FACULTAD DE MATEMÁTICAS

TRABAJO FIN DE GRADO

Cuando Riemann conoce a Bernoulli

Aplicaciones de la Teoría de la Probabilidad en la Teoría de Números

Realizado por
Juan Manuel Hernández de la Hera

Para la obtención del
Doble Grado en Matemáticas y Estadística

Dirigido por
D. Guillermo Curbera Costello

Realizado en el Departamento de
Análisis Matemático

Convocatoria de junio, curso 2021/22

Resumen

El objetivo de este trabajo es recopilar y mostrar, de la forma más cercana posible, las posibles aplicaciones de la Teoría de la Probabilidad en varias cuestiones concernientes a la Teoría de Números. Conforme el lector pase por las páginas de este documento, se dará cuenta de que estas aplicaciones no sólo no son pocas, sino que son realmente ingeniosas. Este trabajo toma como referencia principal el libro de Emmanuel Kowalski, [19], junto con otros textos explicitados en la bibliografía.

Antes de describir el contenido, no está de más avisar de los prerrequisitos necesarios para afrontar los resultados y sus demostraciones. Es recomendable tener algunas nociones básicas de Teoría de la Probabilidad e Inferencia Estadística, pues en muchos casos se tratarán distribuciones básicas (como la normal o la Poisson) y cuestiones de convergencia de variables aleatorias. Además, por supuesto, de algunas nociones de Teoría Analítica de Números. Si no es el caso, todos los resultados empleados de estas áreas se encuentran en los Anexos correspondientes. Por lo tanto, podemos decir que este trabajo está pensado para cualquier persona que tenga interés en la Teoría de Números, pues aquí se puede ver otra perspectiva de las grandes cuestiones que ésta abarca.

Podemos, pues, empezar con la descripción de los contenidos, lo cual haremos con el propio título, que no es casual. En él hacemos referencia a dos representantes de ambas ramas. Por un lado, Bernhard Riemann, cuya hipótesis es, no solo el mayor problema de la Teoría de Números, sino de las Matemáticas en general. Por otro lado, Jacob Bernoulli, cuya variable homónima es la distibución (junto con la normal) más representativa de la Teoría de la Probabilidad. Por último, la propia estructura del título hace referencia a la obra biográfica sobre Richard Feynman, *Cuando un fotón conoce a un electrón*.

En el primer capítulo encontraremos el Teorema de Erdös-Kac, el cual se puede considerar como el pionero de la Teoría Probabilística de Números. Este resultado establece que el número de factores primos de un número entero n sin contar multiplicidad ($\omega(n)$) tomado aleatoriamente del conjunto $\{1, \dots, N\}$ converge, previa renormalización, en ley a la normal estándar $\mathcal{N}$ cuando $N \to +\infty$. Posteriormente, discutiremos la razón de esta renormalización, y daremos el resultado que motivó este teorema, el Teorema de Hardy-Ramanujan. Finalmente, con ayuda del software R, visualizaremos empíricamente la convergencia en ley de $\omega(n)$, tanto renormalizando como sin renormalizar, además de la distribución de $\varphi(n)/n$, donde $\varphi(n)$ es la función totient de Euler.

En el segundo capítulo se incluyen tres teoremas concernientes a la distribución de la función zeta de Riemann. En primer lugar, nos encontraremos los Teoremas de Bohr-Jessen y Bagchi, los cuales tratan la distribución de la función zeta para los $s \in \mathbb{C}$ tales que $\Re(s) \in (1/2, 1]$. A continuación, tenemos el Teorema de Selberg, el cual trata la convergencia justamente sobre la línea crítica, es decir, los $s \in \mathbb{C}$ tales que $\Re(s) = 1/2$. Finalmente, podemos ver una interesante interpretación probabilística de la Hipótesis de Riemann, propuesta por Arnaud Denjoy.

En el tercer capítulo trataremos el sesgo de Chebyshev, y nuestro objetivo a lo largo del capítulo es definir la distribución de Rubinstein-Sarnak, probar su existencia y dar algunas de sus propiedades, como por ejemplo, su función característica o la probabilidad de que tome valores grandes.

El cuarto capítulo sigue una estructura similar al anterior, pero tratando esta vez sobre las sumas de Kloosterman. En particular, nos interesa estudiar los caminos generados por las sumas parciales de éstas. Para ello, definiremos estas sumas, dando algunas de sus propiedades. Posteriormente, estudiaremos su distribución para acabar dando algunas consecuencias de esta distribución, como el soporte de la distribución o la probabilidad de que las sumas parciales tomen valores grandes.

Finalmente, en el quinto capítulo daremos una introducción somera a varios temas en los que también se pueden enlazar la Teoría de la Probabilidad y la Teoría de Números. Aparecerán la equidistribución módulo 1, los espacios entre primos, algunos resultados relacionados con la Teoría de Ratner, un teorema de distribución para las sumas de Rademacher, una breve introducción a los grafos de Ramanujan y el modelo de Cramér junto con algunas de las fallas que presenta. Este capítulo, y por consiguiente, el trabajo, acaba con la mención de algunas ideas más en las que pueden estar involucradas las herramientas probabilísticas.

Para acabar esta introducción, quisiera expresar mi más sincero deseo de que usted, el lector, disfrute tanto al leer este trabajo como yo al escribirlo.

Palabras clave: Probabilidad, Aritmética, Distribución, Convergencia, Zeta, Sumas, Teoría de Números, Ley, Medida, Número primo.

Abstract

The main objective of this project is to collect results arising from Number Theory, which are shown or proven via probabilistic methods. With this aim in mind, the reader may see how ingenious these applications are.

As a prerequisite, it is recommended to be familiar with basic Probability Theory and Statistical Inference. Also, some concepts from Analytic Number Theory will be used, so some background knowledge is required. However, if this is not the case, every result from those branches will appear in the corresponding annex.

About the content, this memoir is divided into five independent chapters. The first one is about the first results concerning the bridge between Number Theory and Probability Theory. We are talking about Erdös-Kac Theorem, which states that the number of prime divisors, without multiplicity, of an entire number n ($\omega(n)$) chosen randomly from the set $\{1, \dots, N\}$ converges in law, renormalizing it first, to the standard Gaussian random variable $\mathcal{N}$ as $N \to +\infty$. After this, we present the Hardy-Ramanujan Theorem, which inspired the previous one. Finally, we will discuss the empirical results of the convergence in law of $\omega(n)$, with and without renormalization, and also the distribution of $\varphi(n)/n$, where $\varphi(n)$ is Euler totient function.

The second chapter contains three fundamental theorems about the distribution of Riemann Zeta function. Firstly, we will present the Bohr-Jessen and Bagchi Theorems. Those deal with the distribution of the Zeta function for $s \in \mathbb{C}$ such that $\Re(s) \in (1/2, 1]$. Next, there is the Selberg Theorem that states the distribution of the Zeta function just on the critical line, i.e., those $s \in \mathbb{C}$ such that $\Re(s) = 1/2$. We finish this chapter with a probabilistic interpretation of the Riemann Hypothesis, by Arnaud Denjoy.

The third chapter is about the Chebyshev Bias, our main goal is to define the Rubinstein-Sarnak distribution, prove its existence and give some properties derived from the distribution result.

The fourth chapter follows the same structure as the third one, being the main object of study in this case the Kloosterman sums. In particular, we are interested in the paths generated by partial Kloostermann sums. For that purpose, we will define those sums and give some of their properties. Then we will study their distribution, and some consequences of it.

Finally, in the last chapter we will give a brief introduction to some other topics where probability has an important role. We will see equidistribution module 1, prime gaps, some results related to Ratner Theory, the distribution of Rademacher Sums, Ramanujan Graphs and Cramér model for primes and some related problems. This chapter ends with a petty list of ideas where probabilistic tools may be helpful.

Keywords: Probability, Arithmetic, Distribution, Convergence, Zeta, Sums, Number Theory, Law, Measure, Prime number.

Agradecimientos

Quiero agradecer a don Guillermo Curbera Costello, tutor de este trabajo, por su inestimable ayuda en la correcta realización de este proyecto.

Quiero agradecer a don José Antonio Prado-Bassas por enseñarme a escribir textos con LaTeX.

Quiero agradecer a don Jorge Francisco Galán Vioque por su permiso y ayuda para el uso del supercomputador del IMUS para la realización de gráficos, sin lo cual no hubiese existido la sección 1.6.

Quiero agradecer a la Biblioteca de la Facultad de Matemáticas de la Universidad de Sevilla por poner a mi disposición los libros, artículos y textos necesarios para poder realizar este trabajo.

Quiero agradecer a mi familia, en especial a mis padres, mi hermano y mi abuela Carmen, por estar siempre ahí apoyándome en todos los proyectos que he querido hacer, y un recuerdo muy especial a mi abuelo Manuel, que allá donde esté, sé que estará orgulloso y lo estará disfrutando.

Quiero agradecer a mis compañeros y compañeras, cuyo apoyo ha sido importante durante estos cinco años como universitario.

Índice general

Índice de figuras

Índice de extractos de código

Dedicado a mi abuelo Manuel y mi abuela Carmen.
¡Abuelos, lo conseguí!

"Para aquellos que no conocen las matemáticas, es difícil sentir la belleza de la naturaleza. Si quieres apreciarla, es necesario aprender el lenguaje en el que habla."
-Richard Feynman

1. Relación entre la Probabilidad y la Teoría (Analítica) de Números

> Puede que Dios no juegue a los dados con el universo, pero algo extraño ocurre con los números primos.
>
> *Paul Erdös*

1.1. Introducción

Este capítulo pretende ser una bienvenida para el lector a la Teoría Probabilística de Números, la rama de las Matemáticas que aborda las cuestiones de la Teoría de Números desde una perspectiva probabilística, es decir, empleando herramientas de la Teoría de la Probabilidad.

Para ello, presentamos uno de los resultados más importantes de la Teoría Probabilística de Números, el Teorema de Erdös-Kac, el cual afirma la normalidad asintótica de un estadístico asociado a la función aritmética $\omega(n)$, cuya forma será también razonada en esta sección. Además, se expondrá otro resultado igualmente importante, el Teorema de Hardy-Ramanujan, pues es considerado como precursor del Teorema de Erdös-Kac. Por último, nos ayudaremos del software estadístico R para discutir, con los gráficos presentes, las cuestiones distribucionales de $\omega(n)$ y de la función totient de Euler normalizada.

Es por todo esto que se recomienda al lector, en caso de no estar familiarizado con la Teoría de la Probabilidad o la Teoría Analítica de Números, visite los Apéndices B y C en caso de cualquier duda concerniente a estos temas.

1.2. Resultados previos

Para desarrollar las ideas y teoremas de este capítulo, precisamos de los siguientes resultados.

Teorema 1.2.1. *Para $N \geq 1$, sea $\Omega_N = \{1, \ldots, N\}$, con la medida de probabilidad uniforme $\mathbb{P}_N$. Fijado un entero $q \geq 1$, se denota $\pi_q : \mathbb{Z} \to \mathbb{Z}/q\mathbb{Z}$ a la aplicación reducción módulo q. Sean X_N las variables aleatorias dadas por*

$$X_N(n) := \pi_q(n), \quad n \in \Omega_N.$$

Se tiene que, con $N \to +\infty$, las variables aleatorias X_N convergen en ley a la medida

de probabilidad uniforme μ_q en $\mathbb{Z}/q\mathbb{Z}$. De hecho, para cualquier función

$$f : \mathbb{Z}/q\mathbb{Z} \to \mathbb{C}$$

se cumple

$$|\mathbb{E}(f(X_N)) - \mathbb{E}(f)| \leq \frac{2}{N}\|f\|_1, \tag{1.1}$$

donde

$$\|f\|_1 = \sum_{a \in \mathbb{Z}/q\mathbb{Z}} |f(a)|.$$

Demostración. Basta con probar (1.1), pues este hecho implicaría la convergencia en ley cuando $N \to +\infty$. Por definición, tenemos:

$$\mathbb{E}(f(X_N)) = \frac{1}{N} \sum_{1 \leq n \leq N} f(\pi_q(n))$$

y

$$\mathbb{E}(f) = \frac{1}{q} \sum_{a \in \mathbb{Z}/q\mathbb{Z}} f(a).$$

La idea es que, entre los enteros $1 \leq n \leq N$, aproximadamente N/q están en cualquier clase residuo a (módulo q), es decir, cada clase residuo a (módulo q) contiene, aproximadamente, N/q de los números enteros anteriormente mencionados. Usando esta aproximación en la primera fórmula, obtenemos la segunda. Veámoslo. Uniendo los enteros de la suma según su clase residuo módulo q correspondiente, esto resulta en

$$\frac{1}{N} \sum_{1 \leq n \leq N} f(\pi_q(n)) = \sum_{a \in \mathbb{Z}/q\mathbb{Z}} f(a) \times \frac{1}{N} \sum_{\substack{1 \leq n \leq N \\ n \equiv a \pmod{q}}} 1$$

La primera suma cuenta, para cada a, la cantidad de enteros $1 \leq n \leq N$ tales que el resto resultante al dividir por q es a. Es por ello que estos enteros n se pueden expresar de la forma $n = mq + a$, para algún $m \in \mathbb{Z}$, siendo a otro entero. Por lo tanto, es suficiente con contar esos enteros $m \in \mathbb{Z}$ que cumplan $1 \leq mq + a \leq N$, o equivalentemente

$$\frac{1-a}{q} \leq m \leq \frac{N-a}{q}$$

Luego, simplemente, estamos contando enteros que pertenecen a un intervalo. Cabe destacar que los extremos de dicho intervalo no son necesariamente (por construcción) enteros. La longitud del intervalo es

$$\frac{N-a}{q} - \frac{1-a}{q} = \frac{N-1}{q}.$$

En general, para cualquier intervalo $[\alpha, \beta]$ con $\alpha \leq \beta$ el número $N_{\alpha,\beta}$, el cual denota el número de enteros en el intervalo $[\alpha, \beta]$, cumple

$$\beta - \alpha - 1 \leq N_{\alpha,\beta} \leq \beta - \alpha + 1.$$

Entonces, el número N_a de valores de m satisface

$$\frac{N-1}{q} - 1 \leq N_a \leq \frac{N-1}{q} + 1. \tag{1.2}$$

Por lo tanto,

$$\left| N_a - \frac{N}{q} \right| \leq 1 + \frac{1}{q}.$$

Sumando en a (en $\mathbb{Z}/q\mathbb{Z}$), podemos deducir que

$$\begin{aligned} \left| \frac{1}{N} \sum_{1 \leq n \leq N} f(\pi_q(n)) - \frac{1}{q} \sum_{a \in \mathbb{Z}/q\mathbb{Z}} f(a) \right| &= \left| \sum_{a \in \mathbb{Z}/q\mathbb{Z}} f(a) \left(\frac{N_a}{N} - \frac{1}{q} \right) \right| \\ &\leq \frac{1+q^{-1}}{N} \sum_{a \in \mathbb{Z}/q\mathbb{Z}} |f(a)| \leq \frac{2}{N} \|f\|_1. \end{aligned}$$

□

Proposición 1.2.2. *Para $N \geq 1$, sea $\Omega_N = \{1, \ldots, N\}$ con la medida de probabilidad uniforme $\mathbb{P}_N$. Sea S un conjunto finito de enteros coprimos por parejas. Conforme $N \to +\infty$, el vector $(\pi_q)_{q \in S}$ visto como vector aleatorio en Ω_N con valores en*

$$X_S = \prod_{q \in S} \mathbb{Z}/q\mathbb{Z},$$

converge en ley a un vector de variables aleatorias independientes y uniformes. De hecho, para cualquier función $f : X_S \to \mathbb{C}$ se tiene

$$\left| \mathbb{E}(f((\pi_q)_{q \in S})) - \mathbb{E}(f) \right| \leq \frac{2}{N} \|f\|_1. \tag{1.3}$$

Demostración. Sea r el producto de los elementos de S. Entonces, por el Teorema Chino del Resto, sabemos que existe un isomorfismo de anillos $f \colon X_S \to \mathbb{Z}/r\mathbb{Z}$ tal que la medida de probabilidad uniforme μ_r en el lado derecho corresponde con el producto de medidas uniformes en X_S. Es por ello que f se puede identificar con una función $g : \mathbb{Z}/r\mathbb{Z} \to \mathbb{C}$, y su esperanza, a su vez, con la esperanza de g respecto de μ_r. Por el Teorema 1.2.1, tenemos

$$\left| \mathbb{E}(f((\pi_q)_{q \in S})) - \mathbb{E}(f) \right| = \left| \mathbb{E}(g(\pi_r)) - \mathbb{E}(g) \right| \leq \frac{2}{N} \|g\|_1.$$

Lo cual concluye la prueba, pues f y g tienen la misma norma ℓ_1.

□

El siguiente resultado nos será de gran utilidad en la última sección del capítulo, pues nos servirá para discutir la convergencia en ley de la función totient de Euler. Para ello, empecemos definiendo esta función.

Definición 1.2.3. *Para $n \geq 1$ se define la función totient de Euler (también llamada función indicatriz) $\varphi(n)$ como el número de enteros positivos menores que n que son primos con n, es decir*

$$\varphi(n) := \sum_{\substack{k=1 \\ (n,k)=1}}^{n} 1.$$

Esta función es multiplicativa y se puede deducir la siguiente expresión de la función

$$\frac{\varphi(n)}{n} = \prod_{p|n} \left(1 - \frac{1}{p}\right),$$

con p primo, para todo $n \geq 1$. Para ver una demostración de ambos hechos, ver la Proposición C.1.7.

Definimos ahora, la siguiente familia de variables aleatorias. Sea $\Omega_N = \{1, \ldots, N\}$ con la medida de probabilidad uniforme. Consideramos las variables aleatorias

$$\mathbf{F}_N(n) := \frac{\varphi(n)}{n}, \ n \ \in \Omega_N.$$

El objetivo es doble, pues buscamos probar que la sucesión $(\mathbf{F}_N)_{N\geq 1}$ converge en ley, y además, encontrar la distribución a la que converge. Para ello, nos ayudaremos de la sucesion de variables aleatorias $(B_p)_p$ Bernoulli independientes indexadas por los números primos, definidas de la siguiente forma, para p un número primo,

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

Proposición 1.2.4. *La sucesión de variables aleatorias $(\mathbf{F}_N)$ converge en ley a la variable aleatoria $\mathbf{F}$ dada por*

$$\mathbf{F} := \prod_{p} \left(1 - \frac{B_p}{p}\right),$$

donde el producto recorre todos los primos y converge casi seguro.

Demostración. Para $M \geq 1$ denotamos por $\mathrm{F}_{N,M}$ a la variable aleatoria en Ω_N definida por

$$\mathrm{F}_{N,M}(n) := \prod_{\substack{p|n \\ p\leq M}} \left(1 - \frac{1}{p}\right).$$

Es natural pensar en esta variable aleatoria como una aproximación de $\mathbf{F}_N$. Por otra parte, fijando M, tendríamos un producto finito, que es mucho más fácil de manejar. Usaremos un lema de "perturbación" para probar la convergencia en ley de la sucesión $(\mathbf{F}_N)_{N\geq 1}$ entendiendo, para ello, el comportamiento de $\mathrm{F}_{N,M}$. El lema que usaremos es la Proposición B.3.2.

En primer lugar, fijamos $M \geq 1$. Dado que el producto anterior solo recorre los primos $p \leq M$, tenemos, por la Proposición 1.2.2, que las variables aleatorias $\mathrm{F}_{N,M}$ convergen en ley con $N \to +\infty$ a la variable aleatoria

$$\mathrm{F}_M = \prod_{p \leq M} \left(1 - \frac{B_p}{p}\right).$$

La primera suposición de la Proposición B.3.2 se satisface. Tenemos que verificar la segunda suposición, que concierne a la aproximación de $\mathbf{F}_N$ por $\mathrm{F}_{N,M}$ en media.

