

# Clase 2: Demostraciones Combinatorias y Definición de Probabilidad.

Harvard Statistics 110: Probability.

## Resumen

Partimos estudiando una técnica para demostrar identidades combinatorias llamada “principio de doble conteo”. Luego nos centramos en la definición formal de las probabilidades.

## 1. Demostraciones combinatorias.

Muchas identidades que surgen en combinatoria pueden ser demostradas usando álgebra, pero ese método no permite entenderlas en un contexto de conteo. Es por ello que en esta sección veremos una técnica llamada “**principio de doble conteo**” que remarca ese aspecto.

El principio de doble conteo señala que **si dos fórmulas enumeran a un mismo conjunto, entonces son iguales**. Es decir, son dos generalizaciones que están contando lo mismo.

**Proposición 1.**

$$\binom{n}{r} = \binom{n}{n-r}$$

**Demostración.** Esta identidad la demostramos algebraicamente en la clase pasada. Trate-mos de comprenderla mejor usando el principio de doble conteo.

Suponga que estamos comprando una copa de helado y nos interesa saber cuántas podemos formar con  $r$  sabores distintos de los  $n$  disponibles. Una manera de enumerarlas es con el coeficiente binomial.

$$\binom{n}{r}$$

Podemos hacer el mismo conteo a partir de los  $n - r$  sabores que no se pueden escoger. El

número de formas de excluirlos es dado por:

$$\binom{n}{n-r}$$

Al contar los  $n - r$  sabores de helado no elegibles, también estamos enumerando a los  $r$  que sí se pueden optar como consecuencia de lo primero, que en total son  $\binom{n}{r}$ . Por lo tanto, es posible concluir que:

$$\binom{n}{r} = \binom{n}{n-r}$$

**Proposición 2.**

$$n \binom{n-1}{r-1} = r \binom{n}{r}$$

**Demostración.** En un curso de  $n$  estudiantes se están formando grupos de trabajo de  $r$  personas los que deben tener un/a representante. ¿Cuántos grupos posibles se pueden crear?

Podemos formar los grupos escogiendo primero a quiénes los representarán. Como solo es una de las  $n$  personas, estas se pueden elegir de  $n$  maneras distintas.

$$\binom{n}{1} = n$$

Al elegir quién liderará un grupo, significa que quedan  $r - 1$  vacantes en cada uno de ellos para  $n - 1$  estudiantes restantes. El número de maneras en que podemos escoger a estos últimos es:

$$\binom{n-1}{r-1}$$

Debido a que la creación de los grupos de trabajo consistió de un proceso de dos tareas, podemos usar la regla del producto para contar el número de maneras de formarlos como:

$$n \binom{n-1}{r-1}$$

También es posible partir creando los grupos antes de elegir quién los liderará. Como consisten de  $r$  de las  $n$  personas, podemos contar las distintas formas de constituirlos como:

$$\binom{n}{r}$$

Luego de formar los grupos, debemos escoger a quiénes los representarán. Como es un/a

estudiante de los  $r$  de cada uno, entonces se pueden elegir de  $r$  maneras distintas.

$$\binom{r}{1} = r$$

Puesto que este proceso para crear los grupos también consistió de dos tareas, por la regla del producto podemos contar las formas de construirlos como:

$$r \binom{n}{r}$$

Acabamos de ver dos formas de contar lo mismo: El número de maneras de conformar grupos de trabajo de  $r$  personas con un total de  $n$  de ellas. En consecuencia, se puede concluir que:

$$n \binom{n-1}{r-1} = r \binom{n}{r}$$

**Proposición 3.** *Identidad de Vandermonde.*

$$\binom{m+n}{r} = \sum_{i=0}^r \binom{m}{i} \binom{n}{r-i}$$

**Demostración.** En una caja hay  $n$  frutas y  $m$  verduras. ¿Cuántos grupos de  $r$  alimentos distintos podemos formar en total?

El tamaño total de todos los alimentos es  $n + m$ . Por lo tanto, una manera de contar estos grupos de  $r$  elementos es:

$$\binom{n+m}{r}$$

Otra forma de armar estos grupos es según la cantidad de frutas o verduras que elijamos. Si tomamos  $r$  de las primeras, entonces no es posible seleccionar alguna de las segundas. Por lo tanto, usando la regla del producto podemos enumerar las maneras de agruparlas como:

$$\binom{m}{0} \binom{n}{r}$$

Por otra parte, si optamos por 1 verdura solo será posible seleccionar  $r - 1$  frutas. En consecuencia, el número de grupos que es posible formar es:

$$\binom{m}{1} \binom{n}{r-1}$$

Se observa el patrón. Para una constante  $0 \leq i \leq r$ , cada grupo de frutas y verduras distintas

se puede formar de

$$\binom{m}{i} \binom{n}{r-i}$$

maneras.

