Tema 2

Cálculo de Probabilidades

2.1 Introducción

El estudio matemático de los fenómenos aleatorios requiere una definición formal de conceptos que permitan estructurar una teoría consistente.

Un **experimento aleatorio** es aquel cuyo resultado no puede predecirse de antemano (aunque se conocen todos los resultados posibles) y que además, repitiéndose en idénticas condiciones iniciales pueden obtenerse resultados distintos.

El conjunto de todos los posibles resultados de un experimento aleatorio se denomina **espacio muestral**, Ω.

2.1.1 Sucesos. Operaciones con sucesos

- Un subconjunto cualquiera del espacio muestral recibe el nombre de **suceso**. Como cualquier subconjunto, los sucesos se designan utilizando letras mayúsculas.

  Existen distintos tipos de sucesos.

  - **Suceso elemental:** Es un subconjunto formado por un único elemento, o un suceso formado por un único resultado.

    Por ejemplo, si consideramos el experimento "Lanzar un dado", con espacio muestral $\Omega = \{1, 2, 3, 4, 5, 6\}$, el suceso "sacar un 2", que con notación conjuntista sería $A = \{2\}$, es un suceso elemental.

  - **Suceso compuesto:** Es un subconjunto formado por más de un elemento, o un suceso formado por varios resultados elementales.

    Por ejemplo, el suceso $B = \{2, 4, 6\}$ ("sacar un número par), es un suceso compuesto.

  - **Suceso imposible:** Es un suceso que no puede obtenerse en ningún caso como resultado del experimento. Se representa con el símbolo $\emptyset$.

    Un ejemplo de suceso imposible es "sacar un siete".

- **Suceso seguro:** Es un suceso que siempre se va a producir, es decir, es el suceso compuesto formado por todos los elementos del espacio muestral, (o todo el espacio muestral).

  Un ejemplo de suceso seguro es "sacar un número mayor o igual que 1".

Decimos que un suceso se verifica cuando el resultado del experimento es alguno de sus elementos. Por ejemplo, si al lanzar el dado obtenemos un 2, se verifican el suceso A y el suceso B, y si obtenemos un 4, se verifica el suceso B.

2.1.2 Operaciones con sucesos

Puesto que los sucesos son subconjuntos, pueden realizarse con ellos las mismas operaciones que con los conjuntos.

Continuando con el ejemplo del experimento consitente en lanzar un dado, consideremos los sucesos siguientes:
$A = \{2\}$ ("sacar un 2")
$B = \{2, 4, 6\}$ ("sacar un número par")
$C = \{1, 3, 5\}$ ("sacar un número impar")
$D = \{4, 5, 6\}$ ("sacar un número mayor o igual que 4", o "sacar al menos un 4")
$F = \{3, 4\}$ ("sacar un 3 o un 4")

- **Intersección de sucesos**

  Dados dos sucesos A y B, se llama **suceso intersección**, $A \cap B$, al formado por los elementos que pertenecen a A y a B.

  Decimos que se verifica el suceso $A \cap B$ cuando se verifican simultáneamente A y B.

  Cuando $A \cap B = \emptyset$, se dice que A y B son **sucesos incompatibles** (ya que no pueden suceder simultáneamente).

  En nuestro caso, dado que $A \subset B$, obviamente $A \cap B = \{2\}$, o equivalentemente, $A \cup B = 2$.

  Otros ejemplos serían:

  - $B \cap C = \emptyset$ (es imposible obtener al mismo tiempo un número par y un número impar).
  - $F \cap D = \{3\}$.
  - $\Omega \cap D = \{4, 6\}$ (podría describirse como el suceso "sacar un cuatro o un número seis", o "sacar un número par mayor que 2").
  - $C \cap D = \{5\}$ (podría describirse como el suceso sacar un número impar mayor que 3").

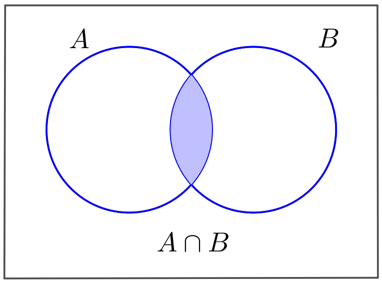


Figura 2.1: Intersección de sucesos

- **Unión de sucesos**

  Dados dos sucesos A y B, se llama **suceso unión**, $A \cup B$, al formado por todos los elementos que pertenecen a A o a B.

  Decimos que se verifica el suceso $A \cup B$ cuando se verifica A o B.

  En nuestro caso, dado que A está contenido en B ($A \subset B$), obviamente $A \cup B = \{2, 4, 6\}$, o equivalentemente, $A \cup B = B$.

  Observación: En general, en el suceso $A \cup B$ hay elementos que pertenecen a A y a B (es decir, al suceso $A \cap B$).

