Estadística Descriptiva y Probabilidad

(Teoría y problemas)

3ª Edición

Autores
I. Espejo Miranda
F. Fernández Palacín
M. A. López Sánchez
M. Muñoz Márquez
A. M. Rodríguez Chía
A. Sánchez Navas
C. Valero Franco

UCA Universidad de Cádiz
Servicio de Publicaciones

Edita: Servicio de Publicaciones de la Universidad de Cádiz
C/ Dr. Marañón, 3
11002 Cádiz

http://www.uca.es/publicaciones

ISBN: 978-84-9828-058-6

Depósito legal:

Parte B

Probabilidad

Estadística Descriptiva y Probabilidad. Teoría y Problemas (Revisión: Febrero 2006)
I. Espejo Miranda, F. Fernández Palacín, M. A. López Sánchez, M. Muñoz Márquez, A. M. Rodríguez Chía, A. Sánchez Navas, C Valero Franco

Introducción a teoría de la probabilidad

La segunda parte de este manual está dedicada a las herramientas por excelencia de la Estadística: la función de probabilidad y la variable aleatoria. Ello no quiere decir que en ocasiones no se planteen problemas específicos de probabilidad, pero lo cierto es que el enfoque con más proyección de la Estadística, el inferencial, no existiría sin dichas herramientas.

La existencia de fenómenos o experimentos no determinísticos, donde el conocimiento de las condiciones en las que éstos se desarrollan no garantizan los resultados, hace imprescindible el uso de una función que asigne niveles de certidumbre a cada uno de los desenlaces del fenómeno, y ahí es donde aparece la probabilidad. Los experimentos o fenómenos que poseen la característica anterior se denominan aleatorios. Intuitivamente, la concreción numérica del fenómeno mediante la asignación de valores con un cierto criterio, da origen a la variable aleatoria. Una correcta proyección de estos conceptos es lo que va a permitir estudiar grandes colectivos a partir de pequeñas partes de ellos, llamadas muestras, dando lugar a lo que se conoce como inferencia estadística.

Esta segunda parte está formada por otros tres capítulos, en el primero de ellos se introduce el concepto de probabilidad, haciéndose una breve incursión por la teoría de conjuntos, dada la existencia de una correspondencia total entre los elementos de esta teoría y los resultados de un experimento o fenómeno aleatorio. Se hace un recorrido por la evolución que a lo largo del tiempo ha tenido la probabilidad comentando sus distintas definiciones; se estudian sus propiedades y se introducen los importantes conceptos de probabilidad condicionada y de independencia. Termina el capítulo con los teoremas de la probabilidad total y de Bayes.

El segundo capítulo está dedicado a la variable aleatoria, que se presenta desde un punto de vista intuitivo y conceptual. El paso desde la función de cuantía en el caso discreto a la función de densidad en el continuo se hace de una forma gráfica y natural. La función de distribución se plantea como una alternativa a la función de densidad, al igual que las funciones generatriz de momentos y la característica, aunque indicando sus utilidades específicas. Desde una óptica más local, la esperanza matemática permitirá definir coeficientes que expresen las singularidades de la distribución, entre los que destacan la media y la varianza. Desde una perspectiva univariable se termina el tema con una alusión al cambio de variables y se da la expresión de la desigualdad de Tchebychev para el caso probabilístico. Para aquellos estudiantes que quieran ver como se generalizan algunos de estos conceptos, se estudia el caso multidimensional, con desarrollo expreso del bidimensional. Tanto para variables discretas como continuas, se dan las expresiones de las funciones de densidad y de distribución. Se analiza de nuevo la dependencia e independencia entre variables. La función esperanza toma aquí su versión n-dimensional y a partir de ella se pueden calcular una infinidad de coeficientes que aportan una visión puntual de la distribución, entre los que cabe destacar los de covarianza y correlación, proyecciones de los que se ven en descriptiva. Por último, se desarrolla el cambio de variables para el caso de dos dimensiones.

En el tercer capítulo se estudian distintas estructuras probabilísticas que modelizan una gran cantidad de situaciones reales, dedicándose especial atención a las distribuciones binomial y Poisson, en el caso discreto, y normal, en el continuo; de forma más somera también se analizan otras distribuciones que se derivan de aquellas. El teorema central del límite servirá para comprobar el papel que juega la distribución normal dentro de la Estadística. Al final del capítulo se estudian algunas distribuciones multivariables.

Estadística Descriptiva y Probabilidad. Teoría y Problemas (Revisión: Febrero 2006)
I. Espejo Miranda, F. Fernández Palacín, M. A. López Sánchez, M. Muñoz Márquez, A. M. Rodríguez Chía, A. Sánchez Navas, C Valero Franco

http://www.uca.es/teloydisren

Capítulo 4

Teoría de la probabilidad

1. Evolución histórica

Como en la mayoría de los descubrimientos, la noción de probabilidad se ha ido desarrollando a lo largo del tiempo en función de la necesidad, de los recursos y de la aportación de los grandes genios que son capaces, en un momento de inspiración, de dar un paso de un siglo.

Es difícil establecer históricamente el nacimiento de la probabilidad, aunque su conceptualización como disciplina matemática es reciente; parece, no obstante, que su origen tiene relación con los juegos de azar. Se consideran como remotos precursores de la teoría de la probabilidad la abundante presencia del hueso astrágalo de oveja o ciervo (antecedente inmediato del dado), en excavaciones arqueológicas con una antiguedad de más de 40.000 años. En épocas más recientes, en las culturas griega, egipcia y romana la afición a los juegos de azar, especialmente mediante la tirada de dados y tablas, estaba ampliamente extendida. En estas civilizaciones el azar se explicaba mediante la voluntad divina.

En el siglo XV, Dante obtiene algunas probabilidades en juegos sencillos de lanzamientos de dados. En el siglo XVI, Cardano con su tratado *Liber de ludo aleae* (Libro de los juegos de azar) y Galileo Galilei en su *Considerazione sopra el giuoco dei dadi* (Consideraciones sobre el juego de dados), describen ciertos juegos de dados y diversos problemas combinatorios, incluso este último, publica un tratado sobre la probabilidad de error.

