TEORÍA DE LA PROBABILIDAD

Introducción

El Cálculo de Probabilidades tuvo su origen en el estudio de algunos problemas relacionados con juegos de dados. La dificultad y al mismo tiempo lo interesante de estos juegos consiste en que al lanzar uno o varios dados, de la manera usual en que esto se realiza, el resultado que se obtiene no está únicamente determinado, sino que es uno de un conjunto de posibles resultados. La teoría se fue desarrollando con la solución de problemas cada vez más complejos y de diversa índole, no únicamente en el ámbito de los juegos con dados, sino abarcando incluso fenómenos naturales.

En situaciones como la del lanzamiento de varios dados, donde existen diferentes posibles resultados, el resultado es únicamente uno de ellos, pero, a priori, no se sabe cuál; de ahí que podemos hacer aseveraciones acerca del resultado que obtendremos, las cuales tienen la característica de que pueden resultar ser verdaderas o falsas una vez que observamos el resultado obtenido. Por ejemplo, al lanzar tres dados, la aseveración "la suma de los números que muestran las caras superiores de los dados, una vez lanzados, es mayor que 10" puede resultar verdadera o falsa cuando se realiza el lanzamiento. A cada una de esas aseveraciones que se pueden hacer acerca del resultado que se obtiene se le llama un **evento**; cuando la aseveración resulta ser verdadera, decimos que el evento ocurre; cuando resulta falsa, decimos que no ocurre.

Considerando que un evento puede ocurrir o no ocurrir, lo que se busca es asignar un número real no negativo a cada evento, el cual indique que tanto se puede esperar que el evento ocurra. A ese número se le llama la probabilidad del evento. De esta forma, la probabilidad la podemos ver como una función no negativa definida sobre la familia de eventos, a la cual llamaremos **función de probabilidad**.

Para fines de aplicar el Cálculo de Probabilidades se introduce el concepto de **experimento aleatorio**, donde un experimento se considera como cualquier proceso que conduzca a un resultado. Algunos experimentos pueden consistir simplemente en la observación de un determinado fenómeno natural y en ir anotando algunas de las características que observamos; por ejemplo, la medición día con día del crecimiento de una planta. En otros experimentos están involucradas acciones nuestras, además de la observación, como en el caso del lanzamiento de un dado.

Un experimento consta de dos partes, por un lado tenemos la descripción del proceso que estamos considerando, especificando las condiciones bajo las cuales se realiza, y por otro lado tenemos lo que consideramos como su resultado.

> **Se dice que un experimento es aleatorio cuando al realizarse, bajo las condiciones que se indiquen, su resultado puede ser cualquiera de un conjunto de resultados posibles. Cabe aclarar que la posibilidad de distintos resultados puede provenir de que las condiciones que se especifican para la realización del experimento incluyen cierta arbitrariedad.**

En la formulación moderna, para modelar matematicamente un experimento aleatorio se busca construir una terna $(\Omega, \Im, P)$, llamada espacio de probabilidad, donde Ω, llamado el espacio muestral del experimento, es el conjunto de sus posibles resultados, $\Im$ es una σ-álgebra de subconjuntos de Ω, los cuales representan a los eventos, y P es una función definida sobre $\Im$, la cual tiene la característica de asignarle a Ω el valor 1. Cada elemento de $\Im$ es entonces un subconjunto del total de posibles resultados del experimento aleatorio en consideración y lo que nos interesa obtener es la probabilidad de que ocurra alguno de los posibles resultados que pertenecen a ese subconjunto. En general, la función P se construye mediante un proceso de extensión: se comienza por asignar probabilidades a algunos eventos y después, utilizando las propiedades de P, se busca extenderla hasta abarcar el mayor número posible de subconjuntos de Ω.

De esta forma, el modelo matemático para el estudio del experimento consta de 3 elementos, el primero es el conjunto Ω, el segundo una familia de subconjuntos de Ω y el tercero una función P, que llamaremos medida de probabilidad, la cual asocia a cada elemento A de $\Im$ un número real que designa la probabilidad de que ocurra alguno de los posibles resultados que pertenecen a A.

Este modelo se fue construyendo durante los primeros 33 años del siglo XX, buscando axiomatizar el Cálculo de Probabilidades. Durante esos 33 años se estableció y consolidó su vínculo con la Teoría de la Medida, la cual surgió en los primeros años del mismo siglo con los trabajos de Émile Borel y Henri Lebesgue. **Es con su formulación axiomática que podemos hablar de una teoría matemática de la probabilidad**, ya que de esta forma se tiene un cuerpo teórico, independiente de los problemas reales específicos que se tratan con el Cálculo de Probabilidades.

La identificación de una función de probabilidad con una medida no surgió de manera automática como algo general aplicable a cualquier fenómeno aleatorio, sino que requirió de un proceso, que llevó varios años, en el cual se mostró que es una alternativa adecuada. **En el centro de este proceso se encuentra el planteamiento de problemas donde se trata de calcular probabilidades de eventos cuya ocurrencia o no ocurrencia depende de una infinidad de observaciones y la aceptación de la σ-aditividad como una propiedad de cualquier función de probabilidad, la cual permite atacar ese tipo de problemas.**

Espacios de Probabilidad

La Teoría de la Probabilidad se utiliza para modelar y estudiar fenómenos aleatorios, los cuales tienen la característica de evolucionar de una manera azaroza o que, aún no siendo azarozo su desarrollo, su estudio puede realizarse pensándolo como si lo fuera. Sin embargo, una vez que se formula la Teoría de la Probabilidad en forma axiomática, los elementos que la componen no requieren de una interpretación práctica. La teoría matemática puede desarrollarse a partir de los axiomas, sin hacer referencia a algún fenómeno aleatorio. Se investiga acerca de las propiedades del modelo matemático y se van introduciendo definiciones de nuevos conceptos, dentro del modelo, enriqueciendo así el modelo mismo y demostrando nuevas propiedades. Nos referiremos a este proceso como el **desarrollo formal de la teoría**. Los nuevos conceptos que se van introduciendo provienen en general de problemas que se plantean en el estudio de algún fenómeno natural y las propiedades que se van encontrando del modelo se utilizan para estudiar el fenómeno en consideración; sin embargo el desarrollo formal de la teoría se va dando sin aludir a los problemas que motivan las definiciones de nuevos conceptos.

En lo que sigue vamos a exponer el desarrollo formal de la Teoría de la Probabilidad. Para ello requerimos de algunos conceptos que son propiamente de teoría de la medida, pero lo que ocurre es que una función de probabilidad es una medida, con la única particularidad de que asigna al conjunto total (llamado en Probabilidad .^es^pacio muestral") el valor 1.

La primera parte será teórica porque es necesario contar con algunos elementos antes de comenzar a resolver problemas.

Al final se encuentra una lista de ejemplos y ejercicios.

Álgebras y σ-álgebras

Definición 1. *Sea $\mathbb{F}$ un conjunto y $\mathcal{G}$ una familia de subconjuntos de $\mathbb{F}$. Diremos que:*

1. *$\mathcal{G}$ es cerrada bajo complementos si $A^c \in \mathcal{G}$ para cualquier $A \in \mathcal{G}$.*
2. *$\mathcal{G}$ es cerrada bajo diferencias propias si $B - A \in \mathcal{G}$ para cualquier pareja $A, B \in \mathcal{G}$ tal que $A \subset B$.*
3. *$\mathcal{G}$ es cerrada bajo uniones (resp. intersecciones) finitas si $\bigcup_{j=1}^{n} A_j \in \mathcal{G}$ (resp. $\bigcap_{j=1}^{n} A_j \in \mathcal{G}$) para cualquier colección finita $A_1, A_2, \ldots, A_n$ de elementos de $\mathcal{G}$.*
4. *$\mathcal{G}$ es cerrada bajo uniones (resp. intersecciones) infinitas numerables si $\bigcup_{j=1}^{\infty} A_j \in \mathcal{G}$ (resp. $\bigcap_{j=1}^{\infty} A_j \in \mathcal{G}$) para cualquier colección infinita numerable $A_1, A_2, A_3, \ldots$ de elementos de $\mathcal{G}$.*

5. *$\mathcal{G}$ es cerrada bajo uniones (resp. intersecciones) monótonas si $\bigcup_{j=1}^{\infty} A_j \in \mathcal{G}$ (resp. $\bigcap_{j=1}^{\infty} A_j \in \mathcal{G}$) para cualquier sucesión creciente (resp. decreciente) $(A_n)_{n\in\mathbb{N}}$, de elementos de $\mathcal{G}$.*

Definición 2 (álgebra). *Sea $\mathbb{F}$ un conjunto. Se dice que una familia $\mathcal{A}$ de subconjuntos de $\mathbb{F}$ es un álgebra si se satisfacen las siguientes condiciones:*

1. *$\mathbb{F} \in \mathcal{A}$.*
2. *$\mathcal{A}$ es cerrada bajo complementos.*
3. *$\mathcal{A}$ es cerrada bajo uniones finitas.*

Definición 3 ($\boldsymbol{\sigma}$-álgebra). *Sea $\mathbb{F}$ un conjunto. Se dice que una familia $\Im$ de subconjuntos de $\mathbb{F}$ es una σ-álgebra si se satisfacen las siguientes condiciones:*

1. *$\mathbb{F} \in \Im$.*
2. *$\Im$ es cerrada bajo complementos.*
3. *$\Im$ es cerrada bajo uniones infinitas numerables.*

Obsérvese que toda σ-álgebra es también un álgebra. En efecto, como $\mathbb{F} \in \Im$ y $\Im$ es cerrada bajo complementos, el conjunto vacío $\emptyset$ pertenece a $\Im$. Así que, si $A_1, A_2, \ldots, A_n$ son elementos de $\Im$, entonces, definiendo $A_j = \emptyset$ para cualquier $j > n$, se tiene:

$$\bigcup_{j=1}^{n} A_j = \bigcup_{j=1}^{\infty} A_j \in \Im$$

Aunque está dicho en las definiciones, enfatizo que un álgebra, o una σ-álgebra, es un conjunto cuyos elementos son conjuntos.

Definición 4. *Sea $\mathbb{F}$ un conjunto y $\mathcal{G}$ una familia de subconjuntos de $\mathbb{F}$. Llamaremos **álgebra generada por** $\mathcal{G}$ a la más pequeña familia de subconjuntos de $\mathbb{F}$ que forme un álgebra y que contenga a todos los elementos de $\mathcal{G}$.*

Ejemplo 1. *Tomemos como $\mathbb{F}$ al conjunto $\mathbb{N}$ de los números naturales.*

a) Definamos:

$A = \{1, 2, 3, 4\}$

$B = \{3, 4, 5, 6\}$

El álgebra $\mathcal{A}$ generada por los conjuntos A y B está dada por:

$\mathcal{A} = \{\emptyset, \{1, 2\}, \{3, 4\}, \{5, 6\}, \{7, 8, 9, 10, \ldots\}, \{1, 2, 3, 4\}, \{1, 2, 5, 6\}, \{1, 2, 7, 8, 9, 10, \ldots\},$

$\{3,4,5,6\}, \{3,4,7,8,9,10,\ldots\}, \{5,6,7,8,9,10,\ldots\}, \{1,2,3,4,5,6\},$

$\{1,2,3,4,7,8,9,10,\ldots\}, \{1,2,5,6,7,8,9,10,\ldots\}, \{3,4,5,6,7,8,9,10,\ldots\}, \mathbb{N}\}$

b) Sean A y B dos subconjuntos distintos de $\mathbb{N}$ y tales que $A \cap B \neq \emptyset$ y $A \cup B \neq \mathbb{N}$.

Con el objeto de tener conjuntos ajenos de tal manera que A_1, A_2 y $\mathbb{N}$ se puedan expresar como unión de algunos de ellos, partiendo de A_1 y A_2 vamos a definir 2^2 conjuntos ajenos, de la siguiente manera:

$D_1 = A \cap B$

$D_2 = A \cap B^c$

$D_3 = A^c \cap B$

$D_4 = A^c \cap B^c$

El álgebra $\mathcal{A}$ generada por los conjuntos A y B está dada por:

$\mathcal{A} = \{\emptyset, D_1, D_2, D_3, D_4, D_1 \cup D_2, D_1 \cup D_3, D_1 \cup D_4, D_2 \cup D_3, D_2 \cup D_4, D_3 \cup D_4,$

$D_1 \cup D_2 \cup D_3, D_1 \cup D_2 \cup D_4, D_1 \cup D_3 \cup D_4, D_2 \cup D_3 \cup D_4, \mathbb{N}\}$

Ejemplo 2. *Tomemos como $\mathbb{F}$ al conjunto $\mathbb{R}$ de los números reales y definamos:*

$A = [1,3]$

$B = [2,4]$

Siguiendo el ejemplo 1b, definamos:

$D_1 = A \cap B = [2,3]$

$D_2 = A \cap B^c = [1,2)$

$D_3 = A^c \cap B = (3,4]$

$D_4 = A^c \cap B^c = (-\infty,1) \cup (4,\infty)$

El álgebra $\mathcal{A}$ generada por los conjuntos A y B está dada por:

$\mathcal{A} = \{\emptyset, [2,3], [1,2), (3,4], (-\infty,1) \cup (4,\infty),$

$[1,3], [2,4], (-\infty,1) \cup [2,3] \cup (4,\infty), [1,2) \cup (3,4], (-\infty,2) \cup (4,\infty), (-\infty,1) \cup (3,\infty),$

$[1,4], (-\infty,3] \cup (4,\infty), (-\infty,1) \cup [2,\infty), (-\infty,2) \cup (3,\infty), \mathbb{R}\}$

Ejemplo 3. *Dados 3 subconjuntos, A, B y C, de un conjunto $\mathbb{F}$, podemos seguir el mismo procedimiento que seguimos para el caso de dos conjuntos. Considerando intersecciones, definimos 2^3 conjuntos:*

$D_1 = A \cap B \cap C$

$D_2 = A \cap B \cap C^c$

$D_3 = A \cap B^c \cap C$

$D_4 = A \cap B^c \cap C^c$

$D_5 = A^c \cap B \cap C$

$D_6 = A^c \cap B \cap C^c$

$D_7 = A^c \cap B^c \cap C$

$D_8 = A^c \cap B^c \cap C^c$

Tomando el conjunto vacío, $\mathbb{F}$, cada uno de los 8 conjuntos, las $\binom{8}{2}$ uniones por parejas, las $\binom{8}{3}$ uniones por ternas, las $\binom{8}{4}$ uniones tomadas de 4 en 4, etcétera, hasta llegar a las $\binom{8}{7}$ uniones tomadas de 7 en 7, obtenemos un total de:

$\binom{8}{0} + \binom{8}{1} + \binom{8}{2} + \binom{8}{3} + \binom{8}{4} + \binom{8}{5} + \binom{8}{6} + \binom{8}{7} + \binom{8}{8} = 2^8 = 256$ *conjuntos.*

Quitando los que se repitan, obtenemos el algebra generada por los conjuntos A, B y C.

