Capítulo 5

Las estructuras básicas de la Combinatoria

Cuando contamos y enumeramos, cuando hacemos combinatoria, aparecen con frecuencia unas estructuras básicas que merecen nombres y análisis específicos. Así, cada vez que, tras meditar detenidamente sobre una cierta cuestión combinatoria, identifiquemos los objetos de interés (por ejemplo, particiones de un conjunto en bloques no vacíos), sabremos que la respuesta será la correspondiente familia de números (para el caso, números de Stirling). Este capítulo será una suerte de muestrario de estas estructuras básicas: explicaremos los contextos en las que aparecen y aprenderemos a contar cuántas de ellas hay, en cada caso.

Prepárese el lector, pues, para una excursión, casi taxonómica, en la que irá descubriendo paulatinamente las principales familias, algunos de los géneros, y hasta alguna que otra especie, que pueblan el hábitat combinatorio: subconjuntos, permutaciones, particiones en bloques, en ciclos, de enteros... En realidad, ya iniciamos esta excursión en el capítulo anterior, en el que presentamos diversos tipos de *listas* (con y sin repetición, circulares, etc.). Recomendamos al lector que tenga presentes los resultados que allí obtuvimos, porque volverán a aparecer en los argumentos combinatorios que siguen.

5.1. Subconjuntos. Coeficientes binómicos

Sea A un conjunto con $n \geq 1$ elementos. Para las cuestiones que nos interesan, los nombres de los elementos de A no desempeñan papel alguno, así que, por concreción y conveniencia, supondremos que A es el conjunto $\{1, \ldots, n\}$. Queremos saber cuántos subconjuntos (sin repetición) distintos de tamaño k podemos extraer de él. Por ejemplo, si $n = 4$ y $k = 2$, hay seis 2-subconjuntos: $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$ y $\{3, 4\}$. Llamemos

$$C(n, k) = \#\{k\text{-subconjuntos extraídos de un conjunto de } n \text{ elementos}\}.$$

El parámetro k, que indica el tamaño de los subconjunto que nos interesan, puede tomar en principio los valores $k = 0, 1, 2, \ldots$ Pero si $k > n$, entonces $C(n, k) = 0$, pues resulta imposible construir un subconjunto con más elementos de los que tiene el conjunto de partida.

Nos situamos en el rango de interés: para un cierto $n \geq 1$ y para cada $0 \leq k \leq n$, y argumentamos de forma indirecta, relacionando listas con conjuntos. El número de *listas* de longitud k y sin repetición que podemos formar con n símbolos es

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

como vimos en la sección 2.2.1. Consideremos ahora una cualquiera de ellas: en sus k posiciones tiene símbolos distintos, así que podremos reordenarla de $k!$ maneras distintas. La observación clave es que cada una de estas $k!$ posibles reordenaciones da lugar, si hacemos caso omiso del orden de presentación de los símbolos, a un *único* conjunto de tamaño k.

Por tanto, podemos relacionar cada $k!$ listas (en las que sí es relevante el orden) con un solo conjunto (en el que el orden no es relevante). Se trata de una aplicación $k!$ a 1, entre el conjunto de las k-listas sin repetición formadas con símbolos $\{1,\dots,n\}$ y la colección de k-subconjuntos de $\{1,\dots,n\}$, que nos permite concluir (apelando al lema 2.1.2) que, para cada $n \geq 1$,

$$C(n,k) = \begin{cases} \dfrac{n!}{k!\,(n-k)!} & \text{si } 0 \leq k \leq n; \\ 0 & \text{si } k > n. \end{cases}$$

con el convenio habitual de que $0! = 1$. Obsérvese que de este análisis combinatorio se deduce que la fracción $\frac{n!}{k!\,(n-k)!}$ es un entero, algo que no es sencillo comprobar algebraicamente[1].

Es tradicional designar al cociente de factoriales de la fórmula anterior con el siguiente símbolo:

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

que se lee "n sobre k". Estos números, los **coeficientes binómicos**[2], están en principio definidos para cada entero positivo n y para cada entero $0 \leq k \leq n$. Note, lector, que

$$\binom{n}{n} = 1 \qquad \text{y} \qquad \binom{n}{0} = 1\,,$$

usando de nuevo que $0! = 1$. Unos valores consistentes con la interpretación combinatoria de los $C(n,k)$: por un lado, como el único conjunto con n elementos que se puede extraer de un conjunto de n elementos es el propio conjunto, tenemos que $C(n,n) = 1$; por otro, $C(n,0) = 1$, pues el conjunto vacío $\emptyset$ es el único conjunto de tamaño cero incluido en $\{1,\dots,n\}$,

En lo que sigue, para evitar una proliferación innecesaria de símbolos, y salvo esporádicas reapariciones de la notación $C(n,k)$, diremos que, dado $n \geq 1$, el número $\binom{n}{k}$ cuenta

- *el número de subconjuntos de tamaño k que podemos extraer del conjunto* $\{1,\dots,n\}$;
- o en lenguaje alternativo, *el número de maneras distintas de seleccionar k símbolos de entre una colección de n*;

con el *convenio adicional*, del que haremos uso con profusión en lo que sigue, de que $\binom{n}{k} = 0$ para los valores de k fuera del rango $0 \leq k \leq n$.

[1] Quizás quiera el lector revisar ahora el ejemplo 1.2.12. O al revés, utilizar que $\binom{n}{k}$ sea un entero para dar una prueba alternativa del enunciado que allí se proponía.

[2] Llamados así porque aparecen en el desarrollo del binomio de Newton (véase la sección 5.1.2).

Veamos una primera (y simpática) aplicación de estos números a un problema de recuento.

EJEMPLO 5.1.3 ***La ley de Murphy y los calcetines***. *Tenemos diez pares de calcetines (distintos) y desaparecen seis calcetines.*

Observe el lector que el número de pares íntegros que pueden quedar está entre 4 y 7. Suponemos aquí que la desaparición de calcetines es fruto y obra del puro azar, y nos interesa saber, por ejemplo, cuál es la *probabilidad* de que queden únicamente 4 pares útiles, que sería el peor caso; o cuán probable es que queden 7, que sería el mejor. Para analizar la cuestión, etiquetamos los calcetines según la pareja a la que pertenezcan y si son derecho o izquierdo:

$$D_1, I_1,\ D_2, I_2,\ \dots, D_{10}, I_{10}\,.$$

Como en el ejemplo 2.2.6, apelaremos a la noción de probabilidad como cociente del número de casos favorables entre el total de casos. Estos casos ("desapariciones") posibles son $\binom{20}{6} = 38\,760$, pues hay que elegir los 6 calcetines que desaparecen de entre los 20.

Para que queden 7 pares útiles, la única posibilidad es que hayan desaparecido tres pares completos. Estos casos "favorables" son $\binom{10}{3} = 120$, pues basta elegir esas tres parejas (de entre las 10 que hay). Así que, si llamamos p_7 a la probabilidad de que queden 7 pares útiles,

$$p_7 = \frac{\binom{10}{3}}{\binom{20}{6}} = \frac{1}{323} \approx 0.31\,\%\,.$$

Si quedan 4 parejas útiles, necesariamente habrá desaparecido un calcetín de cada una de las otras seis. Contamos posibilidades decidiendo primero qué cuatro parejas quedan íntegras ($\binom{10}{4}$ posibilidades), y luego qué seis calcetines desaparecen: uno de la primera pareja que queda (2 posibilidades), otro de la segunda (otras 2 posibilidades), etc. En total, 2^6. Así que

$$p_4 = \frac{\binom{10}{4}\,2^6}{\binom{20}{6}} = \frac{112}{323} \approx 34.67\,\%\,.$$

Así que es ¡112 veces! más probable estar en el caso de menor número (cuatro) de parejas supervivientes que en aquél en el que nos quedan siete. Ley de Murphy, por si lo dudaba.

Aunque en realidad lo más probable es que nos queden cinco pares útiles. Analizamos este caso con el siguiente argumento: elegimos los 5 pares que quedan completos ($\binom{10}{5}$ maneras) y, de los restantes 5 pares, uno ha de desaparecer completo (5 posibilidades), y del resto hemos de elegir qué calcetín desaparece (2^4 maneras). En total,

$$p_5 = \frac{\binom{10}{5}\,5\times 2^4}{\binom{20}{6}} = \frac{168}{323} \approx 52.01\,\%\,.$$

El último caso (6 pares útiles) requeriría elegir los 6 pares íntegros, y de los otros cuatro, elegir los dos que desaparecen completos y tomar un calcetín de las otras dos parejas:

$$p_6 = \frac{\binom{10}{6}\binom{4}{2}\,2^2}{\binom{20}{6}} = \frac{42}{323} \approx 13.00\,\%\,.$$

Nótese el permanente uso de la regla del producto que hemos hecho en estos cálculos. Compruébese también que $p_4 + p_5 + p_6 + p_7 = 1$. ♣

5.1.1. Propiedades de los coeficientes binómicos

Los coeficientes binómicos gozan de un buen número de propiedades que usaremos a menudo en lo que sigue. Vamos ahora a enunciar las más básicas, acompañadas de las correspondientes demostraciones, algunas de corte algebraico, y otras de corte combinatorio. El lector irá descubriendo otras muchas propiedades en los siguientes apartados, y en la colección de ejercicios de esta sección.

A. Simetría y suma total

La primera propiedad que nos disponemos a demostrar es la siguiente:

$$\boxed{\binom{n}{k} = \binom{n}{n-k}} \qquad \text{para } n \geq 1 \text{ y } 0 \leq k \leq n.$$

La razón por la que la denominamos propiedad de "simetría" resultará evidente cuando, unas páginas más adelante, dispongamos los coeficientes binómicos en un triángulo.

La prueba algebraica es directa:

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

Pero es más interesante confirmar esta propiedad utilizando argumentos combinatorios. Llamemos

$$\Gamma_1 = \left\{ \begin{array}{c} \text{subconjuntos de tamaño } k \\ \text{extraídos de } \{1,\dots,n\} \end{array} \right\} \quad \text{y} \quad \Gamma_2 = \left\{ \begin{array}{c} \text{subconjuntos de tamaño } n-k \\ \text{extraídos de } \{1,\dots,n\} \end{array} \right\}.$$

A cada subconjunto B de tamaño k (es decir, incluido en Γ_1) le podemos asociar el subconjunto de tamaño $n-k$ (que estará incluido en Γ_2) formado por todos los elementos de $\{1,\dots,n\}$ que no están en B, esto es, $\{1,\dots,n\} \setminus B$.

Construimos así una aplicación entre los conjuntos Γ_1 y Γ_2 que, como podrá comprobar sin dificultad el lector, es biyectiva. De lo que se deduce que ambos conjuntos han de tener el mismo tamaño. Es decir, que las cantidades

$$|\Gamma_1| = \binom{n}{k} \qquad \text{y} \qquad |\Gamma_2| = \binom{n}{n-k}$$

han de ser iguales.

La segunda propiedad da cuenta del valor de la *suma de todos* los coeficientes binómicos de índice superior fijo: para cada $n \geq 1$,

$$\boxed{\sum_{k=0}^{n} \binom{n}{k} = 2^n}$$

La prueba algebraica directa, manipulando los factoriales que aparecen en la suma, parece una tarea laboriosa. Sin embargo, es sencillo organizar una prueba algebraica por inducción, y animamos al lector a completarla (ejercicio 5.1.1). Veremos también una demostración alternativa, utilizando el teorema del binomio, en la sección 5.1.2.

La prueba combinatoria es más instructiva. Sabemos que $\binom{n}{k}$ cuenta, para cada $0 \le k \le n$, el número de subconjuntos de tamaño k que podemos extraer del conjunto $A = \{1, \dots, n\}$. Llamemos Γ al conjunto de todos los posibles subconjuntos de A. Sabemos (recuérdese el ejemplo 2.2.3) que el conjunto Γ tiene tamaño 2^n. Definamos además

$$\begin{aligned}
\Gamma_0 &= \{\text{subconjuntos de } A \text{ de tamaño } 0\} \\
\Gamma_1 &= \{\text{subconjuntos de } A \text{ de tamaño } 1\} \\
&\ \vdots \\
\Gamma_n &= \{\text{subconjuntos de } A \text{ de tamaño } n\}.
\end{aligned}$$

Es evidente que los conjuntos $\{\Gamma_0, \Gamma_1, \dots, \Gamma_n\}$ constituyen una partición de Γ. Así que, con la regla de la suma, concluimos que

$$|\Gamma| = \sum_{j=0}^{n} |\Gamma_j| \quad ; \qquad \text{es decir,} \qquad 2^n = \sum_{j=0}^{n} \binom{n}{j}.$$

Podemos reinterpretar esta relación en términos de listas. Recuerde el lector la *identificación entre subconjuntos y listas de ceros y unos* del ejemplo 2.2.3: dar un subconjunto de $\{1, \dots, n\}$ es exactamente lo mismo que dar una n-lista con ceros y unos (un 1 en la posición j-ésima de la lista significa que el elemento j está en el subconjunto; un 0, que no). En la fórmula anterior, a la izquierda, aparece 2^n, el número total de n-listas con ceros y unos; y a la derecha, esas listas aparecen clasificadas en función del número de unos que contengan, pues

$$\binom{n}{k} = \#\left\{ \begin{array}{c} \text{listas de longitud } n \text{ formadas con ceros} \\ \text{o unos que tienen exactamente } k \text{ unos} \end{array} \right\},$$

dado que para dar una lista con k unos basta decidir qué posiciones llevan esos unos, y hay $\binom{n}{k}$ maneras de hacerlo.

B. Regla de recurrencia

La siguiente recurrencia para los coeficientes binómicos nos permitirá calcularlos eficientemente. La registramos, para futuras referencias, como:

Lema 5.1.1 (Recurrencia de coeficientes binómicos) *Dado un entero $n \ge 2$, y para cada entero k, con $1 \le k \le n-1$,*

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

Observe, lector, que el rango en el enunciado del lema se ha escogido para que ninguno de los enteros que aparecen dentro de los coeficientes binómicos sea negativo, o el inferior mayor que el superior, o... La prueba algebraica de esta recursión es sencilla y directa:

$$\begin{aligned}
\binom{n-1}{k} + \binom{n-1}{k-1} &= \frac{(n-1)!}{k!(n-1-k)!} + \frac{(n-1)!}{(k-1)!(n-k)!} = \frac{n-k}{n}\frac{n!}{k!(n-k)!} + \frac{k}{n}\frac{n!}{k!(n-k)!} \\
&= \binom{n}{k}\Big[\frac{n-k}{n} + \frac{k}{n}\Big] = \binom{n}{k}.
\end{aligned}$$

Para la comprobación combinatoria, construimos la siguiente partición:

$$\left\{\begin{matrix}\text{Subconjuntos de}\\ \text{tamaño } k \text{ extraídos}\\ \text{de } \{1,\dots,n\}\end{matrix}\right\} = \left\{\begin{matrix}\text{Subconjuntos de tamaño } k\\ \text{extraídos de } \{1,\dots,n\}\\ \text{que contienen al elemento } n\end{matrix}\right\} \cup \left\{\begin{matrix}\text{Subconjuntos de tamaño } k\\ \text{extraídos de } \{1,\dots,n\}\\ \text{que } \textit{no} \text{ contienen al elemento } n\end{matrix}\right\}$$

Por supuesto, la elección del elemento n, el último, para este proceso es totalmente arbitraria (podíamos haber elegido, por ejemplo, el primero). La colección (de conjuntos) de la izquierda, ya lo sabemos, tiene tamaño $\binom{n}{k}$, y la regla de la suma nos permite escribir que

$$\begin{aligned}\binom{n}{k} &= \#\left\{\begin{matrix}\text{Subconjuntos de tamaño } k\\ \text{extraídos de } \{1,\dots,n\}\\ \text{que contienen al elemento } n\end{matrix}\right\} + \#\left\{\begin{matrix}\text{Subconjuntos de tamaño } k\\ \text{extraídos de } \{1,\dots,n\}\\ \text{que } \textit{no} \text{ contienen al elemento } n\end{matrix}\right\}\\ &= \#\left\{\begin{matrix}\text{Subconjuntos de tamaño } k-1\\ \text{extraídos de } \{1,\dots,n-1\}\end{matrix}\right\} + \#\left\{\begin{matrix}\text{Subconjuntos de tamaño } k\\ \text{extraídos de } \{1,\dots,n-1\}\end{matrix}\right\}\\ &= \binom{n-1}{k-1} + \binom{n-1}{k}.\end{aligned}$$

La penúltima igualdad es la clave del argumento, y está basada en un par de biyecciones. Para el primer término argumentamos así: para construir todos los subconjuntos de tamaño k con los elementos $\{1,\dots,n\}$ que contengan al elemento n, basta decidir quiénes son sus $k-1$ acompañantes, es decir, basta elegir $k-1$ elementos del conjunto $\{1,\dots,n-1\}$. Para el segundo término, como los subconjuntos que estamos considerando en este caso no contienen a n, tendremos que escoger los k elementos de entre los del conjunto $\{1,\dots,n-1\}$.

Figura 5.1: Tartaglia

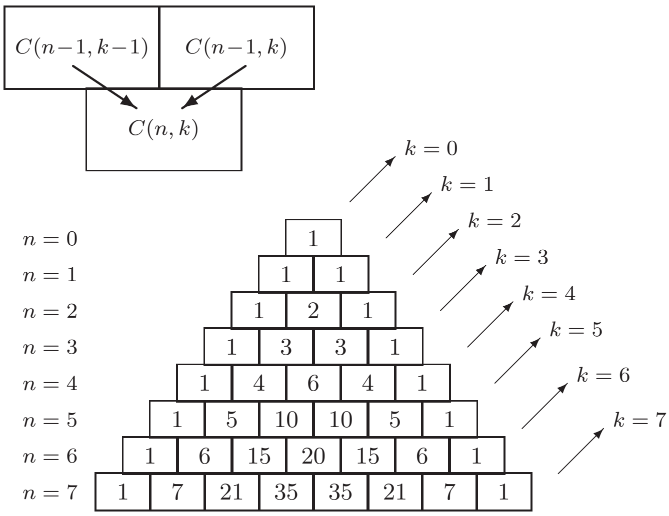
La combinación de esta regla de recurrencia con lo que vamos a llamar los *valores frontera*, a saber,

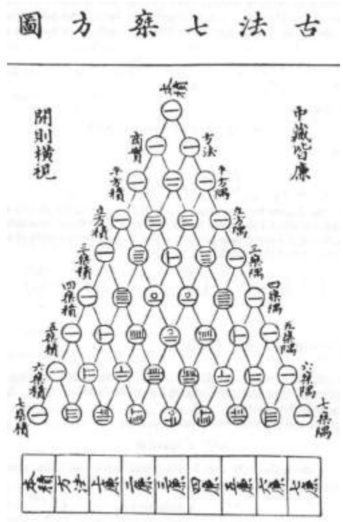
$$\binom{n}{0} = 1 \qquad \text{y} \qquad \binom{n}{n} = 1\,, \qquad \text{para cada } n \geq 1,$$

nos permite calcular y codificar los valores de todos los coeficientes binómicos. Para ello, es costumbre utilizar el llamado **triángulo de Pascal-Tartaglia**[3]: un triángulo formado por casillas que van etiquetadas con dos parámetros, n y k. El parámetro n etiqueta los "pisos" del triángulo, empezando en $n=0$, mientras que el parámetro k, por su parte, marcará la coordenada de las sucesivas diagonales, de nuevo de $k=0$ en adelante. En la casilla de coordenadas n y k situamos el número $C(n,k)$, o indistintamente $\binom{n}{k}$.

[3]A veces sólo triángulo de Tartaglia, a veces sólo triángulo de Pascal. Niccolo Fontana (1499-1557) es más conocido como Tartaglia (tartamudo; o tartaja, más catizo y un punto despectivo). Parece ser que de pequeño fue gravemente herido en la cara por las tropas francesas que ocupaban Brescia, su localidad natal, y que de aquel episodio conservó una gran cicatriz en el rostro y ciertas dificultades para hablar. Tradujo y publicó numerosas obras matemáticas clásicas, como los *Elementos* de Euclides y algunos tratados de Arquímedes. Consiguió, entre otros logros, obtener una fórmula para la resolución de la ecuación cúbica (véase una de las notas al pie en el apartado G1 de la sección 1.1).

Para que todo cuadre, conviene definir $C(0,0) = 1$; una definición consistente con la fórmula de los factoriales, aunque sin aparente significado combinatorio[4]. Los valores en los bordes (las fronteras) del triángulo son del tipo $\binom{n}{0}$ o $\binom{n}{n}$, y por tanto valen siempre 1. Las casillas interiores se rellenan siguiendo la ecuación de recurrencia, cuya interpretación gráfica aparece debajo de estas líneas, a la izquierda: cada coeficiente binómico se obtiene sumando los valores de los dos inmediatamente superiores. Con esta regla, y los valores en los bordes, podemos completar el triángulo, tal como hacemos a continuación (hasta $n = 7$):





Si, por ejemplo, el lector dirige su mirada al piso $n = 6$ y diagonal $k = 2$, encontrará el valor de $C(6,2)$, o bien $\binom{6}{2}$, que es 15. Los coeficientes binómicos eran ya conocidos, en mayor o menor grado, siglos antes de Tartaglia o Pascal, por ejemplo, por Ibn Ezra[5] y por Levi ben Gerson[6], entre los siglos XII y XIV; aunque los matemáticos árabes y chinos[7] también manejaban estos números, como se aprecia en la figura[8] de la derecha, del siglo XIII. El triángulo de Tartaglia, casi un icono cabalista, contiene, veladamente, muchas sucesiones de números de interés[9].

[4] ¿O si lo tiene, en realidad? Alardee el lector de sus habilidades para el razonamiento escolástico-bizantino y justifique combinatoriamente que $C(0,0) = 1$.

[5] Abraham Ben Meir Ibn Ezra (1089-1164), matemático, exégeta y astrólogo. . . ¡español!, nacido en Tudela y muerto en Calahorra, estaba interesado y sabía calcular los coeficientes binómicos con $n = 7$. Porque siete *eran* los cuerpos celestes: el Sol, la Luna, Mercurio, Venus, Marte, Júpiter y Saturno. Como buen astrólogo, a Ibn Ezra le preocupaba saber de cuántas formas se puede estar simultáneamente bajo varios de esos signos. El lector podrá encontrar más información sobre este personaje en el artículo *La astrología combinatoria del rabino Ibn Ezra*, de Doron Zeilberger (La Gaceta de la Real Sociedad Matemática Española **1** (1998), no. 3).

[6] Parece que fue el rabino Levi Ben Gerson (1288-1344) el primero en dar una expresión explícita de los coeficientes binómicos.

[7] Véase la referencia a Al-Karaji del apartado 1.2.2. Aprovechamos aquí para sugerir al lector una excelente referencia en historia de las matemáticas: *Mathematics and its history* (Springer-Verlag, 1991), de J. Stillwell.

[8] La tabla acaba en $n = 8$. Observe el lector los símbolos. ¿Sería capaz de escribir los números del 1 al 99 en estos caracteres chinos?

[9] Por ejemplo, los números triangulares $T_n = \binom{n+1}{2}$ (véase el ejercicio 1.2.3), que están en la diagonal $k = 2$. Véase también la sección 8.3.5. Buen momento para recomendar, de nuevo, la lectura de *El diablo de los números* de Erzensberger.

La regla de recurrencia anterior permite escribir $\binom{n}{k}$ (cuyo índice superior es n) en términos de la suma de dos coeficientes binómicos cuyos índices superiores son $n-1$. Si repetimos el procedimiento, pero para los dos nuevos coeficientes binómicos, llegamos a

$$\begin{aligned}\binom{n}{k} &= \binom{n-1}{k} + \binom{n-1}{k-1} = \left[\binom{n-2}{k} + \binom{n-2}{k-1}\right] + \left[\binom{n-2}{k-1} + \binom{n-2}{k-2}\right] \\ &= \binom{n-2}{k} + 2\binom{n-2}{k-1} + \binom{n-2}{k-2} = \binom{2}{0}\binom{n-2}{k} + \binom{2}{1}\binom{n-2}{k-1} + \binom{2}{2}\binom{n-2}{k-2}.\end{aligned}$$

Ahora $\binom{n}{k}$ se escribe como suma de coeficientes binómicos de índice superior $n-2$. Nótese que los números que los acompañan pueden ser escritos, a su vez, como coeficientes binómicos.

Podríamos iterar el proceso, pero los cálculos serían algo engorrosos, y vale la pena argumentar en general, combinatoriamente. Queremos escribir $\binom{n}{k}$ en términos de coeficientes binómicos cuyo índice superior sea, por ejemplo, $n-l$. Primero declaramos del “primer tipo” a l elementos de entre $\{1,\dots,n\}$, por ejemplo los l primeros, marcando los restantes $n-l$ como del “segundo tipo”. Ahora clasificamos los k-subconjuntos dependiendo del número j (con $0 \le j \le k$) de elementos del primer tipo que contengan, en la siguiente partición:

$$\left\{\begin{array}{c}k\text{-subconjuntos extraídos}\\ \text{de } \{1,\dots,n\}\end{array}\right\} = \bigcup_{j=0}^{k} \left\{\begin{array}{c}k\text{-subconjuntos extraídos de } \{1,\dots,n\}\\ \text{con } j \text{ elementos del primer tipo}\end{array}\right\}.$$

Calculamos el tamaño del conjunto con etiqueta j seleccionando primero qué j elementos del primer tipo están en nuestro subconjunto (hay $\binom{l}{j}$ posibilidades); y luego seleccionando los $k-j$ elementos del segundo tipo que contiene el subconjunto (hay $\binom{n-l}{k-j}$ posibilidades). Aplicando las reglas de la suma y del producto llegamos a:

Lema 5.1.2 (Fórmula de Vandermonde) *Dados enteros $n \ge 1$, $1 \le l \le n$ y $1 \le k \le n$,*

$$\binom{n}{k} = \sum_{j=0}^{k} \binom{l}{j}\binom{n-l}{k-j}.$$

Observe, lector, que en esta suma de Vandermonde[10] el índice j llega, en realidad, hasta el mínimo de l y k. Esto se justifica combinatoriamente con que j ha de ser menor o igual que k, el tamaño del subconjunto, y también menor o igual que l, el número de elementos del primer tipo; y algebraicamente, observando que el primer coeficiente obliga a que $j \le l$, y que en el segundo se requiere que $k-j \ge 0$. También, por cierto, en este segundo coeficiente binómico debe ocurrir que $n-l \ge k-j$, lo que nos da una condición adicional $j \ge k-n+l$. Así que el rango de sumación resulta ser $\text{máx}(0, k-n+l) \le j \le \text{mín}(l,k)$. Aparatosas condiciones

[10]El nombre del matemático y músico francés Alexandre-Théophile Vandermonde (1735-1796) ha quedado asociado al ubicuo *determinante de Vandermonde* (véase el ejemplo 8.1.9). Pese a que Vandermonde fue uno de los pioneros de la teoría de los determinantes, parece ser que jamás consideró el que lleva su nombre. Hay quien conoce esta fórmula como de Zhu–Vandermonde, añadiendo al matemático chino Zhu Shijie (Chu Shi-Chieh), autor del texto “Siyuan Yujian” (El espejo de jade de las cuatro incógnitas), de 1303, del que procede la preciosa ilustración del triángulo de Pascal de la página anterior.

que hemos obviado en esta fórmula (como haremos a menudo en lo que sigue) apelando al ya mencionado convenio por el que declaramos $\binom{n}{k} = 0$ si k no está en el rango $0 \leq k \leq n$, o si aparecen índices negativos. El caso $l = 1$ de la fórmula de Vandermonde es, por cierto, la recurrencia básica del lema 5.1.1.

Finalizamos este apartado con un par de identidades binómicas adicionales a las que nos referiremos genéricamente como "reglas de absorción", en particular homenaje al apartado a) que sigue a continuación en el que, observe, lector, el factor k se absorbe[11] en el coeficiente binómico, dando como resultado la inmediata expulsión del factor n.

