Unidad 1. Probabilidad

El siguiente material es sólo orientativo, está sujeto a correcciones y se irá completando con el intercambio en el aula.

En esta unidad estudiaremos las nociones generales de probabilidad haciendo énfasis en los problemas de combinatoria y nos guiaremos de diferentes textos como Baron (2006), Devore (2018), Maronna (2021), Ross (2015) y Yohai (2006). Las consideraciones históricas tendrán como principal referencia el libro de Hacking (1995).

1. Espacio de probabilidad

Un modelo probabilístico es una descripción matemática de una situación caracterizada por la incerteza denominada *experimento aleatorio*.

Definición 1. *Un experimento aleatorio $\mathcal{E}$ es un experimento del cual se conocen todos los resultados posibles pero no se conoce, a priori, el resultado que ocurrirá si se lleva a cabo. Al conjunto de todos los resultados posibles asociados a $\mathcal{E}$ se lo denomina espacio muestral y se lo denota con Ω.*

La noción intuitiva de evento debe formalizarse adecuadamente y se lleva a cabo a través de la siguiente definición.

Definición 2. *Una colección $\mathcal{F}$ de subconjuntos de Ω se llama σ-álgebra si se satisfacen las siguientes tres condiciones*

a) $\forall A \in \mathcal{F} : \; A^c \in \mathcal{F}$.

b) $\emptyset \in \mathcal{F}$.

c) Si $\{A_i\}_{i \in I}$ es cualquier colección a lo sumo numerable de elementos de $\mathcal{F}$ entonces $\bigcup_{i \in I} A_i \in \mathcal{F}$.

A cada elemento de la σ-álgebra $\mathcal{F}$ le llamaremos evento.

Al par ordenado $(\Omega, \mathcal{F})$ se le llama *espacio medible*.

Algunos eventos tienen un nombre especial: a Ω se le llama "suceso seguro", al "vacío" se le denomina "suceso imposible". Y si $\omega \in \Omega$ y $\{\omega\}$ es un evento, a este último se le denomina "suceso elemental".

Ejemplo 1. *Dado un conjunto no vacío Ω, $\mathcal{F} = \{\emptyset, \Omega\}$ es una σ-álgebra. También la colección "partes" del conjunto Ω definida como*

$$\mathcal{P}(\Omega) = \{E : E \text{ es un subconjunto de } \Omega\}$$

es una σ-álgebra.

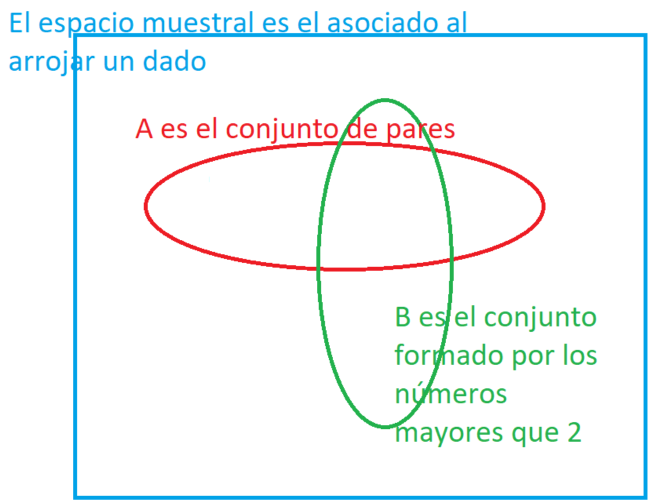

Figura 1: Lanzamiento del dado

Ejemplo 2. *Si $\mathcal{E}$ es el experimento aleatorio consistente en "lanzar un dado" entonces $\Omega = \{1,2,3,4,5,6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$. En la Figura 1 se representa Ω, A el evento "que resulte un número par" y B es el evento "que resulte un número mayor que 2". Así $A = \{2,4,6\}$ y $B = \{3,4,5,6\}$.*

Y ahora llegamos a la *definición axiomática de la probabilidad.*

Definición 3. *Sea $(\Omega, \mathcal{F})$ un espacio medible. A una función*

$$\begin{array}{rcl} P:\mathcal{F} & \to & [0,1] \\ A & \longmapsto & P(A) \end{array}$$

que satisface los siguientes axiomas

a) $P(\Omega) = 1$,

b) Si $\{A_i\}_{i=1}^{\infty}$ es una colección numerable de eventos disjuntos dos a dos ($A_i \cap A_j = \emptyset$ si $i \neq j$) entonces

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right),$$

se la denomina probabilidad. A la terna $(\Omega, \mathcal{F}, P)$ se le llama espacio de probabilidad.

La siguiente proposición da propiedades elementales de la probabilidad.

Proposición 1. *Son válidas las siguientes propiedades*

1) $P(\emptyset) = 0$.

2) Si $\{A_i\}_{i=1}^n$ *es una sucesión finita de eventos disjuntos dos a dos entonces*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P\left(A_i\right).$$

3) Para cualquier evento A *se satisface*

$$P\left(A^c\right) = 1 - P\left(A\right).$$

4) $\forall A, B \in \mathcal{F}$ *si* $A \subseteq B$ *entonces*

$$P\left(B\right) = P\left(A\right) + P\left(B - A\right) \geq P\left(A\right).$$

5) Si $\{A_i\}_{i=1}^n$ *es una sucesión finita de eventos entonces*

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{i=1}^{n} P\left(A_i\right) - \sum_{1\leq i<j\leq n} P\left(A_i \cap A_j\right) + \sum_{1\leq i<j<k\leq n} P\left(A_i \cap A_j \cap A_k\right) + \ldots \quad (1)\\ &+ \ldots (-1)^{n+1} P\left(A_1 \cap \ldots \cap A_n\right) \end{aligned}$$

Demostración

1) Queremos mostrar que $P(\emptyset) = 0$. Notar que si $\forall i \geq 1$ definimos $A_i = \emptyset$ entonces

$$\emptyset = \bigcup_{i=1}^{\infty} A_i$$

y por b) de la definición axiomática (notar que los eventos A_i son conjuntos disjuntos dos a dos) se tiene que

$$P(\emptyset) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right) = \lim_{n\to\infty} \sum_{i=1}^{n} P\left(A_i\right) = \lim_{n\to\infty} \sum_{i=1}^{n} P\left(\emptyset\right) = \lim_{n\to\infty} nP\left(\emptyset\right)$$

y dado que $P(\emptyset)$ es un número positivo o cero entonces debe ser $P(\emptyset) = 0$ de lo contrario tendríamos que $\lim_{n\to\infty} nP\left(\emptyset\right) = \infty$.

2) Si definimos $A_j = \emptyset$ para $j > n$ entonces la propiedad se deduce del axioma b) y la afirmación probada en el ítem anterior.