Ejercicio 1. *Tomando como $\mathbb{F}$ al conjunto $\mathbb{N}$ de los números naturales, definamos A como el conjunto formado por todos los múltiplos de 2 y B como el conjunto formado por todos los múltiplos de 3. Encuentra los 16 conjuntos que forman el álgebra generada por A y B.*

Ejercicio 2. *Tomando como $\mathbb{F}$ al conjunto $\mathbb{R}$ de los números reales, definamos A como el conjunto formado por la unión de todos los intervalos cerrados cuyos extremos sean dos números pares consecutivos y que no se intersecten y B como el conjunto formado por la unión de todos los intervalos cerrados cuyos extremos sean dos números impares consecutivos y que no se intersecten; es decir:*

$A = [2, 4] \cup [6, 8] \cup [10, 12] \cup \cdots$

$B = [1, 3] \cup [5, 7] \cup [9, 11] \cup \cdots$

Encuentra los 16 conjuntos que forman el álgebra generada por A y B.

Ejercicio 3. *Sea $\mathcal{J}$ el conjunto por el vacío y todos los intervalos de la forma $(a, b]$, $(-\infty, c]$ o (d, ∞), donde $a, b, c, d \in \mathbb{R}$ y $a < b$.*

Demuestra que el conjunto $\mathcal{A}$ formado por todos los conjuntos de la forma $\bigcup_{k=1}^{n} J_k$ donde $n \in \mathbb{N}$ y $J_1, \ldots, J_n$ son intervalos en $\mathcal{J}$, ajenos por parejas, es un álgebra de subconjuntos de $\mathbb{R}$.

Definición 5 (Intersección de $\boldsymbol{\sigma}$-álgebras). *Dado un conjunto $\mathbb{F}$ y una familia arbitraria de σ-álgebras de subconjuntos de $\mathbb{F}$, se define la intersección de esas σ-álgebras como la familia de conjuntos que pertenecen a todas ellas.*

Se puede ver fácilmente que la intersección de σ-álgebras de subconjuntos de $\mathbb{F}$ es también una σ-álgebra y, dada una colección arbitraria $\mathcal{B}$ de subconjuntos de $\mathbb{F}$, siempre existe por lo menos una σ-álgebra que contiene a todos los elementos de $\mathcal{B}$, a saber, la formada por todos los subconjuntos de $\mathbb{F}$. Se puede definir entonces una σ-álgebra como la intersección de todas las σ-álgebras de subconjuntos de $\mathbb{F}$ que contienen a todos los elementos de $\mathcal{B}$.

Definición 6 ($\boldsymbol{\sigma}$ álgebra generada por una familia de conjuntos). *Dada una colección $\mathcal{A}$ de subconjuntos de un conjunto $\mathbb{F}$, se define la σ-álgebra generada por $\mathcal{A}$ como la intersección de todas las σ-álgebras que contienen a todos los conjuntos de $\mathcal{A}$. Denotaremos por $\sigma(\mathcal{A})$ a esta σ-álgebra.*

Evidentemente la σ-álgebra generada por $\mathcal{A}$ es la más pequeña σ-álgebra de subconjuntos de $\mathbb{F}$ que contiene a todos los elementos de $\mathcal{A}$.

σ-álgebra de Borel en $\mathbb{R}$

Una σ-álgebra de particular importancia es la de los conjuntos borelianos en $\mathbb{R}$ y, de manera más general, en $\mathbb{R}^n$.

Los conjuntos borelianos deben su nombre a Émile Borel quien los introdujo para caracterizar a los subconjuntos de $\mathbb{R}$ a los cuales se les puede asignar una longitud.

La idea fundamental consiste en que se puede asignar una longitud a todos los subconjuntos de $\mathbb{R}$ que se puedan obtener a partir de los intervalos mediante las operaciones conjuntistas de unión numerable y diferencia. La definición moderna se basa en la generación de σ-álgebras de acuerdo con la definición anterior.

Cuando hablemos de intervalos de números reales, vamos a incluir tanto a los finitos como a los infinitos. Explícitamente, el conjunto de los intervalos de números reales está formado por todos los de los tipos siguientes:

1. Los intervalos con extremos a y b, de cualquier tipo, donde $a, b \in \mathbb{R}$ y $a < b$.

2. Los intervalos de la forma $[a, a]$, $(-\infty, a)$, $(-\infty, a]$, (a, ∞) y $[a, \infty)$, donde a es un número real cualquiera.

3. El intervalo $(-\infty, \infty)$.

De acuerdo con lo anterior, no consideraremos como intervalo al conjunto vacío.

Cuando los dos extremos de un intervalo sean números reales, diremos que el intervalo es finito; en caso contrario, diremos que es infinito.

Definición 7 ($\boldsymbol{\sigma}$-álgebra de Borel en $\mathbb{R}$). *La σ-álgebra de Borel en $\mathbb{R}$, la cual será denotada por $\mathfrak{B}(\mathbb{R})$, es la σ-álgebra de subconjuntos de $\mathbb{R}$ generada por la familia de todos los intervalos de números reales. A los elementos de esa σ-álgebra los llamaremos borelianos de $\mathbb{R}$.*

Proposición 1. *La σ-álgebra de los conjuntos borelianos de $\mathbb{R}$ está generada por cualquiera de las siguientes familias de conjuntos:*

$\mathcal{D}_1$: Los intervalos de la forma $(-\infty, x]$, donde $x \in \mathbb{R}$.

$\mathcal{D}_2$: Los intervalos de la forma $(-\infty, x)$, donde $x \in \mathbb{R}$.

$\mathcal{D}_3$: Los intervalos de la forma $(a, b]$, donde $a, b \in \mathbb{R}$ y $a < b$.

$\mathcal{D}_4$: Los intervalos de la forma (a, b), donde $a, b \in \mathbb{R}$ y $a < b$.

$\mathcal{D}_5$: Los intervalos de la forma $[a, b)$, donde $a, b \in \mathbb{R}$ y $a < b$.

$\mathcal{D}_6$: Los intervalos de la forma $[a, b]$, donde $a, b \in \mathbb{R}$ y $a \leq b$.

Demostración

Vamos a demostrar primero que se tienen las siguientes contenciones:

$$\sigma(\mathcal{D}_1) \subset \sigma(\mathcal{D}_2) \subset \sigma(\mathcal{D}_3) \subset \sigma(\mathcal{D}_4) \subset \sigma(\mathcal{D}_5) \subset \sigma(\mathcal{D}_6) \subset \sigma(\mathcal{D}_1)$$

a) Para cualquier $x \in \mathbb{R}$, se tiene:

$(-\infty, x] = \cap_{n=1}^{\infty} \left(-\infty, x + \frac{1}{n}\right)$

$(-\infty, x) = \cup_{n=1}^{\infty} \left(-n, x - \frac{1}{n}\right]$

Así que:

$$\sigma(\mathcal{D}_1) \subset \sigma(\mathcal{D}_2) \subset \sigma(\mathcal{D}_3)$$

b) Para cualquier pareja $a, b \in \mathbb{R}$ tal que $a < b$, se tiene:

$(a,b]=\cap_{n=1}^{\infty}\left(a,b+\frac{1}{n}\right)$

$(a,b)=\cup_{n=1}^{\infty}\left[a+\frac{1}{n},b\right)$

$[a,b)=\cup_{n=1}^{\infty}\left[a,b-\frac{1}{n}\right]$

Así que:

$$\sigma\left(\mathcal{D}_3\right)\subset\sigma\left(\mathcal{D}_4\right)\subset\sigma\left(\mathcal{D}_5\right)\subset\sigma\left(\mathcal{D}_6\right)$$

c) Para cualquier pareja $a,b\in\mathbb{R}$ tal que $a\leq b$, se tiene:

$[a,b]=(-\infty,b]-(-\infty,a)=(-\infty,b]\cup_{n=1}^{\infty}(-\infty,a-\frac{1}{n}]$

Así que:

$$\sigma\left(\mathcal{D}_6\right)\subset\sigma\left(\mathcal{D}_1\right)$$

Por lo tanto:

$$\sigma\left(\mathcal{D}_1\right)=\sigma\left(\mathcal{D}_2\right)=\sigma\left(\mathcal{D}_3\right)=\sigma\left(\mathcal{D}_4\right)=\sigma\left(\mathcal{D}_5\right)=\sigma\left(\mathcal{D}_6\right)$$

Es decir, $\mathcal{D}_1$, $\mathcal{D}_2$, $\mathcal{D}_3$, $\mathcal{D}_4$, $\mathcal{D}_5$ y $\mathcal{D}_6$ generan la misma $\sigma-$álgebra, la cual denotaremos por $\Im$.

d) Si $a,b\in\mathbb{R}$ y $a<b$, denotemos por I a un intervalo con extremos a y b; entonces:

$$I\in\mathcal{D}_3\cup\mathcal{D}_4\cup\mathcal{D}_5\cup\mathcal{D}_6\subset\Im$$

e) Si a es cualquier número real, entonces:

$[a,a]\in\mathcal{D}_6\subset\Im$

$(-\infty,a)=\cup_{n=1}^{\infty}(-\infty,a-\frac{1}{n}]\in\sigma\left(\mathcal{D}_1\right)=\Im$

$(-\infty,a]\in\mathcal{D}_1\subset\Im$

$(a,\infty)=\mathbb{R}-(-\infty,a]\in\sigma\left(\mathcal{D}_1\right)=\Im$

$[a,\infty)=\mathbb{R}-(-\infty,a)=\mathbb{R}-\cup_{n=1}^{\infty}(-\infty,a-\frac{1}{n}]\in\sigma\left(\mathcal{D}_1\right)=\Im$

Además:

$(-\infty,\infty)=\mathbb{R}\in\Im$

Así que cualquier intervalo pertenece a $\Im$. Por lo tanto, $\mathfrak{B}\left(\mathbb{R}\right)\subset\Im$.

f) Finalmente:

$$\mathcal{D}_1\cup\mathcal{D}_2\cup\mathcal{D}_3\cup\mathcal{D}_4\cup\mathcal{D}_5\cup\mathcal{D}_6\subset\mathfrak{B}\left(\mathbb{R}\right)$$

Por lo tanto, $\Im \subset \mathfrak{B}(\mathbb{R})$.

Podemos concluir entonces que $\Im = \mathfrak{B}(\mathbb{R})$.

■

Funciones finitamente aditivas y σ-aditivas

Vamos a trabajar con el conjunto de números reales extendidos, el cual consiste del conjunto de números reales y dos elementos especiales, $-\infty$ y ∞, con los cuales operaremos bajo las siguientes convenciones:

Si $c \in \mathbb{R}$, entonces:

$-\infty < c < \infty$,

$c - \infty = -\infty$,

$c + \infty = \infty$,

$c(\infty) = \infty$ si $c > 0$,

$c(\infty) = -\infty$ si $c < 0$,

$(0)(\infty) = (0)(-\infty) = 0$,

$\frac{c}{\infty} = \frac{c}{-\infty} = 0$,

$(\infty)(\infty) = \infty + \infty = \infty$,

$\infty - \infty$ e $\frac{\infty}{\infty}$ no están definidos.

$\overline{\mathbb{R}}$ denotará al conjunto $\mathbb{R} \cup \{-\infty, \infty\}$.

El conjunto de números reales $\mathbb{R}$ también será denotado por $(-\infty, \infty)$. El conjunto de números reales no negativos será denotado por $[0, \infty)$, o por $\mathbb{R}^+$, y $\overline{\mathbb{R}}^+$ denotará al conjunto $\mathbb{R}^+ \cup \{\infty\}$.

Ejercicio 4. *Demuestra que la familia de conjuntos de la forma $B \cup \{\infty\}$, $B \cup \{-\infty\}$ y $B \cup \{-\infty, \infty\}$, donde B es un boreliano de $\mathbb{R}$, forma una $\sigma-$álgebra de subconjuntos de $\overline{\mathbb{R}}$.*

Definición 8. *Llamaremos σ-álgebra de Borel en $\overline{\mathbb{R}}$ y la denotaremos por $\mathfrak{B}\left(\overline{\mathbb{R}}\right)$ a la $\sigma-$álgebra de subconjuntos de $\overline{\mathbb{R}}$ generada por la familia de conjuntos de la forma $B \cup \{\infty\}$, $B \cup \{-\infty\}$ y $B \cup \{-\infty, \infty\}$, donde B es un boreliano de $\mathbb{R}$.*

Ejercicio 5. *Demuestra que la σ-álgebra $\mathfrak{B}\left(\overline{\mathbb{R}}\right)$ de los conjuntos borelianos de $\overline{\mathbb{R}}$ está generada por cualquiera de las siguientes familias de conjuntos.*

1. Los intervalos de la forma $[-\infty, x]$, donde $x \in \mathbb{R}$.
2. Los intervalos de la forma $[-\infty, x]$, donde $x \in \overline{\mathbb{R}}$.
3. Los intervalos de la forma $[-\infty, x)$, donde $x \in \mathbb{R}$.
4. Los intervalos de la forma $[-\infty, x)$, donde $x \in \overline{\mathbb{R}}$.