Lema 5.1.3 (Reglas de absorción)

$$\text{a)}\quad k\binom{n}{k} = n\binom{n-1}{k-1}; \qquad \text{b)}\quad \binom{n}{k}\binom{k}{l} = \binom{n}{l}\binom{n-l}{k-l}.$$

DEMOSTRACIÓN. Nótese que a) es el caso particular de b) en el que tomamos $l = 1$. La comprobación algebraica de ambas identidades, usando las fórmulas de los coeficientes binómicos con factoriales, es simple rutina que dejamos como entretenimiento para el lector. Una prueba combinatoria de a) va como sigue: tenemos n personas, de las que hemos de seleccionar a k, y una de éstas tendrá un papel especial; digamos que es el "jefe". El procedimiento de recuento en el que primero elegimos a las k personas, y luego de entre estas k, al jefe, conduce al miembro de la izquierda de a). El procedimiento alternativo en el que primero elegimos al jefe (n posibles maneras) y luego a los $k-1$ acompañantes (de entre las $n-1$ personas restantes), nos da el de la derecha.

La prueba de b) es análoga, pero considerando l jefes. ■

C. Cálculo, tamaño y forma de los coeficientes binómicos

En este apartado discutimos tres cuestiones sobre los números binómicos, a saber: cómo se calculan *eficientemente*, qué tamaño tienen, y también si la sucesión de números $(\binom{n}{k})_{k=0}^{n}$, con n fijo, tiene siempre el aspecto que se puede apreciar en la figura de esta página.

C1. Cálculo. La primera cuestión puede sorprender. ¿Cómo que "cómo se calculan"? ¡Pero si ya disponemos de una *fórmula*!, la que viene escrita en términos de factoriales:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 3\cdot 2\cdot 1}.$$

¡Ah!, pero hemos incluido la coletilla "eficientemente". Para empezar, recuerde el lector la discusión de la sección 3.7, $n!$ es un número enorme en cuanto n sea medianamente grande. Si por ejemplo quisiéramos calcular un modesto $\binom{50}{25}$ con la fórmula anterior, necesitaríamos evaluar los productos de 50 a 26, y de 25 a 1, para luego dividirlos como dice la fórmula anterior. Incluso cancelando los factores que aparecen simultáneamente en numerador y denominador, es una tarea que puede poner en apuros a cualquier ordenador. Sin embargo, la

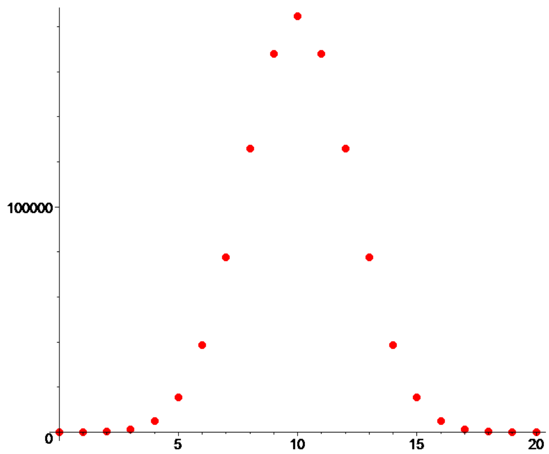
[11] Dejamos que el lector pergeñe una interpretación digestiva análoga para el apartado b).

aplicación reiterada de la recurrencia suele permitir calcularlos de una manera más eficiente, pues sólo requiere un cierto número de sumas. Por cierto, la respuesta es

$$\binom{50}{25} = 126\,410\,606\,437\,752,$$

algo más de 126 billones[12]. Recuerde, lector: número de maneras distintas de seleccionar a 25 personas de un grupo de 50. Una barbaridad. Por cierto, a la vista de estos descomunales números, quizás valga la pena disponer de alguna técnica para *estimar* el tamaño de un coeficiente binómico general. Vea el lector el apartado C3.

C2. Forma. Pese a todo, la fórmula de los factoriales contiene información muy relevante, como vamos a ver en este apartado y en el siguiente, al analizar la forma y el tamaño de los coeficientes binómicos.



Fijamos n. Como ya sabemos, por la propiedad A de esta misma sección, la lista de coeficientes $\binom{n}{k}$ es simétrica con respecto al elemento central (o centrales, si n es impar). Y de hecho los valores de esta lista van creciendo, de izquierda a derecha, hasta llegar al centro (o centros), donde se alcanza el máximo. Es decir,

$$\max_{k=0,\dots,n} \left\{ \binom{n}{k} \right\} = \binom{n}{\lfloor n/2 \rfloor} = \binom{n}{\lceil n/2 \rceil}$$

Y a partir de entonces empiezan a decrecer. Los dos coeficientes binómicos que aparecen a la derecha de la expresión anterior son el mismo si n es par. Véase, por ejemplo, la gráfica con los valores de los sucesivos coeficientes binómicos en el caso $n = 20$.

La comprobación de este crecimiento/decrecimiento es directa. Para empezar, por simetría, basta argumentar en el rango $0 \le k \le \lfloor n/2 \rfloor$ (la mitad "izquierda"). Pongamos, por comodidad, que n es par, digamos $n = 2m$. Queremos probar que

$$\binom{2m}{k-1} < \binom{2m}{k}$$

siempre que $k \le m$. Es decir, que la sucesión es creciente hasta $k = m$. Pero obsérvese que, usando la fórmula de factoriales,

$$\binom{2m}{k-1} < \binom{2m}{k} \iff \frac{(2m)!}{(k-1)!\,(2m-k+1)!} < \frac{(2m)!}{k!\,(2m-k)!} \iff \frac{1}{2m-k+1} < \frac{1}{k},$$

condición que justamente se cumple en el rango seleccionado, $k \le m$. Dejamos que el lector complete los detalles del caso en que n sea impar.

Es buen momento para introducir unas cuantas nociones sobre sucesiones de números que aparecerán repetidamente en lo que sigue.

[12] Sí, gracias por separar el número en miles, que si no...

Definición 5.1.4 Consideremos una sucesión finita $(a_0, a_1, \ldots, a_n)$. Decimos que la sucesión

a) es **creciente** si $a_j \leq a_{j+1}$ para cada $j = 0, \ldots, n-1$; y es **decreciente** si $a_j \geq a_{j+1}$ para cada $j = 0, \ldots, n-1$;

b) es **convexa** si
$$a_k \leq \frac{a_{k-1} + a_{k+1}}{2} \quad \text{para todo } k = 1, \ldots, n-1;$$
y es **cóncava** si las desigualdades anteriores llevan un "$\geq$";

c) es **unimodal** si para un cierto j se cumple que
$$a_0 \leq a_1 \leq \cdots \leq \boxed{a_j} \geq a_{j+1} \geq \cdots \geq a_n.$$

d) es **logcóncava** si
$$a_k^2 \geq a_{k-1} \cdot a_{k+1} \quad \text{para todo } k = 1, \ldots, n-1;$$
y es **logconvexa** si las desigualdades anteriores llevan un "$\leq$".

En cualquiera de estas definiciones se añade el adverbio "estrictamente" cuando los símbolos "$\leq$" o "$\geq$" se sustituyen por "$<$" y "$>$", respectivamente.

La definición a) es natural, y la b) es la análoga, en el contexto de las sucesiones, de la convexidad/concavidad de funciones que vimos en la definición 3.3.3. Como allí, el significado geométrico de la convexidad es que el valor a_k queda por debajo del punto medio entre a_{k-1} y a_{k+1} (de la "cuerda" que une esos dos puntos extremos).

Las otras dos requieren cierta explicación. Una sucesión unimodal[13] crece hasta un cierto punto, su máximo, en el que empieza a decrecer. Podría haber varios máximos, pero en ese caso deberían ser todos iguales y consecutivos. Como casos extremos, son unimodales las sucesiones crecientes (el caso $j = n$ en la definición), las decrecientes ($j = 0$), e incluso las constantes. Como acabamos de ver, los coeficientes binómicos $\binom{n}{k}$ son, para n fijo, una sucesión unimodal, con un único máximo (si n es par), o dos consecutivos e iguales (n impar).

La logconcavidad, como iremos comprobando, es una propiedad bastante habitual en las sucesiones de corte combinatorio. Digamos que, como ocurre en buena parte de los casos de interés, los a_k son todos positivos; entonces, tomando logaritmos en la desigualdad que define la logconcavidad obtenemos que
$$\ln(a_k) \geq \frac{\ln(a_{k-1}) + \ln(a_{k+1})}{2} \quad \text{para todo } k = 1, \ldots, n-1,$$
que nos dice que la sucesión de logaritmos es *cóncava*. Y de ahí, claro, el nombre.

Siguiendo con a_k positivos, podemos reescribir la condición de logconcavidad como
$$\frac{a_k}{a_{k-1}} \geq \frac{a_{k+1}}{a_k} \quad \text{para todo } k = 1, \ldots, n-1,$$

[13] Que tiene una sola *moda*. El nombre proviene del contexto probabilista: cuando los a_j representan probabilidades de que se tomen ciertos valores, la moda es el valor(es) más probable(s).

que nos dice que la sucesión de razones de términos consecutivos a_k/a_{k-1} es decreciente. Caben tres posibilidades: si todos los a_k/a_{k-1} son > 1, entonces la sucesión (a_n) es decreciente; si son todos < 1, la sucesión será creciente. El caso en el que las razones comienzan siendo > 1 para acabar siendo < 1 se corresponde con el habitual de la unimodalidad. Hemos probado:

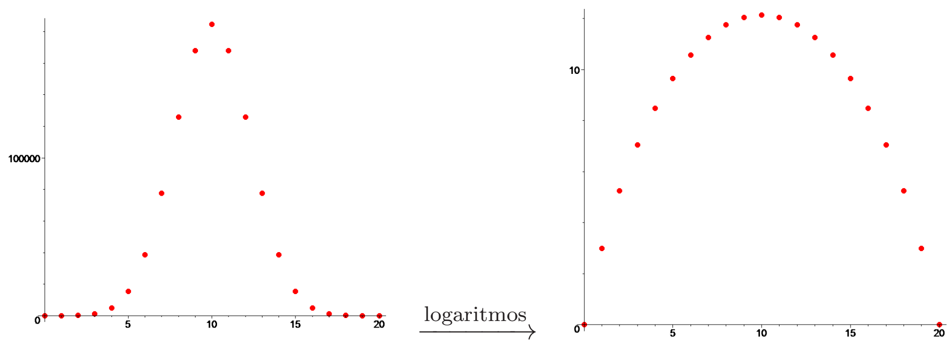
Lema 5.1.5 *Si $(a_0, a_1, \dots, a_n)$ es una sucesión logcóncava de números positivos, entonces es también unimodal.*

Esta observación abre una nueva vía, que será bien útil más adelante, para probar la unimodalidad de sucesiones positivas a través de la comprobación de la logconcavidad. Más aún, un análisis más detallado nos dice que si la sucesión es estrictamente logcóncava (si las desigualdades llevan un $<$ en lugar de un $\leq$), entonces la sucesión sólo puede tener uno o dos máximos; este último caso ocurrirá cuando $a_j/a_{j-1} = 1$ en cierto j.

Para cada entero n positivo, la sucesión $\left(\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}\right)$, los números que habitan un piso dado del triángulo de Tartaglia, es *unimodal*, como ya hemos visto antes, y de hecho (es estrictamente) *logcóncava*, como se comprueba con el siguiente argumento directo: para todo $k = 1, \dots, n-1$,

$$\binom{n}{k}^2 = \frac{n!}{k!\,(n-k)!}\,\frac{n!}{k!\,(n-k)!} = \frac{n-k+1}{k}\,\frac{n!}{(k-1)!\,(n-k+1)!}\,\frac{k+1}{n-k}\,\frac{n!}{(k+1)!\,(n-k-1)!}$$
$$= \frac{(k+1)(n-k+1)}{k(n-k)}\binom{n}{k-1}\binom{n}{k+1} = \left(1+\frac{n+1}{k(n-k)}\right)\binom{n}{k-1}\binom{n}{k+1} > \binom{n}{k-1}\binom{n}{k+1}.$$

Obsérvese que para n fijo, la sucesión de los $\left(\binom{n}{k}\right)$ para $0 \leq k \leq n$ no es cóncava (ni convexa), pero la sucesión de sus logaritmos sí que es cóncava. Véase, en las figuras que siguen, el caso $n = 20$:



Mas adelante (teorema 7.2.22 de Newton) seremos capaces de probar la logconcavidad (y por ende la unimodalidad) de ciertas sucesiones incluso sin disponer de fórmulas explícitas para sus términos.

C3. Estimaciones de tamaño. Nos interesamos, por último, por la cuestión de determinar el orden de magnitud de un coeficiente binómico genérico $\binom{n}{k}$. Bueno, dirá el lector, el tamaño es muy variable, pues por ejemplo $\binom{50}{0} = \binom{50}{50} = 1$, mientras que, como hemos visto antes, $\binom{50}{25}$ es gigantesco.

En este apartado vamos a estimar el tamaño del coeficiente $\binom{n}{\lfloor n/2 \rfloor}$ que ocupa la posición central (o centrales) de una fila del triángulo de Pascal–Tartaglia, y que como hemos visto antes es el mayor de toda la fila. En la sección 5.1.2 siguiente, y con ayuda del teorema del binomio, daremos estimaciones más ajustadas para otros coeficientes binómicos.

Por comodidad de cálculo, y para no andar distinguiendo el caso par del impar, vamos a analizar únicamente coeficientes del tipo $\binom{2n}{n}$, y dejamos al lector que, al terminar, analice el caso impar y unifique resultados usando la notación de suelos y techos.

Para empezar, la suma de todos los coeficientes $\binom{2n}{j}$ vale $2^{2n} = 4^n$, así que cada uno de ellos, y en particular $\binom{2n}{n}$, es menor que 4^n. Como además $\binom{2n}{n}$ es el mayor de todos ellos, tenemos que

$$4^n = \sum_{j=0}^{2n} \binom{2n}{j} < \sum_{j=0}^{2n} \binom{2n}{n} = (2n+1)\binom{2n}{n}.$$

Reuniendo las dos estimaciones, concluimos que

$$\frac{4^n}{2n+1} < \binom{2n}{n} < 4^n,$$

una estimación válida para todo $n \geq 1$, pero no tan precisa como sería deseable. Por ejemplo, para $n = 25$ la cota de la derecha es casi 9 veces más grande que el verdadero valor de $\binom{50}{25}$, mientras que la cota de la izquierda es casi 6 veces menor.

En realidad, cuando n es grande, $\binom{2n}{n}$ no es del orden de $4^n/n$, como parece apuntar la cota de la izquierda, ni como 4^n, como dictaría la de la derecha. Con ayuda de la precisa estimación asintótica de Stirling (sección 3.7) podemos afinar mucho más y afirmar que

$$\binom{2n}{n} = \frac{(2n)!}{n!\,n!} \sim \frac{(2n)^{2n}\, e^{-2n} \sqrt{4\,\pi\, n}}{[n^n\, e^{-n} \sqrt{2\,\pi\, n}]^2} = \frac{1}{\sqrt{\pi n}}\, 4^n \qquad \text{cuando } n \to \infty.$$

Recuerde el lector, de la sección 3.1, que la estimación asintótica anterior quiere decir que

$$\lim_{n\to\infty} \frac{\binom{2n}{n}}{4^n/\sqrt{\pi n}} = 1.$$

De manera que el orden de magnitud (asintótico) correcto lleva, además del factor 4^n, un $\sqrt{n}$ en el denominador, acompañado por cierto de un sumamente misterioso factor $\sqrt{\pi}$. La comparación numérica con nuestro ejemplo habitual da

$$\binom{50}{25} = 126\,410\,606\,437\,752\,, \qquad \text{mientras que} \qquad \frac{4^{25}}{\sqrt{25\,\pi}} \approx 127\,044\,199\,911\,599,$$

que es una aproximación asombrosa, pues aunque ambas cantidades difieren en algo más de medio billón[14], su cociente, que es lo que consideramos en estas estimaciones asintóticas, o mejor, su diferencia porcentual[15], es de apenas un 0.5 %.

[14] ¿Les parece poco?
[15] ¡Ah!, visto así. . .

5.1.2. Sobre el teorema del binomio

FIGURA 5.2: Newton

Nombramos[16] a los números $\binom{n}{k}$ como *coeficientes binómicos* porque aparecen en el desarrollo del *binomio de Newton*[17], esa regla que nos permite escribir (y calcular) el desarrollo de una potencia de una *suma de dos términos*. Una formulación habitual es la siguiente:

Lema 5.1.6 (Binomio de Newton) *Para cualquier natural $n \geq 1$ y cualquier número real x,*

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$$

Observe, lector, que la fórmula anterior es también válida para $n = 0$ por el convenio de que $\binom{0}{0} = 1$.

DEMOSTRACIÓN. Vamos a dar dos demostraciones de esta identidad: una primera por inducción en n, y la otra siguiendo un argumento combinatorio. El ejercicio 5.1.19 propone una prueba alternativa.

El caso $n = 1$ es evidente. Supongamos que la identidad es cierta para un cierto n (y todo número real x), y analicemos el caso $n+1$ como sigue:

$$\begin{aligned}
(1+x)^{n+1} &= (1+x)^n\,(1+x) = (1+x)\sum_{k=0}^{n}\binom{n}{k}x^k = \sum_{k=0}^{n}\binom{n}{k}x^k + \sum_{k=0}^{n}\binom{n}{k}x^{k+1} \\
&= \sum_{k=0}^{n}\binom{n}{k}x^k + \sum_{l=1}^{n+1}\binom{n}{l-1}x^l = \binom{n}{0}x^0 + \sum_{k=1}^{n}\left[\binom{n}{k}+\binom{n}{k-1}\right]x^k + \binom{n}{n}x^{n+1} \\
&= \binom{n+1}{0}x^0 + \sum_{k=1}^{n}\binom{n+1}{k}x^k + \binom{n+1}{n+1}x^{n+1} = \sum_{k=0}^{n+1}\binom{n+1}{k}x^k.
\end{aligned}$$

Observe, lector, cómo en el argumento hemos separado los términos primero y último, y que hemos usado la regla de recurrencia de los coeficientes binómicos y también sus valores frontera.

[16] No sólo es la razón del nombre, sino buena parte de su razón de ser, y desde luego su origen histórico, más ligado a desarrollos algebraicos que a cuestiones combinatorias.

[17] Sir Isaac Newton (1642-1727) es, quizás, el científico más famoso de todos los tiempos; sólo Einstein rivaliza con él en fama y gloria. En sus *Philophiae naturalis principia mathematica*, o simplemente *Principia*, de 1687, que son considerados como el más importante libro científico jamás escrito, estableció los principios básicos de la mecánica, la dinámica de fluidos, el movimiento ondulatorio, dedujo las leyes de Kepler del movimiento de planetas y cometas, etc. Y, claro, se le considera el inventor del cálculo diferencial e integral, junto con Leibniz. Mmm, ¿junto, al tiempo, independientemente?; ¿quién fue antes? Bueno, Newton y Leibniz mantuvieron unas, digamos, no muy elegantes discusiones públicas acerca del asunto. Lagrange decía de Newton que "[...] fue el más grande genio que ha existido", añadiendo que "[...] y también el más afortunado, dado que sólo se puede encontrar una vez un sistema que rija el mundo".

Teólogo, alquimista, inventor de máquinas, filósofo, profesor de la cátedra lucasiana en Cambridge, sucediendo a Barrow: un auténtico gigante. Aunque él mismo dejó dicho (en uno de sus escasos raptos de humildad) que "había visto un poco más lejos era porque estaba subido a hombros de gigantes", en reconocimiento al trabajo anterior de otros científicos.

Vamos a dar ahora una prueba alternativa del binomio en la que se conectan la definición combinatoria del coeficiente binómico $\binom{n}{k}$ (recuerde, lector: número de maneras de escoger k elementos de entre un conjunto de n elementos) con el papel que desempeñan en la fórmula del enunciado: coeficientes del polinomio que se obtiene al desarrollar $(1+x)^n$.

$$n \text{ factores} \begin{cases} (1+x) \\ (1+x) \\ (1+x) \\ \cdot \\ \cdot \\ \cdot \\ (1+x) \\ (1+x) \end{cases} \quad \begin{array}{l} \text{hay que escoger} \\ k \text{ veces la } x \end{array}$$

Fijamos $n \geq 1$. En el esquema de la izquierda hemos escrito n veces el factor $(1+x)$. Nos preguntamos cuántas veces aparecerá el término x^k al multiplicarlos todos, pues ése será el coeficiente de x^k. Pero para que aparezca x^k habrá que tomar k veces la x (y $n-k$ veces el 1, claro). Es decir, hay que elegir las k filas en las que tomamos la x. Lo que se puede hacer de $\binom{n}{k}$ maneras, certificando así la validez de la fórmula del binomio. ■

Como aviso para navegantes[18], señalamos que éste es un primer ejemplo de una *función generatriz*: la función $(1+x)^n$ "genera", al ser desarrollada en potencias de x, la sucesión de números

$$\left(\binom{n}{k}\right)_{k=0}^{\infty} = \left(\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \dots, \binom{n}{n}, 0, 0, 0 \dots\right).$$

Sobre estas cuestiones insistiremos, y mucho, en el capítulo 12 y siguientes.

Si reemplazamos x por x/y en la fórmula del binomio,

$$\left(1 + \frac{x}{y}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{x^k}{y^k} \quad \Longrightarrow \quad \left(\frac{y+x}{y}\right)^n = \sum_{k=0}^{n} \binom{n}{k} x^k \, y^{-k},$$

obtenemos, tras multiplicar por y^n, la siguiente generalización:

$$\boxed{(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k \, y^{n-k}}$$

que en principio, por el argumento utilizado, sería válida sólo para $y \neq 0$. Pero como para $y = 0$ afirma simplemente que $x^n = x^n$, concluimos que es válida para todo y. Quizás el lector quiera probar directamente esta fórmula con un argumento combinatorio como el de antes, escribiendo n veces el factor $(x+y)$, y luego…

¿Y para esto abren una sección?, preguntará el lector exigente: esto del binomio lo llevamos usando desde nuestra más tierna infancia, por ejemplo para escribir que

$$(a+b)^2 = a^2 + 2ab + b^2,$$

sin olvidarnos de ese término cruzado, o en el algo más aparatoso cálculo del cubo[19]:

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$$

[18]Que sin desmayo, viento en popa, a toda vela, surquen las páginas de este libro.

[19]Buen momento, lector, para que reconozca que en muy contadas ocasiones habrá ido más allá del cubo.

Pero no, la fórmula del binomio contiene mucha más información, como iremos comprobando en esta sección. Recomendamos en este punto la siguiente guía de lectura para el lector: el apartado A contiene un buen número de identidades para los coeficientes binómicos que se obtienen directamente de (o tras manipular levemente) la fórmula del binomio; aunque avisamos de que el (sub)apartado A3 es de nivel algo más elevado. Por su parte, el apartado B contiene la generalización del binomio conocida como el *multinomio* de Newton, que trata de potencias de sumas con más de dos sumandos. Lea estos dos apartados, sin dudar.

Si le quedan energías (y ambición), embárquese en las estimaciones para el tamaño de los coeficientes binómicos que obtendremos en el apartado C; en la discusión del apartado D sobre cómo Bernoulli (Jacob) obtuvo una fórmula para la suma de las primeras potencias p-ésimas de números naturales; o incluso en la lectura de los apartados E y F, que contienen dos generalizaciones, muy naturales, de la fórmula del binomio de Newton, a saber: si podemos sustituir el exponente n (entero positivo) por un exponente real cualquiera; y qué se obtendría si en lugar de multiplicar $(1+x)$ por sí mismo n veces, consideráramos n factores del tipo $(1+x_i)$, para ciertos $x_1,\dots,x_n$. Más adelante, en el apartado 5.1.8, veremos otra generalización del binomio de Newton, realmente asombrosa, debida al gran Abel.

Como aperitivo, analizamos la bien conocida regla de Leibniz para la derivación de un producto. Si $f(x)$ y $g(x)$ son dos funciones, entonces

$$(f\cdot g)' = f'\cdot g + f\cdot g'.$$

Ésta es la mencionada regla de Leibniz. En la escritura anterior hemos omitido, por comodidad, y porque nos acabamos de poner en modo decididamente algebraico, la dependencia en la variable x. Para calcular la segunda derivada del producto, basta derivar lo obtenido en el paso anterior para obtener

$$(f\cdot g)'' = f''\cdot g + f'\cdot g' + f'\cdot g' + f\cdot g'' = f''\cdot g + 2\,f'\cdot g' + f\cdot g''$$

Compruebe, querido lector, derivando esta última expresión, que

$$(f\cdot g)''' = f'''\cdot g + 3\,f''\cdot g' + 3\,f'\cdot g'' + f\cdot g''',$$

fórmula en la que aparecen unos coeficientes $1,3,3,1$ de sospechosa cercanía a valores de coeficientes binómicos. Para el caso general, denotando por $f^{(n)}$ a la n-ésima derivada de f (con el convenio de que $f^{(0)}=f$), tenemos la siguiente[20]:

Lema 5.1.7 (Regla de derivación de Leibniz) *Dadas dos funciones f y g, y para cualquier $n\geq 0$,*

$$(f\cdot g)^{(n)} = \sum_{k=0}^{n}\binom{n}{k}\,f^{(k)}\cdot g^{(n-k)}$$

Dejamos como ejercicio para el lector la prueba de esta identidad por inducción en n, el número de derivadas, siguiendo los pasos de la demostración del binomio de Newton original, lema 5.1.6, y usando que $f^{(n+1)}=(f^{(n)})'$ y que la derivada de una suma es la suma de las derivadas. Quizás incluso se anime a pergeñar un argumento simbólico, como el visto para el binomio, para probar esta identidad, cambiando el multiplicar por 1 o por x de entonces por la disyuntiva derivar/no derivar.

[20] Ups, Newton y Leibniz juntos la misma página. La que se va a liar...

A. Jugando con el binomio de Newton

El binomio de Newton:

$$(\star) \quad (1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k\,, \qquad \text{o bien} \qquad (\star\star) \quad (x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k\, y^{n-k},$$

es una identidad válida para todo $x, y \in \mathbb{R}$ y todo $n \geq 0$. Podemos sustituir x (o quizás y) por los valores que nos plazca, sin restricción alguna, podemos derivar estas identidades, quizás integrarlas, a ver qué sale... ¿A qué esperamos?

A1. Sustituyendo x por ± 1. De entre todos los valores (numéricos) por los que puede sustituir x en la expresión $(\star)$, $x = 1$ o $x = -1$ parecen las elecciones más sugerentes. El caso $x = 1$ nos da la ya conocida expresión

$$2^n = \sum_{k=0}^{n} \binom{n}{k}\,,$$

mientras que al tomar $x = -1$ (sin olvidarnos de que el polinomio de la derecha de $(\star)$ tiene término independiente) obtenemos una expresión nueva, que registramos, para futuras referencias, en el siguiente lema.

Lema 5.1.8 *Para todo $n \geq 1$,*

$$\boxed{0 = \sum_{k=0}^{n} (-1)^k \binom{n}{k}}$$

Para $n = 0$, la suma de la derecha vale 1.

De manera que si, para un $n \geq 1$ dado, sumamos todos los coeficientes $\binom{n}{j}$ de una fila del triángulo de Pascal–Tartaglia, obtenemos 2^n. Mientras que si los sumamos, pero *alternados en signo*, la suma vale 0. Esta identidad, algo más inesperada[21], será útil más adelante, por ejemplo en la demostración del principio de inclusión/exclusión (sección 5.1.5), y también en los argumentos del apartado 5.1.7.

Por cierto, si en la expresión $(\star\star)$ tomamos $x = p$ e $y = 1 - p$, obtenemos que

$$1 = \sum_{k=0}^{n} \binom{n}{k} p^k\, (1-p)^{n-k}.$$

Escribir esta expresión con p, en lugar de la habitual x, responde a razones de estética y tradición. Aunque la expresión anterior es válida para todo p real, será especialmente interesante cuando p sea un número entre 0 y 1: una *probabilidad* (capítulo 15 y siguientes). Insistimos, lector: para todo p y todo entero $n \geq 1$. Tome un p cualquiera, por ejemplo $p = 1/3$, y un n, digamos $n = 5$. Entonces la suma $(1/3)^5 + 5 \cdot (1/3)^4 (2/3) + 10 \cdot (1/3)^3 (2/3)^2 + 10 \cdot (1/3)^2 (2/3)^3 + 5 \cdot (1/3)(2/3)^4 + (2/3)^5$ vale... 1. Asombroso.

[21] Note el lector que, si n fuera impar, la identidad sería consecuencia de la simetría los coeficientes binómicos, que se irían cancelando por parejas. En el caso de n par ya no es así, pues hay un número impar de sumandos.