  Cuando se quiere expresar el suceso formado por los elementos que pertenecen a A o a B, pero que no pertenecen a ambos a la vez, se habla de **unión disjunta** de sucesos. $A \overset{\circ}{\cup} B$

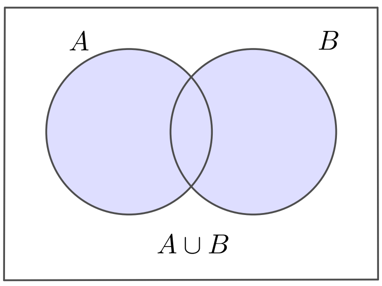


Figura 2.2: Unión de sucesos

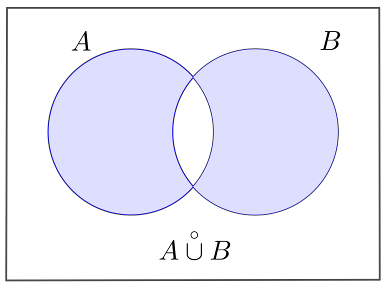


Figura 2.3: Unión disjunta de sucesos

Algunos ejemplos serían:

- $B \cup C = \Omega$ (es decir, el suceso sacar un número par o un número impar coincide con el suceso seguro).
- $A \cup C = \{1, 2, 3, 5\}$ (podría describirse como el suceso "sacar un dos o un número impar").

- $A \cup F = \{2,3,4\}$ (podría describirse como el suceso sacar un número mayor o igual que 2 y menor o igual que 4).
- $B \overset{\circ}{\cup} D = \{2,5\}$

- **Suceso contrario**

  Dado un suceso A, se llama **suceso contrario**, $\overline{A}$, al suceso que se verifica cuando no se verifica A, es decir, al formado por todos los elementos de Ω que no pertenecen a A.

  Esto también podemos expresarlo como $\overline{A} = \Omega \setminus A$.

  En nuestro caso, el suceso contrario a A sería "no sacar un 2", es decir $\overline{A} = \{1,3,4,5,6\}$.

  Otros ejemplos serían:

  - $\overline{B} = \{1,3,5\}$ (lo contrario de "sacar un número par" es "sacar un número impar").
  - $\overline{D} = \{1,2,3\}$ (lo contrario de "sacar un número mayor o igual que 4" es "sacar un número menor que 4").

- **Diferencia de sucesos**

  Dados dos sucesos A y B, se llama **suceso diferencia**, $B \setminus A$, al formado por los elementos que pertenecen a B y no pertenecen a A.

  Decimos que se verifica el suceso $B \setminus A$ cuando se verifica B y no se verifica A.

  Es decir, se tiene la igualdad conjuntista $B \setminus A = B \cap \overline{A}$

  En nuestro caso, obviamente $B \setminus A = \{4,6\}$.

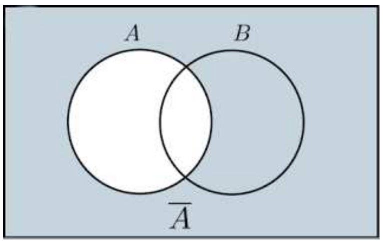


Figura 2.4: Suceso contrario

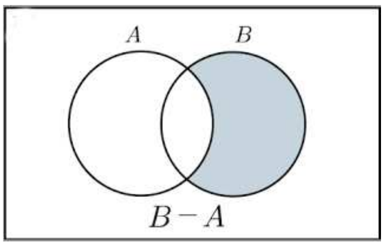


Figura 2.5: Diferencia de sucesos

- **Leyes de Morgan**

  Se utilizan para facilitar la obtención del resultado de operaciones con sucesos:

  - $\overline{A \cup B} = \overline{A} \cap \overline{B}$

    Es decir, si no se obtiene como resultado del experimento un algún elemento de A o de B, es porque no se produjo ningún resultado que sea elemento de A, ni tampoco ningún resultado que sean elemento de B. O equivalentemente, si no se verifica A ni se verifica B, no se verifica el suceso "A o B".

– $\overline{A \cap B} = \overline{A} \cup \overline{B}$
Es decir, si no se obtiene como resultado del experimento ningún elemento que pertenezca a la vez a A y a B, es porque o no se verifica A, o no se verifica B.

2.2 Las leyes del azar: Ley de los grandes números

El cálculo de probabilidades surgió de la necesidad de querer medir la incertidumbre y de querer cuantificar la facilidad con la que un hecho podía producirse. Es decir, de la necesidad de buscar qué leyes rigen el azar.

El concepto de frecuencia relativa nos permite una aproximación intuitiva del concepto de probabilidad. Jacques Bernouilli en el siglo XVIII demostró la "Ley de los grandes números", que establece que la frecuencia relativa de un suceso tiene a estabilizarse en torno a un valor dado cuando el número de pruebas crece indefinidamente. A dicho valor límite se le llamó probabilidad del suceso.

$$\mathcal{P}(A) = \lim_{n \to +\infty} \frac{f_A}{n} = \lim_{n \to +\infty} h_A$$

A esta definición de probabilidad se la conoce como definición frecuentista.

2.3 La Ley de Laplace

La definición frecuentista de probabilidad presentaba un inconveniente de tipo práctico, ya que para conocer la probabilidad de un suceso sería necesario efectuar un gran número de pruebas para obtener realmente un número aproximado.