Definición 9 (Función finitamente aditiva sobre un álgebra). *Sea $\mathbb{F}$ un conjunto y $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$. Se dice que una función no negativa $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}$ es finitamente aditiva si dada cualquier familia finita, $A_1, \ldots, A_n$, de elementos de $\mathcal{A}$ tal que $A_i \cap A_j = \emptyset$ para $i \neq j$, entonces:*

$$\mu(\bigcup_{k=1}^{n} A_k) = \textstyle\sum_{k=1}^{n} \mu(A_k)$$

Obsérvese que, si $\mathcal{A}$ es un álgebra de subconjuntos de $\mathbb{F}$, para probar que una función $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}$ es finitamente aditiva, basta con demostrar que si A y B son dos conjuntos ajenos del álgebra, entonces $\mu(A \cup B) = \mu(A) + \mu(B)$. Teniendo esta propiedad, la aditividad finita se prueba con un razonamiento de inducción.

Definición 10 (Función σ-aditiva sobre un álgebra). *Sea $\mathbb{F}$ un conjunto y $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$. Se dice que una función no negativa $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}$ es σ-aditiva si es finitamente aditiva y dada cualquier familia infinita numerable, $A_1, A_2, \ldots$, de elementos de $\mathcal{A}$ tal que $A_i \cap A_j = \emptyset$ para $i \neq j$ y $\bigcup\limits_{k=1}^{\infty} A_k \in \mathcal{A}$, entonces:*

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \textstyle\sum_{k=1}^{\infty} \mu(A_k)$$

Definición 11 (Función σ-aditiva sobre una σ-álgebra). *Sea $\mathbb{F}$ un conjunto y $\Im$ una σ-álgebra de subconjuntos de $\mathbb{F}$. Se dice que una función no negativa $\mu : \Im \mapsto \overline{\mathbb{R}}$ es σ-aditiva si es finitamente aditiva y dada cualquier familia infinita numerable, $A_1, A_2, \ldots$, de elementos de $\Im$ tal que $A_i \cap A_j = \emptyset$, entonces:*

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \textstyle\sum_{k=1}^{\infty} \mu(A_k)$$

Proposición 2. *Sea $\mathbb{F}$ un conjunto, $\mathcal{A}$ un álgebra de subconjuntos de $\mathbb{F}$ y $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ una función no negativa finitamente aditiva, entonces:*

1. *Si $A, B \in \mathcal{A}$ y $A \subset B$, entonces $\mu(A) \leq \mu(B)$ y, si $\mu(A) < \infty$, $\mu(B - A) = \mu(B) - \mu(A)$.*

2. *Si* $A_1, \ldots, A_n \in \mathcal{A}$, *entonces* $\mu\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} \mu(A_k)$.
3. *Para cualquier pareja* $A, B \in \mathcal{A}$ *tal que* $\mu(A \cap B) < \infty$, *se tiene:*

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

Demostración

1. Sean $A, B \in \mathcal{A}$ tales que $A \subset B$, entonces $B = A \cup (B - A)$, así que $\mu(B) = \mu(A) + \mu(B - A)$. Por lo tanto, $\mu(B - A) = \mu(B) - \mu(A)$ y $\mu(A) \leq \mu(B)$, ya que μ es no negativa.

2. Sean ahora $A_1, \ldots, A_n \in \mathcal{A}$ y definamos $A_0 = \emptyset$, entonces:

$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \mu\left(\bigcup_{k=1}^{n}\left(A_k - \bigcup_{j=0}^{k-1} A_j\right)\right)$

$= \sum_{k=1}^{n} \mu\left(A_k - \bigcup_{j=0}^{k-1} A_j\right) \leq \sum_{k=1}^{n} \mu(A_k)$

3. $A \cup B = A \cup (B - A \cap B)$; así que:

$$\mu(A \cup B) = \mu(A) + [\mu(B) - \mu(A \cap B)] = \mu(A) + \mu(B) - \mu(A \cap B)$$

■

Definición 12. *Sea* $\mathbb{F}$ *un conjunto y* $\mathcal{A}$ *un álgebra de subconjuntos de* $\mathbb{F}$. *Se dice que una función no negativa* $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}$ *es* σ*-subaditiva, o que satisface la propiedad de la subaditividad numerable, si dada cualquier colección infinita* $A_1, A_2, \ldots$ *de elementos de* $\mathcal{A}$ *tales que* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, *entonces:*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

Teorema 1. *Sea* $\mathbb{F}$ *un conjunto,* $\mathcal{A}$ *un álgebra de subconjuntos de* $\mathbb{F}$ *y* $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ *una función no negativa y finitamente aditiva, entonces las siguientes propiedades son equivalentes:*

i. μ *es* σ*-subaditiva.*

ii. μ *es* σ*-aditiva.*

iii. Para cualquier sucesión creciente $(A_n)_{n\in\mathbb{N}}$, *de elementos de* $\mathcal{A}$ *tales que* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, *se tiene* $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\rightsquigarrow\infty} \mu(A_n)$.

iv. Para cualquier sucesión decreciente $(A_n)_{n\in\mathbb{N}}$, *de elementos de* $\mathcal{A}$ *tales que* $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$ *y* $\mu(A_n) < \infty$ *para alguna* $n \in \mathbb{N}$, *se tiene* $\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\rightsquigarrow\infty} \mu(A_n)$.

v. Para cualquier sucesión decreciente $(A_n)_{n\in\mathbb{N}}$, *de elementos de* $\mathcal{A}$ *tales que* $\bigcap_{n=1}^{\infty} A_n = \emptyset$ *y* $\mu(A_n) < \infty$ *para alguna* $n \in \mathbb{N}$,, *se tiene* $\lim_{n\to\infty} \mu(A_n) = 0$.

Demostración

$i \iff ii$

Supongamos que μ es σ-subaditiva y sea $A_1, A_2, \ldots$ una colección numerable de elementos de $\mathcal{A}$, ajenos por parejas y tales que $\bigcup_j A_j \in \mathcal{A}$. Se tiene:

$\mu(\bigcup_j A_j) \geq \sum_{k=1}^{n} \mu(A_j)$ para cualquier $n \in \mathbb{N}$, así que $\mu(\bigcup_j A_j) \geq \sum_{k=1}^{\infty} \mu(A_j)$.

Por lo tanto, μ es σ-aditiva.

Supongamos ahora que μ es σ-aditiva y sea $A_1, A_2, \ldots$ una colección numerable de elementos de $\mathcal{A}$ tales que $\bigcup_j A_j \in \mathcal{A}$. Sea $B_n = A_n - \bigcup_{j=1}^{n-1} A_j$, entonces $\bigcup_j A_j = \bigcup_j B_j$, así que:

$$\mu(\bigcup_j A_j) = \mu(\bigcup_j B_j) = \sum_{k=1}^{\infty} \mu(B_j) \leq \sum_{k=1}^{\infty} \mu(A_j)$$

$ii \Longrightarrow iii$

Supongamos que μ es σ-aditiva y sea $A_1, A_2, \ldots$ una colección de elementos de $\mathcal{A}$ tales que $A_1 \subset A_2 \subset \ldots$ y $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Se tiene:

$$\bigcup_{n=1}^{\infty} A_n = A_1 \cup (A_2 - A_1) \cup (A_3 - A_2) \cup \cdots$$

Así que:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{k=1}^{\infty} \mu(A_k - A_{k-1}) = \lim_{n \rightsquigarrow \infty} \sum_{k=1}^{n} \mu(A_k - A_{k-1}) = \lim_{n \rightsquigarrow \infty} \mu(A_n),$$

donde $A_0 = \emptyset$.

$iii \Longrightarrow i$

Sea $A_1, A_2, \ldots$ una colección numerable de elementos de $\mathcal{A}$ tales que $\bigcup_j A_j \in \mathcal{A}$. Para cada $n \in \mathbb{N}$, definamos $B_n = \bigcup_{j=1}^{n} A_j$, entonces $B_1 \subset B_2 \subset \ldots$ y $\bigcup_j A_j = \bigcup_j B_j$, así que:

$$\mu(\bigcup_j A_j) = \mu(\bigcup_j B_j) = \lim_{n \to \infty} \mu(B_n) \leq \lim_{n \rightsquigarrow \infty} \sum_{j=1}^{n} \mu(A_j) = \sum_{k=1}^{\infty} \mu(A_j)$$

$iii \Longrightarrow iv$ es inmediato tomando complementos.

$iv \Longrightarrow v$ es inmediato pues v es un caso particular de iv.

$v \Longrightarrow ii$

Sea $A_1, A_2, \ldots$ una colección numerable de elementos de $\mathcal{A}$, ajenos por parejas y tales que $\bigcup_j A_j \in \mathcal{A}$. Sea $B_n = \bigcup_{j=1}^{\infty} A_j - \bigcup_{j=1}^{n} A_j$, entonces $B_1 \supset B_2 \supset \ldots$ y $\bigcap_{n=1}^{\infty} B_n = \emptyset$.

Por lo tanto, $\lim_{n \rightsquigarrow \infty} \mu(B_n) = 0$, así que:

$$\mu(\bigcup_{j=1}^{\infty} A_j) = \text{lím}_{n\rightsquigarrow\infty} \mu(\bigcup_{j=1}^{n} A_j) = \text{lím}_{n\rightsquigarrow\infty} \sum_{k=1}^{n} \mu(A_j) = \sum_{k=1}^{\infty} \mu(A_j)$$

■

Medidas sobre álgebras y $\sigma-$álgebras

Definición 13. *Sea $\mathbb{F}$ un conjunto y $\Im$ una σ-álgebra de subconjuntos de $\mathbb{F}$. Se dice que una función no negativa $\mu : \Im \mapsto \overline{\mathbb{R}}$ es una medida si μ es σ-aditiva y $\mu(\emptyset) = 0$.*

Definición 14. *Llamaremos espacio de medida a una terna $(\mathbb{F}, \Im, \mu)$ donde $\mathbb{F}$ es un conjunto, $\Im$ una σ-álgebra de subconjuntos de $\mathbb{F}$ y $\mu : \Im \mapsto \overline{\mathbb{R}}$ una medida.*

En general, una medida se obtiene definiéndola primero para una familia de subconjuntos de un conjunto que no necesariamente forma una σ-álgebra; después se extiende a una familia más grande siguiendo el método de Henri Lebesgue. Lo más común es buscar definirla sobre un álgebra de subconjuntos de un conjunto y después extenderla a la σ-álgebra generada por esa álgebra.

Una medida que fue la base para que se desarrollara una teoría general de la medida es la que definió Lebesgue en el año 1902 y que, en honor a su nombre, se le conoce como medida de Lebesgue. Además esta medida es fundamental en la teoría de la probabilidad.

La medida de Lebesgue la podemos pensar como una extensión del concepto de longitud de un intervalo. Antes de introducirla necesitamos la siguiente definición:

Definición 15. *Diremos que un subconjunto A de $\mathbb{R}$ tiene medida cero si dado cualquier número $\varepsilon > 0$, existe una familia finita o infinita numerable de intervalos abiertos, cuya unión contiene al conjunto A, y tales que la suma de sus longitudes es menor que ε.*

Ahora enunciaremos el resultado que obtuvo Lebesgue, sin demostrarlo para no sobrecargar el texto.

Teorema 2. *Existe una función λ definida sobre la σ-álgebra $\mathfrak{L}(\mathbb{R})$ de subconjuntos de $\mathbb{R}$ generada por los intervalos y los conjuntos de medida cero, la cual tiene las siguientes propiedades:*

a) $\lambda(\emptyset) = 0$

b) $\lambda(I) = \ell(I)$ *para cualquier intervalo* I.

c) Si $E_1, E_2, \ldots$ es una colección infinita numerable de elementos de $\mathfrak{L}(\mathbb{R})$, ajenos por parejas, entonces:

$$\lambda\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \lambda\left(E_j\right)$$

La función λ será llamada la **medida de Lebesgue** sobre $\mathbb{R}$ y a los elementos de $\mathfrak{L}(\mathbb{R})$ los llamaremos **conjuntos Lebesgue medibles**.

Si B es un conjunto Lebesgue medible, la medida λ, restringida a los subconjuntos de B que son Lebesgue medibles, será llamada la medida de Lebesgue sobre B.

Una vez definida la medida λ, Lebesgue desarrolló el tema que le interesaba: definir un concepto de integral con el cual se tuvieran más funciones integrables que con el método de Riemann; además de que el nuevo concepto de integral tuviera mejores propiedades que la integral de Riemann.

