El diablo de los números

Sección a cargo de

Javier Fresán

Teoría probabilística de números*

por

Emmanuel Kowalski

By this art you may contemplate
the variation of the 23 *letters...*
J. L. Borges,
La biblioteca de Babel

I could be bounded in a nutshell and
count myself a king of infinite space
W. Shakespeare,
Hamlet

1. Introducción

¿Qué es la teoría probabilística de números? En sentido amplio, se podría decir que coincide, más o menos, con la teoría analítica de números (y es quizás por esta razón que ciertos aritméticos, para evitar que su trabajo se asocie con el análisis, prefieren el término «estadística aritmética»), pero en este artículo la abordaremos desde un punto de vista más restringido que nos permitirá concentrarnos en algunos aspectos específicos (elegidos por el autor en función de sus gustos personales): entenderemos la teoría probabilística de números como el *estudio del comportamiento asintótico de sucesiones (o familias) de variables aleatorias definidas de modo aritmético*. Este es también el punto de vista de mi libro [16], que los lectores pueden considerar como la extensión y continuación natural de este texto. El artículo [20] de Perret-Gentil es otra excelente referencia que cubre una gran variedad de interacciones recientes entre la teoría de números y la probabilidad.

Para ilustrar los temas en los que nos concentraremos, empezamos enunciando el que se suele considerar el primer gran resultado de la teoría probabilística de números: el teorema de Erdős-Kac [3], publicado en 1940.

*Traducido del inglés por Javier Fresán.

Teorema 1.1 (Erdős-Kac). *Para cada entero $n \geq 1$, sea $\omega(n)$ el número de divisores primos distintos de n. Cuando $N \to +\infty$, las variables aleatorias*

$$n \longmapsto \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}}$$

definidas sobre el espacio de probabilidad finito $\Omega_N = \{1, \ldots, N\}$, con la medida de probabilidad uniforme, convergen en ley a la variable aleatoria gaussiana estándar de esperanza 0 y de varianza 1.

En términos concretos (y como se enunció por primera vez), el teorema significa que para dos números reales cualesquiera $a < b$, se tiene

$$\lim_{N\to+\infty} \frac{1}{N}\left|\left\{n \leq N \mid a < \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}} < b\right\}\right| = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\, dt$$

(y en esta igualdad ni siquiera la existencia del límite es en absoluto obvia). Así, la distribución de probabilidad más importante se puede obtener a partir de un objeto extremadamente simple y en apariencia determinístico como es el número de divisores primos de enteros «aleatorios». Considerar los intervalos $\{1, \ldots, N\}$ cuando $N \to +\infty$ es una forma natural de hablar de enteros «aleatorios», pero no es la única, como veremos más adelante (Ejemplo 2.2).

El apéndice contiene un recordatorio de las nociones básicas de probabilidad, así como de algunos resultados que se usan a lo largo del texto, pero la convergencia en ley es tan esencial que preferimos enunciarla aquí en un contexto bastante general.

Definición 1.2. Sea (X_N) una sucesión de variables aleatorias definidas sobre ciertos espacios de probabilidad[1] y con valores en un espacio métrico fijo T. Sea X una variable aleatoria con valores en T. La sucesión (X_N) converge en ley a X si, para toda función continua y acotada $f\colon T \to \mathbb{C}$, se cumple

$$\lim_{N\to+\infty} \mathbf{E}(f(X_N)) = \mathbf{E}(f(X)).$$

Aunque para la mayoría de los resultados bastará considerar $T = \mathbb{R}$ o un espacio vectorial de dimensión finita, veremos en la Sección 4 que la teoría probabilística de números puede también hacer intervenir teoremas de límites funcionales, en los que el espacio T es un espacio vectorial de dimensión infinita.

En la sección siguiente, esbozamos la demostración del Teorema 1.1, tras haber explicado algunos resultados más antiguos que son también de gran interés. En la Sección 3, presentamos uno de los teoremas más importantes de la teoría probabilística de números: el teorema de Selberg sobre la distribución límite de $\log|\zeta(1/2+it)|$, donde $\zeta(s)$ designa la función zeta de Riemann. Como la prueba completa es bastante delicada, nos contentaremos con motivar el resultado. Se trata de un área en la que la mayor parte de los trabajos más fascinantes se han llevado a cabo en los últimos años, con nuevas conexiones con conceptos sofisticados de la teoría de probabilidad

[1] Que pueden depender de N.

«pura» como las caminatas aleatorias ramificadas o el caos multiplicativo gaussiano. En la Sección 4, explicamos un resultado de naturaleza muy diferente, debido a W. Sawin y al autor, en el que los teoremas límite involucran variables aleatorias con valores en el espacio $\mathcal{C}([0,1])$ de funciones continuas en el intervalo $[0,1]$ y las herramientas aritméticas clave en la demostración, basadas en la versión de Deligne de la hipótesis de Riemann sobre cuerpos finitos [2], son especialmente profundas.

Agradecimientos. Quisiera dar las gracias a J. Fresán por sugerirme que escribiese este artículo y traducirlo al español. La sección sobre el teorema de Selberg está inspirada en una charla que di en el seminario Betty B. en marzo de 2018, como preparación a la de A. Harper en el seminario N. Bourbaki el día siguiente. Agradezco a N. y B. Bourbaki su invitación.

2. Teoría probabilística de números clásica

2.1. El teorema de Schoenberg

Aunque se suele considerar el teorema de Erdős-Kac como el punto de partida de la teoría probabilística de números, se podría argumentar que los resultados anteriores de Schoenberg quizá sean más merecedores de ese título. De hecho, yendo más lejos, se podría decir que el origen filosófico de la teoría probabilística de números es el siguiente hecho elemental, que no solo yace en el corazón tanto de los resultados de Schoenberg como del teorema de Erdős-Kac, sino que además explica por qué es posible que haya interacciones interesantes entre probabilidad y teoría de números.

Teorema 2.1. *Dado un entero positivo q, sea π_q la variable aleatoria sobre Ω_N con valores en $\mathbb{Z}/q\mathbb{Z}$ definida por $\pi_q(n) = n \pmod q$. Sea $\mathcal{Q}$ un conjunto finito de enteros positivos coprimos entre sí y sea Q el producto de los elementos de $\mathcal{Q}$.*

La familia $(\pi_q)_{q\in\mathcal{Q}}$ converge en ley cuando $N \to +\infty$ a una familia de variables aleatorias independientes $(U_q)_{q\in\mathcal{Q}}$, donde U_q está distribuida uniformemente en $\mathbb{Z}/q\mathbb{Z}$. De hecho, toda función

$$f\colon \prod_{q\in\mathcal{Q}} \mathbb{Z}/q\mathbb{Z} \to \mathbb{C}$$

cumple la desigualdad

$$\left|\frac{1}{N}\sum_{n\in\Omega_N} f((\pi_q(n))_q) - \frac{1}{Q}\sum_{x\in\prod\mathbb{Z}/q\mathbb{Z}} f(x)\right| \leq \frac{1}{N}\sum_{x\in\prod\mathbb{Z}/q\mathbb{Z}} |f(x)|.$$

Observación. En otras palabras, estamos ante un caso de la Definición 1.2 con

$$T = \prod_{q\in\mathcal{Q}} \mathbb{Z}/q\mathbb{Z}, \quad X_N(n) = (\pi_q(n))_{q\in\mathcal{Q}} \quad \text{y} \quad X((a_q)_{q\in\mathcal{Q}}) = (a_q)_{q\in\mathcal{Q}}.$$

Que las componentes de X son independientes y se distribuyen uniformemente es más o menos tautológico a partir de la definición de familia independiente (véase la Sección A.1 del apéndice).

En cierto modo, este enunciado es «obvio», y como tal se usa sin mayor explicación en muchos de los tratamientos de la teoría probabilística de números clásica. Cualitativamente, se deduce de la combinación del teorema chino del resto, que afirma que el producto de $\mathbb{Z}/q\mathbb{Z}$ indexado por $q \in \mathcal{Q}$ se puede identificar con $\mathbb{Z}/Q\mathbb{Z}$, y del límite elemental

$$\lim_{N\to+\infty} \frac{1}{N}\big|\{n \in \Omega_N \mid n \equiv a \,(\mathrm{mód}\, Q)\}\big| = \frac{1}{Q}$$

para todo a módulo Q. Sin embargo, casi cualquier intento de obtener generalizaciones no triviales del teorema de Erdős-Kac probablemente pasa por generalizar el Teorema 2.1, lo cual puede hacer intervenir matemáticas muy profundas.

Ejemplo 2.2.

1. Si se sustituyen los enteros entre 1 y N por los valores $p-1$, donde p recorre los números primos entre 1 y N (de nuevo con la medida de probabilidad uniforme en este conjunto de primos), el análogo del Teorema 2.1 está relacionado directamente con cuestiones sobre números primos en progresiones aritméticas que están íntimamente ligadas a la hipótesis de Riemann generalizada.
2. Fijemos un entero $m \geq 2$. Si se sustituyen los enteros entre 1 y N por los enteros $\mathrm{Tr}(g)$, donde g recorre uniformemente las matrices de $\mathrm{SL}_m(\mathbb{Z})$ cuyos coeficientes son todos $\leq N$, el análogo del Teorema 2.1 está íntimamente relacionado con el estudio de las brechas espectrales de subgrupos de congruencia de $\mathrm{SL}_m(\mathbb{Z})$ (véase, por ejemplo, la discusión en [15]), lo cual tiene que ver a su vez con grafos expansores, formas automorfas, la propiedad (T) de Kazhdan...