La mayoría de los autores consideran que la génesis del cálculo de probabilidades, como disciplina matemática, se encuentra en la resolución del problema planteado por el caballero de Mére, un jugador empedernido de la Francia del siglo XVII, a su amigo y matemático B. Pascal (1623-1662), quien mantuvo una abundante correspondencia sobre dicho problema con su colega P. Fermat (1601-1665). El problema consistía en cómo deberían repartirse el dinero de las apuestas depositado en la mesa si los jugadores se vieran obligados a finalizar la partida sin que existiera un ganador. Dicho problema se detalla en el ejercicio 4.1

Ejercicio 4.1 *Dos jugadores de cartas, A y B, apuestan 250€ cada uno, en un juego que vencerá aquel que llegue primero a tres partidas ganadas. El juego se interrumpe cuando A lleva ganadas dos partidas y B una, ¿cómo deberían repartirse el dinero?*

Además de la correspondencia a la que se hacía referencia en el párrafo anterior, se considera fundamental en el nacimiento del cálculo de probabilidades la obra del matemático holandés C. Huygens (1629-1695), quien introduce el concepto de esperanza matemática, como generalización de la media aritmética, en su obra *De ratiocinüs in ludo aleae*, (Del raciocinio en los juegos de azar) en la que además resuelve varios problemas planteados por Pascal y Fermat.

En 1713 Jacques Bernouilli (1654-1705) lega uno de los tratados clave en la construcción de la teoría de la probabilidad, publicado tras su muerte lleva por título *Ars conjectandi* (El arte de conjurar), en él se introduce el término estocástico y se detalla la ley conocida como de ensayos de Bernouilli (primer teorema límite de la teoría demostrado con todo rigor), que enunciado de una forma sencilla dice así: "la frecuencia relativa de un suceso tiende a estabilizarse en torno a un número, a medida que el número de pruebas del experimento crece indefinidamente". En esta época hay que destacar las importantes contribuciones de autores como Abraham de Moivre (1667-1754), Daniel Bernouilli (1700-1782) o Thomas Bayes (1702-1761), entre otros.

A partir del citado periodo, el cálculo de probabilidades adquiere una mayor interrelación con otras ciencias y no se circunscribe exclusivamente a los juegos de azar.

Uno de los matemáticos artífices del asentamiento de las bases de lo que hoy se conoce como probabilidad clásica es Pierre Simon, marqués de Laplace (1749-1827). El momento cumbre fue la publicación en 1812, del importante y extenso tratado *Theorie analitique des probabilités*; en él, aparece la primera definición del concepto de probabilidad, conocida hoy como definición clásica.

Contemporáneo de Laplace, merece ser destacado C. F. Gauss, (1777-1855), quien dedicó parte de su actividad a estudiar la teoría de errores, dando lugar a la ley normal, de la que estimó sus parámetros.

Después de este periodo, el interés por el cálculo de probabilidades fue disminuyendo, llegando prácticamente a desaparecer como disciplina matemática durante el siglo XIX. Esto se debió, por una parte, a la aparición de algunas contradicciones como la que puso de manifiesto el matemático francés Bertrand, y por otra, a una relativa apatía por la probabilidad debido al carácter esencialmente "deterministíco" del siglo XIX. La evolución de la teoría se desplaza hacia el Este, en particular hacia la escuela de San Petesburgo. En ella resaltan las contribuciones de P.L. Tchebychev (1821-1894) y de su discípulo A. Markov (1856-1922).

Sin embargo, no es hasta principios del siglo XX, más concretamente 1933, fecha de la publicación de la obra *Fundamentos* por el matemático ruso A. N. Kolmogorov, cuando la probabilidad pasa a convertirse en una rama más de las matemáticas. En esta obra, Kolmogorov apoyándose en la teoría de conjuntos y en la teoría de la medida, da una definición axiomática del cálculo de probabilidades, como generalización y síntesis de los conocimientos que de la probabilidad se tenían hasta entonces.

2. Conjuntos. Operaciones

Un *conjunto* es una colección en un todo de objetos bien definidos que poseen una o varias propiedades. Cada uno de los objetos del conjunto se llama *elemento.*

La idea de conjunto sólo hace referencia a la presencia de sus elementos y no a ninguna ordenación o repetición de éstos. Los conjuntos pueden venir dados por:

Extensión: Se especifica cada uno de los elementos que pertenece al conjunto.

Descripción: Se da una o varias propiedades que deben cumplir todos los elementos del conjunto. Cualquier elemento que verifique esas propiedades pertenece al conjunto.

Ejemplo 4.1

$$A = \underbrace{\{x \; / \quad x \geq 0, \quad x^2 - x - 2 = 0\}}_{\text{Descripción}} \equiv \overbrace{\{2\}}^{\text{Extensión}}$$

Se dice que dos *conjuntos* son *iguales* cuando están formados por los mismos elementos.

Ejemplo 4.2 *Los conjuntos* $A = \{x \mid x^2 - x - 2 = 0\}$ *y* $B = \{-1, 2\}$ *son iguales.*

2.1. Subconjunto

B es un *subconjunto* de A, ó bien A contiene a B, si todo elemento de B es elemento de A.

Ejemplo 4.3 *Si* $A = \{x \mid x^2 - x - 2 = 0\}$ *y* $B = \{-1\}$ *entonces* $B \subset A$.

2.2. Operaciones entre conjuntos

Unión. La unión de conjuntos es otro conjunto que está formado por todos los elementos de dichos conjuntos.

Ejemplo 4.4 *Sea* $A = \{1, 2, 4\}$, $B = \{2, 3\}$ *entonces se tiene que* $A \cup B = \{1, 2, 3, 4\}$.

Intersección. La intersección de conjuntos es otro conjunto que está formado por los elementos comunes a todos los conjuntos.

Ejemplo 4.5 *Dados* $A = \{1, 2, 4\}$ *y* $B = \{2, 3\}$ *se tiene* $A \cap B = \{2\}$.

2.3. Conjuntos característicos

Conjunto vacío. Es aquel conjunto que no tiene ningún elemento. Se denota por $\emptyset$. Este conjunto se considera como subconjunto de cualquier otro conjunto.