Mostró entonces que si una función no negativa $f : \mathbb{R} \mapsto \overline{\mathbb{R}}$ es tal que, para cualquier $x \in \mathbb{R}$, el conjunto $\{y \in \mathbb{R} : f(y) \leq x\}$ es Lebesgue medible, entonces es posible definir la integral de f con respecto a λ, denotada por $\int_{\mathbb{R}} f d\lambda$, de tal manera que si $(g_n)_{n\in\mathbb{N}}$ es una sucesión no decreciente de funciones no negativas $g_n : \mathbb{R} \mapsto \overline{\mathbb{R}}$, cada una de ellas con la propiedad de que el conjunto $\{y \in \mathbb{R} : g_n(y) \leq x\}$ es Lebesgue medible, entonces la función $g : \mathbb{R} \mapsto \overline{\mathbb{R}}$, definida por $g(y) = \text{lím}_{n\rightsquigarrow\infty} g_n(y)$, tiene la misma propiedad y además:

$$\text{lím}_{n\to\infty} \int_{\mathbb{R}} g_n d\lambda = \int_{\mathbb{R}} \text{lím}_{n\to\infty} g_n d\lambda$$

Este resultado es conocido como el **teorema de la convergencia monótona** y a las funciones $f : \mathbb{R} \mapsto \overline{\mathbb{R}}$ que tienen la propiedad mencionada se les llama **funciones Lebesgue medibles**.

Definición 16. *Diremos que un espacio de medida $(\mathbb{F}, \Im, \mu)$ es completo si $\Im$ contiene a todos los subconjuntos de los conjuntos de medida μ igual a cero.*

Si un un espacio de medida $(\mathbb{F}, \Im_0, \mu)$ no es completo, se puede completar. En efecto, definamos $\Im$ como la familia de conjuntos de la forma $A\cup E$, donde $A \in \Im_0$ y E es un subconjunto de un conjunto $B \in \Im_0$ tal que $\mu(B) = 0$.

Observemos que $\Im_0 \subset \Im$ ya que si $A \in \Im_0$, entonces $A = A \cup \emptyset \in \Im$.

Demostremos que $\Im$ forma una σ-álgebra de subconjuntos de $\mathbb{E}$.

La prueba de que $\mathbb{E} \in \Im$ y que $\Im$ es cerrada bajo uniones numerables es inmediata. Para probar que $\Im$ es cerrada bajo complementos, sea $C = A \cup E \in \Im$, donde $A \in \Im_0$ y E es un subconjunto de un conjunto $B \in \Im_0$ tal que $\mu(B) = 0$. Entonces:

$C^c = (A \cup E)^c = A^c \cap E^c = A^c \cap [E^c \cap (B \cup B^c)]$

$= A^c \cap [(E^c \cap B) \cup (E^c \cap B^c)] = A^c \cap [(E^c \cap B) \cup B^c]$

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

Se tiene:

$A^c \cap B^c \in \Im$

$A^c \cap E^c \cap B \subset B$

Por lo tanto, $C^c \in \Im$.

Así que $\Im$ forma una σ-álgebra de subconjuntos de $\mathbb{E}$, la cual contiene a $\Im_0$.

Extendamos μ a $\Im$ definiendo $\mu\,(A \cup E) = \mu\,(A)$ para cualquier $A \in \Im_0$ y E cualquier subconjunto de un conjunto $B \in \Im_0$ tal que $\mu\,(B) = 0$.

Mostremos ahora que el espacio de medida $(\mathbb{F}, \Im, \mu)$ es completo.

Sea $C \in \Im$ tal que $\mu\,(C) = 0$, entonces $C = A \cup E$, donde $A \in \Im_0$ y $E \subset B$, para algún conjunto $B \in \Im_0$ tal que $\mu\,(B) = 0$.

Entonces:

$A \cup B \in \Im_0$

$\mu\,(A) = \mu\,(A \cup E) = \mu\,(C) = 0$

Así que, $\mu\,(A \cup B) = 0$.

Sea ahora $D \subset C$, entonces:

$$D \subset A \cup E \subset A \cup B$$

Así que $D = \emptyset \cup D \in \Im$.

Por lo tanto, $(\mathbb{F}, \Im, \mu)$ es completo.

Lo anterior también muestra que si $C \in \Im$ y $\mu\,(C) = 0$, entonces C está contenido en un conjunto $G \in \Im_0$ tal que $\mu\,(G) = 0$. Es decir, todo conjunto $C \in \Im$ de medida cero está contenido en un conjunto $G \in \Im_0$ de medida cero.

La terna $(\mathbb{R}, \mathfrak{L}\,(\mathbb{R}), \lambda)$ es un ejemplo de espacio de medida completo.

Espacios de probabilidad

Definición 17. *Llamaremos espacio de probabilidad a una terna $(\Omega, \Im, P)$, donde Ω es un conjunto, $\Im$ una σ-álgebra de subconjuntos de Ω y P una medida sobre $\Im$ tal que $P(\Omega) = 1$, a la cual llamaremos medida de probabilidad. A Ω lo llamaremos el espacio muestral, a los elementos de $\Im$ eventos y a la medida P de un evento A, la probabilidad de A.*

Como todo espacio de medida se puede completar, cuando trabajemos con un espacio de probabilidad $(\Omega, \Im, P)$, asumiremos que es completo.

Antes de su formulación axiomática, el Cáculo de Probabilidades tenía como base dos reglas, una de las cuales era llamada regla de la probabilidad compuesta". En esta segunda regla se hablaba de probabilidades condicionales sin haber definido lo que eso significaba matemáticamente, de manera que se utilizaba únicamente cuando las probabilidades condicionales con las que se trataba tenían un sentido intuitivo. Lo que se hizo entonces fue formular una definición, en lugar de un teorema o una regla.

Definición 18 (Probabilidad condicional). *Sean A y B dos eventos y supongamos $P(A) > 0$, se define la probabilidad condicional de B, dada la ocurrencia de A, $P(B|A)$, mediante la fórmula:*

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

La probabilidad condicional dado un evento A es una nueva medida de probabilidad, la cual asigna el valor 1 a A y el valor 0 a A^c; es decir, se trata de una medida de probabilidad concentrada en A. Se puede decir que al tomar probabilidades condicionales dado un evento A, el espacio muestral se reduce, convirtiéndose A en un nuevo espacio muestral.

Definición 19 (Independencia de eventos). *Diremos que los eventos de una familia no vacía cualquiera $\{A_\gamma\}$, finita o infinita, son estocásticamente independientes si dada cualquier subcolección finita de ellos, $A_{\gamma_1}, \ldots, A_{\gamma_n}$, donde $n \in \mathbb{N}$, se tiene:*

$$P(A_{\gamma_1} \cap \cdots \cap A_{\gamma_m}) = P(A_{\gamma_1}) \cdot P(A_{\gamma_m})$$

Definición 20 (Eventos mutuamente excluyentes). *Diremos que los eventos de una familia no vacía cualquiera $\{A_\gamma\}$, finita o infinita, son mutuamente excluyentes si cualquier par de ellos son conjuntos ajenos.*

Ahora, algunas propiedades simples:

Proposición 3. *Si los eventos de una familia no vacía* $\{A_\gamma\}_{\gamma\in\Gamma}$ *son mutuamente excluyentes y* Γ' *es cualquier subconjunto no vacío de* Γ*, entonces los eventos de la familia* $\{A_\gamma\}_{\gamma\in\Gamma'}$ *son mutuamente excluyentes.*

Proposición 4. *Si los eventos de una familia no vacía* $\{A_\gamma\}_{\gamma\in\Gamma}$ *son independientes y* Γ' *es cualquier subconjunto no vacío de* Γ*, entonces los eventos de la familia* $\{A_\gamma\}_{\gamma\in\Gamma'}$ *son independientes.*

Proposición 5. *Sea* $\{A_1,,\dots,A_n\}$ *una familia de eventos independientes, donde* $n\geq 2$*, y reemplacemos uno de ellos, cualquiera, por su complemento, entonces los eventos de la nueva familia siguen siendo independientes.*

Demostración

Reordenemos los eventos $A_1,,\dots,A_n$ de tal manera que el evento que reemplazamos por su complemento sea A_n.

Sea $\{n_1,\dots,n_k\}$ cualquier subconjunto no vacío del conjunto $\{1,\dots,n\}$, donde los elementos $n_1,\dots,n_k$ están ordenados del menor al mayor y $k\geq 2$. Entonces:

Si $n\notin\{n_1,\dots,n_k\}$, como los eventos de la familia $\{A_1,,\dots,A_n\}$ son independientes, se tiene:

$P\left(A_{n_1}\cap\cdots\cap A_{n_j}\right)=P\left(A_{n_1}\right)\cdots P\left(A_{n_j}\right)$

Si $n\in\{n_1,\dots,n_k\}$, entonces $n_k=n$ y se tiene:

$P\left(A_{n_1}\cap\cdots\cap A_{n_{k-1}}\cap A_n^c\right)$

$=P\left(A_{n_1}\cap\cdots\cap A_{n_{k-1}}-A_{n_1}\cap\cdots\cap A_{n_{k-1}}\cap A_n\right)$

$=P\left(A_{n_1}\cap\cdots\cap A_{n_{k-1}}\right)-P\left(A_{n_1}\cap\cdots\cap A_{n_{k-1}}\cap A_n\right)$

$=P\left(A_{n_1}\right)\cdots P\left(A_{n_{k-1}}\right)-P\left(A_{n_1}\right)\cdots P\left(A_{n_k-1}\right)P\left(A_n\right)$

$=P\left(A_{n_1}\right)\cdots P\left(A_{n_{k-1}}\right)\left[1-P\left(A_n\right)\right]=P\left(A_{n_1}\right)\cdots P\left(A_{n_{k-1}}\right)P\left(A_n^c\right)$

Así que los eventos de la familia $\{A_1,,\dots,A_{n-1},A_n^c\}$ son independientes.

■

Corolario 1. *Sea* $\{A_1,,\dots,A_n\}$ *una familia de eventos independientes y* U *un subconjunto del conjunto* $\{1,\dots,n\}$ *Para cada* $j\in\{1,\dots,n\}$*, definamos:*

$$B_j=\begin{cases}A_j & si\ j\in U\\ A_j^c & si\ j\notin U\end{cases}$$

Entonces, los eventos de la familia $\{B_1, \ldots, B_n\}$ *son independientes.*

Corolario 2. *Sean* $\{A_\gamma\}_{\gamma\in\Gamma}$ *una familia de eventos independientes y* T *un subconjunto de* Γ. *Para cada* $\gamma \in \Gamma$, *definamos:*

$$B_\gamma = \begin{cases} A_\gamma & si\ \gamma \in T \\ A_\gamma^c & si\ \gamma \notin T \end{cases}$$

Entonces, los eventos de la familia $\{B_\gamma\}_{\gamma\in\Gamma}$ *son independientes.*

Teorema 3. *Dado un evento* A *de probabilidad positiva, la función que asigna a cada evento* B *el número real* $P(B \mid A)$, *de acuerdo con la definición, es una medida de probabilidad.*

Demostración

La función así definida es claramente no negativa, además:

$$P(\Omega \mid A) = \frac{P(A\cap\Omega)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

Finalmente, si $\{B_n\}_{n\in\mathbb{N}}$ es una familia de eventos eventos mutuamente excluyentes, entonces:

$$P(\cup_{n=1}^{\infty} B_n \mid A) = \frac{P\left[A\cap(\cup_{n=1}^{\infty} B_n)\right]}{P(A)} = \frac{P\left[\cup_{n=1}^{\infty}(A\cap B_n)\right]}{P(A)} = \frac{\sum_{n=1}^{\infty} P(A\cap B_n)}{P(A)}$$

$$= \sum_{n=1}^{\infty} \frac{P(A\cap B_n)}{P(A)} = \sum_{n=1}^{\infty} P(B_n \mid A)$$

■

La definición de probabilidad condicional implica inmediatamente el siguiente resultado:

Proposición 6 (Regla del producto). *Sean* $A_1, \ldots, A_n$ n *eventos tales que* $P(A_1 \cap \ldots \cap A_{n-1}) > 0$, *entonces:*

$$P(\cap_{k=1}^{n} A_k) = P(A_n|A_1 \cap \ldots \cap A_{n-1}) \cdots P(A_2|A_1)P(A_1)$$

Proposición 7. *Sean* A *un evento tal que* $P(A) = 1$ *y* B *cualquier evento, entonces* $P(B) = P(B \cap A)$.

Demostración

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

Así que:

$P(B) = P(B \cap A) + P(B \cap A^c)$

Pero $B \cap A^c \subset A^c$ y $P(A^c) = 1 - P(A) = 0$; así que $P(B \cap A^c) = 0$. Por lo tanto:

$$P(B)=P\left(B\cap A\right)$$

Corolario 3 (**Regla de la probabilidad total**). *Sean B un evento cualquiera y $A_1, A_2\ldots$ una colección finita o infinita numerable de eventos de probabilidad positiva, mutuamente excluyentes y tales que $P\left(\bigcup_n A_n\right)=1$, entonces:*

$$P(B)=\sum_n P(B\mid A_n)P(A_n)$$

Teorema 4 (**Lema de Borel-Cantelli**). *Sea $A_1, A_2,\ldots$ una sucesión de eventos independientes tales que $\sum_{n=1}^{\infty} P\left(A_n\right)=\infty$ y sea:*

$$A=\{\omega\in\Omega:\omega\in A_n \text{ para una infinidad de valores de } n\}$$

Entonces $P(A)=1$.