El Teorema 2.1 se suele aplicar a un conjunto $\mathcal{Q}$ formado por números primos, en cuyo caso la condición de coprimalidad es automática. Desde el punto de vista probabilístico, lo que es crucial es que aparezcan *sucesiones de variables aleatorias independientes*. Teniendo en cuenta que la independencia es uno de los conceptos más fundamentales de la probabilidad (una de las primeras cosas que la distingue de la teoría de la integración), no debería sorprendernos que existan interacciones entre la probabilidad y la teoría de números.

Gracias al Teorema 2.1, podemos motivar un bello resultado de Schoenberg [23]. Como se verá en nuestro esbozo, dar una demostración completa no requiere más información aritmética, y solo teoría de probabilidad muy elemental.

Recordemos la definición de la función φ de Euler: si n es un entero positivo, $\varphi(n)$ es el número de clases residuales invertibles módulo n o, de modo equivalente, el número de enteros a coprimos con n tales que $0 \leq a \leq n$. Por ejemplo, $\varphi(n) = n-1$ equivale a decir que n es primo. En cierto sentido, el entero $\varphi(n)$ está siempre «relativamente cerca» de n, pero el cociente[2] $\varphi(n)/n$ no tiene límite cuando n tiende a infinito. De hecho, los experimentos muestran enseguida que este cociente varía erráticamente en $[0,1]$: si el lector escribe un entero grande, digamos de 50 dígitos, e intenta adivinar el valor de $\varphi(n)/n$, podría quedar desconcertado ante el resultado. El teorema de Schoenberg proporciona una explicación.

[2]Que se interpreta, probabilísticamente, como la probabilidad respecto a la medida uniforme de que una clase residual sea invertible.

Teorema 2.3 (Schoenberg). *Para cada entero $N \geq 1$, sea F_N la variable aleatoria sobre Ω_N dada por $n \mapsto \varphi(n)/n$. Cuando $N \to +\infty$, la sucesión (F_N) converge en ley a una variable aleatoria F, dada por el producto infinito sobre los números primos*

$$F = \prod_p \Big(1 - \frac{B_p}{p}\Big), \tag{1}$$

donde (B_p) es una sucesión de variables aleatorias independientes de Bernoulli, indexadas por los números primos, tales que, para cada p, se tiene

$$\mathbf{P}(B_p = 1) = \frac{1}{p}, \qquad \mathbf{P}(B_p = 0) = 1 - \frac{1}{p}.$$

Este producto infinito converge casi seguramente.

Esbozo de demostración. El punto de partida es la conocida fórmula

$$\frac{\varphi(n)}{n} = \prod_{p|n} \Big(1 - \frac{1}{p}\Big) \tag{2}$$

para la función de Euler normalizada. Esta igualdad refleja el hecho de que el grupo $(\mathbb{Z}/n\mathbb{Z})^\times$ de elementos invertibles del anillo $\mathbb{Z}/n\mathbb{Z}$ tiene orden $\varphi(n)$ y de que el teorema chino del resto proporciona un isomorfismo de grupos

$$(\mathbb{Z}/n\mathbb{Z})^\times \simeq \prod_{p|n} (\mathbb{Z}/p^{v_p(n)}\mathbb{Z})^\times,$$

donde $v_p(n)$ es el exponente del número primo p en la factorización de n. La fórmula resulta entonces del cálculo

$$\varphi(p^v) = p^v - p^{v-1},$$

de modo que

$$\frac{\varphi(p^v)}{p^v} = 1 - \frac{1}{p}$$

para cualquier número primo p y cualquier entero $v \geq 1$. Reescribamos (2) de forma inteligente como el producto infinito

$$\frac{\varphi(n)}{n} = \prod_p \Big(1 - \frac{b_p(n)}{p}\Big) \tag{3}$$

sobre los números primos, donde $b_p(n)$ es igual a 1 si p divide a n y a 0 si no.

Si restringimos n a Ω_N y vemos b_p como una variable aleatoria, *por definición* y por el Teorema 2.1, estas variables aleatorias convergen en ley cuando $N \to +\infty$ a la variable aleatoria de Bernoulli B_p con «probabilidad de éxito» $1/p$. (Intuitivamente, todo lo que estamos diciendo es que un «entero aleatorio» tiene probabilidad $1/p$ de ser divisible por p, lo cual es lógico.) Además, aplicando de nuevo el Teorema 2.1, se ve que, para cualquier familia finita de números primos distintos $(p_i)_{1 \leq i \leq k}$, las

tuplas $(b_{p_1}(n), \dots, b_{p_k}(n))$ con $n \in \Omega_N$ convergen en ley a la tupla $(B_{p_1}, \dots, B_{p_k})$, en la cual las variables aleatorias de Bernoulli son *independientes*.

Llegado este punto, teniendo en cuenta que la familia (b_p) restringida a Ω_N converge en ley a (B_p) cuando $N \to +\infty$, es natural esperar que (3) admita una variable aleatoria límite dada por el producto infinito (1): basta simplemente con tomar el límite término a término.

Esbozamos una forma elemental de obtener una demostración rigurosa basada en el lema probabilístico A.2 del apéndice. Grosso modo, consiste en encontrar aproximaciones sucesivas de las variables aleatorias, para cada N, tales que cada una de ellas converja en ley y que la diferencia sea pequeña en promedio. En nuestro caso, la elección natural consiste en aproximar por los productos finitos (que veremos también como variables aleatorias sobre Ω_N)

$$\prod_{p \le M} \Big(1 - \frac{b_p(n)}{p}\Big),$$

donde $M \ge 1$ es un parámetro que determina la calidad de las aproximaciones. Como el producto es finito, se sigue inmediatamente del Teorema 2.1 que estas aproximaciones, restringidas a Ω_N, convergen en ley cuando $N \to \infty$ a

$$\prod_{p \le M} \Big(1 - \frac{B_p}{p}\Big).$$

Para comprobar la calidad de las aproximaciones, observamos que

$$\begin{aligned}\Big|\frac{1}{N}\Big(\sum_{n \le N} \frac{\varphi(n)}{n} - \sum_{n \le N} \prod_{p \le M} \Big(1 - \frac{1}{p}\Big)\Big)\Big| &\le \frac{1}{N} \sum_{n \le N} \Big| \prod_{\substack{p \mid n \\ p > M}} \Big(1 - \frac{1}{p}\Big) - 1\Big| \\ &\le \frac{1}{N} \sum_{n \le N} \sum_{d \in D_n} \frac{1}{d},\end{aligned}$$

donde D_n es el conjunto de enteros $d \ge 2$ que dividen a n y que solo tienen factores primos $p > M$. En particular, todos los elementos d de D_n cumplen $M < d \le N$, e intercambiando el orden de las sumas sobre n y sobre d, se obtiene

$$\frac{1}{N} \sum_{n \le N} \sum_{d \in D_n} \frac{1}{d} \le \sum_{M < d \le N} \frac{1}{d} \times \frac{1}{N} \sum_{\substack{n \le N \\ d \mid n}} 1 \le \sum_{d > M} \frac{1}{d^2}.$$

En esta desigualdad, el término más a la derecha es una función $f(N, M)$ de N y M que tiende a 0 cuando $M \to +\infty$ uniformemente con respecto a N (puesto que no depende de N). Por tanto, se cumplen las condiciones del Lema A.2.

Así se termina la prueba del teorema de Schoenberg, salvo por el enunciado sobre la convergencia del producto infinito (1). Este último es una consecuencia simple, por ejemplo, del teorema de las tres series de Kolmogorov (Teorema A.3). □

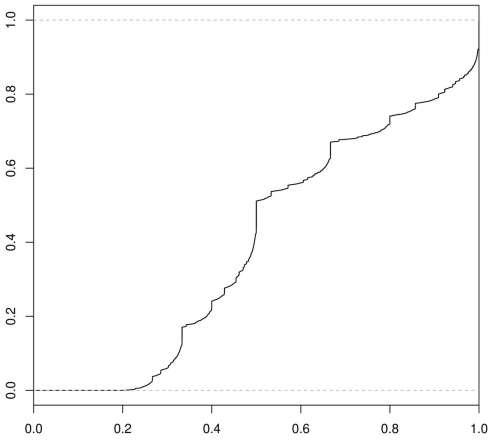

Figura 1: Grafo empírico de la función de distribución de $\varphi(n)/n$ para $n \leq 10^6$.