3) Dado A, como $\Omega = A \cup A^c$, la propiedad se deduce de la demostrada en el ítem 2) y de uno de los axiomas que define la probabilidad.

4) Recordemos que por definición la operación diferencia "de un conjunto B menos un conjunto A" se define como $B - A = B \cap A^c$. De este modo como

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

la propiedad se deduce nuevamente por la del ítem 2).

5) Para probar esta afirmación utilizaremos inducción.

El "paso o caso base" se deduce de la propiedad del ítem 4). En efecto, dados dos eventos A_1 y A_2, esa propiedad nos permite descomponer

$$\begin{aligned} P(A_1) &= P(A_1 \cap A_2) + P(A_1 - (A_1 \cap A_2) \\ P(A_2) &= P(A_1 \cap A_2) + P(A_2 - (A_1 \cap A_2) \end{aligned}$$

y de aquí

$$P(A_1) + P(A_2) - P(A_1 \cap A_2) = P(A_1 \cup A_2).$$

Probemos ahora que, si la propiedad vale para n entonces vale para $n+1$. Consideremos $\{A_i\}_{i=1}^{n+1}$ una sucesión finita de eventos. Luego

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) =& P(\bigcup_{i=1}^{n} A_i \cup A_{n+1}) \\ =& P(\bigcup_{i=1}^{n} A_i) + P(A_{n+1}) - P((\bigcup_{i=1}^{n} A_i) \cap A_{n+1}), \qquad \text{por el caso base.} \end{aligned} \tag{2}$$

Aplicando la hipótesis inductiva a la sucesión de de conjuntos $A_1 \cap A_{n+1}, \ldots, A_n \cap A_{n+1}$ se tiene que:

$$\begin{aligned} -P((\bigcup_{i=1}^{n} A_i) \cap A_{n+1}) = & -P(\bigcup_{i=1}^{n} (A_i \cap A_{n+1})) \\ & = -\sum_{i=1}^{n} P\left(A_i \cap A_{n+1}\right) + \sum_{1 \leq i < j \leq n} P\left(A_i \cap A_j \cap A_{n+1}\right) + \\ & - \sum_{1 \leq i < j < k \leq n} P\left(A_i \cap A_j \cap A_{n+1} \cap A_k\right) + \ldots \\ & + \ldots \underbrace{-(-1)^{n+1}}_{+(-1)^{n+2}} P\left(A_1 \cap \ldots \cap A_n \cap A_{n+1}\right). \end{aligned}$$

Reemplazando esta expresión obtenida y, además, el resultado de aplicar la hipótesis inductiva a $P(\bigcup_{i=1}^{n} A_i)$ (o sea (1)) en (2) y, acomodando convenientemente los términos se obtiene

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) &= \sum_{i=1}^{n+1} P\left(A_i\right) - \sum_{1 \leq i < j \leq n} P\left(A_i \cap A_j\right) + \sum_{1 \leq i < j < k \leq n+1} P\left(A_i \cap A_j \cap A_k\right) + \ldots \\ & + \ldots (-1)^{n+2} P\left(A_1 \cap \ldots \cap A_{n+1}\right). \end{aligned}$$

□

Una sucesión de conjuntos $\{A_n\}_{n\geq 1}$ se dice creciente (decreciente) si $\forall n \geq 1 : A_n \subseteq A_{n+1}$ (si $\forall n \geq 1 : A_{n+1} \subseteq A_n$).

La próxima proposición nos dice que la función de probabilidad es una función "continua" para lo cual es necesario introducir una noción de límite para sucesiones.

Proposición 2. *Dado $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, son válidas las siguientes afirmaciones:*

a) Si $\{A_n\}_{n\geq 1}$ es una sucesión creciente de eventos entonces

$$\lim_{n\to\infty} P(A_n) = P(A)$$

con $A = \bigcup_{n=1}^{\infty} A_n$.

b) Si $\{A_n\}_{n\geq 1}$ es una sucesión decreciente de eventos entonces

$$\lim_{n\to\infty} P(A_n) = P(A)$$

con $A = \bigcap_{n=1}^{\infty} A_n$.

Demostración. Haremos la prueba de a) y dejamos b) como ejercicio. Si nos guiamos por el esquema dibujado en (2) definimos la siguiente sucesión de conjuntos

$$\begin{aligned} B_1 &= A_1 \\ B_2 &= A_2 - A_1 \\ B_n &= A_n - A_{n-1}, \forall n \geq 3 \end{aligned}$$

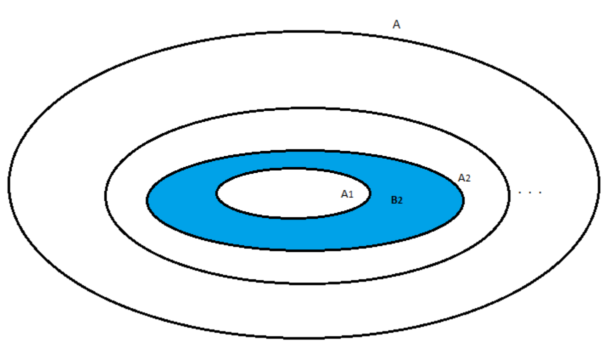

Figura 2: Sucesión creciente de conjuntos

la cual es una sucesión de eventos disjuntos dos a dos y tal que

$$\bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{n} A_i = A_n \tag{3}$$
$$\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i = A \tag{4}$$

Así

$$\begin{aligned}
\lim_{n\to\infty} P(A_n) &= \lim_{n\to\infty} P(\bigcup_{i=1}^{n} B_i) && \text{por (3)}\\
&= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) && \text{los } B_i \text{ son disjuntos dos a dos}\\
&= \sum_{i=1}^{\infty} P(B_i) && \text{por definición de serie}\\
&= P(\bigcup_{i=1}^{\infty} B_i) && \text{usamos el axioma que define la probabilidad}\\
&= P(A) && \text{por (4)}
\end{aligned}$$

como queríamos probar. □

El siguiente resultado nos da una cota para la probabilidad de la unión de eventos. Su demostración se deja como ejercicio y se encuentra en el práctico.

Proposición 3. ***Desigualdad de Bonferroni***

Si $\{A_i\}_{i=1}^{n}$ es una sucesión de eventos entonces

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P\left(A_i\right).$$

2. Espacios discretos

Definición 4. *Un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ se denomina discreto si existe un evento Ω_0 a lo sumo numerable tal que $P(\Omega_0) = 1$.*

De este modo en los espacios de probilidad discretos "toda la masa" está contenida en un conjunto que o bien es finito o bien es infinito y se puede contar.

Un elemento ω de un espacio muestral se llama un átomo si $P(\{w\}) > 0$. Si en la Definición 4 exigimos además que todo elemento $\omega \in \Omega_0$ sea un átomo entonces se cumple que

$$\text{Si } E \text{ es un evento tal que } P(E) = 0 \text{ entonces } E \cap \Omega_0 = \emptyset$$

(¿por qué?).