El número de grupos de frutas y verduras que armemos según la cantidad elegida será siempre distinta entre sí. Por lo tanto, es posible usar la **regla de la adición** para contar el total de estas combinaciones como:

$$\sum_{i=0}^r \binom{m}{i} \binom{n}{r-i}$$

Dado a que esta suma está contando lo mismo que  $\binom{n+m}{r}$ , entonces:

$$\binom{m+n}{r} = \sum_{i=0}^r \binom{m}{i} \binom{n}{r-i}$$

Según Richard Brualdi en *Introductory Combinatorics* (2009: 28), la **regla de la adición** señala que si un conjunto  $S$  puede ser representado como una partición de  $\{A_k\}_{k=1}^n \subseteq S$ , entonces su tamaño está dado por:

$$|S| = \sum_{k=1}^n |A_k|, \text{ donde } A_i \cap A_j = \emptyset \text{ con } i \neq j$$

## 2. Definición formal de probabilidad.

Luego de haber estudiado la definición ingenua de probabilidad y los métodos de conteo que permiten calcularla, avancemos a su versión **formal**. Para ello, introduzcamos el concepto de espacio de probabilidad.

El **espacio de probabilidad** es un conjunto que consiste de tres elementos:

1. Un **espacio muestral**  $S$ .
2. Un **espacio de eventos**  $F$ , que es un conjunto de eventos  $A_i \subseteq S$ .
3. Una **función (o medida) de probabilidad**  $P$ .

Para que una función  $P$  sea catalogada como **de probabilidad**, en la década de 1930 el matemático Andrei Kolmogorov estableció que debe cumplir con los siguientes **axiomas**<sup>1</sup>:

1.  $P(A) \geq 0$ ,  $\forall A \in F$  con  $P(A) \in \mathbb{R}$ .

---

<sup>1</sup>También llamados “axiomas de probabilidad” o “axiomas de Kolmogorov”.

2.  $P(S) = 1$

3. Para  $\{A_k\}_{k=1}^{\infty} \subseteq S$ ,

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k) \iff A_i \cap A_j = \emptyset \text{ cuando } i \neq j$$

Por lo tanto, esta definición establece los **requisitos** que debe cumplir cualquier función para ser una de probabilidad. No restringe a una en particular como ocurre con la que catalogamos como “ingenua” en la clase anterior. Algo interesante es que esta última satisface los tres axiomas, de manera que sigue siendo útil en su contexto.

En probabilidades, cuando dos (o más) eventos  $A_1$  y  $A_2$  son **disjuntos** (i.e, si  $A_1 \cap A_2 = \emptyset$ ) también se dice que son **mutuamente excluyentes**, que es cuando no pueden ocurrir al mismo tiempo. De ahora en adelante, ambos términos serán usados de manera intercambiada.

## 2.1. Propiedades de la función de probabilidad.

De los axiomas de probabilidad que acabamos de ver se extraen las siguientes propiedades.

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

**Demostración.** Sea  $\{A_i\}_{i=1}^{\infty}$  una sucesión de eventos donde  $A_i = \emptyset$ . Puesto que  $\emptyset \cap \emptyset = \emptyset$ , todos los  $A_i$  son disjuntos entre sí. De este modo, por el axioma 3 se tiene que:

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

Como  $A_i = \emptyset$  para  $i = 1, 2, \dots$ , entonces  $\bigcup_{i=1}^{\infty} A_i = \emptyset$ . Por lo tanto,

$$P(\emptyset) = \sum_{i=1}^{\infty} P(\emptyset)$$

Para que la igualdad de arriba sea correcta, debemos buscar un número cuya serie resulte en sí misma. El único número real que cumple con esta característica es el cero. En consecuencia,

$$P(\emptyset) = 0 \quad (\text{Q. E. D})$$

**Propiedad 2.** Para todo  $A \subseteq S$ ,  $P(A^c) = 1 - P(A)$ .

**Demostración.** Esta propiedad la usamos en la clase anterior para resolver el problema del

cumpleaños. Ahora demostrémosla usando lo aprendido en esta sección.