La variable aleatoria F que aparece en el teorema es muy interesante, pues nada parece indicar que se trate de una variable aleatoria «estándar» (gaussiana, Poisson, exponencial, uniforme, etc.). De hecho, pertenece a una clase de variables aleatorias con valores reales cuya ley de probabilidad es *singular* respecto a la medida de Lebesgue, un tipo de objeto «patológico» que aparece normalmente en la teoría de la integración construido a partir de conjuntos de tipo Cantor. Vemos así hasta qué punto la teoría de números puede llegar a ser salvaje e impredecible... De forma más precisa, si definimos la función $f(x) = \mathbf{P}(F \leq x)$ para $x \in \mathbb{R}$ (la *función de distribución de probabilidad* de la variable aleatoria F), se tiene:

Proposición 2.4 (Erdős). *La función f es continua en $\mathbb{R}$, estrictamente creciente en $[0,1]$ y diferenciable casi en todas partes en $\mathbb{R}$ respecto a la medida de Lebesgue. Además, $f'(x) = 0$ para casi todo x respecto a la medida de Lebesgue.*

Ilustramos esta proposición, cuya prueba por Erdős es muy elegante (véase, por ejemplo, [16, Exercise 1.4.4]) con una aproximación del grafo de la función $f(x)$ para $0 \leq x \leq 1$ dada por los valores $F(n)$ con $n \leq 10^6$ (véase la Figura 1).

2.2. El teorema de Erdős-Kac

El método de demostración del teorema de Schoenberg se aplica a muchas otras funciones aritméticas además de a la función de Euler. De hecho, resulta en general más fácil trabajar con funciones f definidas sobre los enteros positivos que cumplen la propiedad de aditividad

$$f(nm) = f(n) + f(m)$$

siempre que n y m sean coprimos, por ejemplo la función $f(n) = \log(\varphi(n)/n)$. Hay resultados generales de distribución para funciones aditivas que permiten recuperar el teorema de Schoenberg como caso particular.

Entre las funciones aditivas más sencillas se encuentra la función ω dada por el número de divisores primos distintos de n. El objetivo del teorema de Erdős-Kac es precisamente entender su distribución cuando n toma valores elegidos al azar uniformemente en $\Omega_N = \{1, \dots, N\}$.

Usando la notación b_p de la sección anterior para la variable aleatoria sobre Ω_N que indica si n es divisible por p o no, podemos escribir

$$\omega(n) = \sum_p b_p(n).$$

Parecería, por tanto, razonable esperar que el comportamiento asintótico de ω fuese similar al de la suma $\sum_p B_p$. El asunto es más complicado que en el caso del Teorema 2.3: sin ningún truncamiento, la suma diverge casi seguramente, por la dirección no trivial del lema de Borel-Cantelli y el hecho de que la suma

$$\sum_p \mathbf{P}(B_p = 1) = \sum_p \frac{1}{p}$$

diverge (un hecho bien conocido sobre la distribución de los primos que se remonta a Euler y que recordaremos en la sección siguiente). Pero observando que todo primo que divide a $n \leq N$ es necesariamente $\leq N$, podemos trabajar con la suma finita

$$W_N = \sum_{p \leq N} B_p$$

y esperar que sea una buena aproximación de la distribución de la función $\omega(n)$ sobre Ω_N para N grande. Esto es, de hecho, lo que ocurre, pero la aproximación no es tan directa como la convergencia en ley (la sucesión (W_N) no converge en ley cuando $N \to +\infty$). Sin embargo, como (B_p) son variables aleatorias independientes de Bernoulli, es un simple ejercicio de probabilidad comprobar que las variables aleatorias renormalizadas

$$\frac{W_N - \mathbf{E}(W_N)}{\sqrt{\mathbf{V}(W_N)}}$$

(que tienen esperanza 0 y varianza 1) convergen en ley a una gaussiana estándar. Por otra parte, para las variables aleatorias W_N se tiene

$$\mathbf{E}(W_N) = \sum_{p \leq N} \frac{1}{p} \sim \log\log N,$$

$$\mathbf{V}(W_N) = \sum_{p \leq N} \frac{1}{p}\Big(1 - \frac{1}{p}\Big) \sim \log\log N,$$

donde la asintótica para la suma de los inversos de los primos se conoce como *fórmula de Mertens* (véase también (10), más abajo, para entender de dónde viene).

Aunque este es un enfoque algo inusual del teorema de Erdős-Kac, es posible hacerlo riguroso, por ejemplo calculando asintóticamente los momentos

$$\frac{1}{N} \sum_{n \leq N} \Big(\sum_{p \leq Q} b_p(n) \Big)^k$$

para todo entero $k \geq 1$ y haciendo una elección conveniente de $Q < N$, lo bastante pequeña para que el cálculo sea factible y lo bastante grande para que el número de factores primos de $n \leq N$ mayores que Q sea despreciable. Si se argumenta de este modo, el hecho aritmético clave vuelve a ser el Teorema (elemental) 2.1.

En su artículo [7], Harper propone otro acercamiento muy probabilístico al teorema de Erdős-Kac, basado en el método de Stein de aproximación gaussiana y de Poisson. Es también posible evitar el truncamiento que involucra al parámetro Q y calcular asintóticamente las funciones características de W_N y de $\omega(n)$ sobre Ω_N. Los primeros en hacerlo fueron Rényi y Turán, y ahondando en la forma precisa del resultado se llega a la llamada «convergencia mod-Poisson» introducida por el autor y Nikeghbali [17], que pone de manifiesto conexiones intrigantes con las permutaciones aleatorias y los polinomios aleatorios sobre cuerpos finitos. (El artículo [5] de A. Granville, muy agradable de leer, contiene un divertido resumen de todos estos resultados. Su cómic en colaboración con J. Granville [6] —algo fallido en opinión del autor— también puede atraer a ciertos lectores).

Ejemplo 2.5. Erdős observó que el hecho de que los enteros $n \leq N$ suelan tener alrededor de $\log\log N$ factores primos[3] tiene una aplicación muy interesante al «problema de la tabla de multiplicar». Consideremos una tabla de multiplicar de enteros $1 \leq n \leq N$, como las que se pueden encontrar para N alrededor de 10 en muchas escuelas. ¿Cuántos enteros aparecen en la tabla de multiplicar? En otras palabras, ¿cuántos enteros $m \leq N^2$ se pueden escribir como un producto $m = ab$ con $1 \leq a, b \leq N$? Si llamamos $m(N)$ ese número, el primer resultado básico es que

$$\lim_{N \to +\infty} \frac{m(N)}{N^2} = 0,$$

es decir, que «la mayoría de» los enteros $n \leq N$ no aparecen en la tabla de multiplicar. Intuitivamente, la razón es que, como la mayoría de los enteros $n \leq N$ tienen alrededor de $\log\log N$ factores primos, la mayoría de los productos ab tienen alrededor de $2\log\log N$ factores primos. Pero este fenómeno es muy atípico en el caso de enteros $m \leq N^2$, que de nuevo deberían tener alrededor de $\log\log N^2 \sim \log\log N$ factores primos. El (sorprendente) orden de magnitud de $m(N)$ fue determinado por Ford [4], tras los trabajos de varios autores, especialmente Tenenbaum; sus argumentos reposan sobre resultados probabilísticos muy sutiles.

3. El teorema de Selberg

En esta sección, presentamos el segundo teorema más famoso de la teoría probabilística de números, que es el punto de partida de las que probablemente sean hoy en día las investigaciones más profundas en esta disciplina. Se trata de un teorema de Selberg [24] sobre el comportamiento límite del logaritmo de la función zeta de Riemann en la llamada recta crítica.

[3] La información necesaria es más fácil de demostrar que el teorema de Erdös-Kac completo.

3.1. La función zeta de Riemann

La función zeta de Riemann es la función meromorfa sobre $\mathbb{C}$ obtenida por continuación analítica de la función definida por la *serie de Dirichlet*

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s}$$

en el conjunto abierto de los números complejos $s \in \mathbb{C}$ tales que $\mathrm{Re}(s) > 1$. En este dominio, la serie converge absoluta y uniformemente en subconjuntos compactos, de lo cual se deduce que la función zeta es holomorfa.

Hay múltiples métodos para obtener la continuación analítica e identificar los polos (el libro clásico [25, Ch. II] de Titchmarsh reseña hasta siete). Uno de los más intrínsecos (por su relación con las formas modulares, en este caso las funciones theta) consiste en utilizar la expresión

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \frac{1}{s(s-1)} + \frac{1}{2}\int_1^{+\infty} (\theta(x)-1)\big(x^{s/2} + x^{(1-s)/2}\big)\frac{dx}{x}, \qquad (4)$$

donde θ es la función theta, dada por

$$\theta(x) = \sum_{n\in\mathbb{Z}} e^{-\pi n^2 x}$$

para cada $x > 0$. Esta fórmula se deduce, con relativa facilidad, de la definición de la función gamma,

$$\Gamma(s) = \int_0^{+\infty} e^{-t} t^s \frac{dt}{t},$$

combinada con la fórmula de sumación de Poisson. Usando el decrecimiento rápido de la función theta cuando $x \to +\infty$, se ve fácilmente que el lado derecho de (4), y por tanto la función zeta de Riemann, es una función meromorfa sobre $\mathbb{C}$ con un único polo simple en $s = 1$ de residuo 1.