Demostración

Para cada $m\in\mathbb{N}$, sea $B_m=\bigcap_{n=m}^{\infty} A_n^c$. Entonces la sucesión de eventos B_m es creciente y $A^c=\bigcup_{m=1}^{\infty} B_m$, así que:

$$P(A^c)=P\left[\bigcup_{m=1}^{\infty} B_m\right]=\lim_{m\to\infty} P\left[B_m\right]$$

Pero, $B_m\subset\bigcap_{n=m}^{m+k} A_n^c$ para cualquier $k\in\mathbb{N}$, así que, $P(B_m)\leq\prod_{n=m}^{m+k}\left[1-P(A_n)\right]$ para cualquier $k\in\mathbb{N}$. Por lo tanto:

$$P(B_m)\leq\lim_{k\rightsquigarrow\infty}\prod_{n=m}^{m+k}\left[1-P(A_n)\right]=0$$

Se concluye entonces que $P(A)=1-P(A^c)=1$. ■

Construcción de espacios de Probabilidad

Ejemplo 4. *Consideremos el experimento aleatorio consistente en lanzar una dado n veces consecutivas, anotando el resultado de cada uno de los lanzamientos.*

Este experimento aleatorio lo podemos modelar utilizando como espacio muestral al conjunto de todos los posibles resultados de los n lanzamientos, es decir, al conjunto Ω formado por todas las colecciones ordenadas de n números naturales que pertenecen al conjunto $\{1,2,3,4,5,6\}$:

$$\Omega=\{(x_1,x_2,\ldots,x_n):x_k\in\{1,2,3,4,5,6\} \text{ para cualquier } k\in\{1,2,\ldots,n\}\}$$

La familia de eventos la podemos considerar como el conjunto potencia de Ω, es decir la familia formada por todos los subconjuntos de Ω.

$$\Im=\mathcal{P}\left(\Omega\right)$$

Finalmente, como medida de probabilidad podemos tomar la que asigna a cada subconjunto de Ω el cociente que resulta de dividir el número de elementos de ese subconjunto entre 6^n, el cual es el número de elementos de Ω.

$$P(A) = \frac{\text{Número de elementos de } A}{6^n}$$

Tendríamos así definido nuestro espacio de probabilidad, el cual podemos pensarlo como modelo matemático del lanzamiento de n dados en forma consecutiva.

Recordemos que el modelo mismo es una abstracción, no requerimos de referinros al lanzamiento del dado para definirlo. Podríamos decir: un ejemplo de espacio de probabilidad es el siguiente: Definamos $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\Im$ como la familia de todos los subconjuntos de Ω y como medida de probabilidad tomemos a la función $P : \Im \to \mathbb{R}$ definida, para cada $A \in \Im$, mediante la relación $P(A) = \sum_{\{\omega \in A\}} p(\omega)$, donde $p(\omega) = \frac{1}{6^n}$ para cualquier $\omega \in \Omega$. El espacio de probabilidad $(\Omega, \Im, P)$ así definido está formado por elementos matemáticos abstractos que se tendrían que tratar como independientes de cualquier fenómeno aleatorio.

Una vez que tenemos definido el espacio de probabilidad de un experimento aleatorio ya queda definida la probabilidad de cualquier evento, así como cualquier probabilidad condicional.

a) ¿Cual es la probabilidad de obtener por lo menos una vez 6 en los n lanzamientos?

b) ¿Cuál es el más pequeño valor de n para el cual la probabilidad de obtener por lo menos una vez 6 en los n lanzamientos es mayor que $\frac{1}{2}$para el cual la probabilidad de obtener por lo menos una vez 6 en los n lanzamientos es mayor que $\frac{1}{2}$.

c) ¿Cual es la probabilidad de obtener por lo menos una vez 6 en los n lanzamientos sabiendo que en todos los lanzamientos se obtiene un número mayor o igual a 2?

Solución

a) Definamos el siguiente evento:

A : No se obtiene 6 en ninguno de los n lanzamientos.

Entonces A^c es el evento "*se obtiene por lo menos una vez* 6 *en los n lanzamientos*" y se tiene:

$$P(A) = \frac{5^n}{6^n} = \left(\frac{5}{6}\right)^n$$

Por lo tanto:

$$P(A^c) = 1 - P(A) = 1 - \left(\frac{5}{6}\right)^n$$

b) Para que $P(A^c) > \frac{1}{2}$ se requiere que $1 - \left(\frac{5}{6}\right)^n > \frac{1}{2}$; es decir, $\left(\frac{5}{6}\right)^n < \frac{1}{2}$. Tomando el logaritno natural en ambos lados de la desigualdad, ésta no se altera ya que la función logaritmo es creciente. Así que tenemos:

$$n \ln\left(\frac{5}{6}\right) < \ln\left(\frac{1}{2}\right) = -\ln 2$$

De aquí se sigue que:

$$n > -\frac{\ln 2}{\ln\left(\frac{5}{6}\right)} = \frac{\ln 2}{\ln\left(\frac{6}{5}\right)} \approx 3.8$$

Así que $n = 4$ es el más pequeño valor de n para el cual la probabilidad de obtener por lo menos una vez 6 en los n lanzamientos es mayor que $\frac{1}{2}$.

c) Definamos los siguientes eventos:

B : Se obtiene por lo menos una vez 6 en los n lanzamientos.

A : En todos los lanzamientos se obtiene un número mayor o igual a 2.

Se tiene:

$P(A) = \frac{5^n}{6^n}$

$P(A \cap B^c) = \frac{4^n}{6^n}$

Así que:

$P(A \cap B) = P(A - A \cap B^c) = P(A) - P(A \cap B^c) = \frac{5^n}{6^n} - \frac{4^n}{6^n} = \frac{5^n - 4^n}{6^n}$

$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{5^n - 4^n}{5^n} = 1 - \left(\frac{4}{5}\right)^n$

Ejercicio 6. *Consideremos el experimento aleatorio consistente en lanzar dos dados n veces consecutivas, anotando el par de resultados de cada uno de los lanzamientos.*

a) Encuentra un espacio de probabilidad que modele este experimento aleatorio.

b) ¿Cual es la probabilidad de obtener por lo menos una vez la pareja $(6, 6)$ en los n lanzamientos?

c) ¿Cuál es el más pequeño valor de n para el cual la probabilidad de obtener por lo menos una vez la pareja $(6, 6)$ en los n lanzamientos es mayor que $\frac{1}{2}$.

Ejemplo 5. *Como ya lo mencionamos, en algunos problemas de probabilidad se hace referencia al concepto de experimento aleatorio y se construye un espacio de probabilidad ad* *hoc para ese experimento, de manera similar a como lo hicimos con el lanzamiento del dado, pero nuevamente, el espacio de probabilidad que se construye tiene que tratarse como independiente de cualquier fenómeno aleatorio.*

Un caso particular de experimento aleatorio es lo que se conoce como **ensayo de Bernoulli**, el cual se define como un experimento aleatorio que admite únicamente dos posibles resultados, a uno de los cuales se le llama éxito y al otro fracaso. Utilizando este concepto, podemos, por ejemplo, definir algunas variables aleatorias de interés o plantearnos algunos problemas de probabilidad que historicamente fueron importantes para el desarrollo de la Teoría de la Probabilidad.

Por, ejemplo, definamos el experimento aleatorio consistente en la realización consecutiva de n ensayos de Bernoulli, cada uno de ellos independiente de los demás y tal que la probabilidad de obtener éxito es igual a un número $p \in [0, 1]$, el cual es el mismo para cada uno de los ensayos. Como resultado del experimento consideremos la $n-$ada de resultados que se obtienen en los n lanzamientos, denotando por 1 a la obtención de éxito y por 0 a la obtención de fracaso.

El espacio muestral Ω_n de este experimento aleatorio lo podemos representar de la siguiente manera:

$$\Omega_n = \{(s_1, s_2, \ldots, s_n) : s_j \in \{0, 1\} \text{ para cualquier } j \in \{1, 2, \ldots, n\}\}$$

Como los ensayos de Bernoulli son independientes y la probabilidad de éxito en cada en sayo es p, la probabilidad de obtener la secuencia $(s_1, s_2, \ldots, s_n)$ es un producto donde cada factor es p o $1-p$, dependiendo de que se obtenga éxito o fracaso.

Como la obtención de éxito lo estamos representando por el número 1 y la de fracaso por 0, si $s_j = 1$, el factor correspondiente es p, mientras que si $s_j = 0$, el factor es $1-p$; esto se puede expresar como $ps_j + (1-p)(1-s_j)$. Así que, para cada $\omega = (s_1, s_2, \ldots, s_n) \in \Omega$, se tiene:

$$P_n(\{\omega\}) = P_n(\{(s_1, s_2, \ldots, s_n)\}) = \prod_{j=1}^{n} [ps_j + (1-p)(1-s_j)]$$

Vamos a mostrar que con esta asignación de probabilidades, obtenemos un espacio de probabilidad $(\Omega_n, \Im_n, P_n)$, donde $\Im_n$ es la $\sigma-$álgebra formada por todos los subconjuntos de Ω_n y P_n es la función, definida para cada subconjunto A de Ω_n de la siguiente manera:

$$P_n(A) = \sum_{\{\omega \in A\}} P_n(\{\omega\})$$

Que P_n es $\sigma-$aditiva y $P_n(\emptyset) = 0$ es inmediato.

Así que únicamente resta probar que $P_n(\Omega_n) = 1$.

Para cada $n \in \mathbb{N}$, definamos $S(n) = \sum_{\{\omega \in \Omega_n\}} P_n(\{\omega\})$. Entonces:

Como $\Omega_1 = \{(0), (1)\}$, se tiene:

$$S(1) = P_1(\{(0)\}) + P_1(\{(1)\}) = (1-p) + p = 1$$

Además, para cualquier $n \in \mathbb{N}$, se tiene:

$\Omega_{n+1} = \{(s_1, s_2, \ldots, s_n, s_{n+1}) : s_j \in \{0,1\} \text{ para cualquier } j \in \{1, 2, \ldots, n, n+1\}\}$

$= \{(s_1, s_2, \ldots, s_n, 0) : s_j \in \{0,1\} \text{ para cualquier } j \in \{1, 2, \ldots, n\}\}$

$\cup \{(s_1, s_2, \ldots, s_n, 1) : s_j \in \{0,1\} \text{ para cualquier } j \in \{1, 2, \ldots, n\}\}$

Así que:

$S(n+1) = \sum_{\{\omega \in \Omega_{n+1}\}} P_{n+1}(\{\omega\}) = (1-p)\, S(n) + pS(n) = S(n)$

Por lo tanto, $S(n) = 1$ para cualquier $n \in \mathbb{N}$; lo cual significa que $P_n(\Omega_n) = 1$.

¿Cuál es la probabilidad de obtener exactamente k éxitos en los n ensayos?

Solución

Hay $\binom{n}{k}$ secuencias $(s_1, s_2, \ldots, s_n) \in \Omega_n$ en las cuales k lugares están ocupados por el número 1 y el resto por el número 0. Así que si, para cada $k \in \{1, 2, \ldots, n\}$ definimos el evento:

A_k : Se obtienen k éxitos en los ensayos,

se tiene:

$$P(A_k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Obtenemos lo que se llama probabilidades de tipo binomial, nombre que proviene del hecho de que los términos $\binom{n}{k} p^k (1-p)^{n-k}$ son los del desarrollo de $[p + (1-p)]^n$:

$$[p + (1-p)]^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

Ejemplo 6. *Una urna contiene 5 bolas numeradas del 1 al 5. Un experimento aleatorio consiste en seleccionar bolas de la urna, una por una, siendo cada elección al azar. En cada elección se anota el número de la bola que se selecciona y se deja ésta fuera de la urna. El experimento concluye cuando se obtiene la bola número 5 y consideramos como resultado la colección ordenada de números que se obtienen.*

a) Construye el espacio de probabilidad de este experimento, es decir la terna $(\Omega, \mathcal{A}, P)$, donde Ω es el espacio muestral, $\mathcal{A}$ la familia de eventos y P la función de probabilidad.

b)Verifica que la función P que definiste satisface las 3 propiedades que la caracterizan.

Solución

Definamos:

$\Omega_1 = \{(5)\}$

$\Omega_2 = \{(1,5), (2,5), (3,5), (4,5)\}$

$\Omega_3 = \{(1,2,5), (1,3,5), (1,4,5), (2,1,5), (2,3,5), (2,4,5),$

$(3,1,5), (3,2,5), (3,4,5), (4,1,5), (4,2,5), (4,3,5)\}$

$\Omega_4 = \{(1,2,3,5), (1,2,4,5), (1,3,2,5), (1,3,4,5), (1,4,2,5), (1,4,3,5),$

$(2,1,3,5), (2,1,4,5), (2,3,1,5), (2,3,4,5), (2,4,1,5), (2,4,3,5),$

$(3,1,2,5), (3,1,4,5), (3,2,1,5), (3,2,4,5), (3,4,1,5), (3,4,2,5),$

$(4,1,2,5), (4,1,3,5), (4,2,1,5), (4,2,3,5), (4,3,1,5), (4,3,2,5)\}$

$\Omega_5 = \{(1,2,3,4,5), (1,2,4,3,5), (1,3,2,4,5), (1,3,4,2,5), (1,4,2,3,5), (1,4,3,2,5),$

$(2,1,3,4,5), (2,1,4,3,5), (2,3,1,4,5), (2,3,4,1,5), (2,4,1,3,5), (2,4,3,1,5),$

$(3,1,2,4,5), (3,1,4,2,5), (3,2,1,4,5), (3,2,4,1,5), (3,4,1,2,5), (3,4,2,1,5),$

$(4,1,2,3,5), (4,1,3,2,5), (4,2,1,3,5), (4,2,3,1,5), (4,3,1,2,5), (4,3,2,1,5)\}$

El espacio muestral del experimento aleatorio puede entonces representarse de la siguiente manera:

$$\Omega = \bigcup_{k=1}^{5} \Omega_k$$

Se tiene entonces:

$$P(\{\omega\}) = \begin{cases} \frac{1}{5} & \text{si } \omega \in \Omega_1 \\ \frac{1}{5\cdot 4} & \text{si } \omega \in \Omega_2 \\ \frac{1}{5\cdot 4\cdot 3} & \text{si } \omega \in \Omega_3 \\ \frac{1}{5\cdot 4\cdot 3\cdot 2} & \text{si } \omega \in \Omega_4 \\ \frac{1}{5\cdot 4\cdot 3\cdot 2\cdot 1} & \text{si } \omega \in \Omega_5 \end{cases}$$

Definamos $\mathcal{A}$ como el conjunto potencia de Ω, es decir, la familia formada por todos los subconjuntos de Ω.