Como  $A^c = \{x \mid x \notin A\}$ , entonces  $A \cap A^c = \emptyset$ . Además,  $A^c \subseteq S$  porque  $A \subseteq S$ . Todo esto conlleva a que  $A \cup A^c = S$ . Ya que es posible, al aplicar el axioma 3 se obtiene lo siguiente:

$$P(A \cup A^c) = P(S) = P(A) + P(A^c)$$

El axioma 2 señala que  $P(S) = 1$ , lo que implica que  $1 = P(A) + P(A^c)$ . Por lo tanto,

$$P(A^c) = 1 - P(A) \quad (\text{Q. E. D})$$

**Propiedad 3.** Sean  $A, B \subseteq S$ . Si  $A \subseteq B$ , entonces  $P(A) \leq P(B)$ .

**Demostración.** Ya que  $A \subseteq B$ , podemos definir que:

$$B = A \cup \{B \cap A^c\}$$

$A \cap \{B \cap A^c\} = \emptyset$ , por lo tanto es posible aplicar el axioma 3 como:

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

Por el axioma 1,  $P(B \cap A^c) \geq 0$ . Esto permite concluir de la igualdad de arriba que cuando  $A \subseteq B$ , entonces:

$$P(A) \leq P(B) \quad (\text{Q. E. D})$$

**Propiedad 4.** Para cualquier  $A \subseteq S$ ,  $0 \leq P(A) \leq 1$ .

**Demostración.** Mediante la propiedad 3 se puede establecer que

$$P(A) \leq P(S)$$

El axioma 1 señala que  $P(A) \geq 0$  y el axioma 2 que  $P(S) = 1$ . Por consiguiente,

$$0 \leq P(A) \leq 1 \quad (\text{Q. E. D})$$

**Propiedad 5.** Sean  $A, B \subseteq S$ .  $P(B \cap A^c) = P(B) - P(A \cap B)$ .

**Demostración.** Asumamos que  $A$  y  $B$  tienen elementos en común. Podemos definir a  $B$  como:

$$B = \{A \cap B\} \cup \{B \cap A^c\}$$

Si cambiamos la operación del lado derecho a una intersección, entonces  $\{A \cap B\} \cap \{B \cap A^c\} =$

$\emptyset$ . Así, por el axioma 3 tenemos que:

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

Dado que  $B = \{A \cap B\} \cup \{B \cap A^c\}$ , entonces  $P(B) = P(\{A \cap B\} \cup \{B \cap A^c\})$ . Al reemplazar en la igualdad de arriba, se obtiene:

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

Que es lo mismo a:

$$P(B \cap A^c) = P(B) - P(A \cap B) \quad (\text{Q. E. D})$$

**Propiedad 6.**  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ , con  $A, B \subseteq S$ .

**Demostración.** Al igual que en la demostración anterior, establezcamos que  $A$  y  $B$  no son eventos disjuntos. Es posible definir su unión como<sup>2</sup>:

$$A \cup B = A \cup \{B \cap A^c\}$$

Debido a que  $A$  y  $\{B \cap A^c\}$  son eventos disjuntos, se puede aplicar el axioma 3 como:

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

En la propiedad 5 vimos que  $P(B \cap A^c) = P(B) - P(A \cap B)$ . Por lo tanto,

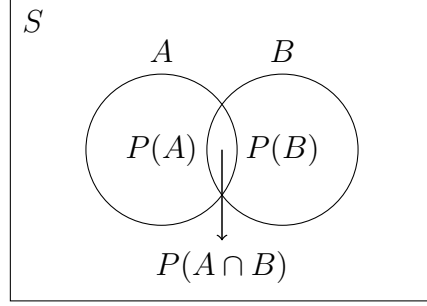
$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (\text{Q. E. D})$$

La propiedad 6 puede ser entendida como una versión más general del axioma 3, donde debemos prevenir calcular dos veces a  $P(A \cap B)$  cuando  $A$  y  $B$  no son mutuamente excluyentes.

A continuación tenemos un diagrama de Venn de los eventos  $A$  y  $B$ , con  $A \cap B \neq \emptyset$ , donde  $P(A)$  y  $P(B)$  son sus áreas.

---

<sup>2</sup>No confundir con lo realizado en la propiedad 3. En esa ocasión,  $A \subseteq B$  y acá no es el caso.



Como se observa arriba, cuando  $A \cap B \neq \emptyset$ , parte de  $P(A)$  ya está en  $P(B)$  y viceversa. Esta área es  $P(A \cap B)$ . Si calculamos a  $P(A \cup B)$  solo como la adición entre  $P(A)$  y  $P(B)$ , el resultado será erróneo porque estaremos duplicando la suma de la intersección. Para evitar este problema, tenemos que restarlo en esta operación.