Para entender el teorema de Selberg, basta de hecho con saber que se puede definir $\zeta(s)$ para $\mathrm{Re}(s) > 0$ (excepto si $s = 1$). Esta propiedad se demuestra de forma elemental, por ejemplo usando la siguiente «sumación por partes»: sea $\{t\}$ la parte fraccionaria de un número real $t \geq 0$, de modo que $\{t\} = t - n$ donde $n \geq 0$ es el mayor entero $n \leq t$. Suponiendo $\mathrm{Re}(s) > 1$, calculamos

$$\begin{aligned}
\sum_{n\geq 1} \frac{1}{n^s} &= s\int_1^{+\infty} \Big(\sum_{1\leq n\leq t} 1\Big) t^{-s-1}\, dt \\
&= s\int_1^{+\infty} (t - \{t\}) t^{-s-1}\, dt \\
&= s\int_1^{+\infty} t^{-s}\, dt - s\int_1^{+\infty} \{t\} t^{-s-1}\, dt \\
&= \frac{s}{s-1} + (\text{función holomorfa para } \mathrm{Re}(s) > 0),
\end{aligned}$$

y esta fórmula proporciona la continuación analítica deseada.

La razón por la cual los teóricos de números se interesan por la función zeta de Riemann es que da acceso a muchas propiedades de los números primos. Esto se debe a la notable *fórmula del producto de Euler*, que afirma que para todo número complejo s tal que $\mathrm{Re}(s) > 1$, se tiene la relación

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1}, \tag{5}$$

en la cual el producto infinito sobre los números primos es absolutamente convergente. Se trata de un hecho bastante simple, y sin embargo esencial, así que explicamos la demostración (en la que veremos que el contenido aritmético de la fórmula es la existencia y unicidad de la factorización de los enteros en números primos). En primer lugar, si para un número real fijo $P \geq 2$ desarrollamos $(1 - p^{-s})^{-1}$ en serie geométrica y multiplicamos las series resultantes para los primos $p \leq P$ obtenemos

$$\prod_{p \leq P} (1 - p^{-s})^{-1} = \prod_{p \leq P} \sum_{k \geq 0} p^{-ks} = \sum_{n \in N_P} n^{-s},$$

donde N_P es el conjunto de enteros $n \geq 1$ únicamente divisibles por primos $\leq P$. Obsérvese que la unicidad de la factorización en números primos implica que cada entero aparece como mucho una vez, mientras que de la existencia de la factorización se deduce que todo entero $n \leq P$ pertenece a N_P. La fórmula de Euler para $\mathrm{Re}(s) > 1$ se obtiene haciendo tender P a infinito.

Observación. Las aplicaciones más sencillas de esta fórmula son demostraciones de que existen infinitos números primos:

- (Euler) Si no, $\zeta(\sigma)$ estaría acotada para todo $\sigma > 1$, lo cual no es cierto.
- (Hacks) Si no, $\zeta(2) = \pi^2/6$ (Euler) sería un número racional, lo cual no es cierto (Lambert).

3.2. El teorema de los números primos

Riemann [22] descubrió cómo explotar sistemáticamente el producto de Euler para obtener fórmulas asintóticas para el número $\pi(x)$ de primos $p \leq x$. Dada la importancia de esta idea, la discutimos brevemente antes de pasar al teorema de Selberg. La intuición no es difícil de explicar. En primer lugar, es más conveniente trabajar con la función

$$\psi(x) = \sum_{\substack{p \text{ primo, } k \geq 1 \\ p^k \leq x}} \log p = \sum_{n \leq x} \Lambda(n),$$

donde Λ es la llamada función de Von Mangoldt, que se anula si n no es una potencia (no trivial) de un número primo y vale $\Lambda(p^k) = \log p$ para todo $k \geq 1$. Las

funciones $\pi(x)$ y $\psi(x)$ se comparan fácilmente, pero la segunda es más sencilla de manejar debido a la fórmula

$$\sum_{n\geq 1} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}, \tag{6}$$

que relaciona directamente la función de Von Mangoldt con la derivada logarítmica de la función zeta de Riemann (para demostrarla basta derivar término a término el producto de Euler). A partir de las propiedades de la función zeta, vemos que (6) es también una función meromorfa, con polos simples localizados como sigue:

- Hay un único polo simple con residuo 1 en $s = 1$.
- Hay polos simples en todos los ceros $s = \rho$ de la función zeta de Riemann con residuo igual al opuesto de la multiplicidad de ρ como cero.

Mediante un cálculo basado en la fórmula de Perron, que es un avatar de la inversión de Fourier, se obtiene (esencialmente tal y como lo hizo Riemann)

$$\psi(x) = \frac{1}{2i\pi} \int_{(3)} -\frac{\zeta'(s)}{\zeta(s)}\, x^s\, \frac{ds}{s},$$

donde la integral está tomada a lo largo de la recta vertical de parte real[4] 3, orientada de la parte imaginaria negativa a la parte imaginaria positiva.

Supongamos que sabemos que existe un número real $0 < c < 1$ tal que todos los ceros de la función zeta de Riemann tienen parte real menor que c (y, además, que la función zeta no crece «muy rápido» cuando la parte imaginaria es grande). En ese caso, aplicando la fórmula integral de Cauchy a un rectángulo de lados paralelos a los ejes, con lados horizontales a altura $\pm T$ y lados verticales de partes reales c y 3, y haciendo tender T a infinito, obtenemos

$$\psi(x) = x + \frac{1}{2i\pi} \int_{(c)} -\frac{\zeta'(s)}{\zeta(s)}\, x^s\, \frac{ds}{s} = x + O(x^c) \tag{7}$$

para todo $x \geq 2$, al menos si la integral sobre la recta de parte real c converge. Si $c < 1$, esta fórmula determina el comportamiento asintótico de $\psi(x)$. El *teorema de los números primos* resulta fácilmente: se tiene

$$\lim_{x\to+\infty} \frac{\pi(x)}{x/\log(x)} = 1, \tag{8}$$

pero la fórmula (7) es mucho más precisa.

Riemann considera a continuación la posibilidad de que todos los ceros en la región $0 < \mathrm{Re}(s) < 1$ tengan parte real[5] 1/2, que es el mejor resultado que uno puede esperar puesto que hay ceros con parte real 1/2 (el propio Riemann calculó

[4] Se puede sustituir el número 3 por cualquier número mayor que 1.

[5] «Es ist sehr wahrscheinlich» (es decir, «es muy probable»).

aproximaciones numéricas de los primeros). Si así fuera, el argumento que hemos esbozado establecería la asintótica

$$\pi(x) = \int_2^x \frac{dt}{\log t} + O(x^{1/2+\varepsilon})$$

para todo $x \geq 2$ y todo $\varepsilon > 0$, con una constante implícita que solo depende de ε.

Sin embargo, esta hipótesis de Riemann sigue hoy en día sin demostración, y tampoco se sabe si la estrategia que hemos descrito funciona tal cual: ¡no se conoce ningún valor $c < 1$ con la propiedad deseada! Fue por tanto un enorme progreso que Hadamard y De la Vallée Poussin consiguieran demostrar independientemente en 1896 que la función zeta de Riemann no se anula si $\mathrm{Re}(s) \geq 1$, y deducir de ello el teorema de los números primos (8), aunque no en una forma tan precisa como (7).

3.3. ¿Por qué considerar la función zeta probabilísticamente?

A simple vista, se podría pensar que una función holomorfa o meromorfa como la función zeta de Riemann es un objeto muy regular y determinístico. Sin embargo, el comportamiento de la función zeta en la parte del plano complejo más relevante para la teoría de números (es decir, $0 \leq \mathrm{Re}(s) \leq 1$) es extremadamente complejo.

Por ejemplo, Riemann ya enunció correctamente que el número de ceros $\rho = \beta + i\gamma$ de la función zeta tales que $0 \leq \beta \leq 1$ y $|\gamma| \leq T$ es asintótico a $\pi^{-1}T\log(T)$ cuando T tiende a infinito. En particular, en una caja de altura 1 en la que la parte imaginaria pertenece a $[T, T+1]$, el número de ceros es en promedio aproximadamente igual a $\pi^{-1}\log(T)$, y por tanto crece cuando T crece. Suponiendo por un momento la hipótesis de Riemann cierta para simplificar la discusión, esto significa que la función $t \mapsto \zeta(1/2+it)$ tiene que oscilar mucho. Como, por otra parte, también se sabe que no está acotada (por ejemplo, gracias a la asintótica

$$\frac{1}{2T}\int_{-T}^{T} |\zeta(\tfrac{1}{2}+it)|^2\, dt \sim \log T$$

cuando $T \to +\infty$ dada por el trabajo de Hardy y Littlewood), concluimos que es prácticamente imposible «adivinar» el valor de $\zeta(1/2+it)$ cuando t es un número real grande «aleatorio», salvo si lo calculamos aproximadamente. La Figura 2 ilustra este punto mediante una representación del módulo $|\zeta(1/2+it)|$ para $t \in [100, 200]$.