Ejemplo 3. *Sea* $\Omega = \{\omega_1, \omega_2, \cdots\}$ *un conjunto a lo sumo numerable,* $\mathcal{F} = \mathcal{P}(\Omega)$ *y sea* $\{p_i\}_{i\geq 1}$ *una sucesión de números en el intervalo* $(0,1)$ *tal que* $\sum_{i=1}^{\infty} p_i = 1$*. Entonces, dado un subconjunto* E *de* Ω *si definimos*

$$P(E) = \sum_{i:\omega_i \in E} p_i \tag{5}$$

entonces $(\Omega, \mathcal{F}, P)$ *es un espacio de probabilidad.*

Tener en cuenta que hay que probar, para garantizar que la definición es correcta, que (5) define una probabilidad (ejercicio). Notar que en la definición debe entenderse que si Ω es finito entonces la sucesión $\{p_i\}_{i\geq 1}$ también lo es y tiene el mismo número de elementos que el espacio muestral.

Definición 5. *Sea* $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $p_i = 1/n$ *para* $i = 1 \ldots, n$ *y* P *como en* (5)*. Entonces, para cualquier evento* E*,*

$$P(E) = \frac{\#E}{\#\Omega} = \frac{\textit{casos favorables}}{\textit{casos posibles}}. \tag{6}$$

A este espacio de probabilidad se le denomina uniforme discreto o de equiprobabilidad.

Ya hemos abordado ejemplos de espacios uniformes discretos, como por ejemplo el asociado al experimento aleatorio $\mathcal{E}$: "arrojar un dado y registrar el número que resulta en su cara superior".

Ejemplo 4. *Considerar el experimento aleatorio* $\mathcal{E}$: *"lanzar sucesivamente una moneda hasta que salga cara". De este modo*

$$\Omega = \{c, xc, xxc, xxxc, \ldots, xxxx\cdots\}$$

(donde $xxxx\cdots$ *denota el resultado "que nunca salga cara").*

Así si, para $i = 1, 2, \ldots$, ω_i *denota el resultado "que salga cara luego de* $i-1$ *cruces" definimos* $p_i = 1/2^i$*. Y, definamos* $\omega_0 = xxxx\cdots$.

En primer lugar notar que $P(\textit{nunca salga cara}) = 0$ *con lo cual el evento* $E = \{\omega_0\}$ *tiene probabilidad cero pero no es el conjunto vacío.*

Por otro lado si A *es el evento "que salga cara en una cantidad par de lanzamientos",*

$$\begin{aligned} P(A) &= \sum_{i=1}^{\infty} \frac{1}{2^{2i}} = \sum_{i=1}^{\infty} \frac{1}{4^i} \\ &= \frac{1}{4} \sum_{i=0}^{\infty} \frac{1}{4^i} = \frac{1}{4} \frac{1}{1-1/4} \\ &= 1/3 \end{aligned}$$

lo que a priori podría contrariar el sentido común.

3. Técnicas de conteo

3.1. Introducción

En esta sección trabajaremos sólo con espacios de equiprobabilidad de tal modo que la probabilidad de un evento es el cociente de "casos favorables" sobre "casos posibles". Ingresaremos así al estudio de problemas de combinatoria.

El siguiente principio es la base para la mayoría de las estrategias de conteo que utilizaremos en la asignatura.

Principio fundamental de conteo o de multiplicación

Consideremos un proceso que consiste en la ejecución de r etapas. Supongamos que:

a) *Hay n_1 resultados posibles en la primera etapa.*

b) *Para cada resultado en la primera etapa hay n_2 resultados posibles en la segunda etapa.*

c) *Más generalmente, para cada sucesión de resultados en las primeras $(i-1)$ etapas, hay n_i resultados para la i-ésima etapa con $i = 2, \ldots, r$.*

Entonces el número total de resultados posibles que puede arrojar el proceso es $n_1 \cdot n_2 \cdots n_r$.

Una representación se da en la Figura 3

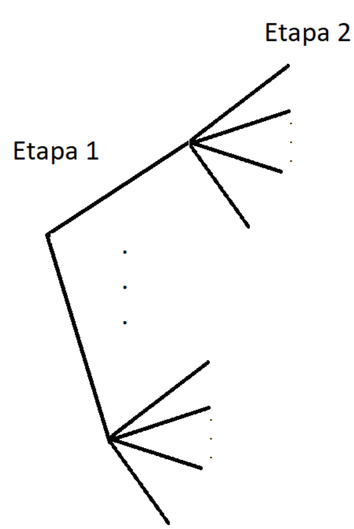

Figura 3: Representación del principio de multiplicación

Observar que este principio, al que también se lo suele denominar “de conteo”, se puede reformular como una proposición en términos de cardinales de espacio producto (¿cómo?).

Ejemplo 5. *En cierta ciudad un número telefónico es una secuencia de 7 dígitos con el primer dígito diferente de 0 y 1. Por el principio de conteo se pueden construir*

$$8 \cdot \underbrace{10 \cdots 10}_{6 \text{ veces}} = 8 \cdot 10^6$$

números telefónicos posibles.

Vamos a asumir que contamos con un *conjunto inicial* con n objetos y de él *seleccionamos al azar* una secuencia de k objetos. Hacemos las siguientes distinciones:

- los objetos pueden ser seleccionados *con o sin reemplazo* y,
- las secuencias pueden ser *distinguibles o indistinguibles*.

Definiremos ahora con precisión los diferentes esquemas de muestreo.

Definición 6. Muestreo con reemplazo *significa que cada objeto muestreado o seleccionado es devuelto al conjunto inicial de tal modo que cada uno de ellos tiene igual probabilidad, $1/n$, de ser seleccionado cada vez.*
Muestreo sin reemplazo *significa que cada objeto seleccionado no es devuelto al conjunto inicial, con lo cual con cada selección el conjunto se reduce en un elemento.*

A continuación le damos precisión a las nociones de *distinguible* e indistinguible involucrando a la noción de *orden*.

Definición 7. *Las secuencias de k objetos se dicen* **distinguibles** *si el muestreo de exactamente los mismos k objetos en un orden diferente establece un resultado diferente (o sea, un elemento distinto del espacio muestral). Las secuencias de objetos se dicen* **indistinguibles** *si el muestreo de exactamente los mismos k objetos en un orden diferente establece el mismo resultado.*

La definición anterior asume que los objetos del conjunto inicial son distinguibles y la diferencia entre distinguible e indistinguible alude a las secuencias de los objetos extraídos del conjunto inicial.

A modo de ejemplo, cuando se escribe un password o clave con símbolos diferentes entonces las claves son distinguibles. En cambio en una encuesta el orden de los sujetos es irrelevante y por lo tanto las secuencias de encuestas no se pueden distinguir unas de otras.