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

En cambio, cuando  $A \cap B = \emptyset$ , no debemos lidiar con sumar dos veces a  $P(A \cap B)$  en la adición entre  $P(A)$  y  $P(B)$  porque  $P(A \cap B) = P(\emptyset) = 0$ . Esto conlleva a que  $P(A \cup B) = P(A) + P(B) - 0 = P(A) + P(B)$ . Es decir, al axioma 3.

## 2.2. Principio de inclusión-exclusión.

Es posible generalizar la propiedad 6 de las probabilidades a una cantidad  $n$  de eventos. La suma resultante se conoce como el **principio de inclusión-exclusión**.

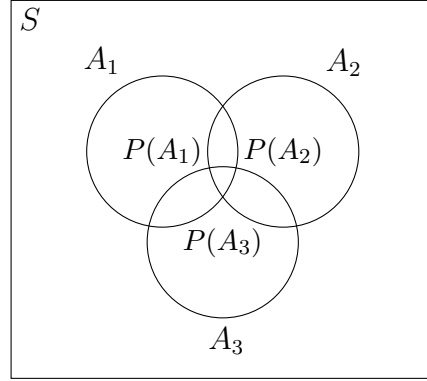
Sea  $\{A_i\}_{i=1}^n$  una sucesión finita de eventos. El principio de inclusión-exclusión señala que:

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) \\ &\quad - \sum_{1 \leq i < j < k < l \leq n} P(A_i \cap A_j \cap A_k \cap A_l) + \dots + (-1)^{n+1} \cdot P\left(\bigcap_{i=1}^n A_i\right) \end{aligned}$$

Como se observa en la fórmula del principio de inclusión-exclusión, hay una alternancia entre sumas y restas. Usemos diagramas de Venn para entender la intuición detrás de este hecho.

Sean  $A_1$ ,  $A_2$  y  $A_3$  tres eventos que no son mutuamente excluyentes, con  $P(A_1)$ ,  $P(A_2)$  y  $P(A_3)$  siendo sus probabilidades.





Si calculamos la probabilidad de la unión entre  $A_1$ ,  $A_2$  y  $A_3$  al igual como lo expresado en la propiedad 6 de las probabilidades, obtenemos lo siguiente:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - [P(A_1 \cap A_2) + P(A_2 \cap A_3) + P(A_1 \cap A_3)]$$

En la resta del lado derecho, efectivamente estamos evitando contar dos veces las áreas  $P(A_1 \cap A_2)$ ,  $P(A_2 \cap A_3)$  y  $P(A_3 \cap A_1)$ . Sin embargo, esa operación excluye totalmente a  $P(A_1 \cap A_2 \cap A_3)$ . Para que vuelva a ser considerada, debemos sumarla.

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - [P(A_1 \cap A_2) + P(A_2 \cap A_3) + P(A_1 \cap A_3)] + P(A_1 \cap A_2 \cap A_3)$$

La intuición del principio de inclusión-exclusión nos muestra, también, la razón de su nombre: Para calcular la probabilidad de la unión de eventos no disjuntos, debemos quitar posibles duplicaciones y añadir aquellas áreas que queden afuera producto de ese proceso.

A continuación veremos un importante ejemplo en probabilidades conocido como el Problema de De Montmort, el que puede ser resuelto a partir del principio de inclusión-exclusión.

### 2.3. Problema de coincidencia de De Montmort.

En 1708, el matemático francés Pierre Rémond de Montmort planteó el siguiente problema:

**Ejemplo.** Suponga que tiene un mazo de  $n$  cartas enumeradas como  $1, 2, \dots, n$  y que, luego de barajarlas, va sacando una tras otra sin devolución mientras cuenta en voz alta “ $1, 2, \dots, n$ ”. Calcule la probabilidad de que coincida con decir el número de la carta sacada.

**Solución.** El conteo en voz alta nos da información sobre la posición de cada carta en el mazo. Como la coincidencia de mencionar el número de la carta sacada puede ocurrir en

cualquiera de las  $n$  que hay en la baraja, podemos definir una sucesión finita de eventos  $\{A_i\}_{i=1}^n$  como:

$$A_i = \text{“La } i\text{-ésima carta sacada tiene escrito el número } i\text{”}$$

En ese sentido, lo que interesa es la ocurrencia de **al menos uno** de los eventos. Por lo tanto, para resolver este problema buscaremos calcular la probabilidad de la unión de ellos.

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

Los eventos  $A_i$  **no son mutuamente excluyentes** porque en el mismo mazo barajado es posible, por ejemplo, que la segunda carta lleve el número 2, la novena el número 9, etc. Es decir, todos pueden ocurrir en el mismo experimento. Por lo tanto, calcularemos a  $P(\bigcup_{i=1}^n A_i)$  mediante el **principio de inclusión-exclusión**.