A2. Derivando. Nuestro siguiente objetivo es evaluar, para $n \geq 1$, la suma

$$\sum_{k=0}^{n} k \binom{n}{k}$$

Observe, lector, que no podemos obtener esta expresión sustituyendo en el binomio por ningún valor fijo de x. Sin embargo, *derivando* (con respecto a x) los dos miembros de la identidad ($\star$), obtenemos

$$n(1+x)^{n-1} = \sum_{k=1}^{n} \binom{n}{k} k x^{k-1} .$$

Note el lector que esta identidad requiere que $n \geq 1$, y fíjese también en que la suma de la derecha empieza en $k = 1$ (pues el caso $k = 0$ es cero). Ahora sí, evaluando en $x = 1$, tenemos

$$\boxed{\sum_{k=0}^{n} k \binom{n}{k} = n\, 2^{n-1}}$$

Por cierto, la expresión anterior es también válida para $n = 0$, con los convenios habituales. Veamos un ejemplo de uso combinatorio de esta identidad.

EJEMPLO 5.1.1 *El "tamaño medio" de los subconjuntos de* $\{1, \ldots, n\}$*.*

Llamemos Γ a la colección de todos los subconjuntos de $\{1, \ldots, n\}$. Ya sabemos que $|\Gamma| = 2^n$. El "tamaño medio" al que nos referimos es la media aritmética de los tamaños de todos los subconjuntos, es decir,

$$\frac{1}{2^n} \sum_{\gamma \in \Gamma} |\gamma| ,$$

donde $|\gamma|$ significa el tamaño de cada subconjunto γ. Escribimos ahora esa suma, que consta de 2^n sumandos, agrupando los subconjuntos de Γ por su tamaño, que será un cierto parámetro k entre 0 y n. El cálculo, pese a la aparatosa notación, es sencillo y directo:

$$\begin{aligned}
\frac{1}{2^n} \sum_{\gamma \in \Gamma} |\gamma| &= \frac{1}{2^n} \sum_{k=0}^{n} \sum_{\substack{\gamma \in \Gamma \\ |\gamma| = k}} |\gamma| = \frac{1}{2^n} \sum_{k=0}^{n} \sum_{\substack{\gamma \in \Gamma \\ |\gamma| = k}} k = \frac{1}{2^n} \sum_{k=0}^{n} k \, \# \left\{ \begin{array}{c} \text{subconjuntos } \gamma \in \Gamma \\ \text{tales que } |\gamma| = k \end{array} \right\} \\
&= \frac{1}{2^n} \sum_{k=0}^{n} k \binom{n}{k} = \frac{1}{2^n} \left(n\, 2^{n-1} \right) = \frac{n}{2} ,
\end{aligned}$$

En Γ hay subconjuntos pequeños (pocos), medianos (muchos) y grandes (pocos), en proporciones que se equilibran para dar que el tamaño medio es $n/2$. Vea el lector un argumento algo más directo el ejercicio 5.1.12. ♣

Registramos también la fórmula que se obtiene tomando $x = -1$:

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k \, k = 0$$

(incluso para $n = 0$), una suma alternada de coeficientes binómicos como la del lema 5.1.8, pero donde éstos van ahora multiplicados por los sucesivos valores de k, y que *también vale* 0.

Por otro lado, si en ($\star\star$), y para $n \geq 1$, derivamos con respecto a x, obtenemos

$$n(x+y)^{n-1} = \sum_{k=1}^{n} \binom{n}{k} k x^{k-1}\, y^{n-k} \qquad \Longleftrightarrow \qquad x\, n(x+y)^{n-1} = \sum_{k=1}^{n} \binom{n}{k} k x^{k}\, y^{n-k},$$

tras multiplicar por un factor x a la derecha y a la izquierda con objeto de volver a acompasar la potencia de la x a la derecha, que había quedado momentáneamente descolocada[22]. Si ahora, en esta última expresión, tomamos $x = p$ e $y = 1 - p$, obtenemos

$$\boxed{np = \sum_{k=1}^{n} k \binom{n}{k} p^k \,(1-p)^{n-k}}$$

De nuevo, una imponente suma (la de la derecha) que vale simplemente np, ¡para cualquier valor de p!, y en particular para cualquier $p \in [0,1]$, que será el caso de mayor interés, cuando digamos que el cálculo anterior significa que la "media de una variable aleatoria binomial de parámetros n y p vale np". El lector ducho en manejos probabilistas entenderá de inmediato el entrecomillado anterior; aquél que no, puede consultar el capítulo 15. Por cierto, el caso $p = 1/2$ de esta expresión es la fórmula encuadrada de la página anterior.

A3. Derivando (más). Esta argucia, consistente en derivar (y luego sustituir), será de uso habitual en el capítulo 12 dedicado a las funciones generatrices, y como ya hemos dejado entrever en un par de ocasiones, resultará particularmente útil en el análisis de cuestiones probabilistas. Pero el lector ambicioso[23] estará ya reclamando que le saquemos, aquí y ahora, más jugo al truco.

Vamos, pues. Fijamos $n \geq 1$ y consideramos M, con $0 \leq M < n$, que contará el número de derivadas que vamos haciendo. Obsérvese que no llegamos a $M = n$. El lector que, ordenadamente, se ponga a derivar a derecha e izquierda del binomio llegará sin dificultad a la expresión siguiente, válida para cualquier $0 \leq M < n$:

$$(\dagger) \qquad n(n-1)\cdots(n-M+1)\,(1+x)^{n-M} = \sum_{j=M}^{n} \binom{n}{j}\, j(j-1)\cdots(j-M+1)\, x^{j-M}$$

Podríamos haber escrito la suma de la derecha empezando en $j = 0$, pues para $j < M$ alguno de los factores del producto $j(j-1)\cdots(j-M+1)$ es con seguridad nulo. El caso $M = 0$ es el propio binomio de Newton. Compruebe el lector que el caso $M = n$ nos daría una identidad $n! = n!$, de escaso interés, razón por la que ha sido excluido desde el principio.

[22] Normal, te cae una derivada encima...

[23] Ambicioso, pero prudente, y que no haya salido corriendo a revisar los capítulos dedicados a funciones generatrices o los de probabilidad.

Podemos, y resulta cómodo y conveniente, reescribir (†) usando coeficientes binómicos. Cerciórese el lector de que, añadiendo un factor $M!$ aquí y allá, (†) es equivalente a

$$(\dagger\dagger)\qquad \binom{n}{M}(1+x)^{n-M} = \sum_{j=M}^{n}\binom{n}{j}\binom{j}{M}x^{j-M}.$$

Tomando, como ya es tradicional en este apartado, $x=1$ y $x=-1$, obtenemos las dos siguientes identidades: para $n\geq 1$ y para todo $0\leq M<n$,

$$\sum_{j=M}^{n}\binom{n}{j}\binom{j}{M} = 2^{n-M}\binom{n}{M} \qquad \text{y} \qquad \sum_{j=M}^{n}\binom{n}{j}\binom{j}{M}(-1)^{j-M} = 0.$$

Los casos $M=0$ y $M=1$ han sido vistos con anterioridad. La de la izquierda, la suma sin signo, es también válida para $M=n$, e incluso para $n=0$. La suma alternada en signo de la derecha, sin embargo, vale 1 cuando $M=n$ (incluyendo el caso $n=0$).

Reconocerá, lector, que son identidades un punto mágicas. La que no tiene signo recuerda vagamente a la identidad de Vandermonde del lema 5.1.2; pero no, no es la misma, pues los índices que varían en los coeficientes binómicos no son los mismos. A la de la derecha nos referiremos más adelante (apartado 5.1.7) como propiedad de *ortogonalidad* de los coeficientes binómicos. Véase en particular el lema 5.1.28. Y todo esto, con simples manipulaciones algebraicas; aunque alguna de estas identidades reaparecerán, claro, como fruto de argumentos combinatorios en páginas sucesivas.

Como colofón y remate de este alarde algebraico, señalamos que de (†) se deduce el siguiente resultado, que por el asombro (¿cuántos van?) que supone elevamos a la categoría de:

Teorema 5.1.9 *Dado $n\geq 1$, para cada $0\leq M<n$,*

$$\sum_{j=0}^{n}\binom{n}{j}(-1)^{j}\,j^{M} = 0.$$

Los casos $M=0$ y $M=1$ de este resultado ya han sido vistos anteriormente. Para $n=0$, la suma de la izquierda valdría 1. En la expresión anterior se podría cambiar también el $(-1)^j$ por $(-1)^{j-M}$, lo que la situaría en forma más paralela a (†).

Por cierto, tiene el lector el derecho (casi la obligación) de preguntarse qué ocurriría si en la suma del teorema 5.1.9 pusiéramos $M=n$, o incluso $M>n$. Nuestro derivativo argumento no da información sobre estos casos. Pues bien, resulta que no sólo tiene sentido la pregunta (el caso $M\geq n$), sino además un significado combinatorio de enorme interés. No se pierda el lector el ejemplo 5.1.1, y las referencias de que allí parten.

Demostración. Fijamos $n\geq 1$. Arrancamos de la expresión (†), evaluada en $x=-1$, que nos dice que para cualquier $0\leq M<n$,

$$(\star)\qquad \sum_{j=M}^{n}\binom{n}{j}(-1)^{j-M}\,j(j-1)\cdots(j-M+1) = 0.$$

El teorema estaría probado si pudiéramos asegurar que cada j^M, con $0 \leq M < n$, se pudiera escribir como *combinación lineal* de los polinomios (en j) siguientes:

$$1, j, j(j-1), j(j-1)(j-2), \ldots, j(j-1)\cdots(j-n+1),$$

pues bastaría introducir esa combinación lineal en la suma alternada del enunciado, reordenar las sumas (cambiando el orden de sumación), que resultarían ser todas nulas por repetida aplicación de $(\star)$. ¿Existe siempre esa combinación lineal? Veamos, por ejemplo, el caso j^2. Escribimos

$$j^2 = \square\, j(j-1) + \square\, j + \square\, 1,$$

donde los símbolos $\square$ deben rellenarse con coeficientes. Por ejemplo, como a la izquierda ha de quedar un j^2, es evidente que el primer cuadrado ha de llevar un 1, y como eso deja un $-j$ adicional, el segundo debe ser un 1, y el tercero un 0. De manera que

$$j^2 = j(j-1) + j\,.$$

Esto es algo general. Para el lector avisado, la lista de polinomios (en j) exhibida arriba es una *base* del espacio vectorial de los polinomios de grado $\leq n-1$, tan legal como la base estándar $\{1, j, j^2, \ldots, j^{n-1}\}$, y que conocida como la base de los *factoriales decrecientes.* El lector interesado puede consultar el apartado B de la sección 5.3.1, donde encontrará una manera ordenada de calcular los coeficientes de esas combinaciones lineales, en términos de una familia conocida como números de Stirling. Aquél más preocupado por la comprobación de que se trata, efectivamente, de una base, puede revisar el capítulo 7.

De hecho, el enunciado del teorema y $(\star)$ son del todo equivalentes, y son a su vez equivalentes a las identidades que se obtienen tomando, en lugar de la base estándar o la de los factoriales decrecientes, cualquier otra base del espacio de los polinomios de grado $\leq n-1$. ■

B. El multinomio de Newton

Interesa ahora obtener una fórmula para el desarrollo de expresiones como $(x+y+z)^n$, o $(x+y+z+t)^n$, etc., en las que en lugar de dos sumandos, como en el binomio, consideramos más. Preparando el terreno, reescribimos el binomio de Newton de la siguiente manera:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{a,b\geq 0,\, a+b=n} \frac{n!}{a!\, b!}\, x^a y^b\,.$$

La de la derecha es una suma doble, con índices a y b, pero en la que sólo se consideran aquellas parejas (a, b) que suman n. Esta escritura de la derecha se justifica observando que en la primera suma, para cada k, los exponentes de x e y suman siempre n.

El caso de tres sumandos, la *fórmula del trinomio*, tiene una estructura bastante similar:

$$(x+y+z)^n = \sum_{a,b,c\geq 0,\, a+b+c=n} \frac{n!}{a!\, b!\, c!}\, x^a y^b z^c = \sum_{a,b,c\geq 0} \binom{n}{a,\ b,\ c} x^a y^b\, .z^c\,,$$

A la derecha hemos utilizado una nueva notación,

$$\binom{n}{a,\,b,\,c} = \begin{cases} \frac{n!}{a!\,b!\,c!} & \text{si } a+b+c=n; \\ 0 & \text{en caso contrario,} \end{cases}$$

que simplifica la escritura de la suma triple, en la que sólo se suman índices $a, b, c \geq 0$ tales que $a + b + c = n$. No sorprenderá al lector que el número anterior sea conocido como un *coeficiente trinómico*. O que la generalización al caso de n sumandos, la **fórmula del multinomio** de Newton, rece como sigue: para cualesquiera números naturales $k, n \geq 1$ y para cualquier lista de números reales $(x_1, \dots, x_k)$,

$$\boxed{(x_1 + x_2 + \cdots + x_k)^n = \sum_{a_1,\dots,a_k \geq 0} \binom{n}{a_1,\, a_2, \dots,\, a_k} {x_1}^{a_1} {x_2}^{a_2} \cdots {x_k}^{a_k}}$$

donde los **coeficientes multinómicos** se definen como

$$\binom{n}{a_1,\, a_2, \dots,\, a_k} = \begin{cases} \dfrac{n!}{a_1!\, a_2! \cdots a_k!} & \text{si } a_1 + a_2 + \cdots + a_k = n; \\ 0 & \text{en caso contrario.} \end{cases}$$

Proponemos al lector dos maneras de comprobar que esta expresión es correcta. La primera consiste en utilizar el principio de inducción, teniendo en cuenta que, en la expresión anterior, hay dos parámetros involucrados, k y n. Dejamos los detalles como ejercicio 5.1.19 para el lector. Como alternativa, daremos una prueba de corte combinatorio, como la que al principio de la sección dábamos para el caso del binomio, aunque habrá de esperar un poco, al apartado 5.1.6, cuando dotemos a estos coeficientes multinómicos de significado combinatorio.

A propósito de las sumas multinómicas que hemos escrito antes, es razonable preguntarse por cuántos sumandos aparecen en ellas. En el caso $k = 2$, el del binomio, hay $n+1$ sumandos. La respuesta, para la versión multinómica general, es que hay tantos sumandos como listas $(a_1, \dots, a_k)$ de números no negativos que suman exactamente n haya, y de éstas hay

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

como veremos en la sección 5.1.3.

Por cierto, en paralelo al binomio de Newton (lema 5.1.6) y la regla de derivación de Leibniz (lema 5.1.7), y tras mirar fijamente la fórmula del desarrollo de $(x_1 + \cdots + x_k)^n$, ¿se anima el lector a escribir una fórmula para

$$(f_1 \cdots f_k)^{(n)},$$

la derivada n-ésima de un producto de k funciones?

C. Coeficientes binómicos y entropía

Retomamos ahora la discusión sobre estimaciones del tamaño de un coeficiente binómico genérico $\binom{n}{k}$, ampliando el análisis que sobre el coeficiente central (y mayor) $\binom{n}{\lfloor n/2\rfloor}$ hicimos en la sección 5.1.1 anterior.

Empezamos con una acotación sencilla, que se sigue de argumentos que ya utilizamos en la estimación de la función factorial en el apartado 3.7.

Lema 5.1.10 *Para todo $n \geq 1$ y todo $1 \leq k \leq n$,*

$$\frac{n^k}{k^k} \leq \binom{n}{k} \leq \frac{n^k}{k!} \leq \frac{(en)^k}{k^k}.$$

La cota (última) de la derecha es peor que la estimación $n^k/k!$, pero cuenta con la ventaja de que no aparece factorial alguno, además del evidente y reconfortante paralelismo con la cota inferior.

A modo de ilustración numérica, consideramos los coeficientes binómicos $\binom{50}{20}$ y $\binom{50}{30}$, cuyo valor por cierto coincide, y para los que el lema anterior nos daría las siguientes estimaciones:

$$\begin{aligned} 9.09 \times 10^7 \approx \Big(\frac{5}{2}\Big)^{20} \leq \binom{50}{20} \leq \Big(\frac{5e}{2}\Big)^{20} \approx 4.41 \times 10^{16}, \\ 4.52 \times 10^6 \approx \Big(\frac{5}{3}\Big)^{30} \leq \binom{50}{30} \leq \Big(\frac{5e}{3}\Big)^{30} \approx 4.83 \times 10^{19}, \end{aligned}$$

Note el lector que el segundo rango es mucho más amplio, y observe de paso cómo estas cotas no captan las propiedades de simetría de los coeficientes binómicos.

En general, las cotas del lema 5.1.10 son muy poco precisas, especialmente cuando k es cercano a n, como el lector podrá apreciar si sigue con detalle el argumento que emplearemos en la demostración. Por insistir en esa imprecisión, para $\binom{2n}{n}$ el lema 5.1.10 daría $2^n \leq \binom{2n}{n} \leq (2e)^n$, que es una peor estimación (por ambos lados) que la de la sección 5.1.1.

Demostración. Comenzamos con las cotas superiores. La primera es inmediata, y se sigue de sustituir cada uno de los k factores del numerador siguiente por el mayor de ellos:

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \leq \frac{n^k}{k!}.$$

La segunda sale de combinar esta última estimación con la (holgada) acotación de e^k en la que nos quedamos únicamente con un término, justamente el término k, de la serie de potencias que la define:

$$e^k = \sum_{j=0}^{\infty} \frac{k^j}{j!} > \frac{k^k}{k!} \qquad \Longrightarrow \qquad \frac{1}{k!} < \frac{e^k}{k^k}.$$

Para la cota inferior observamos primero que, para cualquier m en el rango $0 < m < k \leq n$,

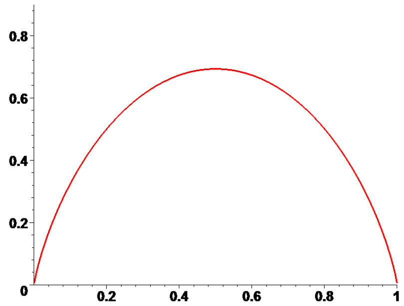
$$\frac{n}{k} \leq \frac{n-m}{k-m},$$

como podrá comprobar fácilmente el lector, lo que nos dice que

$$\frac{n^k}{k^k} = \frac{n}{k} \cdot \frac{n}{k} \cdots \frac{n}{k} \le \frac{n}{k}\,\frac{n-1}{k-1} \cdots \frac{n-(k-1)}{k-(k-1)} = \binom{n}{k},$$

que es la estimación inferior del enunciado. ■

Para afinar más, introducimos la siguiente función, definida para $p \in [0,1]$:



$$H(p) = -p\,\ln(p) - (1-p)\,\ln(1-p),$$

cuyo aspecto representamos a la derecha. Observe el lector que $H(p)$ es simétrica con respecto a $p = 1/2$, esto es, $H(p) = H(1-p)$, y compruebe que alcanza su máximo en $p = 1/2$, y que $H(1/2) = \ln(2)$.

Note también que, para $p \in (0,1)$,

$$e^{nH(p)} = e^{-np\ln(p)}\, e^{-n(1-p)\ln(1-p)} = p^{-np}\,(1-p)^{-n(1-p)}.$$

A esta función nos referiremos como una **entropía**[24]. En términos de esta función, tenemos las siguientes estimaciones para el tamaño de un coeficiente binómico genérico.

Teorema 5.1.11 *Para $n \ge 1$ y $0 \le k \le n$,*

$$\frac{e^{nH(k/n)}}{n+1} \le \binom{n}{k} \le e^{nH(k/n)}.$$

Antes de dar la demostración de este resultado, vamos a compararlo con las cotas del lema 5.1.10. Para empezar, las estimaciones del teorema 5.1.11 son simétricas, y acotan por igual los coeficientes binómicos $\binom{n}{k}$ y $\binom{n}{n-k}$, pues $H((n-k)/n) = H(1-k/n) = H(k/n)$.

Pero además baten ampliamente a las del lema 5.1.10 (bueno, casi siempre). Vemos primero un ejemplo numérico, comparando las estimaciones para $\binom{50}{20}$ y $\binom{50}{30}$. Por un lado, el lema 5.1.10 nos daría, usando que $(5/2)^{20} = e^{20\ln(5/2)}$ y que $(5/3)^{30} = e^{30\ln(5/3)}$,

$$e^{18.33} \le \binom{50}{20} \le e^{38.33} \qquad \text{y} \qquad e^{15.32} \le \binom{50}{30} \le e^{45.32}$$

[24] Entropía: medida del desorden de un sistema. Repase, querido lector, su segunda ley de la termodinámica.

> Si alguien le hace notar que su teoría del Universo contradice las ecuaciones de Maxwell... entonces, peor para las ecuaciones de Maxwell. Si no concuerda con las observaciones... bueno, estos experimentales hacen chapucillas de vez en cuando. Pero si su teoría va en contra de la segunda ley de la termodinámica, entonces, ¡ay!, entonces no le concedo esperanza alguna: el único destino posible para su teoría es caer en la más profunda de las humillaciones.

(Cita del astrofísico inglés Sir Arthur Stanley Eddington). O visite el capítulo 26 dedicado a la transmisión de la información para poner en contexto esta asignación de nombre.

mientras que el teorema 5.1.11 acotaría ambos números por un muchísimo más ajustado

$$\frac{e^{33.31}}{51} \leq \binom{50}{20} = \binom{50}{30} \leq e^{33.31}.$$

Compruebe también el lector que del teorema 5.1.11 se deduce que

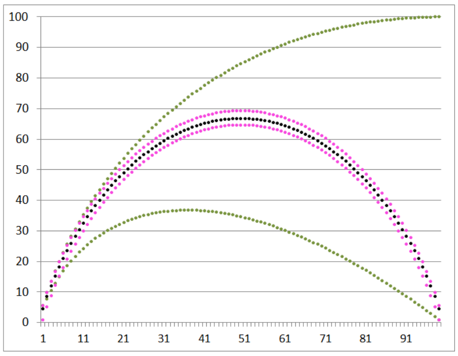
$$\frac{4^n}{n+1} \leq \binom{2n}{n} \leq 4^n,$$

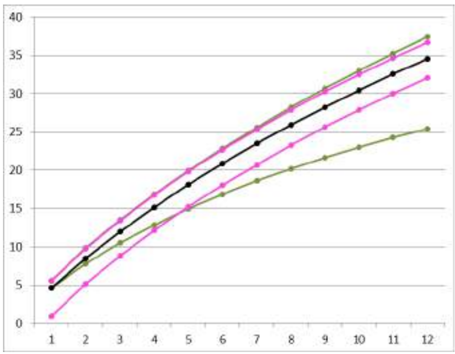
que son las cotas que obtuvimos en la sección anterior.

En general, estas cotas, medidas en términos logarítmicos, son:

$$\begin{array}{lrcl} \text{(lema 5.1.10)} & k \ln\Big(\dfrac{n}{k}\Big) & \leq \ln\dbinom{n}{k} & \leq k \ln\Big(\dfrac{en}{k}\Big), \\ \text{(teorema 5.1.11)} & nH\Big(\dfrac{k}{n}\Big) - \ln(n+1) & \leq \ln\dbinom{n}{k} & \leq nH\Big(\dfrac{k}{n}\Big), \end{array}$$

En la figura que sigue hemos representado estas cuatro cotas para el caso $n = 100$, junto con los verdaderos valores de $\ln(\binom{100}{k})$, que en la gráfica aparecen en negro (tercera contando desde abajo). Las dos gráficas que la envuelven, en malva, son las estimaciones de entropía. Note, lector, el excelente ajuste que proporcionan. En verde, las dos estimaciones del lema 5.1.10. A la derecha, y para el lector puntilloso, presentamos una ampliación del rango de valores de k pequeños, en el que se observa cómo ahí la estimación (inferior) de entropía es ligeramente peor que la inferior del lema 5.1.10. Pero claro, se dirá ese lector inquisitivo, pero al tiempo razonable: para esos valores de k tan pequeños, ¿para qué andan estimando?





En la demostración del teorema 5.1.11 usaremos la siguiente observación, ya de por sí interesante.

Lema 5.1.12 *Fijemos un $q \in (0,1)$ y consideremos la sucesión*

$$a_k = \binom{n}{k} q^k (1-q)^{n-k}, \qquad \textit{para } k = 0, 1, \ldots, n.$$

La sucesión (a_k) es unimodal[25]*. En particular, si nq es un entero, la sucesión tiene máximo en $k = nq$.*

[25] Sí: también es logcóncava, lo que nos daría la unimodalidad como consecuencia inmediata. Esta logconcavidad se comprueba fácilmente a partir del teorema 7.2.22 de Newton.

Observe, lector, que el caso $q = 1/2$ se corresponde, salvo un factor de escala, con la unimodalidad de los coeficientes binómicos que vimos unas páginas atrás.

Demostración. Calculamos

$$a_k - a_{k+1} = \binom{n}{k} q^k (1-q)^{n-k} - \binom{n}{k+1} q^{k+1} (1-q)^{n-k-1} = \binom{n}{k} q^k (1-q)^{n-k} \Big[1 - \frac{n-k}{k+1} \frac{q}{1-q}\Big],$$

de manera que se tendrá que $a_k - a_{k+1} \leq 0$ si y sólo si

$$1 - \frac{n-k}{k+1} \frac{q}{1-q} \leq 0 \qquad \text{es decir, si y sólo si} \qquad k \leq qn + q - 1.$$

De manera que la sucesión es creciente hasta $qn + q - 1$. El mismo argumento nos dice que será decreciente desde ese punto. Y por tanto, unimodal.

Cuando nq es un entero, como $|q-1| < 1$, la sucesión tendrá un máximo en $k = nq$. Note el lector que podría ocurrir, como en el caso de los coeficientes binómicos, que hubiera dos máximos (iguales en valor) consecutivos. ■

Demostración del teorema 5.1.11. Fijamos un k entre 0 y 1, que por comodidad notacional representamos como $k = nq$, para cierto $q \in [0, 1]$; esto es, $q = k/n$. En estos términos, las acotaciones que pretendemos comprobar son

$$\frac{e^{nH(q)}}{n+1} \leq \binom{n}{nq} \leq e^{nH(q)}.$$

Observe, lector, que este enunciado es evidente si fuera $q = 0$ o si fuera $q = 1$.

Para la cota superior basta observar que

$$\binom{n}{nq} q^{nq} (1-q)^{n(1-q)} \leq \sum_{j=0}^{n} \binom{n}{j} q^j (1-q)^{n-j} = 1,$$

donde la última igualdad es consecuencia del teorema del binomio, de lo que se deduce que

$$\binom{n}{nq} \leq q^{-nq} (1-q)^{-n(1-q)} = e^{nH(q)}.$$

Con un poco (pero no mucho) más de trabajo, el lector podrá comprobar (ejercicio 5.1.5) que no sólo $\binom{n}{nq}$, sino la *suma* de los coeficientes binómicos hasta el $\binom{n}{nq}$ se acota por $e^{nH(q)}$.

Veamos ahora la cota inferior. El lema 5.1.12 nos dice que

$$\binom{n}{nq} q^{nq} (1-q)^{n(1-q)}$$

es el mayor de los términos de la sucesión a_k del enunciado (quizás "empatado" con algún otro). Pero entonces, usando el binomio una vez más,

$$1 = \sum_{j=0}^{n} \binom{n}{j} q^j (1-q)^{n-j} \leq (n+1) \binom{n}{nq} q^{nq} (1-q)^{n(1-q)},$$

donde hemos utilizado que hay $n+1$ términos en la suma anterior. Despejando $\binom{n}{nq}$ se obtiene la cota buscada. ■

Planteamos, por último, la cuestión de establecer estimaciones asintóticas para $\binom{n}{k}$, en lugar de cotas. Ya sabemos, recuérdese la sección anterior, que $\binom{2n}{n} \sim 4^n/\sqrt{\pi n}$ cuando $n \to \infty$, asintótico que obtuvimos usando la fórmula de Stirling. ¿Y para un coeficiente binómico general? Note, lector, que hay dos parámetros involucrados, n y k, de forma que habrá primero que decidir si nos interesa la situación en la que n se va a infinito y k queda fijo, o si los dos se van a infinito, en cuyo caso deberíamos prefijar a qué velocidades lo hace cada uno... En el siguiente resultado analizamos dos posibilidades: el asintótico cuando $n \to \infty$ y k está fijo, y el caso en el que ambos parámetros se van a infinito a (esencialmente) la misma velocidad.

