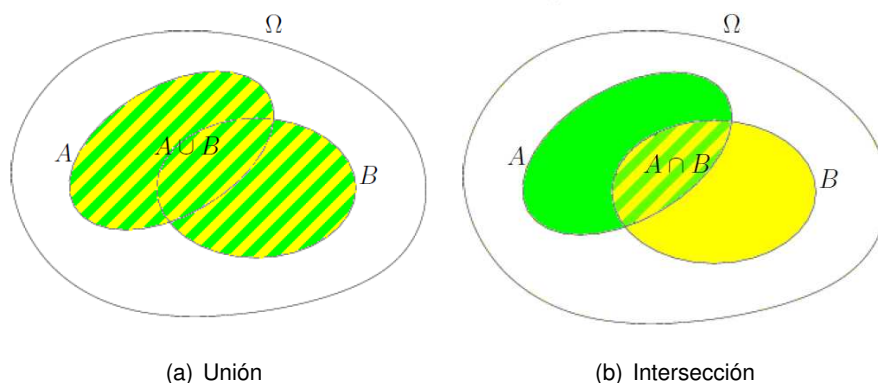


Figura 2. Unión e Intersección de sucesos.

2.3. Intersección de sucesos

Sean $A, B \in \mathcal{A}$ dos sucesos. El suceso intersección se denota como $A \cap B$, y es lo que ocurre cuando sucede A y también sucede B .

La intersección de sucesos es conmutativa ($A \cap B = B \cap A$). Si consideramos el suceso seguro Ω , tenemos que $A \cap \Omega = A$. Además, $A \cap A^c = \emptyset$. La intersección de A y B la representamos en la Figura 2, parte (b).

Tanto la unión como la intersección son asociativas. Si $A, B, C \in \mathcal{A}$, se verifica:

$$A \cup (B \cup C) = (A \cup B) \cup C, \quad \text{y} \quad A \cap (B \cap C) = (A \cap B) \cap C.$$

Además:

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

Ejemplo. La unión $A \cup B = \{2, 4, 6\} \cup \{4, 5, 6\} = \{2, 4, 5, 6\}$, es decir, sacar par o sacar mayor que 3. Por otro lado, la intersección es $A \cap B = \{2, 4, 6\} \cap \{4, 5, 6\} = \{4, 6\}$ (sacar par y sacar mayor que 3).

2.3.1. Sucesos incompatibles

Sean $A, B \in \mathcal{A}$ dos sucesos. Se dice que son **sucesos incompatibles** si $A \cap B = \emptyset$. Un ejemplo de sucesos incompatibles se representa en la Figura 3. Por definición, A y A^c (en general, un suceso y su complementario) son incompatibles.

2.4. Suceso diferencia

Sean $A, B \in \mathcal{A}$ dos sucesos. El suceso diferencia se define como el que ocurre si sucede A y no sucede B y se denota por $A \setminus B = A \cap B^c$. En la Figura 3 se representa el suceso diferencia entre A y B . Se puede observar que el suceso complementario de A se define como la diferencia $\Omega \setminus A$.

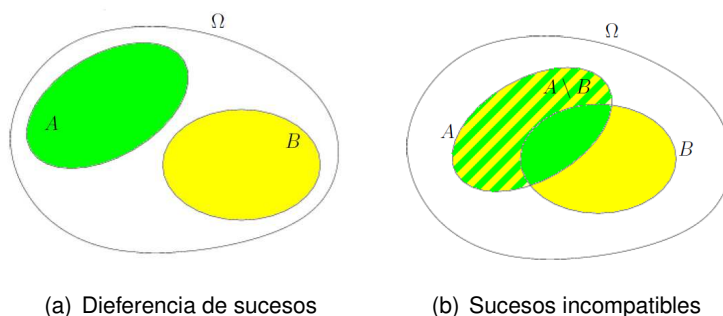


Figura 3. Sucesos incompatibles y diferencia de sucesos.

Ejemplo. En el ejemplo, la diferencia $A \setminus B$ sería el resultado de sacar par y no sacar mayor que 3. Es decir: $A \setminus B = A \cap B^c = \{2, 4, 6\} \cap \{1, 2, 3\}^c = \{2\}$.

2.5. Leyes de De Morgan

Las **Leyes de De Morgan** (Augustus De Morgan, 1806-1871) permiten intercambiar la unión y la intersección a través del complementario, y serán útiles en el cálculo de probabilidades de sucesos.