El espacio muestral es finito y consiste de las  $n$  permutaciones de las  $n$  cartas, que en total son  $P(n, n) = n!$  ya que el mazo fue barajado. Por este mismo hecho se puede asumir que todos sus elementos son igualmente probables de ocurrir, lo que permite usar la definición ingenua de las probabilidades.

Comencemos con las probabilidades para cada evento  $A_i$ . La cantidad de resultados favorables para ellos es  $P(n-1, n-1) = (n-1)!$ , porque debemos contar las permutaciones de las cartas, pero manteniendo fijo una de ellas para una posición. Por lo tanto,

$$P(A_i) = \frac{(n-1)!}{n!}$$

Luego necesitamos ver las probabilidades de las intersecciones. Para el caso de dos eventos  $A_i$  y  $A_j$ , para  $1 \leq i < j \leq n$ , el número de ocurrencia de ambos al mismo tiempo es  $P(n-2, n-2) = (n-2)!$ , ya que debemos mantener ocupadas las posiciones  $i$  y  $j$  para las cartas de los mismo números.

$$P(A_i \cap A_j) = \frac{(n-2)!}{n!}$$

De este modo, para tres eventos  $A_i$ ,  $A_j$  y  $A_k$ , la probabilidad de su intersección es:

$$P(A_i \cap A_j \cap A_k) = \frac{(n-3)!}{n!}$$

Así, podemos generalizar para la intersección de los  $n$  eventos  $A_i$  que:

$$P\left(\bigcap_{i=1}^n A_i\right) = \frac{1}{n!}$$

Con esto en cuenta, escribamos la fórmula para la unión de los eventos  $A_i$ .

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n \frac{(n-1)!}{n!} - \sum_{i < j} \frac{(n-2)!}{n!} + \sum_{i < j < k} \frac{(n-3)!}{n!} - \sum_{i < j < k < l} \frac{(n-4)!}{n!} \\ &\quad + \dots + (-1)^{n+1} \cdot \frac{1}{n!} \end{aligned}$$

Como se observa, en la primera suma estamos contando  $n$  veces la misma expresión. Para la segunda, es la fracción para 2 grupos de los  $n$ . En la tercera, corresponde a 3 de  $n$ , etc. Teniendo esto en cuenta, podemos hacer el siguiente reemplazo.

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= n \cdot \frac{(n-1)!}{n!} - \binom{n}{2} \cdot \frac{(n-2)!}{n!} + \binom{n}{3} \cdot \frac{(n-3)!}{n!} - \binom{n}{4} \cdot \frac{(n-4)!}{n!} \\ &\quad + \dots + (-1)^{n+1} \cdot \frac{1}{n!} \end{aligned}$$

Al simplificar, obtenemos lo siguiente:

$$P\left(\bigcup_{i=1}^n A_i\right) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \dots + (-1)^{n+1} \cdot \frac{1}{n!}$$

Considere la serie de Maclaurin de  $\exp(x)$ .

$$\exp(x) = \sum_{i=1}^{\infty} \frac{x^i}{i!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Para  $x = -1$ ,

$$\exp(-1) = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} \dots$$

Si  $n \rightarrow \infty$  en  $P(\bigcup_{i=1}^n A_i)$ , la suma desde su segundo término en adelante converge a la serie de Maclaurin de  $-\exp(-1)$ . Al reemplazarlo, obtenemos que:

$$\lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n A_i\right) = 1 - \exp(-1) \approx 0.632$$

Es interesante que  $P(\bigcup_{i=1}^n A_i)$  no converja a 1 o a 0 mientras crece la cantidad de cartas.

Esto se puede explicar por el hecho de que aumenta el número de ubicaciones posibles para cada carta, lo que hace disminuir y estabilizar la probabilidad de que hayan coincidencias con sus números a  $\approx 0.632$  ( $\approx 63.2\%$ ).

## 2.4. Reconociendo operaciones entre conjuntos.

En ocasiones es difícil detectar las operaciones entre conjuntos que se están realizando en un problema matemático. Por esta razón, en la siguiente tabla se resumen algunas expresiones que las vinculan a ellas, para facilitar ese proceso.

Frase	Operación
$A$ o $B$	$A \cup B$
$A$ y $B$	$A \cap B$
No $A$	$A^c$
$A$ o $B$ , pero no ambos	$(A \cap B^c) \cup (A^c \cap B)$
Al menos uno de los $A_1, A_2, \dots$	$A_1 \cup A_2 \cup \dots$
Todos los $A_1, A_2, \dots$	$A_1 \cap A_2 \cap \dots$
$A$ implica $B$	$A \subseteq B$