3.2. Permutaciones, combinaciones y particiones

La combinación de los esquemas da lugar a los siguientes tipos básicos de conteo.

3.2.1. Permutaciones con reemplazo

Las posibles secuencias distinguibles de k objetos extraídos de un conjunto de n objetos son denominadas *permutaciones*.

Permutaciones con reemplazo.
Cuando el muestreo es con reemplazo el número total de permutaciones es

$$\underbrace{n \cdots n}_{k \text{ veces}} = n^k \tag{7}$$

y se denota con $P_r(n, k)$.

Y la fórmula se deduce inmediatamente del principio de multiplicación.

Ejemplo 6. *Si el conjunto inicial de objetos es* $\{1, 2, 3\}$*, las permutaciones de* $k = 2$ *elementos con reemplazo son*

$$11, 12, 13, 21, 22, 23, 31, 32, 33,$$

habiendo un total de $P_r(3, 2) = 3^2 = 9$ *de ellas.*

Ejemplo 7. *De un alfabeto consistente de 10 dígitos, 26 letras minúsculas y 26 mayúsculas podemos crear*

$$P_r(62, 8) = 62^8 = 218340105584896 \approx 218 \textit{ trillones}$$

de claves diferentes con 8 caracteres. Entonces si un programa espía tiene una velocidad de inspección de 1 millón de claves por segundo, le tomará casi 7 años en chequear todas las claves. A esta velocidad

$$P(\textit{averiguar tu clave en una semana}) = \frac{60 \cdot 60 \cdot 24 \cdot 7 \cdot 10^6}{62^8} \approx 0.00277$$

3.2.2. Permutaciones sin reemplazo

Si no hay reemplazo el número total de permutaciones se denota con $P(n, k)$. Comencemos con un ejemplo.

Ejemplo 8. *Si el conjunto inicial de objetos es* $\{1, 2, 3\}$*, las permutaciones de* $k = 2$ *elementos sin reemplazo son*

$$12, 13, 21, 23, 31, 32,$$

habiendo un total de $P(3, 2) = 3 \cdot 2 = 6$ *de ellas.*

Utilizando nuevamente el principio de conteo de deduce la siguiente expresión

Permutaciones sin reemplazo.
Cuando el muestreo es sin reemplazo el número total de permutaciones es

$$P(n,k) = n \cdot (n-1) \cdots (n-k+1) = \frac{n!}{(n-k)!}. \tag{8}$$

Recordar que el factorial de un número natural n es $n! =: n \cdot (n-1) \cdot (n-2) \cdots 1$ y si $n = 0$ se define entonces $0! =: 1$.

Ejemplo 9. *Por ejemplo ¿de cuántas formas se pueden sentar 10 estudiantes en una clase de 15 sillas? La respuesta es, de* $P(15, 10) = 15 \cdot 14 \cdots 6 = 110897286400$ *formas. O sea mayor que ciento diez mil millones de formas, el cual es un número bastante grande para comprobarlo empíricamente en nuestra clase.*

Nota. Es importante observar que el número de permutaciones sin reemplazo es igual al número de posibles ubicaciones de k objetos distinguibles en n celdas o compartimentos.

3.2.3. Combinaciones sin reemplazo

En las combinaciones no se releva el orden en las secuencias de k elementos. Además consideremos el caso en que los elementos no pueden ser seleccionados más de una vez.

Ejemplo 10. *Si el conjunto inicial de objetos es* $\{1, 2, 3\}$*, las permutaciones de* $k = 2$ *elementos sin reemplazo son*

$$12, 21$$
$$13, 31$$
$$23, 32$$

En rojo hemos marcado las secuencias que debemos eliminar del conteo si consideramos combinaciones.

En otros términos, estamos buscando determinar la totalidad de subconjuntos de tamaño k construidos a partir de n objetos. En el ejemplo $\{1, 2\}$ y $\{2, 1\}$ forman el mismo subconjunto como así también $\{1, 3\}$ y $\{3, 1\}$ y finalmente $\{2, 3\}$ y $\{3, 2\}$ también son iguales.

La respuesta al interrogante general es la siguiente:

Subconjuntos de un conjunto o combinaciones sin reemplazo.
El número total de combinaciones sin reemplazo al que denotaremos, $C(n, k)$, es igual a

$$C(n, k) = \frac{P(n, k)}{P(k, k)} = \frac{n!/(n-k)!}{k!} = \frac{n!}{(n-k)!k!}. \tag{9}$$

Otra notación es el número combinatorio $\binom{n}{k}$.

¿Cómo se deduce (9)? Si tuviésemos en cuenta el orden, la cantidad total de secuencias ordenadas y considerando que no hay reemplazo es $P(n, k)$. De este modo por cada subconjunto (secuencia no ordenada) contado en $C(n, k)$ hay $k! = P(k, k)$ permutaciones sin reemplazo correspondientes (en el ejemplo, para el 12 hay $2 \cdot 1 = 2!$ secuencias ordenadas, para el 13 hay $2 \cdot 1 = 2!$ secuencias ordenadas, etc).

Para el ejemplo anterior la cantidad total de combinaciones sin reemplazo es

$$\binom{3}{2} = (3!/(3-2)!)/2! = 3$$

De modo general $C(n, k) \cdot P(k, k) = P(n, k)$ que es la expresión que queríamos validar.

Ejemplo 11. *Si un software inspeciona y detecta subconjuntos de 3 carpetas de un total de 10, deberá realizar*

$$\binom{10}{3} = \frac{10!}{7!3!} = 120$$

inspecciones para agotar la nómina de subconjuntos.

3.2.4. Combinaciones con reemplazo

En estas combinaciones el orden no importa y cada objeto puede ser muestreado más de una vez. Un típico ejemplo de este contexto es el siguiente

Ejemplo 12. *Supongamos que una copa de helado tiene 3 bochas, cada una con un sabor elegido de un total de* 5. *Entonces el postre con bochas "vainilla, chocolate, chocolate" es el mismo que el que tiene "chocolate, vainilla, chocolate".*

Otro ejemplo, consideremos las combinaciones de $k = 3$ elementos de los primeros $n = 3$ números naturales en negro y, en rojo las que no contamos porque son equivalentes

$$\begin{array}{ll} 111 & \\ 112 & 121\ 211 \\ 113 & 131\ 311 \\ 122 & 212\ 221 \\ \ldots & \\ 333 & \end{array}$$

En términos generales:

Combinaciones con reemplazo.
El número total de combinaciones con reemplazo de k elementos construidas a partir de n elementos, $C_r(n,k)$, es

$$C_r(n,k) = \binom{n+k-1}{k}. \tag{10}$$

Continuación del Ejemplo 12.