La primera solución a este problema fue planteada por Laplace, que definió la probabilidad de un suceso A como el número de resultados favorables a la realización de A, dividido entre el número de resultados posibles del experimento. Es decir:

$$\mathcal{P}(A) = \frac{\text{Número de casos favorables}}{\text{Número de casos favorables}}$$

Para poder calcular probabilidades de sucesos por medio de la Ley de Laplace, deben satisfacerse las siguientes condiciones:

- El número de posibles resultados del experimento (número de sucesos elementales del espacio muestral) debe ser finito.

- Todos los posibles resultados del experimento deben tener la misma probabilidad. Es decir, todos los sucesos elementales del experimento deben ser equiprobables (postulado de indiferencia o equiprobabilidad)

2.4 Definición axiomática de probabilidad

A principios del siglo XX, utilizando conceptos de Teoría de Conjuntos y Teoría de la Medida, fue posible la construcción de un modelo teórico para el estudio de la probabilidad.

Así, teniendo en cuenta que una probabilidad es un límite de frecuencias relativas, debería tener las mismas propiedades que las frecuencias relativas:

- $h_\Omega = 1$
- $h_A \geq 0 \, \forall \, A \subset \Omega$
- $\forall \, A, \, B \subset \Omega, \, A \cap B = \emptyset \implies h_{A \cup B} = h_A + h_B$

Estas propiedades motivaron lo que se conoce como definición axiomática de probabilidad de Kolmogorov.

2.4.1 Espacio de Probabilidad

El modelo matemático correspondiente a un experimento aleatorio es la terna $(\Omega, \mathcal{A}, \mathcal{P})$, llamada **Espacio de Probabilidad**, donde:

- Ω es el **Espacio Muestral**, el conjunto de todos los resultados posibles del experimento.
- $\mathcal{A}$ es la σ**-Álgebra de sucesos**, es decir, una colección de subconjuntos de Ω verificando los axiomas siguientes:

  i) $\emptyset \in \mathcal{A}$

  ii) $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$

  iii) $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A} \Rightarrow \bigcup\limits_{n \in \mathbb{N}} A_n \in \mathcal{A}$

- $\mathcal{P}$ es una función de conjunto verificando los siguientes axiomas **(axiomas de Kolmogorov):**

  i) $\mathcal{P}(A) \geq 0 \quad \forall A \in \mathcal{A}$

  ii) $\mathcal{P}(\Omega) = 1$

  iii) $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A} \, / \, A_i \cap A_j = \emptyset \quad \forall i \neq j \Rightarrow \mathcal{P}(\bigcup\limits_n A_n) = \sum\limits_n \mathcal{P}(A_n)$

2.4.2 Propiedades de la probabilidad

Como consecuencia de los anteriores axiomas, se verifican las siguientes propiedades:

- $\mathcal{P}(\emptyset) = 0$ (la probabilidad del suceso imposible es cero).
- $\mathcal{P}(\overline{A}) = 1 - \mathcal{P}(A)$ (las probabilidades de un suceso y su contrario siempre suman 1).
- $\mathcal{P}(A) \leq 1$ para cualquier suceso A del espacio muestral.
- $\mathcal{P}(B \setminus A) = \mathcal{P}(B \cap \overline{A}) = \mathcal{P}(B) - \mathcal{P}(A \cap B)$.

  En efecto, $B = (B \setminus A) \cup (A \cap B)$ con $B \setminus A$ y $A \cap B$ disjuntos. Por tanto, por el axioma III:

  $$\mathcal{P}(B) = \mathcal{P}(B \setminus A) + \mathcal{P}(A \cap B) \implies \mathcal{P}(B \setminus A) = \mathcal{P}(B) - \mathcal{P}(A \cap B)$$

- $\mathcal{P}(A \mathring{\cup} B) = \mathcal{P}(A \setminus B) + \mathcal{P}(B \setminus A)$

  Basta observar que $A \mathring{\cup} B = (A \setminus B) \cup (B \setminus A)$, y que $(A \setminus B) \cap (B \setminus A) = \emptyset$

- $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$ para cualquier pareja de sucesos A y B.

  Razonando como en la propiedad anterior, $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$, siendo esos tres conjuntos disjuntos. Por tanto:

  $$\mathcal{P}(A \cup B) = \mathcal{P}(A \setminus B) + \mathcal{P}(B \setminus A) + \mathcal{P}(A \cap B) \implies \mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$$

  Esta propiedad es una generalización del axioma III, y además puede extenderse el resultado para más de dos sucesos. Por ejemplo:

  $$\mathcal{P}(A_1 \cup A_2 \cup A_3) = \mathcal{P}(A_1) + \mathcal{P}(A_2) + \mathcal{P}(A_3) - \mathcal{P}(A_1 \cap A_2) - \mathcal{P}(A_1 \cap A_3) - \mathcal{P}(A_2 \cap A_3) + \mathcal{P}(A_1 \cap A_2 \cap A_3)$$

  El resultado puede aplicarse para cualquier colección finita de sucesos $\{A_i\}_{i=1}^{n}$:

  $$\mathcal{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathcal{P}(A_i) - \sum_{i \neq j}^{n} \mathcal{P}(A_i \cap A_j) + \sum_{i \neq j \neq k}^{n} \mathcal{P}(A_i \cap A_j \cap A_k) - \ldots (-1)^{n-1} \mathcal{P}\left(\bigcap_{i=1}^{n} A_i\right)$$

- Si $A \subset B$, entonces $\mathcal{P}(A) \leq \mathcal{P}(B)$.

