NOCIONES DE PROBABILIDAD

FELIPE OSORIO

El objetivo de este capítulo es presentar algunos elementos básicos de teoría de probabilidades y que permitirá el desarrollo de los capítulos posteriores. El material es introductorio y mayores detalles pueden ser hallados, por ejemplo, en los libros de Casella y Berger (2002), Chung (1974) y Hoel et al. (1971). Iniciaremos la presentación de este capítulo recordando algunas definiciones de sumas y productorios, así como teoría de conjuntos y elementos de análisis combinatorio (Charalambides, 2002), también conocido como técnicas de conteo. Esta instancia ofrece una oportunidad para incorporar ejemplos utilizando el ambiente para cálculo estadístico R (R Core Team, 2025) y posteriormente se presenta la definición axiomática de probabilidad de Kolmogorov así como los conceptos de probabilidad condicional e indepencia. Ambos conceptos claves para el análisis de datos.

1. Elementos de teoría de conjuntos

Un conjunto es una colección bastante general de objetos o números, que en este contexto serán llamados elementos. Indicaremos a un conjunto por letras mayúsculas. Decimos que a es un elemento del conjunto A, o bien, que a pertenece a A y escribimos $a \in A$, en caso contrario escribimos $a \notin A$. Para los objetos contenidos en A usaremos la notación

$$A = \{a_1, a_2, \dots, a_n, \dots\},$$

es decir es necesario listar o conocer los elementos de A, mientras que

$$A = \{a : a \text{ tiene la propiedad } P\},$$

donde P es una característica que define los elementos de A.

Ejemplo 1.1. Considere los siguientes conjuntos:

$$A = \{1, 2, \dots, n\}, \qquad B = \{2, 4, 6, 8, \dots\}, \qquad C = \{x : 0 < x < 1\}.$$

Tambien existe conjuntos bastante comunes, tales como:

$$\begin{aligned}
&\mathbb{R} \text{ es el conjunto de todos los números reales,}\\
&[a, b] = \{x : a \leq x \leq b\} \text{ es un intervalo cerrado en } \mathbb{R},\\
&(a, b) = \{x : a < x < b\} \text{ es un intervalo abierto en } \mathbb{R},\\
&(a, b] = \{x : a < x \leq b\} \text{ es un intervalo semi-cerrado en } \mathbb{R}.
\end{aligned}$$

Conjuntos pueden ser de naturaleza muy diversa, en el siguiente ejemplo se considera una colección de funciones.

Ejemplo 1.2. Sea $f : \mathbb{R} \to \mathbb{R}$ y considere

$$A = \{f : f(x) = ax + b, a, b \in \mathbb{R}\},$$

denota el conjunto de todas las rectas en el plano.

Un conjunto también puede representar una colección de conjuntos. Sea A y B dos conjuntos, entonces $\mathcal{C} = \{A, B\}$ es un conjunto (o colección) de conjuntos.

Ejemplo 1.3. Sea $A = \{0, 1\}$ y $B = \{1, 2, \dots, n\}$. Entonces,

$$\mathcal{C} = \{A, B\} = \{\{0, 1\}, \{1, 2, \dots, n\}\},$$

es decir $\mathcal{C}$ contiene dos elementos. El conjunto $\{0, 1\}$ y el conjunto $\{1, 2, \dots, n\}$. Note que $\{\{0, 1\}, \{1, 2, \dots, n\}\} \neq \{0, 1, 2, \dots, n\}$. El primero es un conjunto con 2 elementos, mientras el último es un conjunto con $n + 1$ elementos.

Dos conjuntos son de particular interés, el *conjunto universal*, es decir aquél conjunto que contiene todos los objetos bajo consideración, y que será denotado por Ω, mientras que el *conjunto vacio* o *nulo*, denotado por $\varnothing$ corresponde al conjunto que no contiene elementos.

Ejemplo 1.4. Considere $\Omega = \mathbb{R}$ el conjunto de números reales y sea,

$$A = \{x \in \mathbb{R} : x^2 - 2x + 2 = 0\}.$$

De este modo, $A = \varnothing$, pues la ecuación cuadrática $x^2 - 2x + 2 = 0$ no tiene raíces reales.

Un conjunto B es llamado *subconjunto* de A si y sólo si para todo $x \in B$, entonces $x \in A$, en cuyo caso anotamos $B \subseteq A$. Si $B \subseteq A$ y existe $x \in A$ tal que $x \notin B$, entonces se dice que B es un *subconjunto propio* de A y anotamos $B \subset A$. Cuando $B \subseteq A$ y $A \subseteq B$ entonces los conjuntos A y B consisten de los mismos elementos, y en este caso $A = B$.

Ejemplo 1.5. Suponga que $A = \{1, 2, 3, 4, 5\}$ y $B = \{1, 3, 5\}$, en este caso $B \subset A$. Mientras que si $A = \{0, 1\}$ y $B = \{0, 1\}$, entonces $A = B$. Por otro lado, considere $A = \{x : x \geq 0\}$ y $B = \{x : x > 1\}$, luego $B \subset A$.

Evidentemente, tenemos que $\varnothing \subseteq A$ para todo A, mientras que $A \subseteq \Omega$. Más aún $\varnothing \subset \Omega$ y $\Omega \subseteq \Omega$. Además, Ω es un concepto relativo. Por ejemplo, podríamos definir $\Omega = [0, 1]$ y considerar todos los subintervalos en $[0, 1]$.

Un par de elementos a y b (no necesariamente diferentes) donde a es el primer elemento mientras que b es el segundo es llamado *par ordenado* y escribimos (a, b). En general una n-upla ordenada es $(a_1, a_2, \dots, a_n)$ y es usual llamar a a_i la i-ésima coordanada de la n-upla.

El *producto cartesiano* de los conjuntos A y B denotado por $A \times B$ es definido como

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

Esta definición puede ser extendida a n conjuntos $A_1, A_2, \dots, A_n$ como

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \dots, a_n) : a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n\}.$$

En particular, si $A_1 = A_2 = \cdots = A_n = A$, entonces el producto cartesiano es denotado por A^n. Por ejemplo $\mathbb{R}^2$ es el conjunto de todos los pares ordenados de números reales.

La *unión* de dos conjuntos A y B consiste de todos los elementos en A o en B, es decir

$$A \cup B = \{x : x \in A, \ \text{o } x \in B\}.$$

Esta definición se extiende a n conjuntos $A_1, \ldots, A_n$ como

$$A_1 \cup A_2 \cup \cdots \cup A_n = \{x \in \Omega : x \in A_k \text{ para al menos un } k \in \{1, 2, \ldots, n\}\},$$

y más generalmente para una familia de conjuntos $\{A_i, i \in I\}$, con I un conjunto de índices,

$$\bigcup_{i \in I} A_i = \{x \in \Omega : x \in A_i \text{ para al menos un subíndice } i \in I\}.$$

Evidentemente esta última expresión puede ser usada para una secuencia infinita de conjuntos $A_1, A_2, \ldots$, es decir $\bigcup_{n=1}^{\infty} A_n$, la cual debe satisfacer

$$x \in A_1 \text{ o } x \in A_2 \text{ o } \ldots \quad \Longleftrightarrow \quad x \text{ está en al menos uno de } A_1, A_2, \ldots.$$

Ejemplo 1.6. Sea $A = \{1, 2\}$ y $B = \{1, 5\}$, entonces $A \cup B = \{1, 2, 5\}$. Ahora, si $A = (0, 1]$ y $B = (\frac{1}{2}, \infty)$, entonces $A \cup B = (0, \infty)$.