La respuesta a ¿cuántos tipos de postre hay? es

$$C_r(n,k) = \binom{5+3-1}{3} = 35.$$

Notar que si, en cambio, existe la restricción de que ningún sabor se puede repetir, entonces el número de postres posibles es

$$C(5,3) = \binom{5}{3} = 10$$

$\square$

Para demostrar la validez de la fórmula (10), sin pérdida de generalidad, consideremos las combinaciones con reemplazo de k elementos construidas a partir del conjunto inicial $\{1,\ldots,n\}$. Estas combinaciones pueden describirse como el conjunto de las secuencias $i_1 \ldots i_k$ de números seleccionados del conjunto inicial y satisfaciendo $i_1 \leq \ldots \leq i_k$.

A cada secuencia $i_1 \ldots i_k$ le asociamos la secuencia $(i_1+0)(i_1+1)\ldots(i_k+k-1)$ construida a partir de la anterior sumando a cada dígito consecutivamente el $0,1,\ldots k-1$, representado a continuación

$$\begin{array}{rcl}
\underbrace{1\ldots1}_{\text{k veces}} & \longmapsto & 1\,2\,3\ldots k-1\,k \\
\underbrace{1\ldots1}_{\text{k-1 veces}}\ 2 & \longmapsto & 1\,2\,3\ldots k-1\,k+1 \\
\underbrace{1\ldots1}_{\text{k-1 veces}}\ 3 & \longmapsto & 1\,2\,3\ldots k-1\,k+2 \\
\ldots & & \\
\underbrace{n\ldots n}_{\text{k veces}} & \longmapsto & n\,n+1\ldots n+k-1,
\end{array}$$

con lo cual el problema se convierte en el equivalente de hallar todos los subconjuntos de tamaño k de los primeros $n+k-1$ números naturales. Así el número total, $C_r(n,k)$, de combinaciones con reemplazo de k elementos construidas a partir de n elementos es

$$C_r(n,k) = \binom{n+k-1}{k}. \tag{11}$$

3.2.5. Particiones

Como vimos, una combinación de k elementos seleccionados sin reemplazo de un total de n puede ser vista como una partición del conjunto inicial en dos, una parte contiene k elementos y la otra $n-k$. Más generalmente, supongamos que a los n elementos los queremos asignar sin importar el orden a r grupos $G_1, \ldots G_r$ tal que para cada i el grupo G_i posea n_i elementos y $\sum_{i=1}^{r} n_i = n$. Por el principio de multiplicación hay

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-n_2-\ldots-n_{n_{r-1}}}{n_r} = \frac{n!}{n_1!\cdots n_r!} \tag{12}$$

asignaciones posibles. A (12) se la denota con $\binom{n}{n_1,\ldots,n_r}$ y se le llama **coeficiente multinomial**.

Ejemplo 13. *Cuatro estudiantes de Río Cuarto y doce estudiantes del sur de Córdoba se dividen en cuatro grupos con 4 personas por grupo. ¿Cuál es la probabilidad que al dividirse, en cada grupo haya un estudiante de Río Cuarto?*

Si denotamos con A el evento cuya probabilidad queremos hallar entonces por el principio de conteo y por la definición de coeficiente multinomial

$$\#A = 4!\frac{12!}{3!3!3!3!}.$$

Además como

$$\#\Omega = \binom{16}{4,4,4,4} = \frac{16!}{4!4!4!4!},$$

entonces

$$P(A) = \frac{\#A}{\#\Omega} = \frac{12 \cdot 8 \cdot 4}{15 \cdot 14 \cdot 13}$$

Un típico ejemplo de utilización números combinatorios es el de los anagramas.

Ejemplo 14. ***Anagramas***

¿Cuantas palabras pueden formarse cambiando de lugar las letras en la palabra "TATTOO"? Notar que hay seis posiciones a ser completadas con las letras. El problema puede ser visto como un problema de partición de las posiciones: asignar tres posiciones al "grupo de la letra T", dos posiciones al "grupo de la letra O" y, finalmente, una posición al "grupo de la letra A". Hay

$$\binom{6}{1,2,3} = \frac{6!}{1!2!3!} = 60$$

formas de hacerlo. En consecuencia, la probabilidad de elegir al azar la palabra "TATTOO" del total de letras es $1/60$.

Otra forma de calcular la probabilidad es distinguiendo todas las letras. Esto es, etiquetamos las letras de la palabra "TATTOO" y formamos el conjunto $\{T_1, A, T_2, T_3, O_1, O_2\}$. Luego, hay un total de 6! secuencias de letras que se pueden formar con las letras del conjunto y hay 3!2!1! formas de escribir la palabra "TATTOO" (por ejemplo "$T_1AT_2T_3O_1O_2$", "$T_2AT_1T_3O_1O_2$", etc). Así, la probabilidad de seleccionar la palabra "TATTOO" es

$$\frac{3!2!1!}{6!} = \frac{1}{60},$$

igual a la calculada arriba.

Este ejemplo muestra que de acuerdo a cómo se elija contar los cardinales del espacio muestral y del evento en cuestión pueden ser diferentes según la forma de conteo elegida. No obstante la probabilidades del evento en un espacio y en otro son las mismas.

4. Probabilidad condicional

Las probabilidades condicionales nos proveen un camino para razonar acerca del resultado de un experimento contanto con información parcial.

Definición 8. *Dado un espacio de probabilidad* $(\Omega, \mathcal{F}, P)$ *y dos eventos* A *y* B, *definimos la probabilidad condicional de* A *dado* B *como*

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{13}$$

provisto que $P(B) > 0$.

Dado un evento B fijo con $P(B) > 0$ es simple demostrar que la aplicación

$$\begin{array}{rcl} P_{|B} : \mathcal{F} & \to & [0,1] \\ A & \longmapsto & P_{|B}(A) = P(A|B) \end{array}$$

define una probabilidad.

Ejemplo 15. *Arrojamos dos dados. Dado que el primer dado mostró un tres ¿cuál es la probabilidad que la suma total de los resultados aparecidos en ambas caras exceda seis?*

Para responder a la pregunta denotemos con A el evento "que la suma exceda seis" y B el evento "que el primer dado muestre un 3". De este modo

$$P(A|B) = \frac{3/36}{6/36} = 1/2.$$

Ejemplo 16. *Si una familia posee dos hijos ¿cuál es la probabilidad de que ambos sean varones dado que al menos uno es varón?*

Para contestar a esta pregunta, consideremos $\Omega = \{MM, MV, VM, VV\}$ y asumimos que cada suceso elemental tiene probabilidad $1/4$ de ocurrir. Luego,

$$P\left(\{VV\} \,|\, \{MV\} \cup \{VM\} \cup \{VV\}\right) = \frac{1/4}{3/4} = 1/3.$$

Y si nos preguntamos ¿cuál es la probabilidad que ambos sean varones dado que el más joven es varón? entonces la respuesta es

$$P\left(\{VV\} \,|\, \{MV\} \cup \{VV\}\right) = \frac{1/4}{2/4} = 1/2.$$

Recordemos que una partición de Ω es una colección $\{B_i\}_{i\in I}$ de conjuntos no vacíos y disjuntos dos a dos tal que $\bigcup_{i\in I} B_i = \Omega$.