Ejemplo 4.6 *Tomando* $A = \{1, 2, 4\}$ *y* $B = \{3, 5\}$ *queda* $A \cap B = \emptyset$.

Conjunto universal. Es aquel formado por la totalidad de los elementos del mismo tipo. Todo conjunto se puede considerar como subconjunto de él. Se denota por Ω.

Conjuntos disjuntos. Si dos conjuntos A y B no poseen elementos comunes se dicen que son disjuntos, es decir, $A \cap B = \emptyset$.

Partición. Se dice que los conjuntos $A_1, A_2, A_3, \ldots, A_n$ forman una partición o un sistema completo de sucesos, si son disjuntos dos a dos y la unión de todos ellos es el conjunto universal, es decir:

1. $A_i \cap A_j = \emptyset \quad \forall i \neq j$
2. $A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n = \Omega$

Conjunto complementario. El conjunto complementario de un conjunto A, es el conjunto formado por todos los elementos del conjunto universal que no están en A. Se denota por $\bar{A}$.

2.4. Propiedades de las operaciones entre conjuntos

1. La unión e intersección de conjuntos son operaciones conmutativas y asociativas.

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

   b) $(A \cup B) \cup C = A \cup (B \cup C), \quad (A \cap B) \cap C = A \cap (B \cap C)$

2. Se verifica la distributiva de cada operación respecto a la otra.

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

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

3. El operador complementario verifica:

   a) $A \cap \bar{A} = \emptyset$ y $A \cup \bar{A} = \Omega$

   b) $\bar{\Omega} = \emptyset$ y $\bar{\emptyset} = \Omega$

   c) $\overline{\overline{A}} = A$

4. Se cumplen las llamadas Leyes de Morgan:

   a) $\overline{A \cup B} = \overline{A} \cap \overline{B}$

   b) $\overline{A \cap B} = \overline{A} \cup \overline{B}$

3. Álgebra de sucesos

Hay determinados experimentos en los que las situaciones o causas determinan perfectamente los resultados o efectos, como por ejemplo ciertas situaciones físicas. Sin embargo, existen otros experimentos en los que en las mismas condiciones se obtienen resultados diferentes, como ocurre en los juegos de azar. Los primeros experimentos reciben el nombre de *determinísticos*, mientras que los segundos se conocen como *aleatorios*.

Se parte del experimento de lanzar un dado al aire y observar el número de puntos que figura en la cara superior. Los números 1, 2, 3, 4, 5, 6 son los sucesos elementales posibles. También se puede considerar sucesos compuestos como conseguir par, trío,..., formados por la unión de sucesos elementales. Si se prolonga indefinidamente esta sucesión de pruebas con sus resultados, se llega al conjunto potencialmente infinito de todas las pruebas asociadas a un experimento aleatorio que se llama universo, población o colectivo asociado al mismo.

Para el estudio de un fenómeno determinista se hace preciso la constatación de ciertas regularidades. En el caso de fenómenos aleatorios estas regularidades aparecen al considerar un gran número de pruebas.

Ejemplo 4.7 *Si obtengo m veces el valor 3 en n tiradas de un dado la frecuencia del suceso 3 será $\frac{m}{n}$.*

El hecho de que la frecuencia de un suceso tienda a aproximarse a un número fijo al aumentar el número de pruebas se ha denominado "ley de azar" o "ley de estabilidad" de las series estadísticas.

La noción de probabilidad como valor límite ideal de estas frecuencias es la base del modelo matemático apropiado para el estudio de estos fenómenos. Su teoría constituye el cálculo de probabilidades.

3.1. Espacio muestral. Sucesos

Se llama *espacio muestral* al conjunto de todos los resultados posibles de un experimento aleatorio. Se denota por Ω.

Los espacios muestrales pueden ser:

Finitos. Como, por ejemplo, el espacio muestral correspondiente al experimento aleatorio del lanzamiento de un dado una única vez, que consta de seis elementos que corresponden a los seis resultados posibles.

Infinitos numerables. Como, por ejemplo, el resultante de la contabilización del número de veces que hay que tirar una moneda hasta que aparezca por primera vez cara, donde el espacio muestral está compuesto por los números naturales, $\{1, 2, 3, 4, \ldots, n, \ldots\}$

Continuos. Como, por ejemplo, el que se obtiene al medir, en radianes, el ángulo que forma una aguja, que se lanza sobre una superficie plana, con una determinada dirección prefijada, donde el espacio muestral es $[0, 2\pi]$.

Se denomina *suceso* a todo subconjunto del espacio muestral, es decir A es un suceso si $A \subseteq \Omega$. Los *sucesos elementales* son aquellos que constan de un único elemento.

Al realizar un experimento aleatorio se dice que se ha verificado el suceso A, si el resultado obtenido pertenece a A.

3.2. Relaciones entre sucesos

Implicación. Un suceso A implica otro suceso B cuando siempre que se verifique A se verifica B.

Ejemplo 4.8 *En el lanzamiento de un dado se consideran los sucesos: A =\{Obtener un 4\} y B =\{Obtener un múltiplo de 2 \}, entonces A implica B.*

Suceso contrario. Dado un suceso A, se define el *suceso contrario* de A y se denota por $\bar{A}$, como aquel que se verifica si y sólo si no se verifica A.

Ejemplo 4.9 *Si en el lanzamiento de un dado $A = \{1, 2\}$, entonces el suceso contrario viene dado por $\bar{A} = \{3, 4, 5, 6\}$.*

Unión de sucesos. Se dice que el suceso $C = A \cup B$ se verifica, si y sólo si se verifica A , B o ambos.

Ejemplo 4.10 *En el lanzamiento de un dado se describen los siguientes sucesos:*

A = \{Obtener una puntuación menor o igual que tres \}

B = \{Obtener una puntuación par\}

Entonces $A \cup B$ = \{Obtener 1, 2, 3, 4, 6\}

Suceso seguro. El *suceso seguro*, que se denota por Ω, es aquel que siempre se verifica. Para cualquier suceso A siempre se cumple que $\Omega = A \cup \bar{A}$.

Intersección de sucesos. Se dice que el suceso $C = A \cap B$ se verifica, si y sólo si se verifican A y B.