La *intersección* de dos conjuntos A y B es el conjunto que incluye los elementos comunes a ambos conjuntos y es denotado por $A \cap B$, esto es

$$A \cap B = \{x : x \in A, \text{ y } x \in B\}.$$

Esta definición se extiende a n conjuntos $A_1, \ldots, A_n$ como

$$A_1 \cap A_2 \cap \cdots \cap A_n = \{x \in \Omega : x \in A_k \text{ para todos los subíndices } k \in \{1, 2, \ldots, n\}\},$$

y más generalmente para una familia de conjuntos $\{A_i, i \in I\}$, como:

$$\bigcap_{i \in I} A_i = \{x \in \Omega : x \in A_i \text{ para todos los subíndices } i \in I\}.$$

Para una secuencia infinita de conjuntos $A_1, A_2, \ldots$, la intersección $\cap_{n=1}^{\infty} A_n$ debe satisfacer

$$x \in A_1 \text{ y } x \in A_2 \text{ y } \ldots \quad \Longleftrightarrow \quad x \text{ pertenece a todo } A_1, A_2, \ldots.$$

Ejemplo 1.7. Para $A = \{1, 2\}$ y $B = \{1, 5\}$, entonces $A \cap B = \{1\}$. Mientras que, si $A = (0, 1]$ y $B = (\frac{1}{2}, \infty)$, entonces $A \cap B = (\frac{1}{2}, 1]$.

El *complemento* (con respecto al conjunto universal Ω) de un conjunto A es definido como:

$$A^c = \{x \in \Omega : x \notin A\}.$$

Ejemplo 1.8. Suponga $\Omega = [0, 1]$ y $A = [0, \frac{1}{2})$, luego $A^c = [\frac{1}{2}, 1]$. Por otro lado, sea $\Omega = \mathbb{R}$ y $A = \mathbb{R}$, entonces $A^c = \varnothing$.

Además, es fácil notar que

$$(A^c)^c = A, \qquad \Omega^c = \varnothing, \qquad \varnothing^c = \Omega.$$

La *diferencia* del conjunto B con el conjunto A es definida como el conjunto de los elementos de A que no pertenecen a B, esto es,

$$A - B = \{x \in \Omega : x \in A, x \notin B\}.$$

Podemos apreciar que,

$$A - B = A \cap B^c, \qquad A^c = \Omega - A.$$

Propiedad 1.9. Sean A, B y C conjuntos definidos en Ω. Tenemos que,

(i) *Asociatividad de la unión e intersección*:

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

(ii) *Distributividad de la intersección con respecto a la unión y de la unión con respecto a la intersección*

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

(iii) *Conmutatividad de la unión e intersección*

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A.$$

(iv) *El conjunto vacío* $\varnothing$ *es el elemento neutro para la unión*

$$A \cup \varnothing = \varnothing \cup A = A.$$

(v) *El conjunto universal* Ω *es el elemento neutro para la intersección*

$$A \cap \Omega = \Omega \cap A = A.$$

(vi) *El complemento satisface*

$$A \cap A^c = \varnothing, \qquad A \cup A^c = \Omega.$$

(vii) $A \cup A = A$, $A \cap A = A$, $A \cup \Omega = \Omega$, $A \cap \varnothing = \varnothing$.

(viii) *Leyes de De Morgan*

$$(A \cup B)^c = A^c \cap B^c, \qquad (A \cap B)^c = A^c \cup B^c.$$

Las fórmulas de De Morgan pueden ser extendidas para n conjuntos $A_1, \dots, A_n$, como:

$$(A_1 \cup A_2 \cup \cdots \cup A_n)^c = A_1^c \cap A_2^c \cap \cdots \cap A_n^c,$$
$$(A_1 \cap A_2 \cap \cdots \cap A_n)^c = A_1^c \cup A_2^c \cup \cdots \cup A_n^c,$$

y para una familia de conjuntos $\{A_i, i \in I\}$, como:

$$\Big(\bigcup_{i \in I} A_i\Big)^c = \bigcap_{i \in I} A_i^c, \qquad \Big(\bigcap_{i \in I} A_i\Big)^c = \bigcup_{i \in I} A_i^c.$$

Dos conjuntos A y B se dicen *disjuntos* si ellos no tienen elementos en común, y escribimos $A \cap B = \varnothing$, Más generalmente, los conjuntos $A_1, \dots, A_n$ se dicen *disjuntos por pares* (o mutuamente excluyentes) si

$$A_i \cap A_j = \varnothing, \qquad i \neq j,$$

con $\{i, j\}$ desde el conjunto de índices $\{1, 2, \dots, n\}$.

Ejemplo 1.10. Sea $A = \{x : x > 1\}$ y $B = \{x : x < 0\}$. Entonces A y B son disjuntos.

Ejemplo 1.11. Sea A y B dos conjuntos en Ω. Entonces,

$$(A - B) \cap (B - A) = (A \cap B^c) \cap (B \cap A^c) = (A \cap A^c) \cap (B \cap B^c) = \varnothing,$$

es decir $A - B$ y $B - A$ son disjuntos.

Definición 1.12. Una colección de conjuntos $\{A_1, \dots, A_n\}$ es una *partición* del conjunto A, si satisface las condiciones:

(i) $\{A_1, \dots, A_n\}$ son disjuntos por pares.
(ii) $\bigcup_{i=1}^n A_i = A$.

Propiedad 1.13. Suponga que $\{A_1, \dots, A_n\}$ es una partición de A, y suponga $B \subseteq A$. Entonces $\{B \cap A_1, \dots, B \cap A_n\}$ es una partición de $B \cap A$. En efecto,

$$(B \cap A_i) \cap (B \cap A_j) = B \cap (A_i \cap A_j) = \varnothing, \qquad \forall\, i \neq j,$$

y

$$\bigcup_{i=1}^n (B \cap A_i) = B \cap \Big(\bigcup_{i=1}^n A_i\Big) = B \cap A.$$

Ejemplo 1.14. La familia de conjuntos $A_i = [i, i+1)$, para $i = 0, 1, 2, \dots$ forman una partición de $[0, \infty)$.

Ejemplo 1.15. A y A^c son una partición de Ω, pues

$$A \cap A^c = \varnothing, \qquad A \cup A^c = \Omega.$$

Definición 1.16. La secuencia de conjuntos $\{A_n\}_{n \geq 1}$, se dice monótona, si:

(i) $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$. Es decir, $\{A_n\}$ es creciente ($A_n \subseteq A_{n+1}$, $\forall\, n \in \mathbb{N}$) y anotamos $A_n \uparrow$.
(ii) $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$. Es decir, $\{A_n\}$ es decreciente ($A_n \supseteq A_{n+1}$, $\forall\, n \in \mathbb{N}$) y anotamos $A_n \downarrow$.

Definición 1.17. El límite de una secuencia monótona se define como:

(i) Si $A_n \uparrow$, entonces $\lim_{n\to\infty} = \bigcup_{n=1}^\infty A_n$.
(ii) Si $A_n \downarrow$, entonces $\lim_{n\to\infty} = \bigcap_{n=1}^\infty A_n$.

2. Técnicas de conteo

El análisis combinatorio es el estudio sistemático del conteo de todas las posibles agrupaciones desde un conjunto de objetos. Esta tarea es equivalente contar los elementos de un conjunto finito. A continuación presentaremos ideas preliminares para posteriormente introducir diversas técnicas de conteo.

Los conjuntos A y B se dicen *equivalentes* y anotamos $A \sim B$ si y solo si existe una función biyectiva de A a B.

Ejemplo 2.1. El conjunto $A = \{a_1, a_2, \dots, a_n\}$ es equivalente al conjunto $B = \{1, 2, \dots, n\}$ con $f(a_k) = k$ para todo $k = 1, 2, \dots, n$, es una función biyectiva desde A a B.

Un conjunto A es llamado *finito*, con n elementos si y solo si es equivalente al conjunto $\{1, 2, \dots, n\}$. Asumiremos que $\varnothing$ es finito con cero elementos. Un conjunto que no es finito de dice *infinito*, mientras que un conjunto A se denomina *infinito numerable* si y solo si es equivalente al conjunto de números naturales $\mathbb{N} = \{1, 2, \dots\}$. Un conjunto A es llamado *numerable* (o contable) si es finito o infinito numerable, en caso contrario se dice *no numerable*.