Lema 5.1.13 a) *Para k fijo,*

$$\binom{n}{k} \sim \frac{n^k}{k!} \qquad \textit{cuando } n \to \infty.$$

b) *Sea (k_n) una sucesión de enteros tal que $k_n \sim qn$ cuando $n \to \infty$, para cierto $q \in (0,1)$ fijo. Entonces, cuando $n \to \infty$,*

$$\binom{n}{k_n} \sim \frac{e^{nH(q)}}{\sqrt{2\pi n q(1-q)}} \qquad y \qquad \ln\binom{n}{k_n} \sim nH(q).$$

El primer apartado, fijando por ejemplo $k=7$, sugiere aproximar $\binom{100}{7}$ por $100^7/7!$. El error (relativo) cometido sería de aproximadamente un 20 %. Pero la aproximación de $\binom{1000}{7}$ por $1000^7/7!$ daría ya sólo un error del 2.1 %.

En cuanto al segundo apartado, está escrito por comodidad en términos de una genérica sucesión k_n con cierta propiedades asintóticas. El lector con mayor inclinación por la concreción puede optar por escoger $k_n = \lfloor nq \rfloor$, y escribir así que, cuando $n \to \infty$,

$$\binom{n}{\lfloor nq \rfloor} \sim \frac{e^{nH(q)}}{\sqrt{2\pi n q(1-q)}} \qquad \text{y} \qquad \ln\binom{n}{\lfloor nq \rfloor} \sim nH(q).$$

En el caso $q = 1/2$, la estimación de la izquierda nos da

$$\binom{n}{\lfloor n/2 \rfloor} \sim \frac{e^{nH(1/2)}}{\sqrt{2\pi n/4}} = \frac{2^n}{\sqrt{\pi n/2}} \qquad \text{cuando } n \to \infty,$$

que coincide con la estimación asintótica para el coeficiente binómico central (y mayor) que obtuvimos unas páginas atrás. Pero tomando por ejemplo $q = 1/3$, obtenemos

$$\binom{n}{\lfloor n/3 \rfloor} \sim \frac{e^{nH(1/3)}}{\sqrt{2\pi n 2/9}} = \frac{3^n}{4^{n/3}}\,\frac{3}{2\sqrt{\pi n}} \qquad \text{cuando } n \to \infty.$$

Demostración. La parte a) es directa, pues basta observar que, para k fijo,

$$\frac{\binom{n}{k}}{n^k/k!} = \frac{n(n-1)\cdots(n-k+1)}{n^k} = \Big(1-\frac{1}{n}\Big)\Big(1-\frac{2}{n}\Big)\cdots\Big(1-\frac{k-1}{n}\Big) \longrightarrow 1 \quad \text{cuando } n \to \infty.$$

Para la parte b), observamos que, por las estimaciones de entropía del teorema 5.1.11,

$$nH(k_n/n) - \ln(n+1) \le \ln \binom{n}{k_n} \le nH(k_n/n),$$

que tras dividir por $nH(q)$ da

$$\frac{H(k_n/n)}{H(q)} - \frac{\ln(n+1)}{nH(q)} \le \frac{\ln \binom{n}{k_n}}{nH(q)} \le \frac{H(k_n/n)}{H(q)}.$$

Dado que la función de entropía $H(p)$ es continua, se tiene que $H(k_n/n)/H(q) \to 1$ cuando $n \to \infty$. Para rematar el argumento, observamos que $\ln(n+1)/n \to 0$ cuando $n \to \infty$.

Para la otra estimación usamos que, cuando $n \to \infty$, $k_n \sim nq$ y $n - k_n \sim np$ (donde, por comodidad notacional, escribimos $p = 1 - q$), además de la aproximación de Stirling, para escribir que, cuando $n \to \infty$,

$$\binom{n}{k_n} = \frac{n!}{k_n!\,(n-k_n)!} \sim \frac{n^n\, e^{-n}\,\sqrt{2\pi n}}{k_n^{k_n}\, e^{-k_n}\,\sqrt{2\pi k_n}\;(n-k_n)^{n-k_n}\, e^{-(n-k_n)}\,\sqrt{2\pi(n-k_n)}}$$
$$= \frac{n^n}{k_n^{k_n}\,(n-k_n)^{n-k_n}\,\sqrt{2\pi\, k_n(n-k_n)}} \sim \frac{n^n}{(nq)^{nq}\,(np)^{np}\,\sqrt{2\pi\, nqp}} = \frac{1}{q^{nq}\,p^{np}\,\sqrt{2\pi\, nqp}} = \frac{e^{nH(q)}}{\sqrt{2\pi nqp}}.$$

Observe el lector que de esta última estimación asintótica se puede obtener la del logaritmo, pero no al revés. Consulte el ejercicio 3.7.4. ■

Para concluir esta entrópica discusión, señalamos que, en ciertos contextos relacionados con codificación (véase el capítulo 26), es más habitual usar la función de entropía

$$H_2(q) = -q\,\log_2(q) - (1-q)\,\log_2(1-q),$$

usando logaritmos en base 2, en lugar de la base natural del número e. Esta función tiene un aspecto muy similar a la de $H(q)$ y su máximo está en $H_2(1/2) = 1$. Observe también el lector que $e^{nH(q)} = 2^{nH_2(q)}$.

D. Jacob Bernoulli y las sumas de potencias

Retomamos aquí la discusión que sobre las sumas de potencias iniciamos en el apartado 1.2.1. Interesa obtener una fórmula, o un procedimiento de cálculo eficaz, para las cantidades

$$S_p(n) = \sum_{k=1}^{n} k^p\,, \qquad \text{para enteros } p \ge 0 \text{ y } n \ge 1.$$

Ya sabemos, por ejemplo, que

$$S_0(n) = n\,, \quad S_1(n) = \frac{n(n+1)}{2}\,, \quad S_2(n) = \frac{n(n+1)(2n+1)}{6} \quad \text{y} \quad S_3(n) = \frac{n^2(n+1)^2}{4}.$$

Véase el citado apartado 1.2.1 y alguno de sus ejercicios.

Para analizar $S_p(n)$ con generalidad, aplicamos un truco telescópico, similar al que llevamos a cabo para el caso $p = 2$ en el ejemplo 1.2.4. Para un p fijo,

$$\sum_{j=1}^{n} \big[(j+1)^{p+1} - j^{p+1}\big] = (n+1)^{p+1} - 1,$$

por cancelación telescópica. Por otro lado, usando el binomio de Newton,

$$\begin{aligned}\sum_{j=1}^{n} \big[(j+1)^{p+1} - j^{p+1}\big] &= \sum_{j=1}^{n} \Big[\sum_{k=0}^{p+1} \binom{p+1}{k} j^k - j^{p+1}\Big] = \sum_{j=1}^{n}\sum_{k=0}^{p} \binom{p+1}{k} j^k \\ &= \sum_{k=0}^{p} \binom{p+1}{k} \sum_{j=1}^{n} j^k = \sum_{k=0}^{p} \binom{p+1}{k} S_k(n);\end{aligned}$$

note, lector, el cambio de orden de sumación efectuado. Igualando los dos cálculos obtenemos que, para $p \geq 0$,

$$(n+1)^{p+1} - 1 = \sum_{k=0}^{p} \binom{p+1}{k} S_k(n)\,;$$

o mejor, que para $p \geq 1$,

$$(\star) \qquad S_p(n) = \frac{(n+1)^{p+1} - 1}{p+1} - \frac{1}{p+1} \sum_{k=0}^{p-1} \binom{p+1}{k} S_k(n)\,,$$

tras despejar del sumatorio de arriba el término con $S_p(n)$.

La expresión $(\star)$, y de hecho el argumento que lleva a ella, se debe, al parecer, a Blaise Pascal. Es una expresión muy interesante. Por un lado, es una regla recursiva para calcular cada $S_p(n)$ si se conocen los anteriores $S_k(n)$ para $k < p$. Como ejemplo, le aconsejamos y animamos al lector a que la use para comprobar que

$$S_4(n) = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}.$$

El lector que con ánimo haya seguido el consejo (y haya llegado a buen puerto), reconocerá que la tarea ha sido trabajosa: escribir todos los términos, expandir la potencia de $(n+1)$, luego simplificar aquí y allá, para finalmente. . . Y se preguntará si no existe un método más directo (y rápido) de obtener ese mismo resultado.

Como paso previo, observamos que la expresión $(\star)$ nos dice que, para una potencia fija p y como función de n, la expresión $S_p(n)$

- es un polinomio (en n) de grado $p+1$,
- sin término independiente,
- y tal que el coeficiente de n^{p+1} es $1/(p+1)$.

Sí, todo esto, lector, está incluido en la recursión $(\star)$. Compruébelo, como corresponde, por inducción.

Y si es un polinomio con esas características, ¿no existirá una fórmula para calcular sus coeficientes? Escribimos los primeros casos, en busca de, quizás, inspiración:

$$
\begin{array}{lllllll}
S_0(n) = & n & & & & & \\
S_1(n) = & \frac{1}{2}\,n^2 & +\frac{1}{2}\,n & & & & \\
S_2(n) = & \frac{1}{3}\,n^3 & +\frac{1}{2}\,n^2 & +\frac{1}{6}\,n & & & \\
S_3(n) = & \frac{1}{4}\,n^4 & +\frac{1}{2}\,n^3 & +\frac{1}{4}\,n^2 & & & \\
S_4(n) = & \frac{1}{5}\,n^5 & +\frac{1}{2}\,n^4 & +\frac{1}{3}\,n^3 & -\frac{1}{30}\,n & & \\
S_5(n) = & \frac{1}{6}\,n^6 & +\frac{1}{2}\,n^5 & +\frac{5}{12}\,n^4 & -\frac{1}{12}\,n^2 & & \\
S_6(n) = & \frac{1}{7}\,n^7 & +\frac{1}{2}\,n^6 & +\frac{1}{2}\,n^5 & -\frac{1}{6}\,n^3 & +\frac{1}{42}\,n & \\
S_7(n) = & \frac{1}{8}\,n^8 & +\frac{1}{2}\,n^7 & +\frac{7}{12}\,n^6 & -\frac{7}{24}\,n^4 & +\frac{1}{12}\,n^2 & \\
S_8(n) = & \frac{1}{9}\,n^9 & +\frac{1}{2}\,n^8 & +\frac{2}{3}\,n^7 & -\frac{7}{15}\,n^5 & +\frac{2}{9}\,n^3 & -\frac{1}{30}\,n \\
S_9(n) = & \frac{1}{10}\,n^{10} & +\frac{1}{2}\,n^9 & +\frac{3}{4}\,n^8 & -\frac{7}{10}\,n^6 & +\frac{1}{2}\,n^4 & -\frac{3}{20}\,n^2 \\
S_{10}(n) = & \frac{1}{11}\,n^{11} & +\frac{1}{2}\,n^{10} & +\frac{5}{6}\,n^9 & -n^7 & +n^5 & -\frac{1}{2}\,n^3 & +\frac{5}{66}\,n
\end{array}
$$

Figura 5.3: Jacob Bernoulli

¿Ve cosas, lector? Seguro que sí[26]. Algunas son bastante evidentes: ese coeficiente $1/(p+1)$ del término de grado mayor, que ya habíamos anunciado, o el más misterioso $1/2$ como coeficiente común del siguiente término. Para la tercera columna, si tiene a bien escribir los coeficientes con denominador común 12, descubrirá la secuencia $2, 3, 4, \ldots$ Note también la regularidad de signos por columnas, y esas columnas vacías alternadas.

Entra en acción Jacob Bernoulli[27]. Este asunto del cálculo de las sumas de potencias se había convertido en un objetivo de (y casi una competición entre) los matemáticos de la época, y un Bernoulli no podía quedarse al margen, ni del objetivo, ni sobre todo de la competición. Así que se puso a mirar fijamente la tabla anterior. Mírela usted también, lector, aunque tendrá que hacerlo con los ojos de Bernoulli[28]. Si lo hace de esa manera, quizás sea capaz de proponer la asombrosa fórmula para los coeficientes de $S_p(n)$ que recogemos como teorema 5.1.15, y en cuyo enunciado aparece una famosa familia de números que llevan el nombre, como corresponde, de su creador.

[26]Si también oye voces, avísenos, y ya le derivamos al servicio correspondiente.

[27]Jacob Bernoulli (1654-1705) fue el primer matemático de la gran saga de los Bernoulli, que ocupó la cátedra de matemáticas de la Universidad de Basilea durante más de cien años. Estudió Teología, aunque sus intereses derivaron pronto hacia las matemáticas, la física y la astronomía. Tutor de su hermano Johann (véase su nota biográfica en el apartado 12.2.4), acabaría enemistado con él, lo que en realidad fue una constante en la familia de los Bernoulli de Basilea; vea el lector interesado su árbol genealógico en el apartado 9.2. Además de sus aportaciones a la probabilidad (recogidas en el *Ars Conjectandi* de 1713, publicado póstumamente), trabajó en cálculo infinitesimal, cálculo de variaciones, álgebra, geometría, series infinitas, etc. Numerosos objetos matemáticos han sido bautizados en su honor: ecuación de Bernoulli, lemniscata de Bernoulli, números de Bernoulli, etc.

[28]Es un decir. Si fuera menester, póngase peluca.

Definición 5.1.14 La sucesión (B_n) de **números de Bernoulli** viene dada por la recurrencia

$$B_0 = 1, \qquad \sum_{j=0}^{n-1} \binom{n}{j} B_j = 0 \qquad \text{para cada } n \geq 2.$$

La anterior es una definición de los números de Bernoulli a través de una regla recursiva. Quizás al lector le quede más claro si la (re)escribe como

$$B_n = -\frac{1}{n+1} \sum_{j=0}^{n-1} \binom{n+1}{j} B_j \qquad \text{para cada } n \geq 1.$$

Los primeros valores de esta sucesión son

$$1, -\tfrac{1}{2}, \tfrac{1}{6}, 0, -\tfrac{1}{30}, 0, \tfrac{1}{142}, 0, -\tfrac{1}{30}, 0, \tfrac{5}{66}, 0, -\tfrac{691}{2730}, 0, \tfrac{7}{6}, 0, \ldots$$

Nótese que, salvo B_1, que vale $-1/2$, todos los números de Bernoulli de índice impar son nulos, esto es, $B_{2n+1} = 0$ para cada $n \geq 0$. Mire, por cierto, la tabla de la página anterior y repare en esas equiespaciadas columnas vacías.

Vamos ya con la anunciada fórmula:[29]

Teorema 5.1.15 (Fórmula de Faulhaber-Bernoulli) *Para cada $p \geq 0$,*

$$S_p(n) = \sum_{k=1}^{n} k^p = \sum_{j=0}^{p} \left[\frac{1}{p+1} (-1)^j \binom{p+1}{j} B_j \right] n^{p+1-j}.$$

Las expresiones encerradas entre corchetes son las prometidas fórmulas para los coeficientes del polinomio $S_p(n)$; aunque, observe el lector, que cada una de ellas es el coeficiente del término en n^{p+1-j}, de manera que por ejemplo el caso $j = 0$ se corresponde con el coeficiente de mayor grado, n^{p+1}.

Se cuenta que, tras descubrir la fórmula, Jacob Bernoulli, en plan decididamente retador, se jactaba de haber sido capaz de calcular el valor de la suma

$$1^{10} + 2^{10} + 3^{10} + \cdots + 1000^{10}$$

en sólo unos minutos. El asunto tiene mérito: el resultado es el gigantesco número[30]

$$91\,409\,924\,241\,424\,243\,424\,241\,924\,242\,500.$$

[29] Hemos decidido que Jacob Bernoulli comparta el crédito de esta fórmula con el matemático alemán Johann Faulhaber (1580-1635). No sabemos qué tal le sentará a Jacob, con ese carácter que tiene. Sobre todo, porque al parecer el bueno de Faulhaber se limitó a exhibir, unos años antes, los 17 primeros casos (lo que tampoco está mal), y a hacer unas cuantas observaciones y conjeturas sobre los coeficientes generales en su *Academia Algebrae* de 1631.

[30] ¿Se atreve a leerlo en castellano, querido lector? Diga con nosotros: noventa y un quintillones, cuatrocientos nueve mil novecientos veinticuatro cuatrillones, doscientos cuarenta y un mil cuatrocientos veinticuatro trillones, doscientos cuarenta y tres mil… Uff, termine usted, s'il vous plaît.

Aunque a decir verdad, Jacob se aprovechaba, el muy pillín, de ciertos malabares numéricos. Primero, claro, disponía de la fórmula del teorema 5.1.15, que en el caso $p = 10$ permite reducir el cálculo original (que tiene $n = 1000$ sumandos) a la suma de sólo 11 sumandos; o menos, porque

$$S_{10}(n) = \frac{1}{11}\,n^{11} + \frac{1}{2}\,n^{10} + \frac{5}{6}\,n^9 - n^7 + n^5 - \frac{1}{2}\,n^3 + \frac{5}{66}\,n,$$

aprovechando que los números de Bernoulli de índice impar son cero a partir de B_3. Y además, astutamente, escogía $n = 1000$, es decir, que el cálculo involucra potencias de 10 que, con nuestra aritmética habitual, se suman relativamente rápido. Aún así, ¡bien por Jacob!

Demostración del teorema 5.1.15. Fijamos p y escribimos

$$S_p(n) = \sum_{k=1}^{n} k^p = \sum_{j=0}^{p} H(p,j)\, n^{p+1-j}.$$

El objetivo es probar que los coeficientes $H(p,j)$ siguen la fórmula del enunciado. Observamos primero que $S_p(n) - S_p(n-1) = n^p$, y luego usamos la expresión de cada S_p, el binomio de Newton, y un cambio en el orden de sumación como sigue:

$$\begin{aligned}
n^p = S_p(n) - S_p(n-1) &= \sum_{j=0}^{p} H(p,j)\,\big[n^{p+1-j} - (n-1)^{p+1-j}\big] \\
&= \sum_{j=0}^{p} H(p,j)\,\Big[n^{p+1-j} - \sum_{k=0}^{p+1-j} \binom{p+1-j}{k} (-1)^{p+1-j-k}\, n^k\Big] \\
&= \sum_{j=0}^{p} H(p,j) \sum_{k=0}^{p-j} \binom{p+1-j}{k} (-1)^{p-j-k}\, n^k = \sum_{k=0}^{p} \Big[\sum_{j=0}^{p-k} \binom{p+1-j}{k} H(p,j)\,(-1)^{p-j-k}\Big]\, n^k.
\end{aligned}$$

Ahora, comparando coeficientes de estos dos polinomio en la variable n, concluimos que la suma

$$\sum_{j=0}^{p-k} \binom{p+1-j}{k} H(p,j)\,(-1)^{p-j-k}$$

vale 1 si $k = p$, mientras que vale 0 en los restantes casos $0 \le k < p$.

El caso $k = p$ nos da que $H(p,0) = 1/(p+1)$, como ya sabíamos, mientras que en los restantes casos, tenemos que

$$(\star) \qquad \sum_{j=0}^{p-k} \frac{(p+1-j)!}{(p+1-j-k)!}\, H(p,j)\,(-1)^j = 0 \qquad \text{para cada } 0 \le k < p,$$

tras eliminar unos cuantos factores que no dependen del índice j. En realidad podríamos terminar aquí[31], porque la relación anterior es una recurrencia que permite calcular sucesivamente todos los $H(p,j)$. Por ejemplo, tomar $k = p-1$ en $(\star)$ daría

$$0 = \frac{(p+1)!}{2!}\, H(p,0) - \frac{p!}{1!}\, H(p,1) \quad\Longrightarrow\quad 0 = \frac{p!}{2} - p!\, H(p,1) \quad\Longrightarrow\quad H(p,1) = \frac{1}{2},$$

[31] Se nota que no somos Bernoulli.

valor que ya había quedado patente en la tabla de valores de $S_p(n)$ que vimos al comienzo de esta discusión.

El genio de Jacob Bernoulli consistió en observar que los números de esa tabla eran, salvo por un danzante signo y unos factores de escala *fijos* en cada columna, como coeficientes binómicos. Siguiendo su intuición, y aprovechando que $H(p,0) = 1/(p+1)$, proponemos que los $H(p,j)$ se escriban como sigue:

$$H(p,j) = \frac{1}{p+1}\,(-1)^j\binom{p+1}{j}\,G(p,j),$$

donde los $G(p,j)$ dependen, por ahora, también de p, aunque al final del argumento veremos que no es necesario considerar tal dependencia. Obsérvese que $G(p,0) = 1$. Llevando esto a la expresión $(\star)$ obtenemos que, para cada $0 \le k < p$,

$$0 = \sum_{j=0}^{p-k} \frac{(p+1-j)!}{(p{+}1{-}j{-}k)!}\,\frac{1}{p+1}\,\frac{(p+1)!}{j!\,(p+1-j)!}\,G(p,j) \quad\Longrightarrow\quad 0 = \sum_{j=0}^{p-k} \frac{1}{j!\,(p{+}1{-}j{-}k)!}\,G(p,j),$$

tras las pertinentes simplificaciones. O mejor aún, llamando s a la diferencia $p-k$, que para cada $1 \le s \le p$,

$$0 = \sum_{j=0}^{s} \frac{1}{j!\,(s+1-j)!}\,G(p,j) \quad\Longrightarrow\quad (\dagger)\quad 0 = \sum_{j=0}^{s} \binom{s+1}{j}\,G(p,j),$$

Ya casi hemos terminado. Para p fijo, la recurrencia $(\dagger)$, arrancando con $G(p,0) = 1$, nos da la sucesión de números de Bernoulli hasta p; compárese con la definición 5.1.14. Por ejemplo, el caso $s = 1$ daría

$$0 = \binom{2}{0}G(p,0) + \binom{2}{1}G(p,1) \quad\Longrightarrow\quad G(p,1) = -\frac{1}{2},$$

que es B_1. Y así sucesivamente, de manera que $G(p,j) = B_j$ para cada j entre 1 y p. Pero si tomáramos un p' distinto, por ejemplo mayor, $(\star)$ daría los números de Bernoulli hasta p'. La conclusión es que $G(p,j) = B_j$ para cada p y cada posible j, lo que da fin a la prueba. ■

E. El binomio de Newton para exponentes reales

Rescribimos ahora el binomio de Newton, válido para cualquier $n \in \mathbb{N}$ (además de para todo $x \in \mathbb{R}$), de la siguiente manera:

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\,x^k = \sum_{k=0}^{\infty} \frac{n\,(n-1)\cdots(n-k+1)}{k!}\,x^k,$$

cancelando un factor $(n-k)!$. Observe lector, que el coeficiente de x^k, $n\,(n-1)\cdots(n-k+1)/k!$ lleva k factores en el numerador, y otros tantos en el denominador (los factores de $k!$). El caso $k = 0$ es algo especial, pues el coeficiente es simplemente un 1. A partir de $k = 1$ la expresión cobra más sentido: $n/1!$ para $k = 1$, $n(n-1)/2!$ para $k = 2$, etc., y así hasta el caso $k = n$, en el que se obtiene $n(n-1)\cdots 2\cdot 1/n! = 1$. Aunque la suma original es finita, ampliamos el rango de sumación hasta infinito porque, para $k > n$, todos los coeficientes son 0. Por ejemplo, $k = n+1$ daría $n(n-1)\cdots 2\cdot 1\cdot 0/(n+1)! = 0$.

El teorema general del binomio de Newton que enunciamos a continuación afirma que la expresión anterior es cierta sustituyendo el entero positivo n por *cualquier número real* α:

Teorema 5.1.16 (binomio de Newton) *Para todo* $\alpha \in \mathbb{R}$, *si* $|x| < 1$,

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}\, x^k.$$

¡Atención!, la expresión anterior es una *serie de potencias*, y por eso añadimos el rango de valores en el que tiene sentido, que en el caso que nos ocupa es $|x| < 1$. Repase el lector dubitativo el apartado 4.3 sobre series de potencias. Si α es un entero positivo, entonces esa serie deviene en simple polinomio, porque los coeficientes son 0 a partir de un cierto momento (para $k > \alpha$), y en ese caso la expresión (el binomio de Newton habitual) es válida para todo $x \in \mathbb{R}$. Pero en general no ocurre así. Por ejemplo, para $\alpha = 1/2$ se tiene que

$$\begin{aligned}(1+x)^{1/2} = \sqrt{1+x} &= \sum_{k=0}^{\infty} \frac{1/2\cdot(1/2-1)\cdots(1/2-k+1)}{k!}\, x^k \\ &= 1 + \frac{1/2}{1!}\, x + \frac{1/2\cdot(1/2-1)}{2!}\, x^2 + \frac{1/2\cdot(1/2-1)(1/2-2)}{3!}\, x^3 + \cdots \\ &= 1 + \frac{1}{2}\, x - \frac{1}{8}\, x^2 + \frac{1}{16}\, x^3 - \frac{5}{128}\, x^4 + \cdots\end{aligned}$$

Compruebe el lector, sacando 1/2 repetidas veces como factor común en la la fórmula de los coeficientes, contando signos negativos, etc., que el coeficiente de x^k de la serie anterior sigue la siguiente fórmula para $k \geq 2$:

$$\frac{(-1)^{k-1}}{2^k}\, \frac{1\cdot 3\cdot 5\cdot (2k-3)}{k!},$$

y que en particular son coeficientes que no se anulan.

Quizás el lector pueda pensar que esta generalización del binomio de Newton es únicamente un resultado formal. Pero la posibilidad que ofrece de obtener expresiones explícitas para los coeficientes de las series de Taylor asociadas a funciones como, por ejemplo, $\sqrt{1+x}$, será extremadamente útil en algunas aplicaciones combinatorias. Aunque habremos de esperar un poco[32], al capítulo 12, para descubrir algunas de ellas. También allí daremos la demostración del teorema 5.1.16.

F. Productos, funciones simétricas... y medias simétricas

El teorema del binomio da cuenta del desarrollo de un producto del tipo

$$(1+x)^n = (1+x) \overset{n \text{ veces}}{\cdots} (1+x)\,.$$

Planteamos ahora la cuestión de codificar el desarrollo de productos del tipo

$$(\star) \qquad (1+x_1)\cdot(1+x_2)\cdots(1+x_n),$$

donde $x_1, \ldots, x_n$ son números dados, pero no necesariamente iguales. Obsérvese que si todos los x_i fueran iguales (a x, por ejemplo) estaríamos en el caso del binomio.

[32] ¿Un poco? Que estamos en el capítulo 5, ¿cómo miden éstos el tiempo?, ¿topológicamente?

El desarrollo (en el orden más razonable) del producto ($\star$) da lugar, primero, a un 1, luego a cada uno de los sumandos del tipo x_i, después a todos los posibles sumandos $x_i x_j$ (con $i \neq j$ y sin repetirse), luego a los sumandos formados por los posibles tríos $x_i x_j x_k$, etc. Estas combinaciones son viejas conocidas; aparecieron en la discusión sobre medias de la sección 3.6. Se trata de las llamadas funciones simétricas elementales:

$$s_0(x_1,\dots x_n) = 1 \qquad \text{y} \qquad s_k(x_1,\dots x_n) = \sum_{i_1<\cdots<i_k} x_{i_1}\cdots x_{i_k} \quad \text{para cada } k=1,\dots,n.$$

Cada s_k suma precisamente todos los productos de los x_i tomados en grupos de k. Por ejemplo, $s_1(x_1,\dots,x_n) = x_1+\cdots+x_n$, $s_2(x_1,\dots,x_n) = x_1x_2+x_1x_3+\cdots+x_{n-1}x_n$, etc. La última es simplemente $s_n(x_1,\dots,x_n) = x_1\cdots x_n$. Nótese que cada s_k contiene justamente $\binom{n}{k}$ sumandos y que $s_k(x,\dots,x) = \binom{n}{k}x^k$ para cualquier $k \geq 0$.

De manera que el producto ($\star$) se puede escribir como

$$(\star\star) \qquad \prod_{i=1}^{n}(1+x_i) = \sum_{k=0}^{n} s_k(x_1,\dots,x_n).$$

Si todos los x_i son iguales a x, cada s_k se transforma en $\binom{n}{k}x^k$, y la expresión anterior es justamente el binomio de Newton.