Para cada $A \in \mathcal{A}$, definamos:

$$P(A) = \sum_{\{\omega \in A\}} P(\{\omega\})$$

Obviamente esta función P es no negativa y finitamente aditiva. Además:

$\# (\Omega_1) = 1$

$\# (\Omega_2) = 4$

$\# (\Omega_3) = 4 \cdot 3$

$\# (\Omega_4) = 4 \cdot 3 \cdot 2$

$\# (\Omega_5) = 4 \cdot 3 \cdot 2 \cdot 1$

Así que:

$P (\Omega_1) = \frac{1}{5}$

$P (\Omega_2) = 4 \cdot \left(\frac{1}{5 \cdot 4}\right) = \frac{1}{5}$

$P (\Omega_3) = 4 \cdot 3 \cdot \left(\frac{1}{5 \cdot 4 \cdot 3}\right) = \frac{1}{5}$

$P (\Omega_4) = 4 \cdot 3 \cdot 2 \cdot \left(\frac{1}{5 \cdot 4 \cdot 3 \cdot 2}\right) = \frac{1}{5}$

$P (\Omega_5) = 4 \cdot 3 \cdot 2 \cdot 1 \cdot \left(\frac{1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}\right) = \frac{1}{5}$

Por lo tanto:

$$P (\Omega) = \sum_{k=1}^{5} P (\Omega_k) = 5 \left(\frac{1}{5}\right) = 1$$

Ejercicio 7. *Se realizan consecutivamente ensayos de Bernoulli independientes, con probabilidad de éxito* p *en cada ensayo, hasta que se presente alguna de las dos siguientes situaciones: 1) se obtiene éxito por segunda ocasión; 2) se realiza el cuarto ensayo. Considera como resultado del experimento la colección ordenada de los resultados que se obtienen en los ensayos de Bernoulli que se realizan.*

a) Construye el espacio de probabilidad de este experimento, es decir la terna $(\Omega, \mathcal{A}, P)$, donde Ω es el espacio muestral, $\mathcal{A}$ la familia de eventos y P la función de probabilidad.

b)Verifica que la función P que definiste satisface las 3 propiedades que la caracterizan (es no negativa, σ-aditiva y $P (\Omega) = 1$).

Ejemplo 7. *Una urna contiene* 6 *bolas numeradas del* 1 *al* 6*. Un experimento aleatorio consiste en seleccionar bolas de la urna, una por una, siendo cada elección al azar. En cada elección se anota el número de la bola que se selecciona y se deja ésta fuera de la urna. El* *experimento concluye cuando se obtiene la bola marcada con el número 1 y consideramos como resultado del experimento a la colección ordenada de números que se obtienen.*

a) Construye un espacio de probabilidad que represente a este experimento, es decir una terna $(\Omega, \mathcal{A}, P)$, donde Ω es el espacio muestral, $\mathcal{A}$ la familia de eventos y P la función de probabilidad.

b)Verifica que la función P que definiste satisface las 3 propiedades que la caracterizan.

Solución

Para $k \in \{1, 2, 3, 4, 5, 6\}$, definamos Ω_k como el conjunto de posibles resultados que terminan en la k-ésima elección.

Obsérvese que el número de elementos de Ω_k es igual a $\frac{5!}{(5-k+1)!} = \frac{5!}{(6-k)!}$.

Además, los posibles resultados que pertenecen a Ω_k son equiprobables.

Finalmente, si $\omega \in \Omega_k$, entonces:

$$P(\{\omega\}) = \frac{1}{\frac{6!}{(6-k)!}} = \frac{(6-k)!}{6!}$$

Definamos $\Omega = \bigcup_{k=1}^{6} \Omega_k$ y $\mathcal{A}$ como el conjunto potencia de Ω, es decir, la familia formada por todos los subconjuntos de Ω.

Para cada $A \in \mathcal{A}$, definamos:

$$P(A) = \sum_{\{\omega \in A\}} P(\{\omega\})$$

donde convenimos en que una suma de los elementos de un conjunto vacío es igual a 0.

Obviamente esta función P es no negativa. Además, para cualquier $k \in \{1, 2, 3, 4, 5, 6\}$, se tiene:

$$P(\Omega_k) = \frac{5!}{(6-k)!} \frac{(6-k)!}{6!} = \frac{1}{6}$$

Así que:

$$P(\Omega) = \sum_{k=1}^{6} P(\Omega_k) = 6\left(\frac{1}{6}\right) = 1$$

Demostremos ahora que P es finitamente aditiva.

Sean $A, B \in \mathcal{A}$ dos conjuntos ajenos.

Si $A = \emptyset$, entonces $P(A) = 0$ y $A \cup B = B$, así que:

$$P\left(A\cup B\right)=\sum_{\{\omega\in B\}}P\left(\{\omega\}\right)=P\left(B\right)=P\left(A\right)+P\left(B\right)$$

De la misma manera, si $B=\emptyset$, se tiene $P\left(A\cup B\right)=P\left(A\right)=P\left(A\right)+P\left(B\right)$.

Supongamos ahora que A y B son distintos del vacío y sean $A=\{\alpha_1,\alpha_2,\ldots,\alpha_i\}$ y $B=\left\{\beta_1,\beta_2,\ldots,\beta_j\right\}$. Entonces:

$$A\cup B=\left\{\alpha_1,\alpha_2,\ldots,\alpha_i,\beta_1,\beta_2,\ldots,\beta_j\right\}$$

Así que:

$$P\left(A\cup B\right)=\sum_{k=1}^{i}P\left(\{\alpha_k\}\right)+\sum_{k=1}^{j}P\left(\{\beta_k\}\right)=P\left(A\right)+P\left(B\right)$$

Por lo tanto, P es finitamente aditiva.

Sea ahora $A_1,A_2,\ldots$ una colección infinita numerable de elementos de $\mathcal{A}$, ajenos por parejas. Si en esa colección hubiera una infinidad de conjuntos distintos del vacío, la unión de ellos sería un conjunto infinito contenido en Ω, lo cual es imposible ya que Ω es finito. Por lo tanto, en la colección $A_1,A_2,\ldots$, o bien todos son conjuntos vacíos, o bien únicamente hay una colección finita de ellos los cuales son distintos del vacío.

Si todos los conjuntos de la colección $A_1,A_2,\ldots$ son iguales al conjunto vacío, entonces $\bigcup_{k=1}^{\infty}A_k=\emptyset$, así que $\sum_{k=1}^{\infty}P\left(A_k\right)=0=P\left(\bigcup_{k=1}^{\infty}A_k\right)$.

Si no todos los conjuntos de la colección $A_1,A_2,\ldots$ son iguales al conjunto vacío, sean $A_{j_1},A_{j_2},\ldots,A_{j_n}$ los únicos conjuntos de la colección que son distintos del vacío. Entonces

$$P\left(\bigcup_{k=1}^{\infty}A_k\right)=P\left(\bigcup_{k=1}^{n}A_{j_k}\right)=\sum_{k=1}^{n}P\left(A_{j_k}\right)=\sum_{k=1}^{\infty}P\left(A_k\right)$$

Por lo tanto, P es σ-aditiva.

Así que la terna $(\Omega,\mathcal{A},P)$ es un espacio de probabilidad correspondiente al experimento aleatorio dado.

Cálculo de probabilidades

Lo expuesto anteriormente es la base para calcular probabilidades; en particular, es importante saber trabajar con la regla de la probabilidad total ya que se requiere utilizarla en la solución de muchos problemas.

El método general para calcular la probabilidad de un evento consiste en expresarlo como una unión de eventos mutuamente excluyentes cuyas probabilidades sean más sencillas de calcular.

Ejemplo 8. *Calcula el número de lanzamientos de tres dados que se requieren para que la probabilidad de obtener por lo menos en una ocasión* 3 *seises sea mayor que* $\frac{1}{2}$.

Solución

Definamos:

A: Se obtiene por lo menos una tercia de cincos en n lanzamientos.

Entonces:

$$P(A) = 1 - \left(\tfrac{215}{216}\right)^n$$

Así que $P(A) \geq \frac{1}{2}$ si y sólo si $n \geq -\frac{\ln 2}{\ln \frac{215}{216}} = 149.4$.

Por lo tanto, se requieren por lo menos 150 lanzamientos.

Ejercicio 8. *Calcula el número de lanzamientos de cinco dados que se requieren para que la probabilidad de obtener por lo menos en una ocasión* 5 *seises sea mayor que* $\frac{1}{2}$.

Ejercicio 9. *En una caja hay 3 fichas, una blanca, una negra y una amarilla. Tres personas, A, B y C, juegan el siguiente juego: cada una elige una ficha al azar, con la condición de que quien obtenga la ficha amarilla acumula un punto. Una vez definido quién acumula el punto, regresan las fichas a la caja y vuelven a seleccionar al azar una ficha cada uno, bajo las mismas condiciones. Convienen en que el juego lo gana quien acumule primero* 5 *puntos. Cuando A tiene acumulados 3 puntos, B 3 puntos y C 4 puntos, ¿cuál es la probabilidad que cada jugador tiene de ganar el juego?*

Ejemplo 9. *Dos personas A y B juegan a lanzar consecutivamente una moneda balanceada con la condición de que cada vez que se obtenga cara, A gana un punto, mientras que B lo gana cuando se obtiene cruz. Convienen en que gana el juego quien obtenga primero n puntos. Supongamos que en un momento dado al jugador A le faltan r puntos para ganar y al jugador B s puntos. Si el juego continúa partiendo de esa situación, ¿cuál es la probabilidad que tiene cada jugador de ganar el juego?*

Solución

Este problema tiene una gran importancia desde el punto de vista histórico pues con base en él se desarrollaron algunos métodos de solución de problemas de probabilidad. El ejercicio consiste en justificar la siguiente regla, la cual fue encontrada por Pascal:

La regla está basada en el llamado triángulo aritmético, el cual se construye de la siguiente manera:

$$\begin{array}{ccccccccccccccccc}
&&&&&&&&1&&&&&&&&\\
&&&&&&&1&&1&&&&&&&\\
&&&&&&1&&2&&1&&&&&&\\
&&&&&1&&3&&3&&1&&&&&\\
&&&&1&&4&&6&&4&&1&&&&\\
&&&1&&5&&10&&10&&5&&1&&&\\
&&1&&6&&15&&20&&15&&6&&1&&\\
&1&&7&&21&&35&&35&&21&&7&&1&\\
1&&8&&28&&56&&70&&56&&28&&8&&1\\
\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots
\end{array}$$

Cada elemento de un renglón se obtiene sumando los dos elementos adyacentes del renglón anterior.

Entonces, decía Pascal, si a A le faltan r puntos y a B s puntos, para determinar la probabilidad que cada jugador tiene de ganar el juego, tómese del triángulo aritmético el renglón que tenga $r+s$ elementos. Si H es la suma de los primeros s elementos y G es la suma de los r últimos, las probabilidades de que A y B ganen el juego son $\frac{H}{2^{r+s-1}}$ y $\frac{G}{2^{r+s-1}}$, respectivamente.

Solución

Observemos que en la situación dada, para que alguno de los dos gane el juego necesitarían lanzar la moneda a lo más $r+s-1$ veces. Así que consideremos el experimento consistente en jugar ese número de lanzamientos. Este experimento tiene un total de 2^{r+s-1} posibles resultados, los cuales son equiprobables y cada uno de ellos se puede representar mediante una serie de letras del tipo $(P, P, Q, B, Q, Q, \ldots)$, en donde P y Q significa que el jugador P y Q ganan el correspondiente lanzamiento, respectivamente. De estas series, aquellas que tengan por lo menos r veces la letra P representan resultados favorables al triunfo de P y aquellas que tengan por lo menos s veces la letra Q representan resultados favorables al triunfo de Q.

Pero de entre todas las posibles series de $r+s-1$ letras, el número de aquéllas en las que hay exactamente k veces la letra P es igual al número de combinaciones de $r+s-1$ objetos tomadas de k en k, es decir $\binom{r+s-1}{k}$. Por lo tanto, el total de resultados favorables al triunfo de P es:

$$\binom{r+s-1}{r} + \binom{r+s-1}{r+1} + \cdots + \binom{r+s-1}{r+s-1}$$

y el total de resultados favorables al triunfo de Q es:

$$\binom{r+s-1}{s} + \binom{r+s-1}{s+1} + \cdots + \binom{r+s-1}{r+s-1}$$

Entonces si p es la probabilidad de que P gane y q la probabilidad de que Q gane, tenemos:

$$p = \frac{\binom{r+s-1}{r} + \binom{r+s-1}{r+1} + \cdots + \binom{r+s-1}{r+s-1}}{2^{r+s-1}}$$

$$q = \frac{\binom{r+s-1}{s}+\binom{r+s-1}{s+1}+\cdots+\binom{r+s-1}{r+s-1}}{2^{r+s-1}}$$

Por otra parte, observemos que $\binom{1}{0} = \binom{1}{1} = 1$, así que, gracias a la relación $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$, el triángulo aritmético puede expresarse de la siguiente manera:

$$\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & \binom{1}{0} & & \binom{1}{1} & & & & & \\
 & & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & \\
 & & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & & \\
 & & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & & \\
 & \binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5} & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{array}$$

De esta manera, si consideramos el renglón del triángulo aritmético con $r+s$ elementos, la suma de los primeros s elementos resulta ser:

$$H = \binom{r+s-1}{0} + \binom{r+s-1}{1} + \cdots + \binom{r+s-1}{s-1} = \binom{r+s-1}{r} + \binom{r+s-1}{r+1} + \cdots + \binom{r+s-1}{r+s-1}$$

La suma de los últimos r elementos resulta ser:

$$G = \binom{r+s-1}{s} + \binom{r+s-1}{s+1} + \cdots + \binom{r+s-1}{r+s-1}$$

Y entonces, efectivamente, de acuerdo al resultado obtenido previamente, las probabilidades de que A y B ganen el juego son $\frac{H}{2^{r+s-1}}$ y $\frac{G}{2^{r+s-1}}$, respectivamente.