El número de elementos de un conjunto finito A, denotado por $\mathsf{N}(A)$, es llamado *cardinal* de A. En caso que Ω sea un conjunto finito, su cardinalidad será denotada como $\mathsf{N}(\Omega) = N$.

Los principios fundamentales del conteo corresponden al *Principio de multiplicación* y el *Principio de adición*. En lo que sigue consideraremos conjuntos finitos. A partir de las definiciones anteriores podemos destacar el siguiente lema.

Lema 2.2. *Si A y B son conjuntos finitos y equivalentes, entonces*
$$\mathsf{N}(A) = \mathsf{N}(B).$$

De este modo, el cardinal de un conjunto finito A puede ser determinado desde un conjunto finito B, equivalente a A, cuya cardinalidad sea conocida. A continuación, revisamos algunas propiedades básicas del cardinal, que se pueden reestablecer como los principios del conteo.

Resultado 2.3. *Si A y B son conjuntos finitos, entonces*
$$\mathsf{N}(A \times B) = \mathsf{N}(A)\,\mathsf{N}(B).$$

La cardinalidad del producto Cartesiano de más que dos conjuntos finitos se deduce desde el resultado anterior, en efecto considere el resultado siguiente.

Corolario 2.4. *Si $A_1, A_2, \dots, A_k$ son conjuntos finitos, entonces*
$$\mathsf{N}(A_1 \times A_2 \times \cdots \times A_k) = \mathsf{N}(A_1)\,\mathsf{N}(A_2)\cdots\mathsf{N}(A_k).$$

Estos resultados son conocidos frecuentemente como el principio de multiplicación, el cual se puede exponer como sigue:

Resultado 2.5 (Principio de Multiplicación)**.** *Suponga que el conjunto A_1 contiene n_1 elementos, el conjunto A_2 contiene n_2 elementos, ..., y el conjunto A_k contiene n_k elementos. Entonces el número de maneras de escoger un objeto desde cada uno de los k conjuntos es:*
$$n_1 \cdot n_2 \cdots n_k.$$

Ejemplo 2.6. En cierto medioambiente, existe 14 especies de mosca de la fruta, 17 especies de polillas y 13 especies de mosquitos. Se desea determinar el número de formas en que se puede escoger una especie de cada tipo. En efecto, tenemos $n_1 = 14$, $n_2 = 17$ y $n_3 = 13$. De este modo, por el principio de multiplicación sigue que existe $14 \cdot 17 \cdot 13 = 3094$ maneras diferentes de seleccionar una especie de cada tipo.

Ahora, considere el siguiente resultado,

Resultado 2.7. *Si A y B son conjuntos finitos y disjuntos, entonces*
$$\mathsf{N}(A \cup B) = \mathsf{N}(A) + \mathsf{N}(B).$$

Basado en el resultado anterior, se puede mostrar que para A y B subconjuntos de un conjunto universal Ω, entonces:
$$\mathsf{N}(A^c) = \mathsf{N}(\Omega) - \mathsf{N}(A),$$
mientras que,
$$\mathsf{N}(A - B) = \mathsf{N}(A) - \mathsf{N}(A \cap B),$$
y particularmente para $B \subseteq A$,
$$\mathsf{N}(A - B) = \mathsf{N}(A) - \mathsf{N}(B).$$

El siguiente resultado se conoce como principio de adición y corresponde a una extensión del Resultado 2.7.

Corolario 2.8. *Si $A_1, A_2, \dots, A_k$ son conjuntos finitos y mutuamente excluyentes, entonces*
$$\mathsf{N}(A_1 \cup A_2 \cup \cdots \cup A_k) = \mathsf{N}(A_1) + \mathsf{N}(A_2) + \cdots + \mathsf{N}(A_k).$$

Esto permite establecer el siguiente principio.

Resultado 2.9 (Principio de Adición). *Si un elemento ω_i puede ser seleccionado en n_i maneras diferentes para $i = 1, 2, \dots, k$ y la selección de ω_i excluye la selección simultánea de ω_j, para $i, j = 1, 2, \dots, k$, $i \neq j$. Entonces cualquiera de los elementos ω_1 o ω_2 o $\dots$ o ω_k, puede ser seleccionado en*

$$n_1 + n_2 + \cdots + n_k,$$

maneras.

Ejemplo 2.10. Considere el siguiente experimento: *Lanzar un dado o una moneda.* De este modo, los resultados posibles del experimento son $6 + 2 = 8$.

En análisis combinatorio debemos distinguir entre conjuntos *ordenados* y *desordenados.* En un conjunto ordenado, el orden es relevante, mientras que no lo será para un conjunto desordenado. Para ilustrar, considere el siguiente ejemplo.

Ejemplo 2.11. Sea $A = \{1, 2, 3\}$, el listado de todos los subconjuntos ordenados de tamaño 2, corresponde a

$$(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2),$$

mientras que la lista de todos los subconjuntos desordenados de tamaño 2 consiste en:

$$\{1, 2\}, \{1, 3\}, \{2, 3\}.$$

Note que en este caso el conjunto $\{2, 1\}$ es idéntico al conjunto $\{1, 2\}$, de modo que no se incluye en el listado.

Definición 2.12. Sea A un conjunto finito con n elementos. Una k-upla ordenada $(a_1, a_2, \dots, a_k)$ con $a_r \in A$, $r = 1, 2, \dots, k$, es llamada una k-permutación del conjunto A, o bien una k-permutación de n.

Ejemplo 2.13. Para el conjunto $A = \{1, 2, 3\}$ de 3 elementos tenemos que las 2-permutaciones son las siguientes:

$$(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2).$$

Note que, cualquier permutación (a_i, a_j) del conjunto $A = \{a_1, a_2, a_3\}$ puede ser construída como sigue:

- Seleccionamos el primer elemento a_i desde el conjunto A con 3 elementos, y
- Seleccionamos el segundo elemento a_j, el cual debe ser diferente de a_i, desde el conjunto $B = A - \{a_i\}$ de 2 elementos.

De este modo, usando el principio de multiplicación, el número de 2-permutaciones de 3 es $3 \cdot 2 = 6$. En efecto, $6 = \mathsf{N}(C)$, donde

$$C = \{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\}.$$

Definición 2.14. El número de k-permutaciones de n, denotadas por $P_{n,k}$ o bien $(n)_k$, es dada por

$$P_{n,k} := (n)_k = n(n-1)\cdots(n-k+1).$$

Ejemplo 2.15. Suponga que se desea calcular cuantas “palabras”de cuatro letras pueden ser formadas (no necesariamente en español) a partir de la palabra “válido”. Note que lo que se desea es obtener el número de todas las 4-permutaciones de 6, es decir:

$$P_{6,4} = 6 \cdot 5 \cdot 4 \cdot 3 = 360.$$

Para el caso particular en que $k = n$, tenemos que el número de permutaciones de n, es dada por

$$P_{n,n} = n(n-1)\cdots 2 \cdot 1.$$

Definición 2.16. El producto de todos los enteros desde 1 a n, es llamado n *factorial* y se denota como $n!$, es decir

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-1)n = \prod_{j=1}^{n} j.$$

Ejemplo 2.17. ¿De cuantas formas diferentes podemos ubicar 5 personas en 5 asientos?

$$(5)_5 = 5! = 120.$$

Note que podemos escribir el número $P_{n,k} = (n)_k$ como:

$$\begin{aligned} P_{n,k} &= n(n-1)\cdots(n-(k-1)) \\ &= \frac{n(n-1)\cdots(n-(k-1))(n-k)(n-(k+1))\cdots 2\cdot 1}{(n-k)(n-(k+1))\cdots 2\cdot 1} = \frac{n!}{(n-k)!}, \end{aligned}$$

para $k = 1, 2, \dots, n$, y $n = 1, 2, \dots$. Asumiremos que $0! = 1$, lo que permite escribir

$$\begin{aligned} &P_{n,0} = (n)_0 = 1, \qquad n = 0, 1, \dots, \\ &P_{0,0} = 1, \end{aligned}$$

y además,

$$P_{n,n-1} = P_{n,n} = n!,$$

mientras que, para $k > n$, $P_{n,k} = (n)_k = 0$. Es fácil notar que el factorial $n!$ crece rápidamente conforme n crece. Por ejemplo,

$$2! = 2, \quad 3! = 6, \quad 4! = 24, \quad 5! = 120, \quad 6! = 720, \quad 7! = 5040, \dots$$

Una aproximación útil es dada a continuación.