Con unas pequeñas manipulaciones adicionales, podemos también dar una expresión cerrada y conveniente para un producto del tipo

$$\prod_{i=1}^{n}(x-x_i),$$

donde ahora consideraremos a x como una variable, y a los $x_1,\dots,x_n$ como números dados. La observación pertinente es que las funciones simétricas cumplen una propiedad de *homogeneidad* (de grado k): para $\alpha \in \mathbb{R}$, $s_k(\alpha x_1,\dots,\alpha x_n) = \alpha^k s_k(x_1,\dots,x_n)$. Compruébelo el lector. Usando esto y ($\star\star$), podemos escribir

$$\begin{aligned}\boxed{\prod_{i=1}^{n}(x-x_i)} &= x^n \prod_{i=1}^{n}\Big(1+\frac{-x_i}{x}\Big) = x^n \sum_{k=0}^{n} s_k(-x_1/x,\dots,-x_n/x) \\ &= x^n \sum_{k=0}^{n} \frac{(-1)^k}{x^k}\; s_k(x_1,\dots,x_n) = \boxed{\sum_{k=0}^{n}(-1)^k\; s_k(x_1,\dots,x_n)\, x^{n-k}}\end{aligned}$$

Una expresión que no está nada mal. El lector familiarizado con manejos de polinomios habrá ya observado que el producto $\prod_{i=1}^{n}(x-x_i)$ es la manera natural de escribir un polinomio de grado n y mónico (es decir, tal que el coeficiente de x^n, el término de mayor grado, es 1) si los números x_i son sus *raíces*. La fórmula anterior nos da los *coeficientes* del polinomio en términos de las funciones simétricas evaluadas en las citadas raíces. Sacaremos partido de esta identidad algo más adelante (apartado 7.2.4).

Aquí vamos a emplear estas identidades para probar las desigualdades de Newton–Maclaurin para las medias simétricas, que presentamos al lector en el lema 3.6.4, y que recordamos ahora. Dada una lista de números *no negativos* $(x_1, \dots, x_n)$, las cantidades

$$S_k(x_1, \dots, x_n) = \frac{1}{\binom{n}{k}} s_k(x_1, \dots, x_n) \qquad \text{para cada } k = 0, 1, \dots, n$$

cumplen que

$$S_1(x_1, \dots, x_n) \geq S_2(x_1, \dots, x_n)^{1/2} \geq S_3(x_1, \dots, x_n)^{1/3} \geq \cdots \geq S_n(x_1, \dots, x_n)^{1/n}.$$

Las cantidades primera y última,

$$S_1(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \text{y} \qquad S_n(x_1, \dots, x_n)^{1/n} = \sqrt[n]{x_1 \cdots x_n}$$

son las habituales medias aritmética y geométrica de los números $x_1, \dots, x_n$.

En la demostración usaremos una propiedad que verifican ciertas sucesiones logcóncavas. Recordamos, definición 5.1.4, que una sucesión $(a_0, a_1, \dots, a_n)$ es logcóncava si

$$a_k^2 \geq a_{k-1} \cdot a_{k+1} \qquad \text{para todo } k = 1, \dots, n-1.$$

Lema 5.1.17 *Sea $(a_0, a_1, \dots, a_n)$ una sucesión de números no negativos, logcóncava y tal que $a_0 = 1$. Entonces*

$$a_1 \geq a_2^{1/2} \geq a_3^{1/3} \geq \cdots \geq a_n^{1/n}.$$

El lema 5.1.17 es aplicable, entre otros muchos ejemplos, a la sucesión de coeficientes binómicos $\binom{n}{k}$, que es logcóncava, como ya comprobamos en el apartado 5.1.1, y cumple que $\binom{n}{0} = 1$. Por ejemplo, si tomamos $n = 6$, se tiene que

$$\binom{6}{1} \geq \binom{6}{2}^{1/2} \geq \binom{6}{3}^{1/3} \geq \binom{6}{4}^{1/4} \geq \binom{6}{5}^{1/5} \geq \binom{6}{6}^{1/6}.$$

Verifíquelo numéricamente, lector. ¡Vaya con la logconcavidad! Por cierto, la condición de que $a_0 = 1$, en apariencia menor, resulta ser crucial, en el sentido de que sin ella pudiera no ser cierta la conclusión del lema 5.1.17. Como ejemplo, considere el lector las sucesiones

$$(1, 3, 3, 1), \qquad (10, 30, 30, 10) \qquad \text{y} \qquad \Big(\frac{1}{10}, \frac{3}{10}, \frac{3}{10}, \frac{1}{10}\Big);$$

la primera está formada por los números $\binom{3}{k}$, para $k = 0, 1, 2, 3$, mientras que las otras dos se obtienen de ésta por cambio de escala: multiplicar y dividir por 10, respectivamente. Las tres son sucesiones logcóncavas, pero mientras que las dos primeras cumplen la conclusión del lema 5.1.17, la tercera no, puesto que $a_1 = 3/10 = 0.3$ es menor que $a_2^{1/2} = \sqrt{3/10} \approx 0.55$.

Demostración del lema 5.1.17. Supongamos que la sucesión $(a_0 = 1, a_1, \dots, a_n)$ cumple

$$(\star) \qquad a_k^2 \geq a_{k-1} \cdot a_{k+1} \qquad \text{para todo } k = 1, \dots, n-1.$$

El caso $k = 1$ de $(\star)$ nos dice que

$$a_0 \cdot a_2 \leq a_1^2 \quad \Longrightarrow \quad a_2^{1/2} \leq a_1,$$

usando que $a_0 = 1$, lo que nos da la primera desigualdad de la cadena.

Ahora, usando los casos $k = 1$ y $k = 2$ de $(\star)$ conjuntamente,

$$(a_0 \cdot a_2)\,(a_1 \cdot a_3)^2 \le a_1^2\,a_2^4 \quad\Longrightarrow\quad a_0\,a_1^2\,a_2\,a_3^2 \le a_1^2\,a_2^4 \quad\Longrightarrow\quad a_3^2 \le a_2^3 \quad\Longrightarrow\quad a_3^{1/3} \le a_2^{1/2}.$$

Observe, lector, cuán hábilmente hemos elegido la combinación (producto) de coeficientes a_i en el arranque. Luego escribiríamos, usando $(\star)$ con $k = 1, 2, 3$,

$$(a_0 \cdot a_2)\,(a_1 \cdot a_3)^2\,(a_2 \cdot a_4)^3 \le a_1^2\,a_2^4\,a_3^6 \quad\Longrightarrow\quad a_0\,a_1^2\,a_2^4\,a_3^2\,a_4^3 \le a_1^2\,a_2^4\,a_3^6,$$

de donde, tras cancelar factores, se obtiene que $a_4^3 \le a_3^4$, que nos da el siguiente caso. Y así en general, porque de que

$$(a_0 \cdot a_2)\,(a_1 \cdot a_3)^2\,(a_2 \cdot a_4)^3 \cdots (a_{k-2} \cdot a_k)^{k-1} \cdot (a_{k-1} \cdot a_{k+1})^k \le a_1^2\,a_2^4\,a_3^6 \cdots a_{k-1}^{2k-2}\,a_k^{2k}$$

se deduce que $a_{k+1}^k \le a_k^{k+1}$. ■

Volvemos al asunto que nos ocupa, que es la demostración de las desigualdades de Newton–Maclaurin para las medias simétricas del lema 3.6.4, y que se completaría si comprobáramos que la sucesión de medias simétricas cumple las hipótesis del lema 5.1.17. Como $S_0(x_1, \ldots, x_n) = 1$, sólo resta verificar que:

Teorema 5.1.18 (Newton) *Para cualquier lista $(x_1, \ldots, x_n)$ de números reales no negativos, la sucesión de medias simétricas $(S_k(x_1, \ldots, x_n))_{k=0}^n$ es logcóncava. Es decir,*

$$S_k^2(x_1, \ldots, x_n) \ge S_{k-1}(x_1, \ldots, x_n) \cdot S_{k+1}(x_1, \ldots, x_n) \qquad \text{para todo } k = 1, \ldots, n-1.$$

La igualdad se alcanza si y sólo todos los x_i son iguales.

Demostración. El objetivo es probar que, para cualquier entero $n \ge 2$ y para cualquier lista $\mathcal{F} = (x_1, \ldots, x_n)$ de números reales no negativos,

$$(\star) \qquad S_k^2(\mathcal{F}) \ge S_{k-1}(\mathcal{F}) \cdot S_{k+1}(\mathcal{F}) \qquad \text{para cada } k = 1, \ldots, n-1.$$

Haremos la prueba de $(\star)$ por inducción en n, que es el tamaño de la lista $\mathcal{F}$, y para ello necesitaremos dos ingredientes: por un lado, probar el caso $k = 1$ de $(\star)$ para cualquier $n \ge 2$, no tanto para arrancar la inducción en sí (aunque también), sino sobre todo para (curiosamente) dar cuenta del caso $k = n - 1$, el último, que se escapa del argumento de inducción; y, por otro, una observación sobre polinomios y sus derivadas que será la clave en la prueba por inducción propiamente dicha.

1) Fijamos $n \ge 2$, una lista $\mathcal{F} = (x_1, \ldots, x_n)$ cualquiera de números no negativos, y consideramos el caso $k = 1$ de $(\star)$, que se escribe, aprovechando que $S_0(\mathcal{F}) = 1$, de la siguiente manera:

$$(\dagger) \qquad S_1^2(\mathcal{F}) \ge S_2(\mathcal{F}),$$

es decir,

$$\Big(\frac{1}{n}\sum_{i=1}^n x_i\Big)^2 \ge \frac{1}{\binom{n}{2}} \sum_{1\le i<j\le n} x_i x_j\,, \qquad \text{o bien} \qquad (n-1)\Big(\sum_{i=1}^n x_i\Big)^2 \ge 2n \sum_{1\le i<j\le n} x_i x_j,$$

tras simplificar y reordenar factores. Ahora, desarrollando el cuadrado de que aparece en la última expresión, la desigualdad es equivalente a

$$(n-1)\Big(\sum_{i=1}^{n} x_i^2 + 2\sum_{1\leq i<j\leq n} x_i x_j\Big) \geq 2n \sum_{1\leq i<j\leq n} x_i x_j \iff (n-1)\sum_{i=1}^{n} x_i^2 - 2\sum_{1\leq i<j\leq n} x_i x_j \geq 0.$$

Finalmente, la última desigualdad (y por tanto también (†)) es equivalente a (la claramente cierta) desigualdad

$$\sum_{1\leq i<j\leq n} (x_i - x_j)^2 \geq 0,$$

como se comprueba escribiendo el lado izquierdo como

$$\sum_{1\leq i<j\leq n} (x_i - x_j)^2 = \sum_{1\leq i<j\leq n} (x_i^2 + x_j^2) - 2\sum_{1\leq i<j\leq n} x_i x_j,$$

y luego observando que en la primera suma doble de la derecha aparece cada sumando x_i^2 justamente $n-1$ veces. Lector, hágase una tabla, si fuera menester, para convencerse.

2) Llamemos $p_{\mathcal{F}}(x)$ al polinomio de grado n mónico cuyas raíces son los números reales de la lista $\mathcal{F}$. Ya hemos visto que podemos escribir

$$p_{\mathcal{F}}(x) = \prod_{i=1}^{n}(x - x_i) = \sum_{k=0}^{n}(-1)^k \binom{n}{k} S_k(\mathcal{F})\, x^{n-k}.$$

El teorema de Rolle[33] nos dice que $p'_{\mathcal{F}}(x)$, el polinomio que se obtiene al derivar $p_{\mathcal{F}}(x)$, tiene también raíces reales, aunque una menos, claro. Llamemos $\mathcal{F}' = (y_1, \dots, y_{n-1})$ a esas raíces.

Podemos, por un lado, obtener una expresión para $p'_{\mathcal{F}}(x)$ derivando el desarrollo de $p_{\mathcal{F}}(x)$:

$$\frac{p'_{\mathcal{F}}(x)}{n} = \sum_{k=0}^{n-1}(-1)^k \binom{n}{k} S_k(\mathcal{F})\, \frac{n-k}{n}\, x^{n-k-1}.$$

Note el lector que hemos dividido por n en ambos lados de la expresión resultante.

Por otro lado, el polinomio $p'_{\mathcal{F}}(x)$, que es de grado $n-1$, cumple que el coeficiente de x^{n-1} es n. De manera que $p'_F(x)/n$ es mónico, y como tal, admitirá un desarrollo análogo a los anteriores:

$$\frac{p'_{\mathcal{F}}(x)}{n} = \prod_{i=1}^{n-1}(x - y_i) = \sum_{k=0}^{n-1}(-1)^k \binom{n-1}{k} S_k(\mathcal{F}')\, x^{n-k-1}.$$

Comparando coeficientes, y observando que

$$\binom{n}{k}\frac{n-k}{n} = \frac{n!}{k!\,(n-k)!}\,\frac{n-k}{n} = \frac{(n-1)!}{k!\,(n-k-1)!} = \binom{n-1}{k},$$

[33] Corra a buscar su libro de Cálculo, lector.

deducimos que

$$(\ddagger) \qquad S_k(\mathcal{F}) = S_k(\mathcal{F}') \qquad \text{para cada } k = 0, 1, \dots, n-1.$$

3) Tras estos preliminares, arrancamos la prueba de ($\star$) por inducción en n, el tamaño de la lista $\mathcal{F}$. El paso $n = 2$ inicial,

$$S_1^2(x_1, x_2) \geq S_0(x_1, x_2) \cdot S_2(x_1, x_2),$$

es inmediato, pues está incluido en (†), o más directamente, porque significa que $(x_1+x_2)/2 \geq \sqrt{x_1 \cdot x_2}$, que es la habitual desigualdad entre la media aritmética y la geométrica de dos números (nótese que esta última es consecuencia de que $(x_1 - x_2)^2 \geq 0$).

Supongamos entonces que la condición ($\star$) es cierta para cualquier lista de longitud $j \leq n-1$, y consideremos una lista $\mathcal{F} = (x_1, \dots, x_n)$ de longitud n. Asociada a la lista $\mathcal{F}$ tenemos la lista $\mathcal{F}'$ descrita antes, que tiene un elemento menos. Para ella se cumplirá, por hipótesis de inducción, que

$$S_k^2(\mathcal{F}') \geq S_{k-1}(\mathcal{F}') \cdot S_{k+1}(\mathcal{F}') \qquad \text{para cada } k = 1, \dots, n-2.$$

Pero la expresión (‡) nos dice que todo lo que aparece en la línea anterior es cierto si cambiamos $\mathcal{F}'$ por $\mathcal{F}$.

4) La demostración habría terminado si no fuera porque falta por comprobar un caso,

$$S_{n-1}^2(\mathcal{F}) \geq S_{n-2}(\mathcal{F}) \cdot S_n(\mathcal{F}),$$

que no se recoge en las desigualdades anteriores, y que hay que probar por separado. Como vamos a ver, esta desigualdad se sigue de (†), aunque no aplicada a la lista $\mathcal{F} = (x_1, \dots, x_n)$, sino a la lista de los *recíprocos* $(1/x_1, \dots, 1/x_n)$.

Observe el lector que en $S_{n-1}(\mathcal{F})$ aparecen todos los posibles productos que contienen a todas las variables menos una. Y en $S_{n-2}(\mathcal{F})$, los productos que contienen a todas las variables menos dos. Por su parte, $S_n(\mathcal{F})$ es, simplemente, $(x_1 \cdots x_n)$.

Reescribimos la expresión anterior con una notación especial,

$$\Big(\frac{1}{n} \sum_{j=1}^n x_1 \cdots \widehat{x}_j \cdots x_n\Big)^2 \geq \frac{1}{\binom{n}{2}} \Big[\sum_{j<k} x_1 \cdots \widehat{x}_j \cdots \widehat{x}_k \cdots x_n\Big] \cdot (x_1 \cdots x_n),$$

donde las variables $\widehat{x}$ denotan justamente las que *faltan* en cada producto. Si alguno de los x_i fuera 0, la desigualdad anterior sería evidente, pues a la derecha tendríamos un 0, y una cantidad no negativa a la izquierda. Así que podemos suponer que $x_1 \cdots x_n \neq 0$. Dividiendo entonces por $(x_1 \cdots x_n)^2$ ambos lados de la desigualdad, el lector puede comprobar que nos queda la expresión

$$\Big(\frac{1}{n} \sum_{j=1}^n \frac{1}{x_j}\Big)^2 \geq \frac{1}{\binom{n}{2}} \sum_{j<k} \frac{1}{x_j \, x_k},$$

que no es sino

$$S_1^2(1/x_1, \dots, 1/x_n) \geq S_2(1/x_1, \dots, 1/x_n).$$

desigualdad que es cierta, como ya vimos en (†). ■

5.1.3. Algunas aplicaciones combinatorias de los coeficientes binómicos

En las últimas páginas los coeficientes binómicos han estado pavoneándose sin recato, exhibiendo algunas de sus curiosas propiedades aquí y allá. Es hora ya de ponerlos a trabajar. En este apartado, y en los dos siguientes (5.1.4 y 5.1.5), analizamos unas cuantas cuestiones combinatorias en cuya respuesta intervienen, de una manera u otra, los coeficientes binómicos.

Las primeras de estas aplicaciones tienen que ver con contar el **número de soluciones de la ecuación diofántica**

$$x_1 + x_2 + \cdots + x_k = n.$$

Aquí, los datos son n y k. El adjetivo *diofánticas*[34] hace alusión a que las soluciones de la ecuación anterior han de ser números enteros no negativos. Es decir, una solución de la ecuación anterior es una lista $(x_1, \dots, x_n)$ de números enteros no negativos que suman n. En las aplicaciones habituales, a los x_j se les suelen exigir condiciones adicionales, como ser todos mayores o iguales que 1, o similares.

El enunciado anterior es pura abstracción que recoge los ingredientes fundamentales de diversos problemas combinatorios, algunos de los cuales pasamos a enunciar.

Cuestión 1. Composiciones del entero n de longitud k.

Como vimos en la sección 2.2.3, una composición de n es una manera de escribir n como suma (ordenada) de números naturales. La composición tendrá longitud k si hay exactamente k sumandos. Así que, si llamamos $x_1, \dots, x_k$ a estos sumandos, contar el número de estas composiciones es lo mismo que calcular el número de soluciones de

$$\left\{ \begin{array}{c} x_1 + x_2 + \cdots + x_k = n \\ x_1 \geq 1, \dots, x_k \geq 1 \end{array} \right\}$$

Cuestión 2. Distribuciones de n bolas *idénticas* en k cajas *numeradas*.

Es ésta la primera vez que aparece en este texto el lenguaje de las distribuciones de bolas en cajas, un lenguaje que, como se irá viendo, resulta extremadamente útil en la representación de múltiples cuestiones. En esta jerga es habitual tratar con bolas idénticas (indistinguibles), o con bolas numeradas (distinguibles); y lo mismo para las cajas. De nuevo, esta representación es una abstracción que puede adaptarse a diversos problemas combinatorios. En la sección 5.4 resumiremos la casuística que irá apareciendo en las páginas que siguen.

Nos interesa aquí saber de cuántas maneras distintas se pueden distribuir las n bolas (idénticas) en las k cajas (numeradas). Note, lector, que como las bolas son idénticas, lo único relevante para ese recuento es decidir *cuántas* bolas van en cada caja. Así que, llamando x_j al número de bolas que va en cada caja j, contar el número de distribuciones es lo mismo (con biyección por medio) que contar el número de soluciones de

$$\left\{ \begin{array}{c} x_1 + x_2 + \cdots + x_k = n \\ x_1 \geq 0, \dots, x_k \geq 0 \end{array} \right\} \qquad \text{o quizás de} \qquad \left\{ \begin{array}{c} x_1 + x_2 + \cdots + x_k = n \\ x_1 \geq 1, \dots, x_k \geq 1 \end{array} \right\}$$

[34]Estudiaremos este tipo de ecuaciones con detalle en el capítulo 6, en especial en la sección 6.1.3. Aquí no nos interesaremos por cómo resolverlas, sino por saber cuántas soluciones distintas tienen.

si es que permitimos cajas vacías (caso de la izquierda), o si no permitimos que queden cajas vacías (caso de la derecha). Por supuesto, el lector puede pergeñar otras restricciones sobre las distribuciones de bolas, que se traducen de manera inmediata en condiciones sobre los x_j.

Cuestión 3. Multiconjuntos de tamaño k extraídos de $\{1,\dots,n\}$.

Claro, ¡no podían faltar!, dirá el lector: si en las listas distinguíamos entre aquellas en las que se permite repetición de las que no, ¿por qué no hacer lo mismo con los conjuntos? En un subconjunto, de manera natural, la repetición de símbolos está prohibida. Y aunque suene algo forzado, definiremos un **multiconjunto** de tamaño k extraído de $\{1,\dots,n\}$ como una colección de k símbolos escogidos de $\{1,\dots,n\}$ donde se permite que aparezca varias veces cada símbolo.

Antes de caer en el desánimo por la vertiginosa proliferación de problemas con y sin repetición, con y sin orden, con y sin... observe el lector (o medite un rato hasta convencerse, levitando místicamente si fuere necesario) que, para describir un multiconjunto, lo único relevante es decidir *cuántas* veces aparece cada símbolo. Es decir, podemos representar un multiconjunto general de la siguiente manera:

$$\{1,\overset{x_1\text{ veces}}{\dots},1,2,\overset{x_2\text{ veces}}{\dots},2,\dots,n,\overset{x_n\text{ veces}}{\dots},n\}\,, \quad \longleftrightarrow \quad \{1_{x_1},2_{x_2},\dots,n_{x_n}\}\,.$$

Los x_j son enteros no negativos, y su suma ha de valer k (que es el tamaño total del multiconjunto). Así que el número de multiconjuntos (de tamaño k y con símbolos en $\{1,\dots,n\}$) coincide con el número de soluciones de

$$\#\text{ soluciones de}\left\{\begin{array}{c} x_1+x_2+\cdots+x_n=k \\ x_1\ge 0\,,\dots\,,x_n\ge 0\end{array}\right\},$$

que es un problema como los anteriores, salvo que los papeles de n y k se han intercambiado con respecto a ellos.

Una vez convencidos de que vale la pena dar solución a la cuestión sobre soluciones de ecuaciones diofánticas, pues de esa manera daríamos solución, de una tacada, a estas tres cuestiones, nos ponemos con ello.

Empezamos el análisis recurriendo al lenguaje de las composiciones. Revise el lector el argumento que utilizamos en la sección 2.2.3, y que entonces nos permitió deducir que el número total de composiciones de n era 2^{n-1}: una composición se obtiene colocando los n unos en fila y decidiendo, en cada hueco entre ellos (hay $n-1$) si se para (y se suma todo lo que se haya ido acumulando) o se sigue adelante. Es decir, es lo mismo que una lista de longitud $n-1$ formada con dos símbolos. Si ahora la composición ha de tener k sumandos, querrá decir que en la lista anterior ha de aparecer exactamente $k-1$ veces el separador que representa el "para y suma" (¿por qué $k-1$?). De manera que nos basta con escoger en qué $k-1$ huecos (de los $n-1$ posibles) van estos separadores. Las biyecciones implícitas en este argumento nos permiten deducir que

$$\begin{aligned}\#\left\{\begin{array}{c}\text{composiciones de lon-}\\ \text{gitud } k \text{ del número } n\end{array}\right\} &= \#\left\{\begin{array}{c}\text{formas de elegir } k-1 \text{ posiciones (para los separadores)}\\ \text{de entre } n-1 \text{ (los huecos a nuestra disposición)}\end{array}\right\}\\ &= \#\left\{\begin{array}{c}\text{subconjuntos distintos de tamaño } k-1\\ \text{extraídos del conjunto } \{1,\dots,n-1\}\end{array}\right\} = \binom{n-1}{k-1}.\end{aligned}$$

Y de forma inmediata, vista la traducción a ecuaciones diofánticas, que

$$\boxed{\#\text{ soluciones de }\left\{\begin{array}{c} x_1+x_2+\cdots+x_k=n \\ x_1\geq 1,\dots,x_k\geq 1\end{array}\right\} = \binom{n-1}{k-1}}$$

Sigamos adelante: en las cuestiones combinatorias descritas al principio aparecía también el problema de contar el número de soluciones de

$$\left\{\begin{array}{c} x_1+\cdots+x_k=n \\ x_i\geq 0\end{array}\right\}.$$

Por ejemplo, al contar las distribuciones de n bolas idénticas en k cajas numeradas, si es que permitimos cajas vacías.

Vamos a transformar este problema en uno del tipo anterior (con restricciones sobre las variables del tipo ≥ 1) con una sencilla biyección de corte algebraico. A cada lista $(x_1,\dots,x_k)$ solución del problema anterior le asociaremos una lista $(y_1,\dots,y_k)$ con la receta $y_j=x_j+1$ para cada j. Es claro que los y_j son números ≥ 1. Y su suma vale lo que sumaban los x_j más el número de unos añadidos, es decir, $n+k$. Esta biyección nos permite concluir que

$$\boxed{\#\text{ sols. de}\left\{\begin{array}{c} x_1+\cdots+x_k=n \\ x_i\geq 0\end{array}\right\}} = \#\text{ sols. de}\left\{\begin{array}{c} y_1+\cdots+y_k=n+k \\ y_i\geq 1\end{array}\right\} = \boxed{\binom{n+k-1}{k-1}}$$

Quizás el lector quiera reescribir este argumento en términos de bolas en cajas (quitando una bola de cada caja si es que todas tienen al menos una; o, en sentido contrario, añadiendo una bola a cada caja en una distribución general para obtener una con cajas no vacías). Por cierto que el número de k-multiconjuntos con símbolos $\{1,\dots,n\}$ de la cuestión 3 resulta ser, leyendo adecuadamente la fórmula anterior, $\binom{n+k-1}{n-1}$.

Visto el éxito conseguido, nos venimos arriba y nos animamos a plantear una versión más general del problema diofántico, como es calcular el número de soluciones de

$$\left\{\begin{array}{c} x_1+x_2+\cdots+x_k=n \\ x_1\geq p_1,\dots,x_k\geq p_k\end{array}\right\}.$$

Alternativamente, contar las distribuciones de n bolas idénticas en k cajas numeradas en las que en la caja 1 van al menos p_1 bolas, en la caja 2 al menos p_2, etc. Los datos aquí son n, k y unos enteros no negativos $p_1,\dots,p_k$.

Resolvemos la cuestión con un truco como el de antes, transformándola en una que ya sabemos contar, como por ejemplo aquella en el que las restricciones son del tipo ≥ 0. Para ello, empleamos la biyección $y_j=x_j-p_j$ para cada $j=1,\dots,k$. Los y_j son enteros mayores o iguales que 0 (pues los $x_j\geq p_j$) que suman

$$\sum_{j=1}^k y_j=\sum_{j=1}^k (x_j-p_j)=\sum_{j=1}^k x_j-\sum_{j=1}^k p_j=n-\sum_{j=1}^k p_j\,.$$

De manera que

$$\#\text{ sols de}\left\{\begin{array}{c} x_1+\cdots+x_k=n \\ x_1\geq p_1,\ldots,x_k\geq p_k \end{array}\right\} = \#\text{ sols de}\left\{\begin{array}{c} y_1+\cdots+y_k=n-\sum_{j=1}^k p_j \\ y_1\geq 0,\ldots,y_k\geq 0 \end{array}\right\},$$

y, aplicando la fórmula de arriba concluimos lo siguiente, que elevamos a a categoría de lema para ulteriores referencias.

Lema 5.1.19 *Dados $n, k \geq 1$ y unos enteros no negativos $p_1, \ldots, p_k$,*

$$\#\textit{ soluciones de}\left\{\begin{array}{c} x_1+\cdots+x_k=n \\ x_1\geq p_1,\ldots,x_k\geq p_k \end{array}\right\} = \binom{n+k-1-\sum_{j=1}^k p_j}{k-1}$$

La fórmula es aparatosa, pero de aplicación directa, si es que conocemos los valores de k, n y $p_1, \ldots, p_k$. Para recordarla, le proponemos al lector la siguiente regla mnemotécnica: el coeficiente binómico tiene, como índice superior, el valor total de la suma (que en el enunciado es n), más el número (k) de sumandos, menos 1, y menos lo que sumen las restricciones (los p_j); y como índice inferior, el número de sumandos menos 1.