Es por ello altamente recomendable considerar las variaciones de los valores de la función zeta de modo *probabilístico* para intentar entender su comportamiento «en promedio». Esto es también muy natural desde el punto de vista de las aplicaciones aritméticas, ya que es bastante raro que un valor individual de la función zeta represente un papel decisivo. Se trata a menudo de cuestiones de integrales que hacen intervenir la función zeta, o de la localización de todo un conjunto de ceros en lugar de uno solo. En muchos casos, se evita así el recurso a la hipótesis de Riemann. El primer ejemplo espectacular es la demostración de Hoheisel de que existe $c < 1$ tal que siempre hay un número primo entre x y $x + x^c$ para cualquier valor lo bastante grande de x. La estrategia más obvia parecería necesitar que todos los ceros tengan

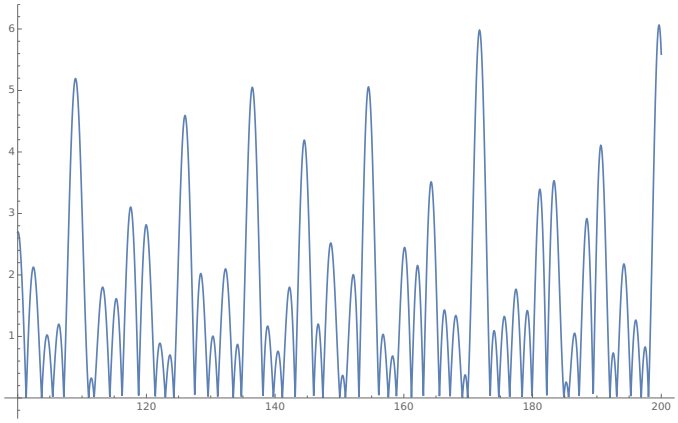

Figura 2: Grafo de $|\zeta(\frac{1}{2}+it)|$ para $100 \leq t \leq 200$.

parte real $\leq c$, pero se entendió que hay sitio para algunas excepciones potenciales, siempre y cuando no sean «demasiadas», véase, por ejemplo, [11, Ch. 10].

El resultado probabilístico más importante sobre la función zeta de Riemann es un teorema debido a Selberg [24] en[6] 1946. Para simplificar, consideraremos solo el módulo de zeta, aunque Selberg demostró un resultado similar (y de la misma importancia) para el argumento.

Teorema 3.1 (Selberg). *Para cada número real $T \geq 1$, consideremos la variable aleatoria sobre el intervalo $[-T, T]$, con la medida de Lebesgue normalizada $\frac{1}{2T}\,dt$, dada por la fórmula*

$$t \mapsto \begin{cases} \log|\zeta(1/2+it)|/\sqrt{\frac{1}{2}\log\log T} & \text{si } \zeta(1/2+it) \neq 0, \\ 0 & \text{si no.} \end{cases}$$

Estas variables aleatorias convergen en ley a la gaussiana estándar cuando $T \to +\infty$.

En términos concretos, el teorema significa que, para dos números reales cualesquiera $a < b$, el siguiente límite existe y toma el valor indicado:

$$\lim_{T\to+\infty} \frac{1}{2T}\,\lambda\left(\left\{t \in [-T,T] \;\middle|\; a < \frac{\log|\zeta(1/2+it)|}{\sqrt{\frac{1}{2}\log\log T}} < b\right\}\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}\,dt \quad (9)$$

(en esta fórmula, λ designa la medida de Lebesgue). Selberg demostró (9) calculando asintóticamente los momentos del logaritmo de la función zeta de Riemann en la recta crítica, es decir, probando que para todo entero $k \geq 0$ se tiene

$$\lim_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T}\left(\frac{\log|\zeta(1/2+it)|}{\sqrt{(\log\log T)/2}}\right)^k dt = \begin{cases} \frac{k!}{2^{k/2}(k/2)!} & \text{si } k \text{ es par,} \\ 0 & \text{si } k \text{ es impar.} \end{cases}$$

[6] No es el primero: lo preceden los trabajos de Bohr, Jessen y otros sobre los valores de la función $\zeta(\sigma+it)$ para $\sigma > 1/2$.

Como es bien conocido, el lado derecho de esta fórmula es igual a $\mathbf{E}(N^k)$ para la variable aleatoria gaussiana estándar N, es decir,

$$\mathbf{E}(N^k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} t^k e^{-t^2/2}\, dt.$$

El Teorema 3.1 resulta entonces del *método de los momentos*. Al demostrarlo, no parece que Selberg tuviera conocimiento de esta interpretación probabilística.

Las pruebas modernas (tales como la de Radziwiłł y Soundararajan [21]) hacen hincapié en la naturaleza probabilística del resultado identificando una aproximación de la función zeta de Riemann con un objeto de naturaleza similar a una gaussiana, como explicamos de forma intuitiva a continuación. Pero debemos enfatizar que este teorema es considerablemente más profundo y difícil de demostrar que, pongamos, el teorema de Erdős-Kac.

El Teorema 3.1 es el enunciado fundamental que justifica el uso de ciertos argumentos heurísticos relativos a la distribución de los valores de la función zeta de Riemann en la recta crítica. Por ejemplo, como el factor de normalización $\sqrt{(\log\log T)/2}$ tiende a infinito con T, vemos que la mayor parte de los valores $\zeta(1/2+it)$ son o bien muy grandes o bien muy pequeños: tomando, por ejemplo, $a = -1/10$ y $b = 1/10$ en (9), se sigue que al menos el 92 % de los números reales $t \in [-T, T]$ cumplen una de las dos desigualdades

$$|\zeta(1/2+it)| \geq \exp\Big(\tfrac{1}{10}\sqrt{\tfrac{1}{2}\log\log T}\Big) \quad \text{o} \quad |\zeta(1/2+it)| \leq \exp\Big(-\tfrac{1}{10}\sqrt{\tfrac{1}{2}\log\log T}\Big).$$

En el primer caso, la función zeta es «muy grande», y, en el segundo, «muy pequeña», puesto que estas cotas también tienden a infinito o a cero cuando $T \to +\infty$.

3.4. Justificación del teorema de Selberg

Podemos dar una explicación heurística del teorema de Selberg que es muy natural desde el punto de vista probabilístico.

Paso 1. Aunque el producto de Euler (5) diverge en la recta $\mathrm{Re}(s) = 1/2$, se podría esperar que, de algún modo, siguiera existiendo una aproximación de la función $\zeta(1/2+it)$ dada por un producto parcial

$$\zeta(1/2+it) \approx \prod_{p \leq P} \big(1 - p^{-1/2-it}\big)^{-1},$$

para $|t| \leq T$ y un valor conveniente de P. Tal aproximación es muy delicada y solo puede ser cierta en un sentido probabilístico, con un valor de P tal que $\log P \sim \log T$.

Paso 2. Usando la serie de Taylor de $\log(1-z)$ alrededor de $z = 0$, se esperaría entonces una aproximación de tipo

$$\log|\zeta(1/2+it)| \approx \mathrm{Re}\Big(\sum_{p \leq P} p^{-1/2-it}\Big),$$

ya que las contribuciones de orden superior (que hacen intervenir p^{-1-it}, $p^{-3/2-it}$, etc.) son despreciables, puesto que las series correspondientes son o bien absolutamente convergentes o bien «casi convergentes».

Paso 3. Se podría entonces usar el siguiente resultado, que proporciona en este caso la conexión crucial entre teoría de números y probabilidad (representa más o menos el mismo papel que el Teorema 2.1 en la prueba del teorema de Erdős-Kac):

Teorema 3.2. *Sea* $\mathbf{T}$ *el producto infinito de copias del círculo unidad en* $\mathbb{C}$ *indexado por los números primos. Para cada número real* $T \geq 1$*, sea* X_T *la variable aleatoria sobre el intervalo* $[-T, T]$*, con la medida de Lebesgue normalizada* $\frac{1}{2T}\lambda$*, y con valores en* $\mathbf{T}$*, dada por la fórmula*

$$X_T(t) = (2^{-it}, 3^{-it}, \ldots) = (p^{-it})_p \in \mathbf{T}.$$

Las variables aleatorias X_T *convergen en ley cuando* $T \to +\infty$ *a una variable aleatoria* $U = (U_p)$ *cuyas componentes* U_p *son todas independientes y están distribuidas uniformemente en el círculo unidad.*

Observación. Otra forma de expresar este teorema consiste en observar que $\mathbf{T}$ es un grupo abeliano compacto, de modo que el límite del enunciado es también la medida de probabilidad de Haar en este grupo.

Se trata de nuevo de un resultado relativamente elemental. En primer lugar, se comprueba que basta con demostrar la fórmula

$$\lim_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} \varphi(X_T(t))\, dt = \mathbf{E}(\varphi(U))$$

cuando φ es una función sobre $\mathbf{T}$ de la forma

$$\varphi((x_p)) = \prod_p x_p^{m_p}$$

para ciertos enteros $m_p \in \mathbb{Z}$, todos nulos salvo un número finito (es un caso del «criterio de Weyl»). Se tiene entonces $\varphi(X_T(t)) = r^{-it}$, con r el número racional

$$r = \prod_p p^{m_p},$$

y una integración directa da

$$\lim_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} \varphi(X_T(t))\, dt = \begin{cases} 1 & \text{si } m_p = 0 \text{ para todo } p, \\ 0 & \text{si no,} \end{cases}$$

puesto que $r = 1$ si y solo si todos los m_p son cero, por la unicidad de la factorización en números primos, que vuelve a ser el ingrediente aritmético fundamental. Otro cálculo simple muestra que la esperanza $\mathbf{E}(\varphi(U))$ viene dada por la misma fórmula.

Paso 4. Basándonos en este resultado, es natural esperar que, probabilísticamente hablando, $\log|\zeta(1/2+it)|$ esté, para todo $|t|\leq T$, «cerca» de $\mathrm{Re}(Y_P)$, donde

$$Y_P=\sum_{p\leq P}\frac{U_p}{\sqrt{p}},$$

y las variables aleatorias (U_p) son, como en el paso 3, independientes y se distribuyen uniformemente en el círculo unidad.