Ejemplo 10. *Dada una población formada exclusivamente por hermanos gemelos, consideremos el experimento aleatorio consistente en seleccionar primero una pareja de hermanos gemelos al azar y después, también al azar, uno de los hermanos de esa pareja. Supongamos que la probabilidad de que los hermanos gemelos seleccionados sean del mismo sexo es igual a* $0{,}7$ *y que la probabilidad de que la persona seleccionada, en la segunda parte del experimento, sea de sexo masculino es igual a* $0{,}4$*. Sabiendo que la persona seleccionada en la segunda parte del experimento es de sexo masculino, ¿cuál es la probabilidad de que su hermano también lo sea?*

Solución

El espacio muestral del experimento aleatorio descrito está dado por:

$$\Omega = \{(MM)M, (FF)F, (MF)M, (MF)F\}$$

Se sabe que:

$P\left(\{(MM)M, (MF)M\}\right) = 0{,}4$

$P(\{(MM)M,(FF)F\}) = 0{,}7$

$P(\{(MF)M\}) = P(\{(MF)F\})$

Así que:

$P(\{(MF)M\}) = P(\{(MF)F\}) = 0{,}15$

$P(\{(MM)M\}) = 0{,}25$

$P(\{(FF)F\}) = 0{,}35$

Definamos los eventos:

A: El hermano gemelo seleccionado es de sexo masculino.

B: El hermano del gemelo seleccionado es de sexo masculino.

Entonces:

$A = \{(MM)M,(MF)M\}$

$B = \{(MM)M,(MF)F\}$

$P(B \mid A) = \frac{P(A\cap B)}{P(A)} = \frac{25}{40} = \frac{5}{8} = 0{,}625$

Otro método:

Definamos el evento:

C: Los dos hermanos gemelos tienen el mismo sexo.

Entonces:

$P(A \cap B) = P(A \cap C) = P(A) - P(A \cap C^c)$

$= P(A) - P(A \mid C^c)P(C^c) = 0{,}4 - (0{,}5)(0{,}3) = 0{,}25$

$P(B \mid A) = \frac{P(A\cap B)}{P(A)} = \frac{25}{40} = \frac{5}{8} = 0{,}625$

Ejemplo 11. *Dos jugadores, A y B, juegan a lanzar sucesivamente un dado balanceado, con la condición de que el ganador será el primero que obtenga cinco o seis. A tiene primero dos turnos y después B tres turnos; si ninguno de los dos obtiene cinco o seis, continúan lanzando el dado, A dos veces, después B tres veces; y así sucesivamente. ¿Qué probabilidad tiene cada jugador de ganar el juego?*

Solución

Denotemos por A, y B a las probabilidades de que A y B ganen el juego, respectivamente.

$P(A) = \frac{5}{9} + \left(\frac{2}{3}\right)^5 \frac{5}{9} + \left(\frac{2}{3}\right)^{10} \frac{5}{9} + \cdots = \frac{5}{9} \frac{1}{1-\left(\frac{2}{3}\right)^5} = \frac{5}{9} \frac{243}{211} = \frac{135}{211} = 0{,}63981$

$P(B) = \left(\frac{2}{3}\right)^2 \frac{19}{27} + \left(\frac{2}{3}\right)^7 \frac{19}{27} + \left(\frac{2}{3}\right)^{12} \frac{19}{27} + \cdots = \left(\frac{2}{3}\right)^2 \frac{19}{27} \frac{1}{1-\left(\frac{5}{6}\right)^6} = \frac{76}{243} \frac{243}{211} = \frac{76}{211} = 0{,}36019$

Ejemplo 12. *Consideremos el experimento aleatorio consistente en la realización consecutiva de una infinidad de ensayos de Bernoulli, cada uno de ellos independiente de los demás y tal que la probabilidad de obtener éxito es igual a* $p \in (0,1)$. *Como resultado del experimento consideremos la sucesión de resultados que se obtienen en la infinidad de ensayos, denotando por* 1 *a la obtención de éxito y por* 0 *a la obtención de fracaso.*

a) ¿Cuál es la probabilidad de obtener un número infinito de éxitos al realizar el experimento?

b) Para $k \in \{0, 1, 2, \ldots\}$, ¿cuál es la probabilidad de que se obtengan k fracasos antes del primer éxito?

c) Para $k \in \{0, 1, 2, \ldots\}$ y $r \in \mathbb{N}$, ¿cuál es la probabilidad de que que se obtengan k fracasos antes del $r-$simo éxito?

Solución

El espacio muestral Ω de este experimento aleatorio lo podemos representar de la siguiente manera:

$$\Omega = \left\{ (s_n)_{n\in\mathbb{N}} : s_j \in \{0,1\} \text{ para cualquier } j \in \mathbb{N} \right\}$$

a) Para cada $n \in \mathbb{N}$, definamos el evento:

A_n : Se obtiene éxito en el $n-$simo ensayo.

Se tiene $\sum_{n=1}^{\infty} P(A_n) = \infty$; así que, por el lema de Borel Cantelli:

$$P(\{\omega \in \Omega : \omega \in A_n \text{ para una infinidad de valores de } n\}) = 1$$

b) Para cada $k \in \{0, 1, 2, \ldots\}$, definamos el evento:

B_k : Se obtienen k fracasos antes del primer éxito.

Se tiene entonces:

$$P(B_k) = p(1-p)^k$$

c) Para cada $k \in \{0, 1, 2, \ldots\}$ y $r \in \mathbb{N}$, definamos el evento:

C_k : Se obtienen k fracasos antes del antes del $r-$simo éxito.

C_k ocurre si y sólo si en los primeros $k+r-1$ ensayos hay k fracasos y $r-1$ éxitos, después de cual se obtiene éxito. Así que:

$$P\left(C_k\right)=\binom{k+r-1}{k}p^{r-1}\left(1-p\right)^k p=\binom{k+r-1}{k}p^r\left(1-p\right)^k$$

Ejemplo 13. *Supongamos que una prueba de sangre de una cierta enfermedad contagiosa es tal que las probabilidades de un falso negativo (la prueba es negativa para una persona que tiene la enfermedad) y de un falso positivo (la prueba es positiva para una persona que no tiene la enfermedad) son* 0,02 *y* 0,03*, respectivamente. Supongamos que cuando la prueba resulta negativa en una persona seleccionada al azar, la probabilidad de que la persona esté contagiada es igual a* 0,0015*. Si la prueba resulta positiva en una persona seleccionada al azar, ¿cuál es la probabilidad de que la persona esté contagiada?*

Solución

Definamos los siguientes eventos:

A: La persona está contagiada

B: La prueba es positiva

Sabemos:

$P(B\mid A^c)=0{,}03$

$P(B^c\mid A)=0{,}02$

$P(A\mid B^c)=0{,}0015$

Sea $p=P(A)$, se tiene entonces:

$\frac{15}{10000}=P(A\mid B^c)=\frac{P(B^c\cap A)}{P(B^c)}=\frac{P(B^c|A)P(A)}{P(B^c|A)P(A)+P(B^c|A^c)P(A^c)}=\frac{(0{,}02)p}{(0{,}02)p+(0{,}97)(1-p)}=\frac{2p}{2p+97(1-p)}$

Así que, $p=\frac{291}{4285}=0{,}067911$.

$P(B)=P(B\mid A)P(A)+P(B\mid A^c)P(A^c)=(0{,}98)\,p+(0{,}03)\,(1-p)$

$=\frac{98}{100}\frac{291}{4285}+\frac{3}{100}\left(1-\frac{291}{4285}\right)=\frac{81}{857}=0{,}094516$

$P(A\mid B)=\frac{P(B\cap A)}{P(B)}=\frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^c)P(A^c)}=\frac{\frac{98}{100}\frac{291}{4285}}{\frac{98}{100}\frac{291}{4285}+\frac{3}{100}\left(1-\frac{291}{4285}\right)}$

$=\frac{(98)(291)}{(98)(291)+3(3994)}=\frac{4753}{6750}=0{,}70415$

Ejercicio 10. *Supongamos que una prueba de sangre de una cierta enfermedad contagiosa es tal que las probabilidades de un falso negativo y de un falso positivo son ambas iguales a* 0,01*. Supongamos además que la probabilidad de que una persona elegida al azar esté contagiada es igual a* p*. ¿Para qué valores de* p*, están contagiadas más del* 90 % *de la personas para las cuales la prueba resulta positiva?*

Ejemplo 14. *La probabilidad de que un documento se encuentre en alguno de los* 8 *cajones de un escritorio es igual a* $\frac{1}{4}$*. Si el documento se encuentra en el escritorio, es igualmente probable que se encuentre en cualquiera de sus* 8 *cajones. Se busca el documento en* 3 *de los cajones y no se encuentra. Calcula la probabilidad de encontrarlo en alguno de los* 5 *cajones restantes.*

Solución

Definamos los eventos:

A : El documento no se encuentra en los primeros 3 cajones.

B : El documento se encuentra en el cuarto, quinto, sexto, séptimo u octavo cajón.

C : El documento se encuentra en el escritorio.

Se pide encontrar $P(B \mid A)$.

Se tiene:

$P(C) = \frac{1}{4}$

$P(A \mid C) = \frac{5}{8}$

$P(A \mid C^c) = 1$

$P(A \cap B \mid C) = \frac{5}{8}$

$P(A \cap B \mid C^c) = 0$

$P(A) = P(A \mid C)P(C) + P(A \mid C^c)P(C^c) = \frac{5}{8}\frac{1}{4} + \frac{3}{4} = \frac{29}{32}$

$P(A \cap B) = P(A \cap B \mid C)P(C) + P(A \cap B \mid C^c)P(C^c) = \frac{5}{8}\frac{1}{4} = \frac{5}{32}$

Así que:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{5}{32}}{\frac{29}{32}} = \frac{5}{29} = 0{,}17241$$

Ejercicio 11. *Se tienen* 4 *escritorios, cada uno con dos cajones. Los dos primeros escritorios contienen una moneda de oro en cada cajón, el tercero contiene una moneda de plata en cada* *cajón y el cuarto contiene una moneda de oro en un cajón y una de plata en el otro. Se elige un escritorio al azar, se abre un cajón al azar y después el otro. Si en el primer cajón se encuentra una moneda de oro, ¿cuál es la probabilidad de que el segundo cajón también contenga una moneda de oro?*

Ejemplo 15. *Una compañía de seguros estima que en una cierta población, el* 30 % *de los individuos son propensos a tener accidentes. Estima además que la probabilidad de que una persona propensa a tener dicho accidente lo tenga realmente, en un periodo de un año, es igual a* 0,4*, mientras que esta probabilidad se reduce a* 0,2 *para las personas no propensas. a) ¿Cuál es la probabilidad de que un nuevo asegurado tenga un accidente durante el primer año de la póliza?. Si un asegurado tiene un accidente durante el primer año de la póliza, ¿cuál es la probabilidad de que b) esa persona sea propensa a tener un accidente? y c) el asegurado tenga un accidente durante el segundo año de la póliza?*

Solución

Definamos los eventos.

A: El asegurado es propenso a tener un accidente.

B_1: El asegurado tiene un accidente durante el primer año.

B_2: El asegurado tiene un accidente durante el segundo año.

a. $P(B_1) = P(B_1 \mid A)P(A) + P(B_1 \mid A^c)P(A^c)$

$= (0{,}4)(0{,}3) + (0{,}2)(0{,}7) = 0{,}26$

b. $P(A \mid B_1) = \frac{P(B_1|A)P(A)}{P(B_1)} = \frac{(0{,}4)((0{,}3)}{0{,}26} = 0{,}4615$

c. $P(B_2 \mid B_1) = P(B_2 \mid A)P(A \mid B_1) + P(B_2 \mid A^c)P(A^c \mid B_1)$

$= (0{,}4)(0{,}4615) + (0{,}2)(0{,}5385) = 0{,}2923$

La solución detallada de este problema ilustra el concepto de probabilidad condicional.