De nuevo animamos al lector a que reinterprete el argumento biyectivo en términos de quitar o añadir bolas en cajas. Observe el lector que si $\sum_j p_j > n$ no hay soluciones de la ecuación diofántica del lema 5.1.19, pero de esto ya da cuenta el convenio habitual de los coeficientes binómicos.

La fórmula anterior contiene, como casos particulares, los dos resultados obtenidos antes:

$$\begin{array}{rc} \text{si } p_1=p_2=\cdots p_k=0, \text{ entonces la respuesta es} & \displaystyle\binom{n+k-1}{k-1}; \\ \text{y si } p_1=p_2=\cdots=p_k=1, \text{ entonces es simplemente} & \displaystyle\binom{n-1}{k-1}. \end{array}$$

El lector ambicioso estará ya esperando la siguiente generalización: dadas dos listas de enteros no negativos $(p_1, \ldots, p_k)$ y $(q_1, \ldots, q_k)$, contar el número de soluciones de

$$\left\{\begin{array}{c} x_1+x_2+\cdots+x_k=n \\ p_1\leq x_1\leq q_1,\ldots,p_k\leq x_k\leq q_k \end{array}\right\}.$$

Ahora hay restricciones sobre los x_j por arriba y por debajo. En la formulación de bolas y cajas, cada caja j llevaría entre p_j y q_j bolas. Aunque el aspecto del problema es muy similar al de los anteriores, en esta ocasión, como veremos, no dispondremos de una fórmula "sencilla" en términos de los parámetros involucrados, pues la solución pasa por aplicar el siempre aparatoso principio de inclusión/exclusión. Lo ilustramos con un ejemplo.

Ejemplo 5.1.1 *El número de soluciones de la ecuación diofántica $x_1+x_2+x_3=50$, donde $0\leq x_1\leq 10$, $5\leq x_2\leq 35$ y $0\leq x_3\leq 20$.*

Iniciamos el análisis, para facilitar los cálculos, poniendo todas las cotas inferiores a cero:

$$y_1=x_1\geq 0\,,\quad y_2=x_2-5\geq 0\,,\quad y_3=x_3\geq 0\,.$$

Con esta transformación, el problema pasa a ser el de contar el número de soluciones de

$$(\star) \qquad \left\{ \begin{array}{c} y_1 + y_2 + y_3 = 45 \\ 0 \leq y_1 \leq 10\,,\ 0 \leq y_2 \leq 30\,,\ 0 \leq y_3 \leq 20 \end{array} \right\}.$$

Pasaremos ahora al complementario y aplicaremos el principio de inclusión/exclusión. Definimos a continuación los conjuntos de soluciones de interés, al tiempo que iremos registrando sus tamaños, usando la fórmula del lema 5.1.19. Primero, el conjunto en el que se hallan las soluciones de $(\star)$ y en el que definimos su complementario:

$$\mathcal{X} = \left\{ \text{soluciones de } \left\{ \begin{array}{c} y_1 + y_2 + y_3 = 45 \\ y_i \geq 0 \end{array} \right\} \right\} \Longrightarrow |\mathcal{X}| = \binom{45+3-1}{3-1} = \binom{47}{2}.$$

Obsérvese que una lista de $\mathcal{X}$ no es solución de $(\star)$, bien porque $y_1 \geq 11$, bien porque $y_2 \geq 31$, o bien porque $y_3 \geq 21$. Ésas son todas las posibilidades. De manera que si definimos los conjuntos $\mathcal{A}$, $\mathcal{B}$ y $\mathcal{C}$ (de listas "prohibidas")

$$\mathcal{A} = \left\{ \begin{array}{c} \text{soluciones} \\ \text{de} \end{array} \left\{ \begin{array}{c} y_1 + y_2 + y_3 = 45 \\ y_1 \geq 11\,, y_2 \geq 0\,, y_3 \geq 0 \end{array} \right\} \right\} \Rightarrow |\mathcal{A}| = \binom{45+3-11-1}{3-1} = \binom{36}{2},$$

$$\mathcal{B} = \left\{ \begin{array}{c} \text{soluciones} \\ \text{de} \end{array} \left\{ \begin{array}{c} y_1 + y_2 + y_3 = 45 \\ y_1 \geq 0\,, y_2 \geq 31\,, y_3 \geq 0 \end{array} \right\} \right\} \Rightarrow |\mathcal{B}| = \binom{45+3-31-1}{3-1} = \binom{16}{2},$$

$$\mathcal{C} = \left\{ \begin{array}{c} \text{soluciones} \\ \text{de} \end{array} \left\{ \begin{array}{c} y_1 + y_2 + y_3 = 45 \\ y_1 \geq 0\,, y_2 \geq 0\,, y_3 \geq 21 \end{array} \right\} \right\} \Rightarrow |\mathcal{C}| = \binom{45+3-21-1}{3-1} = \binom{26}{2},$$

el número de soluciones de $(\star)$ será

$$|\mathcal{X}| - |\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}|.$$

Calculamos $|\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}|$ con el principio de inclusión/exclusión, lo que exige conocer, además de los tamaños de cada conjunto, que ya hemos calculado, el tamaño de las intersecciones de dos y tres conjuntos. Por ejemplo,

$$\mathcal{A} \cap \mathcal{B} = \left\{ \begin{array}{c} \text{soluciones} \\ \text{de} \end{array} \left\{ \begin{array}{c} y_1 + y_2 + y_3 = 45 \\ y_1 \geq 11\,, y_2 \geq 31\,, y_3 \geq 0 \end{array} \right\} \right\} \Rightarrow |\mathcal{A} \cap \mathcal{B}| = \binom{45+3-42-1}{3-1} = \binom{5}{2}.$$

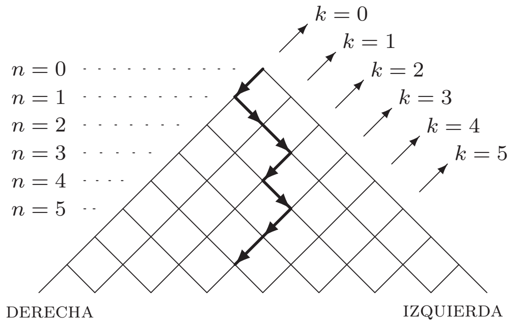
De manera análoga, se obtiene que $|\mathcal{A} \cap \mathcal{C}| = \binom{15}{2}$. La intersección $\mathcal{B} \cap \mathcal{C}$ resulta ser vacía, porque las condiciones $y_2 \geq 31$ e $y_3 \geq 21$ exceden el valor total de la suma, 45; aunque instamos al lector a que aplique fielmente la fórmula con coeficientes binómicos y compruebe que obtiene un índice superior negativo, y por tanto que $|\mathcal{B} \cap \mathcal{C}| = 0$. Como consecuencia de esto, la intersección tres a tres será también vacía. El resultado final es

$$\#\,\text{soluciones} = \binom{47}{2} - \left[\binom{36}{2} + \binom{16}{2} + \binom{26}{2}\right] + \left[\binom{5}{2} + \binom{15}{2}\right] = 121. \qquad \clubsuit$$

Más adelante, por ejemplo en el análisis de *particiones de enteros* del apartado 5.3.3, aparecerán ecuaciones diofánticas de naturaleza algo distinta a las vistas aquí, y para cuyo estudio resulta especialmente conveniente utilizar el lenguaje de las funciones generatrices (capítulo 12 y siguientes).

5.1.4. El caminante de Pólya y los coeficientes binómicos

En la red que aparece dibujada a la derecha, cada nodo se identifica con unas coordenadas (n, k): las filas se etiquetan con n y las diagonales con k, tal y como se indica en la figura. Nuestro objetivo es estudiar los posibles caminos desde $(0, 0)$ hasta un cierto (n, k) tales que, en cada paso, desde cada nodo sólo se puede pasar a los dos nodos que están inmediatamente debajo. La longitud del camino será el número de pasos dados.



Para cada par de valores n y k, llamaremos $\mathrm{Cam}(n, k)$ al número de caminos distintos que podemos trazar desde el nodo $(0, 0)$ a un nodo cualquiera de coordenadas (n, k). Distintos, por supuesto, significa que las sucesiones de pasos de que constan se diferencian en alguno de ellos. Para fijar ideas, al describir estos caminos hablaremos de "paso a la derecha" y "paso a la izquierda" adoptando el punto de vista de un caminante que circulara por la red (así que es justamente la orientación *contraria* a la del lector que lee estas páginas). Esta formulación, que como veremos es muy rica y elegante, se debe al matemático húngaro Pólya[35].

Figura 5.4: Pólya

Observe el lector que $\mathrm{Cam}(n, 0) = 1$, para cualquier $n \geq 0$, porque sólo hay un camino que lleve del punto $(0, 0)$ al $(n, 0)$ (el que consiste en dar siempre pasos a la derecha). Análogamente, $\mathrm{Cam}(n, n) = 1$ (dar sólo pasos a la izquierda). El objetivo es encontrar, si es que la hay, una fórmula cerrada para $\mathrm{Cam}(n, k)$ en términos de n y k.

Quizás el lector, inspirado por el aspecto de la figura[36], pueda sospechar que estos números guardan relación con los coeficientes binómicos. Y así es. De hecho, para cada par de valores de los parámetros n y k,

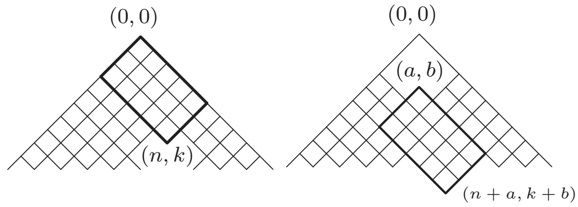
$$\mathrm{Cam}(n, k) = \binom{n}{k}.$$

La comprobación es directa: un camino de $(0, 0)$ a (n, k) es una lista de longitud n formada por dos símbolos, D e I (derecha e izquierda), de manera que hay exactamente k símbolos D. Y listas de ésas hay $\binom{n}{k}$, las maneras de elegir la ubicación de los k símbolos D. Esta reinterpretación de los coeficientes binómicos nos va a permitir una serie de identidades usando argumentos gráficos, como los que siguen.

[35] George Pólya (1887-1985) es uno de los miembros más destacados de la magnífica escuela matemática húngara del siglo XX. Su actividad se desarrolló en diversos lugares, primero en su Budapest natal (hasta 1912), luego en Göttingen (1913) y Zürich, desde 1914 hasta 1940 (en 1924 estuvo en Inglaterra, trabajando con Hardy y Littlewood; el libro *Inequalities* es uno de los frutos de esta colaboración). La siguiente etapa de esta suerte de paseo aleatorio vital (una estupenda biografía suya se titula justamente *The random walks of George Pólya*, Gerard Alexanderson, MAA, 2000) fue la Universidad de Stanford en Estados Unidos, a donde emigró en 1940, y en la que permanecería hasta el final de sus días. El inquieto Pólya trabajó en muy diversos campos de las matemáticas: teoría de números, combinatoria, análisis real y complejo, probabilidad, etc. Además de por esta actividad investigadora, Pólya es famoso por sus reflexiones sobre la actividad matemática y la didáctica de las matemáticas: libros como *How to solve it* (1945), *Mathematics and plausible reasoning* (1954) o *Mathematical Discovery* (1962) han sido auténticos éxitos editoriales.

[36] De innegable parecido con el triángulo de Pascal–Tartaglia. El título de la sección también ayuda.

La primera observación tiene que ver con **traslaciones de caminos**. Como se observa en la figura, existe una biyección entre los caminos de $(0,0)$ a (n,k) y los que van de (a,b) a $(n+a,k+b)$, pues simplemente hay que tomar cada camino y desplazar su punto de arranque al (a,b). Esto nos dice que

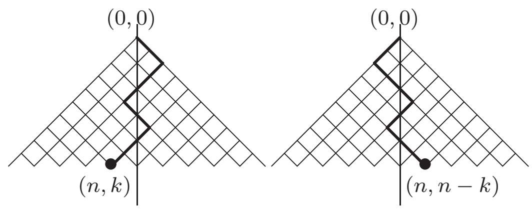


$$\binom{n}{k} = \#\big\{\text{caminos } (0,0) \to (n,k)\big\} = \#\big\{\text{caminos } (a,b) \to (a+n,b+k)\big\}.$$

La propiedad de simetría de los coeficientes binómicos,

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

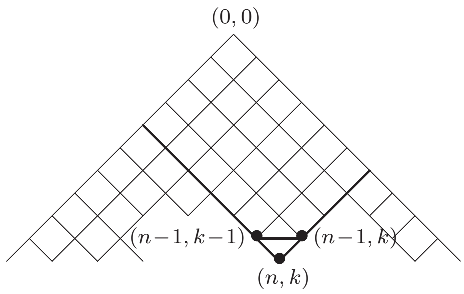
se corresponde, en este lenguaje, con un argumento de **reflexión**. Cada camino $(0,0) \to (n,k)$ se puede reflejar (con respecto al eje vertical que pasa por $(0,0)$) para obtener un camino de $(0,0)$ a $(n,n-k)$. Esta reflexión es una biyección entre el conjunto de caminos $(0,0) \to (n,k)$ y el de caminos $(0,0) \to (n,n-k)$.



La regla de recursión

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

se obtiene **clasificando** los caminos hasta (n,k) dependiendo del vértice inmediatamente superior por el que pasen, para obtener una partición del conjunto total de caminos:

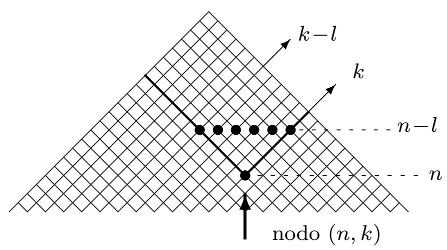


$$\left\{\begin{matrix}\text{caminos}\\ (0,0)\to(n,k)\end{matrix}\right\} = \left\{\begin{matrix}\text{caminos } (0,0)\to(n-1,k-1)\\ \text{y luego izquierda}\end{matrix}\right\} \cup \left\{\begin{matrix}\text{caminos } (0,0)\to(n-1,k)\\ \text{y luego derecha}\end{matrix}\right\}.$$

El argumento concluye observando que en los dos conjuntos de la derecha basta con contar de cuántas maneras se puede llegar a los puntos $(n-1,k-1)$ y $(n-1,k)$, respectivamente.

Ejemplo 5.1.2 *Clasificando caminos por barreras horizontales.*

Clasificamos los caminos hasta (n,k) según el nodo de la barrera horizontal de un nivel anterior $n-l$ por el que pasan (véase el dibujo). Los nodos de la barrera tienen coordenadas $(n-l, k-l+m)$, donde $m = 0,1,\dots,l$. Esto da lugar a una partición del conjunto de caminos, pues *todos* los caminos hasta (n,k) han de pasar *necesariamente* por uno (y sólo uno) de los nodos de la barrera.



Es decir,

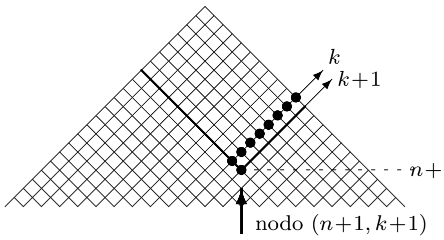
$$\left\{ \begin{array}{c} \text{caminos} \\ (0,0) \to (n,k) \end{array} \right\} = \bigcup_{m=0}^{l} \left\{ \begin{array}{c} \text{caminos consistentes en ir } (0,0) \to (n-l, k-l+m) \\ \text{y luego ir } (n-l, k-l+m) \to (n,k) \end{array} \right\}$$

A la izquierda tenemos $\binom{n}{k}$ caminos distintos. Del punto $(0,0)$ al punto $(n-l, k-l+m)$ se puede ir de $\binom{n-l}{k-l+m}$ maneras distintas. Y hay tantos caminos desde $(n-l, k-l+m)$ hasta (n,k) como entre $(0,0)$ y $(l, l-m)$, como podrá comprobar el lector si aplica cuidadosamente la regla de traslación de caminos que vimos antes. Aplicando las reglas de la suma y el producto, se llega a que

$$\binom{n}{k} = \sum_{m=0}^{l} \binom{l}{l-m}\binom{n-l}{k-l+m} \overset{[j=l-n]}{=} \sum_{j=0}^{l} \binom{l}{j}\binom{n-l}{k-j},$$

que es la fórmula de Vandermonde (lema 5.1.2). ♣

Ejemplo 5.1.3 *Clasificando caminos por barreras diagonales.*



Para un cierto nodo de la red, que por comodidad tomaremos de coordenadas $(n+1, k+1)$, consideramos la barrera diagonal que incluye a los nodos de de coordenadas (j,k), con $j = k, \dots, n$ (véase el dibujo). Como antes, queremos clasificar los caminos de $(0,0)$ a $(n+1, k+1)$ dependiendo de su paso por la barrera. La clasificación más natural sería

$$\left\{ \begin{array}{c} \text{caminos} \\ (0,0) \to (n,k) \end{array} \right\} = \bigcup_{j=k}^{n} \left\{ \begin{array}{c} \text{caminos consistentes en ir } (0,0) \to (j,k) \\ \text{y luego ir } (j,k) \to (n+1, k+1) \end{array} \right\},$$

pero, ¡atención!, no se trataría de una partición, pues hay caminos que están en más de uno de los conjuntos de la derecha. Por ejemplo, un camino que llegue a, digamos, el nodo (j,k), luego baje un paso por la barrera, para luego abandonarla y llegar al nodo final, se contaría, al menos, dos veces en la clasificación anterior. Para resolver esta dificultad, vamos a clasificarlos en función del momento (nodo) en el que se "abandone" la barrera:

$$\left\{ \begin{array}{c} \text{caminos} \\ (0,0) \to (n,k) \end{array} \right\} = \bigcup_{j=k}^{n} \left\{ \begin{array}{c} \text{caminos } (0,0) \to (n+1, k+1) \text{ tales que } (j,k) \\ \text{es el \textit{último} nodo de la barrera por el que pasan} \end{array} \right\}$$

Compruebe el lector que ahora cada camino se cuenta una (y sólo una) vez, lo que nos permite utilizar la regla de la suma. Contamos ahora cuántos caminos contiene cada bloque de la partición. Si el nodo (j,k) es el último de la barrera que se toca, entonces el recorrido, más allá de ese punto, está ya fijado: paso a la izquierda y luego bajada hasta el nodo $(n+1, k+1)$. Combinando todas estas observaciones, concluimos que

$$\boxed{\binom{n+1}{k+1}} = \sum_{j=k}^{n} \binom{j}{k} = \boxed{\sum_{j=0}^{n} \binom{j}{k}}$$

Quizás el lector quiera atreverse a subir un poco más arriba esa barrera diagonal (digamos l pasos, en lugar de 1), y tras la pertinente clasificación, obtener la siguiente fórmula:

$$\binom{n}{k} = \sum_{j=k-l}^{n-l} \binom{j}{k-l}\binom{n-j-1}{l-1}.$$

Compare, lector, esta expresión con la fórmula de Vandermonde del ejemplo anterior, y repare en las diferencias, en particular en que el índice de sumación j aparece en la parte *superior* de los coeficientes binómicos.

Quizás, a la vista de estos dos ejemplos con barreras horizontal y diagonal, eche en falta el lector el análisis de otra barrera diagonal, pero ahora dispuesta paralela al borde de la derecha, y compuesta por nodos de la forma $\binom{r+j}{j}$, con $j \geq 0$ y r fijo. Pero obsérvese que, al sumar unos cuantos de estos términos, por ejemplo $n+1$ de ellos, obtenemos

$$\sum_{j=0}^{n} \binom{r+j}{j} = \sum_{j=0}^{n} \binom{r+j}{r} \overset{[k=r+j]}{=} \sum_{k=r}^{r+n} \binom{k}{r} = \sum_{k=0}^{r+n} \binom{k}{r},$$

un tipo de suma que hemos aprendido a calcular al principio de este mismo ejemplo: el resultado concreto, para los parámetros anteriores, es $\binom{r+n+1}{r+1}$. Compruébelo, lector, y observe que estas barreras paralelas al borde derecho no producen nuevas identidades porque son reflexión especular de las correspondientes paralelas al borde izquierdo; la simetría de los coeficientes binómicos hace el resto. ♣

Viene al caso sugerir al lector de que lleve un registro ordenado de las identidades binómicas que han ido apareciendo, tanto en esta sección, como en la dedicada al binomio de Newton. Son unas cuantas ya. En particular, del último ejemplo escribimos el resultado de sumar coeficientes binómicos cuando variamos el índice superior (diagonal fija), y lo comparamos con la habitual suma en la que se varía el índice inferior (piso fijo): para $k \leq n$,

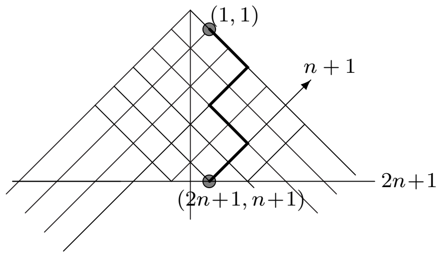
$$\sum_{j=0}^{n} \binom{k}{j} = 2^k \qquad \text{y} \qquad \sum_{j=0}^{n} \binom{j}{k} = \binom{n+1}{k+1}.$$

Note, lector, que la primera suma se extiende hasta $j = k$, y que la segunda arranca en $j = k$.

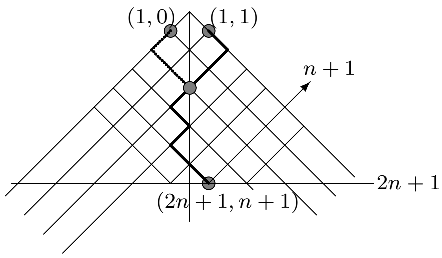
Ejemplo 5.1.4 *Una fórmula para los números de Catalan C_n.*

De entre las diversas aplicaciones combinatorias de los números de Catalan C_n, de las que vimos varias en la sección 2.3, recuperamos, para el análisis que nos conducirá a la fórmula de los C_n, la del ejemplo 2.3.2. Recuerde, lector: C_n cuenta el número de listas $(x_1, \dots, x_{2n})$ formadas por n símbolos $+1$ y n símbolos -1, de manera que las sumas parciales, esto es, x_1, $x_1 + x_2$, $x_1 + x_2 + x_3$, etc, sean siempre no negativas.

Cada una de estas listas se corresponde con un camino en la red protagonista de esta sección, que como resulta conveniente haremos arrancar en el punto de coordenadas $(1, 1)$, y donde interpretamos por ejemplo cada $+1$ como un paso a la izquierda (recordemos, izquierda para el caminante), y cada -1, como un paso a la derecha.



Como la lista tiene tantos símbolos $+1$ (pasos a la izquierda) como símbolos -1 (derecha), el tal camino acaba en el punto de coordenadas $(2n+1, n+1)$. Además, como las sumas parciales son siempre no negativas, tendremos que en cada instante, el caminante habrá dado tantos (o más) pasos a la izquierda como a la derecha. En resumen, *nunca se puede tocar el eje vertical*, y las listas de interés se corresponden con caminos como el que exhibimos en la figura.



El número total de caminos posibles desde $(1,1)$ hasta $(2n+1, n+1)$ es $\binom{2n}{n}$. Pero sólo nos incumbe contar aquéllos que no tocan la línea vertical. Llega aquí la idea brillante[37]: un argumento de reflexión. Vamos a evaluar el tamaño del complementario, los caminos de $(1,1)$ a $(2n+1, n+1)$ que *sí* tocan el eje vertical. Imagine el lector uno de ellos, como el que aparece en el dibujo de la derecha, y localice la *primera vez* que alcanza el eje vertical. Refleje ahora este primer tramo, y compruebe que a cada uno de estos caminos (de $(1,1)$ a $(2n+1, n+1)$ que tocan el eje) le corresponde uno desde $(1,0)$ a $(2n+1, n+1)$.

Pero al revés también, pues un camino de $(1,0)$ a $(2n+1, n+1)$ está obligado a cruzar el eje vertical. Compruebe, lector, que la coletilla "la primera vez que toca el eje" es justamente la condición que permite establecer una biyección entre los dos tipos de caminos. Ahora, de $(1,0)$ a $(2n+1, n+1)$, con el argumento de traslación habitual, sabemos que hay tantos como de $(0,0)$ a $(2n, n+1)$; esto es, $\binom{2n}{n+1}$ caminos. Así que, finalmente, deducimos que

$$C_n = \binom{2n}{n} - \binom{2n}{n+1} \qquad \text{para cada } n \geq 0.$$

Dejamos que el lector gestione algebraicamente esta expresión, jugando con los factoriales, y se convenza de que se puede reescribir como sigue:

$$\boxed{C_n = \frac{1}{n+1}\binom{2n}{n}} \qquad (C_n) = (1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796\ldots)$$

A la derecha hemos escrito los primeros términos de la sucesión. ♣

Como estamos ya tan lanzados (y duchos) en la evaluación de sumas que involucran coeficientes binómicos, nos va a permitir el lector que rematemos este apartado con un cálculo adicional.

Ejemplo 5.1.5 *Unas sumas (sumamente) misteriosas.*

Para k fijo, consideramos la sucesión de coeficientes binómicos de la diagonal k del triángulo de Pascal-Tartaglia, que para este ejemplo vamos a escribir como $\binom{j+k}{k}$, para cada $j \geq 0$.

[37]Idea que, por consiguiente, no es original de los autores.

Fijamos ahora $n \geq k$ y sumamos primeros $n-k+1$ números de la sucesión. Como hemos visto en el ejemplo 5.1.3,

$$\sum_{j=0}^{n-k} \binom{j+k}{k} = \binom{n+1}{k+1}.$$

En particular, tomando $n = 2k$, tenemos que

$$\sum_{j=0}^{k} \binom{j+k}{k} = \binom{2k+1}{k+1}, \qquad \text{o, en otros términos,} \qquad \sum_{j=0}^{k} \frac{\binom{j+k}{k}}{\binom{2k+1}{k+1}} = 1.$$

Observe, lector: $k+1$ números positivos que suman 1 (probabilidades, diremos más adelante), que hemos obtenido "ponderando" la sucesión inicial $\binom{j+k}{k}$ con el factor fijo $\binom{2k+1}{k+1}$.

Vamos ahora a ponderar de otra manera, tomando $\binom{j+k}{k}/2^{j+k}$. Nótese que ahora la ponderación de cada coeficiente binómico depende del contador j. El resultado que vamos a ver es que

$$(\star) \qquad \sum_{j=0}^{k} \binom{j+k}{k} \frac{1}{2^{j+k}} = 1$$

(de nuevo $k+1$ números positivos que suman 1). Y como remate sorpresivo, resulta que

$$(\star\star) \qquad \sum_{j=0}^{\infty} \binom{j+k}{k} \frac{1}{2^{j+k}} = 2$$

(nótese que ahora sumamos hasta infinito), lo que nos dice que al sumar hasta k tenemos 1, y que el resto de la serie, hasta ∞, acumula otro 1. Sí, lector, las sumas de $(\star)$ y $(\star\star)$ valen 1 y 2, respectivamente, *para cualquier valor de* k. Asómbrese.

La manera más directa de probar $(\star\star)$ utiliza que, para cualquier $k \geq 0$,

$$(\dagger) \qquad \frac{x^k}{(1-x)^{k+1}} = \sum_{j=0}^{\infty} \binom{j+k}{k} x^{j+k}, \qquad \text{para } |x| < 1,$$

una serie de Taylor que obtuvimos en el ejemplo 4.3.6. Tome ahora, lector, $x = 1/2$ a ambos lados de la expresión, y tendrá $(\star\star)$.

Para la identidad $(\star)$ vamos a dar un argumento de corte probabilista. Sí, ya sabemos que los capítulos dedicados a la probabilidad vienen más tarde, pero como ya hemos hecho en alguna ocasión, apelamos a los conocimientos previos del lector. Si no es el caso, mírese el capítulo 15.

El argumento analiza un torneo entre dos equipos, A y B, que juegan partidos entre ellos, de manera que gana el primero que sume $k+1$ victorias. Cada jugador gana un partido con probabilidad $p = 1/2$. Observe el lector que como mínimo el torneo ha de constar de $k+1$ partidos, y que el torneo más largo no puede superar los $2k+1$ partidos.