Escribimos $\mathrm{E}_{N,M} = \mathbf{F}_N - \mathrm{F}_{N,M}$. La esperanza de $|\mathrm{E}_{N,M}|$ viene dada por

$$\mathbb{E}_N(|\mathrm{E}_{N,M}|) = \frac{1}{N} \sum_{n \leq N} \left| \prod_{p|n} \left(1 - \frac{1}{p}\right) - \prod_{\substack{p|n \\ p \leq M}} \left(1 - \frac{1}{p}\right) \right| \leq \frac{1}{N} \sum_{n \leq N} \left| \prod_{\substack{p|n \\ p > M}} \left(1 - \frac{1}{p}\right) - 1 \right|.$$

Para cierto n, podemos ver que la cantidad

$$\prod_{\substack{p|n \\ p > M}} \left(1 - \frac{1}{p}\right) - 1$$

está acotada por la suma de los inversos de los enteros $d \geq 2$ que son libres de cuadrados (ver Definición A.1.1), dividen a n y tienen todos sus factores primos mayores que M. Sea D_n el conjunto formado por esos enteros. En particular, siempre se tiene que $\mathrm{M} < \mathrm{d} \leq \mathrm{N}$ si $\mathrm{n} \in D_n$. Entonces

$$\mathbb{E}_N(|\mathrm{E}_{N,M}|) \leq \frac{1}{N} \sum_{n \leq N} \sum_{d \in D_n} \frac{1}{d} \leq \sum_{M < d \leq N} \frac{1}{d} \times \frac{1}{N} \sum_{\substack{n \leq N \\ n \equiv 0 (mod\ d)}} 1 \leq \sum_{M < d \leq N} \frac{1}{d^2} \leq \frac{1}{M},$$

para todo $N \geq M$. La segunda suposición de la Proposición B.3.2 se deduce de manera inmediata, y concluimos que $(\mathbf{F}_N)_{N \geq 1}$ converge en ley, y su que su límite es el límite en ley $\mathbf{F}$ de las variables aleatorias F_M con $M \to +\infty$. Para concluir la prueba, basta con comprobar que el producto aleatorio

$$\prod_p \left(1 - \frac{B_p}{p}\right), \tag{1.4}$$

que recorre los números primos, converge casi seguro, y tiene la misma distribución que $\mathbf{F}$. La convergencia casi segura se tiene por el Teorema B.3.8, aplicado al logaritmo del producto anterior, que es la suma

$$\sum_p Y_p, \qquad Y_p := \log\left(1 - \frac{B_p}{p}\right),$$

de variables aleatorias independientes. Es evidente que $Y_p \leq 0$ y solo toma los valores 0 (con probabilidad $1 - \frac{1}{p}$) y $\log(1 - \frac{1}{p})$ (con probabilidad $\frac{1}{p}$), por lo que

$$\mathbb{E}(Y_p) = \frac{1}{p}\log\left(1 - \frac{1}{p}\right) \sim -\frac{1}{p^2},$$
$$\mathbb{V}(Y_p) = \mathbb{E}(Y_p^2) - \mathbb{E}(Y_p)^2 = \frac{1}{p}\log\left(1 - \frac{1}{p}\right)^2 - \frac{1}{p^2}\log\left(1 - \frac{1}{p}\right)^2 \ll \frac{1}{p^3},$$

que implica, por el Teorema B.3.1, que la serie aleatoria $\sum_p Y_p$ converge casi seguro, y por lo tanto también lo hace su exponencial, que es el producto (1.4). Ahora, con esta convergencia casi segura, es inmediato que la distribución del producto aleatorio es también la distribución de $\mathbf{F}$.

□

1.3. Relación natural. Teorema de Erdös-Kac

Mostramos, ahora, un resultado clave de este capítulo, el Teorema de Erdös-Kac, el resultado más conocido de la Teoría Probabilística de Números. Presentamos este teorema en dos formulaciones equivalentes, siendo la primera la formulación original dada en el artículo de Paul Erdös y Mark Kac [7], mientras que en la segunda se utilizan términos más próximos a la Teoría de la Probabilidad. Veámoslos.

Teorema 1.3.1 (Erdös-Kac)**.** *Para todo $n \in \mathbb{N}$, sea $\omega(n)$ el número de divisores primos de n, sin contar multiplicidad. Entonces, para cualesquiera $a, b \in \mathbb{R}$, se cumple*

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

Teorema 1.3.2 (Erdös-Kac, reformulación)**.** *Para $N \geq 1$, sea $\Omega_N = \{1, \ldots, N\}$ con la medida de probabilidad uniforme $\mathbb{P}_N$. Sea $\mathbf{X}_N$ la variable aleatoria*

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

definida en Ω_N para todo $N \geq 3$. Entonces $\{\mathbf{X}_N\}_{N\geq 3}$ converge en ley a una variable aleatoria normal estándar, es decir, a una variable aleatoria normal de media 0 y varianza 1.

Observación 1.3.1. Este resultado recuerda, como se verá a continuación, al Teorema Central del Límite. Veamos cómo. En primer lugar, sabemos que la función $\omega(n)$ es aditiva. Reescribámosla como:

$$\omega(n) = \sum_p \mathbf{B}_p(n)$$

donde $\mathbf{B}_p(n)$ son variables aleatorias Bernoulli definidas sobre Ω_N como las funciones características del evento $p|n$. Usando la Proposición 1.2.2, la conjetura natural probabilística para el límite anterior sería (si existiera) la serie

$$\sum_p \mathbf{B}_p,$$

donde las $\mathbf{B}_p$ son variables aleatorias Bernoulli independientes. Pero esta serie diverge casi seguro. De hecho:

$$\sum_p \mathbb{E}(\mathbf{B}_p) = \sum_p \frac{1}{p}.$$

(a) Paul Erdös

(b) Mark Kac

Figura 1.1: Los dos matemáticos que propusieron el Teorema del que trata esta sección.

Esta serie es divergente por la estimación dada por Mertens (Teorema C.2.3), la cual es

$$\sum_{p \leq N} \frac{1}{p} = \log\log \mathrm{N} + \mathrm{O}(1),$$

lo cual se cumple para todo $N \geq 3$, por lo que la divergencia se deduce del Teorema B.3.1, o bien como aplicación del Lema B.3.3.

Se puede afinar la fórmula de ω, viendo que para cada $n \in \Omega_N$, n no tiene divisores primos mayores que N, por lo que se tiene

$$\omega(n) = \sum_{p \leq N} \mathbf{B}_p(n),$$

por lo que se puede esperar que la distribución probabilística de ω sea parecida a

$$\sum_{p \leq N} \mathbf{B}_p. \tag{1.5}$$

THE GAUSSIAN LAW OF ERRORS IN THE THEORY OF ADDITIVE NUMBER THEORETIC FUNCTIONS.* [1]

By P. Erdös and M. Kac.

The present paper concerns itself with the applications of statistical methods to some number-theoretic problems. Recent investigations of Erdös and Wintner [2] have shown the importance of the notion of statistical independence in number theory; the purpose of this paper is to emphasize this fact once again.

It may be mentioned here that we get as a particular case of our main theorem the following result:

If $\nu(m)$ denotes the number of prime divisors of m, and K_n the number of those integers from 1 up to n for which $\nu(m) < \lg\lg n + \omega\sqrt{2\lg\lg n}$ (ω an arbitrary real number), then

$$\lim_{n\to\infty} \frac{K_n}{n} = \pi^{-\frac{1}{2}} \int_{-\infty}^{\omega} \exp(-u^2)\,du.$$

This theorem refines some known results of Hardy, Ramanujan [3] and Erdös.[4]

1. In what follows p will denote a prime and ω will denote a real number.

Let $f(m)$ be an additive number-theoretic function, so that $f(mn) = f(m) + f(n)$ if $(m, n) = 1$. Suppose that $f(p^\alpha) = f(p)$ and $|f(p)| \leqq 1$. Obviously

$$f(m) = \sum_{p|m} f(p).$$

Furthermore put $\sum_{p<n} p^{-1}f(p) = A_n$ and $(\sum_{p<n} p^{-1}f^2(p))^{\frac{1}{2}} = B_n$. Then our main theorem may be stated as follows:

---

* Received December 7, 1939.

[1] A preliminary account appeared in the *Proceedings of the National Academy*, vol. 25 (1939), pp. 206-207.

[2] P. Erdös and A. Wintner, "Additive arithmetic functions and statistical independence," *American Journal of Mathematics*, vol. 61 (1939), pp. 713-722.

[3] Srinivasa Ramanujan, *Collected Papers* (1927), pp. 262-275.

[4] P. Erdös, "Note on the number of prime divisors of integers," *Journal of the London Mathematical Society*, vol. 12 (1937), pp. 308-314.

Figura 1.2: Primera página del artículo original de Erdös y Kac, donde presentaron el resultado homónimo.

Ésta última es una suma de variables aleatorias independientes (aunque no idénticamente distribuidas) , cuyo comportamiento asintótico está bien definido. De hecho, una aplicación sencilla del Teorema Central del Límite nos permite deducir que la familia de variables aleatorias

$$\frac{\sum_{p\leq N}\mathbf{B}_p - \sum_{p\leq N} p^{-1}}{\sqrt{\sum_{p\leq N} p^{-1}(1-p^{-1})}} \tag{1.6}$$

converge en ley a la variable aleatoria normal estándar.

Demostración. Probaremos la convergencia en ley mediante el Método de los momentos B.3.4. Para ello, procederemos de la siguiente forma:

- Veamos que, usando el Teorema 1.2.1, para todo $k \geq 0$, se cumple

  $$\mathbb{E}_N[\mathbf{X}_N^k] = \mathbb{E}[\mathbf{X}_N^k] + o(1),$$

  donde $(\mathbf{X}_N)$ es la misma variable aleatoria renormalizada descrita anteriormente, es decir

  $$\mathbf{X}_N = \frac{\mathbf{Z}_N - \mathbb{E}[\mathbf{Z}_N]}{\sqrt{\mathbb{V}[\mathbf{Z}_N]}},$$

  donde

  $$\mathbf{Z}_N = \sum_{p\leq N} \mathbf{B}_p. \tag{1.7}$$

- Como mencionamos anteriormente, aplicamos el Teorema Central del Límite a la sucesión $(\mathbf{X}_N)$, viendo así que converge en ley a una variable aleatoria normal estándar $\mathcal{N}$.

- Se sigue que

  $$\lim_{N\to+\infty} \mathbb{E}[\mathbf{X}_N^k] = \mathbb{E}[\mathcal{N}^k]$$

  y por el Método de los momentos B.3.4, se concluye que $\mathbf{X}_N$ converge en ley a $\mathcal{N}$.

Paso 1.- Buscamos simplificar la suma en (1.7) para controlar los términos de error para el siguiente paso. Consideremos las variables aleatorias $\mathbf{B}_p$ en Ω_N definidas anteriormente, es decir, $\mathbf{B}_p(n) = 1$ si $p|n$ y $\mathbf{B}_p(n) = 0$ en otro caso. Sea

$$\sigma_N = \sum_{p\leq N} \frac{1}{p}.$$

Recordemos que $\sigma_N \to +\infty$ conforme $N \to +\infty$. Entonces definimos

$$Q = N^{1/(\log\log N)^{\frac{1}{3}}}, \tag{1.8}$$

y

$$\tilde{\omega}(n) = \sum_{\substack{p|n \\ p\leq Q}} 1 = \sum_{p\leq Q} \mathbf{B}_p(n), \qquad \tilde{\omega_0}(n) = \sum_{p\leq Q} \left(\mathbf{B}_p(n) - \frac{1}{p}\right).$$

Ambas vistas como variables aleatorias sobre Ω_N. En primer lugar, se tiene que

$$\tilde{\omega}(n) \leq \omega(n) \leq \tilde{\omega}(n) + (\log\log N)^{1/3}.$$

Esto se debe a que, si $\alpha > 0$ y si $p_1, \ldots, p_m$ son primos $\geq N^\alpha$ dividiendo a $n \leq N$, entonces se cumple

$$N^{m\alpha} \leq p_1 \cdots p_m \leq N,$$

y por lo tanto $m \leq \alpha^{-1}$.

En segundo lugar, para cualquier $N \geq 1$ y cualquier $n \in \Omega_N$ tenemos, por definición de σ_N la identidad

$$\tilde{\omega_0}(n) = \tilde{\omega}(n) - \sum_{p\leq Q} \frac{1}{p} = \omega(n) - \sigma_N + \mathrm{O}(\log\log N)^{1/3}. \tag{1.9}$$

Por la fórmula de Mertens (Teorema C.2.3)

$$\sum_{p\leq x} \frac{1}{p} = \log\log x + \mathrm{O}(1)$$

y la definición de σ_N se ve que

$$\sum_{p\leq Q} \frac{1}{p} = \sum_{p\leq N} \frac{1}{p} + \mathrm{O}(\log\log\log N) = \sigma_N + \mathrm{O}(\log\log\log N).$$

Ahora, definimos

$$\tilde{\mathbf{X}}_N(n) = \frac{\tilde{\omega_0}(n)}{\sqrt{\sigma_N}}$$

como variables aleatorias sobre Ω_N. Probaremos que $\tilde{\mathbf{X}}_N$ converge en ley a $\mathcal{N}$. Usando el Lema B.3.6 en (1.9), obtenemos que las variables aleatorias

$$n \mapsto \frac{\omega(n) - \sigma_N}{\sqrt{\sigma_N}}$$

convergen en ley a $\mathcal{N}$. Finalmente, aplicando el mismo Lema una vez más usando la fórmula de Mertens (Teorema C.2.3) obtenemos el Teorema de Erdös-Kac.

Falta por probar la convergencia de $\tilde{\mathbf{X}}_N$. Para ello, fijamos un entero $k \geq 0$. El objetivo es probar el límite

$$\mathbb{E}_N[\tilde{\mathbf{X}}_N^k] \longrightarrow \mathbb{E}[\mathcal{N}^k] \tag{1.10}$$

con $N \to +\infty$. Una vez se tenga probado para todo $k \geq 0$, entonces, por el Método de los momentos B.3.4, se tendría la convergencia en ley de $(\mathbf{X}_N)$ a la normal estándar $\mathcal{N}$.

Paso 2.- Empezamos la prueba de (1.10). Para ello, usamos la definición de $\tilde{\omega}_0(n)$ y expandimos la k-ésima potencia en $\mathbb{E}_N[\tilde{\mathbf{X}}_N^k]$ para llegar a

$$\mathbb{E}_N[\tilde{\mathbf{X}}_N^k] = \frac{1}{\sigma_N^{k/2}} \sum_{p_1 \leq Q} \cdots \sum_{p_k \leq Q} \mathbb{E}_N \left(\left(\mathbf{B}_{p_1} - \frac{1}{p_1} \right) \cdots \left(\mathbf{B}_{p_k} - \frac{1}{p_k} \right) \right).$$

Para simplificar, hemos omitido N de los subíndices de las variables aleatorias $\mathbf{B}_{p_i}$. Lo importante es que la variable aleatoria

$$\left(\mathbf{B}_{p_1} - \frac{1}{p_1} \right) \cdots \left(\mathbf{B}_{p_k} - \frac{1}{p_k} \right) \tag{1.11}$$

se puede expresar como $f(\pi_q)$ para algún módulo $q \geq 1$ y alguna $f : \mathbb{Z}/q\mathbb{Z} \to \mathbb{C}$, por lo que el Teorema 1.2.1 puede ser aplicado a cada sumando. Para ser más preciso, el valor en $n \in \Omega_N$ de la variable (1.11) solo depende de la clase residuo x de n en $\mathbb{Z}/q\mathbb{Z}$ donde q es el mínimo común múltiplo de $p_1, \ldots, p_k$. De hecho, este valor es igual a $f(x)$, donde

$$f(x) = \left(\delta_{p_1}(x) - \frac{1}{p_1} \right) \cdots \left(\delta_{p_k}(x) - \frac{1}{p_k} \right),$$

donde δ_{p_i} denotan las funciones características de las clases residuos módulo q que son 0 módulo p_i. Es evidente que $|f(x)| < 1$, pues es el producto de términos menores o iguales a 1 con alguno menor estrictamente que 1, por lo que se tiene la cota

$$\|f\|_1 \leq q.$$

De esto, por el Teorema 1.2.1, obtenemos

$$\left| \mathbb{E}_N \left(\left(\mathbf{B}_{p_1} - \frac{1}{p_1} \right) \cdots \left(\mathbf{B}_{p_k} - \frac{1}{p_k} \right) \right) - \mathbb{E}(f) \right| \leq \frac{2q}{N} \leq \frac{2Q^k}{N}.$$

Pero, por definición de f, también se tiene

$$\mathbb{E}(f) = \mathbb{E} \left(\left(B_{p_1} - \frac{1}{p_1} \right) \cdots \left(B_{p_k} - \frac{1}{p_k} \right) \right),$$

donde los (B_p) forman una sucesión de variables aleatorias Bernoulli independientes con $\mathbb{P}(B_p = 1) = 1/p$ (B_p para p divide a q se realiza concretamente como la función característica δ_p en $\mathbb{Z}/q\mathbb{Z}$ con la medida de probabilidad uniforme). Por lo tanto, se deduce

$$\begin{aligned}
\mathbb{E}_N[\tilde{\mathbf{X}}_N^k] &= \frac{1}{\sigma_N^{k/2}} \sum_{p_1 \leq Q} \cdots \sum_{p_k \leq Q} \left\{ \mathbb{E} \left(\left(B_{p_1} - \frac{1}{p_1} \right) \cdots \left(B_{p_k} - \frac{1}{p_k} \right) \right) + \mathrm{O}(Q^k N^{-1}) \right\} \\
&= \left(\frac{\tau_N}{\sigma_N} \right)^{k/2} \mathbb{E}[X_N^k] + \mathrm{O}(Q^{2k} N^{-1}) \\
&= \left(\frac{\tau_N}{\sigma_N} \right)^{k/2} \mathbb{E}[X_N^k] + \mathrm{o}(1),
\end{aligned}$$

este último paso por la elección tomada de (1.8), donde

$$X_N = \frac{1}{\sqrt{\tau_N}} \sum_{p \leq Q} \left(B_p - \frac{1}{p} \right)$$

y

$$\tau_N = \sum_{p \leq Q} \frac{1}{p}\left(1 - \frac{1}{p}\right) = \sum_{p \leq Q} \mathbb{V}(B_p).$$

Paso 3.- Concluimos la prueba. Observamos que la versión del Teorema central del límite en el Teorema B.3.7 es aplicable a las variables aleatorias B_p, y de hecho implica la convergencia en ley a $\mathcal{N}$. Pero es más, la sucesión X_N satisface la hipótesis de integrabilidad uniforme en el recíproco del Método de los momentos (ver Teorema B.3.5) y, en particular, se tiene

$$\mathbb{E}[X_N^k] \longrightarrow \mathbb{E}[\mathcal{N}^k].$$

Como $\tau_N \sim \sigma_N$ por la fórmula de Mertens (Teorema C.2.3), se deduce que $\mathbb{E}_N[\tilde{\mathbf{X}}_N^k]$ también converge a $\mathbb{E}[\mathcal{N}^k]$, lo cual era nuestro objetivo desde el paso 1 (1.10). □

Observación 1.3.2. La prueba original del teorema se debe a Erdös y Kac en 1939, publicada en [7]. La prueba anterior sigue el trabajo de Granville y Soundararajan [12] y de Billingsley [2].

1.4. Razón de ser de la renormalización en el teorema de Erdös-Kac

Como mencionamos en la Observación 1.3.1, si estamos familiarizados con la Inferencia Estadística, el Teorema 1.3.2 nos puede recordar, por su forma, a una aplicación del Teorema Central del Límite.

Aunque no es tan sencillo (tal y como vimos en su demostración), sí podemos calcular la esperanza y varianza de $\omega(n)$, tal y como se hace en [16].

Sea $\Omega_N = \{1, \ldots, N\}$ con la medida de probabilidad uniforme. Para $1 \leq n \leq N$, queremos calcular la esperanza y la varianza de $\omega(n)$.

Empezaremos por la esperanza. Para ello, definimos las variables aleatorias

$$X_m = \begin{cases} 0 & \text{si } m \nmid n, \\ 1 & \text{si } m | n. \end{cases} \tag{1.12}$$

Entonces, tenemos que

$$\mathbb{E}(\omega(n)) = \mathbb{E}\left(\sum_{p \leq N} X_p\right) = \sum_{p \leq N} \mathbb{E}(X_p), \tag{1.13}$$

que se tiene por las propiedades de la esperanza de variables aleatorias. Calculemos la esperanza de una de las X_m cualquiera.

$$\mathbb{E}(X_m) = \mathbb{P}(X_m = 1) = \frac{1}{N} \sum_{\substack{n \leq N \\ m|n}} 1 = \frac{1}{N}\left\lfloor \frac{N}{m} \right\rfloor = \frac{1}{N}\left(\frac{N}{m} + O(1)\right) = \frac{1}{m} + O\left(\frac{1}{N}\right). \tag{1.14}$$

Teniendo esto en cuenta, podemos calcular de forma explícita (1.13).

$$\sum_{p\leq N}\mathbb{E}(X_p)=\sum_{p\leq N}\frac{1}{p}+\sum_{p\leq N}O\left(\frac{1}{N}\right).$$

Por el Teorema C.2.3

$$\sum_{p\leq N}\frac{1}{p}=\log\log N+O(1).$$

Por lo tanto

$$\mathbb{E}(\omega(n))=\log\log N+O(1) \tag{1.15}$$

Ya tenemos la esperanza. Para la varianza, nos será de utilidad, además de ciertas propiedades de la esperanza y la varianza, la desigualdad de Chebyshev (Teorema B.3.9).