Observación 2.18 (Fórmula de Stirling)*.* Si n es un entero positivo,

$$n! \approx e^{-n} n^{n+1/2} \sqrt{2\pi}.$$

En efecto, es posible mostrar que,

$$\lim_{n\to\infty} \frac{n!}{e^{-n} n^{n+1/2}} = \sqrt{2\pi}.$$

Resultado 2.19. *El número $n!$ de permutaciones de n, satisface la relación de recurrencia:*

$$n! = n(n-1)!, \qquad n = 1, 2, \dots,$$

con condición inicial $0! = 1$.

Ejemplo 2.20. En R se dispone de las funciones `factorial` y `lfactorial` para el cálculo del factorial y su logaritmo, respectivamente. Adicionalmente, podemos llevar a cabo el cálculo de $P_{n,k}$ mediante la función `prod`, por ejemplo:

```
# 4-permutaciones de 6
> prod(6:3)
[1] 360
```

```
# cálculo usando factoriales
> factorial(6) / factorial(6 - 4)
[1] 360
```

Por ejemplo, podemos escribir una función en R para evaluar la formula de Stirling, como:

```
# fórmula de Stirling
> stirling <- function(n) {
+   val <- sqrt(2 * pi) * exp(-n) * n^(n + .5)
+   exact <- factorial(n)
+   rel <- abs(val - exact) / exact
+   attr(val, "error") <- rel
+   val
}

# aproximación de 5!
> stirling(5)
[1] 118.0192
attr(,"rel")
[1] 0.016507
```

En algunos casos, no todos los objetos que se desea permutar pueden ser distinguidos. Por ejemplo, existe $3! = 6$ permutaciones de las 3 letras AB$\underline{\text{B}}$, a saber:

$$\text{AB}\underline{\text{B}},\ \text{A}\underline{\text{B}}\text{B},\ \text{BA}\underline{\text{B}},\ \text{B}\underline{\text{B}}\text{A},\ \underline{\text{B}}\text{AB},\ \underline{\text{B}}\text{BA}.$$

Sin embargo, las dos B's *no son distinguibles*. En efecto, las únicas permutaciones distinguibles son ABB, BAB y BBA. De este modo, los distintos ordenamientos son

$$\frac{3!}{2!\,1!} = \frac{6}{2} = 3,$$

pues disponemos un total de 3 letras, dos de ellas B y una A.

Para generalizar el concepto anterior, suponga un conjunto de n elementos particionado en r subconjuntos conteniendo $k_1, k_2, \dots, k_r$ elementos con $n = k_1 + k_2 + \cdots + k_r$. El número de permutaciones de n objetos que incluye $k_1, k_2, \dots, k_r$ objetos indistinguibles, es dado en el siguiente resultado.

Resultado 2.21. *El número de permutaciones de n tipos de elementos con $k_1, k_2, \dots, k_r$ elementos, respectivamente, es dado por*

$$M_{k_1,k_2,\dots,k_r} = \frac{n!}{k_1!k_2!\cdots k_r!}, \qquad n = k_1 + k_2 + \cdots + k_r.$$

Observación 2.22. Es frecuente escribir $M_{k_1,k_2,\dots,k_r}$ usando el *coeficiente multinomial*,

$$\binom{n}{k_1, k_2, \dots, k_r} = \frac{n!}{k_1!k_2!\cdots k_r!}.$$

Ejemplo 2.23. Calcular el número de permutaciones distintas con las cifras $\{4, 7, 3, 4, 7, 7, 3\}$. De este modo,

$$\binom{7}{2, 2, 3} = \frac{7!}{2!\,2!\,3!} = \frac{5040}{2 \cdot 2 \cdot 6} = 210.$$

Para motivar ideas considere el siguiente ejemplo,

Ejemplo 2.24. Suponga que se lanza 3 veces una moneda y registramos si el resulta es cara (C) o sello (S). De este modo, obtenemos:

$$(C,C,C),(C,C,S),(C,S,C),(S,C,C),(C,S,S),(S,C,S),(S,S,C),(S,S,S).$$

Es decir, tenemos $2^3 = 8$ permutaciones con repetición.

El siguiente resultado presenta el número de permutaciones con repeticiones.

Resultado 2.25. *El número de k-permutaciones de n con repetición, denotada como $U_{n,k}$ es dada por:*

$$U_{n,k} = n^k.$$

Ejemplo 2.26. Suponga el conjunto $A = \{1,2,3\}$ y considere todas las 2-permutaciones de A con repetición

$$(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3).$$

Es decir, el total del 2-permutaciones de 3 cuando las repeticiones son permitidas es dada por

$$U_{3,2} = 3^2 = 9.$$

Ejemplo 2.27. Considere un grupo de k personas. ¿Cuántas listas posibles se pueden realizar con los días de sus cumpleaños? En efecto, tenemos

$$\underbrace{365 \cdot 365 \cdots 365}_{k \text{ veces}} = 365^k.$$

Definición 2.28. Sea $A = \{a_1, a_2, \dots, a_n\}$ un conjunto finito de n elementos. Una colección (desordenada) de k elementos $\{a_1, a_2, \dots, a_r\}$ con $a_r \in A$, $r = 1, 2, \dots, k$ es llamada una k-combinación del conjunto A, o bien una k-combinación de n.

Ejemplo 2.29. Considere el conjunto $A = \{1,2,3,4\}$, las 2-combinaciones de A son las siguientes:

$$\{1,2\},\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}.$$

En efecto, $\{1,2\}$ y $\{2,1\}$ corresponden al mismo conjunto y sólo difieren en el orden. Mientras que, las 3-combinaciones de A corresponden a

$$\{1,2,3\},\{1,2,4\},\{1,3,4\},\{2,3,4\}.$$

En R la función `combn` permite listar todas las combinaciones posibles tomadas desde un conjunto con n elementos. Para el conjunto A, podemos listar las 2-combinaciones y 3-combinaciones, como

```
# conjunto de índices
> a <- 1:4
> a
[1] 1 2 3 4

# 2-combinaciones de 4
> combn(a, m = 2)
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1    1    2    2    3
[2,]    2    3    4    3    4    4
```

```
# 3-combinaciones de 4
> combn(a, m = 3)
     [,1] [,2] [,3] [,4]
[1,]    1    1    1    2
[2,]    2    2    3    3
[3,]    3    4    4    4
```

Resultado 2.30. *El número de k-combinaciones de n, denotada por $C_{n,k}$ o bien $\binom{n}{k}$ es dada por*

$$C_{n,k} := \binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}.$$

Ejemplo 2.31. Un comité está conformado por 12 miembros y el quorum mínimo para su funcionamiento es de 8 miembros. Deseamos saber en cuantas maneras podemos constituir el comité de manera tal que tengamos el quorum mínimo. De ahí que, existe

$$\binom{12}{8} = \frac{12!}{8!\,4!} = \frac{8!\cdot 9\cdot 10\cdot 11\cdot 12}{8!\,4!} = \frac{11880}{24} = 495,$$

maneras de escoger 8 miembros. Por otro lado, suponga que deseamos conocer en cuantas formas podemos tener quorum. Como tenemos quorum cuando tenemos 8, 9, 10, 11 o 12 miembros, el número de maneras en que se tiene quorum es:

$$\binom{12}{8} + \binom{12}{9} + \binom{12}{10} + \binom{12}{11} + \binom{12}{12} = 495 + 220 + 66 + 12 + 1 = 794,$$

maneras.