1. Primera Ley de De Morgan (Figura 4): $(A \cup B)^c = A^c \cap B^c$, para $A, B \in \mathcal{A}$.
2. Segunda Ley de De Morgan (Figura 5): $(A \cap B)^c = A^c \cup B^c$, para $A, B \in \mathcal{A}$.

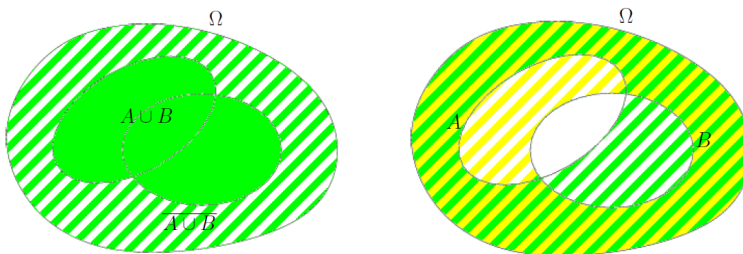


Figura 4. Primera Ley de De Morgan: complementario de la unión como intersección de complementarios.

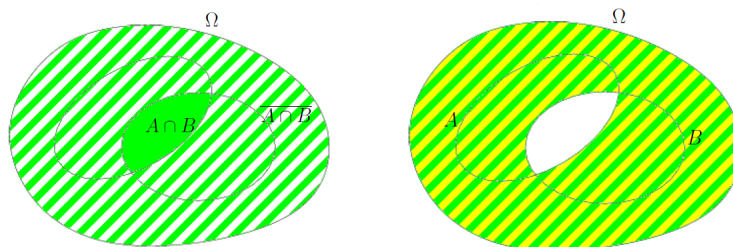


Figura 5. Segunda Ley de De Morgan: complementario de la intersección como unión de complementarios.

Ejemplo. En el ejemplo, tenemos que: $A \cup B = \{2, 4, 5, 6\} \Rightarrow (A \cup B)^c = \{1, 3\}$. Podemos comprobar que este resultado coincide con: $A^c \cap B^c = \{1, 3, 5\} \cap \{1, 2, 3\} = \{1, 3\}$. Para la segunda ley, podríamos hacer una comprobación similar. Así: $(A \cap B)^c = \{1, 2, 3, 5\}$ y $A^c \cup B^c = \{1, 3, 5\} \cup \{1, 2\} = \{1, 2, 3, 5\}$.

2.6. Sistema exhaustivo. Sistema completo

Para la aplicación de los resultados que introduciremos en las siguientes secciones, es importante considerar dos tipos de sistemas de sucesos: los sistemas exhaustivos y, dentro de estos, los sistemas completos.

- **Sistema exhaustivo de sucesos:** una familia de sucesos $\{A_1, \dots, A_n : A_i \in \mathcal{A}\}$ es un sistema exhaustivo de sucesos si $\bigcup_{i=1}^n A_i = \Omega$, es decir, si su unión cubre el espacio muestral.
- **Sistema completo de sucesos:** una familia de sucesos $\{A_1, \dots, A_n : A_i \in \mathcal{A}\}$ es un sistema completo si es exhaustivo, y además $A_i \cap A_j = \emptyset$, para todo $i \neq j$. En la Figura 6, la familia $\{A_1, A_2, A_3, A_4\}$ es un sistema exhaustivo, que además es completo.

Ejemplo. Si definimos $C = \{1, 3\}$, tenemos que la colección de conjuntos $\{A, B, C\}$ es un sistema exhaustivo, ya que su unión coinciden con Ω (todos los posibles resultados). Sin embargo, no es un conjunto completo de sucesos, puesto que A y B tienen intersección no vacía.

3. Asignación y definición de probabilidad

En el estudio de experimentos aleatorios no solo es importante conocer los posibles resultados sino también saber con qué probabilidad ocurre cada uno de ellos. Es por ello que, una vez definidos los po-

sibles sucesos se les debe asignar una probabilidad. Introduciremos primero la definición frecuentista y la asignación de Laplace, que serán las que consideraremos desde el punto de vista práctico.