Lema 1. ***Ley de la probabilidad total***

Consideremos $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $\{B_i\}_{i=1}^n$ una partición finita del espacio muestral tal que $\forall i : P(B_i) > 0$. Entonces

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i).$$

Demostración: en clase.

Intuitivamente, este lema afirma que dada una partición del espacio muestral en un número finito de escenarios, la probabilidad de que un evento A ocurra es un promedio ponderado o pesado de su probabilidad condicional bajo cada escenario y en el que cada escenario es pesado por su propia probabilidad.

Ejemplo 17. *Consideremos dos urnas, la I y la II, con pelotas de colores. La urna I contiene dos de color blanco y 3 azules y la urna II contiene 3 pelotas blancas y 4 azules. Si el experimento consiste en "extraer al azar una pelota de la urna I, luego colocarla en la urna II y entonces extraer a la azar de la urna II una pelota y registrar su color" entonces ¿cuál es la probabilidad de que la segunda pelota extraída sea de color azul?.*

Para dar una respuesta a esta pregunta denotemos con A_2 : el evento "que la segunda pelota sea azul" y con A_1 : el evento "que la primera pelota sea azul". Por la Ley de la probabilidad total

$$P(A_2) = P(A_2|A_1)P(A_1) + P(A_2|A_1^c)P(A_1^c) = \frac{5}{8}\frac{3}{5} + \frac{1}{2}\frac{2}{5} = \frac{23}{40}$$

El teorema (o regla) de Bayes se interesa por la reversión del orden del condicionamiento de los eventos.

Teorema 1. *Consideremos un $(\Omega, \mathcal{F}, P)$ y $\{B_i\}_{i=1}^n$ una partición finita del espacio muestral tal que $\forall i : P(B_i) > 0$ y consideremos un evento arbitrario A. Entonces, para cualquier j:*

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^n P(A|B_i)P(B_i)}.$$

Demostración: ejercicio.

En la Figura 4 se representa un tipo de inferencia en la que está implícito el Teorema de Bayes. Un médico observa, a través de imágenes producidas por un tomógrafo, una sombra en determinado órgano, es el evento A. El quiere estimar la verosimilitud de tres eventos mutuamente exclusivos y que serían causas potenciales exhaustivas: la causa 1, B_1, es la presencia de un tumor maligno, la causa 2, B_2 se corresponde con un tumor benigno y la 3, B_3, otras causas específicas diferentes de la presencia de un tumor. Dado que el médico observa una sombra la regla de Bayes da las *probabilidades a posteriori de cada una de las causas* a partir de las *probabilidades a priori*. Por ejemplo, interesa determinar

$$P(\underbrace{\text{haya un tumor maligno}}_{B_1} | \underbrace{\text{se observó una mancha}}_{A})$$

a partir de (1) utilizando

$$\underbrace{P(A|B_i)P(B_i)}_{\text{probabilidades a priori}} \quad \text{y } P(B_i),\ i = 1, 2, 3$$

Ejemplo 18. *Se asume que un test de detección de un tipo de gripe es correcto el 95 % de las veces; es decir si la persona tiene la enfermedad, el resultado es positivo con probabilidad 0.95 y, si la persona no tiene la enfermedad, el test da negativo con probabilidad 0.95. Una persona seleccionada al azar de cierta población tiene una probabilidad de portar la enfermedad de 0.001. Si para esta persona el test da positivo, ¿cuál es la probabilidad de que esté enferma?*

Solución: ejercicio del Práctico 1.

Definición 9. *Dados dos eventos A y B diremos que son independientes si y sólo si $P(B|A) = P(B)$.*

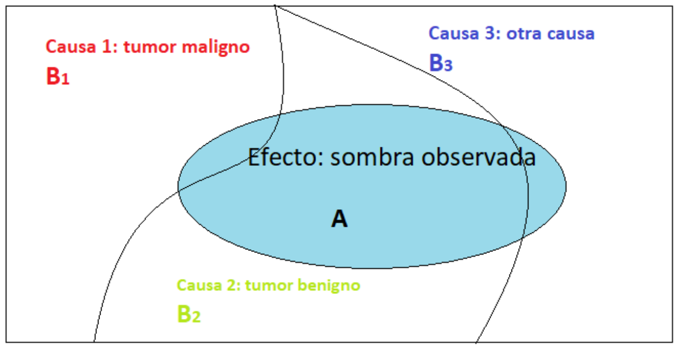

Figura 4: Ejemplo de un contexto de inferencia en el que está involucrado el Teorema de Bayes.

Notar que esta definición es equivalente a que se satisfaga $P(A|B) = P(A)$ y también a que $P(A \cap B) = P(A) \cdot P(B)$. Esto motiva la siguiente definición más general

Definición 10. *Sea $\{A_i\}_{i \in I}$ una familia de eventos. Diremos que*

a) los eventos de la familia son independientes si

$$\forall J \subseteq I, J \text{ finito} : P\left(\bigcap_{i=1}^{J} A_j\right) = \prod_{i=1}^{J} P\left(A_j\right)$$

b) los eventos son mutuamente independientes si

$$\forall i \neq j : P\left(A_i \cap A_j\right) = P\left(A_i\right) P\left(A_j\right)$$

Es claro que si una familia de eventos satisface a) entonces satisface b) pero la implicación recíproca no es válida en general.

Ejemplo 19. *Consideremos el siguiente espacio muestral el $(\Omega, \mathcal{F}, P)$ de equiprobabilidad donde*

$$\Omega = \{abc, acb, cab, cba, bca, bac, aaa, bbb, ccc\}$$

y consideremos el experimento aleatorio de seleccionar al azar un triplete de letras. Sea A_k el evento "que la k-ésima letra resulte una a" con $k = 1, 2, 3$. Entonces si $I = \{1, 2, 3\}$, los eventos de la familia $\{A_i\}_{i \in I}$ son mutuamente independientes pero no independientes.

5. Espacios no numerables

Hasta el momento hemos estudiado espacios en los que la totalidad de la probabilidad se acumula en un conjunto finito o bien en un conjunto a lo sumo numerable.

Consideremos ahora $\Omega = [a, b]$ con $a < b$ y para A contenido en Ω definamos

$$P(A) = \frac{\text{longitud de } A}{b - a}.$$

Usualmente $a = 0$ y $b = 1$. A estos espacios se los denomina también *uniformes* (no discretos) y equivale a "elegir al azar un punto de $\omega \in \Omega$".