Ejemplo 4.11 *En el lanzamiento de un dado, si se tiene:*
$A = \{$*Obtener una puntuación menor o igual que tres*$\}$
$B = \{$*Obtener una puntuación par*$\}$
Entonces $A \cap B = \{$*Obtener un 2*$\}$.

Suceso imposible. *Suceso imposible* es aquel que no se puede verificar nunca, se denota $\emptyset$.

Sucesos incompatibles. Dos *sucesos* son *incompatibles* cuando al verificarse uno de ellos no se verifica el otro, o equivalentemente, cuando su intersección es el suceso imposible.

3.3. Álgebra de Boole

Una familia $\mathcal{A}$ de subconjuntos de Ω tiene estructura de *Álgebra* o *Álgebra de Boole* sobre Ω, si verifica:

1. $\Omega \in \mathcal{A}$
2. $\forall A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$
3. $\forall A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$

El par $(\Omega, \mathcal{A})$ se dirá *espacio medible finito.*

Una familia $\mathcal{A}$ de subconjuntos de Ω se dice que tiene estructura de σ-*álgebra* sobre Ω, si

1. $\Omega \in \mathcal{A}$

2. $\forall \{A_i\}_{i\in I\!N} \Rightarrow \bigcup_{i\in I\!N} A_i \in \mathcal{A}$
3. $\forall A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$

El par $(\Omega, \mathcal{A})$ se dirá *espacio medible.*

Obsérvese que con esta definición la intersección de sucesos del álgebra y el conjunto $\emptyset$ pertenecen al álgebra.

Se establece una correspondencia biunívoca entre conjuntos y sucesos que viene dada por la tabla 4.1.

Cálculo de Probabilidades	Teoría de Conjuntos
Suceso Seguro (Espacio muestral)	Conjunto Universal
Suceso Elemental	Punto del Conjunto Universal
Suceso	Subconjunto
Sucesos Incompatibles	Conjuntos Disjuntos
Unión de Sucesos	Unión de Conjuntos
Suceso Imposible	Conjunto Vacío
Suceso Contrario	Conjunto Complementario
Intersección de Sucesos	Intersección de Conjuntos
Sistema Completo	Partición

Tabla 4.1: Cálculo de probabilidades y teoría de conjuntos

4. Distintas definiciones del concepto de probabilidad

Continuando con el estudio de un experimento aleatorio y una vez que se han definido los sucesos se aprecia la necesidad de definir alguna medida que cuantifique la incertidumbre o la asiduidad de que un determinado suceso se obtenga al realizar un experimento aleatorio, a tal medida se le denomina *probabilidad.*

La dificultad de dar una definición del concepto de probabilidad sin objeciones o limitaciones, queda reflejada por los diferentes intentos realizados a lo largo de la historia para encontrar una definición de dicho concepto.

De la introducción histórica que se ha elaborado se desprende la existencia de tres definiciones del concepto de probabilidad que a continuación se discuten.

Definición clásica, debida a Laplace: La probabilidad de un suceso A es el cociente entre el número de casos favorables al suceso y el número de casos posibles.

Los inconvenientes de definir la probabilidad de esta forma son:

- No es válida cuando los sucesos elementales no son equiprobables.
- A veces no es posible contar.

Definición frecuentista, debida a Bernouilli: La probabilidad de un suceso es el valor límite de su frecuencia relativa al repetir indefinidamente la experimentación.

Los inconvenientes de definir así la probabilidad son los siguientes:

- Desde el punto de vista del análisis no puede interpretarse el límite anterior por la imposibilidad de fijar el número de repeticiones.
- En algunas ocasiones no es posible realizar una experimentación indefinida.
- Las condiciones bajo las cuales se realiza la experimentación pueden variar a lo largo del tiempo y, con ellas, las frecuencias relativas.

Para evitar los inconvenientes de ambas definiciones, además de un gran número de paradojas y dificultades surgidas a comienzos del presente siglo, se hizo necesaria una profunda revisión del concepto de probabilidad utilizando las herramientas más precisas del momento: *La teoría de conjuntos*, desarrollada principalmente por Borel, y la potente *Teoría de la medida*, debida a Lebesgue. A la luz de estas teorías, la probabilidad empieza a entenderse como una medida de la incertidumbre, con propiedades similares a las medidas de longitud, tiempo, etc.

Una concepción más operativa es definir la probabilidad como una medida personal de la incertidumbre de un suceso, basada en aquellos experimentos previos, que con la información disponible, se consideren indistinguibles o intercambiables. En situaciones repetitivas, cuando exista una amplia experiencia, la probabilidad viene determinada por la frecuencia relativa, mientras que, en otros casos, depende de distintos tipos de información.

Todo lo anterior llevó a Kolmogorov a introducir axiomáticamente el concepto de Probabilidad.

Definición axiomática de probabilidad, debida a Kolmogorov: Dado $(\Omega, \mathcal{A})$ un espacio medible finito, una función sobre $\mathcal{A}$, $P : \mathcal{A} \longrightarrow \mathbb{R}$ se dice *medida de probabilidad finita* o simplemente *probabilidad*, si cumple los siguientes axiomas:

A1. La probabilidad de un suceso es siempre mayor o igual a cero,

$$P(A) \geq 0$$

A2. La probabilidad del suceso seguro es igual a la unidad,

$$P(\Omega) = 1$$

A3. La probabilidad de la unión de dos sucesos incompatibles es igual a la suma de las probabilidades de cada uno de ellos,

$$\forall A, B \in \mathcal{A} \mid A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$$

Definición axiomática de probabilidad: Dado $(\Omega, \mathcal{A})$ un espacio medible, una función sobre $\mathcal{A}$, $P : \mathcal{A} \longrightarrow \mathbb{R}$ se dice *medida de probabilidad infinita, de Kolmogorov* o simplemente *probabilidad*, si cumple los siguientes axiomas:

A1. La probabilidad de un suceso es siempre mayor o igual a cero,

$$P(A) \geq 0$$

A2. La probabilidad del suceso seguro es igual a la unidad,

$$P(\Omega) = 1$$

A3. La probabilidad de la unión de sucesos incompatibles es igual a la suma de las probabilidades de cada uno de ellos,

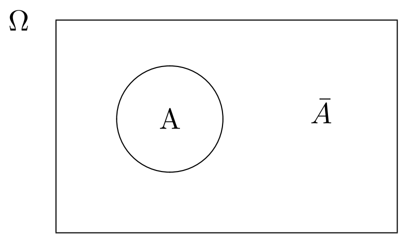
$$\forall \{A_i\}_{i\in I\!N} \mid A_i \cap A_j = \emptyset, \forall i \neq j \qquad P(\bigcup_{i\in I\!N} A_i) = \sum_{i\in I\!N} P(A_i)$$

El gran inconveniente de estas definiciones es que no dan un método para el cálculo de probabilidades, por lo que en la práctica hay que basarse en las definiciones clásica y frecuentista.