Para j tal que $0 \leq j \leq k$, veamos cuán probable es que A gane el torneo (justamente) en el partido $k+1+j$. Si ése es el caso, el partido $k+1+j$ ha de ser para A, y en los $k+j$

anteriores, A habrá ganado k, mientras que B se habrá apuntado j. Esto nos dice que la probabilidad de que gane A en el partido $k+1+j$ es

$$\binom{j+k}{k}\frac{1}{2^{k+1+j}}$$

(hemos usado que en cada partido se decide el ganador con probabilidad $1/2$); y, sumando en todos los posibles valores de j, la probabilidad de que A gane el torneo es

$$\frac{1}{2}\sum_{j=0}^{k}\binom{j+k}{k}\frac{1}{2^{j+k}}.$$

Por simetría, el número anterior es también la probabilidad de que B gane el torneo. Finalmente, como o bien gana A, o bien gana B, la suma de esas dos probabilidades ha de valer 1, lo que nos da ($\star$).

Repare, lector, en la sucesión de pequeños milagros que han ido ocurriendo en el argumento por usar $1/2$. Por ejemplo, si en la serie de (†) hubiéramos tomado $x = p$, con p un cierto número entre 0 y 1, pero distinto de $1/2$, el resultado dependería de k. ♣

5.1.5. Coeficientes binómicos y principio de inclusión y exclusión

El místico principio de inclusión/exclusión, que recogimos en su momento como lema 2.3.2, permite calcular el tamaño de la unión de una colección finita de conjuntos $\{A_1, A_2 \dots, A_n\}$ mediante la expresión

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{j=1}^{n}(-1)^{j+1}\alpha_j,$$

donde cada α_j es la suma de los tamaños de todas las posibles intersecciones de j conjuntos escogidos de entre los $A_1, \dots, A_n$:

$$\begin{aligned}
\alpha_1 &= |A_1| + |A_2| + \cdots + |A_n| \\
\alpha_2 &= |A_1 \cap A_2| + |A_1 \cap A_3| + \cdots + |A_{n-1} \cap A_n| \\
&\vdots \\
\alpha_n &= |A_1 \cap A_2 \cap \cdots \cap A_n|.
\end{aligned}$$

El número de sumandos que aparece en cada α_j es un coeficiente binómico. Por ejemplo, α_1 consta de $\binom{n}{1}$ sumandos, tantas cono maneras de escoger 1 elemento de entre n hay; α_2 consta de $\binom{n}{2}$ sumandos (formas de escoger 2 de entre n), etc. En general, α_j tendrá $\binom{n}{j}$ términos.

Conocer este dato es muy útil porque, como veremos más adelante (véanse los ejemplos del apartado C), en ocasiones *todas* las intersecciones de j conjuntos, para cada j, serán del mismo tamaño; esta simetría permite simplificar algunas fórmulas. Además, da lugar a la siguiente demostración (alternativa a la prueba por inducción que vimos en el lema 2.3.2) del principio de inclusión/exclusión con un argumento de doble recuento.

A. Demostración del principio de inclusión/exclusión

Antes de arrancar la prueba, recordamos, del lema 5.1.8, que la suma (completa y alternada en signo) de coeficientes binómicos vale 0:

$$\sum_{j=0}^{k}(-1)^j\binom{k}{j}=0\,,\qquad \text{y por tanto que}\qquad (\star)\quad 1=\sum_{j=1}^{k}(-1)^{j+1}\binom{k}{j},$$

tras pasar todos los términos, menos el primero, a la derecha de la igualdad.

Vamos con la demostración en sí. Llamamos $A=\cup_{j=1}^{n}A_j$ y denotamos los elementos de A por $\omega_1,\omega_2,\dots$ Construimos una tabla con columnas etiquetadas por los elementos ω_j de A. Etiquetando las filas aparecen primero los conjuntos $A_1,\dots,A_n$, luego las intersecciones dos a dos, $A_1\cap A_2$, $A_1\cap A_3$, $\dots A_{n-1}\cap A_n$; después, las tres a tres, las cuatro a cuatro…y así, hasta la intersección de todos los A_j. Hay, pues, 2^n-1 filas.

	ω_1	ω_2	ω_3	ω_4	$\cdots$
A_1	1	1	1	1	$\cdots$
A_2	1	1	1	0	$\cdots$
A_3	1	1	0	1	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
A_n	1	0	0	1	$\cdots$
$A_1\cap A_2$	-1	0	-1	0	$\cdots$
$A_1\cap A_3$	-1	-1	0	-1	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$A_{n-1}\cap A_n$	-1	0	0	-1	$\cdots$
$A_1\cap A_2\cap A_3$	1	1	0	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$A_{n-2}\cap A_{n-1}\cap A_n$	1	0	0	1	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$A_1\cap\cdots\cap A_n$	$(-1)^{n+1}$	0	0	0	$\cdots$

Los registros de la matriz van a ser -1, 0 y 1. En la columna etiquetada por un cierto elemento ω_j, pondremos un 0 si ω_j no pertenece al conjunto que etiquete la fila. En caso contrario, distinguiremos si se trata de una intersección de un número *impar* de conjuntos (en cuyo caso pondremos un $+1$) o de una intersección de un número *par* (escribiremos un -1). Por ejemplo, habría un 1 en las casillas correspondientes si ω_j estuviera en A_3, en $A_2\cap A_5\cap A_6$, o en $A_2\cap A_3\cap A_7\cap A_8\cap A_9$, mientras que sería un -1 si perteneciera, por ejemplo, a $A_1\cap A_3$, a $A_2\cap A_3\cap A_5\cap A_n$, etc. La tabla tendría un aspecto como el de la figura. Sumamos ahora las entradas de la matriz, primero por filas y luego por columnas.

La suma por filas nos da la suma alternada en signos del principio de inclusión/exclusión, pues justamente hemos diseñado la matriz para eso. Veámoslo. La fila etiquetada con A_1 contiene $|A_1|$ unos, la de A_2, $|A_2|$ unos, etc. Cada fila etiquetada con $A_i\cap A_j$ contiene $|A_i\cap A_j|$ signos "-1". Y así sucesivamente. En total, la suma por filas nos da

$$(\star)\qquad \sum_{i=1}^{n}|A_i|-\sum_{1\le i<j\le n}|A_i\cap A_j|+\cdots\,.$$

La suma por columnas requiere un análisis más cuidadoso. Fijamos un $\omega\in A$. Digamos que ese elemento ω está en *exactamente* k de los A_j, donde k está en el rango $1\le k\le n$. Entonces, en su columna habrá exactamente k unos en las filas etiquetadas por los conjuntos $A_1,\dots,A_n$. Pero entonces ω pertenecerá a exactamente $\binom{k}{2}$ intersecciones dos a dos, justamente las que involucren a los k conjuntos a los que pertenece ω. Así que habrá $\binom{k}{2}$

signos "-1" en las filas de las intersecciones dos a dos. De igual manera, estará en $\binom{k}{3}$ de las intersecciones tres a tres, etc. El último signo no nulo estará en las filas de las intersecciones k a k: será un $(-1)^{k+1}$, justo en la fila etiquetada por la intersección de todos los A_j en los que esté ω; más allá, todo ceros. En total, la suma de la columna de ω valdrá

$$\binom{k}{1} - \binom{k}{2} + \binom{k}{3} + \cdots + (-1)^{k+1}\binom{k}{k} = \sum_{j=1}^{k}(-1)^{j+1}\binom{k}{j} = 1,$$

usando ($\star$). Así que la suma de las entradas de la columna del elemento ω vale exactamente 1. Como el argumento es válido para cualquier ω (aunque en cada caso se tendrá un valor de k distinto), concluimos que cada columna de la matriz suma 1. Y como hay $|A| = |A_1 \cup \cdots \cup A_n|$ columnas, resulta que la suma ($\star$) vale $|A|$. ■

Consulte el lector un par de generalizaciones del principio del inclusión/exclusión en el ejemplo 5.1.1 y en el ejercicio 5.1.23.

B. Las desigualdades de Bonferroni

La fórmula del principio de inclusión/exclusión,

$$\Big|\bigcup_{j=1}^{n} A_j\Big| = \underbrace{\sum_{i=1}^{n}|A_i|}_{\alpha_1} - \underbrace{\sum_{1\le i<j\le n}|A_i\cap A_j|}_{\alpha_2} + \cdots + (-1)^{n+1}\underbrace{|A_1\cap\cdots\cap A_n|}_{\alpha_n} = \sum_{j=1}^{n}(-1)^{j+1}\alpha_j$$

exige calcular los valores de los sucesivos α_j, una tarea en general muy aparatosa, pues hay una cantidad enorme de sumandos, $2^n - 1$, que calcular.

Nos gustaría saber el error que se comete si nos limitamos a sumar (con el signo correspondiente) los primeros α_j. Para empezar, es casi directo comprobar por inducción que

$$|A_1 \cup \cdots \cup A_n| \le \sum_{j=1}^{n}|A_j| = \alpha_1\,,$$

apoyándose en el caso $n = 2$, en el que claramente $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2| \le |A_1| + |A_2|$. De manera que, si nos quedamos simplemente con el primer término (las sumas de los tamaños individuales), acotamos la suma completa por arriba. Esta desigualdad es a veces conocida como *desigualdad de Boole*.

También por inducción, pero con algo más de trabajo, se puede probar que $|A_1\cup\cdots\cup A_n| \ge \alpha_1 - \alpha_2$. Dejamos al lector los detalles como ejercicio 5.1.22. Ahora, al quedarnos con los dos primeros términos, se acota la suma completa por debajo.

Esta situación, en la que se va acotando (una vez por exceso, la siguiente por defecto) la suma completa al ir incorporando (con el signo adecuado) los sucesivos α_j, es general. Son las llamadas **desigualdades de Bonferroni**[38], que exhibiremos en el teorema 5.1.21, y que vamos a deducir del siguiente resultado, que afirma que el error cometido al truncar la suma del principio de inclusión/exclusión está acotado por el tamaño del primer término despreciado.

[38]El matemático italiano Carlo Emilio Bonferroni (1892-1960) presentó sus desigualdades en un contexto probabilístico, y con vistas a su aplicación a la ciencia actuarial, a la que dedicó parte de su vida.

Proposición 5.1.20 *Para cada $2 \leq t \leq n$,*

$$\Big|\sum_{j=1}^{n}(-1)^{j+1}\alpha_j - \sum_{j=1}^{t-1}(-1)^{j+1}\alpha_j\Big| = \Big|\sum_{j=t}^{n}(-1)^{j+1}\alpha_j\Big| \leq \alpha_t\,.$$

Es decir, para cada $2 \leq t \leq n$,

$$-\alpha_t + \sum_{j=1}^{t-1}(-1)^{j+1}\alpha_j = \Big|\bigcup_{j=1}^{n} A_j\Big| \leq \alpha_t + \sum_{j=1}^{t-1}(-1)^{j+1}\alpha_j$$

El lector atento advertirá cierta similitud entre este resultado y el teorema 4.2.12 de Leibniz sobre series alternadas. Y de hecho se podría demostrar como aquél si ocurriera que $\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \cdots$; es decir, si la sucesión (α_j) fuera *decreciente*. Pero no es el caso, en general.

Miremos, por ejemplo, α_1 y α_2. En el primero se suman términos $|A_j|$, y en el segundo tamaños de intersecciones $|A_i \cap A_j|$, que son, en principio, más pequeños; pero de los primeros sumamos n, mientras que de los segundos sumamos más, $\binom{n}{2}$. De manera que, en ese equilibrio entre tamaño de los sumandos y número de sumandos, podría ocurrir que α_1 fuera menor que α_2. Como ejemplo sencillo, consideremos los seis conjuntos siguientes: $A_1 = \{1,2\}$, $A_2 = \{1,3\}$, $A_3 = \{1,4\}$, $A_4 = \{1,5\}$, $A_5 = \{1,6\}$ y $A_6 = \{1,7\}$. Obsérvese que en este caso $\alpha_1 = \binom{6}{1} \cdot 2 = 12$, mientras que $\alpha_2 = \binom{6}{2} \cdot 1 = 15$.

Bien, habrá que gestionar esta dificultad para conseguir demostrar la proposición 5.1.20, y lo haremos estimando cuidadosamente el valor de sumas alternadas (pero incompletas) de coeficientes binómicos. Pero antes, veamos cómo podemos deducir las desigualdades de Bonferroni a partir de la proposición 5.1.20. Partimos de la ya conocida

$$|A_1 \cup \cdots \cup A_n| \leq \alpha_1,$$

es decir, $|A_1 \cup \cdots \cup A_n| - \alpha_1 \leq 0$. Ahora, del caso $t = 2$ de la proposición 5.1.20 deducimos que

$$\big||A_1\cup\cdots\cup A_n| - \alpha_1\big| \leq \alpha_2 \quad\Longrightarrow\quad \alpha_1 - |A_1\cup\cdots\cup A_n| \leq \alpha_2 \quad\Longrightarrow \boxed{|A_1 \cup \cdots \cup A_n| \geq \alpha_1 - \alpha_2}$$

que es la segunda desigualdad de la cadena. Nótese cómo nos libramos del valor absoluto en la primera implicación, usando el caso anterior. Ahora el caso $t = 3$ nos da la tercera desigualdad:

$$\begin{aligned} \big||A_1 \cup \cdots \cup A_n| - (\alpha_1 - \alpha_2)\big| \leq \alpha_3 \quad &\Longrightarrow \quad |A_1 \cup \cdots \cup A_n| - \alpha_1 + \alpha_2 \leq \alpha_3 \\ &\Longrightarrow \quad \boxed{|A_1 \cup \cdots \cup A_n| \leq \alpha_1 - \alpha_2 + \alpha_3}, \end{aligned}$$

donde quitamos de nuevo el valor absoluto usando el caso anterior. Y así, sucesivamente:

Teorema 5.1.21 (Desigualdades de Bonferroni) *Dada una colección $A_1, \ldots, A_n$ de conjuntos, se tiene que*

$$|A_1 \cup \cdots \cup A_n| \leq \sum_{j=1}^{k}(-1)^{j+1}\alpha_j \quad \text{si } k \text{ impar}; \qquad |A_1 \cup \cdots \cup A_n| \geq \sum_{j=1}^{k}(-1)^{j+1}\alpha_j \quad \text{si } k \text{ par.}$$

Con cada α_j que se añade (con su signo), nos acercamos cada vez más al valor real del tamaño de la unión; y estas aproximaciones se van alternando: una por exceso, la siguiente por defecto. Sólo al final, al sumar todos los α_j con sus signos correspondientes, recuperamos el valor exacto de $|A_1 \cup \cdots \cup A_n|$. Recordemos, además, que podemos acotar el error (en valor absoluto) que se comete en cada una de ellas por el valor de α_{k+1}, el primer término despreciado.

Queda pendiente la demostración de la proposición 5.1.20, en la que usaremos dos resultados, que tienen interés por sí mismos. El primero recoge una identidad sobre sumas (quizás) *incompletas* y alternadas en signo de coeficientes binómicos:

Lema 5.1.22 *Dado $k \geq 1$,*

$$\sum_{j=0}^{t} (-1)^j \binom{k}{j} = (-1)^t \binom{k-1}{t} \qquad \text{para todo } t \geq 0.$$

El caso $t = k$ de este resultado es el lema 5.1.8, que constataba que la suma *completa* y alternada en signo de coeficientes binómicos vale 0. Observe, lector, que si $t > k$, los dos miembros de esta identidad anterior valen 0.

Demostración. La identidad anterior, que depende de dos parámetros, k y t, se prueba con un argumento de inducción doble (apartado 1.2.2) en el que se utiliza la regla de recurrencia de los coeficientes binómicos (dos veces) como sigue:

$$\begin{aligned}\sum_{j=0}^{t} (-1)^j \binom{k}{j} &= \sum_{j=0}^{t} (-1)^j \Big[\binom{k-1}{j} + \binom{k-1}{j-1}\Big] = \sum_{j=0}^{t} (-1)^j \binom{k-1}{j} + \sum_{m=0}^{t-1} (-1)^{m+1} \binom{k-1}{m} \\ &= (-1)^t \binom{k-2}{t} - (-1)^{t-1} \binom{k-2}{t-1} = (-1)^t \Big[\binom{k-2}{t} + \binom{k-2}{t-1}\Big] = (-1)^t \binom{k-1}{t}.\end{aligned}$$

En la tercera igualdad se ha usado la hipótesis de inducción, y por el camino hemos cambiado de índice en una suma. Revíselo el lector y complete los detalles del argumento de inducción, probando los primeros casos, etc. ■

Sean unos conjuntos $A_1, \dots, A_n$, y llamemos $A = \cup_{j=1}^{n} A_j$. Tenemos, por un lado, la sucesión de números $(\alpha_1, \alpha_2, \dots, \alpha_n)$ habituales del principio de inclusión/exclusión: cada α_j suma los tamaños de las intersecciones j a j de los conjuntos $A_1, \dots, A_n$. Añadimos, pues así resulta conveniente, $\alpha_0 = |A|$.

Consideramos una sucesión adicional de números, $(\beta_0, \beta_1, \dots, \beta_n)$, donde β_k cuenta el número de elementos de A que están en *exactamente* k de los conjuntos $A_1, \dots, A_n$. Estas cantidades aparecieron implícitamente en la demostración del principio de inclusión/exclusión de esta misma sección. Obsérvese que $\beta_0 = 0$. El siguiente resultado incluye una fórmula para calcular cada α_j en términos de los β_k.

Lema 5.1.23 *Para cada $j = 0, \dots, n$,*

$$\alpha_j = \sum_{k=j}^{n} \binom{k}{j} \beta_k.$$

Demostración. Para empezar, nótese que el caso $j=0$,

$$\alpha_0 = |A| = \sum_{k=0}^{n} \beta_k,$$

sale directamente de aplicar la regla de la suma.

Retroceda ahora, querido lector, un par de páginas, a la demostración del principio de inclusión/exclusión. Sitúese allí, frente a la lustrosa matriz con las columnas etiquetadas por los elementos de A y las filas etiquetadas con todas las posibles intersecciones de los A_i. ¿Está ya? Bien, ahora rellene esa matriz como allí se indicaba, pero sin hacer distingos entre los $+1$ y -1. Es decir, use únicamente ceros (si el elemento ω que etiqueta la columna no está en la intersección que rotula la fila) y unos (en caso contrario).

Fíjese ahora en el bloque de las filas que contienen a todas las intersecciones j a j. El número de unos que aparece en esas filas es justamente α_j. Miramos ahora las columnas. Digamos que un elemento ω está en exactamente k de los $A_1, \dots, A_n$. Entonces, en su columna, y en el bloque de filas seleccionado, aparecerán $\binom{k}{j}$ unos, pues ése es el número de posibles intersecciones j a j de entre los k conjuntos a las que pertenece ω. Por supuesto, ese número de unos podría ser 0, si $k<j$, pero de eso ya da cuenta el coeficiente binómico.

Repitiendo el argumento para cada ω y recordando la definición de los β_k concluimos que

$$\alpha_j = \sum_{k=j}^{n} \binom{k}{j} \beta_k.$$

para cada $j \le n$ (incluso para $j=0$, como se argumentó antes y por separado). ■.

Demostración de la proposición 5.1.20. Seguimos con los conjuntos $A_1, \dots, A_n$, con $A = \cup_{j=1}^{n} A_j$. Las sucesiones $(\alpha_0, \alpha_1, \dots, \alpha_n)$ y $(\beta_0, \beta_1, \dots, \beta_n)$ son las del lema anterior.

Consideramos las sumas

$$S_t = \sum_{j=1}^{t} (-1)^{j+1} \alpha_j \,, \quad \text{para cada } t = 1, \dots, n.$$

El principio de inclusión/exlusión nos dice que $S_n = |A|$. Queremos probar que $|S_n - S_t| \le \alpha_{t+1}$ para cada $t = 1, \dots, n-1$.

Volvemos al argumento de la prueba del principio de inclusión/exclusión. Ahora mantenemos la matriz como allí estaba, con entradas 0, $+1$ y -1. Obsérvese que $S_n - S_t$ es la suma de las entradas de las filas de esa matriz desde las intersecciones $t+1$ a $t+1$ en adelante.

Consideramos un $\omega \in A$ que esté en, digamos, k de los conjuntos $A_1, \dots, A_n$. Si $k \le t$, entonces en ese bloque de filas sólo aparecerán ceros en su columna. Si $k \ge t+1$, ω estará en $\binom{k}{t+1}$ de las intersecciones $t+1$ a $t+1$, en $\binom{k}{t+2}$ de las $t+2$ a $t+2$, etc., y finalmente en $\binom{k}{k}$ de las intersecciones k a k. De manera que, incluyendo los signos, la columna de ω sumará

$$\sum_{j=t+1}^{k} (-1)^{j+1} \binom{k}{j} = \sum_{j=0}^{k} (-1)^{j+1} \binom{k}{j} - \sum_{j=0}^{t} (-1)^{j+1} \binom{k}{j} = (-1)^t \binom{k-1}{t},$$

donde hemos usado el lema 5.1.22 para evaluar las dos sumas de la expresión intermedia: la primera, como es "completa", vale cero, y la segunda vale el número que aparece a la derecha.

Ahora, considerando todos los $\omega \in A$, concluimos que

$$S_n - S_t = \sum_{k=t+1}^{n} (-1)^t \binom{k-1}{t} \beta_k.$$

Compruebe ahora el lector, por simple manipulación algebraica de la fórmula de factoriales, que si $k \geq t+1$,

$$\binom{k-1}{t} \leq \binom{k}{t+1}$$

(si $k \leq t$, ambos coeficientes binómicos son nulos), y concluya, usando el lema 5.1.23, que

$$|S_n - S_t| = \Big| \sum_{k=t+1}^{n} (-1)^t \binom{k-1}{t} \beta_k \Big| \leq \sum_{k=t+1}^{n} \binom{k-1}{t} \beta_k \leq \sum_{k=t+1}^{n} \binom{k}{t+1} \beta_k = \alpha_{t+1}\,. \qquad \blacksquare$$

C. Algunas aplicaciones

Como ya hemos mencionado, en ciertas ocasiones el principio de inclusión/exclusión da lugar a fórmulas (relativamente) manejables. Los dos siguientes ejemplos recogen dos aplicaciones combinatorias especialmente interesantes, como son contar el número de aplicaciones sobreyectivas y calcular el número de los llamados "desbarajustes", un tipo especial de permutaciones. El lector podrá encontrar otra notable aplicación en la sección 6.5.2, en la fórmula para la llamada *función ϕ de Euler*, una de las funciones estrella de la aritmética.

EJEMPLO 5.1.1 *El recuento del número de* ***aplicaciones sobreyectivas*** *entre los conjuntos $\mathcal{X} = \{1,\dots,n\}$ e $\mathcal{Y} = \{1,\dots,k\}$.*

Recuérdese, del ejemplo 2.2.5, que podemos identificar aplicaciones de $\mathcal{X} \to \mathcal{Y}$ con n-listas formadas con los símbolos $\{1,\dots,k\}$. Informalmente, una aplicación es sobreyectiva si no se "salta" ningún elemento de $\mathcal{Y}$; o, con más precisión, si para todo $y \in \mathcal{Y}$ existe al menos un $x \in \mathcal{X}$ tal que $f(x) = y$. En otros términos, es una n-lista en la que aparecen todos los k símbolos. En el argumento que sigue alternaremos estos dos lenguajes de aplicaciones y listas.

Para empezar, el número de aplicaciones sobreyectivas entre $\mathcal{X} = \{1,\dots,n\}$ e $\mathcal{Y} = \{1,\dots,k\}$ es 0 si $n < k$, pues en este caso resulta imposible "cubrir" el conjunto imagen. Pongámonos, entonces, en el caso en el que $n \geq k$.

Vamos a pasar al complementario: las aplicaciones que no sean sobreyectivas, o bien se saltarán el elemento 1, o bien el 2, etc., así hasta el k. Así que, si definimos los conjuntos

$$A_j = \{\text{aplicaciones } \mathcal{X} \to \mathcal{Y} \text{ que se saltan el elemento } j\}\,, \qquad \text{para cada } j = 1,\dots,k\,,$$

el número de aplicaciones sobreyectivas resulta ser $k^n - |A_1 \cup A_2 \cup \cdots \cup A_k|$.

Calculemos, en primer lugar, el tamaño de cada uno de los A_j. Una aplicación que esté en A_j no toma el valor j como imagen. En términos de listas, es una n-lista en la que podemos utilizar cualquier elemento de $\mathcal{Y}$ menos el símbolo j. Esto nos dice que

$$|A_j| = (k-1)^n\,, \quad \text{para cada } j = 1,\dots,k.$$

Si una aplicación está en $A_i \cap A_j$, entonces no tomará como imagen ni a i ni a j. Es decir, la lista correspondiente se formará con cualesquiera de los otros $k-2$ símbolos. Por tanto,

$$|A_i \cap A_j| = (k-2)^n\,, \quad \text{para cada } i \neq j.$$

El mismo argumento sirve para las intersecciones tres a tres, cuatro a cuatro, etc. Como el número de sumandos en cada suma de la expresión del principio de inclusión/exclusión viene dado por un coeficiente binómico, concluimos que, si $|\mathcal{X}| = n$ e $|\mathcal{Y}| = k$, con $n \geq k$,

$$\begin{aligned}&\#\{\text{aplicaciones sobreyectivas } \mathcal{X} \to \mathcal{Y}\} = k^n - |A_1 \cup A_2 \cup \cdots \cup A_k| \\ &= k^n - \left[\binom{k}{1}(k-1)^n - \binom{k}{2}(k-2)^n + \cdots \pm \binom{k}{k}(k-k)^n\right] = \sum_{j=0}^{k-1}(-1)^j \binom{k}{j}(k-j)^n,\end{aligned}$$

expresión que podemos reescribir, tras el pertinente arreglo cosmético, como

$$\#\{\text{aplicaciones sobreyectivas } \{1,\dots,n\} \to \{1,\dots,k\}\} = \sum_{j=0}^{k}\binom{k}{j}(-1)^{k-j}\,j^n$$

para cada $n \geq k$. Viene al caso recordar ahora el lema 5.1.9, que si el lector tiene a bien revisar, nos dice que la suma de la derecha *vale* 0 cuando $n < k$. Lo que cierra felizmente el círculo, pues resulta que la suma de la derecha cuenta el número de aplicaciones sobreyectivas entre $\{1,\dots,n\}$ y $\{1,\dots,k\}$ incluso cuando $n < k$, dando 0 en ese caso.

Aún así, la fórmula de arriba es tremenda –lo reconocemos, lector–, y de aparente poca utilidad como método de cálculo. Y eso que nos habían prometido –nos recordará, puntilloso, el lector– que estábamos en uno de esos casos de amable aplicación del principio de inclusión/exclusión, en el que todas las intersecciones del mismo tipo tienen el mismo tamaño. En la sección 5.3.1 retomaremos esta cuestión para obtener un método de cálculo alternativo del número de aplicaciones sobreyectivas, apoyándonos en una de las más nobles familias de números de la combinatoria: los números de Stirling (de segunda especie). El lector impaciente puede ya saltar hasta allí, si le place. ♣

EJEMPLO 5.1.2 ***Desbarajustes.***

Las permutaciones de $\{1,\dots,n\}$ son, recordemos, las n-listas sin repetición formadas con ese conjunto de símbolos. Es decir, todas las $n!$ posibles reordenaciones de los símbolos $\{1,\dots,n\}$. O, alternativamente, las aplicaciones biyectivas de $\{1,\dots,n\}$ en sí mismo. En la sección 5.2 nos ocuparemos de estos objetos, de su estructura interna y de diversos problemas combinatorios que aparecen al estudiarlos.

Nos centramos aquí en un tipo muy especial de permutaciones, a las que llamaremos desbarajustes: son las n-listas en las que ningún símbolo está en su posición "natural": es decir, el 1 no está en la posición primera, el 2 no está en la segunda, etc. En el lenguaje de las aplicaciones, son las biyecciones que no fijan elemento alguno.

Nos interesa hallar una fórmula explícita para el número de desbarajustes, al que nos referiremos como D_n. O mejor, para $D_n/n!$, que representa la proporción que los desbarajustes ocupan entre todas las permutaciones. Esto es, la “probabilidad” de que si escogemos una permutación al azar, ésta sea un desbarajuste.