En R se encuentra disponible la función `choose` para el cálculo de combinatorios, es decir podemos obtener los resultados anteriores usando:

```
# quorum mínimo
> choose(12, 8)
[1] 495

# quorum, i.e. 8,9,10,11 o 12 miembros
> accum <- 0
> for (i in 8:12) {
+  accum <- accum + choose(12, i)
+ }
> accum
[1] 794
```

El número $C_{n,k} = \binom{n}{k}$ es conocido como *coeficiente binomial*, y satisface la siguiente relación

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

A continuación se presenta una propiedad muy relevante del coeficiente binomial.

Resultado 2.32 (Triángulo de Pascal)**.** *El número $C_{n,k} = \binom{n}{k}$, de k-combinaciones de n, satisface la relación de recurrencia*

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1},$$

para $k = 1, 2, \dots,\ n = 1, 2, \dots,$ con condiciones iniciales

$$\binom{n}{0} = 1, \qquad n = 0, 1, \dots,$$
$$\binom{n}{k} = 0, \qquad k > n.$$

Corolario 2.33. *El número $C_{n,k} = \binom{n}{k}$, de k-combinaciones de n, satisface la relación de recurrencia*

$$\binom{n}{k} = \sum_{r=k}^{n} \binom{r-1}{k-1}, \qquad k = 1, 2, \dots, n, n = 1, 2, \dots.$$

Las definiciones anteriores permiten escribir el siguiente resultado.

Resultado 2.34 (Fórmula del binomio de Newton)**.** *Sea a, b dos número reales y n un entero positivo. Entonces*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Ejemplo 2.35. Un caso especial del teorema del binomio es

$$\begin{aligned} 2^n = (1+1)^n &= \sum_{k=0}^{n} \binom{n}{k} \\ &= \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n-1} + \binom{n}{n}. \end{aligned}$$

Esta fórmula es interesante pues representa el número de todos los subconjuntos posibles que se pueden formar desde un conjunto con n elementos.

3. Fundamentos de Probabilidad

El objetivo de esta sección es caracterizar el concepto de *medir un conjunto*. A partir de este momento consideraremos que es posible describir, usando elementos de teoría de conjuntos, el resultado de un experimento aleatorio. En efecto, considere la siguiente definición.

Definición 3.1. El conjunto Ω, de todos los resultados posibles de un experimento aleatorio es llamado *espacio muestral*.

Ejemplo 3.2. Considere los siguientes experimentos aleatorios.

(a) Lanzar una moneda, en cuyo caso $\Omega = \{C, S\}$.
(b) Lanzar un dado. De este modo, $\Omega = \{1, 2, 3, 4, 5, 6\}$.
(c) Jugar al "cachipún"(piedra/papel/tijeras), así $\Omega = \{$papel, piedra,tijeras$\}$.
(d) Duración de un artículo eléctrico. En este caso $\Omega = [0, \infty)$.

Definición 3.3. Un evento (o suceso) es cualquier colección de resultados posibles de un experimento aleatorio, esto es, cualquier subconjunto de Ω (incluyendo al propio Ω).

Sea $A \subseteq \Omega$, diremos que A ocurre si $\omega \in A$ con $\omega \in \Omega$ un resultado asociado a un experimento aleatorio. Evidentemente, $\omega \notin A$ si y sólo si A no ocurre.

Adicionalmente necesitamos definir una familia de conjuntos, tal que todo evento $A \subseteq \Omega$, pertenezca al *espacio de eventos*, denotado por $\mathcal{F}$. Es decir, $\mathcal{F}$ es conjunto de todos los subconjuntos posibles de Ω. Esto es requerido pues, para todo evento $A \in \mathcal{F}$ deseamos asociar un número entre cero y uno llamado probabilidad de A. Considere las siguientes definiciones.

Definición 3.4. Una colección de subconjuntos de Ω es llamado *σ-álgebra* y es denotada por $\mathcal{F}$ si satisface las propiedades:

(a) $\varnothing \in \mathcal{F}$.
(b) Si $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$.
(c) Si $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

Note que $\varnothing \subset \Omega$ y $\Omega = \varnothing^c$, así por la Propiedad (a) y (b) sigue que $\Omega \in \mathcal{F}$. Además, si $A_1, A_2, \dots \in \mathcal{F}$ entonces $A_1^c, A_2^c, \dots \in \mathcal{F}$ y de este modo, $\bigcup_{i=1}^{\infty} A_i^c \in \mathcal{F}$. Por las leyes de De Morgan, tenemos

$$\Big(\bigcup_{i=1}^{\infty} A_i^c\Big)^c = \bigcap_{i=1}^{\infty} A_i.$$

Es decir, tenemos que $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

Observación 3.5. Asociado a un espacio muestral Ω puede haber muchas σ-álgebras. Por ejemplo, la colección $\{\varnothing, \Omega\}$ es σ-álgebra (minimal).

Ejemplo 3.6. Considere $\Omega = \{1, 2, 3\}$. Los siguientes subconjuntos de Ω, ¿son σ-álgebras?

$$\mathcal{F}_1 = \{\varnothing, \Omega, \{1\}, \{2, 3\}\},$$
$$\mathcal{F}_2 = \{\varnothing, \Omega, \{1\}, \{2\}, \{1, 3\}, \{2, 3\}\}.$$

Claramente se verifica que $\mathcal{F}_1$ es una σ-álgebra, mientras que $\mathcal{F}_2$ no.

Definición 3.7. Sea Ω un espacio muestral y sea $\mathcal{F}$ un σ-álgebra, decimos que el par $(\Omega, \mathcal{F})$ es un espacio medible. Si $A \in \mathcal{F}$ decimos que A es medible.

Definición 3.8 (Probabilidad)**.** Dado un espacio muestral Ω y un σ-álgebra asociada $\mathcal{F}$, una *función de probabilidad* $\mathsf{P} : \mathcal{F} \to \mathbb{R}$, satisface:

(a) $\mathsf{P}(A) \geq 0$, para todo $A \in \mathcal{F}$.
(b) $\mathsf{P}(\Omega) = 1$.
(c) Si $A_1, A_2, \dots$ son mutuamente excluyentes, entonces

$$\mathsf{P}\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mathsf{P}(A_i).$$

Sea $(\Omega, \mathcal{F})$ un espacio de medida y P una medida de probabilidad definida en $\mathcal{F}$. Entonces $(\Omega, \mathcal{F}, \mathsf{P})$ se denomina *espacio de probabilidad*.

Resultado 3.9. *Si* P *es una función de probabilidad y* A *es cualquier conjunto en* $\mathcal{F}$*, entonces*

(a) $\mathsf{P}(\varnothing) = 0$.
(b) $\mathsf{P}(A) \leq 1$.
(c) $\mathsf{P}(A^c) = 1 - \mathsf{P}(A)$.

Demostración. Primero considere (c). Como A y A^c son una partición de Ω, sigue que
$$\mathsf{P}(A \cup A^c) = \mathsf{P}(\Omega) = 1,$$
además $A \cap A^c = \varnothing$ son disjuntos, luego
$$\mathsf{P}(A \cup A^c) = \mathsf{P}(A) + \mathsf{P}(A^c) = 1.$$
Como $\mathsf{P}(A^c) \geq 0$, (b) sigue desde (c). Finalmente, para probar (a) note que $\Omega = \Omega \cup \varnothing$ y como $\Omega \cap \varnothing = \varnothing$, tenemos
$$1 = \mathsf{P}(\Omega) = \mathsf{P}(\Omega \cup \varnothing) = \mathsf{P}(\Omega) + \mathsf{P}(\varnothing),$$
de este modo $\mathsf{P}(\varnothing) = 0$. □

Resultado 3.10. *Si* P *es una función de probabilidad y* $A, B \in \mathcal{F}$. *Entonces,*

(a) $\mathsf{P}(B \cap A^c) = \mathsf{P}(B) - \mathsf{P}(A \cap B)$.
(b) $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$.
(c) *Si* $A \subseteq B \Rightarrow \mathsf{P}(A) \leq \mathsf{P}(B)$.