2.5 Probabilidad Condicionada. Teorema de la Probabilidad Compuesta.

Antes de la definición axiomática debida a Kolmogorov, se consideró la probabilidad de un suceso A como límite de la frecuencia relativa del suceso. (Si el experimento se realiza N veces y el suceso A se verifica N_A veces, la frecuencia relativa de A es el cociente $f_A = \frac{N_A}{N}$)

El problema que vamos a tratar en este tema, es el de formalizar la idea intuitiva de que la información aportada por el hecho de que haya ocurrido un suceso A, ha de ser recogida cambiando el espacio de probabilidad de partida. Veámoslo con un ejemplo.

Consideremos un experimento aleatorio consistente en lanzar un dado, y el espacio muestral $\Omega = \{1, 2, 3, 4, 5, 6\}$. Sea B el suceso "ser par", es decir $B = \{2, 4, 6\}$. Suponiendo que todos los casos son equiprobables, por la ley de Laplace $\mathcal{P}(B) = \frac{3}{6} = \frac{1}{2}$

Supongamos ahora que disponemos de la siguiente información: 'En el lanzamiento se obtuvo un número mayor que 3', es decir se verificó el suceso $A = \{4, 5, 6\}$. En esta situación, la probabilidad de que el número haya sido par es ahora $\frac{2}{3}$.

2.5.1 Probabilidad Condicionada

Definición

- Llamamos **probabilidad del suceso B condicionada por A**, y la denotamos por $\mathcal{P}(B/A)$, a la probabilidad de que ocurra B sabiendo que ha ocurrido A.

- Está definida siempre que $\mathcal{P}(A) > 0$ (es decir, si $\mathcal{P}(A) = 0$ no se define).

La noción de condicionalidad nace de la necesidad de querer utilizar la información que aporta saber que ha ocurrido A, para medir la probabilidad de la ocurrencia de B.

Así, supongamos que realizamos un experimento N veces. Si quisiéramos aproximar la probabilidad de un suceso B, usaríamos su frecuencia relativa $f_B = \frac{N_B}{N}$.

Ahora bien, si realizamos el experimento y tenemos información de que se verificó un suceso A, para calcular la probabilidad de ocurrencia de B usando dicha información, lo natural es recurrir al cociente $f_{B/A} = \frac{N_{A\cap B}}{N_A}$, que mide la frecuencia de realización de B con respecto la frecuencia de veces que se presenta el suceso A:

$$f_{B/A} = \frac{N_{A\cap B}}{N_A} = \frac{\frac{N_{A\cap B}}{N}}{\frac{N_A}{N}} = \frac{f_{A\cap B}}{f_A}$$

Es decir:

$$f_{B/A} = \frac{f_{A\cap B}}{f_A}$$

Teniendo en cuenta la idea de probabilidad como límite de frecuencias relativas, tomando límites en la igualdad anterior resulta que la probabilidad del suceso B condicionada por A se calcula mediante la expresión:

$$\mathcal{P}(B/A) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(A)} \quad \text{si} \quad \mathcal{P}(A) > 0$$

Este hecho motiva la siguiente definición formal.

Definición

- Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad, y $A \in \mathcal{A}$ un suceso con $p(A) > 0$, se llama **probabilidad condicionada por A** a la función de conjunto

$$\begin{array}{rccl} \mathcal{P}(\cdot/A): & \mathcal{A} & \longmapsto & [0,1] \\ & B & \rightarrow & \mathcal{P}(B/A) = \dfrac{\mathcal{P}(B \cap A)}{\mathcal{P}(A)} \end{array}$$

- Al valor $\mathcal{P}(B/A)$ se le llama **probabilidad de B condicionada por A**.

Veamos que $\mathcal{P}(\cdot/A)$ es en efecto una medida de probabilidad sobre $(\Omega, \mathcal{A})$ comprobando que verifica los axiomas de Kolmogorov:

i) $p(B/A) \geq 0 \quad \forall B \in \mathcal{A}$, pues $\mathcal{P}(A \cap B) \geq 0$ al ser $\mathcal{P}$ una probabilidad sobre $(\Omega, \mathcal{A})$

ii) $\mathcal{P}(\Omega/A) = 1$. En efecto

$$\mathcal{P}(\Omega/A) = \frac{\mathcal{P}(\Omega \cap A)}{\mathcal{P}(A)} \underbrace{=}_{\Omega \cap A = A} \frac{\mathcal{P}(A)}{\mathcal{P}(A)} = 1$$

iii) Sea $(B_n)_{n \in \mathbb{N}} \in \mathcal{A}$ con $B_i \cap B_j = \emptyset \quad \forall i \neq j$, entonces $\mathcal{P}(\bigcup\limits_n B_n/A) = \sum\limits_n \mathcal{P}(B_n/A)$, en efecto

$$\mathcal{P}(\bigcup_n B_n/A) = \frac{\mathcal{P}\big((\bigcup\limits_n (B_n \cap A)\big)}{\mathcal{P}(A)} \underbrace{=}_{*} \frac{\sum\limits_n \mathcal{P}(B_n \cap A)}{\mathcal{P}(A)} = \sum_n \mathcal{P}(B_n/A)$$

$$(*) \quad (B_i \cap A) \cap (B_j \cap A) = \emptyset \quad \forall i \neq j$$

Tenemos así un nuevo espacio de probabilidad $\big(\Omega, \mathcal{A}, \mathcal{P}(\cdot/A)\big)$ en el que se recoge la información proporcionada por la realización del suceso A.