Paso 5. La cuestión del comportamiento de la variable aleatoria Y_P forma parte de la teoría de probabilidad estándar: la teoría de sumas de variables aleatorias independientes, el reino del teorema central del límite. La esperanza de Y_P es cero, puesto que cada U_p tiene esperanza cero, y la varianza $\sigma_P=\mathbf{E}(Y_P^2)$ de Y_P cumple

$$\sigma_P^2=\sum_{p\leq P}\frac{1}{p}$$

ya que las variables U_p son independientes. Recordemos la asintótica

$$\sigma_P^2=\sum_{p\leq P}\frac{1}{p}\sim\log\log P \tag{10}$$

cuando $P\to+\infty$, que era una herramienta importante en la demostración del teorema de Erdős-Kac. Ahora estamos en condiciones de observar que (10) se deduce por sumación por partes del teorema de los números primos[7].

Como la varianza tiende a $+\infty$ cuando $P\to+\infty$, una versión conveniente (por ejemplo, la de Lyapunov) del teorema central del límite implica que Y_P/σ_P converge en ley a una gaussiana compleja estándar. Por tanto, la parte real de Y_P/σ_P converge en ley a su parte real, que es una variable aleatoria gaussiana (real) centrada con varianza $1/2$. Esto significa que $\mathrm{Re}(Y_P/(\frac{1}{\sqrt{2}}\sigma_P))$ converge en ley a la gaussiana estándar, y así se obtiene el teorema de Selberg, siempre suponiendo que se pueda justificar la aproximación del paso anterior.

3.5. Otras conexiones entre la función zeta y la probabilidad

Cerramos esta sección con algunos ejemplos más de los aspectos probabilísticos de la función zeta de Riemann. También nos gustaría indicar que esta función es solo un ejemplo (el más sencillo) de una importante clase de objetos similares conocidos como funciones L, que están asociados a diversos objetos aritméticos (cuerpos de números, variedades algebraicas, formas automorfas, etc.) y cuyas propiedades, ya sean conocidas o conjeturales, codifican algunos de los aspectos más profundos y fascinantes de la teoría de números moderna.

[7] Aunque se trata de un resultado más sencillo, que fue demostrado por Mertens usando métodos elementales mucho antes del teorema de los números primos.

El teorema de Bagchi y el teorema de Voronin

El teorema de Selberg se ocupa del estudio probabilístico de la función zeta de Riemann en la recta crítica $\mathrm{Re}(s) = 1/2$. Es, por supuesto, también interesante saber qué ocurre en otras regiones, y Bagchi demostró un bello resultado en esta dirección (aunque es mucho más sencillo que el teorema de Selberg).

Sea $D \subset \mathbb{C}$ un disco cerrado de centro $3/4$ y de radio $r < 1/4$, de modo que D esté contenido en la banda vertical $\{s \in \mathbb{C} \mid 1/2 < \mathrm{Re}(s) < 1\}$. Consideremos el espacio de Banach $H(D)$ formado por las funciones $f\colon D \to \mathbb{C}$ que son holomorfas en el interior de D, junto con la norma $\|f\| = \sup_{d\in D} |f(z)|$. Para cada $T \geq 1$, la función que asocia a $t \in [-T, T]$ la traslación de la function zeta por t, es decir, la función holomorfa $s \mapsto \zeta(s + it)$ sobre D, es una variable aleatoria sobre $[-T, T]$ con valores en $H(D)$. El teorema de Bagchi afirma que estas variables aleatorias convergen en ley cuando $T \to +\infty$ a la serie de Dirichlet aleatoria

$$\sum_{n\geq 1} \frac{U_n}{n^s} = \prod_p (1 - U_p p^{-s})^{-1},$$

cuyos coeficientes U_n están definidos multiplicativamente a partir de una sucesión (U_p) con las propiedades del Teorema 3.2, a saber

$$U_n = \prod_p U_p^{n_p} \quad \text{para} \quad n = \prod_p p^{n_p}.$$

Además, calculando el soporte de esta serie de Dirichlet aleatoria, Bagchi redemostró el llamado «teorema de universalidad» de Voronin, que afirma que, para toda función $f \in H(D)$ que no se anula en el interior de D y para todo $\varepsilon > 0$, se tiene

$$\liminf_{T\to+\infty} \frac{1}{2T}\, \lambda\big(\{t \in [-T, T] \mid \|\zeta(\cdot + it) - f(\cdot)\| < \varepsilon\}\big) > 0.$$

Conexiones conjeturales con la teoría de matrices aleatorias

Además del modelo probabilístico sugerido por el Teorema 3.1 (una gaussiana centrada con varianza $(\log\log T)/2$ para el módulo de $\zeta(1/2+it)$ cuando $|t| \leq T$), hay un abundante corpus de evidencia (incluyendo evidencias numéricas) que relaciona la distribución y las propiedades de $\zeta(1/2 + it)$ con las matrices aleatorias unitarias de gran tamaño N (es decir, con elementos del grupo unitario $\mathrm{U}_N(\mathbb{C})$ que se distribuyen según la medida de probabilidad de Haar en este grupo compacto), con $N \sim \log T$. De ser cierta, la naturaleza de esta aproximación sigue siendo misteriosa, pero conduce por ejemplo a conjeturas muy precisas sobre los momentos

$$\frac{1}{2T} \int_{-T}^{T} |\zeta(1/2 + it)|^k \, dt,$$

para $k > 0$ real (e incluso para k complejo con $\mathrm{Re}(k) \geq 0$), gracias al trabajo de Keating y Snaith. El seminario Bourbaki de Ph. Michel [19] contiene una excelente presentación de estas relaciones conjeturales.

Conexiones con el caos multiplicativo gaussiano

Algunos de los trabajos más profundos de la teoría probabilística de números en estos últimos años se han dedicado al estudio de aspectos más finos de la distribución de la función zeta de Riemann en la recta crítica. Uno de los objetivos ha sido entender una conjetura de Fedorov, Hiary y Keating sobre la distribución del máximo de $\zeta(1/2+it)$ cuando t varía en un intervalo de longitud 1 (y toma valores elegidos al azar uniformemente en $[-T,T]$ o $[T,2T]$ con $T\to+\infty$). Esto conduce a relaciones con objetos tales como los campos log-correlados, las caminatas aleatorias ramificadas o el caos multiplicativo gaussiano. Invitamos al lector que quiera saber más al respecto a consultar el seminario Bourbaki de Harper [8], que discute los trabajos de Najnudel y Arguin-Belius-Bourgade-Radziwiłł-Soundararajan, o su reciente prepublicación [9] para los últimos desarrollos.

4. Caminos de Kloosterman

Nuestro último ejemplo de teoría probabilística de números es un resultado reciente de W. Sawin y el autor [18] en el que aparecen objetos probabilísticos y aritméticos muy diferentes de los anteriores y cuya prueba reposa sobre la relación fascinante con aspectos profundos de la geometría aritmética, como la forma fuerte de la hipótesis de Riemann sobre cuerpos finitos de Deligne [2] y las posteriores e igualmente notables elaboraciones en los trabajos de Katz. Además, este resultado conduce a imágenes intrigantes como las de las Figuras 3 y 4.

Los objetos aritméticos en cuestión son *sumas exponenciales* finitas: las llamadas *sumas de Kloosterman* $\mathrm{Kl}(a,b;p)$ módulo números primos. Para definirlas, usamos la notación $e(z)=\exp(2i\pi z)$ para un número complejo z. Sea p un número primo y sean a y b enteros coprimos con p. Dado un entero x coprimo con p, designamos por $\bar{x}$ el inverso de x módulo p, es decir, la clase residual módulo p tal que $x\bar{x}\equiv 1$ (mód p). Obsérvese que la cantidad $e(\bar{x}/p)$ está bien definida puesto que, al cambiar de representante de $\bar{x}$ en $\mathbb{Z}$, la exponencial queda multiplicada por un factor $e(k)=1$ para un cierto entero $k\in\mathbb{Z}$. Sea

$$\mathrm{Kl}(a,b;p)=\frac{1}{\sqrt{p}}\sum_{1\le x\le p-1} e\Big(\frac{ax+b\bar{x}}{p}\Big) \tag{11}$$

(enseguida explicaremos la presencia del factor de normalización $1/\sqrt{p}$).

Resulta difícil exagerar la importancia de estas sumas exponenciales en la teoría analítica de números, ya sea al resolver ecuaciones diofánticas, al estudiar los números primos en progresiones aritméticas o en la teoría de formas automorfas (los artículos [10, 14] presentan muchas de sus notables aplicaciones). Un modo simple de motivarlas, hasta cierto punto, es observar que la integral

$$\int_0^{+\infty}\exp(-ax-b/x)\,dx,$$

que es el análogo «continuo» natural de (11), no es sino la función de Bessel, cuya importancia en física y matemática aplicada es de sobra conocida.