Demostración. Para notar (a) considere que para A y B conjuntos cualquiera
$$B = (B \cap A) \cup (B \cap A^c)$$
(pues $B = B \cap (A \cup A^c) = B \cap \Omega$). Luego,
$$\mathsf{P}(B) = \mathsf{P}(\{B \cap A\} \cup \{B \cap A^c\}) = \mathsf{P}(B \cap A) + \mathsf{P}(B \cap A^c).$$
En efecto, $(B \cap A) \cap (B \cap A^c) = B \cap (A \cap A^c) \cap B = \varnothing$. Para probar (b), note que
$$(A \cup B) = A \cup (B \cap A^c) = (A \cup B) \cap (A \cap A^c) = (A \cup B) \cap \Omega.$$
Además,
$$A \cap (B \cap A^c) = (A \cap A^c) \cap B = \varnothing \cap B = \varnothing.$$
Por tanto, sigue que
$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B \cap A^c) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B),$$
por parte (a). Si $A \subseteq B$ entonces $A \cap B = A$. De este modo, usando (a) tenemos
$$0 \leq \mathsf{P}(B \cap A^c) = \mathsf{P}(B) - \mathsf{P}(A),$$
y (c) es verificado. □

Como $\mathsf{P}(A \cup B) \leq 1$, reagrupando términos tenemos
$$\mathsf{P}(A \cap B) \geq \mathsf{P}(A) + \mathsf{P}(B) - 1,$$
la desigualdad anterior es un caso particular de la desigualdad de Bonferroni.

Ejemplo 3.11. En una cierta zona urbana, el 60 % de los propietarios están suscritos al diario y el 80 % está suscrito al servicio de cable, mientras que el 50 % está suscrito a ambos servicios. Si un propietario es seleccionado al azar. ¿Cuál es la probabilidad de que esté suscrito al menos a uno de estos servicios?

Defina el evento A : un propietario está suscrito al diario, y B : un propietario está suscrito al servicio de cable. Se desea calcular la probabilidad
$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B) = 0{,}6 + 0{,}8 - 0{,}5 = 0{,}9.$$
Además, ¿cuál es la probabilidad de que esté suscrito sólo a uno de los dos servicios?
$$\begin{aligned}\mathsf{P}[(A \cap B^c) \cup (A^c \cap B)] &= \mathsf{P}(A \cap B^c) + \mathsf{P}(A^c \cap B) = \mathsf{P}(A) + \mathsf{P}(B) - 2\,\mathsf{P}(A \cap B)\\ &= 0{,}6 + 0{,}8 - 2 \cdot 0{,}5 = 0{,}4.\end{aligned}$$

Resultado 3.12. *Si* P *es una función de probabilidad. Entonces,*

$$\mathsf{P}(A) = \sum_{i=1}^{\infty} \mathsf{P}(A \cap C_i),$$

para cualquier partición $C_1, C_2, \dots$.

Demostración. Dado que $C_1, C_2, \dots$ forman una partición, tenemos $C_i \cap C_j = \varnothing$ para todo $i \neq j$ y $\Omega = \bigcup_{i=1}^{\infty} C_i$. De ahí que

$$A = A \cap \Omega = A \cap \Big(\bigcup_{i=1}^{\infty} C_i \Big) = \bigcup_{i=1}^{\infty} (A \cap C_i).$$

De este modo,

$$\mathsf{P}(A) = \mathsf{P}\Big(\bigcup_{i=1}^{\infty} (A \cap C_i) \Big) = \sum_{i=1}^{\infty} \mathsf{P}(A \cap C_i),$$

pues, dado que los C_i's son disjuntos, tambien lo es la secuencia $\{A \cap C_i\}_{i=1}^{\infty}$. □

3.1. Espacios muestrales finitos. En esta sección consideraremos que

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\},$$

para caracterizar $\mathsf{P}(A)$ supondremos eventos elementales, es decir $A = \{\omega_i\}$ y definimos $p_i = \mathsf{P}(\{\omega_i\})$ la probabilidad de $\{\omega_i\}$ tal que,

(a) $p_i \geq 0$, $i = 1, \dots, n$.
(b) $p_1 + p_2 + \cdots + p_n = 1$.

Suponga ahora que $A = \{\omega_{j_1}, \dots, \omega_{j_r}\}$ está formado por r elementos de Ω, luego

$$\mathsf{P}(A) = p_{j_1} + \cdots + p_{j_r}.$$

Adicionalmente, supondremos que cada $\{\omega_i\}$ es igualmente probable. Entonces,

$$p_i = \mathsf{P}(\{\omega_i\}) = \frac{1}{n}.$$

Luego, para un evento $A = \{\omega_{j_1}, \dots, \omega_{j_r}\}$ sigue que

$$\mathsf{P}(A) = \frac{r}{n},$$

o bien,

$$\mathsf{P}(A) = \frac{\mathsf{N}(A)}{\mathsf{N}(\Omega)},$$

con $\mathsf{N}(A)$ la cardinalidad del conjunto A. Debemos resaltar que esta *no* es una definición general, sino que apropiada *sólo* bajo el supuesto de espacios muestrales finitos y equiprobables.

Ejemplo 3.13. Se lanza un dados dos veces. Considere los siguientes eventos:

A : la suma de los resultados es menor o igual a 3.
B : el resultado del primer lanzamiento es impar.

Primeramente, describiremos el espacio muestral asociado al experimento, es decir:

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

Luego,

$$\mathsf{P}(A) = \frac{\mathsf{N}(A)}{\mathsf{N}(\Omega)} = \frac{3}{36} = \frac{1}{12}, \qquad \mathsf{P}(B) = \frac{\mathsf{N}(B)}{\mathsf{N}(\Omega)} = \frac{1}{2}.$$

Además,

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B) = \frac{1}{12} + \frac{18}{36} - \frac{2}{36} = \frac{19}{36}.$$
$$\mathsf{P}(A^c) = 1 - \mathsf{P}(A) = 1 - \frac{1}{12} = \frac{11}{12}.$$

Ejemplo 3.14. En un lote de 50 ampolletas hay 2 defectuosas. Se extraen 5 de ellas al azar sin reemplazo. Hallar la probabilidad de que al menos una sea defectuosa. De este modo, considere el evento A : existe al menos una ampolleta defectuosa en las 5 extracciones. De ahí que A^c : corresponde al evento: ninguna ampolleta es defectuosa. Entonces,

$$\begin{aligned}\mathsf{P}(A^c) &= \frac{48}{50}\frac{47}{49}\frac{46}{48}\frac{45}{47}\frac{44}{46} = 0{,}9600 \cdot 0{,}9592 \cdot 0{,}9583 \cdot 0{,}9574 \cdot 0{,}9565 \\ &= 0{,}8082.\end{aligned}$$

Así, $\mathsf{P}(A) \approx 0{,}2$. Note que una manera alternativa de calcular esta probabilidad es usando combinatorios. En efecto,

$$\mathsf{P}(A^c) = \frac{\binom{48}{5}\binom{2}{0}}{\binom{50}{5}} = \frac{1\,712\,304}{2\,118\,760} = 0{,}8082.$$

Ejemplo 3.15 (Problema del cumpleaños)**.** Para introducir ideas, suponga que existe 50 personas en una sala y que deseamos determinar la probabilidad de que al menos un par de personas tengan la misma fecha de cumpleaños.[1] Para abordar el problema, sea A el evento de que 2 personas tengan cumpleaños en días *diferentes*, entonces:

$$\mathsf{P}(A) = \frac{365}{365} \cdot \frac{364}{365}.$$

Considere ahora que escogemos una tercera persona, es decir, para el evento B : que 3 personas tengan cumpleaños en días diferentes, es dado por

$$\mathsf{P}(A) = \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365}.$$

Suponga, un poco más generalmente, un grupo de k personas y suponga que se desea calcular la probabilidad de que al menos 2 personas cumplan años el mismo día. Primeramente defina el evento A_k : k personas tienen sus cumpleaños en fechas diferentes. Como el año tiene 365 días, el número de cumpleaños posibles es 365^k. Así,

$$\mathsf{P}(A_k) = \frac{(365)_k}{365^k} = \frac{365 \cdot 364 \cdots (365 - k + 1)}{365^k} = \frac{365!}{(365-k)!}\,\frac{1}{365^k},$$

y por la regla del complemento,

$$\mathsf{P}(A_k^c) = 1 - \frac{(365)_k}{365^k}.$$

Es decir, por ejemplo

k	10	15	20	25	30	35	40	45	50
$\mathsf{P}(A_k^c)$	0.117	0.253	0.411	0.569	0.706	0.814	0.891	0.941	0.970

[1]No consideraremos el 29 de febrero para este problema.