Probabilidad. Definición de Laplace (método clásico). Supongamos que Ω tiene un número finito de sucesos y que todos los sucesos $\omega \in \Omega$ tienen la misma probabilidad. La probabilidad del suceso A , que denotaremos por $\mathbb{P}(A)$, se calcula como el cociente entre *casos favorables* y *casos posibles*.

Ejemplo. El experimento del lanzamiento del dado es un claro ejemplo donde todos los sucesos tienen la misma probabilidad y donde, por tanto, podemos aplicar la regla de Laplace para su asignación. En concreto, la probabilidad de cada uno de los sucesos elementales es $1/6$ (1 caso favorable para 6 casos posibles). Tenemos también que $\mathbb{P}(A) = 3/6 = 0.5$ y $\mathbb{P}(B) = 3/6 = 0.5$.

Probabilidad. Definición frecuentista. Si repetimos n veces un experimento aleatorio y n_A es el número de veces que ocurre el suceso A , la frecuencia relativa de este suceso se define como $fr(A) = n_A/n$, donde $0 \leq fr(A) \leq 1$. Al hacer n grande, esta frecuencia se estabiliza, y se define la probabilidad del suceso A como el límite de las frecuencias.

Ejemplo. Consideremos el siguiente ejemplo de Fisher (1940) donde presentaba en una tabla de contingencia los datos de color de pelo y ojos de escolares escoceses. Podemos calcular:

	rubio	pelirrojo	castaño	oscuro	negro	
claros	688	116	584	188	4	1580
azules	326	38	241	110	3	718
castaños	343	84	909	412	26	1774
oscuros	98	48	403	681	85	1315
Total	1455	286	2137	1391	118	5387

Tabla 1. Color de pelo y ojos (Fisher, 1940)

- $\mathbb{P}(\text{"ojos castaños"}) = 1774/5387$, $\mathbb{P}(\text{"rubio"}) = 1455/5387$.
- $\mathbb{P}(\text{"rubio"} \text{ y } \text{"ojos castaños"}) = 343/5387$.
- $\mathbb{P}(\text{"rubio"} \text{ o } \text{"castaño"}) = (1455 + 2137)/5387$.
- $\mathbb{P}(\text{"rubio"} \text{ o } \text{"ojos castaños"}) = (1455 + 1774 - 343)/5387 = 2886/5387$.

3.1. Definición axiomática de la probabilidad

Consideremos el par (Ω, \mathcal{A}) , que denominaremos espacio probabilizable. Se dice que \mathbb{P} es una probabilidad sobre (Ω, \mathcal{A}) si verifica los siguientes axiomas:

- (i) La probabilidad del espacio muestral es uno: $\mathbb{P}(\Omega) = 1$.
- (ii) La probabilidad de la unión de sucesos *mutuamente disjuntos* es igual a la suma de las probabilidades de cada uno: si $\{A_1, A_2, \dots\}$ son tales que $A_i \cap A_j = \emptyset$ si $i \neq j$:

$$\mathbb{P}(A_1 \cup A_2 \cup \dots) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

- (iii) Todo suceso tiene probabilidad entre cero y uno: $0 \leq \mathbb{P}(A) \leq 1$, para todo $A \in \mathcal{A}$.

Podemos comprobar que las siguientes propiedades de la probabilidad \mathbb{P} se derivan de los axiomas:

- La probabilidad del vacío es nula: $\mathbb{P}(\emptyset) = 0$.
- La probabilidad del complementario se obtiene como: $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
- Si $A \subset B$, entonces $\mathbb{P}(A) \leq \mathbb{P}(B)$ y $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$.
- Para dos sucesos cualesquiera, la probabilidad de la unión es:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Ejemplo. En el ejemplo de los datos de Fisher, puedes comprobar que:

$$\mathbb{P}(\text{"rubio" o "ojos marrones"}) = \mathbb{P}(\text{"rubio"}) + \mathbb{P}(\text{"ojos marrones"}) - \mathbb{P}(\text{"rubio" y "ojos marrones"})$$

3.2. Probabilidad condicionada

La probabilidad de un suceso A condicionada a un suceso B es la probabilidad de que ocurra A sabiendo que B sí ha ocurrido. Esta situación se puede ilustrar en experimentos aleatorios se realizan en dos o más fases: sabiendo lo sucedido en una de las fases, la probabilidad de los sucesos en la otra fase es una probabilidad condicionada. También se puede ver claramente en tablas de contingencia como la del ejemplo anterior sobre el color de pelo y ojos de los escolares escoceses.