Proposición 1.4.1. *Sean X e Y dos variables aleatorias independientes. Entonces, se cumple*

$$\mathbb{E}(XY)=\mathbb{E}(X)\mathbb{E}(Y).$$

Por lo que, a su vez, se tiene

$$\mathbb{V}(X+Y)=\mathbb{E}((X+Y)^2)-\mathbb{E}(X+Y)^2=\mathbb{V}(X)+\mathbb{V}(Y).$$

En general, para $X_1,\ldots,X_n$ variables aleatorias independientes se cumple

$$\mathbb{V}(X_1+\cdots+X_n)=\mathbb{V}(X_1)+\cdots+\mathbb{V}(X_n).$$

Demostración. Empezaremos con la propiedad de la esperanza, demostrando para el caso de dos variables aleatorias discretas, pues el caso continuo es análogo(basta con intercambiar sumatorio por integral y función de probabilidad por función de densidad). Aplicando el concepto de independencia de variables aleatorias se tiene

$$\begin{aligned}\mathbb{E}(XY)&=\sum_x\sum_y xy\mathbb{P}[X=x,Y=y]=\sum_x\sum_y xy\mathbb{P}[X=x]\mathbb{P}[Y=y]\\&=\sum_x x\mathbb{P}[X=x]\sum_y y\mathbb{P}[Y=y]=\mathbb{E}(X)\mathbb{E}(Y).\end{aligned}$$

Para la varianza, probaremos la segunda expresión, pues la primera es un caso particular de ésta. Sabemos que

$$\mathbb{V}(X_1+\ldots+X_n)=\sum_{i=1}^{n}\mathbb{V}(X_i)+2\sum_{1\leq i<j\leq n}cov(X_i,X_j).$$

Sabiendo esto, ya lo tendríamos probado, pues la covarianza de dos variables aleatorias independientes es 0, por lo que el segundo sumando sería nulo, quedando solamente el primero, como queríamos probar. □

Sabiendo esto, podemos empezar. Consideremos las variables aleatorias indexadas por los primos definidas en (1.12). Podemos observar que $X_p^2=X_p$, pues $0^2=0$ y $1^2=1$. Luego

$$\mathbb{V}(X_p)=\mathbb{E}(X_p^2)-\mathbb{E}(X_p)^2=\mathbb{E}(X_p)-\mathbb{E}(X_p)^2=\mathbb{E}(X_p)-O(\frac{1}{p^2}).$$

Si las X_p fueran independientes, se podría proceder de la siguiente forma

$$\sum_{p\leq N}\mathbb{V}(X_p)=\sum_{p\leq N}\mathbb{E}(X_p)-\sum_{p\leq N}\frac{1}{p^2}=\sum_{p\leq N}\mathbb{E}(X_p)-O(1)=\log\log N+O(1)$$

Desgraciadamente, esto no es así. Tomemos p_1,p_2 con $p_1,p_2\geq\sqrt{N}$. Entonces, X_{p_1} y X_{p_2} no pueden ser simultaneamente 1 (ver Proposición A.1.2), lo cual no ocurre con variables aleatorias independientes. Sin embargo, solo es necesario que la propiedad de la esperanza se cumpla de forma aproximada. Sean p_1,p_2 con $p_1\neq p_2$

$$\begin{aligned}\mathbb{E}(X_{p_1}X_{p_2})=\mathbb{E}(X_{p_1p_2})&=\frac{1}{p_1p_2}+O\left(\frac{1}{N}\right),\\ \mathbb{E}(X_{p_1})\mathbb{E}(X_{p_2})&=\left(\frac{1}{p_1}+O\left(\frac{1}{N}\right)\right)\left(\frac{1}{p_2}+O\left(\frac{1}{N}\right)\right)=\frac{1}{p_1p_2}+O\left(\frac{1}{N}\right).\end{aligned}$$

Por lo tanto

$$\mathbb{V}(X_{p_1}+X_{p_2})=\mathbb{E}((X_{p_1}+X_{p_2})^2)-\mathbb{E}(X_{p_1}+X_{p_2})^2=\mathbb{V}(X_{p_1})+\mathbb{V}(X_{p_2})+O(1/N).$$

De lo cual se puede deducir

$$\mathbb{V}(\sum_{p\leq M}X_p)=\sum_{p\leq M}\mathbb{V}(X_p)+O\left(\frac{M^2}{N}\right)$$

para todo M. Convenientemente, podemos escoger M de tal forma que el término de error $O(M^2/N)$ sea pequeño, como por ejemplo, $M=N^{1/3}$. Por el Teorema B.3.9 llegamos a que

$$\mathbb{P}(|X-\mathbb{E}(X)|\geq x)\leq\frac{\log\log N+O(1)}{x^2}\tag{1.16}$$

para $X=\sum_{p\leq N^{1/3}}X_p$. Veamos, ahora, la diferencia entre X y $\omega(n)$. Un número $n\leq N$ no puede tener más de dos divisores primos $>N^{1/3}$. Por lo tanto, $|X-\omega(n)|$ nunca es más de 2. Se tiene, por tanto, que

$$\mathbb{P}(|\omega(n)-\log\log n|\geq x)\leq\frac{\log\log N+O(1)}{(x+O(1))^2}.\tag{1.17}$$

Equivalentemente

$$\mathbb{P}\left(|\omega(n)-\log\log N|\geq t\sqrt{\log\log N}\right)\leq\frac{1+O(1/\sqrt{\log\log N})}{t^2}\tag{1.18}$$

para todo $t\geq 1$.

Observación 1.4.1. Tanto el resultado (1.18) como la prueba que hemos presentado se deben a Turán. Hardy y Ramanujan, a su vez, habían dado antes una prueba más compleja de un resultado algo más débil.

1.5. Primeros pasos en la cuantificación del desvío respecto a la media

Como mencionamos en la Observación 1.4.1, existe un resultado anterior al Teorema 1.3.2, publicado por Hardy y Ramanujan en 1917 [13], el cual puede ser considerado como precursor del Teorema 1.3.2.

Pero antes de probar el teorema, necesitamos introducir algunos conceptos.

Definición 1.5.1 (Orden normal)**.** *Sean $f(n), F(n)$ dos funciones aritméticas. Decimos que $f(n)$ tiene orden normal $F(n)$ si $f(n)$ es aproximadamente $F(n)$ para casi todo n (ver Definición A.2.1). Es decir, para todo $\varepsilon > 0$ se cumple*

$$\lim_{x\to\infty} \frac{\left|\left\{n \le x : (1-\varepsilon)F(n) < f(n) < (1+\varepsilon)F(n)\right\}\right|}{x} = 1.$$

Entonces decimos que el orden normal de $f(n)$ es $F(n)$.

Teorema 1.5.2. *A partir de la Definición 1.5.1, se deduce que el orden normal de $\omega(n)$ y $\Omega(n)$ es $\log\log N$. Más precisamente:*

$$\sum_{n\le x} \omega(n) = x\log\log x + B_1 x + o(x), \tag{1.19}$$

$$\sum_{n\le x} \Omega(n) = x\log\log x + B_2 x + o(x), \tag{1.20}$$

donde B_1 es la constante dada por

$$B_1 = \gamma + \sum_{p\le x}\left(\log\left(1-\frac{1}{p}\right) + \frac{1}{p}\right),$$

donde, a su vez, γ es la constante de Euler-Mascheroni, y B_2 es la constante dada por

$$B_2 = B_1 + \sum_{p\le x} \frac{1}{p(p-1)}.$$

Demostración. En primer lugar, escribimos

$$S_1 = \sum_{n\le x}\omega(n) = \sum_{n\le x}\sum_{p\mid n} 1 = \sum_{p\le x}\left[\frac{x}{p}\right],$$

pues hay exactamente $[x/p]$ números enteros positivos menores o iguales que x que son divisibles por p. Quitando los corchetes, nos queda

$$S_1 = \sum_{p\le x}\frac{x}{p} + O\{\pi(x)\} = x\log\log x + B_1 x + o(x) \tag{1.21}$$

por el Teorema C.2.3. De la misma forma

$$S_2 = \sum_{n\le x}\Omega(n) = \sum_{n\le x}\sum_{p^m\mid n} 1 = \sum_{p^m\le x}\left[\frac{x}{p^m}\right]. \tag{1.22}$$

Por lo que

$$S_2 - S_1 = \sum{}' \left[\frac{x}{p^m} \right],$$

donde el sumatorio anterior recorre las potencias de primos $p^m \leq x$ con $m \geq 2$. Quitando los corchetes, el error cometido es menor que

$$\sum{}' 1 \leq \sum{}' \frac{\log p}{\log 2} = \frac{\psi(x) - \vartheta(x)}{\log 2} = o(x),$$

a lo cual se llega aplicando el Teorema C.2.4. Por lo tanto

$$S_2 - S_1 = x \sum{}' p^{-m} + o(x).$$

La serie

$$\sum_{m \geq 2} \sum_p \frac{1}{p^m} = \sum_p \left(\frac{1}{p^2} + \frac{1}{p^3} + \cdots \right) = \sum_p \frac{1}{p(p-1)} = B_2 - B_1$$

es convergente y entonces

$$\sum{}' p^{-m} = B_2 - B_1 + o(1)$$

cuando $x \to +\infty$. Por tanto

$$S_2 - S_1 = (B_2 - B_1)x + o(x)$$

y (1.20) se deduce de (1.21). □

Teorema 1.5.3 (Hardy-Ramanujan)**.** *Sea $f(n)$ la función $\omega(n)$ o $\Omega(n)$. Para todo $\delta > 0$, la cantidad de $n \leq x$ tales que se cumple*

$$|f(n) - \log\log n| > (\log\log n)^{\frac{1}{2}+\delta}$$

es $o(x)$.Se deduce que el orden normal de $\omega(n)$ y $\Omega(n)$ es $\log\log N$.

Demostración. Basta probar que el número de enteros positivos n que cumplen

$$|\omega(n) - \log\log x| > (\log\log x)^{\frac{1}{2}+\delta} \tag{1.23}$$

es $o(x)$. El cambio anterior es justificable, pues la distinción entre $\log\log n$ y $\log\log x$ no tiene importancia. Como

$$\log\log x - 1 \leq \log\log n \leq \log\log x$$

cuando $x^{\frac{1}{e}} \leq n \leq x$, por lo que $\log\log n$ es prácticamente $\log\log x$ para todos esos valores de n, mientras que el número del resto de valores de n es

$$O(x^{\frac{1}{e}}) = o(x).$$

Solo necesitamos considerar el caso $f(n) = \omega(n)$. Como $\omega(n) \geq \Omega(n)$ y, por (1.19) y (1.20),

$$\sum_{n \leq x} \{\Omega(n) - \omega(n)\} = O(x).$$

Por lo que el número de enteros positivos menores que x que cumplen

$$\Omega(n) - \omega(n) > (\log\log x)^{\frac{1}{2}}$$

es

$$O\left(\frac{x}{(\log\log x)^{\frac{1}{2}}}\right) = o(x).$$

Por lo que ambas posibilidades para $f(n)$ son equivalentes. Consideremos el número de pares de factores primos p, q $(p \neq q)$ de n, contando como distintas las parejas (p, q) y (q, p). Hay $\omega(n)$ posibles valores para p, y para cada uno de estos, $\omega(n) - 1$ valores posibles de q. Por lo tanto

$$\omega(n)(\omega(n) - 1) = \sum_{\substack{pq|n \\ p\neq q}} 1 = \sum_{pq|n} 1 - \sum_{p^2|n} 1.$$

Sumando sobre todos los $n \leq x$, tenemos

$$\sum_{n\leq x}\{\omega(n)\}^2 - \sum_{n\leq x}\omega(n) = \sum_{n\leq x}\left(\sum_{pq|n} 1 - \sum_{p^2|n} 1\right) = \sum_{pq\leq x}\left[\frac{x}{pq}\right] - \sum_{p^2\leq x}\left[\frac{x}{p^2}\right].$$

Observemos el segundo sumando.

$$\sum_{p^2\leq x}\left[\frac{x}{p^2}\right] \leq \sum_{p^2\leq x}\frac{x}{p^2} \leq x\sum_{p}\frac{1}{p^2} = O(x),$$

ya que la última suma es convergente. Veamos ahora el primer sumando.

$$\sum_{pq\leq x}\left[\frac{x}{pq}\right] = x\sum_{pq\leq x}\frac{1}{pq} + O(x).$$

Ahora, usando (1.19), tenemos

$$\sum_{n\leq x}\{\omega(n)\}^2 = x\sum_{pq\leq x}\frac{1}{pq} + O(x\log\log x). \tag{1.24}$$

Vemos también que

$$\left(\sum_{p\leq\sqrt{x}}\frac{1}{p}\right)^2 \leq \sum_{pq\leq x}\frac{1}{pq} \leq \left(\sum_{p\leq x}\frac{1}{p}\right)^2, \tag{1.25}$$

ya que, si $pq \leq x$ entonces $p < x$ y $q < x$, mientras que si $p \leq \sqrt{x}$ y $q \leq \sqrt{x}$, entonces $pq \leq x$. Las sumas derecha e izquierda en (1.25) son ambas

$$(\log\log x + O(1))^2 = (\log\log x)^2 + O(\log\log x)$$

y por lo tanto

$$\sum_{n\leq x}\{\omega(n)\}^2 = x(\log\log x)^2 + O(x\log\log x). \tag{1.26}$$

Se sigue que

$$\begin{aligned}\sum_{n\leq x}\{\omega(n)-\log\log x\}^2 &= \sum_{n\leq x}\{\omega(n)\}^2 - 2\log\log x\sum_{n\leq x}\omega(n) + [x](\log\log x)^2\\ &= x(\log\log x)^2 + O(x\log\log x) - 2\log\log x\{x\log\log x + O(x)\}\\ &\quad + \{x+O(1)\}(\log\log x)^2 = x(\log\log x)^2 - 2x(\log\log x)^2\\ &\quad + x(\log\log x)^2 + O(x\log\log x) = O(x\log\log x)\end{aligned} \tag{1.27}$$

por (1.19) y (1.26). Si hay más de νx enteros positivos menores o iguales que x que satisfacen (1.23) con $f(n)=\omega(n)$, entonces

$$\sum_{n\leq x}\{\omega(n)-\log\log x\}^2 \geq \nu x(\log\log x)^{1+2\delta},$$

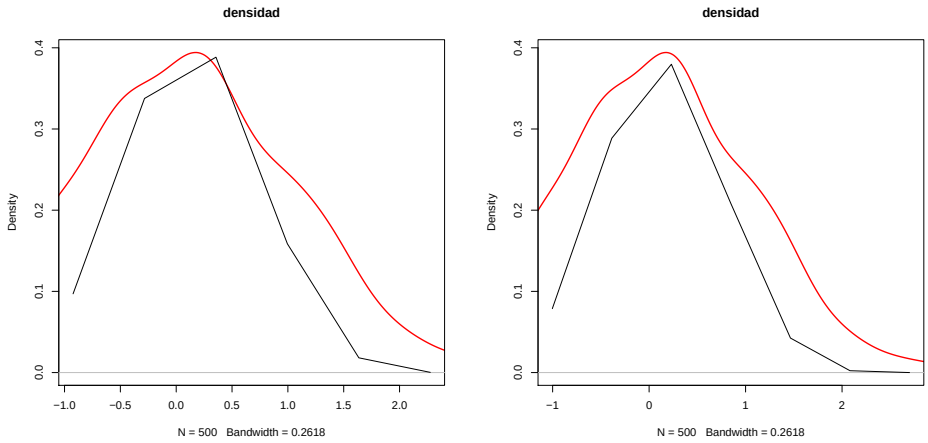
lo cual contradice (1.27) para x suficiente grande, y es cierto para todo positivo ν. Por lo tanto, el numero de enteros positivos que satisfacen (1.23) es $o(x)$. □

1.6. Teoría vs Práctica: visualización de resultados con software R

Habiendo visto en las secciones anteriores algunos de los resultados que abrieron las puertas al uso de herramientas probabilísticas en la Teoría de Números, vamos a visualizar estos resultados de forma experimental. Para ello, dispondremos de varios gráficos generados en R, un lenguaje de programación muy empleado, junto con Python, en el mundo de la Estadística. Los códigos de estos gráficos están disponibles en el Anexo D.

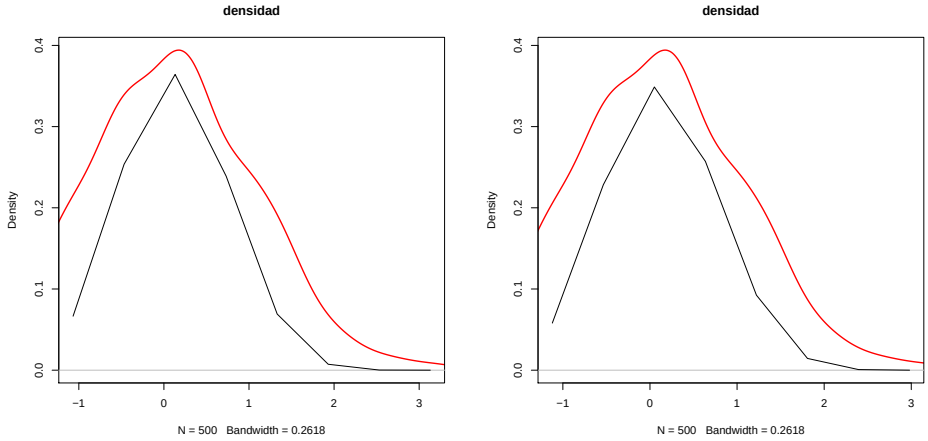
Empezaremos con el Teorema de Erdös-Kac, el Teorema 1.3.2. Recordemos que este teorema establece que, renormalizando la función $\omega(n)$, la función que para un número entero $n>1$ nos devuelve el número de factores primos contados sin multiplicidad, y considerando n como una variable aleatoria uniforme discreta en $\Omega_N=\{1,\ldots,N\}$ converge en ley a la normal estándar $\mathcal{N}$. Si queremos comprobar esto de manera visual, debemos comparar la "densidad" empírica generada por la variable definida en el Teorema 1.3.2 (para un N suficientemente grande) con la función de densidad de una normal estándar $\mathcal{N}$. Recordemos que la función de densidad es un concepto atribuido a las variables aleatorias continuas, mientras que la variable aleatoria que estudiamos es discreta, es por eso que lo entrecomillamos. A partir de este momento, omitiremos las comillas cuando hablemos de densidades, aunque tratemos con distribuciones discretas.

Vemos que el resultado es coherente con el Teorema de Erdös-Kac, pues conforme N crece, la curva negra se aproxima cada vez más a la roja. Esto significa que la distribución empírica se asemeja, conforme N crece, a la función de distribución de la normal estándar $\mathcal{N}$, lo cual era lo esperado. Como curiosidad, los números que aparecen en la parte inferior de las gráficas son referentes a la muestra generada para construir la función de distribución de la normal. El primer dato es el número de elementos de la muestra, que en este caso es $n=500$. Por otro lado, el segundo es una medida de la



(a) $N = 10^5$

(b) $N = 10^6$



(c) $N = 10^7$

(d) $N = 10^8$

Figura 1.3: Comparación de la función de densidad empírica (en negro) frente a la función de densidad de una normal estándar $\mathcal{N}$ (en rojo).

bondad de ajuste gráfico de la distribución de la muestra con respecto a la distribución teórica, aunque no ofrece ninguna información pues este dato varía según la escala que se use en el gráfico.

A pesar de lo tentador que pueda resultar, no se debe comprobar la normalidad de los datos con tests estadísticos, pues no queremos ver si estos datos provienen de una normal. Puede resultar llamativa esta aclaración, pero es necesaria si el lector proviene del mundo del análisis de datos.

Recordando la forma de la variable aleatoria que se presenta en el Teorema 1.3.2, nos puede surgir una duda bastante natural. Si bien la convergencia en ley se tiene si renormalizamos, ¿qué ocurre si no lo hacemos? ¿Sigue habiendo algún tipo de convergencia? Y si la hay, ¿es a la misma o a otra variable aleatoria distinta?

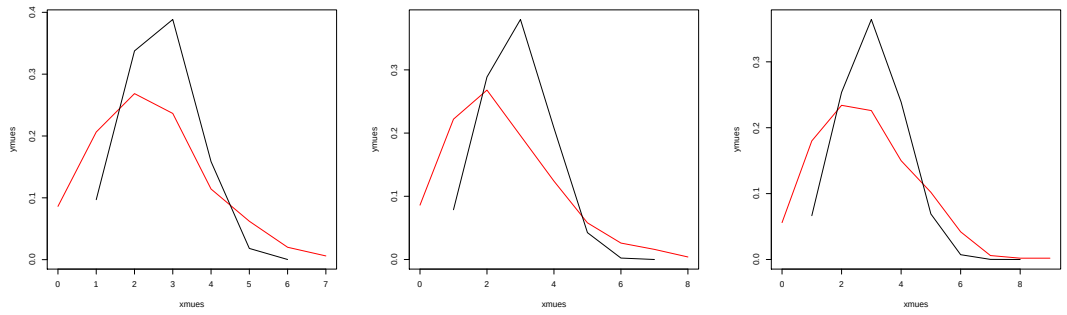
La respuesta es que sí, este problema se puede seguir estudiando desde una perspectiva probabilística, aunque las herramientas a usar no son precisamente sencillas ni genéricas. En este trabajo expondremos una de las maneras de abordar este problema. Para empezar, recordemos que una distribución de Poisson con parámetro real $\lambda \geq 0$ tiene por función de probabilidad

$$\mathbb{P}(x = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

para cualquier entero $k \geq 0$. Podemos llegar, mediante inducción y el Teorema de los Números Primos, a la fórmula asintótica

$$\frac{1}{N}|\{n \leq N \ : \ \omega(n) = k\}| \sim \frac{1}{(k-1)!}\frac{(\log\log N)^{k-1}}{\log N} = e^{-\log\log N}\frac{(\log\log N)^{k-1}}{(k-1)!},$$

para cualquier entero fijado $k \geq 1$. La expresión anterior parece indicar que una aproximación decente para $\omega(n)$ podría ser una Poisson de parámetro $\log\log N$. Por lo tanto, el Teorema de Erdös-Kac sería una consecuencia de la Proposición B.3.34. En la figura 1.4 podemos ver que una aproximación de este tipo no sería mala idea.