El siguiente fragmento de código construye la Figura 1:

```
# secuencia con el número de personas
> k <- 1:100
# probabilidad de que k personas cumplan anos en fechas diferentes
> diferentes <- function(k) {
+   p <- choose(365, k) * factorial(k) / 365^k
+   p
+ }
# construyendo la tabla anterior
> probs <- 1 - diferentes(k)
> cuales <- seq(10, 50, by = 5)
> probs[cuales]
[1] 0.1169 0.2529 0.4114 0.5687 0.7063 0.8144 0.8912 0.9410 0.9704

# despliega el gráfico en la Figura 1
> plot(k, probs, type = "l", xlab = "Número de personas",
+       ylab = "Probabilidad")
```

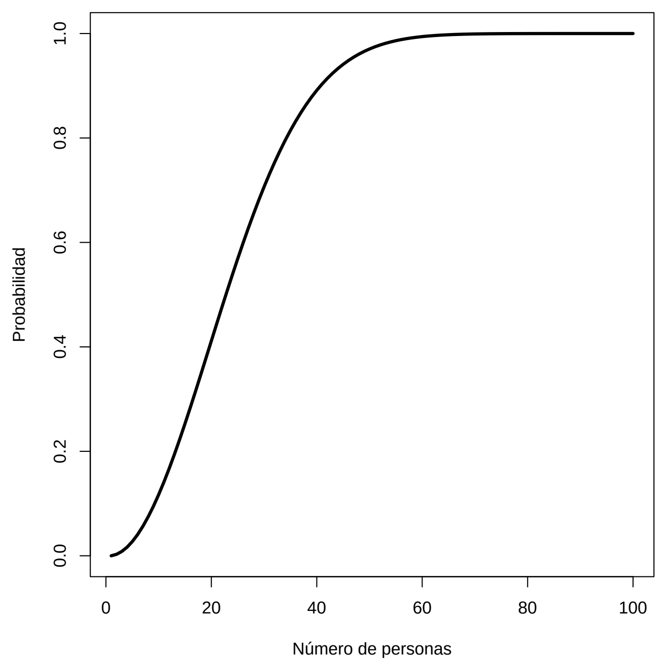

Figura 1. Probabilidad de que dos personas en un grupo tengan cumpleaños el mismo día como una función del número de personas en el grupo.

3.2. Probabilidad condicional. Para introducir ideas, considere un lote con 80 artículos sin defectos y 20 defectuosos y suponga que se selecciona 2 artículos (a) con substitución, y (b) sin substitución. Defina los eventos:

$$\begin{aligned} A &= \{\text{el 1er artículo es defectuoso}\}, \\ B &= \{\text{el 2do artículo es defectuoso}\}. \end{aligned}$$

Cuando escogemos *con* substitución, tenemos:

$$\mathsf{P}(A) = \mathsf{P}(B) = \frac{20}{100} = \frac{1}{5}.$$

Cuando escogemos *sin* substitución, tenemos que:

$$\mathsf{P}(A) = \frac{20}{100} = \frac{1}{5},$$

pero, ¿cambia $\mathsf{P}(B)$?

Definición 3.16. Si A y B son dos eventos en Ω y $\mathsf{P}(B) > 0$, entonces la *probabilidad condicional* de A dado B, escrito $\mathsf{P}(A|B)$ es

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

Note que $\mathsf{P}(B|B) = 1$, es decir, B *"actua"* como Ω. En efecto, como $A = A \cap \Omega$, tenemos

$$\mathsf{P}(A) = \mathsf{P}(A|\Omega) = \frac{\mathsf{P}(A \cap \Omega)}{\mathsf{P}(\Omega)}.$$

La ocurrencia de A es calibrada con relación a B. En particular, si $A \cap B = \varnothing$, entonces

$$\mathsf{P}(A|B) = \mathsf{P}(B|A) = 0.$$

En el ejemplo anterior, se desea calcular $\mathsf{P}(B|A) = 19/99$, pues si A ya ha ocurrido sólo quedan 19 defectuosos entre los 99 artículos.

Reexpresando la probabilidad condicional, tenemos

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

o bien

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

Las expresiones anteriores permiten "contornar" cálculos complicados, usando[2]

$$\mathsf{P}(A|B) = \frac{\mathsf{P}(B|A)\,\mathsf{P}(A)}{\mathsf{P}(B)}.$$

Observación 3.17. El espacio de probabilidad definido por $\mathcal{F} \cap B$ permite notar que $\mathsf{P}(A|B)$ es una función de probabilidad, es decir satisface:

(a) $\mathsf{P}(A|B) \geq 0$.
(b) $\mathsf{P}(\Omega|B) = 1$.
(c) Para $\{A_n\}_{n\geq 1}$ sucesión disjunta

$$\mathsf{P}\Big(\bigcup_{n=1}^{\infty} A_n \Big| B\Big) = \sum_{n=1}^{\infty} \mathsf{P}(A_n|B).$$

[2]Esto es un caso particular del Teorema de Bayes.

Resultado 3.18 (Probabilidad total)**.** *Sea* $(\Omega, \mathcal{F}, \mathsf{P})$ *un espacio de probabilidad y sea* $C_1, C_2, \dots,$ *una partición contable de* Ω *tal que* $\mathsf{P}(C_i) \geq 0$, $\forall i$. *Entonces, para todo* $A \in \mathcal{F}$,

$$\mathsf{P}(A) = \sum_{i=1}^{\infty} \mathsf{P}(A_i|C_i)\,\mathsf{P}(C_i).$$

Demostración. Como los C_i's forman una partición tenemos que

$$A = A \cap \Omega = A \cap \Big(\bigcup_{i=1}^{\infty} C_i\Big) = \bigcup_{i=1}^{\infty} (A \cap C_i).$$

Además,

$$\mathsf{P}(A) = \sum_{i=1}^{\infty} \mathsf{P}(A \cap C_i) = \sum_{i=1}^{\infty} \mathsf{P}(A|C_i)\,\mathsf{P}(C_i).$$

□

Resultado 3.19 (Teorema de Bayes)**.** *Sea* $(\Omega, \mathcal{F}, \mathsf{P})$ *espacio de probabilidad y sea* $\{C_i\}_{i\geq 1}$ *partición contable de* Ω *con* $\mathsf{P}(C_i) \geq 0$, $\forall i$. *Entonces, para todo* $A \in \mathcal{F}$, *tenemos que*

$$\mathsf{P}(C_i|A) = \frac{\mathsf{P}(A|C_i)\,\mathsf{P}(C_i)}{\sum_{k=1}^{\infty} \mathsf{P}(A|C_k)\,\mathsf{P}(C_k)},$$

siempre que $\mathsf{P}(A) > 0$.