En el planteamiento clásico del problema[39], se han escrito n cartas y preparado los n sobres con las direcciones correspondientes, cada uno al lado de su carta. Pero alguien los ha descolocado, de manera que no queda más remedio que introducir las cartas en los sobres al azar. Hecho esto, ¿cuál será la probabilidad de no acertar ninguna? Esta probabilidad, obsérvese, es justamente $D_n/n!$.

Antes de empezar a analizar el problema, quizás el lector quiera meditar un momento sobre la cuestión y apelar a su intuición para, al menos, arriesgar una respuesta aproximada: ¿una probabilidad cercana a 0, cercana a 1? ¿Qué ocurre cuando el número de cartas y sobres es muy grande? Parece difícil que ninguna carta caiga en su sobre... ¿o no?

Argumentamos, como antes, pasando al complementario y usando el principio de inclusión/exclusión. Una permutación no es un desbarajuste si al menos uno de los símbolos está en su posición natural. Así que consideramos los siguientes conjuntos:

$$A_j = \{\text{permutaciones de } \{1,\dots,n\} \text{ con el símbolo } j \text{ en la posición } j\}, \quad \text{para } j = 1,\dots,n$$

(en el lenguaje de las aplicaciones, cada A_j contendría a las biyecciones que fijan el correspondiente elemento j). Como hay $n!$ permutaciones en total,

$$D_n = n! - |A_1 \cup \cdots \cup A_n|\,.$$

Obsérvese que $|A_j| = (n-1)!$ para cada j, porque A_j está formado por las listas con el símbolo j en la posición j, y para contarlas bastará ordenar (permutar) los $n-1$ símbolos restantes. Por otro lado, todas las intersecciones dos a dos tienen tamaño $(n-2)!$, pues en una intersección dos a dos están las listas con dos símbolos ya colocados en sus correspondientes posiciones (sólo hay que permutar los restantes). Y así con el resto de los casos.

Aplicando el principio de inclusión/exclusión obtenemos que

$$D_n = n! - \left[\binom{n}{1}(n-1)! - \binom{n}{2}(n-2)! + \cdots + (-1)^n\binom{n}{n}(n-n)!\right] = \sum_{j=0}^{n}\binom{n}{j}(n-j)!(-1)^j.$$

Esto es, simplificando un poco, para cada $n \ge 1$,

$$\boxed{D_n = n!\sum_{j=0}^{n}\frac{(-1)^j}{j!}}\,, \qquad \text{o bien} \qquad \frac{D_n}{n!} = \sum_{j=0}^{n}\frac{(-1)^j}{j!}.$$

Esta fórmula daría además que $D_0 = 1$; y aunque D_0 no tiene sentido combinatorio, se suele aceptar, por conveniencia, que toma ese valor 1. Se trata, compare el lector, de una expresión bastante más sencilla que la del caso de las aplicaciones sobreyectivas. Pero

[39]Una versión alternativa habla de hombres que dejan sus sombreros a la entrada de una fiesta para luego recogerlos al azar, y en ella se trata de estimar cuán probable es que ninguno reciba su sombrero. Una versión viejuna, en todo caso. Aunque pensándolo bien, lo de las cartas y los sobres tampoco es la gran modernidad.

más aún, la cantidad $\sum_{j=0}^{n}(-1)^j/j!$, que en principio depende de n, es, a todos los efectos numéricos y si n es grande, casi indistinguible[40] de $1/e$, como ya se discutió en la sección 3.2.

Así que la probabilidad $D_n/n!$ de obtener un desbarajuste es prácticamente independiente de n (si n es suficientemente grande). Pero más aún, es una probabilidad "grande": un 36.78 %, mayor que un tercio. Quizás más de lo que hubiéramos apostado al principio.

¿Cómo se explica este (un tanto inesperado) fenómeno? Argumentamos probabilísticamente, apelando a los conocimientos o intuiciones[41] del lector. Llamemos B_j a los complementarios de los A_j en el conjunto de las permutaciones. Es decir, B_j contiene a las permutaciones de $\{1,\dots,n\}$ en las que el símbolo j no está en la posición j. La "proporción" de un A_j era $(n-1)!/n!$, de forma que

$$\text{prob}(B_j) = 1 - \text{prob}(A_j) = 1 - \frac{(n-1)!}{n!} = 1 - \frac{1}{n}\,.$$

La intersección de todos los conjuntos B_j es, precisamente, el conjunto de los desbarajustes. Si los B_j *fueran independientes* entre sí, entonces tendríamos que

$$\text{prob}\Big(\bigcap_{j=1}^{n} B_j\Big) = \prod_{j=1}^{n}\text{prob}(B_j) = \prod_{j=1}^{n}\Big(1-\frac{1}{n}\Big) = \Big(1-\frac{1}{n}\Big)^n \approx \frac{1}{e} \qquad \text{si } n \text{ es grande.}$$

La "independencia", de cuyo significado exacto nos ocuparemos en el capítulo 15, significa, informalmente, que, por ejemplo, el que el símbolo j esté fuera de su posición no "influye" en que el símbolo k esté fuera de su posición. No es este el caso si consideramos permutaciones de $\{1,2\}$, porque si el 1 está fuera de su posición entonces el 2 ha de estar necesariamente descolocado. Para tres símbolos, el que el símbolo 1 no esté en su posición influye en que, por ejemplo, el 2 no esté en la segunda posición, aunque algo "menos". Pero cuando n se hace cada vez más grande, esta influencia se va diluyendo. Es decir, los B_j son (aproximadamente) independientes, lo que explica que la probabilidad de obtener un desbarajuste sea $\approx 1/e$. El lector interesado puede encontrar un argumento más riguroso en el ejercicio 5.1.25. ♣

Como complemento al análisis de los desbarajustes del ejemplo anterior, analizamos a continuación permutaciones que fijan un cierto número de elementos.

Ejemplo 5.1.3 *Permutaciones que fijan un cierto número de elementos y número medio de puntos fijos de una permutación.*

Para cada $k = 0,1,\dots,n$, llamamos $D_n(k)$ al número de permutaciones de $\{1,\dots,n\}$ que fijan *exactamente* k elementos. Obsérvese que $D_n(n) = 1$, pues sólo hay una lista con todos los elementos en su posición (una única permutación que fija todos los elementos, la identidad); y que $D_n(0)$ coincide con D_n, el número de desbarajustes de $\{1,\dots,n\}$ del ejemplo anterior.

Para obtener una fórmula para $D_n(k)$, argumentamos como sigue: seleccionamos primero qué k elementos quedan fijos, para luego permutar los restantes formando un desbarajuste,

[40] Por ejemplo, las primeras cifras decimales de $1/e$ son 0.3678794412, mientras que al sumar hasta $j = 10$ obtenemos 0.3678794643. El lector interesado podrá encontrar un argumento preciso en el ejercicio 5.1.24.

[41] Quizás el lector quiera consultar antes el capítulo 15.

lo que garantiza que no queden fijos más de los k elementos elegido en el primer paso. Esto nos dice que

$$(\star) \qquad D_n(k) = \binom{n}{k} D_{n-k} = \frac{n!}{k!} \sum_{j=0}^{n-k} \frac{(-1)^j}{j!},$$

usando la fórmula de los desbarajustes del ejemplo anterior. En la expresión anterior hemos usado la conveniente definición $D_0 = 1$.

La fórmula $(\star)$ permite estimar cuán frecuentes son las permutaciones en función del número de símbolos que fijan. Ya sabemos, del ejemplo anterior, que para n es grande, el número de desbarajustes es $D_n(0) = D_n \approx n!/e$. Ahora también podemos afirmar que[42]

$$D_n(1) \approx \frac{n!}{e}, \qquad D_n(2) \approx \frac{n!}{2! \cdot e}, \qquad D_n(3) \approx \frac{n!}{3! \cdot e}, \qquad D_n(4) \approx \frac{n!}{4! \cdot e} \dots$$

Quizás el lector quiera comprobar la similitud con la situación del ejemplo 2.3.1, que trataba listas (no necesariamente de longitud n) sin repetición con n símbolos.

Así que, si n es grande, aproximadamente un 36.78 % de las permutaciones de $\{1, \dots, n\}$ son desbarajustes, otro 36.78 % fijan un único elemento, un 18.39 % fijan dos elementos, y un 3.06 % fijan tres. Observe, lector, que ya llevamos más del 95 %.

No resulta extraño, pues, que una permutación fije, *en media*, muy pocos elementos. Uno, de hecho. Para comprobarlo, observamos primero que

$$n! = \sum_{k=0}^{n} D_n(k),$$

por simple aplicación de la regla de la suma. Pero además,

$$\sum_{k=1}^{n} k\, D_n(k) = \sum_{j=0}^{n-1} (n-j) \binom{n}{j} D_j = \sum_{j=0}^{n-1} n \binom{n-1}{j} D_j = n \sum_{j=0}^{n-1} \binom{n-1}{j} D_j = n\,(n-1)! = n!$$

En la segunda igualdad hemos usado la identidad $(n-k)\binom{n}{k} = n\binom{n-1}{k}$, cuya comprobación algebraica es inmediata. De manera que, como se anunciaba antes,

$$\frac{1}{n!} \sum_{k=0}^{n} k\, D_n(k) = 1,$$

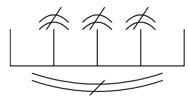
que significa que, *en media*, el número de puntos fijos de una permutación es 1. Note, lector, y participe del asombro: ¡sea cual sea n! ♣

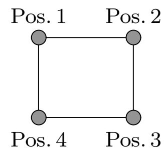
D. Listas con restricciones y principio de inclusión/exclusión

Como ya se ha mencionado en alguna ocasión, y como está comprobando repetidamente el lector, un buen número de cuestiones combinatorias tienen que ver con listas "con restricciones". Habitualmente, estas restricciones impiden contar el número de listas con una simple aplicación de la regla del producto, y hace falta apoyarse en otro tipo de técnicas.

[42] Para el lector (muy) avisado: son, muy aproximadamente, y tras dividir por $n!$, las probabilidades de una variable de Poisson de parámetro 1.

Los dos anteriores son buenos ejemplos. En las aplicaciones sobreyectivas, se trata de contar n-listas con los símbolos $\{1, \dots, k\}$ en las que aparecen *todos* los símbolos posibles. Los desbarajustes son permutaciones (n-listas sin repetición) de $\{1, \dots, n\}$) en las que *ningún* símbolo aparece en su posición natural. En ambos casos tenemos cierto número de condiciones simultáneas (k en el primer caso, n en el segundo), es decir, una intersección de condiciones, que tras pasar al complementario devienen en unión de condiciones (las complementarias de las originales). Ahí es donde entra en escena el principio de inclusión/exclusión.





Otro tipo de restricciones, muy frecuentes, son aquellas en las no se pueden repetir símbolos en ciertas posiciones. Ya nos hemos encontrado con algún caso: en el del ejemplo 2.2.8 queríamos contar las 4-listas con símbolos del conjunto $\{1, \dots, n\}$ en las que en posiciones consecutivas aparecen símbolos distintos y, además, en las que no haya el mismo símbolo en las posiciones primera y cuarta. A la derecha exhibimos un par de representaciones gráficas del problema: en la primera figura aparece la lista con sus prohibiciones; en la segunda, cada posición de la lista etiqueta un *vértice* y cada prohibición se corresponde con un arco o *arista*. Esto es lenguaje de grafos, que desarrollaremos en el capítulo 9. Sólo lo usamos aquí para representar de manera sencilla, sin más aparato teórico. Observe el lector cómo el esquema con vértices y arcos describe la información de manera más clara que el dibujo de la lista con prohibiciones.

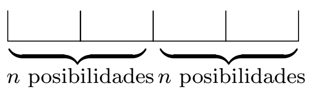
Al comienzo de la sección 2.3 resolvimos esta cuestión separando en dos casos y aplicando la regla de la suma: hay $n(n-1)(n^2-3n+3)$ listas distintas. Pero la estructura del problema, unas cuantas restricciones simultáneas, sugiere abordarlo usando el principio de inclusión/exclusión (tras pasar al complementario).

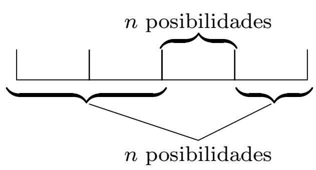
Veamos. Sea A el conjunto de las listas de interés y X el conjunto de las 4-listas con los símbolos $\{1, \dots, n\}$, de las que hay n^4. El complementario de A en X se puede escribir como

$$X \setminus A = \bigcup_{i=1}^4 A_i\,, \quad \text{donde} \quad \begin{cases} A_1 = \{\text{4-listas con } 1^{\text{a}} = 2^{\text{a}}\},\ A_2 = \{\text{4-listas con } 2^{\text{a}} = 3^{\text{a}}\} \\ A_3 = \{\text{4-listas con } 3^{\text{a}} = 4^{\text{a}}\},\ A_4 = \{\text{4-listas con } 4^{\text{a}} = 1^{\text{a}}\} \end{cases}$$

una unión de conjuntos (no disjuntos). Podemos calcular $|A|$, vía el principio de inclusión/exclusión, si evaluamos el tamaño de todas las intersecciones involucradas. Para empezar, $|A_j| = n^3$ para cada $j = 1, \dots, 4$, pues en cada caso hay dos posiciones que deben llevar el mismo símbolo. Para las intersecciones dos a dos se tiene que $|A_i \cap A_j| = n^2$ para cada $i \neq j$, aunque, ¡atención!, el análisis es algo más delicado. Los dos siguientes esquemas analizan los casos $A_1 \cap A_3$ y $A_1 \cap A_4$, que resultan tener el mismo tamaño, n^2, aunque por "razones" distintas!:

$|A_1 \cap A_3| = n^2$





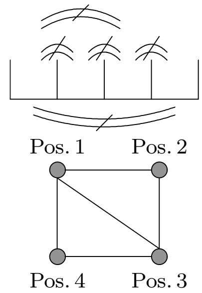
$|A_1 \cap A_4| = n^2$

Compruebe el lector que las intersecciones tres a tres son todas de tamaño n, que la intersección de los cuatro conjuntos también tiene tamaño n, y concluya que

$$|A| = |X| - \Big|\bigcup_{i=1}^4 A_i\Big| = \binom{4}{0}n^4 - \binom{4}{1}n^3 + \binom{4}{2}n^2 - \binom{4}{3}n + \binom{4}{4}n = n(n-1)(n^2-3n+3)\,,$$

que, por supuesto, coincide con el resultado visto en el apartado 2.3.

Observe, lector, cómo en el ejemplo anterior todas las intersecciones k a k tenían el mismo tamaño. Pero atienda al siguiente ejemplo, en el que el conjunto de restricciones es el que se representa a la derecha: hay una más que antes, representada por la arista diagonal. Ahora, el conjunto de listas prohibidas será la unión de los conjuntos



$$
\begin{aligned}
&A_1 = \{\text{listas con } 1^{\text{a}} = 2^{\text{a}}\},\ A_2 = \{\text{listas con } 2^{\text{a}} = 3^{\text{a}}\}\\
&A_3 = \{\text{listas con } 3^{\text{a}} = 4^{\text{a}}\},\ A_4 = \{\text{listas con } 1^{\text{a}} = 3^{\text{a}}\},\\
&A_5 = \{\text{listas con } 1^{a} = 4^{a}\}\,.
\end{aligned}
$$

Compruebe el lector que $|A_j| = n^3$ para cada j, y que $|A_i \cap A_j| = n^2$ para cada $i \neq j$. Veamos las intersecciones tres a tres, de las que hay $\binom{5}{3} = 10$. Por ejemplo, $|A_1 \cap A_2 \cap A_3| = n$, porque estar en A_1 exige tener el mismo símbolo en las dos primeras posiciones; estar en A_2 hace que la tercera lleve ese símbolo común; y estar en A_3 exige que ese símbolo aparezca también en la cuarta posición. Solo hay que decidir, pues, el símbolo común a toda la lista. La mayor parte de estas intersecciones 3 a 3 tienen tamaño n.

Sin embargo, $|A_1 \cap A_2 \cap A_4| = n^2$. Por estar en A_1 y en A_2, las tres primeras posiciones llevan el mismo símbolo; pero estar en A_4 no añade información, pues exige un símbolo común para la primera y tercera posiciones. Así que hay que elegir un símbolo para las tres primeras posiciones, y otro para la cuarta (n^2 posibilidades). También se tiene que $|A_3 \cap A_4 \cap A_5| = n^2$; compruébelo, lector, y note que estos dos casos se corresponden con configuraciones en las que las aristas correspondientes abarcan tres vértices en triángulo (véase el dibujo).

Si el lector se esmera analizando todos los casos, comprobará que las intersecciones cuatro a cuatro y la intersección cinco a cinco tienen todas el mismo tamaño, n, así que

$$
\begin{aligned}
\#\ \text{listas}\ &= |\mathcal{X}| - |A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5| = \binom{5}{0} n^4 - [\alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 - \alpha_5]\\
= \binom{5}{0} n^4 - \binom{5}{1} n^3 + \binom{5}{2} n^2 - \left[\binom{5}{3} - 2\right] n - 2n^2 + \binom{5}{4} n - \binom{5}{5} n = n(n-1)(n-2)^2.
\end{aligned}
$$

En este ejemplo, por cierto, podemos aplicar directamente la regla del producto: n posibilidades para la primera posición, $n-1$ para la segunda, $n-2$ para la tercera (recuérdese la arista diagonal en la representación con grafos) y otra vez $n-2$ (porque en las posiciones 1 y 3 van símbolos distintos) para la cuarta posición, para obtener $n(n-1)(n-2)^2$. ¡Vaya!

El procedimiento promete, pero también se anuncia como extraordinariamente aparatoso. Más adelante, en la sección 10.2.3, añadiremos algo de lenguaje (de grafos y colores) para desarrollar una manera ordenada de aplicar el principio de inclusión/exclusión en este contexto, que daremos en llamar el *polinomio cromático* de un grafo.

5.1.6. Coeficientes multinómicos

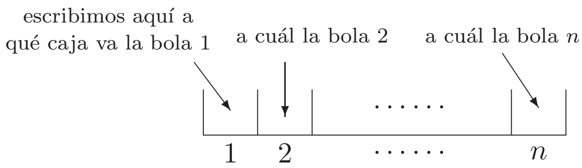
Los coeficientes multinómicos

$$
\binom{n}{a_1, a_2, \dots, a_k} = \frac{n!}{a_1!\, a_2!\, \dots a_k!}
$$

aparecieron asaz fugazmente en el apartado 5.1.2, al tratar, claro, el multinomio de Newton. La expresión de arriba lleva implícito el convenio según el cual el coeficiente multinómico es 0 si los enteros no negativos $a_1, \dots, a_k$ no suman exactamente n.

Vamos a ver ahora que estos coeficientes multinómicos dan respuesta a dos cuestiones combinatorias, y luego veremos por qué aparecen en el desarrollo del multinomio. Se trata de contar

- el número de formas de distribuir n bolas *numeradas* en k cajas *numeradas* de forma que haya a_1 bolas en la primera caja, a_2 en la segunda, etc., donde $a_1 + \cdots + a_k = n$;
- o el número de listas de longitud n formadas con los símbolos $\{1, \dots, k\}$ con exactamente a_1 unos, a_2 doses, etc. De nuevo, los números $a_1, \dots, a_k$ cumplen que $a_1 + \cdots + a_k = n$.



Las dos cuestiones son equivalentes: distribuir las n bolas en k cajas con las restricciones citadas exige decidir a qué caja va cada bola. Es decir, formar una lista de n posiciones (una por bola), en las que indicamos a qué caja va cada bola. Si queremos a_1 bolas en la caja 1, a_2 en la caja 2, etc., entonces en la lista han de aparecer exactamente a_1 unos, a_2 doses, etc. Y viceversa: dada la lista, recuperar la distribución de bolas en en cajas es inmediato.

Ya tuvimos, en la sección 5.1.3, un primer contacto con el lenguaje de bolas en cajas. Aunque allí tratábamos distribuciones de bolas *idénticas* en cajas numeradas; ahora, al considerar a las bolas como *numeradas*, esto es, *distinguibles*, no basta, como allí, con informar de *cuántas* bolas van en cada caja; hay que señalar también *cuáles* son.

Si la cuestión se hubiera planteado sin restricciones sobre el número de bolas por caja (o sobre el número de veces que aparece cada símbolo en la lista), el análisis sería directo: bastaría decidir a qué caja va cada bola; o cuál es el símbolo de cada posición de la lista. Lo que, por la regla del producto, se puede hacer de k^n maneras. Es otra manera de entender las listas (en este caso, de longitud n y k símbolos) o las aplicaciones $\mathcal{X} \to \mathcal{Y}$ si es que $|\mathcal{X}| = k$ e $|\mathcal{Y}| = n$ (revise el lector el ejemplo 2.2.5 y cambie allí los papeles de n y k).

Nuestro problema requiere un análisis diferente. Vemos primero un ejemplo sencillo, que nos dará la pista para el argumento general. Digamos que hay que repartir n bolas numeradas en 3 cajas numeradas, con a_1 bolas en la primera caja, a_2 en la segunda y a_3 en la tercera, donde $a_1 + a_2 + a_3 = n$. Contamos cuántas distribuciones hay con el siguiente proceso:

1. Elegimos *cuáles* (recordemos que las bolas son distinguibles) son las a_1 bolas que van a la caja 1. Esto se puede hacer de $\binom{n}{a_1}$ formas distintas.
2. Elegimos las bolas de la segunda caja de entre las $n - a_1$ bolas que quedan; tenemos $\binom{n-a_1}{a_2}$ maneras de hacerlo.
3. A la caja 3 irán las restantes, así que no habrá que decidir nada en este paso. Aunque también podríamos argumentar como en los dos pasos anteriores y observar que las bolas de la caja 3 se pueden elegir de $\binom{n-a_1-a_2}{a_3} = \binom{a_3}{a_3} = 1$ forma.

La regla del producto nos dice entonces que el número de posibles distribuciones es

$$\binom{n}{a_1}\binom{n-a_1}{a_2}\binom{a_3}{a_3} = \frac{n!}{a_1!(n-a_1)!}\frac{(n-a_1)!}{a_2!(n-a_1-a_2)!}\frac{(n-a_1-a_2)!}{a_3!0!} = \frac{n!}{a_1!\,a_2!\,a_3!}\,.$$

El lector podrá comprobar que el mismo argumento se puede aplicar al caso general, para obtener que el número de distribuciones de n bolas distinguibles en k cajas numeradas con a_1 bolas en la primera caja, a_2 en la segunda, etc., donde $a_1 + a_2 + \cdots + a_k = n$, viene dado por el coeficiente multinómico

$$\binom{n}{a_1, a_2, \ldots, a_k} = \frac{n!}{a_1!\,a_2!\,\ldots a_k!}\,.$$

Note el lector que el caso $k = 2$ es algo especial, pues si $a_1 + a_2 = n$, entonces $a_2 = n - a_1$, así que

$$\binom{n}{a_1, a_2} = \frac{n!}{a_1!\,a_2!} = \frac{n!}{a_1!\,(n-a_1)!} = \binom{n}{a_1}\,,$$

y recuperamos el coeficiente binómico habitual. Una cierta ambigüedad de notación que confiamos en que no cause confusión[43].

EJEMPLO 5.1.4 *Se pretende repartir los* 25 *empleados de una empresa en* 4 *grupos de trabajo: el de relación con los clientes debe constar de* 4 *personas; el de desarrollo de proyectos, de* 6 *personas;* 7 *personas irían al grupo de contabilidad, mientras que las* 8 *restantes trabajarían en tareas de organización interna. ¿De cuántas maneras se pueden estructurar estos grupos?*

En el lenguaje de bolas y cajas, queremos distribuir 25 bolas numeradas (las personas) en 4 cajas numeradas (los grupos de trabajo), con 4 en la primera, 6 en la segunda, 7 en la tercera y 8 en la cuarta. Respuesta:

$$\binom{25}{6,4,7,8} = \frac{25!}{4!\,6!\,7!\,8!} = 4\,417\,238\,826\,000 \quad \text{formas distintas,}$$

casi cuatro billones y medio, una cantidad por cierto enorme. ♣

¿Y por qué aparecen estos coeficientes multinómicos en el multinomio de Newton,

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{a_1,\ldots,a_k \geq 0} \binom{n}{a_1,\, a_2, \ldots,\, a_k} {x_1}^{a_1} {x_2}^{a_2} \cdots {x_k}^{a_k} \quad ?$$

Para argumentar combinatoriamente, habríamos de contar cuántas veces aparece cada término $x_1^{a_1} \cdots x_k^{a_k}$ en el desarrollo de $(x_1 + \cdots + x_k)^n$. El lector que construya un esquema como el que aparecía al principio de la sección 5.1.2 dedicada al teorema del binomio, poniendo en columna los n factores $(x_1 + \cdots + x_k)$, se convencerá que para que aparezca $x_1^{a_1} \cdots x_k^{a_k}$ hay que escoger a_1 veces el x_1, a_2 veces el x_2, etc. Lo que nos da el coeficiente multinómico correspondiente.

[43] Por cierto, los coeficientes binómicos cuentan conjuntos, ¡pero los multinómicos no cuentan multiconjuntos! Véase la sección 5.1.3.

Cerramos este breve apartado insistiendo en la interpretación en términos de listas, no tanto por el resultado, que ya es conocido, sino por una manera de argumentar que es útil en muchas ocasiones. Queremos contar cuántas n-listas podemos formar con los símbolos $\{1,\dots,k\}$ de manera que haya a_1 unos, a_2 doses, etc., donde $a_1+\cdots+a_k=n$. Ya sabemos que la respuesta está en el coeficiente multinómico correspondiente.

Pero argumentemos de la siguiente manera: hagamos que los símbolos sean *distinguibles*. ¿Cómo que distinguibles?, ¿pero de qué manera vamos a poder distinguir, por ejemplo, dos unos entre sí?, protestará el lector, alarmado ante este nuevo artificio. Pues justamente etiquetando de cierta manera cada símbolo. Para ello consideramos el nuevo conjunto de símbolos

$$1_1, 1_2, \dots, 1_{a_1}, 2_1, \dots, 2_{a_2}, \dots, k_1, \dots, k_{a_k}\,,$$

en el hay tantos "clones" del símbolo 1 como nos indica el valor de a_1, tantos del 2 como indique a_2, etc. Ahora, en total hay n símbolos distinguibles, que podemos disponer en una n-lista de $n!$ maneras distintas. Pero, para contrarrestar la diferenciación ficticia y auxiliar que hemos introducido, por ejemplo, entre los unos, deberemos dividir por un factor $a_1!$ (construya el lector explícitamente la aplicación $a_1!$ a 1 que hay detrás de este comentario). Lo mismo deberemos hacer con los doses, con los treses, etc. Así llegamos, de nuevo, al resultado

$$\frac{n!}{a_1!\,a_2!\cdots a_k!}=\binom{n}{a_1,\,a_2,\dots,a_k}.$$

EJEMPLO 5.1.5 *¿Cuántas listas de longitud* 8 *formadas por* 2 *aes,* 3 *bes y* 3 *ces hay?*

Las listas están formadas por los símbolos a,a,b,b,b,c,c,c, pero por supuesto no todas las 8! permutaciones posibles dan lugar a listas distintas. Aplicamos de nuevo el truco citado antes: hacemos distinguibles los símbolos, $a_1,a_2,b_1,b_2,b_3,c_1,c_2,c_3$, y formamos las 8! permutaciones posibles. Obtenemos así, por ejemplo, listas como

$$\boxed{a_1 \mid a_2 \mid b_1 \mid b_2 \mid b_3 \mid c_3 \mid c_2 \mid c_1} \quad \text{y} \quad \boxed{a_2 \mid a_1 \mid b_1 \mid b_2 \mid b_3 \mid c_3 \mid c_2 \mid c_1}$$

que, al borrar los subíndices de las aes, dan lugar a la misma, (a,a,b,b,b,c,c,c). Solucionamos esta proliferación espuria de listas dividiendo por los factoriales correspondientes:

$$\frac{8!}{2!\,3!\,3!}=\binom{8}{3,\,2,\,3}=\frac{8\times7\times6\times5\times4}{2\times2\times3}=560\,.$$

♣

5.1.7. Fórmulas de inversión binómica

Vamos a ver ahora cómo muchos de los resultados que hemos visto hasta ahora pueden entenderse como ejemplos de aplicación de unas fórmulas de "inversión" que involucran coeficientes binómicos. El plural aquí es relevante: habrá dos, una "inferior" y otra "superior".