(a) $N = 10^5$ (b) $N = 10^6$ (c) $N = 10^7$

Figura 1.4: Comparación de la función de densidad empírica (en negro) frente a la función de densidad de una Poisson de parámetro adecuado (en rojo).

Para darle un fundamento riguroso a estas ideas, debemos dar una definición precisa, que, por la variabilidad del parámetro de la Poisson, no puede ser una propiedad

de convergencia al uso. Harper [15] fue el primero en hacerlo. Llegó a una expresión explícita para una cota superior para la variación total entre una versión truncada de $\omega(n)$ en Ω_N y una Poisson adecuada. La versión truncada de $\omega(n)$ sería

$$\sum_{\substack{p|n \\ p\leq Q}} 1, \quad \textit{donde } Q = N^{1/(3\log\log N)^2}$$

y la Poisson Po_N cuyo parámetro es

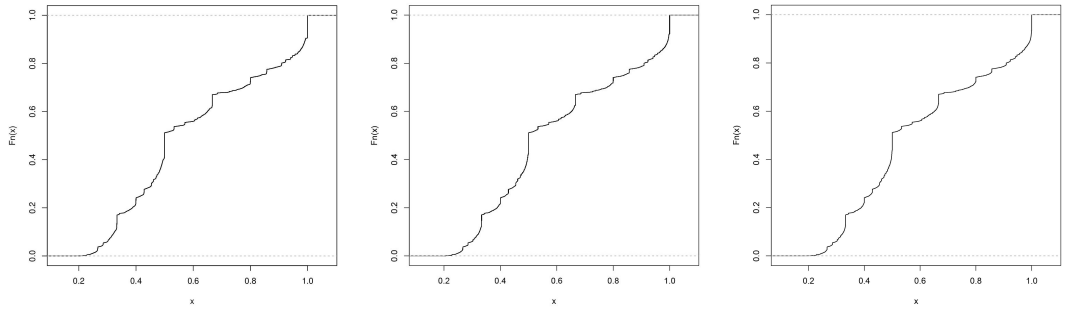
$$\lambda_N = \sum_{p\leq Q} \frac{1}{N}\left\lfloor \frac{N}{p} \right\rfloor$$

(por lo que la Fórmula de Mertens implica que $\lambda_N \sim \log\log N$). Harper probó que, para cualquier subconjunto de enteros no negativos A, se cumple que

$$\left| \mathbb{P}_N\left(\sum_{\substack{p|n \\ p\leq Q}} 1 \in A \right) - \mathbb{P}(Po_N \in A) \right| \ll \frac{1}{\log\log N},$$

y es más, que la velocidad de decrecimiento $(\log\log N)^{-1}$ es la mejor posible. Para probar este resultado, se necesitan herramientas probabilísticas mucho más avanzadas de las que se usan en la prueba del Teorema 1.3.2.

Por último, veremos la distribución de la función totient de Euler, tal y como se ve en la Proposición 1.2.4.



(a) $N = 10^5$ (b) $N = 10^6$ (c) $N = 10^7$

Figura 1.5: Función de distribución de $\varphi(n)/n$, con $n \in \Omega_N$.

Gracias a estos gráficos, podemos obtener algo de información de la forma de la distribución límite F.

2. Distribución de los valores de la función $\zeta(s)$ de Riemann

> Si yo me despertara después de haber dormido durante mil años, mi primera pregunta sería: ¿Ha sido demostrada la hipótesis de Riemann?
>
> *David Hilbert*

2.1. Introducción

En este capítulo se tratarán diversas propiedades de una de las funciones más conocidas y estudiadas de las matemáticas, la función $\zeta(s)$ de Riemann. Si bien es cierto que en su estudio son de vital importancia las herramientas analíticas, a lo largo de este capítulo se dará una perspectiva probabilística de este tema.

Con este objetivo en mente, se empezarán tratando tres de los teoremas fundamentales en el estudio probabilístico de la función $\zeta(s)$ de Riemann, en los cuales estudiaremos la distribución de la función $\zeta(s)$ fijando la parte real de s, tal y como se ve en los Teoremas de Bohr-Jessen y Bagchi, además de estudiar el soporte de las variables aleatorias mediante el resultado alcanzado por Bagchi y Voronin. De forma más particular, estudiaremos también qué ocurre cuando fijamos $\Re(s) = 1/2$, con el teorema de Selberg. Además, veremos dos métodos de aproximación, mediante productos de Euler y la aproximación polinómica de Dirichlet. Para acabar, trataremos también una interpretación probabilística del que es, *con probabilidad 1*, el problema más famoso de las matemáticas actuales, la Hipótesis de Riemann.

Al igual que en el capítulo anterior, se precisan de ciertos conocimientos previos en Teoría de la Probabilidad, Teoría Analítica de Números y Análisis Matemático. Si no es el caso, cualquier resultado auxiliar empleado será explicitado en el Anexo correspondiente para consulta y/o curiosidad del lector.

A modo de aviso al lector, en las pruebas del Teorema 2.2.3 se omiten las pruebas de algunas proposiciones empleadas, pues escapan en dificultad y temática de los objetivos de este trabajo. Aun así, se pueden encontrar referencias a las pruebas de estos resultados.

2.2. Teoremas de Bohr-Jessen y Bagchi

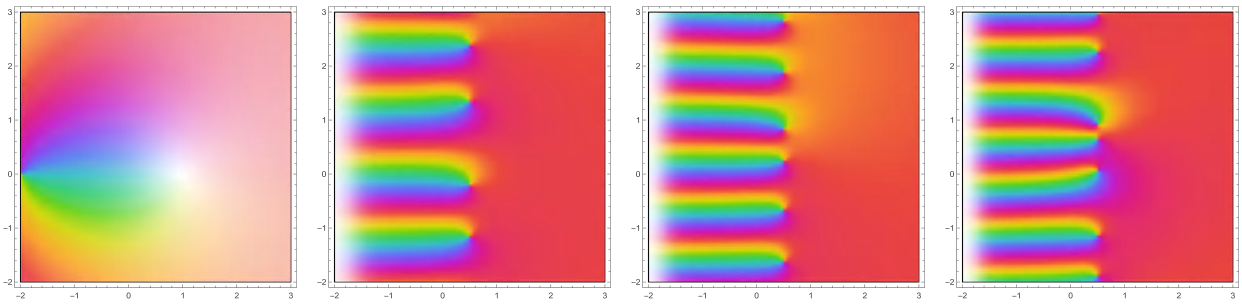
En primer lugar, debemos conocer algunas propiedades de la función $\zeta(s)$ de Riemann. tanto su definición como algunas propiedades claves se encuentran en el Anexo C.

Como mencionamos anteriormente, la Hipótesis de Riemann es el problema que más llama la atención, tanto de matemáticos como de otros tipos de científicos. Es por ello que, a pesar de lo tentador que puede resultar empezar a estudiar qué ocurre cuando $\Re(s) = 1/2$, comenzaremos a trabajar algo más general, cuando $\Re(s) \geq 1/2$. En particular, será mas interesante cuando $\Re(s) \leq 1$. Aún asi, el caso $\Re(s) = 1/2$ será estudiado con más detenimiento en la próxima sección.

Prosiguiendo con lo anterior, definiremos las variables aleatorias que nos serán de gran utilidad en el desarrollo de esta sección.

Definición 2.2.1. *Sea τ un número real fijo, y sea $T \geq 2$ otro número real. Sea $\Omega_T = [-T, T]$ con la medida de probabilidad uniforme $dt/2T$. Se definen las variables aleatorias $Z_{\tau,T}$ como*

$$Z_{\tau,T}(t) := \zeta(\tau + it)$$



(a) t=0 (b) t=1000 (c) t=5000 (d) t=10000

Figura 2.1: $\zeta(s+it)$, con s en el rectángulo determinado por los vértices $-2-2i, 3+3i$, donde los colores codifican el argumento de los valores obtenidos.

El comportamiento asintótico de estas variables aleatorias viene dado por el Teorema de Bohr-Jessen.

Teorema 2.2.2 (Bohr-Jessen)**.** *Sea $\tau > 1/2$ un número real fijo. Consideramos las variables aleatorias $Z_{\tau,T}$ definidas anteriormente. Entonces existe una medida de probabilidad μ_τ en $\mathbb{C}$ tal que $Z_{\tau,T}$ converge a ella cuando $T \to +\infty$. Es más, el soporte de μ_τ es compacto si $\tau > 1$ y es igual a $\mathbb{C}$ si $\tau \in (1/2, 1]$.*

Para poder probar este resultado, necesitamos presentar primero el resto de resultados de esta sección.

Teorema 2.2.3 (Bagchi)**.** *Sea $\tau > 1/2$ un número real fijo. Si $1/2 < \tau < 1$, sea $r > 0$ tal que*

$$D = \{s \in \mathbb{C} \ : \ |s - \tau| \leq r\} \subset \{s \in \mathbb{C} \ : \ 1/2 < \Re(s) < 1\}$$

Figura 2.2: Harald Bohr, uno de los matemáticos que formuló el Teorema principal de esta sección.

y si $\tau \geq 1$ *sea* D *cualquier subconjunto compacto de* $\{s \in \mathbb{C} \ : \ \Re(s) \geq 1\}$ *tal que* $\tau \in D$. *Consideremos las variables aleatorias* $H(D)-$*valuadas* $Z_{D,T}$ *definidas por*

$$t \mapsto (s \mapsto \zeta(s+it))$$

en Ω_T. *Sea* $(X_p)_p$ *una sucesión de variables aleatorias indexadas por los primos, independientes e identicamente distribuidas, con distribución uniforme en el círculo unidad. Entonces se tiene la convergencia en ley* $Z_{D,T} \to Z_D$ *con* $T \to +\infty$, *donde* Z_D *es el producto de Euler aleatorio dado por*

$$Z_D(s) := \prod_p (1 - p^{-s}X_p)^{-1}.$$

Corolario 2.2.4. *Sea* $\tau > 1/2$ *un número real fijo. Cuando* $T \to +\infty$, *las variables aleatorias del Teorema 2.2.2 convergen en ley a la variable aleatoria* $Z_D(\tau)$, *donde* D *es un disco*

$$D = \{s \in \mathbb{C} \ : \ |s - \tau| \leq r\}$$

contenido en el interior de la banda crítica, si $\tau < 1$, *o bien cualquier subconjunto compacto de* $\{s \in \mathbb{C} \ : \ \Re(s) \geq 1\}$ *tal que* $\tau \in D$.

Teorema 2.2.5 (Bagchi-Voronin)**.** *Sea* $1/2 < \tau < 1$ *y sea* $r > 0$ *tal que*

$$D = \{s \in \mathbb{C} \ : \ |s - \tau| \leq r\} \subset \{s \in \mathbb{C} \ : \ 1/2 < \Re(s) < 1\}.$$

El soporte de Z_D *contiene*

$$H(D)^\times = \{f \in H(D) \ : \ f(z) \neq 0 \text{ para todo } z \in D\}$$

y es igual a $H(D)^\times \cup \{0\}$. *En particular, para cualquier función* $f \in H(D)^\times$, *y para cualquier* $\varepsilon > 0$, *existe* $t \in \mathbb{R}$ *tal que*

$$\sup_{s \in D} |\zeta(s+it) - f(s)| < \varepsilon. \tag{2.1}$$

Ahora sí, procedemos a probar el Teorema 2.2.2.

Demostración. Fijamos D como en el Corolario 2.2.4. Claramente tenemos que

$$Z_{\tau,T} = \zeta_{D,T}(\tau)$$

o, equivalentemente, $Z_{\tau,T} = e_\tau \circ \zeta_{D,T}$, donde $e_\tau(f) = f(\tau)$. Esta aplicación es continua en $H(D)$, luego por composición se tiene que la convergencia en ley $Z_{D,T} \to Z_D$ del Teorema 2.2.3 implica la convergencia en ley de $Z_{\tau,T}$ a la variable aleatoria $e_\tau \circ Z_D$, que es $Z_D(\tau)$. Para probar la última parte del teorema, basta aplicar el Teorema 2.2.5 para ver que es obvio que si $1/2 < \tau < 1$, el soporte de $Z_D(\tau)$ es $\mathbb{C}$. □

Nuestro siguiente paso es probar el Teorema 2.2.3, para lo cual nos es necesario exponer algunos resultados auxiliares. Además, nos faltaría probar el Teorema 2.2.5. Sin embargo, la prueba de este resultado se escapa, tanto en herramientas como en complejidad, de los objetivos de este trabajo, por lo que no será incluida. No obstante, si el lector lo desea, puede ver una prueba parcial e incluso una aplicación de este resultado en [19], Sección 3.3. Procedemos, pues, con los resultados auxiliares para la demostración del Teorema 2.2.3.

Proposición 2.2.6. *Para $T \geq 0$, sean $X_T = (X_{p,T})_p$ la sucesión de variables aleatorias en Ω_T dadas por*

$$t \mapsto (p^{-it})_p.$$

Entonces X_T converge en ley con $T \to +\infty$ a la sucesión $X = (X_p)_p$ de variables aleatorias independientes, cada una de las cuales está distribuida uniformemente en $\mathbf{S}^1$.

Demostración. Para probar este resultado, veamos las sucesiones $(X_{p,T})_p$ y $(X_p)_p$ como variables aleatorias en Ω_T, tomando valor en el producto infinito

$$\hat{\mathbf{S}}^1 = \prod_p \mathbf{S}^1$$

de copias del círculo unidad indexadas por los números primos. Con esta interpretación, $(X_p)_p$ es una medida de probabilidad de Haar (ver Definición B.3.35) en el grupo $\hat{\mathbf{S}}^1$. Esto nos permite probar la convergencia mediante el Criterio de Weyl (ver Teorema B.3.23), la Proposición es equivalente con la siguiente propiedad:

$$\lim_{T\to+\infty} \mathbb{E}_T(\chi(X_{p,t})) = 0 \tag{2.2}$$

para cualquier caracter unitario continuo no trivial $\chi\colon \hat{\mathbf{S}}^1 \to \mathbf{S}^1$. Una propiedad de los grupos compactos ($\hat{\mathbf{S}}^1$ lo es) es que para cualquier caracter del tipo anterior existe un conjunto finito no vacío de primos S, y para cada $p \in S$ existe algún entero $m_p \in \mathbb{Z}\backslash\{0\}$ tal que

$$\chi(z) = \prod_{p\in S} z_p^{m_p}$$

para cualquier $z = (z_p)_p \in \hat{\mathbf{S}}^1$. Entonces tenemos por definición

$$\mathbb{E}_T(\chi(X_{pT})) = \frac{1}{2T}\int_{-T}^{T} \prod_{p\in S} p^{-itm_p}\, dt = \frac{1}{2T}\int_{-T}^{T} r^{-it}\, dt$$

donde $r > 0$ es el número racional dado por

$$r = \prod_{p \in S} p^{m_p}.$$

Dado que $r \neq 1$ (al ser S no vacío y $m_p \neq 0$), tenemos que $\mathbb{E}_T(\chi(X_{pT})) \to 0$ cuando $T \to +\infty$ por el Lema B.3.11. □

Lema 2.2.7. *Sean (X_n) las variables aleatorias definidas por*

$$X_n := \prod_{p|n} X_p^{v_p(n)},$$

donde las $(X_p)_p$ están definidas en el Teorema 2.2.3, y $v_p(n)$ es el exponente correspondiente a p en la descomposición en factores primos de n. Sea $Z(s)$ la serie de Dirichlet aleatoria $\sum X_n n^{-s}$ definida y holomorfa casi seguro cuando $\Re(s) > 1/2$. Para cualquier $\sigma_1 > 1/2$, se cumple

$$\mathbb{E}(|Z(s)|) \ll 1 + |s|$$

uniformemente para todo s tal que $\Re(s) \geq \sigma_1$.

Demostración. La serie

$$\sum_{n \geq 1} X_n n^{-\sigma_1}$$

converge casi seguro. De esto se sigue que las sumas parciales

$$S_u := \sum_{n < u} X_n n^{-\sigma_1}$$

están acotadas casi seguro. Por el Lema A.2.2, se tiene que para cualquier s con parte real $\sigma > \sigma_1$, se tiene que

$$Z(s) = (s - \sigma_1) \int_1^{+\infty} \frac{S_u}{u^{s-\sigma_1+1}}\, du,$$

donde la integral converge casi seguro. Entonces

$$|Z(s)| \leq (1 + |s|) \int_1^{+\infty} \frac{|S_u|}{u^{\sigma-\sigma_1+1}}\, du.$$

Por el Teorema de Fubini para funciones no negativas y la desigualdad de Cauchy-Schwarz se tiene que

$$\begin{aligned}
\mathbb{E}(|Z(s)|) &\leq (1 + |s|) \int_1^{+\infty} \mathbb{E}(|S_u|) \frac{du}{u^{\sigma-\sigma_1+1}} \\
&\leq (1 + |s|) \int_1^{+\infty} \mathbb{E}(|S_u|^2)^{1/2} \frac{du}{u^{\sigma-\sigma_1+1}} \\
&= (1 + |s|) \int_1^{+\infty} \mathbb{E}(\sum_{n \leq u} \frac{1}{n^{2\sigma_1}})^{1/2} \frac{du}{u^{\sigma-\sigma_1+1}}
\end{aligned}$$

usando la ortonormalidad de las X_n. El integrando es $\ll u^{-\frac{1}{2}-\sigma}$, luego la integral converge uniformemente para $\sigma \geq \sigma_1$. □

Ahora sí, estamos en disposición de probar el Teorema 2.2.3.

Demostración. Por la Proposición B.3.10, basta probar que para cualquier función acotada y Lipschitziana $f : H(D) \to \mathbb{C}$, se cumple

$$\mathbb{E}_T(f(Z_{D,T})) \to \mathbb{E}(f(Z_D))$$

cuando $T \to +\infty$. Usaremos la expansión en serie de Dirichlet de Z_D como en la Proposición B.3.19. Dado que D es fijo, se puede omitir de la notación para simplificar, denotándose $Z_T = Z_{D,T}$ y $Z = Z_D$. Sea $N \geq 1$ un entero fijado a elegir después. Denotamos

$$Z_{T,N} = \sum_{n \geq 1} n^{-s-it} \varphi\left(\frac{n}{N}\right)$$

visto como una variable aleatoria para $t \in [-T, T]$. Y

$$Z_N = \sum_{n \geq 1} X_n n^{-s} \varphi\left(\frac{n}{N}\right)$$

las sumas parciales suavizadas de las series de Dirichlet como en las Proposiciones B.3.20 y B.3.21. Entonces escribimos

$$\begin{aligned} |\mathbb{E}_T(f(Z_T)) - \mathbb{E}(f(Z))| \leq\ & |\mathbb{E}_T(f(Z_T) - f(Z_{T,N}))| \\ & + |\mathbb{E}_T(f(Z_{T,N})) - \mathbb{E}(f(Z_N))| + |\mathbb{E}(f(Z_N) - f(Z))|. \end{aligned}$$

Dado que f es una función Lipschitziana en $H(D)$, existe una constante $C \geq 0$ tal que

$$|f(x) - f(y)| \leq \|x - y\|_\infty$$

para cualesquiera $x, y \in H(D)$. Por lo tanto tenemos

$$\begin{aligned} |\mathbb{E}_T(f(Z_T)) - \mathbb{E}(f(Z))| \leq\ & C\mathbb{E}_T(\|Z_T - Z_{T,N}\|_\infty) + \\ & + |\mathbb{E}_T(f(Z_{T,N})) - \mathbb{E}(f(Z_N))| + C\mathbb{E}(\|Z_N - Z\|_\infty). \end{aligned}$$

Fijemos $\varepsilon > 0$. Las Proposiciones B.3.20 y B.3.21 juntas muestran que existe algún $N \geq 1$ y alguna constante $C_1 \geq 0$ tal que

$$\mathbb{E}_T(\|Z_T - Z_{T,N}\|_\infty) < \varepsilon + \frac{C_1 N}{T}$$

para todo $T \geq 1$ y

$$\mathbb{E}(\|Z_N - Z\|_\infty) < \varepsilon.$$

Fijemos ese valor concreto de N. Por composición y por la Proposición 2.2.6, las variables aleatorias $Z_{N,T}$, que son polinomios de Dirichlet, convergen en ley a Z_N cuando $T \to +\infty$. como $N/T \to 0$ también uando $T \to +\infty$, deducimos que para cualquier T suficientemente grande, se tiene

$$|\mathbb{E}_T(f(Z_T)) - \mathbb{E}(f(Z))| < 4\varepsilon,$$

con lo que concluye la prueba. □

2.3. Teorema de Selberg

Como se dijo anteriormente, en esta sección buscaremos dilucidar qué ocurre en los puntos de la línea crítica, es decir, en los puntos de la forma $s = \frac{1}{2} + it$. De hecho el Teorema 2.2.2 falla para $\tau = 1/2$, lo que demuestra la complejidad del comportamiento de la función $\zeta(s)$ de Riemann sobre la línea crítica. No obstante, tenemos un Teorema límite previa normalización, debido a Selberg, que arroja algo de luz en el asunto. Este teorema se aplica no directamente a la función, sino al logaritmo de ésta. Para poder expresarlo, debemos definir el significado de $\log \zeta(\frac{1}{2} + it)$ y las variables aleatorias que intervienen en el teorema.