Demostración. Tenemos que

$$\mathsf{P}(C_i|A)\,\mathsf{P}(A) = \mathsf{P}(A|C_i)\,\mathsf{P}(C_i),$$

así

$$\mathsf{P}(C_i|A) = \frac{\mathsf{P}(A|C_i)\,\mathsf{P}(C_i)}{\mathsf{P}(A)}, \qquad \mathsf{P}(A) > 0,$$

desde el Teorema de probabilidad total, sigue que

$$\mathsf{P}(C_i|A) = \frac{\mathsf{P}(A|C_i)\,\mathsf{P}(C_i)}{\sum_{k=1}^{\infty} \mathsf{P}(A|C_k)\,\mathsf{P}(C_k)}.$$

□

Ejemplo 3.20. Considere un lote de 20 artículos defectuosos y 80 sin defectos, desde los que se escoge 2 artículos sin reemplazo. Sea

$$\begin{aligned} A &= \{\text{el 1er artículo es defectuoso}\}, \\ B &= \{\text{el 2do artículo es defectuoso}\}. \end{aligned}$$

Para calcular $\mathsf{P}(B)$ podemos hacer

$$\begin{aligned} \mathsf{P}(B) &= \mathsf{P}(B|A)\,\mathsf{P}(A) + \mathsf{P}(B|A^c)\,\mathsf{P}(A^c) = \frac{19}{99}\frac{1}{5} + \frac{20}{99}\frac{4}{5} \\ &= \frac{1}{5}\frac{1}{99}(19 + 20 \cdot 4) = \frac{1}{5}. \end{aligned}$$

3.3. Independencia estadística. Hasta el momento hemos manipulado situaciones donde los eventos A y B están relacionados. A continuación se introduce un concepto clave en estadística que permite caracterizar que dos o más eventos no tienen relación.

Definición 3.21. Sea $(\Omega, \mathcal{F}, \mathsf{P})$ un espacio de probabilidad y sean $A, B \in \mathcal{F}$. Se dice que A y B son *independientes* si y sólo si

$$\mathsf{P}(A \cap B) = \mathsf{P}(A)\,\mathsf{P}(B).$$

Naturalmente podemos entender la independencia del siguiente modo: "la ocurrencia de un evento B no tiene efecto en la probabilidad de otro evento A". Es decir,

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

Note tambien que,

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

es decir, la ocurrencia de A *no* tiene efecto en B.

Ejemplo 3.22. Se tienen tres urnas, la primera con 2 bolas blancas y dos bolas negras, la segunda con dos bolas blancas y una negra y la tercera con tres bolas negras y una blanca.

Suponga que se extrae una bola de cada urna, ¿cuál es la probabilidad de que las tres bolas sean blancas? Para llevar a cabo este cálculo, defina el evento B_i : extraer una bola blanca de la i-ésima urna, $i = 1, 2, 3$. De este modo, asumiendo independencia obtenemos

$$\mathsf{P}(B_1 \cap B_2 \cap B_3) = \mathsf{P}(B_1)\,\mathsf{P}(B_2)\,\mathsf{P}(B_3) = \frac{1}{2}\,\frac{2}{3}\,\frac{1}{4} = \frac{1}{12}.$$

Ahora suponga que se extrae una bola de una urna al azar, ¿cuál es la probabilidad de que sea blanca? Defina los eventos,

U_i : la bola extraída proviene de la i-ésima urna, $i = 1, 2, 3$.
B : extraer una bola blanca.

Luego,

$$\mathsf{P}(B) = \sum_{i=1}^{3} \mathsf{P}(B|U_i)\,\mathsf{P}(U_i) = \frac{1}{2}\,\frac{1}{3} + \frac{2}{3}\,\frac{1}{3} + \frac{1}{4}\,\frac{1}{3} = \frac{17}{36}.$$

Si se sabe que la bola extraída es de color blanco, ¿cuál es la probabilidad de que haya sido extraída de la urna 1? En efecto, deseamos calcular

$$\mathsf{P}(U_1|B) = \frac{\mathsf{P}(U_1 \cap B)}{\mathsf{P}(B)} = \frac{\mathsf{P}(B|U_1)\,\mathsf{P}(U_1)}{\mathsf{P}(B)} = \frac{\frac{1}{2}\frac{1}{3}}{\frac{17}{36}} = \frac{6}{17}.$$

Resultado 3.23. *Si A y B son independientes, entonces los siguientes pares también son independientes*

(a) A *y* B^c.
(b) A^c *y* B.
(c) A^c *y* B^c.

Demostración. Para probar (a), note que

$$\begin{aligned}\mathsf{P}(A \cap B^c) &= \mathsf{P}(A) - \mathsf{P}(A \cap B) = \mathsf{P}(A) - \mathsf{P}(A)\,\mathsf{P}(B) = \mathsf{P}(A)(1 - \mathsf{P}(B)) \\ &= \mathsf{P}(A)\,\mathsf{P}(B^c).\end{aligned}$$

Partes (b) y (c) son análogas y se dejan de ejercicio para el lector. □

Definición 3.24. Una colección de eventos $A_1, A_2, \ldots, A_n$ es *mutuamente independiente* si para cualquier subcolección $A_{i_1}, \ldots, A_{i_k}$, tenemos

$$\mathsf{P}\Big(\prod_{j=1}^{k} A_{i_j}\Big) = \prod_{j=1}^{k} \mathsf{P}(A_{i_j}).$$

Ejemplo 3.25. Se lanzan 3 dados de distinto color: blanco, rojo y negro ¿Cuál es la probabilidad de que el dado blanco salga 3 y los otros dos no? Considere A, B y C los eventos

$$\begin{aligned}A &: \text{resultado del dado blanco es 3,} \\ B &: \text{resultado del dado rojo es 3,} \\ C &: \text{resultado del dado negro es 3,}\end{aligned}$$

tenemos $\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = 1/6$ y se pide calcular

$$\mathsf{P}(A \cap B^c \cap C^c) = \mathsf{P}(A)\,\mathsf{P}(B^c)\,\mathsf{P}(C^c) = \frac{1}{6}\frac{5}{6}\frac{5}{6} = \frac{25}{216}.$$

Ejemplo 3.26. Suponga que se lanza un dado en dos oportunidades de forma consecutiva. Note que los resultados posibles pueden ser expresados por el conjunto,

$$\begin{aligned}\{&(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\ &(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\ &(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),\\ &(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\ &(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),\\ &(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}.\end{aligned}$$

Ahora, suponga X una variable asociada al evento: *La suma de las caras de un dado en dos lanzamientos consecutivos.* Podemos escribir la siguiente tabla con las probabilidades asociadas a cada uno de los resultados que puede adoptar X, es decir:

x	2	3	4	5	6	7	8	9	10	11	12
$\mathsf{P}(X = x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Asimismo, podemos considerar el siguiente gráfico, donde se aprecia que las probabilidades asociadas a variable X adoptan un comportamiento gobernado por una ley.

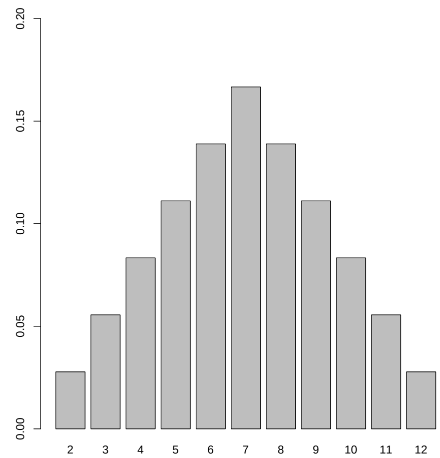

Figura 2. Gráfico de la función de probabilidades asociada a la variable X.

Referencias

Casella, G., Berger, R.L. (2002). *Statistical Inference (2nd Ed.)*. Duxbury, Pacific Grove.

Charalambides, C.A. (2002). *Enumerative Combinatorics* Chapman & Hall, Boca Raton.

Chung, K.L. (1974). *A Course in Probability Theory (2nd Ed.)*. Academic Press, New York.

Hoel, P.G., Port, S.C., Stone, C.J. (1971). *Introduction to Probability Theory* Houghton Mifflin Company, Boston.

R Core Team (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: www.R-project.org

Instituto de Estadística, Universidad de Valparaíso, Chile
URL: 0000-0002-4675-5201
Email address: felipe.osorio@uv.cl