5. Propiedades de la función de probabilidad

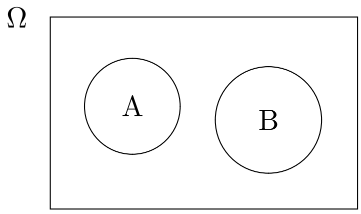
Como consecuencia de los axiomas se pueden deducir una serie de propiedades de la función de probabilidad, destacando las que se describen a continuación.

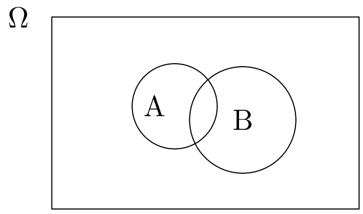
1. La probabilidad del suceso $\bar{A}$ es igual a uno menos la probabilidad del A,
$$P(\bar{A}) = 1 - P(A)$$



2. La probabilidad del suceso imposible es cero,
$$P(\emptyset) = 0$$

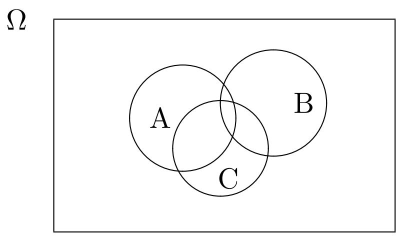
3. Si A y B son dos sucesos cualesquiera se verifica que:
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



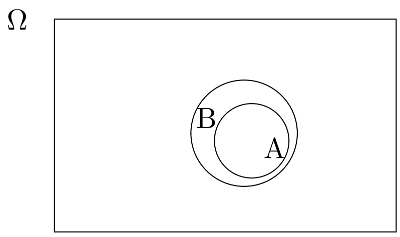


Si se tienen tres sucesos, A, B y C, la propiedad anterior tiene la forma siguiente:

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



4. Si $A \subseteq B$ entonces $P(A) \leq P(B)$



Ejercicio 4.2 *Demuestre las propiedades anteriores.*

6. Probabilidad condicionada. Independencia

La probabilidad de un determinado suceso en un experimento aleatorio puede verse modificada si se posee alguna información antes de la realización del experimento. Por ejemplo, si se consideran los alumnos de una clase de Estadística, la probabilidad de sacar aleatoriamente una alumna rubia será diferente de la de sacar una alumna rubia del grupo de las alumnas, es decir, si se parte del conocimiento de que el alumno escogido sea del sexo femenino. Para modelar este tipo de situaciones en las que se parte de una información a priori, se define el concepto de probabilidad condicionada.

Si $P(B) > 0$, la *probabilidad condicionada* de que se realice A si B se realiza, $P(A/B)$, viene definida por el cociente:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Análogamente, se define la probabilidad condicionada de B respecto a A. Utilizando conjuntamente ambos resultados se obtiene que:

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

Ejemplo 4.12 *El 60 % de los alumnos de una clase de Estadística son chicas y se sabe que el 30 % de las chicas son rubias. ¿Cuál es la probabilidad de escoger un alumno de la clase que sea chica y rubia?*

Para resolver este ejemplo se consideran los siguientes sucesos:

$$\begin{array}{rcl} F & = & \{\textit{ser de sexo femenino}\} \\ M & = & \{\textit{ser de sexo masculino}\} \\ R & = & \{\textit{tener pelo rubio}\}. \end{array}$$

La probabilidad pedida es:

$$P(R \cap F) \;=\; P(R/F)P(F) = 0'3 \cdot 0'6 = 0'18$$

Teorema 4.1 *Si $P(B) > 0$, entonces $P(\cdot/B)$ es una probabilidad:*

A1. $P(A/B) \geq 0 \quad \forall A$

A2. $P(\Omega/B) = 1$

A3. *Para cualquier sucesión de sucesos disjuntos* $\{A_i\}_{i=1}^{n/\infty}$*, se verifica:*

$$P(\bigcup_{i=1}^{n/\infty} A_i/B) = \sum_{i=1}^{n/\infty} P(A_i/B)$$

Teorema 4.2 *Dados* n *sucesos,* $A_1, A_2, \ldots, A_n$*, se tiene:*

$$\begin{aligned} & P(A_1 \cap A_2 \cap ... \cap A_n) = \\ & = \quad P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \cdot \ldots \\ & \quad \ldots \cdot P(A_n/A_1 \cap A_2 \cap ... \cap A_{n-1}) \end{aligned}$$

Ejemplo 4.13 *Siguiendo con el ejemplo 4.12 se sabe que el 40 % de las chicas rubias usan gafas. ¿Cuál es la probabilidad de que al escoger aleatoriamente un alumno de la clase resulte ser una chica rubia con gafas?*

Se define el suceso $G = \{\text{usar gafas}\}$ *y la probabilidad pedida es:*

$$\begin{aligned} P(F \cap R \cap G) &= P(G/(R \cap F))P(R/F)P(F) \\ &= 0'4 \cdot 0'3 \cdot 0'6 = 0'072. \end{aligned}$$

Ejercicio 4.3 *Demuestre los teoremas anteriores.*

7. Dependencia e independencia

En el epígrafe anterior se introdujo el concepto de probabilidad condicionada, debido a que la probabilidad de un determinado suceso se ve alterada por la información de que se dispone a priori. Sin embargo, puede suceder que dicha información no altere la probabilidad de ocurrencia de ese suceso, es decir, el que ocurra el suceso es independiente de la información.