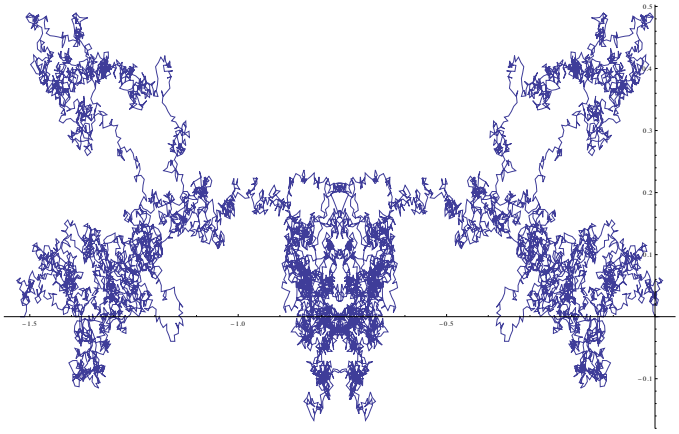

Figura 3: Camino de Kloosterman para $p = 10007$ y $a = b = 1$.

Las sumas de Kloosterman son objetos de apariencia simple: sumas finitas de raíces de la unidad. Sin embargo, como se ve en la Figura 3, el proceso de sumación es extremadamente complejo. Esta figura muestra, en el caso $a = b = 1$ y $p = 10007$, el «camino» tomado al sumar los términos

$$e\Big(\frac{x+\bar{x}}{10007}\Big)$$

(normalizados por $1/\sqrt{10007}$) cuando x va de 1 a 10006. Dicho de otro modo, los vértices del «polígono» son las sumas parciales sucesivas

$$\frac{1}{\sqrt{10007}} \sum_{1\le x\le y} e\Big(\frac{x+\bar{x}}{10007}\Big),$$

para $0 \le y \le 10006$, unidas por segmentos de línea.

Es evidente que este dibujo evoca algún fenómeno aleatorio, y el artículo [18] describe con precisión en qué sentido. Para enunciar su resultado principal, definimos, para un primo p y enteros a y b invertibles módulo p, una función continua

$$\mathcal{K}_p(a,b)\colon [0,1] \to \mathbb{C}$$

con la propiedad siguiente: para todo entero $0 \le y \le p$, se tiene

$$\mathcal{K}_p(a,b)\Big(\frac{y}{p}\Big) = \frac{1}{\sqrt{p}} \sum_{1\le x\le y} e\Big(\frac{ax+b\bar{x}}{p}\Big),$$

y los valores en el intervalo $y/p \le t \le (y+1)/p$ se obtienen por interpolación lineal. Por tanto, la imagen de $\mathcal{K}_p(a,b)$ es un camino del estilo del representado en la Figura 3. Pensaremos en la aplicación $(a,b) \mapsto \mathcal{K}_p(a,b)$ como una variable aleatoria, definida sobre el espacio de probabilidad (finito) $\mathbb{F}_p^\times \times \mathbb{F}_p^\times$ con la medida de probabilidad uniforme y con valores en el espacio de Banach $\mathcal{C}([0,1])$ de las funciones continuas complejas en el intervalo $[0,1]$.

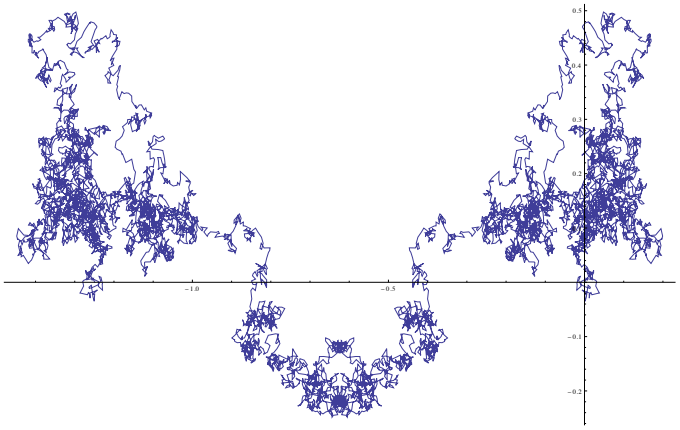

Figura 4: Una muestra de la serie de Fourier aleatoria.

Teorema 4.1 (Kowalski-Sawin). *Sea $(S_h)_{h\in\mathbb{Z}}$ una sucesión de variables aleatorias independientes con valores en $[-2,2]$ que se distribuyen según la medida de Sato-Tate*

$$\frac{1}{\pi}\sqrt{1-\frac{x^2}{4}}\,dx.$$

Se define la serie de Fourier aleatoria

$$X(t) = tS_0 + \sum_{\substack{h\in\mathbb{Z}\\ h\neq 0}} \frac{e(ht)-1}{2i\pi h} S_h, \tag{12}$$

donde la suma se entiende como el límite cuando $H\to+\infty$ de las sumas parciales simétricas sobre $|h|\leq H$. La serie X define una variable aleatoria[8] *con valores en $\mathcal{C}([0,1])$ y las variables aleatorias $\mathcal{K}_p$ convergen en ley a X cuando $p\to+\infty$.*

Para ilustrar el comportamiento de la serie de Fourier aleatoria X, la Figura 4 presenta una aproximación de una muestra de esta serie (obtenida calculando la suma parcial sobre $|h|\leq 10000$, con coeficientes aleatorios que simulan la sucesión (S_h)).

Como antes, intentaremos motivar el resultado. Para ello comenzamos explicando el factor de normalización: según un resultado profundo demostrado por A. Weil en 1948 como consecuencia de la hipótesis de Riemann para las curvas sobre cuerpos finitos, las sumas de Kloosterman cumplen $|\operatorname{Kl}(a,b;p)|\leq 2$ (además, toman valores reales, lo cual es fácil de demostrar calculando su conjugado complejo) y no es posible reemplazar la constante 2 por ninguna función de p que tienda a 0 en la desigualdad. Así, las sumas de Kloosterman también se pueden estudiar probabilísticamente.

Tomemos la suma parcial sobre $1\leq x\leq y$ y apliquemos un principio estándar del análisis armónico, el *método de compleción*. Se trata de desarrollar la función

[8] Es decir, la serie converge casi seguramente uniformemente en $[0,1]$.

característica φ_y del intervalo de sumación en una serie de Fourier discreta módulo p:

$$\varphi_y(x) = \sum_{-p/2<h\le p/2} \alpha_y(h;p) e\Big(\frac{hx}{p}\Big),$$

donde

$$\alpha_y(h;p) = \frac{1}{p} \sum_{1\le x\le y} e\Big(-\frac{hx}{p}\Big),$$

y, después, sustituir esta expresión en la suma parcial para obtener

$$\begin{aligned}\frac{1}{\sqrt{p}} \sum_{1\le x\le y} e\Big(\frac{ax+b\bar{x}}{p}\Big) &= \frac{1}{\sqrt{p}} \sum_{-p/2<h\le p/2} \alpha_y(h;p) \sum_{1\le x\le p-1} e\Big(\frac{(a+h)x+b\bar{x}}{p}\Big) \\ &= \sum_{-p/2<h\le p/2} \alpha_y(h;p)\,\mathrm{Kl}(a+h,b;p). \qquad (13)\end{aligned}$$

Por medio de una simple interpretación como sumas de Riemann se ve fácilmente que $\alpha_y(0;p) = y/p$ y

$$\lim_{p\to+\infty} \alpha_y(h;p) = \int_0^{y/p} e(ht)\,dt = \frac{e(hy/p)-1}{2i\pi h}$$

para $h \neq 0$. Esto indica que el lado derecho de (13) es reminescente del lado derecho de (12), evaluado en $t = y/p$, si se sustituyen las variables aleatorias S_h por las variables aleatorias sobre $\mathbb{F}_p^\times \times \mathbb{F}_p^\times$ dadas por

$$(a,b) \mapsto \mathrm{Kl}(a+h,b;p).$$

Llegado este punto, el Teorema 4.1 no es difícil de entender a partir de un resultado que es esencialmente una consecuencia de los trabajos profundos de Katz [12] sobre la distribución de las sumas de Kloosterman y que enunciamos de forma algo imprecisa como sigue: la familia $(\mathrm{Kl}(a+h,b;p))_h$ «converge en ley» a una familia de variables aleatorias independientes distribuidas a la Sato-Tate. El hecho de que, para h fijo, las variables aleatorias $\mathrm{Kl}(a+h,b;p)$ convergen en ley a la medida de Sato-Tate fue demostrado por Katz en [12], y en otro trabajo [13] desarrolló los principios necesarios para deducir que las diferentes sumas de Kloosterman desplazadas son asintóticamente independientes.

5. Conclusión

Confiamos en que este artículo haya ilustrado las bellas conexiones que existen entre la teoría de números y la probabilidad. Como dijimos al comienzo, se trata de un área en gran expansión, en la que las relaciones no cesan de estrecharse y las interacciones se enriquecen cada vez más. ¡Es de esperar que muchas bellas matemáticas nazcan de esta relación en los años venideros!

A. Nociones de probabilidad

A.1. Espacios de probabilidad, variables aleatorias e independencia

Un *espacio de probabilidad* es una terna $(\Omega, \Sigma, \mathbf{P})$ donde Ω es un conjunto, Σ es una σ-álgebra sobre Ω y $\mathbf{P}$ es una medida de probabilidad sobre Σ (es decir, una medida positiva tal que $\mathbf{P}(\Omega) = 1$).