La población total se divide en dos partes, los propensos a tener un accidente y los no propensos. Estando dentro de la población de propensos, el 40 % tiene algún accidente durante el primer año, mientras que, estando dentro de la población de no propensos, el 20 % tiene algún accidente durante el primer año; por lo tanto se tiene:

$P(\{A \cap B_1\}) = (0{,}3)\,(0{,}4)$

$P(\{A \cap B_1^c\}) = (0{,}3)\,(0{,}6)$

$P(\{A^c \cap B_1\}) = (0{,}7)\,(0{,}2)$

$P(\{A^c \cap B_1^c\}) = (0{,}7)\,(0{,}8)$

Para el segundo año la situación es similar: Estando dentro de la población de propensos, el 40 % tiene algún accidente durante el segundo año, mientras que, estando dentro de la población de no propensos, el 20 % tiene algún accidente durante el segundo año; por lo tanto se tiene:

$P(\{A \cap B_1 \cap B_2\}) = (0{,}3)\,(0{,}4)\,(0{,}4) = 0{,}048$

$P(\{A \cap B_1 \cap B_2^c\}) = (0{,}3)\,(0{,}4)\,(0{,}6) = 0{,}072$

$P(\{A \cap B_1^c \cap B_2\}) = (0{,}3)\,(0{,}6)\,(0{,}4) = 0{,}072$

$P(\{A \cap B_1^c \cap B_2^c\}) = (0{,}3)\,(0{,}6)\,(0{,}6) = 0{,}108$

$P(\{A^c \cap B_1 \cap B_2\}) = (0{,}7)\,(0{,}2)\,(0{,}2) = 0{,}028$

$P(\{A^c \cap B_1 \cap B_2^c\}) = (0{,}7)\,(0{,}2)\,(0{,}8) = 0{,}112$

$P(\{A^c \cap B_1^c \cap B_2\}) = (0{,}7)\,(0{,}8)\,(0{,}2) = 0{,}112$

$P(\{A^c \cap B_1^c \cap B_2^c\}) = (0{,}7)\,(0{,}8)\,(0{,}8) = 0{,}448$

El espacio muestral del experimento aleatorio que define este problema puede representarse como el conjunto formado por las 8 intersecciones anteriores. Así que tenemos el espacio muestral y la probabilidad de cada uno de los posibles resultados del experimento; es decir, tenemos completo el modelo de probabilidad para este problema. Cualquier probabilidad relativa a los eventos A, B_1 y B_2 queda ya completamente determinada. Por ejemplo:

$P(B_2 \mid B_1) = \frac{P(B_1 \cap B_2)}{P(B_1)} = \frac{0{,}048+0{,}028}{0{,}048+0{,}072+0{,}028+0{,}112} = \frac{0{,}076}{0{,}26} = 0{,}292\,31$

Ejercicio 12. *Una compañía de seguros estima que en una cierta población, el* 20 % *de los individuos son propensos a tener accidentes. Estima además que la probabilidad de que una persona propensa a tener dicho accidente lo tenga realmente, en un periodo de un año, es igual a* 0,3*, mientras que esta probabilidad se reduce a* 0,15 *para las personas no propensas. Si un asegurado tiene un accidente durante el primer año de la póliza, ¿cuál es la probabilidad de que tenga un accidente durante el segundo año?*

Ejemplo 16. *Cada una de* n *bolas se coloca al azar en una de* n *cajas. Encuentra la probabilidad de que exactamente una de las cajas quede vacía.*

Solución

Definamos los eventos:

A: Únicamente la caja 1 queda vacía.

A_i: Únicamente la caja 1 queda vacía y quedan dos bolas en la caja i ($i \in \{2, \dots, n\}$).

B: Exactamente una de las cajas queda vacía.

Entonces:

$P(A_i) = \frac{\binom{n}{2}(n-2)!}{n^n}$

$P(A) = P(A_2) + \cdots + P(A_n) = \frac{(n-1)\binom{n}{2}(n-2)!}{n^n} = \frac{\binom{n}{2}(n-1)!}{n^n}$

$P(B) = nP(A) = \frac{n\binom{n}{2}(n-1)!}{n^n} = \frac{\binom{n}{2}n!}{n^n}$

Ejemplo 17. *Sabiendo que, al lanzar un dado* 4 *veces, se obtiene* 6 *en más de una ocasión, encuentra la probabilidad de obtener exactamente dos veces* 6 *en los* 4 *lanzamientos.*

Solución

Definamos los eventos:

A : Se obtiene 6 en más de una ocasión.

B : Se obtiene exactamente dos veces 6 en los 4 lanzamientos.

Se tiene:

$P(A) = 1 - \left(\frac{5}{6}\right)^4 - \binom{4}{1}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^3 = 1 - \frac{625}{1296} - \frac{500}{1296} = \frac{171}{1296}$

$P(A \cap B) = P(B) = \binom{4}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2 = \frac{150}{1296}$

Así que:

$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{150}{171} = \frac{50}{57} = 0{,}87719$

Ejercicio 13. *Una urna contiene dos bolas rojas, una azul y una blanca. Una segunda urna contiene* 2 *bolas negras y* 2 *rojas. Se eligen dos bolas al azar de la primera urna y las bolas rojas que se obtengan se agregan a la segunda urna. Después de eso, se elige una bola al azar de la segunda urna. Sabiendo que entre las bolas seleccionadas de la primera urna hay por lo menos una roja, ¿cuál es la probabilidad de que la bola seleccionada de la segunda urna sea roja?*

Ejemplo 18. *Un médico llegó a la conclusión de que la probabilidad de que su paciente, el Sr. Flores, padezca una cierta enfermedad es igual a* $0{,}65$. *El médico sigue una regla según la cual cuando la probabilidad de que un paciente padezca dicha enfermedad es* $0{,}8$ *o más entonces* *recomienda cirugía, pero si la probabilidad es menor a* 0,8 *entonces recomienda un estudio, el cual es costoso y doloroso. Por tal motivo, el médico recomienda al Sr. Flores hacerse el estudio, resultando positivo. Dicho estudio tiene la característica de que, si el resultado no es alterado por otros padecimientos del paciente, resulta positivo en todos los casos en que la persona padece la enfermedad y negativo en los casos en que no la padece. El resultado del estudio indicaba al médico que debe de recomendar cirugía, sin embargo, después de realizado el estudio, el Sr. Flores informa al médico que padece diabetes, lo cual no había contemplado el médico hasta ese momento. Esa información pone en duda al médico pues el estudio que le practicaron al Sr. Flores resulta positivo en* 40 % *de los casos en que el paciente no padece la enfermedad pero es diabético. Tomando en cuenta esta información, ¿cuál es la probabilidad de que el Sr. Flores padezca la enfermedad?*

Solución

Definamos los eventos:

A: El Sr. Flores padece la enfermedad.

B: El estudio resulta positivo.

Se tiene:

$P(A) = 0{,}65$

$P(B \mid A) = 1$

$P(B \mid A^c) = 0{,}4$

Así que:

$$P(A \mid B) = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^c)P(A^c)} = \frac{,65}{0{,}65+(0{,}4)(0{,}35)} = 0{,}82278$$

Ejemplo 19. *Una cierta fruta crece en dos regiones X y Y, las cuales en ocasiones están infectadas con una plaga que puede dañar la fruta. Es sabido que la probabilidad de que la región X esté infectada es igual a $\frac{1}{10}$, la probabilidad de que la región Y lo esté es igual a $\frac{1}{5}$ y la probabilidad de que ninguna de las dos regiones esté infectada es igual a $\frac{4}{5}$. Definamos los siguientes eventos:*

A: La región X está infectada con la plaga.

B: La región Y está infectada con la plaga.

¿Son A y B independientes?

Solución

$P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{1}{10} + \frac{1}{5} - \frac{1}{5} = \frac{1}{10}$

$P(A)P(B) = \left(\frac{1}{10}\right)\left(\frac{1}{5}\right) = \frac{1}{50}$

Por lo tanto, A y B no son independientes.

Ejercicio 14. *Una cierta fruta crece en dos regiones X y Y, las cuales en ocasiones están infectadas con una plaga que puede dañar la fruta. Es sabido que la probabilidad de que únicamente la región X esté infectada es igual a $\frac{1}{10}$, la probabilidad de que únicamente la región Y lo esté es igual a $\frac{7}{15}$ y la probabilidad de que ninguna de las dos regiones esté infectada es igual a $\frac{7}{30}$. Definamos los siguientes eventos:*

A: La región X está infectada con la plaga.

B: La región Y está infectada con la plaga.

¿Son A y B independientes?

Ejercicio 15. *En una urna hay N tarjetas numeradas del 1 al N. Se van seleccionando al azar las tarjetas de la urna, una a una y sin reemplazo, hasta que se seleccionan todas. Para $k \in \{1, \ldots, N\}$, sea A_k el evento "el número de la k-ésima tarjeta seleccionada es mayor que los números de las seleccionadas previamente". Determina si los eventos A_2 y A_3 son independientes.*

Ejemplo 20. *Dos de tres prisioneros son elegidos al azar para ser liberados. El prisionero A pide al guardia que investigue los nombres seleccionados y que le diga uno de ellos que no sea él mismo. Supongamos que el guardia acepta hacer lo que le pide el prisionero y que, en caso de que A no vaya a ser liberado, le dirá el nombre del prisionero B con probabilidad p y el del prisionero C con probabilidad $1 - p$. Consideremos entonces los siguientes eventos:*

A: El prisionero A es seleccionado para ser liberado.

B: El guardia informa al prisionero A que el prisionero B va a ser liberado.

¿Son A y B independientes?

Solución

El espacio muestral del experimento aleatorio descrito está dado por:

$$\Omega = \{(A, B), (A, C), (B, C), (C, B)\}$$

donde la pareja de letras indica los nombres de los prisioneros que son seleccionados para ser liberados y la segunda letra de la pareja indica el nombre que dice el guardia al prisionero A.

Se tiene:

$P[(A,B)] = \frac{1}{3}$

$P[(A,C)] = \frac{1}{3}$

$P[(B,C)] = \frac{1}{3}(1-p)$

$P[(C,B)] = \frac{1}{3}p$

Por lo tanto:

$P(A) = \frac{2}{3}$

$P(B) = \frac{1}{3} + \frac{1}{3}p = \frac{1}{3}(1+p)$

$P(A \cap B) = \frac{1}{3}$

$P(A)P(B) = \frac{2}{9}(1+p)$

$P(A \cap B) = P(A)P(B) \iff \frac{1}{3} = \frac{2}{9}(1+p) \iff p = \frac{1}{2}$

Así que A y B son independientes si y sólo si $p = \frac{1}{2}$.

Ejercicio 16. *(El dilema del prisionero). En una cárcel hay 3 prisioneros, A, y C, con historiales similares.En un momento dado, los tres solicitan el indulto a un tribunal. Sin conocerse más detalles, llega la información al prisionero A de que han concedido el indulto a 2 de los 3 prisioneros, eligiéndolos al azar. El prisionero A conoce a uno de los miembros del tribunal, quien le da la opción de preguntarle por uno de los nombres de los prisioneros que serán indultados, con la condición de que no le dirá si él será indultado, pero le dice también que si son B y C quienes van a ser indultados, le dirá el nombre de uno de ellos eligiéndolo al azar. Pensando un poco, A determina que si no hace la pregunta, entonces la probabilidad de ser uno de los indultados es* $\frac{2}{3}$*, mientras que si la hace, obtendrá una respuesta y entonces la probabilidad de ser él el otro indultado será* $\frac{1}{2}$*. Por ello,concluye que es mejor no hacer la pregunta, porque, si la hace, sea cual sea la respuesta, la consecuencia será que disminuirá la probabilidad de que él sea uno de los indultados.¿Son correctos los cálculos que hace el prisionero A?*

Ejemplo 21. *Se quiere seleccionar a una persona de entre* n*. Para ello, cada una lanza una moneda y se conviene en que si alguna obtiene un resultado distinto al de las otras, entonces esa persona es la seleccionada; de otra manera se vuelven a lanzar las monedas bajo las mismas condiciones. Para* $k \in \mathbb{N}$*, encuentra la probabilidad de que el número de lanzamientos (de las* n *personas) que se realizan hasta que alguna persona es seleccionada es igual a* k*.*

Solución

Llamando éxito a la obtención de cara al lanzar una moneda, cada lanzamiento de las n monedas representa la realización de n ensayos de Bernoulli con probabilidad de éxito igual a $\frac{1}{2}$. La persona es seleccionada en un lanzamiento únicamente si se obtiene exactamente un éxito o bien exactamente un fracaso, cuya probabilidad está dada por $2n\left(\frac{1}{2}\right)^n$. Así que los diferentes intentos de seleccionar a la persona constituyen una sucesión de ensayos de Bernoulli con probabilidad de éxito $p = 2n\left(\frac{1}{2}\right)^n$.

Para cada $k \in \mathbb{N}$, definamos:

A_k : El número de lanzamientos que se realizan hasta que alguna persona es seleccionada es igual a k.

Entonces:

$$P(A_k) = \left[1 - 2n\left(\tfrac{1}{2}\right)^n\right]^{k-1} 2n\left(\tfrac{1}{2}\right)^n = \tfrac{n}{2^{n-1}}\left(1 - \tfrac{n}{2^{n-1}}\right)^{k-1}$$

Ejercicio 17. *Tres caballos A, B y C participan en una carrera. El evento "A vence a B" se designa por AB, "A vence a C" por AC, "B vence a C" por BC, "A vence a B y B vence a C" por ABC, y así sucesivamente. Se sabe que $P(AB) = \frac{2}{3}$, $P(AC) = \frac{2}{3}$ y $P(BC) = \frac{1}{2}$. Además $P(ABC) = P(ACB)$ y $P(CAB) = P(CBA)$. Finalmente, se sabe que la probabilidad de que haya un empate entre cualquier par de ellos (o entre los 3) es igual a cero y que alguno de los caballos gana la carrera.*

a) ¿Cuál es la probabilidad que cada caballo tiene de ganar la carrera?

b) ¿Son AB, AC y BC independientes?

Ejercicio 18. *Se lanza un dado consecutivamente hasta obtener el número* 6 *por segunda ocasión. Considera como resultado del experimento a la sucesión de números que se obtienen hasta que se obtiene el número* 6 *por segunda ocasión.*

Para $k, n \in \mathbb{N}$, definamos los siguientes eventos:

A_k : El segundo 6 se obtiene en el k-ésimo lanzamiento.

B_n : El segundo 6 se obtiene en alguno de los primeros n lanzamientos.

a) Calcula $P(A_k)$ para cualquier $k \in \mathbb{N}$.

b) Demuestra que $\sum_{k=1}^{\infty} P(A_k) = 1$.

c) Demuestra que, para cualquier $n \in \mathbb{N}$:

$$P(B_n) = 1 - \tfrac{1}{5}n\left(\tfrac{5}{6}\right)^n - \left(\tfrac{5}{6}\right)^n$$