Este esquema se generaliza a $\Omega \subset I\!R^n$ con $n \geq 2$ y donde la longitud es reemplazada por área, volumen, etc.

Nota *Un problema importante, que no será abordado en este curso, es cómo se define la σ-álgebra. Estaríamos tentados a considerar $\mathcal{F} = \mathcal{P}(\Omega)$ pero el inconveniente es que a no todo conjunto se le puede asignar una longitud, un área o un volumen. Son conjuntos no-medibles. Este problema se estudia en asignaturas como teoría de la medida.*

Ejemplo 20. *El 10 por ciento de la superficie de una esfera S está coloreada de azul y el resto es roja. Consideremos el experimento aleatorio de seleleccionar al azar un cubo inscripto en la esfera.*

Si definimos A_j el evento "que el j-ésimo vértice de un cubo seleccionado al azar sea azul" y teniendo en cuenta que $P(A_j) = \frac{1}{10}$ entonces, por la desigualdad de Bonferroni,

$$P\left(\bigcup_{i=1}^{8} A_j\right) \leq \sum_{i=1}^{8} P(A_j) = \frac{8}{10},$$

con lo cual al menos el 20 por ciento de los cubos seleccionados al azar tienen todos los vértices rojos.

Hemos mostrado que, independientemente de la manera en que los colores se distribuyen sobre la superficie de la esfera, es posible inscribir al cubo en la esfera con todos los vértices rojos.

En las unidades sucesivas surgirán muchos espacios no numerables y ejemplos de aplicaciones.

6. Probabilidad y frecuencia relativa

Una manera de definir la probabilidad de un evento es a través de la frecuencia relativa de aparición en sucesivas realizaciones del experimento muestral. Supongamos que un experimento cuyo espacio muestral es Ω es repetido exactamente bajo las mismas condiciones indefinidamente y definamos para un evento E, $n(E)$ el número de veces que E ocurrió en las primeras n repeticiones del experimento. Entonces se define la probabilidad del evento E como

$$P\,(E) = \lim_{n\to\infty} \frac{n(E)}{n}. \tag{14}$$

En otros términos, la probabilidad de un evento es el límite de la sucesión de frencuencias relativas de aparición.

Aunque esta definición de probabilidad es intuitiva tiene serios inconvenientes para ser sostenida. Nos preguntamos ¿cómo aseguramos que tal límite existe para una sucesión infinita de realizaciones del experimento? E incluso, si existiese ¿cómo sabemos que para otra sucesión infinita de realizaciones del experimento el límite sea el mismo?.

Estas objeciones hacen que sea necesario asumir una definición axiomática de probabilidad como la que dimos al inicio. No obstante mantendremos (14) como una constatación empírica de la definición de probabilidad.

7. Complementos: juego de bridge, bufandas en el guardarropa y búsqueda del tesoro.

Cerramos esta unidad compartiendo algunos problemas de resolución más compleja que los que hemos abordado hasta ahora. Como verán son problemas de enunciados muy simples.

Problema 1. *En el juego de bridge participan cuatro jugadores formando dos parejas adversarias. Se juega con un naipe que tiene 52 cartas agrupadas en cuatro palos: corazones ($\heartsuit$), rombos ($\diamondsuit$), tréboles ($\clubsuit$) y picas ($\spadesuit$) y, para cada palo las cartas son $A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K$. En cada mano se reparten las 52 cartas, 13 cada jugador.*

Si A es el evento "que en una mano cada jugador reciba un as", mostrar que la probabilidad de que A ocurra al menos una vez en 7 manos es de aproximadamente $1/2$.

Solución. Para el jugador 1 hay $\binom{52}{13}$ instancias de reparto de las 13 cartas, para el segundo $\binom{39}{13}$ y así sucesivamente. Luego, hay

$$\binom{52}{13,13,13,13} = \frac{52!}{(13!)^4}$$

formas de repartir las 52 cartas a los 4 jugadores por mano.

Por otro lado, hay 4! posibilidades de repartir los aces de tal modo que cada jugador reciba uno y $\binom{48}{12,12,12,12} = 48!/(12!)^4$ posibilidades de repartir las restantes $52-4$ cartas restantes entre los cuatro jugadores. Luego

$$P(A) = \frac{4!48!/(12!)^4}{52!/(13!)^4} \approx 1/10.$$

Sea ahora B_i el evento "que A ocurra por primera vez en la i-ésima mano" y B el evento "que A ocurra al menos una vez en 7 manos".

Notar que

$$B = \bigcup_{i=1}^{n} B_i \tag{15}$$

y, por la independencia entre manos se tiene que

$$\begin{aligned} P(B_i) &= P\left(A^c \text{ ocurra en la mano } 1, \ldots, A^c \text{ ocurra en la mano } i-1, A \text{ ocurra en la mano } i\right) \\ &= \left(P\left(A^c\right)\right)^{i-1} P(A). \end{aligned} \tag{16}$$

En consecuencia, de (15) y (16)

$$P(B) = \sum_{i=1}^{7} P(B_i) = \sum_{i=1}^{7} \left(\frac{9}{10}\right)^{i-1} \frac{1}{10} \approx \frac{1}{2}. \tag{17}$$

Problema 2. *En una fiesta N personas dejan sus bufandas en el guardarropa. Las bufandas se mezclan y cada persona, al finalizar la fiesta, selecciona al azar una bufanda. Nos preguntamos*

a) *¿Cuál es la probabilidad de que ninguna persona seleccione su propia bufanda?*

b) *¿Cuál es la probabilidad de que exactamente k personas seleccionen su propia bufanda?*

Solución.

a) En relación a la primera pregunta, el sentido común diría que cuanto más grande es el número de personas mayor es la probabilidad de que ninguna de ellas seleccione su propio sombrero y que esa probabilidad debería acercarse a 1 si N tiende a infinito. Veamos si es así.

   Cada resultado del experimento puede ser representado como una N-upla $(i_1, i_2, \ldots, i_N)$ en el que la j-ésima coordenada corresponde al número de bufanda elegida por la j-ésima persona. De ese modo, la j-ésima persona elige su bufanda correcta si y sólo si $i_j = j$.

   Denotemos ahora con E_i el evento "que la i-ésima persona seleccione correctamente la bufanda, con lo cual la probabilidad de que ninguna persona seleccione su propia bufanda es el complemento de la unión de los eventos $E_i, i = 1 \ldots, N$.

   Recordemos que

$$\begin{aligned} P\left(\bigcup_{i=1}^{N} E_i\right) &= \sum_{i=1}^{N} P(E_i) - \sum_{i_1<i_2} P(E_{i_1} \cap E_{i_2}) + \ldots \\ &\quad +(-1)^{n+1} \sum_{i_1<i_2<\ldots i_n} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_n}) \\ &\quad + \cdots + (-1)^{N+1} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_N}) \end{aligned} \tag{18}$$

   y computemos el término general $\displaystyle\sum_{i_1<i_2<\ldots i_n} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_n})$.