Dado un espacio topológico T, una variable aleatoria X sobre Ω con valores en T es una aplicación $X\colon \Omega \to T$ que es medible cuando se dota T de la σ-álgebra de Borel: para cualquier abierto U de T, la preimagen $X^{-1}(U)$ pertenece a Σ. La *ley* de X es la medida de probabilidad μ_X sobre T imagen de $\mathbf{P}$ por X. Se suele escribir $\mathbf{P}(X \in U)$ en lugar de $\mu_X(U)$. Si un conjunto U satisface $\mathbf{P}(X \in U) = 1$, se dice que el evento que U representa sucede *casi seguramente*.

Designamos por $\mathbf{E}(X)$ la esperanza y por $\mathbf{V}(X)$ la varianza de una variable aleatoria con valores complejos, dadas, cuando existen, por

$$\mathbf{E}(X) = \int_\Omega X \, d\mathbf{P}, \qquad \mathbf{V}(X) = \mathbf{E}(|X - \mathbf{E}(X)|^2).$$

La naturaleza precisa del «espacio muestral» Ω se deja a menudo sin especificar: lo que importa es que se puedan definir familias convenientes de variables aleatorias con propiedades varias, especialmente la independencia, que recordamos a continuación.

Definición A.1. Dado un conjunto arbitrario (posiblemente infinito) I, se dice que una familia $(X_i)_{i\in I}$ de variables aleatorias con valores en T es independiente si, para todo subconjunto finito $J \subset I$ y para cualquier elección de abiertos $U_j \subset T$ para $j \in J$, se cumple

$$\mathbf{P}(X_j \in U_j \text{ para } j \in J) = \prod_{j\in J} \mathbf{P}(X_j \in U_j).$$

El siguiente teorema básico de existencia basta para muchas aplicaciones de la teoría de la probabilidad: dada una sucesión (μ_N) de medidas de probabilidad en, pongamos, $\mathbb{R}$, existe un espacio de probabilidad $(\Omega, \Sigma, \mathbf{P})$ y una familia (X_N) de variables aleatorias independientes con valores reales sobre Ω de leyes μ_N.

A.2. Variables aleatorias especiales

Sea p un número real tal que $0 \le p \le 1$. Una *variable aleatoria de Bernoulli* con probabilidad de éxito p es una variable aleatoria con valores reales tal que

$$\mathbf{P}(B = 1) = p, \qquad \mathbf{P}(B = 0) = 1 - p$$

(dicho de otro modo, la aplicación $B\colon \Omega \to \{0, 1\}$ es la función característica de un conjunto de medida p).

Se dice que una variable aleatoria con valores reales X es una *variable aleatoria gaussiana* de esperanza $m \in \mathbb{R}$ y de varianza $\sigma^2 > 0$ si

$$\mathbf{P}(X \in U) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_U e^{-(x-m)^2/(2\sigma^2)} \, dx$$

para cualquier abierto U. Si $m = 0$, se dice que X está *centrada*.

A.3. Algunos enunciados útiles

En la Sección 2.1 usamos el siguiente lema:

Lema A.2. *Sea T un espacio de Banach de dimensión finita y sean $(X_N)_{N\geq 1}$ y $(X_{N,M})_{N\geq M\geq 1}$ variables aleatorias con valores en T. Definamos*

$$E_{N,M} = X_N - X_{N,M}$$

y supongamos que se cumplen las siguientes condiciones:

1. *Para cada $M \geq 1$, las variables aleatorias $(X_{N,M})_{N\geq M}$ convergen en ley a una variable aleatoria Y_M.*
2. *Existe una función $f(N, M)$ tal que*

$$\mathbf{E}(\|E_{N,M}\|) \leq f(N, M)$$

y tal que $f(N, M) \to 0$ cuando $M \to +\infty$ uniformemente con respecto a $N \geq M$.

Entonces, las sucesiones (X_N) y (Y_M) convergen en ley cuando $N \to +\infty$ y tienen la misma distribución límite.

Para la demostración, véase [1, Th. 3.2] (que supone que uno sabe de antemano que (Y_M) converge en ley) o [16, Prop. B.4.4].

También nos referimos al teorema de las tres series de Kolmogorov (véase, por ejemplo, [1, Th. 22.8]):

Teorema A.3. *Sea $(X_N)_{N\geq 0}$ una sucesión de variables aleatorias reales* independientes *sobre Ω. Definamos*

$$\widetilde{X}_N = \begin{cases} X_N & \text{si } |X_N| \leq 1, \\ 0 & \text{si no.} \end{cases}$$

La serie $\sum_{N\geq 0} X_N$ converge casi seguramente si y solo si las tres series

$$\sum_{N\geq 0} \mathbf{E}(\widetilde{X}_N), \qquad \sum_{N\geq 0} \mathbf{V}(\widetilde{X}_N^2), \qquad \sum_{N\geq 0} \mathbf{P}(|X_N| > 1)$$

son convergentes.

Referencias

[1] P. Billingsley, *Convergence of probability measures*, 2.ª edición, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, New York, 1999.

[2] P. Deligne, La conjecture de Weil, II, *Inst. Hautes Études Sci. Publ. Math.* **52** (1980), 137–252.

[3] P. Erdős y M. Kac, The Gaussian law of errors in the theory of additive number theoretic functions, *Amer. J. Math.* **62** (1940), 738–742.

[4] K. Ford, The distribution of integers with a divisor in a given interval, *Ann. of Math. (2)* **168** (2008), 367–433.

[5] A. Granville, The anatomy of integers and permutations, 2008, `https://dms.umontreal.ca/~andrew/PDF/Anatomy.pdf`.

[6] A. Granville y J. Granville, *Prime suspects. The anatomy of integers and permutations (illustrated by R. J. Lewis)*, Princeton Univ. Press, NJ, 2019.

[7] A. Harper, Two new proofs of the Erdős-Kac Theorem, with bound on the rate of convergence, by Stein's method for distributional approximations, *Math. Proc. Cambridge Philos. Soc.* **147** (2009), 95–114.

[8] A. Harper, The Riemann zeta function in short intervals, *Séminaire Bourbaki*, 71ème année, Exposé 1159, en prensa, *ArXiv:1904.08204*.

[9] A. Harper, On the partition function of the Riemann zeta function, and the Fyodorov-Hiary-Keating conjecture, *prepublicación*, 2019, *ArXiv:1906.05783*.

[10] D. R. Heath-Brown, Arithmetic applications of Kloosterman sums, *Nieuw Arch. Wiskd. (5)* **1** (2000), 380–384.

[11] H. Iwaniec y E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications **53**, American Mathematical Society, Providence, RI, 2004.

[12] N. M. Katz, *Gauss sums, Kloosterman sums and monodromy groups*, Annals of Mathematics Studies **116**, Princeton University Press, Princeton, NJ, 1988.

[13] N. M. Katz, *Exponential sums and differential equations*, Annals of Mathematics Studies **124**, Princeton University Press, Princeton, NJ, 1990.

[14] E. Kowalski, Poincaré and analytic number theory, *The scientific legacy of Poincaré*, 73–85, Hist. Math. **36**, Amer. Math. Soc., Providence, RI, 2010.

[15] E. Kowalski, Crible en expansion, *Séminaire Bourbaki*, 63ème année, Exposé 1028, *Astérisque* **348** (2012), 17–64.

[16] E. Kowalski, *Arithmetic randonnée: an introduction to probabilistic number theory*, Cambridge Studies in Advanced Mathematics, en prensa.

[17] E. Kowalski y A. Nikeghbali, Mod-Poisson convergence in probability and number theory, *Int. Math. Res. Not. IMNR* **2010**, no. 18, 3549–3587.

[18] E. Kowalski y W. Sawin, Kloosterman paths and the shape of exponential sums, *Compos. Math.* **152** (2016), 1489–1516.

[19] Ph. Michel, Répartition des zéros des fonctions L et matrices aléatoires, *Séminaire Bourbaki*, 53ème année, Exposé 887, *Astérisque* **282** (2002), 211–248.

[20] C. Perret-Gentil, Some recent interactions of probability and number theory, *Eur. Math. Soc. Newsl.* **111** (2019), 13–20.

[21] M. Radziwiłł y K. Soundararajan, Selberg's central limit theorem for $\log|\zeta(1/2+it)|$, *Enseign. Math.* **63** (2017), no. 1–2, 1–19.

[22] B. Riemann, Über die Anzahl der Primzahlen unter einer gegebenen Grösse, 671–680, *Monatsber. Akad. Berlin*, 1859. Reproducido en *Gesammelte Mathe-*

matische Werke, Wissenschaftlicher Nachlass und Nachträge (R. Narasimhan, ed.), 177–185, Springer-Verlag, Berlin, 1990.

[23] I. Schoenberg, Über die asymptotische Verteilung reeller Zahlen mod 1, *Math. Z.* **28** (1928), no. 1, 171–199.

[24] A. Selberg, Contributions to the theory of the Riemann zeta function, *Arch. Math. Naturvid.* **48** (1946), no. 5, 89–155.

[25] E. C. Titchmarsh, *The theory of the Riemann zeta function*, 2.ª edición, Oxford Univ. Press, 1986.

Emmanuel Kowalski, ETH Zürich, D-MATH, Rämistrasse 101, 8092 Zürich, Suiza
Correo electrónico: kowalski@math.ethz.ch
Página web: https://people.math.ethz.ch/~kowalski/