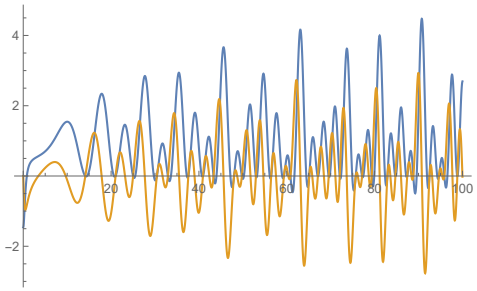


Figura 2.3: Parte real (azul) e imaginaria (naranja) de la función zeta de Riemann sobre la línea crítica cuando $t \in [0, 100]$.

Definición 2.3.1. *Se definen las variables aleatorias L_T sobre Ω_T como $L_T(t) = 0$ si $\zeta(\frac{1}{2} + it) = 0$, en cualquier otro caso*

$$L_T(t) = \log \zeta\left(\frac{1}{2} + it\right),$$

donde el logaritmo de la zeta es el de la única rama que es holomorfo en la franja

$$\{s = \sigma + iy \in \mathbb{C} \ : \ \sigma > \frac{1}{2} - \delta, \ |y - t| \leq \delta\}$$

para algún $\delta > 0$, y satisface que $\log \zeta(\sigma + it) \to 0$ cuando $\sigma \to +\infty$.

Veamos, ahora, el enunciado del Teorema de Selberg.

Teorema 2.3.2 (Selberg)**.** *Con la notación de la Definición 2.3.1, las variables aleatorias*

$$\frac{L_T}{\sqrt{\frac{1}{2}\log\log T}}$$

en Ω_T convergen en ley cuando $T \to +\infty$ a la variable aleatoria normal estándar compleja.

Probaremos parcialmente este resultado, pues solo consideraremos la parte real del $\log \zeta(\frac{1}{2} + it)$, es decir, $\log|\zeta(\frac{1}{2} + it)|$. Por lo que debemos redefinir las variables aleatorias de 2.3.1.

Figura 2.4: Atle Selberg, matemático que formuló el Teorema principal de esta sección.

Definición 2.3.3. *Se definen las variables aleatorias L_T sobre Ω_T como $L_T(t) = 0$ si $\zeta(\frac{1}{2} + it) = 0$, en cualquier otro caso*

$$L_T(t) = \log \left| \zeta \left(\frac{1}{2} + it \right) \right|.$$

Al trabajar con el módulo, podemos olvidarnos de la elección de la rama del logaritmo complejo. Nuestro objetivo pasa a ser probar el siguiente resultado.

Teorema 2.3.4 (Selberg)**.** *Con la notación de la Definición 2.3.3, las variables aleatorias*

$$\frac{L_T}{\sqrt{\frac{1}{2}\log\log T}}$$

en Ω_T convergen en ley cuando $T \to +\infty$ a la variable aleatoria normal estándar real.

El esquema a seguir para probar este Teorema es el siguiente.

1. Aproximar L_T por otra variable aleatoria $\tilde{L}_T$ dada por $t \to \log|\zeta(\sigma_0 + it)|$, para σ_0 suficientemente cerca de $1/2$ y dependiendo de T.

2. Para Z_T dada por $t \to \zeta(\sigma_0 + it)$, tal que $\log|Z_T| = \tilde{L}_T$, una aproximación de su inversa dada por un polinomio corto de Dirichlet, de la forma

$$D_T(s) = \sum_{n \geq 1} a_T(n)\mu(n)n^{-s}$$

donde $a_T(n)$ se anula a partir de algún n. A partir de aquí se obtiene una aproximación de L_T con $-\log|D_T|$.

3. Obtener una aproximación de $|D_T|$ mediante un polinomio corto de Euler, de la forma $\exp(-\Re(P_T))$, donde

$$P_T(t) = \sum_{p^k \leq X} \frac{1}{k} \frac{1}{p^{k(\sigma_0 + it)}}. \tag{2.3}$$

En este punto, L_T se aproxima por $\Re(P_T)$.

4. El último paso es probar la convergencia en ley a la normal estándar de

$$\frac{\Re(P_T)}{\sqrt{\frac{1}{2}\log\log T}}$$

cuando $T \to +\infty$.

Para poder probar este resultado, antes debemos definir los parámetros que intervendrán en la prueba, además de dar algunos resultados auxiliares que nos serán de gran utilidad.

Definición 2.3.5. *Sea $T \geq e^{e^2}$. Se denota por*

$$\varrho_T := \sqrt{\frac{1}{2}\log\log T} \geq 1, \tag{2.4}$$

al factor normalizante en el Teorema 2.3.4. Definimos también

$$W := (\log\log\log T)^4 \sim (\log \varrho_T)^4, \quad \sigma_0 := \frac{1}{2} + \frac{W}{\log T} = \frac{1}{2} + O\left(\frac{(\log \varrho_T)^4}{\log T}\right), \tag{2.5}$$

$$X := T^{1/(\log\log\log T)^2} = T^{1/\sqrt{W}}, \quad Y := T^{1/(\log\log T)^2} = T^{4/\varrho_T^4} \leq X \tag{2.6}$$

Teniendo ya los parámetros fijados, empezamos con los resultados auxiliares.

Proposición 2.3.6. *Se tiene*

$$\mathbb{E}_T\left(|L_T - \tilde{L}_T|\right) = o(\varrho_T)$$

cuando $T \to +\infty$.

Demostración. Usaremos la factorización de Hadamard de la función zeta de Riemann (ver C.3.6). Sea $t \in \Omega_T$ tal que no hay cero de $\zeta(s)$ con ordenada t. Tenemos

$$\log|\zeta(\sigma_0 + it)| - \log|\zeta(\frac{1}{2} + it)| = \Re\left(\int_{\frac{1}{2}}^{\sigma_0} \frac{\zeta'}{\zeta}(\sigma + it)\, d\sigma\right) = \int_{\frac{1}{2}}^{\sigma_0} \Re\left(\frac{\zeta'}{\zeta}(\sigma + it)\right) d\sigma.$$

Para cualquier σ con $\frac{1}{2} \leq \sigma \leq \sigma_0$, tenemos que

$$-\frac{\zeta'}{\zeta}(\sigma + it) = \sum_{|t - \Im\varrho| < 1} \frac{1}{\sigma + it - \varrho} + O(\log(2 + |t|)),$$

por la Proposición C.3.6, donde la suma es sobre los ceros de $\zeta(s)$, contados con multiplicidad, tales que $|\sigma + it - \varrho| < 1$. Fijamos, ahora, $t_0 \in \Omega_T$ e integramos sobre t tal que $|t - t_0| \leq 1$. Esto nos lleva a

$$\int_{t_0-1}^{t_0+1} \left|\log|\zeta(\sigma_0 + it)| - \log|\zeta(\tfrac{1}{2} + it)|\right| dt \leq \sum_{|\Im\varrho - t_0| < 1} \int_{t_0-1}^{t_0+1} \int_{\frac{1}{2}}^{\sigma_0} \left|\Re\left(\frac{1}{\sigma + it - \varrho}\right)\right| dt d\sigma.$$

Podemos trabajar con la integral final de la siguiente forma

$$\int_{\frac{1}{2}}^{\sigma_0} \left|\Re\left(\frac{1}{\sigma+it-\varrho}\right)\right| dt \leq \int_{\mathbb{R}} \left|\Re\left(\frac{1}{\sigma+it-\varrho}\right)\right| dt \leq \int_{\mathbb{R}} \frac{|\sigma-\beta|}{(\sigma-\beta)^2+(t-\gamma)^2}\, dt = \pi$$

para todos σ y ϱ. Por lo que tenemos que

$$\frac{1}{T}\int_{|t-t_0|\leq 1} \left|\log|\zeta(\tfrac{1}{2}+it-\varrho)| - \log|\zeta(\sigma_0+it-\varrho)|\right| dt \ll (\sigma_0-\tfrac{1}{2})\frac{m(t_0)}{T},$$

donde $m(t_0)$ es el número de ceros ϱ tal que $|t_0-\Im(\varrho)|\leq 1$. Esto es $\ll \log(2+|t_0|)$ por la Proposición C.3.6. Finalmente, sumando las cotas

$$\frac{1}{T}\int_{|t-t_0|\leq 1} \left|\log|\zeta(\tfrac{1}{2}+it-\varrho)| - \log|\zeta(\sigma_0+it-\varrho)|\right| dt \ll (\sigma_0-\tfrac{1}{2})\frac{\log(2+|t_0|)}{T},$$

sobre una partición de Ω_T en $\ll T$ intervalos de longitud 2, deducimos que

$$\mathbb{E}_T(|L_T-\tilde{L}_T|) \ll (\sigma_0-\tfrac{1}{2})\log T = W.$$

Tenemos que $W = o(\varrho_T)$, por lo que cuncluye la prueba. □

Para proseguir, debemos definir más parámetros y los polinomios de Dirichlet que intervienen.

Definición 2.3.7. *Con las condiciones previamente descritas, definimos las siguientes constantes*

$$m_1 := 100\log\log T \sim \varrho_T, \quad m_2 := 100\log\log\log T \sim \log\varrho_T.$$

Sea $b_T(n)$ la función característica de del conjunto de enteros positivos libres de cuadrados (A.1.1) tales que todos sus factores primos son menores o iguales que Y y tienen a lo más m_1 factores primos. Sea $c_T(n)$ la función característica de del conjunto de enteros positivos libres de cuadrados tales que todos sus factores primos p satisfacen $Y < p \leq X$ y tienen a lo más m_2 factores primos. Les asociamos los siguientes polinomios de Dirichlet

$$B(s) := \sum_{n\geq 1}\mu(n)b_T(n)n^{-s}, \quad C(s) := \sum_{n\geq 1}\mu(n)c_T(n)n^{-s}$$

para $s\in\mathbb{C}$. El coeficiente $a_T(n)$ es la convolución de Dirichlet

$$\begin{aligned}\sum_{de=n} b_T(d)c_T(e)\mu(d)\mu(e) &= \sum_{\substack{de=n\\(d,e)=1}} b_T(d)c_T(e)\mu(d)\mu(e)\\ &= \mu(n)\sum_{\substack{de=n\\(d,e)=1}} b_T(d)c_T(e) = \mu(n)a_T(n).\end{aligned}$$

Por su construcción, vemos que $a_T(n)$ es la función característica de los enteros positivos libres de cuadrados tales que todos sus factores primos son menores o iguales que X, poseen a lo más m_1 factores primos p tales que $p\leq Y$ y a lo más m_2 factores

primos p tales que $Y < p \leq X$. Por tanto, $a_T(n) = 0$ salvo cuando n es pequeño, concretamente

$$n \leq Y^{100 \log\log T} X^{100 \log\log T} = T^c$$

donde

$$c = \frac{100}{\log\log T} + \frac{100}{\log\log\log T} \to 0, \quad T \to +\infty.$$

Finalmente, definimos la variable aleatoria aritmética

$$D_T = D(\sigma_0 + it). \tag{2.7}$$

Proposición 2.3.8 (Aproximación polinómica de Dirichlet)**.** *La diferencia $Z_T D_T$ converge a 1 en L^2, es decir*

$$\lim_{T\to+\infty} \mathbb{E}_T\left(|1 - Z_T D_T|^2\right) = 0.$$

Proposición 2.3.9 (Aproximación polinómica de Euler)**.** *Las variables aleatorias $D_T \exp(-P_T)$ convergen a 1 en probabilidad, es decir, para todo $\varepsilon > 0$, se tiene*

$$\lim_{T\to+\infty} \mathbb{P}_T\left(|D_T \exp(-P_T) - 1| > \varepsilon\right) = 0.$$

En particular, $\mathbb{P}_T(D_T = 0)$ tiende a 0 cuando $T \to +\infty$

Observación 2.3.1. La prueba de las Proposiciones 2.3.8 y 2.3.9 se desarrolla en las secciones 2.4 y 2.5, respectivamente.

Proposición 2.3.10 (Producto gaussiano de Euler)**.** *Las variables aleatorias $\varrho_T^{-1} P_T$ convergen en ley cuando $T \to +\infty$ a la normal estándar compleja. En particular, las variables aleatorias*

$$\frac{\Re(P_T)}{\sqrt{\frac{1}{2}\log\log T}}$$

convergen en ley a la normal estándar real.

Demostración. Tenemos que $P_T = Q_T + R_T$, donde Q_T es la contribución de los primos, y R_T es la aportación de los cuadrados y potencias superiores de los primos. En primer lugar, afirmamos que R_T está uniformemente acotado en L^2 para todo T. De hecho, usando el Lema B.3.22 tenemos que

$$\begin{aligned}\mathbb{E}_T(|R_T|^2) &= \mathbb{E}_T\left(\left|\sum_{k\geq 2}\sum_{p\leq X^{1/k}} \frac{1}{k} p^{-k\sigma_0} p^{-kit}\right|^2\right) \\ &= \sum_{\substack{p^k \leq X \\ k\geq 2}} \frac{1}{k^2} p^{-2k\sigma_0} + O\left(\sum_{k,l\geq 2}\sum_{\substack{p^k, q^l \leq X \\ p\neq q}}\right) \frac{1}{kl}(pq)^{-2\sigma_0+\frac{1}{2}} \ll 1 + \frac{X^2 \log X}{T} \ll 1\end{aligned}$$

ya que $X \ll T^\varepsilon$ para todo $\varepsilon > 0$. A partir de ahora, basta probar que Q_T/ϱ_T converge en ley a una normal estándar compleja $\mathcal{N}$. Para ello, computaremos los momentos

$$\mathbb{E}_T(Q_T^k \overline{Q_T^l})$$

para $k, l \geq 0$ enteros y comparamos con el correspondiente momento de la variable aleatoria

$$Q_T = \sum_{p \leq X} p^{-\sigma_0} X_p.$$

Aplicando el Lema B.3.11, vemos que

$$\mathbb{E}_T(Q_T^k \overline{Q_T^l}) = \mathbb{E}_T(Q_T^k \overline{Q_T}^l) + O\left(\frac{1}{T} \sum_{m \neq n} (mn)^{-\sigma_0 + 1/2}\right)$$

donde la suma en el término de error recorre los enteros m (resp. n) con a lo más k factores primos, contados con multiplicidad, siendo todos ellos $\leq X$ (resp. con a lo más l factores primos, contados con multiplicidad, siendo todos ellos $\leq X$). Por lo tanto el término del error es

$$\ll \frac{1}{T}\left(\sum_{p \leq X} 1\right)^{k+l} \ll \frac{X^{k+l}}{T}.$$

A continuación, vemos que

$$\mathbb{V}(Q_T) = \sum_{p \leq X} p^{-2\sigma_0} \mathbb{V}(X_p^2) = \frac{1}{2} \sum_{p \leq X} p^{-2\sigma_0}.$$

Podemos computar esta suma dividiéndola en dos rangos: $p \leq Y$ y $Y < p \leq X$. La segunda suma sería

$$\ll \sum_{Y < p \leq X} \frac{1}{p} = \log\left(\frac{\log X}{\log Y}\right) + O(1) \ll \log\log\log T$$

por el Teorema C.2.3 y (2.5). Por otro lado, para $p \leq Y = T^{1/(\log\log T)^2}$, tenemos que

$$p^{-2\sigma_0} = p^{-1} \exp\left(-2\frac{\log p}{\log T} W\right) = p^{-1}\left(1 + O\left(\frac{W}{(\log\log T)^2}\right)\right),$$

que, con (2.5), implica que $\mathbb{V}(Q_T) \sim \frac{1}{2}\log\log T = \varrho_T^2$ cuando $T \to +\infty$. Finalmente, por el Teorema Central del Límite, $Q_T/\mathbb{V}(Q_T)$ y por tanto Q_T/ϱ_T convergen en ley a una normal estándar compleja, con convergencia de los momentos (ver Teoremas B.3.7 y B.3.5). Luego el resultado queda probado por el método de los momentos, ya que $X^{k+l}/T \to 0$ cuando $T \to +\infty$. □

Ahora sí, procedemos con la prueba del Teorema 2.3.4.

Demostración. Hasta que la Proposición 2.3.10 se use, esta prueba es esencialmente una variante del hecho que la convergencia en probabilidad implica la convergencia en ley, y que la convergencia en L^1 o L^2 implica la convergencia en probabilidad. Empecemos fijando una variable normal estándar $\mathcal{N}$. Sea $f : \mathbb{R} \to \mathbb{R}$ una función acotada Lipschitziana, y sea $C \geq 0$ un número real tal que

$$|f(x) - f(y)| \leq C|x - y|, \quad |f(x)| \leq C, \quad \textit{para } x, y \in \mathbb{R}.$$

Consideremos la diferencia

$$\left|\mathbb{E}_T\left(f\left(\frac{L_T}{\varrho_T}\right)\right) - \mathbb{E}(f(\mathcal{N}))\right|,$$

y debemos demostrar que tiende a 0 cuando $T \to +\infty$. Estimaremos esta cantidad mediante las aproximaciones que dimos arriba.

$$\begin{aligned}
&\left|\mathbb{E}_T\left(f\left(\frac{L_T}{\varrho_T}\right)\right) - \mathbb{E}(f(\mathcal{N}))\right| \\
&\quad \leq \mathbb{E}_T\left(\left|f\left(\frac{L_T}{\varrho_T}\right) - f\left(\frac{\tilde{L}_T}{\varrho_T}\right)\right|\right) + \mathbb{E}_T\left(\left|f\left(\frac{\tilde{L}_T}{\varrho_T}\right) - f\left(\frac{\log|D_T|^{-1}}{\varrho_T}\right)\right|\right) + \\
&\quad \mathbb{E}_T\left(\left|f\left(\frac{\log|D_T|^{-1}}{\varrho_T}\right) - f\left(\frac{\Re(P_T)}{\varrho_T}\right)\right|\right) + \left|\mathbb{E}_T\left(f\left(\frac{\Re(P_T)}{\varrho_T}\right)\right) - \mathbb{E}(f(\mathcal{N}))\right|,
\end{aligned} \tag{2.8}$$

y ahora discutiremos cada uno de los cuatro términos del miembro derecho de la desigualdad (tomaremos $|D_T|^{-1}$ como 0 cuando $D_T = 0$). El primer término lo podemos manejar fácilmente usando la Proposición 2.3.6, teniéndose que

$$\mathbb{E}_T\left(\left|f\left(\frac{L_T}{\varrho_T}\right) - f\left(\frac{\tilde{L}_T}{\varrho_T}\right)\right|\right) \leq \frac{C}{\varrho_T}\mathbb{E}_T\left(|L_T - \tilde{L}_T|\right) \to 0$$

cuando $T \to +\infty$. Para el segundo término, sea $A_T \subset \Omega_T$ el suceso

$$\{D_T = 0\} \cup \{|\tilde{L}_T - \log|D_T|^{-1}| > 1/2\},$$

y A_T' su complementario. Como $\log|Z_T| = \tilde{L}_T$, entonces tenemos que

$$\mathbb{E}_T\left(\left|f\left(\frac{\tilde{L}_T}{\varrho_T}\right) - f\left(\frac{\log|D_T|^{-1}}{\varrho_T}\right)\right|\right) \leq 2C\mathbb{P}_T(A_T) + \frac{C}{2\varrho_T}.$$

La Proposición 2.3.8 implica que $\mathbb{P}_T(A_T) \to 0$ (la convergencia a 1 de $Z_T D_T$ en L^2 implica la convergencia en probabilidad, equivalente a la convergencia a 0 en probabilidad para el logaritmo del módulo) y por tanto

$$\mathbb{E}_T\left(\left|f\left(\frac{\tilde{L}_T}{\varrho_T}\right) - f\left(\frac{\log|D_T|^{-1}}{\varrho_T}\right)\right|\right) \to 0$$

cuando $T \to +\infty$. Para el tercer miebro, distinguimos entre el suceso

$$B_T = \{|\log|D_T \exp(P_T|)| > 1/2\}$$

y su complementario, teniendo como antes

$$\mathbb{E}_T\left(\left|f\left(\frac{\log|D_T|^{-1}}{\varrho_T}\right) - f\left(\frac{\Re(P_T)}{\varrho_T}\right)\right|\right) \leq 2C\mathbb{P}_T(B_T) + \frac{C}{2\varrho_T},$$

que tiende a 0 cuando $T \to +\infty$ por la Proposición 2.3.9. Finalmente, la Proposición 2.3.10 implica que

$$\left|\mathbb{E}_T\left(f\left(\frac{\Re(P_T)}{\varrho_T}\right)\right) - \mathbb{E}(f(\mathcal{N}))\right| \to 0$$

cuando $T \to +\infty$, lo cual concluiye la prueba del teorema. □

Observación 2.3.2. Esta prueba fue propuesta en [26].