   Para ello notar que el evento $E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_n}$: "que las personas $i_1, i_2, \ldots, i_n$ seleccionen sus propias bufandas" puede ocurrir de $(N-n)(N-n-1)\ldots 3 \cdot 2 \cdot 1$ formas posibles ya que de las restantes $N-n$ personas la primera puede elegir de un total de $N-n$ bufandas,la segunda de $N-n-1$ bufandas y así sucesivamente. De este modo el cardinal del evento en cuestión es

$$\# E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_n} = (N-n)!.$$

   Así

$$\begin{aligned}
\sum_{i_1<i_2<\ldots i_n} P\left(E_{i_1}\cap E_{i_2}\cap\ldots\cap E_{i_n}\right) &= \sum_{i_1<i_2<\ldots i_n}\frac{(N-n)!}{N!}\\
&= \binom{N}{n}\frac{(N-n)!}{N!}\\
&= \frac{1}{n!}
\end{aligned}$$

ya que hay $\binom{N}{n}$ términos en la suma $\sum\limits_{i_1<i_2<\ldots i_n}$. Reemplazando en (18) obtenemos

$$P\left(\bigcup_{i=1}^{N} E_i\right) = 1-\frac{1}{2!}+\frac{1}{3!}-\cdots+(-1)^{N+1}\frac{1}{N!}$$

y por lo tanto

$$\begin{aligned}
P\left(\text{ninguna persona seleccione su bufanda }\right) &= 1-P\left(\bigcup_{i=1}^{N} E_i\right)\\
&= \frac{1}{2!}-\frac{1}{3!}-\cdots+(-1)^{N}\frac{1}{N!} \qquad (19)
\end{aligned}$$

Observar, por otro lado, que el polinomio de Taylor de orden N de la función $f(x)=\mathrm{e}^x$ alrededor de $a=0$ es

$$\begin{aligned}
P_N(x;a) &= \sum_{k=0}^{N}(x-a)^k f^{(k)}(a)/k!\\
&= \frac{(x-a)^0}{0!}\mathrm{e}^a-\frac{(x-a)^1}{1!}\mathrm{e}^a+\frac{(x-a)^2}{2!}\mathrm{e}^a+\ldots\frac{(x-a)^N}{N!}\mathrm{e}^a\\
&= 1-\frac{x^1}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^N}{N!}
\end{aligned}$$

donde en la segunda igualdad utilizamos que $f^{(k)}(a)$ (la derivada de orden k de f evaluada en a) es e^a y en la última igualdad hicimos $a=0$. Si evaluamos el polinomio en -1 obtenemos

$$P_N(-1;0)=1-1+\frac{1}{2!}-\frac{1}{3!}+\ldots(-1)^N\frac{1}{N!}$$

que coincide con (19). De este modo

$$\begin{aligned}
P\left(\text{ninguna persona seleccione su bufanda}\right) &= 1-1+\frac{1}{2!}-\frac{1}{3!}+\ldots(-1)^N\frac{1}{N!}\\
&= P_N(-1,0)\longrightarrow f(-1)=\mathrm{e}^{-1}, N\to\infty.
\end{aligned}$$

Así, concluimos que

$$P\,(\text{ninguna persona seleccione su bufanda }) \underbrace{\approx}_{N \text{ grande}} \mathrm{e}^{-1}$$

con lo cual nuestro sentido común falló ya que $\mathrm{e}^{-1} \approx 0.36788 \ll 1$.

b) Y ahora nos preguntamos ¿cuál es la probabilidad de que exactamente (o sólo) k personas seleccionen su propia bufanda? La respuesta es (dar los detalles): si E_k es el evento "que exactamente k personas reciban su bufanda" y Ω es el espacio muestral entonces

$$\begin{aligned} P(E_k) &= \frac{1}{\#\Omega}\#E_k \\ &= \frac{1}{N!}\underbrace{\binom{N}{k}(N-k)!\left(\frac{1}{2!}-\frac{1}{3!}+\cdots+(-1)^{N-k}\frac{1}{(N-k)!}\right)}_{\text{cardinal del evento exactamente } k \text{ reciban su bufanda}} \end{aligned}$$

Luego

$$\begin{aligned} P(E_k) &= \frac{1}{k!}\left(\frac{1}{2!}-\frac{1}{3!}+\cdots+(-1)^{N-k}\frac{1}{(N-k)!}\right) \\ &\approx \mathrm{e}^{-1}/k! \text{ si } N \text{ es grande} \end{aligned}$$

para $k = 0, 1, 2, \ldots$.

Problema 3. *Conocemos que un tesoro está localizado en una de dos ciudades, la primera y la segunda, con probabilidades β y $1-\beta$ respectivamente ($0 < \beta < 1$). Buscamos en la primera ciudad y si el tesoro está allí lo encontramos con probabilidad $p > 0$. Mostrar que el evento de no encontrar el tesoro en la primera ciudad sugiere que el tesoro está en la segunda ciudad.*

Solución: Para formalizar el problema consideremos los eventos siguientes

$$\begin{aligned} A &: \text{“que el tesoro esté en la segunda ciudad”} \\ B &: \text{“que no encontremos el tesoro en la primera ciudad”} \end{aligned}$$

y mostremos que

$$P(A|B) > P(A) \tag{20}$$

La desigualdad (20) afirma que el hecho de no haber encontrado el tesoro en la primera ciudad aumenta las chances de que esté en la segunda ciudad (es importante tener en cuenta que el hecho de que no encontremos el tesoro en la primera ciudad no implica que no esté allí).

Utilizando el Lema 1 o Ley de la probabilidad total obtenemos

$$P(B) = P(A^c)P(B|A^c) + P(A)P(B|A) = \beta \cdot (1-p) + (1-\beta) \cdot 1$$

y como además $A \subset B$ (y por ende $A \cap B = A$) podemos concluir que

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1-\beta}{\beta(1-p)+(1-\beta)} = \frac{1-\beta}{1-\beta p} > 1-\beta = P(A).$$

Referencias

Baron, M. (2006). *Probability and Statistics for Computer Scientists*. Chapman and Hall.

Devore, J. (2018). *Fundamentos de Probabilidad y Estadística*. Cengage.

Hacking, I. (1995). *El surgimiento de la probabilidad*. Gedisa.

Maronna, R. (2021). *Probabilidad y Estadística Elementales para Ciencias e IngenierÃa*. UBA y UNLP.

Ross, S. (2015). *A First Course in Probability*. Pearson.

Yohai, V. (2006). *Notas de Probabilidades y Estadistica*. Notas de Clase, UBA.