Por ser $\mathcal{P}(\cdot/A)$ una medida de probabilidad sobre $(\Omega, \mathcal{A})$ verifica todas la propiedades de cualquier probabilidad. Citamos a continuación las más significativas sin proceder a su demostración.

- $\mathcal{P}(\emptyset/A) = 0$

- $\mathcal{P}(\overline{B}/A) = 1 - \mathcal{P}(B/A)$
- Si $B_1 \subset B_2$ entonces $\mathcal{P}(B_1/A) \leq \mathcal{P}(B_2/A)$
- $\mathcal{P}(B_1 \cup B_2/A) = \mathcal{P}(B_1/A) + \mathcal{P}(B_2/A) - \mathcal{P}(B_1 \cap B_2/A)$
- Si $B \subset A$ entonces $\mathcal{P}(B/A) = \frac{\mathcal{P}(B)}{\mathcal{P}(A)}$
- Si $A \subset B$ entonces $\mathcal{P}(B/A) = 1$
- Si $A \cap B = \emptyset$ entonces $\mathcal{P}(B/A) = 0$

La comprobación de éstas es inmediata utilizando la definición de $\mathcal{P}(\cdot/A)$

2.5.2 Teorema de la Probabilidad Compuesta

Utilizando asimismo la definición de probabilidad condicionada puede obtenerse una forma de hallar la probabilidad de la intersección de dos sucesos. Nos la propociona el siguiente teorema.

Teorema Sea $(\Omega, \mathcal{A}, \mathcal{P})$ un espacio de probabilidad, y $A \in \mathcal{A}$ un suceso con $\mathcal{P}(A) > 0$, entonces:

$$\forall B \in \mathcal{A} \quad \mathcal{P}(A \cap B) = \mathcal{P}(A) \cdot \mathcal{P}(B/A)$$

La demostración es inmediata aplicando la definición de probabilidad condicionada.
El anterior teorema admite la siguiente generalización.

Teorema Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ con $\mathcal{P}(\bigcap\limits_{i=1}^{n-1} A_i) > 0$, entonces

$$\mathcal{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathcal{P}(A_1) \cdot \mathcal{P}(A_2/A_1) \cdot \mathcal{P}(A_3/A_1 \cap A_2) \cdot \ldots \cdot \mathcal{P}\left(A_n/\bigcap_{i=1}^{n-1} A_i\right)$$

La demostración es muy sencilla, se obtiene por inducción en n

2.6 Independencia de sucesos

En general resulta bastante obvio que $\mathcal{P}(B/A) \neq \mathcal{P}(B)$. Es decir, saber que se verifica A modifica nuestras expectativas de que se verifique B.

Cuando se tiene que $\mathcal{P}(B/A) = \mathcal{P}(B)$ se interpreta que la probabilidad de que se presente B es independiente de la ocurrencia de A. En este caso, en el teorema de probabilidad compuesta se obtiene $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cdot \mathcal{P}(B)$, lo cual motiva la siguiente definición formal.

Definición

- Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad. Dos sucesos A y B son **independientes** si $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cdot \mathcal{P}(B)$.
- En caso de no darse la anterior igualdad A y B se dirán sucesos **dependientes**.

Es importante observar que en la definición de independencia (o dependencia) de sucesos no se imponen restricciones sobre los valores de $\mathcal{P}(A)$ o $\mathcal{P}(B)$ (mientras que sí se hacía en la definición de probabilidad condicionada)

Por otro lado hay que destacar el carácter recíproco del concepto de independencia.

2.6.1 Propiedades

Con respecto a la dependencia o independencia de sucesos se tienen las siguientes propiedades:

- Si $\mathcal{P}(A) > 0$ entonces A y B son independientes si y sólo si $\mathcal{P}(B/A) = \mathcal{P}(B)$:

  Por definición de probabilidad condicionada $\mathcal{P}(B/A) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(A)}$ (por hipótesis $\mathcal{P}(A) > 0$)

  Como A y B son independientes si y sólo si $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cdot \mathcal{P}(B)$ (por definición de independencia) sustituyendo se obtiene $\mathcal{P}(B/A) = \mathcal{P}(A)$

  (Análogamente si $\mathcal{P}(B) > 0$ la independencia de A y B equivale a $\mathcal{P}(A/B) = \mathcal{P}(A)$)

  Es importante destacar la hipótesis $\mathcal{P}(A) > 0$ pues si $\mathcal{P}(A) = 0$ no está definida la función $\mathcal{P}(\cdot/A)$. El caso $\mathcal{P}(A) = 0$ queda resuelto en la siguiente propiedad.