2.4. Aproximación polinómica de Dirichlet

El objetivo de esta sección es probar la Proposición 2.3.8, es decir, probaremos que

$$\mathbb{E}_T\left(|1 - Z_T D_T|^2\right)$$

converge a 0 cuando $T \to +\infty$. Para ello, en primer lugar usaremos la siguiente fórmula de aproximación

$$\zeta(\sigma_0 + it) = \sum_{1 \leq n \leq T} n^{-\sigma_0 - it} + O\left(\frac{T^{1-\sigma_0}}{|t|+1} + T^{-1/2}\right)$$

para $t \in \Omega_T$. Multiplicando por D_T, obtenemos

$$\mathbb{E}_T(Z_T D_T) = \sum_{\substack{m \geq 1 \\ n \leq T}} a_T(m)\mu(m)\mathbb{E}_T((mn)^{-\sigma_0}) + O\left(T^{1/2}\sum_{m \geq 1} a_T(m)\mathbb{E}_T((|t|+1)^{-1}) + T^{-1/2}\sum_{m \geq 1} a_T(m)m^{-\sigma_0}\right).$$

Recordemos que $|a_T(n)| \leq 1$ para todo n, y $a_T(n) = 0$ salvo que $n \ll T^{\varepsilon}$ para cualquier $\varepsilon > 0$. Por el Lema B.3.11, se tiene que

$$\begin{aligned}\mathbb{E}_T(Z_T D_T) &= 1 + O\left(\frac{1}{T}\sum_{\substack{n \leq T \\ m \neq n}} a_T(m)(mn)^{-\sigma_0}(\log mn)\right) + O(T^{-1/2+\varepsilon}) \\ &= 1 + O(T^{-1/2+\varepsilon})\end{aligned}$$

para cualquier $\varepsilon > 0$. Por lo tanto, basta probar que

$$\lim_{T \to +\infty} \mathbb{E}_T(|Z_T D_T|^2) = 1.$$

Expandimos el cuadrado medio usando la fórmula para D_T, obteniendo

$$\mathbb{E}_T(|Z_T D_T|^2) = \sum_{m,n} \frac{\mu(m)\mu(n)}{(mn)^{\sigma_0}} a_T(m)a_T(n)\mathbb{E}_T\left(\left(\frac{m}{n}\right)^{it}|Z_T|^2\right).$$

Por la Proposición C.3.5, obtenemos una fórmula asintótica para $\mathbb{E}_T\left(\left(\frac{m}{n}\right)^{it}|Z_T|^2\right)$, la cual es

$$\begin{aligned}\mathbb{E}_T\left(\left(\frac{m}{n}\right)^{it}|Z_T|^2\right) &= \zeta(2\sigma_0)\left(\frac{(m,n)^2}{mn}\right)^{\sigma_0} \\ &\quad + \zeta(2-2\sigma_0)\left(\frac{(m,n)^2}{mn}\right)^{1-\sigma_0}\mathbb{E}_T\left(\frac{|t|}{2\pi}\right)^{1-2\sigma_0} + O(\text{mín}(m,n)T^{-\sigma_0+\varepsilon}).\end{aligned}$$

para cualquier $\varepsilon > 0$, donde la esperanza es realmente la integral

$$\int_{-T}^{T}\left(\frac{|t|}{2\pi}\right)^{1-2\sigma} dt,$$

recordemos que (m,n) es el máximo común divisor de m y n. Por las propiedades de $a_T(n)$, el término de error se puede manejar sin demasiadas complicaciones, pues es a lo más

$$T^{-\sigma_0+\varepsilon}\sum_{m,n}(m,n)^{-\sigma_0}a_T(m)a_T(n)\min(m,n) \leq T^{-\sigma_0+\varepsilon}\left(\sum_m m^{1/2}a_T(m)\right)^2 \ll T^{-\sigma_0+2\varepsilon}$$

para cualquier $\varepsilon > 0$. Por lo tanto, solo nos es necesario manejar los términos principales, los cuales reescribimos como

$$\zeta(2\sigma_0)M_1 + \zeta(2-2\sigma_0)\mathbb{E}_T\left(\left(\frac{|t|}{2\pi}\right)^{1-2\sigma_0}\right)M_2, \tag{2.9}$$

donde

$$M_1 = \sum_{m,n}\frac{\mu(m)\mu(n)}{(mn)^{2\sigma_0}}a_T(m)a_T(n)(m,n)^{2\sigma_0}$$

y M_2 es el otro término. Por la estructura multiplicativa de $a_T(n)$, podemos expresar M_1 como $M_1 = M_1'M_1''$, donde

$$M_1' = \sum_{m,n}\frac{\mu(m)\mu(n)}{[m,n]^{2\sigma_0}}b_T(m)b_T(n),$$

$$M_1'' = \sum_{m,n}\frac{\mu(m)\mu(n)}{[m,n]^{2\sigma_0}}c_T(m)c_T(n).$$

Ahora, comparamos M_1' con una suma similar $\tilde{M}_1'$, donde las $b_T(m), b_T(n)$ son reemplazadas por funciones características de enteros cuyos factores primos son todos menores o iguales a Y, eliminando la restricción que supone tener a lo más m_1 factores primos distintos. Tenemos, por el Lema C.1.8

$$\tilde{M}_1 = \prod_{p\leq Y}\left(1-\frac{1}{p^{2\sigma_0}}\right).$$

La diferencia $M_1' - \tilde{M}_1'$ puede ser acotada superiormente por

$$2e^{-m1}\sum_{m,n}\frac{|\mu(m)\mu(n)|}{[m,n]^{2\sigma_0}}e^{\Omega(m)}$$

donde la suma recorre los enteros cuyos factores primos son todos menores o iguales que Y (este paso es un caso de lo que se conoce en inglés como *Rankin's trick*, la condición $\Omega(m) > m_1$ se obtiene acotando su función característica por la función no negativa $e^{\Omega(m)-m_1}$). Por el Lema C.1.8 y la Proposición C.2.5, es a lo más

$$2(\log T)^{-100}\prod_{p\leq Y}\left(1+\frac{1+2e}{p}\right) \ll (\log T)^{-90}.$$

Por lo tanto

$$M_1' \sim \prod_{p\leq Y}(1-p^{-2\sigma_0})$$

cuando $T \to +\infty$. De forma similar se trabaja con M_1'', llegando a

$$M_1'' \sim \prod_{Y<p\leq X} (1 - p^{-2\sigma_0}),$$

y por tanto

$$M_1 \sim \zeta(2\sigma_0) \prod_{p\leq X} (1 - p^{-2\sigma_0}) = \prod_{p>X} (1 - p^{-2\sigma_0}).$$

Ahora, por la elección anterior de parámetros, mediante el Teorema de los Números Primos (Teorema C.2.6) obtenemos la siguiente cota

$$\sum_{p>X} p^{-2\sigma_0} \ll \int_X^{+\infty} \frac{1}{t^{2\sigma_0}} \frac{dt}{\log t} \ll \frac{X^{1-2\sigma_0}}{(2\sigma_0 - 1)\log X} = \frac{X^{1-2\sigma_0}}{2\sqrt{W}} \leq \frac{1}{2\sqrt{W}}.$$

Dado que esto tiende a 0 cuando $T \to +\infty$, se sigue que

$$\prod_{p>X} (1 - p^{-2\sigma_0}) = \exp\left(\sum_{p>X} \left(\frac{1}{p^{2\sigma_0}} + O\left(\frac{1}{p^{4\sigma_0}}\right)\right)\right) = \exp\left(-\sum_{p>X} p^{-2\sigma_0}\right)(1 + o(1))$$

converge a 1 cuando $T \to +\infty$. Solamente falta comprobar que M_2 tiende a 0 cuando $T \to +\infty$. Tenemos

$$M_2 = \sum_{m,n} \frac{\mu(m)\mu(n)}{mn} a_T(m) a_T(n) (m,n)^{2-2\sigma_0} = \sum_{m,n} \frac{\mu(m)\mu(n)}{[m,n]^{2-2\sigma_0}} a_T(m) a_T(n) (mn)^{1-2\sigma_0}.$$

La forma de proceder es muy similar a la llevada a cabo para M_1: factorizamos $M_2 = M_2' M_2''$, donde M_2' tiene coeficientes b_T en vez de a_T, y M_2'' tiene coeficientes c_T en vez de a_T. Aplicando el Lema C.1.8 y el *Rankin's Trick* a ambos factores llegamos a

$$M_2 \sim \prod_{p\leq X} \left(1 + \frac{1}{p^{2-2\sigma_0}} \left(-\frac{1}{p^{2\sigma_0} - 1} - \frac{1}{p^{2\sigma_0 - 1}} + \frac{1}{p^{4\sigma_0 - 2}}\right)\right) = \prod_{p\leq X} \left(1 - \frac{2}{p} + \frac{1}{p^{2\sigma_0}}\right).$$

Deducimos que la contribución de M_2 a (2.9) es

$$\sim \zeta(2 - 2\sigma_0) \mathbb{E}_T \left(\left(\frac{|t|}{2\pi}\right)^{1-2\sigma_0}\right) \prod_{p\leq X} \left(1 - \frac{2}{p} + \frac{1}{p^{2\sigma_0}}\right).$$

Dado que $\zeta(s)$ tiene un polo en $s = 1$ con residuo 1, la última expresión cumple

$$\ll \frac{T^{1-2\sigma_0}}{2\sigma_0 - 1} \prod_{p\leq X} \left(1 - \frac{1}{p}\right) \ll \frac{T^{1-2\sigma_0}}{(2\sigma_0 - 1)\log X}.$$

En términos del parámetro W, como $2\sigma_0 - 1 = 2W/\log T$ y $X = T^{1/\sqrt{W}}$, el término de la derecha es simplemente $\exp(-2W) W^{-1/2}$, y esto tiende a 0 cuando $T \to +\infty$, lo cual concluye la prueba.

2.5. Aproximación por el producto de Euler

En esta sección buscaremos probar la Proposición 2.3.9. En otras palabras, necesitamos probar que $D_T \exp(P_T)$ converge a 1 en probabilidad. Para ello, descomponemos P_T como

$$P_T = Q_T + R_T$$

donde Q_T es la aportación a (2.3) de las potencias primas $p^k \leq Y$.

Además, para cada entero $m \geq 0$, se denota por $\exp_m$ al polinomio de Taylor centrado en 0 de grado m de la función exponencial, es decir

$$\exp_m(z) = \sum_{j=0}^{m} \frac{z^j}{j!}.$$

Ahora, presentaremos el siguiente lema.

Lema 2.5.1. *Sea $z \in \mathbb{C}$ y $m \geq 0$. Si $m \geq 100|z|$, entonces*

$$\exp_m(z) = e^z + O(\exp(-m)) = e^z(1 + O(\exp(-99|z|))).$$

Demostración. Como $j! \geq (j/e)^k$ para todo $j \geq 0$ y $|z| \leq m/100$, la diferencia $e^z - \exp_m(z)$ es a lo más

$$\sum_{j>m} \frac{(m/100)^j}{j!} \leq \sum_{j>m} \left(\frac{em}{100j}\right)^j \ll \exp(-m)$$

□

Definición 2.5.2. *Con las condiciones y parámetros de la Definición 2.3.7, definimos*

$$E_T := \exp_{m_1}(-Q_T), \quad F_T := \exp_{m_2}(-R_T).$$

Además, se tiene que $D_T = B_T C_T$, con

$$B_T(t) = \sum_{n\geq 1} b_T(n)\mu(n)n^{-\sigma_0 - it}, \quad C_T(t) = \sum_{n\geq 1} c_T(n)\mu(n)n^{-\sigma_0 - it},$$

donde las b_T, c_T fueron definidas en la Definición 2.3.7.

La idea de la prueba es que, al ser Q_T y R_T no muy grandes, E_T y F_T son buenas aproximaciones de $\exp(-Q_T)$ y $\exp(-R_T)$, respectivamente. Por otro lado, por la forma de E_T, es posible probar que E_T es muy parecido a B_T, y similar para F_T y C_T. Juntándolo todo, llegaremos a la conclusión.

En primer lugar, observamos que por el Lema B.3.11, tenemos que

$$\mathbb{E}_T(|Q_T|^2) \ll \varrho_T, \quad \mathbb{E}_T(|R_T|^2) \ll \log \varrho_T.$$

La desigualdad de Markov (ver Proposición B.3.12) implica que $\mathbb{P}_T(|Q_T| > \varrho_T)$ tiende a 0 cuando $T \to +\infty$. Ahora, por el Lema 2.5.1, cuando $|Q_T| \leq \varrho_T$, se cumple

$$E_T = \exp(-Q_T)\left(1 + O((\log T)^{-99})\right).$$

De forma similar, la probabilidad $\mathbb{P}_T(|R_T| > \log \varrho_T)$ tiende a 0, y cuando $|Q_T| \leq \log \varrho_T$, tenemos

$$E_T = \exp(-R_T)\left(1 + O((\log\log T)^{-99})\right).$$

Antes de seguir, introducimos otro lema.

Lema 2.5.3. *Sea $t \in \Omega_T$. Entonces*

$$E_T(t) = \sum_{n\geq 1} \alpha(n) n^{-\sigma_0+it},$$

donde los coeficientes $\alpha(n)$ son nulos salvo cuando $n \leq Y^{m_1}$ y n solo tiene factores primos menores o iguales que Y. Es más, $|\alpha(n)| \leq 1$ para todo n, y $\alpha(n) = \mu(n)b_T(n)$ si n tiene a lo más m_1 factores primos contados con multiplicidad, y si no hay potencia prima p^k que divida a n tal que $p^k > Y$.

Demostración. Dado que

$$E_T = \exp_{m_1}(-Q_T) = \sum_{j=0}^{m_1} \frac{(-1)^j}{j!}\left(\sum_{p^k\leq Y} \frac{1}{kp^{k(\sigma_0+it)}}\right)^j,$$

obtenemos, expandiendo la $j-ésima$ potencia, una expresión del tipo deseado, con coeficientes

$$\alpha(n) = \sum_{0\leq j\leq m_1} \frac{(-1)^j}{j!} \sum_{\substack{p_1^{k_1}\dots p_j^{k_j}=n \\ p_i^{k_i}\leq Y}} \frac{1}{k_1\dots k_j}.$$

Vemos por su expresión que $\alpha(n)$ es 0 salvo cuando $n \leq Y^{m_1}$ y n solo tiene factores primos menores o iguales que Y. Supongamos que n tiene un número de factores primos contados con multiplicidad menor o igual que m_1, y que ninguna potencia de primo $p^k > Y$ divide a n. Entonces podemos extender la suma definiendo $\alpha(n)$ para todo $j \geq 0$, y eliminar las condiciones $p_i^{k_i} \leq Y$, tal que

$$\alpha(n) = \sum_{j\geq 0} \frac{(-1)^j}{j!} \sum_{p_1^{k_1}\dots p_j^{k_j}=n} \frac{1}{k_1\dots k_j}.$$

Pero podemos ver que esos son los coeficientes de n^{-s} en la expansión de

$$\exp\left(-\sum_{k\geq 1}\frac{1}{k}p^{-ks}\right) = \exp(-\log\zeta(s)) = \frac{1}{\zeta(s)} = \sum_{n\geq 1}\frac{\mu(n)}{n^s},$$

restringiéndonos a $\Re(s) > 1$. Esto significa que, para esos enteros n, tenemos que $\alpha(n) = \mu(n) = \mu(n)b_T(n)$. Finalmente, para cualquier $n \geq 1$, tenemos que

$$|\alpha(n)| \leq \sum_{j\geq 0}\frac{1}{j!} \sum_{p_1^{k_1}\dots p_j^{k_j}=n} \frac{1}{k_1\dots k_j} = 1,$$

ya que el lado derecho de la igualdad se corresponde a los coeficientes de n^{-s} en $\exp(\log\zeta(s)) = \zeta(s)$. □

Para el siguiente paso, afirmamos que

$$\mathbb{E}_T\left(|E_T - B_T|^2\right) \ll (\log T)^{-60}, \tag{2.10}$$

$$\mathbb{E}_T\left(|F_T - C_T|^2\right) \ll (\log\log T)^{-60}. \tag{2.11}$$

Empezaremos probando la primera estimación. Para ello, definimos $\delta(n) := \alpha(n) - \mu(n)b_T(n)$ para todo $n \geq 1$. Tenemos, entonces

$$\mathbb{E}_T\left(|E_T - B_T|^2\right) = \mathbb{E}_T\left(\left|\sum_{n\geq 1}\frac{\delta(n)}{n^{\sigma_0+it}}\right|^2\right),$$

que estimamos usando el Lema B.3.22. La contribución de los términos fuera de la diagonal es $\ll \frac{1}{T}\sum_{m,n\leq Y^{m_1}}|\delta(m)\delta(n)|(mn)^{\frac{1}{2}-\sigma_0} \leq \frac{4}{T}\left(\sum_{m\leq Y^{m_1}}1\right)^2 \ll T^{-1+\varepsilon}$ para todo $\varepsilon > 0$, por lo que es despreciable. El término diagonal es

$$M = \sum_{n\geq 1}\frac{|\delta(n)|^2}{n^{2\sigma_0}} \leq \sum_{n\geq 1}\frac{|\delta(n)|^2}{n}.$$

Por el Lema 2.5.3, tenemos que $\delta(n)$ es 0, salvo que n tenga más de m_1 factores primos contados con multiplicidad, o es divisible por una potencia prima $p^k > Y$ (y necesariamente $p \leq Y$, ya que δ tiene soporte en los enteros solo divisibles por esos primos). La aportación de los enteros que satisfacen la primera propiedad es a lo más

$$\sum_{\substack{\Omega(n)>m_1\\ p|n\Rightarrow p\leq Y}}\frac{1}{n}.$$

Usaremos el *Rankin's trick* una vez más para acotar esto superiormente. Para cualquier número real fijo $\nu > 1$, tenemos

$$\sum_{\substack{\Omega(n)>m_1\\ p|n\Rightarrow p\leq Y}}\frac{1}{n} \leq \nu^{-m_1}\prod_{p\leq Y}\left(1 + \frac{\nu}{p} + \dots\right) \ll \nu^{-m_1}(\log Y)^{\nu} \leq (\log T)^{-100\log\nu+\nu},$$

por la Proposición C.2.5. Tomando $\nu = e^{2/3} \leq 2$, por ejemplo, tendríamos que la aportación es $\ll (\log T)^{-60}$. La aportación de los enteros divisibles por $p^k > Y$ es a lo más

$$\left(\sum_{\substack{p\leq Y\\ p^k>Y}}\frac{1}{p^k}\right)\left(\sum_{p|n\Rightarrow p\leq Y}\frac{1}{n}\right) \leq \frac{1}{Y}\left(\sum_{\substack{\sqrt{Y}<p^k\leq Y\\ k\geq 2}}1\right)\prod_{p\leq Y}\frac{1}{1-p^{-1}} \ll Y^{-1/2}(\log Y),$$

lo que es aún menor. Esto concluye la prueba de (2.10). La prueba de la segunda estimación es similar, con una consideración adicional con la que manejarse. De hecho, razonando como en el Lema 2.5.3, obtenemos la expresión

$$F_T(t) = \sum_{n\geq 1}\beta(n)n^{-\sigma_0+it},$$

para $t \in \Omega_T$, donde los coeficientes $\beta(n)$ son nulos salvo cuando $n \leq X^{m_2}$ y n solo tiene factores primos menores o iguales que X. Es más, $|\beta(n)| \leq 1$ para todo n, y $\beta(n) = \mu(n)c_T(n)$ si n tiene a lo más m_2 factores primos contados con multiplicidad, y si no hay potencia prima p^k que divida a n tal que $Y < p^k \leq X$. Usando esto, y definiendo ahora $\delta(n) := \beta(n) - \mu(n)c_T(n)$, tenemos por el Lema B.3.22 la siguiente cota

$$\mathbb{E}_T\left(|F_T - C_T|^2\right) \ll \sum_{\substack{n\geq 1\\ \delta(n)\neq 0}} \frac{1}{n^{2\sigma_0}} \leq \sum_{\substack{n\geq 1\\ \delta(n)\neq 0}} \frac{1}{n}.$$

Pero los enteros que satisfacen $\delta(n) \neq 0$ son necesariamente de uno de los siguientes tipos.

(1) Aquellos que cumplen $c_T(n) = 1$, los cuales o bien satisfacen $\Omega(n) > m_2$, o bienson divisibles por una potencia prima $p^k > X$. La aportación de estos enteros se maneja de manera similar a (2.10), y es $\ll (\log\log T)^{-60}$.