Sean $A, B \in \mathcal{A}$ dos sucesos con $\mathbb{P}(B) > 0$. Formalmente, se define la probabilidad del suceso A condicionada al suceso B como:

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

Ejemplo. En el ejemplo de los datos de Fisher, podemos calcular:

- Probabilidad de que alguien con ojos azules sea rubio:

$$\mathbb{P}(\text{Rubio}|\text{Ojos azules}) = 326/718.$$

- Probabilidad de que alguien con ojos oscuros sea rubio:

$$\mathbb{P}(\text{Rubio}|\text{Ojos oscuros}) = 98/1315.$$

3.3. Sucesos independientes

Los sucesos A y B son independientes si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Esto es equivalente a decir que $\mathbb{P}(A|B) = \mathbb{P}(A)$, si $\mathbb{P}(B) \neq 0$. Además, si A y B son independientes, entonces los siguientes pares también lo son: A y B^c ; A^c y B ; A^c y B^c .

4. Resultados importantes

4.1. Regla del producto

A partir de la definición de probabilidad condicionada, dados dos sucesos $A, B \in \mathcal{A}$ con $\mathbb{P}(A) > 0$ ¹, podemos obtener la probabilidad de la intersección como:

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

Del mismo modo, si tuviésemos tres sucesos $A, B, C \in \mathcal{A}$, de tal modo que $\mathbb{P}(A \cap B) > 0$:

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

e igualmente, si $\mathbb{P}(A \cap B \cap C) > 0$:

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

Por tanto, la probabilidad de la intersección, se puede obtener como producto de las probabilidades condicionadas por los sucesos *anteriores*.

¹Observa que en este caso, el suceso que condiciona es A .

Si esto lo generalizamos al caso de n sucesos, obtenemos la Regla del Producto. Para establecer este resultado, es necesario que la intersección de los $(n - 1)$ sucesos primeros tenga probabilidad positiva. Esto es debido a que, para poder condicionar, necesitamos un suceso de probabilidad no nula. Además, si la probabilidad de la intersección de los $(n - 1)$ primeros es positiva, también lo será la probabilidad de las intersecciones de menos sucesos.

Consideremos el conjunto de sucesos $\{A_1, \dots, A_n\}$, tales que la probabilidad de la intersección $\mathbb{P}(\cap_{i=1}^{n-1} A_i) > 0$. Entonces:

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

4.2. Teorema de las Probabilidades Totales

Si $\{A_1, \dots, A_n\}$ es un conjunto completo de sucesos, con $\mathbb{P}(A_i) > 0$, para todo $i = 1, \dots, n$ y $B \in \mathcal{A}$ es un suceso cualquiera, entonces:

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

Una representación gráfica del resultado se muestra en la Figura 6, donde podemos ver que el suceso B puede representarse como unión de cada una de sus intersecciones con los A_i . Como los A_i son disjuntos, la probabilidad de B puede escribirse como suma de las probabilidades de las intersecciones $\mathbb{P}(B \cap A_i)$, y como $\mathbb{P}(A_i) > 0$ la probabilidad de cada intersección puede obtenerse a partir de la probabilidad condicionada $\mathbb{P}(B|A_i)$.

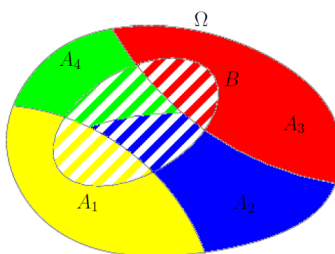


Figura 6. Ilustración del teorema de las Probabilidades Totales.

El Teorema de Probabilidades Totales se aplica para conocer la probabilidad de un suceso, cuando tenemos definido un conjunto completo de sucesos. Sin embargo, puede ocurrir que conozcamos la probabilidad del suceso B y que nos interese, a partir de ella, obtener información sobre los sucesos que forman el conjunto completo, es decir, obtener $\mathbb{P}(A_i|B)$. Para ello utilizamos el Teorema de Bayes.

4.3. Teorema de Bayes

Si $\{A_1, \dots, A_n\}$ es un conjunto completo de sucesos, con $\mathbb{P}(A_i) > 0$, para todo $i = 1, \dots, n$ y $B \in \mathcal{A}$ es un suceso con $\mathbb{P}(B) > 0$, entonces:

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