Se dice que dos sucesos A y B son *independientes* si el conocimiento de la ocurrencia de uno no modifica la probabilidad de aparición del otro. O sea, $P(B/A) = P(B)$.

Ejercicio 4.4 *Compruebe que dos sucesos A y B son independientes si y sólo si $P(A \cap B) = P(A)P(B)$.*

La definición de independencia se puede generalizar a un número finito de sucesos. $A_1, A_2, \ldots, A_n$, se dicen mutuamente independientes si:

$$\begin{aligned}
&P(A_i \cap A_j) = P(A_i)P(A_j) \quad \forall i, \forall j \neq i \\
&P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k) \quad \forall i, \forall j \neq i, \forall k \neq i, j \\
&\qquad \vdots \\
&P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)
\end{aligned}$$

Ejemplo 4.14 *Continuando con el ejemplo 4.12, si se sabe que el 10 % de los alumnos de la clase escriben con la mano izquierda. ¿Cuál es la probabilidad de escoger aleatoriamente una chica que escriba con la mano izquierda?*

Sea el suceso D definido por D={escribir con la mano izquierda}. Admitiendo que los sucesos D y F son independientes, la probabilidad solicitada es:

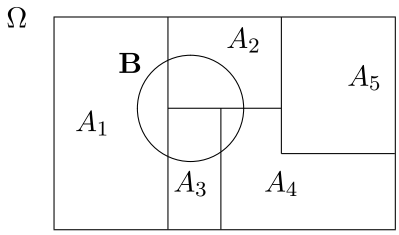
$$P(F \cap D) = P(F)P(D) = 0'6 \cdot 0'1 = 0'06.$$

8. Teorema de la probabilidad total. Teorema de Bayes

8.1. Teorema de la probabilidad total

Se considera un experimento que se realiza en dos etapas, en la primera se supone que los posibles sucesos, $A_1, A_2, \ldots, A_n$, constituyen un sistema completo, de tal forma que son conocidas las *probabilidades a priori*, $P(A_1), P(A_2), \ldots, P(A_n)$. Mientras que en la segunda etapa los resultados posibles, B_j, tienen probabilidades desconocidas que dependen de lo que ocurre en la primera etapa. Si se conocen las probabilidades condicionadas $P(B/A_i)$ para un cierto suceso B y cada A_i se verifica que:

$$P(B) = \sum_{i=1}^{n} P(B/A_i)P(A_i)$$



La demostración de la igualdad anterior se basa en que al ser $A_1, A_2, \ldots, A_n$, una partición de Ω y B un elemento cualquiera del álgebra de sucesos $\mathcal{A}$, se tienen las siguientes igualdades:

$$\begin{aligned} B &= B \cap \Omega \\ &= B \cap (A_1 \cup A_2 \cup .... \cup A_n) \end{aligned}$$

por las propiedades distributiva y conmutativa:

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup .... \cup (A_n \cap B)$$

Como los sucesos $\{A_i\}_{i=1}^n$ son incompatibles, también lo son los sucesos $\{A_i \cap B\}_{i=1}^n$, por lo tanto se puede aplicar el axioma A3:

$$\begin{aligned} P(B) &= P((A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_n \cap B)) \\ &= P(A_1 \cap B) + P(A_2 \cap B) + ... + P(A_n \cap B) \\ &= \sum_{i=1}^{n} P(A_i \cap B) = \sum_{i=1}^{n} P(B/A_i)P(A_i) \end{aligned}$$

Se obtiene de esta forma la igualdad buscada.

Ejemplo 4.15 *Continuando con el ejemplo 4.12; si se sabe que el 20 % de los chicos son rubios. ¿Cuál es la probabilidad de escoger aleatoriamente una persona rubia?*

La probabilidad solicitada es:

$$\begin{aligned} P(R) &= P(R/M)P(M) + P(R/F)P(F) \\ &= 0'2 \cdot 0'4 + 0'3 \cdot 0'6 = 0'26. \end{aligned}$$

8.2. Teorema de Bayes

En los epígrafes anteriores hemos dado la idea intuitiva de probabilidad condicionada como la probabilidad de que ocurra un suceso sabiendo que ha ocurrido con anterioridad otro determinado suceso. Sin embargo, también se puede plantear la probabilidad de que se haya dado un determinado suceso sabiendo que como resultado final del experimento se ha obtenido otro determinado suceso.

En las mismas hipótesis del teorema anterior se tiene que:

$$P(A_k/B) = \frac{P(B/A_k)P(A_k)}{\sum_{i=1}^{n} P(B/A_i)P(A_i)}, \qquad k = 1, \ldots, n$$

Observe que se intenta calcular una probabilidad "antinatura", pues se pretende expresar lo que ocurre antes, A_k, en función de lo que ocurre después, B. De todas formas, lo anterior tiene sentido porque en algunas ocasiones se conoce el resultado final de un experimento, pero se desconocen algunos de los pasos intermedios, en los que se está interesado. El teorema de Bayes resuelve esta cuestión, llevando el cálculo de las probabilidades a un terreno más natural, expresando las probabilidades a posteriori, $P(A_i/B)$, en función de las *verosimilitudes*, $P(B/A_i)$.

Aplicando la definición de probabilidad condicionada:

$$P(A_k/B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(A_k)P(B/A_k)}{P(B)}$$

Y por el teorema de la probabilidad total:

$$= \frac{P(A_k)P(B/A_k)}{\displaystyle\sum_{i=1}^{n} P(A_i)P(B/A_i)}$$

con lo que queda demostrado el teorema.

Ejemplo 4.16 *En el ejemplo 4.12, si se sabe que se ha elegido a una persona rubia, ¿cuál es la probabilidad de que sea chica?*

La probabilidad solicitada, a la vista del resultado del ejemplo 4.15 es

$$P(F/R) = \frac{P(R/F)P(F)}{P(R)} = \frac{0'3 \cdot 0'6}{0'26} = 0'69.$$

9. Ejercicios

9.1. Ejercicio resuelto

4.1 En una determinada ciudad se ha cometido un asesinato. De la investigación se encarga un detective, que tiene 5 sospechosos entre los que se encuentra el asesino. Se sabe que el detective trabaja con un pequeño margen de error, de forma que la probabilidad de creer inocente al verdadero asesino es de 0'05 y la probabilidad de creer culpable a una persona inocente es de 0'08. Si el detective cree que una persona es culpable, ¿cuál es la probabilidad de que esa persona sea el asesino?