- Si $\mathcal{P}(A) = 0$ o $\mathcal{P}(A) = 1$ entonces A y B son sucesos independientes:

  Si $\mathcal{P}(A) = 0$, como $A \cap B \subset A$ entonces $\mathcal{P}(A \cap B) \leq \mathcal{P}(A)$, es decir $\mathcal{P}(A \cap B) = 0$. Así, trivialmente $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cdot \mathcal{P}(B)$.

  Si $\mathcal{P}(A) = 1$ entonces $\mathcal{P}(\overline{A}) = 0$. Como $B = (B \cap A) \cup (B \cap \overline{A})$ siendo la unión disjunta, tenemos $\mathcal{P}(B) = \mathcal{P}(B \cap A) + \mathcal{P}(B \cap \overline{A})$. Vimos que $\mathcal{P}(\overline{A}) = 0$ implica $\mathcal{P}(B \cap \overline{A}) = 0$ y así $\mathcal{P}(B \cap A) = \mathcal{P}(B) = \mathcal{P}(B) \cdot \mathcal{P}(A)$ ya que $\mathcal{P}(A) = 1$.

- Dados dos sucesos A y B, las siguientes afirmaciones son equivalentes:

  a) A y B son independientes

  b) A y $\overline{B}$ son independientes

  c) $\overline{A}$ y B son independientes

  d) $\overline{A}$ y $\overline{B}$ son independientes

  En efecto:

  $a) \Rightarrow b)$

- Si $\mathcal{P}(A) = 0$ ó $\mathcal{P}(A) = 1$ ya está demostrado en 2.6.1
- Si $\mathcal{P}(A) > 0$ $\mathcal{P}(\overline{B}/A) = 1 - \mathcal{P}(B/A) \underbrace{=}_{\mathcal{P}(B/A)=\mathcal{P}(B)} 1 - \mathcal{P}(B) = \mathcal{P}(\overline{B})$

$b) \Rightarrow c)$, $c) \Rightarrow d)$, $d) \Rightarrow a$ se hacen de la misma forma.

- Si los sucesos A y B son **incompatibles** ($A \cap B = \emptyset$) con $\mathcal{P}(A) > 0$ y $\mathcal{P}(B) > 0$ entonces son dependientes:

  Si $A \cap B = \emptyset$ entonces $\mathcal{P}(A \cap B) = 0$. Como por hipótesis $\mathcal{P}(A) \cdot \mathcal{P}(B) > 0$, entonces $\mathcal{P}(A \cap B) \neq \mathcal{P}(A) \cdot \mathcal{P}(B)$.

  La explicación intuitiva es muy clara, pues por la incompatilidad si ocurre A necesariamente no puede ocurrir B, luego el hecho de que haya ocurrido A proporciona la información de la no ocurrencia de B. Es decir, la ocurrencia de B depende de la *no* ocurrencia de A.

  Recíprocamente, si A y B son independientes con $\mathcal{P}(A) > 0$ y $\mathcal{P}(B) > 0$, según el razonamiento anterior no pueden ser incompatibles.

  Aunque suele asociarse incompatibilidad con independencia, la incompatibilidad implica una de las relaciones de dependencia más estrechas entre sucesos.

2.6.2 Sucesos mutuamente independientes

Definicion Sea $(\Omega, \mathcal{A}, \mathcal{P})$ un espacio de probabilidad, y $\{A_1, \ldots, A_n\}$ una colección de sucesos $\mathcal{A}$. Se dice que los **sucesos** de la colección son **mutuamente independientes** si para cualquier subconjunto $\{A_{i_1}, \ldots, A_{i_k}\}$ con $1 \leq k \leq n$, $i_l \neq i_m$ $\forall l, m \in \{1, \ldots, k\}$ se tiene

$$\mathcal{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \mathcal{P}(A_{i_1}) \cdot \ldots \cdot \mathcal{P}(A_{i_k})$$

2.7 Aplicaciones de la probabilidad compuesta: Experimentos Compuestos

Los resultados relativos al teorema de la probabilidad compuesta son frecuentemente aplicados cuando se trabaja con varios experimentos aleatorios sucesivos, pues en ocasiones las condiciones iniciales de cada uno de ellos pueden venir alteradas por los resultados del experimento anterior. Si ello es así, se dice que los experimentos son **dependientes**, y en caso contrario se dirán **independientes**.

Supongamos que realizamos dos experimentos aleatorios con espacios de probabilidad asociados $(\Omega_1, \mathcal{A}_1, \mathcal{P}_1)$ y $(\Omega_2, \mathcal{A}_2, \mathcal{P}_2)$ respectivamente.