(2) Aquellos con $c_T(n) = 0$ pero $\beta(n) \neq 0$, ya que

$$\beta(n) = \sum_{0\leq j\leq m_2} \frac{(-1)^j}{j!} \sum_{\substack{p_1^{k_1}\dots p_j^{k_j}=n\\ Y<p_ik_i\leq X}} \frac{1}{k_1\dots k_j},$$

tal que cada entero n tiene al menos una factorización $n = p_1^{k_1}\dots p_j^{k_j}$ para algún $j \leq m_2$, donde cada potencia prima está entre Y y X. Como $c_T(n) = 0$, o bien $\Omega(n) > m_2$, o bien n tiene algún factor primo mayor que X, o bien n tiene un factor primo menor o igual que Y. Las primeras dos posibilidades se trabajan igual que en (2.10), pero la tercera es diferente. Llamemos a su aportación N. Tenemos

$$N = \sum_{0\leq j\leq m_2} N_j,$$

donde

$$N_j = \sum_{p\leq Y} \sum_{\substack{p^k p_1^{k_1}\dots p_j^{k_j}=n\\ Y<p_ik_i\leq X}} \frac{1}{n}$$

es la contribución de los enteros con factorización de longitud $j+1$ en producto de potencias de primos situados entre Y y X. Por multiplicatividad, tenemos que

$$N_J \leq \left(\sum_{p\leq Y}\sum_{Y<p^k\leq X}\frac{1}{p^k}\right)\left(\sum_{p\leq X}\sum_{Y<p^k\leq X}\frac{1}{p^k}\right)^{j-1}.$$

Consideremos el primer factor. Para un primo dado $p \leq Y$, sea l el menor entero tal que $p^l > Y$. La suma sobre k es entonces

$$\sum_{Y<p^k\leq X}\frac{1}{p^k} \leq \frac{1}{p^l} + \frac{1}{p^{l+1}} + \dots \ll \frac{1}{p^l} \leq \frac{1}{Y},$$

por lo que el primer factor es $\ll \pi(Y)/Y \ll (\log Y)^{-1}$. Por otro lado, para el segundo factor tenemos

$$\sum_{p\leq X}\sum_{Y<p^k\leq X}\frac{1}{p^k} = \sum_{p\leq Y}\sum_{Y<p^k\leq X}\frac{1}{p^k} + \sum_{Y<p\leq X}\sum_{Y<p^k\leq X}\frac{1}{p^k} \ll \frac{\pi(Y)}{Y} + \sum_{Y<p\leq X}\sum_{Y<p^k\leq X},$$

donde hemos usado la cota del primer factor. Para un pripo $Y < p \leq X$, la última suma sobre k es

$$\frac{1}{p} + \frac{1}{p^2} + \cdots \ll \frac{1}{p},$$

y, por tanto, la suma sobre p eso

$$\sum_{Y<p\leq X} \frac{1}{p} = \log\left(\frac{\log X}{\log Y}\right) + O(1) = \log\log\log T + O(1),$$

usando los valores de X e Y y el Teorema C.2.3. Por tanto, la estimación final es

$$N \ll \frac{1}{\log Y}(\log\log\log T)^{m_2} \ll (\log\log T)(\log\log\log T)^{m_2}(\log T)^{-1} \to 0$$

cuando $T \to +\infty$, por lo que (2.11) es cierta.

Con (2.10) y (2.11) probadas, podemos terminar la prueba de la Proposición 2.3.9. Excepto para conjuntos de medida que tiende a 0 cuando $T \to +\infty$, tenemos

$$B_T = E_T + O((\log T)^{-25}), \quad E_T = \exp(-Q_T)\left(1 + O((\log T)^{-99})\right)$$

$$\frac{1}{\log T} \leq \exp(-Q_T) \leq (\log T)$$

(donde la primera propiedad se deduce de (2.10)), y por tanto

$$B_T = \exp(-Q_T)\left(1 + O((\log T)^{-20})\right)$$

de nuevo fuera de un conjunto de medida que tiende a 0. De forma similar, usando (2.11), tenemos

$$C_T = \exp(-R_T)\left(1 + O((\log T)^{-20})\right)$$

fuera de un conjunto de medida que tiende a 0. Multiplicando ambas igualdades obtenemos

$$D_T = \exp(-P_T)\left(1 + O((\log T)^{-20})\right)$$

con probabilidad tendiendo a 1 cuando $T \to +\infty$, lo que concluye la prueba.

2.6. Interpretación probabilística de la Hipótesis de Riemann

Una de las razones por las que la Hipótesis de Riemann (ver Conjetura C.3.4) es tan compleja es que no hay un argumento de plausibilidad, aunque sea poco riguroso, que indique que la Hipótesis de Riemann tenga que ser cierta. Este hecho da cierta importancia a la interpretación probabilística de Denjoy, que aunque pueda parecer

Figura 2.5: Arnaud Denjoy.

absurda si se considera con cuidado, arroja algo de luz en la plausibilidad de la Hipótesis de Riemann.

Supongamos que tenemos una moneda perfecta (sin trucar), y la lanzamos N veces, con N suficientemente grande. Por el Teorema B.3.13, la probabilidad de que el número de caras que salgan se desvíe del valor esperado por más de $KN^{1/2}$ es similar a

$$\int_{-(2K^2/\pi)^{1/2}}^{(2K^2/\pi)^{1/2}} \exp(-\pi x^2)\, dx,$$

es decir, esta integral coincide con el límite de la probabilidad anterior cuando $N \to +\infty$. Luego si restamos el número de caras con el de cruces, la probabilidad de que ese resultado sea menor que $2KN^{1/2}$ en valor absoluto es similar a

$$2\int_{0}^{(2K^2/\pi)^{1/2}} \exp(-\pi x^2)\, dx,$$

y por tanto la probabilidad de que sea menor que $N^{1/2+\varepsilon}$ para algún $\varepsilon > 0$ fijado es aproximadamente

$$2\int_{0}^{N^{\varepsilon}(2\pi)^{1/2}} \exp(-\pi x^2)\, dx.$$

El hecho de que esta cantidad se aproxime a 1 se puede expresar como *con probabilidad 1, el número de caras menos el número de cruces crece menos rápido que* $N^{1/2+\varepsilon}$.

Consideremos ahora un entero n suficientemente grande libre de cuadrados. Entonces $\mu(n) = \pm 1$. Es quizás razonable decir que $\mu(n)$ será más o menos uno *con la misma probabilidad*, pues n tendrá normalmente un gran número de factores (la densidad de los primos $1/\log x$ se acerca a 0) y no parece haber razón alguna para sospechar que es más probable que n tenga un número par o impar de factores primos. Es más, por el mismo argumento, es quizás razonable pensar que sucesivas evaluaciones de $\mu(n)$ son *independientes*, pues saber el valor de $\mu(n)$ no da información de los valores de μ para otros n distintos. Pero entonces, la evaluación de $M(x) := \sum \mu(n)D(x/n)$, donde $D(x)$ vale 1 si $x > 1$, $1/2$ *si* $x = 1$ y 0 en otro caso, seria como lanzar una moneda por cada entero libre de cuadrados menor o igual que x y restar el número de caras

y cruces. Como hemos visto anteriormente, para cualquier $\varepsilon > 0$, el resultado de este experimento para un N suficientemente grande es, con probabilidad casi 1, menor que el número de tiradas elevado a $\frac{1}{2} + \varepsilon$ y *a fortiori* menor que $x^{1/2+\varepsilon}$. Luego estas asunciones sobre los valores de $\mu(n)$ llevan a la conclusión, aparentemente ridícula, que $M(x) = O(x^{1/2+\varepsilon})$ con probabilidad 1, y por tanto, la Hipótesis de Riemann es cierta con probabilidad 1. Esto se debe a que la Conjetura $M(x) = O(x^{1/2+\varepsilon})$ es una formulación equivalente a la Hipótesis de Riemann.

Para más información, consultar [3], capítulos 11 y 12.

3. Desviación o sesgo de Chebyshev

> Las preguntas más importantes de la vida, de hecho, no son en su mayoría más que problemas de probabilidad.
>
> *Pierre Simon Laplace*

3.1. Introducción

En este capítulo presentaremos una introducción al estudio del sesgo de Chebyshev, una rama de la Teoría de Números que trata la diferencia entre las cantidades de primos p tales que $p \equiv 3(mod\ 4)$ o $p \equiv 1(mod\ 4)$. Concretamente, Chebyshev observó que existen más primos del primer tipo que del segundo. Desgraciadamente, solamente publicó una pequeña nota de la que no se puede saber qué probó, o simplemente conjeturó, ya que no volvió a publicar nada referente al tema. Sin embargo, si denotamos por $\pi(x; q, a)$ a la cantidad de primos p menores o iguales que x tales que $p \equiv a(mod\ q)$, para $x = 26861$, tenemos

$$\pi(x; 4, 3) = 1472 < 1473 = \pi(x; 4, 1),$$

hecho que fue observado por Leech en 1957. De hecho, se ha probado que existen infinitos cambios de signo en la diferencia $\pi(x; 4, 3) - \pi(x; 4, 1)$.

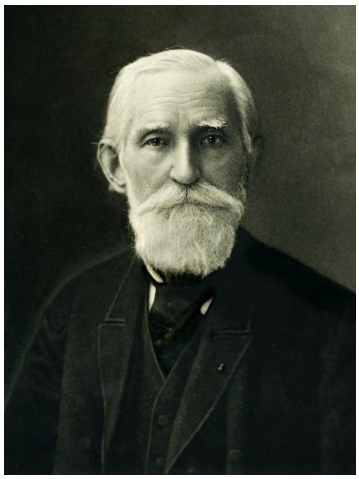


Figura 3.1: Pafnuti Chebyshev, célebre matemático conocido por sus notables aportaciones y por la cantidad de diferentes formas existentes de escribir su apellido.

Es por ello que, en este capítulo, definiremos la distribución de Rubinstein-Sarnak, además de probar su existencia. Discutiremos, a su vez, la hipótesis de simplicidad generalizada, junto con algunos resultados relacionados. Por último, mostraremos algunos resultados que son fruto de la aplicación de los teoremas anteriores.

3.2. Distribución de Rubinstein-Sarnak: definición y existencia

Para poder estudiar el problema anterior, consideramos para $X \geq 1$ el espacio probabilístico $\Omega_X = [1, X]$, con la medida de probabilidad

$$\mathbb{P}_X = \frac{1}{\log X}\frac{dx}{x}. \tag{3.1}$$

Definimos ahora las variables aleatorias que intervendrán en esta sección.

Definición 3.2.1. *Sea $q \geq 1$. Definimos una variable aleatoria en Ω_X, con valores en el espacio vectorial $C_{\mathbb{R}}((\mathbb{Z}/q\mathbb{Z})^\times)$ de funciones reales en el conjunto de elementos invertibles de $\mathbb{Z}/q\mathbb{Z}$, definiendo $N_{X,q}(x)$, para $x \in \Omega_X$, como la función*

$$N_{X,q}(x)(a) := \frac{\log x}{\sqrt{x}}\big(\varphi(q)\pi(x; q, a) - \pi(x)\big), \tag{3.2}$$

para $a \in (\mathbb{Z}/q\mathbb{Z})^\times$, donde φ es la función totient de Euler (ver Anexo C).

Vemos que cualquier información que tengamos de $N_{X,q}$ nos será de utilidad para comparar la cantidad de primos menores o iguales que X en cualquier familia de clases residuo invertibles módulo q.

A partir de ahora, consideraremos q como fijo, por lo que se simplificará la notación de $N_{X,q}$ a N_X, y similar en el resto de casos que involucren a q.

Observación 3.2.1.

1. Si $q = 4$, entonces $(\mathbb{Z}/q\mathbb{Z})^\times = \{1, 3\}$, por lo que, para $x \in \Omega_X$ la función $N_{X,4}$ viene dada por

$$1 \mapsto \frac{\log x}{\sqrt{x}}\big(\varphi(4)\pi(x; 4, 1) - \pi(x)\big), \quad 3 \mapsto \frac{\log x}{\sqrt{x}}\big(\varphi(4)\pi(x; 4, 3) - \pi(x)\big).$$

2. Por el Teorema C.2.7, se tiene que

$$\pi(x; q, a) \sim \frac{1}{\varphi(q)}\pi(x).$$

3. El factor de normalización $\log x/\sqrt{x}$ es el que fue propuesto por el propio Chebyshev.

Ahora sí, estamos en condiciones de enunciar el Teorema clave de esta sección.

Teorema 3.2.2 (Rubinstein-Sarnak)**.** *Sea $q \geq 1$. Asumamos la Hipótesis de Riemann generalizada módulo q (ver Conjetura C.4.2). Entonces, las funciones aleatorias $N_{X,q}$ convergen en ley a la función aleatoria N_q, cuyo soporte está contenido en el hiperplano*

$$H_q = \left\{ f \colon (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{R} \ : \sum_{a \in (\mathbb{Z}/q\mathbb{Z})^\times} f(a) = 0 \right\}. \tag{3.3}$$

A N_q se le denomina distribución de Rubinstein-Sarnak módulo q.

Chebyshev's Bias

Michael Rubinstein and Peter Sarnak

CONTENTS

The title refers to the fact, noted by Chebyshev in 1853, that primes congruent to 3 modulo 4 seem to predominate over those congruent to 1. We study this phenomenon and its generalizations. Assuming the Generalized Riemann Hypothesis and the Grand Simplicity Hypothesis (about the zeros of the Dirichlet L-function), we can characterize exactly those moduli and residue classes for which the bias is present. We also give results of numerical investigations on the prevalence of the bias for several moduli. Finally, we briefly discuss generalizations of the bias to the distribution to primes in ideal classes in number fields, and to prime geodesics in homology classes on hyperbolic surfaces.

1. INTRODUCTION

Dirichlet [1837] proved that for any a and q with $(a,q) = 1$ there are infinitely many primes p with $p \equiv a \bmod q$, and that they are roughly equidistributed amongst these residue classes. We denote the set of such residue classes by A_q. It was later proved by Hadamard and de la Vallée Poussin that the number $\pi(x,q,a)$ of primes $p \le x$ with $p \equiv a \bmod q$ has the behavior

$$\pi(x,q,a) \sim \frac{\mathrm{Li}(x)}{\varphi(q)} \sim \frac{1}{\varphi(q)} \frac{x}{\log x}$$

as $x \to \infty$, where $\varphi(q) = |A_q|$ is the Euler phi function and

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t}.$$

Chebyshev noted in 1853 that there are many more primes congruent to 3 than 1 modulo 4. Much has been written about this since then, but we have found the literature to be a little confused and inaccurate. We have, therefore, tried our best to cite below the original sources where appropriate. A

1058-6458/94 $0.50 per page

Figura 3.2: Primera página del artículo original de Rubinstein y Sarnak, donde presentaron el resultado homónimo.

La prueba de este resultado, la cual haremos desgranándolo en resultados más simples de demostrar, depende de dos cosas:

1. Podemos representar las funciones aleatorias aritméticas N_X como combinaciones de $x \mapsto x^{i\gamma}$, donde las γ son las ordenadas de los ceros de las L-funciones módulo q.
2. Una vez lo tengamos hecho, el Teorema B.3.14 implica la convergencia en ley de cualquier función de este tipo.

Para un caracter de Dirichlet χ módulo q, definimos las variables aleatorias ψ_X en Ω_X como

$$\psi_\chi(x) := \frac{1}{\sqrt{x}} \sum_{n \leq x} \Lambda(n)\chi(x)$$

para $x \in \Omega_X$, donde $\Lambda(n)$ es la función de von Mangoldt.

(a) Michael Rubinstein

(b) Peter Sarnak

Figura 3.3: Los dos matemáticos que propusieron la distribución que trata esta sección

El siguiente lema es un paso clave para expresar N_X en términos de caracteres de Dirichlet.

Lema 3.2.3. *Se tiene que*

$$N_{X,q} = m_q + \sum_{\chi \pmod q} \psi_X \bar{\chi} + E_{X,q},$$

donde el sumatorio recorre los caracteres de Dirichlet no triviales módulo q y $E_{X,q}$ converge a 0 en probabilidad cuando $X \to +\infty$

Demostración. Por la ortogonalidad de los caracteres de Dirichlet módulo q (Proposición C.4.3), tenemos que

$$\varphi(q)\pi(x; q, a) = \sum_{\chi \pmod q} \overline{\chi(a)} \sum_{p \leq x} \chi(p),$$

por lo que

$$\frac{\log x}{\sqrt{x}}\big(\varphi(q)\pi(x;q,a)-\pi(x)\big)=\sum_{\substack{\chi \pmod q\\ \chi \text{ no trivial}}}\overline{\chi(a)}\frac{\log x}{\sqrt{x}}\sum_{p\le x}\chi(p)+O\bigg(\frac{\log x}{\sqrt{x}}\bigg)$$

para $x\ge 2$, donde el término de error cuenta en los primos p que dividen a q. Ahora necesitamos encontrar la conexión de la suma anterior con ψ_X. Recordemos que la función de von Mangoldt difiere poco de la función característica de los primos multiplicada por la función logaritmo. La suma de esta función es la variable aleatoria definida por

$$\theta_\chi(x):=\frac{1}{\sqrt{x}}\sum_{p\le x}\chi(p)\log(p)$$

para $x\in\Omega_X$. Se relaciona a ψ_χ mediante

$$\theta_\chi(x)-\psi_\chi(x)=-\frac{1}{\sqrt{x}}\sum_{k\ge 2}\sum_{p^k\le x}\chi(p^k)\log(p)=-\frac{1}{\sqrt{x}}\sum_{k\ge 2}\sum_{p^k\le x}\chi(p)^k\log(p).$$

Podemos ver que la aportación de $k\ge 3$ es muy escasa, pues el exponente k es a lo más de tamaño $\log x$, y $|\chi(p)|\le 1$ para todo primo p, por lo que está acotada por

$$\bigg|\frac{1}{\sqrt{x}}\sum_{\substack{k\ge 2\\ k\ge 3}}\sum_{p^k\le x}\chi(p)^k\log(p)\bigg|\le\frac{1}{\sqrt{x}}\sum_{3\le k\le\log x}\log x(x)^{1/k}\ll\frac{(\log x)^2}{x^{1/6}},$$

donde la constante implícita es absoluta.

Para $k=2$, tenemos dos casos. Si χ^2 es el caracter trivial, entonces

$$\frac{1}{\sqrt{x}}\sum_{p\le\sqrt{x}}\chi(p)^2\log(p)=\frac{1}{\sqrt{x}}\sum_{\substack{p\le\sqrt{x}\\ p\mid q}}\log(p)=1+O\bigg(\frac{1}{\log x}\bigg)$$

por el Teorema de los Números Primos en progresiones aritméticas. Si χ^2 es no trivial, entonces, por la misma razón tenemos

$$\frac{1}{\sqrt{x}}\sum_{p\le\sqrt{x}}\chi(p)^2\log(p)\ll\frac{1}{\log x}.$$

Por lo tanto, se tiene que

$$\theta_\chi(x)=\psi_\chi(x)-\delta_{\chi^2}+O\bigg(\frac{1}{\log x}\bigg)\tag{3.4}$$

donde δ_{χ^2} es 1 si χ^2 es trivial, y 0 en cualquier otro caso. Por sumación por partes tenemos

$$\sum_{p\le x}\chi(p)=\frac{1}{\log x}\sum_{p\le x}\chi(p)\log p+\int_2^x\bigg(\sum_{p\le t}\chi(p)\bigg)\frac{dt}{t(\log t)^2}$$

para cualquier caracter de Dirichlet χ módulo q, por lo que

$$\frac{\log x}{\sqrt{x}}\big(\varphi(q)\pi(x;q,a)-\pi(x)\big)=\sum_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}}\overline{\chi(a)}\theta_\chi(x)+\frac{\log x}{\sqrt{x}}\int_2^x\frac{\theta_\chi(t)}{t^{1/2}(\log t)^2}\,dt+O\bigg(\frac{\log x}{\sqrt{x}}\bigg). \tag{3.5}$$

Empezamos resolviendo la integral por un caracter no trivial χ. Tenemos que $\theta_\chi(x)=\psi_\chi(x)+O(1/\log x)$ si $\chi^2\neq 1_q$, lo que implica que

$$\int_2^x\frac{\theta_\chi(t)}{t^{1/2}(\log t)^2}\,dt=\int_2^x\frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}\,dt+O\bigg(\frac{x^{1/2}}{(\log x)^3}\bigg)$$

ya que

$$\int_2^x\frac{1}{t^{1/2}(\log t)^2}\,dt\ll\frac{x^{1/2}}{(\log x)^2}.$$

Si χ^2 es trivial, tenemos un término adicional constante $\theta_\chi(x)-\psi_\chi(x)=1+O(1/\log x)$, y llegamos a

$$\begin{aligned}\int_2^x\frac{\theta_\chi(t)}{t^{1/2}(\log t)^2}\,dt&=\int_2^x\frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}\,dt+\int_2^x\frac{1}{t^{1/2}(\log t)^2}\,dt+O\bigg(\frac{x^{1/2}}{(\log x)^3}\bigg)\\&=\int_2^x\frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}\,dt+O\bigg(\frac{x^{1/2}}{(\log x)^2}\bigg).\end{aligned}$$

En cualquiera de los casos anteriores, tenemos que

$$\frac{\log x}{\sqrt{x}}\int_2^x\frac{\theta_\chi(t)}{t^{1/2}(\log t)^2}\,dt=\frac{\log x}{\sqrt{x}}\int_2^x\frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}\,dt+O\bigg(\frac{1}{\log x}\bigg).$$

Si usamos la cota $\psi_\chi(x)\ll(\log x)^2$ en la integral restante, obtenemos

$$\frac{\log x}{\sqrt{x}}\int_2^x\frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}\,dt\ll\log x,$$

que es excesivamente grande, por lo que tenemos que usar la integración de forma no trivial. Por el Corolario C.4.4, tenemos

$$\int_2^x\psi_\chi(t)\,dt\ll x$$

para todo $x\geq 2$. Usando integración por partes deducimos que

$$\frac{\log x}{\sqrt{x}}\int_2^x\frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}\,dt\ll\frac{\log x}{\sqrt{x}}\bigg(\frac{x}{x^{1/2}(\log x)^2}+\int_2^x\frac{t^{1/2}}{(\log t)^2}\,dt\bigg)\ll\frac{1}{\log x}.$$

Finalmente, transformamos el primer término de (3.5) para expresarlo en términos de $\psi_\chi(x)$, como en (3.4). Para cualquier elemento $a \in (\mathbb{Z}/q\mathbb{Z})^\times$ y $x \in \Omega_X$, tenemos que

$$\begin{aligned}\frac{\log x}{\sqrt{x}}\big(\varphi(q)\pi(x;q,a) - \pi(x)\big) &= -\sum_{\substack{\chi^2=1\\ \chi\neq 1}} \overline{\chi(a)} + \sum_{\substack{\chi \pmod q\\ \chi \text{ no trivial}}} \overline{\chi(a)}\psi_\chi(x) + O\left(\frac{1}{\log x}\right)\\ &= m_q(a) + \sum_{\substack{\chi \pmod q\\ \chi \text{ no trivial}}} \overline{\chi(a)}\psi_\chi(x) + O\left(\frac{1}{\log x}\right)\end{aligned}$$

donde la constante implícita depende de q. Como el término de error es $\ll (\log x)^{-1}$ para $x \in \Omega_X$, converge a 0 en probabilidad, lo que concluye la prueba. □

Como E_X tiende a 0 en probabilidad y m_q es una función fija, por el Corolario B.3.15 se tiene que el Teorema 3.2.2 se tiene de la convergencia en ley de las funciones aleatorias

$$M_{X,q} = \sum_{\substack{\chi \pmod q\\ \chi \text{ no trivial}}} \overline{\chi(a)}\psi_\chi(x).$$

Ahora expresamos estas funciones en términos de ceros de L-funciones. Como apunte, en las sumas que los involucren, los ceros son contados con multiplicidad.