Solución: Para la resolución del problema se definen los siguientes sucesos:

$$\begin{aligned} A &= \{\text{ser asesino}\} \\ I &= \{\text{ser enjuiciado inocente}\} \\ C &= \{\text{ser enjuiciado culpable}\} \end{aligned}$$

De esta forma se tiene que:

a) Hay un asesino de 5 sospechosos, por tanto, la probabilidad de que una persona elegida al azar sea el asesino es:

$$P(A) = \frac{1}{5} \qquad P(\bar{A}) = \frac{4}{5}$$

b) La probabilidad de creer inocente al verdadero asesino, es la probabilidad de ser enjuiciado inocente condicionada a que es el asesino, es decir, $P(I/A) = 0'05$. Además, se sabe que $P(C/A) = 1 - P(I/A) = 0'95$.

c) La probabilidad de creer culpable a una persona inocente, es la probabilidad de ser enjuiciado culpable condicionada a que no es el asesino, es decir, $P(C/\bar{A}) = 0'08$.

En el problema se pide la probabilidad de que una persona asesina haya sido enjuiciada culpable, es decir, la probabilidad de que una persona sea asesina condicionada a que ha sido enjuicida culpable. Por tanto, la probabilidad requerida es $P(A/C)$, para calcular dicha probabilidad se recurre al teorema de Bayes,

$$\begin{aligned} P(A/C) &= \frac{P(C/A)P(A)}{P(C/A)P(A) + P(C/\bar{A})P(\bar{A})} \\ &= \frac{0'95 \cdot 0'2}{0'95 \cdot 0'2 + 0'08 \cdot 0'8} \\ &= \frac{0'19}{0'254} = 0'748 \end{aligned}$$

9.2. Ejercicios propuestos

4.1. En una encuesta sobre las preferencias entre dos productos, realizada sobre un conjunto de 300 mujeres y 400 hombres, se han obtenido los siguientes resultados:

Producto	Hombres	Mujeres	Total
A	225	180	405
B	175	120	295

a) Represente la situación utilizando un diagrama de Venn.

b) Imagine que la encuesta ofrece información referida a dos conjuntos de edad, los menores y los mayores de 50 años. ¿Sería posible la representación incluyendo esta nueva información? De ser afirmativa la respuesta, represéntela.

4.2. Un estudiante de Estadística se dispone a realizar un estudio sobre el tipo y las condiciones de la comida que su madre le sirve a diario. Para ello establece las siguientes clasificaciones:

Estado de sal	Salada, normal, sosa
Temperatura	Caliente, fría
Tipo de alimento	Carne, pescado, verduras, pastas

Obtenga, utilizando un diagrama de árbol, el espacio muestral del tipo y las condiciones de las comidas.

4.3. Imagine que tenemos los sucesos A, B y C, exprese en lenguaje de la teoría de conjuntos las siguientes operaciones entre ellos:

a) Ocurren A y al menos uno de los otros dos.
b) Ocurren A y uno sólo de los otros dos.
c) Ocurre uno de los tres, pero no dos a la vez.
d) Ocurre, al menos, uno de los tres.
e) Ocurre C, pero no lo hacen ni A ni B.
f) Ocurren al menos dos de los tres.
g) Ocurren exactamente dos de los tres.
h) No ocurre ninguno de los tres.

4.4. Los alumnos de una determinada carrera se encuentran distribuidos en 5 cursos, de forma que en cada uno de los dos últimos cursos hay la mitad de alumnos que en cada uno de los tres primeros. Se pide que se calcule la probabilidad de que al escoger al azar a un alumno:

a) éste sea de cuarto.
b) le queden menos de tres cursos para acabar.

4.5. Represente el espacio muestral resultante al lanzar dos dados de distinto color y calcule la probabilidad de obtener una suma de siete puntos.

4.6. Represente el espacio muestral resultante al lanzar dos dados del mismo color y calcule la probabilidad de obtener una suma de siete puntos. Compare el resultado con el obtenido en el ejercicio anterior.

4.7. Un jugador lanza tres veces una moneda, si obtiene tres caras gana 100€ si obtiene una o dos caras gana 10€ y si no obtiene ninguna cara pierde 160€. ¿Es justo el juego?

4.8. Juan y Pedro juegan a una variante del juego de los chinos. Cada uno de ellos tiene tres chinos pudiendo seleccionar en una mano ninguno, uno, dos o los tres. A una señal los dos muestran los chinos seleccionados. Juan gana 10€ si sus chinos coinciden con los de Pedro o hay una diferencia de un único chino, mientras que Pedro gana 15€ en el resto de casos.

a) Calcule la probabilidad de que gane Juan.

b) ¿Qué cantidades deben ganar cada uno para que el juego sea justo?

4.9. Calcule la probabilidad de que tres alumnos seleccionados aleatoriamente en una clase cumplan años en meses consecutivos.

4.10. Calcule la probabilidad que tiene un ladrón que ha robado una tarjeta de un cajero automático de acertar con la clave, sabiendo que ésta tiene cuatro dígitos y que si no acierta en tres intentos el cajero se tragará la tarjeta.

4.11. Imagine que se encuentra un procedimiento que genera aleatoria e indefinidamente letras y signos de puntuación. ¿Cuál es la probabilidad de que un cierto momento escriba la novela *"El Quijote"*?

4.12. Se parte de que $P(A) = 0'3$, $P(B) = 0'4$ y $P(A \cap B) = 0'1$, obtenga:

a) $P(\bar{A} \cap \bar{B})$
b) $P(\bar{A} \cap B)$
c) $P(A - B)$
d) $P(A/B)$

4.13. Se considera el conjunto universal $\Omega = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1,\ 0 \leq y \leq 1\}$ y los sucesos $A_1 = \{(x, y) \in \Omega : 0 \leq x \leq \frac{3}{4}\}$, $A_2 = \{(x, y) \in \Omega : \frac{1}{2} \leq x \leq 1\,; \frac{1}{4} \leq y \leq \frac{3}{4}\}$ y $A_3 = \{(x, y) \in \Omega : 0 \leq x \leq \frac{1}{2}\}$.

a) Pruebe que la función $P(A) = \text{Área}(A)$, $\forall A \subseteq \Omega$ es una función de probabilidad.

b) Calcule $P(A_1 \cup A_2 \cup A_3)$, $P(\bar{A}_1 \cap \bar{A}_2)$ y $P(A_1 \cap (A_2 \cup A_3))$.