- Si los experimentos aleatorios son independientes (por ejemplo, cuando son físicamente independientes) sobre $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \times \mathcal{A}_2)$ podemos considerar la probabilidad:

$$\mathcal{P}(A_1 \times A_2) = \mathcal{P}_1(A_1) \cdot \mathcal{P}_2(A_2)$$

- Si los experimentos aleatorios son dependientes sobre $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \times \mathcal{A}_2)$, se considerará la medida de probabilidad:

$$\mathcal{P}(A_1 \times A_2) = \mathcal{P}_1(A_1) \cdot \mathcal{P}_2(A_2/A_1)$$

El caso más sencillo de experimentos independientes es el dado por la repetición de un experimento dos veces en las mismas condiciones, por ejemplo, la extracción de dos bolas de una urna con reemplazamiento, lanzar dos veces un dado, ...

Un ejemplo típico de experimentos dependientes es el consistente en extraer bolas de una urna sin reemplazamiento.

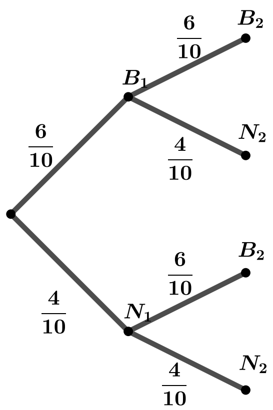
Para la resolución de problemas de experimentos compuestos es habitual el empleo de diagramas de árbol, como veremos en los dos siguientes ejemplos.

1. Se tiene una urna con seis bolas blancas y cuatro negras. Se extraen al azar dos bolas con reemplazamiento, y se desea construir el espacio de probabilidad asociado al experimento.

   Podemos considerar este experimento como la composición de dos experimentos aleatorios consistentes en extraer una bola de la urna.

   El espacio muestral es $\Omega = \{B_1B_2,\ B_1N_2,\ N_1B_2,\ N_1N_2\}$, y para asignar las probabilidades de cada uno de los sucesos elementales, consideramos el siguiente árbol de probabilidad.

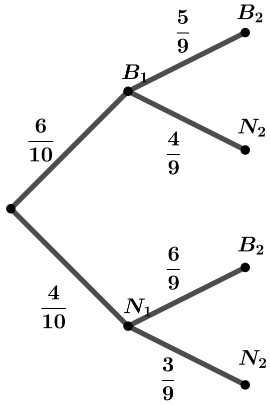
   Como vemos, al haber reemplazamiento, la composición de la urna no varía, y en el diagrama en árbol esto se traduce en que las probabilidades de cada una de las ramas no dependen del origen de la ramificación (ver figura 2.6).



$$\mathcal{P}(B_1 \cap B_2) = \mathcal{P}(B_1) \cdot \mathcal{P}(B_2) = \frac{36}{100}$$
$$\mathcal{P}(B_1 \cap B_2) = \mathcal{P}(B_1) \cdot \mathcal{P}(B_2) = \frac{24}{100}$$
$$\mathcal{P}(B_1 \cap N_2) = \mathcal{P}(B_1) \cdot \mathcal{P}(N_2) = \frac{24}{100}$$
$$\mathcal{P}(N_1 \cap N_2) = \mathcal{P}(N_1) \cdot \mathcal{P}(N_2) = \frac{16}{100}$$

Figura 2.6: Experimentos Independientes

2. Si realizamos ahora el experimento anterior, pero sin reemplazamiento de las bolas, la composición de la urna cambia en la segunda extracción, con lo cual, las probabilidades de las ramas correspondientes a la segunda extracción no permanecen invariantes (ver figura 2.7).



$$\mathcal{P}(B_1 \cap B_2) = \mathcal{P}(B_1) \cdot \mathcal{P}(B_2/B_1) = \frac{30}{90}$$
$$\mathcal{P}(B_1 \cap B_2) = \mathcal{P}(B_1) \cdot \mathcal{P}(B_2/B_1) = \frac{24}{90}$$
$$\mathcal{P}(B_1 \cap N_2) = \mathcal{P}(B_1) \cdot \mathcal{P}(N_2/B_1) = \frac{24}{90}$$
$$\mathcal{P}(N_1 \cap N_2) = \mathcal{P}(N_1) \cdot \mathcal{P}(N_2/N_1) = \frac{12}{90}$$

Figura 2.7: Experimentos Dependientes

2.8 Teorema de la Probabilidad Total. Teorema de Bayes.

Para terminar el tema vamos a dar dos teoremas muy importantes, el teorema de la probabilidad total y el teorema de Bayes.

2.8.1 Sistema Completo de Sucesos

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ una familia numerable de sucesos. Se dice que constituye un **sistema completo de sucesos** si $A_i \cap A_j = \emptyset \quad \forall i \neq j$ y $\bigcup\limits_{n\in\mathbb{N}} A_n = \Omega$.