4.14. Sea una clase de estadística en la que un 20 % de los varones son rubios y un 50 % de las mujeres rubias. Si se sabe que el 30 % de la clase son varones, se pide:

a) La probabilidad de escoger aleatoriamente de la clase un varón rubio.

b) La probabilidad de escoger aleatoriamente una persona rubia de entre todos los alumnos.

c) La probabilidad de que una persona que se ha elegido aleatoriamente sea varón sabiendo que su pelo es rubio.

4.15. ¿Cuál es la probabilidad de que al tirar tres dados honrados salgan números diferentes?

4.16. Se tienen dos barajas de cartas de forma que la primera tiene 30 cartas rojas, 10 blancas y 2 negras y la segunda tiene 20 cartas rojas, 10 blancas y 12 negras. Se lanza una moneda, si sale cara se escogen tres cartas de la primera y una de la segunda; si sale cruz se escoge una carta de la primera y tres de la segunda. Calcule la probabilidad de que de las cuatro cartas extraídas dos sean blancas y las otras dos rojas.

4.17. Un estudio sobre los niveles de audiencia de diferentes cadenas de radio arrojó que el 50 % de la población escuchaba Radio A, el 40 % Radio B y el 30 % Radio C. Además, se obtuvo que el 20 % escuchaba Radio A y Radio B, el 10 % Radio A y Radio C y el 5 % Radio B y Radio C, finalmente sólo el 2 % escuchaba las tres cadenas.

a) ¿Qué porcentaje de la población escuchaba alguna cadena?

b) ¿Qué porcentaje de la población escuchaba una sola cadena?

4.18. En un programa de televisión existe una prueba que consiste en ordenar cronológicamente cinco inventos. El número de aciertos es el número de coincidencias entre las posiciones correctas y las ordenadas por el concursante. ¿Cuál es la probabilidad de que un concursante tenga al menos un acierto, sabiendo que realiza la ordenación de los inventos al azar?

4.19. Se consideran dos sucesos cualesquiera A y B, se pide:

a) Pruebe que si A y B son independientes entonces $\bar{A}$ y B, A y $\bar{B}$, y, $\bar{A}$ y $\bar{B}$ también lo son.

b) Demuestre que si $P(A) = 0$ entonces A y B son independientes.

c) Si A y B son dos sucesos disjuntos, ¿lo son también $\bar{A}$ y $\bar{B}$?

4.20. Se considera un equipo deportivo en octavos de final de una competición, que tiene una probabilidad de pasar a las siguientes fases de $\frac{4}{5}$, $\frac{3}{4}$ y $\frac{2}{3}$, respectivamente, y de $\frac{1}{2}$ de ganar la final si accede a ella, ¿cuál es la probabilidad de que gane la competición?

4.21. Un jugador de baloncesto tiene una probabilidad de encestar un lanzamiento desde una cierta posición de $\frac{1}{4}$.

a) ¿Cuál es la probabilidad de encestar tres lanzamientos consecutivos?

b) ¿Cuál es la probabilidad de que en cinco lanzamientos enceste al menos tres?

4.22. De una urna con tres bolas blancas y dos negras se extrae una bola y a continuación se lanza un dado, de forma que se introducen en la urna tantas bolas del mismo color que la extraída como el resultado obtenido al lanzar el dado. ¿Cuál es la probabilidad de que, una vez realizada esta operación, al extraer dos nuevas bolas, éstas tengan el mismo color?

4.23. Dos amigos son alumnos de la asignatura de Estadística de forma que cuando uno falta le pasa los apuntes al otro. Se sabe que el primero va a asistir a un 80 % de las clases y el segundo a un 40 %, de forma independiente. ¿Cuál es la probabilidad de que los amigos tengan todos los apuntes de clase?

4.24. Se considera una urna en la que hay 4 dados, de forma que en el primero 3 caras son unos y las restantes son doses, en el segundo 4 caras son unos y el resto doses, en el tercero 5 caras son unos y la otra un dos y en el cuarto 2 caras son unos y el resto doses.

a) ¿Cuál es la probabilidad de que al elegir un dado al azar y lanzarlo se obtenga un uno?

b) Se coge al azar un dado de la urna y al lanzarlo se obtiene un uno, ¿cuál es la probabilidad de que sea el cuarto dado?

4.25. De una urna que contiene cinco bolas blancas y tres negras, se extraen al azar cuatro bolas que se introducen en otra urna vacía, de esta urna se sacan aleatoriamente dos bolas que resultan ser una blanca y una negra. ¿Cuál es la probabilidad de que de las cuatro bolas pasadas, dos fueran blancas y las otras dos negras?

4.26. En una piscina de una piscifactoría se han introducido alevines de dos variedades de una especie en las siguientes cantidades y proporciones de machos y hembras:

Variedad	Cantidad	% machos
A	1000	7
B	1500	6

A continuación se escoge un alevín, ¿cuál es la probabilidad de que pertenezca a la variedad A, sabiendo que es hembra?

4.27. Una factoría produce un cierto artículo en tres cadenas de montaje. La cadena A fabrica el 50 % del total, la cadena B el 30 % y la C el 20 %, con porcentajes de defectuosos 0'03, 0'04 y 0'05 respectivamente. Un cliente decide analizar la calidad del producto para lo que selecciona una unidad al azar, ¿qué probabilidad hay de que dicha unidad resulte ser defectuosa?

4.28. Un niño guarda tres cajas con chocolatinas, en la primera tiene dos chocolatinas negras y una blanca, en la segunda dos negras y dos blancas y en la tercera dos blancas y una negra. En un despiste suyo, su hermana pequeña le ha cogido una chocolatina blanca, ¿cuál es la probabilidad de que la haya cogido de la primera caja?