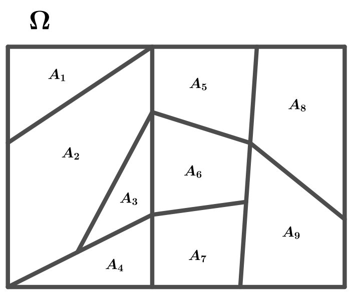


Figura 2.8: Sistema completo de sucesos

2.8.2 Teorema de la Probabilidad Total

Teorema Si tenemos un sistema completo de sucesos $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$, con $\mathcal{P}(A_n) > 0 \quad \forall n \in \mathbb{N}$, entonces:

$$\mathcal{P}(B) = \sum_{n=1}^{+\infty} \mathcal{P}(B/A_n) \cdot \mathcal{P}(A_n) \quad \forall B \in \mathcal{A}$$

Demostración:
Teniendo en cuenta las siguientes igualdades de conjuntos

$$B = B \cap \Omega = B \cap (\bigcup_{n\in\mathbb{N}} A_n) = \bigcup_{n\in\mathbb{N}} (B \cap A_n) \quad \text{con } (B \cap A_i) \cap (B \cap A_j) = \emptyset \quad \forall i \neq j$$

se tiene la que la probabilidad del suceso B es

$$\mathcal{P}(B) = \sum_{n\in\mathbb{N}} \mathcal{P}(B \cap A_n)$$

Y como además $\mathcal{P}(A_n) > 0 \quad \forall n \in \mathbb{N}$, y $\mathcal{P}(B \cap A_n) = \mathcal{P}(A_n) \cdot \mathcal{P}(B/A_n) \quad \forall n \in \mathbb{N}$, entonces:

$$\mathcal{P}(B) = \sum_{n\in\mathbb{N}} \mathcal{P}(A_n) \cdot \mathcal{P}(B/A_n)$$

2.8.3 Teorema de Bayes

Teorema Si tenemos un sistema completo de sucesos $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$, con $\mathcal{P}(A_n) > 0 \quad \forall n \in \mathbb{N}$, entonces:

$$\mathcal{P}(A_i/B) = \frac{\mathcal{P}(A_i) \cdot \mathcal{P}(B/A_i)}{\sum_{n=1}^{+\infty} \mathcal{P}(B/A_n) \cdot \mathcal{P}(A_n)} \quad \forall B \in \mathcal{A}/\mathcal{P}(B) > 0,\ \forall i \in \mathbb{N}$$

Demostración:
Si suponemos que $\mathcal{P}(B) > 0$, entonces $\forall i \in \mathbb{N}$ se tiene que

$$\mathcal{P}(A_i/B) = \frac{\mathcal{P}(A_i \cap B)}{\mathcal{P}(B)} = \frac{\mathcal{P}(B/A_i) \cdot \mathcal{P}(A_i)}{\mathcal{P}(B)}$$

(por definición de probabilidad condicionada, $\mathcal{P}(A_i \cap B) = \mathcal{P}(B/A_i) \cdot \mathcal{P}(A_i)$).
Sustituyendo el denominador por la fórmula de la Probabilidad Total tenemos

$$\mathcal{P}(A_i/B) = \frac{\mathcal{P}(A_i) \cdot \mathcal{P}(B/A_i)}{\sum_{n=1}^{+\infty} \mathcal{P}(B/A_n) \cdot \mathcal{P}(A_n)} \quad \forall B \in \mathcal{A} \ / \ \mathcal{P}(B) > 0,\ \forall i \in \mathbb{N}$$

Observación El teorema de Bayes admite la siguiente interpretación:

Supongamos estamos estudiando algún fenómeno aleatorio, y que los sucesos A_n son considerados como hipótesis sobre alguna cuestión relacionada. Si tenemos cierta información reflejada por las probabilidades $\mathcal{P}(A_n) > 0$, $n \in \mathbb{N}$, y se produce el suceso B, entonces el teorema de Bayes nos informa de la probabilidad de que un A_n concreto haya sido la causa de la ocurrencia de B.

Es decir, es un teorema que nos proporciona un método de medición de las probabilidades de las causas a partir de los efectos observados.

Por esta razón, a las probabilidades $\mathcal{P}(A_n)$ se las denomina **probabilidades a priori**, a las probabilidades $\mathcal{P}(A_n/B)$ **probabilidades a posteriori**, y a las $\mathcal{P}(B/A_n)$ **verosimilitudes**.

Una de las aplicaciones más frecuentes del teorema de Bayes que vamos a ver es el problema de los tests.

Ejemplo Supongamos una población de individuos en la que el 1% tiene una enfermedad. Empleamos un test que permite detectar la patología en una persona enferma 8 de cada 10 (test positivo), y detectar la no existencia de la enfermedad en un individuo sano 9 de cada 10 veces (test negativo) ¿Cuál será la probabilidad de que un individuo esté realmente enfermo si dio positivo en el test?

Sean los sucesos $A = \{\text{sujeto enfermo}\}$ y $B = \{\text{test positivo}\}$. Por el teorema de Bayes

$$\mathcal{P}(A/B) = \frac{\mathcal{P}(B/A)\mathcal{P}(A)}{\mathcal{P}(B/A)\mathcal{P}(A) + \mathcal{P}(B/\overline{A})\mathcal{P}(\overline{A})} = \frac{\frac{8}{10} \cdot \frac{1}{100}}{\frac{8}{10} \cdot \frac{1}{100} + \frac{1}{10} \cdot \frac{99}{100}} = \frac{8}{107}$$

Así, $\mathcal{P}(A/B) = 0.07476$ es la probabilidad de que un sujeto esté realmente enfermo cuando el test dio positivo.