Denotamos por I_X a la variable identidad, $x \mapsto x$ en Ω_X, por lo que para un número complejo s, la variable aleatoria I_X^s es la función $x \mapsto x^s$ en Ω_X. Además, cuando tenemos una función aleatoria X en $(\mathbb{Z}/q\mathbb{Z})^\times$, y una variable aleatoria no negativa Y, el significado de la expresión $X = O(Y)$ es que $\|X\| = O(Y)$, donde la norma es la norma euclídea, es decir

$$\|X\|^2 = \sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times} \Big|X(a)\Big|^2.$$

Lema 3.2.4. *Tenemos que*

$$M_{X,q} = -\sum_{\substack{\chi \pmod q\\ \chi \text{ no trivial}}} \left(\sum_{|\gamma|\le X} \frac{I_X^{i\gamma}}{\frac{1}{2}+i\gamma}\right)\overline{\chi} + O\left(\frac{(\log X)^2}{X^{1/2}}\right)$$

donde γ se sitúa entre las ordenadas de los ceros de $L(s,\chi)$, contados con multiplicidad, y la constante implícita depende de q.

Demostración. El ingrediente principal es la fórmula explícita de la Teoría de los Números Primos, que se puede expresar como

$$\psi_\chi = -\sum_{\substack{L(\beta+i\gamma)=0\\ |\gamma|\le X}} \frac{I_X^{\beta-\frac{1}{2}+i\gamma}}{\beta+i\gamma} + O\left(\frac{I_X^{1/2}(\log X)^2}{X}\right)$$

donde la suma es sobre los ceros de las L-funciones de Dirichlet con $0 \le \beta \le 1$, contados con multiplicidad (ver Teorema C.4.5). Bajo la suposición de la Hipótesis de Riemann

Generalizada módulo q, siempre tenemos que $\beta = \frac{1}{2}$, y esta fórmula implica que

$$\psi_\chi = -\sum_{|\gamma|\leq X} \frac{I_X^{i\gamma}}{\frac{1}{2}+i\gamma} + O\left(\frac{(\log X)^2}{X^{1/2}}\right)$$

con la suma recorriendo los caracteres (cuyo número es $\varphi(q) - 1 \leq q$), se tiene la fórmula. □

Probabilísticamente, tenemos una combinación lineal finita de variables aleatorias $I_X^{i\gamma}$. El nexo con la Teoría de la Probabilidad y la existencia de la distribución de Rubinstein-Sarnak viene dado por la siguiente Proposición.

Proposición 3.2.5. *Sea $k \geq 1$ un entero. Sea F un conjunto finito de números reales y sea $(\alpha(t))_{t\in F}$ una familia de elementos de $\mathbb{C}^k$. Los vectores aleatorios*

$$\sum_{t\in F} I_X^{it}\alpha(t)$$

en Ω_X converge en ley cuando $X \to +\infty$.

Demostración. Después de aplicar una pequeña traslación, esto es una consecuencia directa del Teorema B.3.14. De hecho, consideremos el vector

$$z = \left(\frac{t}{2\pi}\right)_{t\in F} \in \mathbb{R}^F.$$

Por el Teorema B.3.14, las medidas de probabilidad μ_Y en $(\mathbb{R}/\mathbb{Z})^F$ definidas para $Y > 0$ como

$$\mu_Y(A) = \frac{1}{Y}\Big|\{y \in [0,Y] \ : \ yz \in A\}\Big|,$$

para cualquier conjunto medible A, converge en ley a una medida de probabilidad de Haar μ en el subgrupo T de $(\mathbb{R}/\mathbb{Z})^F$ generado por las clases módulo $\mathbb{Z}^F$ de los elementos yz, donde y se mueve en $\mathbb{R}$.

Extendemos el isomorfismo $\theta \mapsto e(\theta)$ de $\mathbb{R}/\mathbb{Z}$ en $\mathbf{S}^1$ por componentes para definir un isomorfismo de $(\mathbb{R}/\mathbb{Z})^F$ en $(\mathbf{S}^1)^F$. Para cualquier función continua f en $(\mathbf{S}^1)^F$, vemos que

$$\begin{aligned}\int_{(\mathbb{R}/\mathbb{Z})^F} f(e(v))\, d\mu_Y(v) &= \frac{1}{Y}\int_0^Y f(e(yz))\, dy \\ &= \frac{1}{Y}\int_0^Y f((e^{ity})_{t\in F})\, dy \\ &= \frac{1}{Y}\int_0^{e^Y} f((x^{it})_{t\in F})\frac{dx}{x} \\ &= \mathbb{E}_X(f((I_X^{it})_{t\in F}))\end{aligned}$$

para $X = e^Y$, después del cambio de variable $x = e^y$. Luego el vector $(I_X^{it})_{t\in F}$ converge en ley cuando $X \to +\infty$ a la imagen de μ por $v \mapsto e(v)$. Se acaba la demostración por composición con la aplicación continua de $(\mathbf{S}^1)^F$ en $\mathbb{C}^k$ definida por

$$(z_t)_{t\in F} \mapsto \sum_{t\in F} z_t\alpha(t),$$

aplicando la Proposición B.3.16. □

De esta prueba, podemos mejorar el resultado anterior.

Corolario 3.2.6. *Con la notación y suposiciones de la Proposición 3.2.5, los vectores aleatorios*

$$\sum_{t\in F} I_X^{it}\alpha(t)$$

convergen en ley cuando $X \to +\infty$ a

$$\sum_{t\in F} I_t\alpha(t),$$

donde $(I_t)_{t\in F}$ es una variable aleatoria con valores en $(\mathbf{S}^1)^F$ con distribución dada por la medida de probabilidad de Haar de la clausura del subgrupo de $(\mathbf{S}^1)^F$ generado por todos los elementos $(x^{it})_{t\in F}$ para $x \in \mathbb{R}$.

Sea $T \geq 2$ un parámetro. Del Lema 3.2.4 y la Proposición 3.2.5 se deduce que, para $X \geq T$, se tiene

$$M_{X,q} = N_{X,T,q} + \sum_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}} \Big(\sum_{T\leq|\gamma|\leq X} \frac{I_X^{i\gamma}}{\frac{1}{2}+i\gamma}\Big)\overline{\chi} + O\Big(\frac{(\log X)^2}{\sqrt{(X)}}\Big), \tag{3.6}$$

donde

$$N_{X,T,q} = -\sum_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}} \Big(\sum_{|\gamma|\leq T} \frac{I_X^{i\gamma}}{\frac{1}{2}+i\gamma}\Big)\overline{\chi}$$

son funciones aleatorias que convergen en ley cuando $X \to +\infty$ para cualquier $T \geq 2$ fijo. El siguiente lema nos ayudará a comprobar que el término restante en esta aproximación es pequeño.

Lema 3.2.7. *Sea $k \geq 1$ un entero. Sea F un conjunto numerable de números reales y sea $(\alpha(t))_{t\in F}$ una familia de elementos de $\mathbb{C}^k$. Supongamos que las siguientes condiciones se cumplen para cualesquiera $T \geq 2$ y $t_0 \in \mathbb{R}$:*

$$\sum_{t\in F} \|\alpha(t)\|^2 |t|^{1/2}\log(1+|t|) < +\infty, \tag{3.7}$$

$$\sum_{\substack{t\in F\\ |t|\geq T}} \frac{\|\alpha(t)\|}{t^{1/4}} \ll \frac{(\log T)^4}{T^{1/4}}, \tag{3.8}$$

$$\big|\{t\in F \ : \ |t-t_0|\leq 1\}\big| \ll \log(1+|t_0|). \tag{3.9}$$

Entonces tenemos que

$$\lim_{\substack{T\leq X\\ T\to+\infty}} \Big\| \sum_{\substack{t\in F\\ |t|\geq T}} I_X^{it}\alpha(t)\Big\|_{L^2} = 0,$$

donde el límite es sobre las parejas (T, X) tales que $T \leq X$ y T tiende a infinito.

En este aserto, usamos el espacio de Hilbert $L^2(\Omega_X;\mathbb{R}^k)$ de funciones $f\colon \Omega_X \to \mathbb{R}^k$, con norma definida por

$$\|f\|_{L^2}^2 := \mathbb{E}_X(\|f\|^2)$$

para $f \in L^2(\Omega_X;\mathbb{R}^k)$.

Demostración. Destacamos, en primer lugar, que la computación exacta de la integral da

$$\mathbb{E}_X(I_X^{i(t_1-t_2)}) = \frac{1}{\log X}\frac{X^{i(t_1-t_2)-1}}{t_1-t_2}$$

para $t_1 \neq t_2$, tenemos la cota general

$$|\mathbb{E}_X(I_X^{i(t_1-t_2)})| \leq \min\left(1, \frac{1}{\log X}\frac{2}{|t_1-t_2|}\right). \tag{3.10}$$

Usaremos esta cota para ganar algo de flexibilidad. Todas las sumas que involucran a t, t_1 y t_2 están restringidas a $t \in F$. Supongamos que $2^5 \leq T \leq X$. Tenemos

$$\left\| \sum_{\substack{t\in F\\ |t|\geq T}} I_X^{it}\alpha(t) \right\|_{L^2} = \mathbb{E}_X\left(\left\| \sum_{T\leq|t|\leq X} I_X^{it}\alpha(t) \right\|^2\right) = \sum_{T\leq|t_1|,|t_2|\leq X} \alpha(t_1)\cdot\alpha(t_2)\mathbb{E}_X(I_X^{i(t_1-t_2)}).$$

Reescribimos esta suma como $S = S_1 + S_2$, donde S_1 es la aportación de los términos donde $|t_1-t_2| \leq |t_1t_2|^{1/4}$, y S_2 el resto. De hecho, supongamos que $|t_2| > 2|t_1|$. Tenemos

$$|t_2| \leq |t_1-t_2| + |t_1| \leq |t_1t_2|^{1/4} + \frac{1}{2}|t_2|,$$

por lo que $|t_2| \leq 2|t_1t_2|^{1/4}$, lo que implica que $|t_2| \leq 2^{4/3}|t_1|^{1/3}$, además

$$2|t_1| < |t_2| \leq 2^{4/3}|t_1|^{1/3},$$

lo que a su vez implica que $T \leq |t_1| < \sqrt{2}$, llegando a una contradicción. Cambiando los roles de t_1 y t_2, vemos a su vez que $|t_1| \leq 2|t_2|$. En particular, se sigue que

$$|t_2-t_1| \leq |t_1t_2|^{1/4} \leq 2|t_1|^{1/2}, \quad |t_2-t_1| \leq |t_1t_2|^{1/4} \leq 2|t_2|^{1/2}.$$

Para $T \geq \sqrt{2}$, tenemos que

$$\begin{aligned}
|S_1| &\leq \sum_{\substack{T\leq|t_1|,|t_2|\leq X\\ |t_2-t_1|\leq|t_1t_2|^{1/4}}} |\alpha(t_1)\cdot\alpha(t_2)| \\
&\leq \frac{1}{2}\sum_{\substack{T\leq|t_1|,|t_2|\leq X\\ |t_2-t_1|\leq|t_1t_2|^{1/4}}} (\|\alpha(t_1)\|^2 + (\|\alpha(t_2)\|^2) \\
&\leq \sum_{T\leq|t_1|\leq X} \|\alpha(t_1)\|^2 \sum_{\substack{T\leq|t_2|\leq X\\ |t_2-t_1|\leq 2|t_1|^{1/2}}} 1 + \sum_{T\leq|t_2|\leq X} \|\alpha(t_2)\|^2 \sum_{\substack{T\leq|t_1|\leq X\\ |t_2-t_1|\leq 2|t_2|^{1/2}}} 1 \\
&\ll \sum_{T\leq|T|\leq X} \|\alpha(t)\|^2|t|^{1/2}\log(1+|t|)
\end{aligned}$$

por (3.9). Esta cantidad tiende a 0 cuando $T \to +\infty$, ya que la serie converge sobre todo t por (3.7). Para la suma S_2, tenemos que

$$|S_2| \leq \frac{2}{\log X} \sum_{\substack{T \leq |t_1|, |t_2| \leq X \\ |t_2 - t_1| > |t_1 t_2|^{1/4}}} \frac{\|\alpha(t_1)\| \|\alpha(t_2)\|}{|t_1 - t_2|} \leq \frac{2}{\log X} \sum_{\substack{T \leq |t_1|, |t_2| \leq X \\ |t_2 - t_1| > |t_1 t_2|^{1/4}}} \frac{\|\alpha(t_1)\| \|\alpha(t_2)\|}{|t_1 t_2|^{1/4}},$$

y por lo tanto

$$|S_2| \leq \frac{2}{\log X} \sum_{T \leq |t_1| \leq X} \frac{\|\alpha(t_1)\|}{|t_1|^{1/4}} \sum_{T \leq |t_2| \leq X} \frac{\|\alpha(t_2)\|}{|t_2|^{1/4}} \ll \frac{1}{\log X} \frac{(\log T)^4}{T^{1/2}},$$

por (3.8), por lo que ya queda probado el Lema. □

Ahora concluiremos la prueba del Teorema 3.2.2. Aplicamos el Lema 3.2.7 al conjunto de ordenadas γ de ceros de alguna $L(s, \chi)$, para χ un caracter no trivial módulo q, y a

$$\alpha(\gamma) = \sum_{\substack{\chi (mod\ q) \\ \chi\ no\ trivial \\ L(\frac{1}{2} + i\gamma) = 0}} \frac{1}{\frac{1}{2} + i\gamma} \overline{\chi}$$

para $\gamma \in F$, viendo $\alpha(\gamma)$ como un vector en $\mathbb{C}^{(\mathbb{Z}/q\mathbb{Z})^\times}$, y teniendo en cuenta la mutliplicidad del cero $\frac{1}{2} + i\gamma$ para cualquier caracter χ tal que $L(\frac{1}{2} + i\gamma, \chi) = 0$. Necesitamos comprobar los tres supuestos del Lema. Por la fórmula asintótica de von Mangoldt (Proposición C.4.6), sabemos que (3.9) es válida para los ceros de una L-función módulo q fija, con una constante implícita que depende de q, y por tanto se mantiene para F.

En siguiente lugar tenemos que

$$\|\alpha(\gamma)\| \leq \sum_{\substack{\chi (mod\ q) \\ \chi\ no\ trivial \\ L(\frac{1}{2} + i\gamma) = 0}} \frac{1}{|\frac{1}{2} + i\gamma|} \|\overline{\chi}\| = \varphi(q)^{1/2} \sum_{\substack{\chi (mod\ q) \\ \chi\ no\ trivial \\ L(\frac{1}{2} + i\gamma) = 0}} \frac{1}{|\frac{1}{2} + i\gamma|} \leq \frac{\varphi(q)^{3/2}}{|\frac{1}{2} + i\gamma|} \tag{3.11}$$

por una estimación sencilla del número de caracteres en los que $\frac{1}{2} + i\gamma$ puede ser cero. La condición (3.7) se tiene de (3.11), dado que tenemos

$$\sum_{L(\frac{1}{2} + i\gamma, \chi) = 0} \frac{1}{|\frac{1}{2} + i\gamma|^{1+\varepsilon}} < +\infty \tag{3.12}$$

para todo $\varepsilon > 0$ y para cualquier $\chi (mod\ q)$, y de nuevo la condición (3.8) es consecuencia de (3.11) y (3.12). De (3.6), concluimos que para $X \geq T \geq 2$ tenemos

$$M_X = N_{X,T} + E'_{X,T}$$

donde

$$E'_{X,T} = \sum_{\substack{\chi (mod\ q) \\ \chi\ no\ trivial}} \left(\sum_{T \leq |\gamma| \leq X} \frac{I_X^{i\gamma}}{\frac{1}{2} + i\gamma} \right) \overline{\chi} + O\left(\frac{(\log X)^2}{\sqrt{X}} \right).$$

Estas funciones aleatorias convergen a 0 en L^2, por lo que también lo hacen en L^1, por el Lema 3.2.7. Por la Proposición B.3.2, concluimos que las funciones aleatorias M_X convergen en ley, y que su límite cuando $T \to +\infty$ es el mismo que el límite de

$$-\sum_{\substack{\chi (mod\ q)\\ \chi\ no\ trivial}} \left(\sum_{|\gamma|\leq T} \frac{I_X^{i\gamma}}{\frac{1}{2}+i\gamma} \right) \overline{\chi}.$$

Para terminar la prueba, falta probar que el soporte de N_q está contenido en el hiperplano (3.3). Notemos antes que

$$\begin{aligned} \sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times} N_X(x)(a) &= \frac{\log x}{\sqrt{x}} \sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times} (\varphi(q)\pi(x;q,a) - \pi(x)) \\ &= \frac{\log x}{\sqrt{x}} \sum_{\substack{p\leq x\\ p (mod\ q)\notin(\mathbb{Z}/q\mathbb{Z})^\times}} 1 \ll \frac{\log x}{\sqrt{x}} \end{aligned}$$

para todo $x \in \Omega_X$, ya que a lo más una cantidad finita de primos son no congruentes con algún $a \in (\mathbb{Z}/q\mathbb{Z})^\times$. Luego las variables aleatorias

$$\sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times} N_X(a)$$

converge en probabilidad a 0, y por el Corolario B.3.17, se tiene que el soporte de N_q está contenido en en el núcleo de la forma lineal

$$f \mapsto \sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times} f(a),$$

es decir, (3.3).

3.3. La hipótesis de simplicidad generalizada

El siguiente paso que debemos dar es obtener una expresión explícita de la distribución N_q. De hecho, sabemos por la demostración del Teorema 3.2.2, que N_q es la ley límite cuando $T \to +\infty$ de las variables aleatorias que son a su vez distribuciones límite cuando $X \to +\infty$ de la función aleatoria dada por la suma finita

$$m_q - \sum_{\substack{\chi (mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{L(\frac{1}{2}+i\gamma,\chi)=0\\ |\gamma|\leq T}} \frac{I_X^{i\gamma}}{\frac{1}{2}+i\gamma} \overline{\chi(a)},$$

que converge por la Proposición 3.2.5. La prueba de esta Proposición nos muestra que el límite, en principio, es computable. Precisamente, sea X_T el conjunto de duplas (χ, γ), donde χ recorre los caracteres de Dirichlet módulo q no triviales y γ recorre las ordenadas de los ceros no triviales de $L(s, \chi)$ con $|\gamma| \leq T$. Entonces, por el Corolario 3.2.6, tenemos que

$$N_{q,T} = m_q - \sum_{\substack{\chi (mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{L(\frac{1}{2}+i\gamma,\chi)=0\\ |\gamma|\leq T}} \frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma} \overline{\chi(a)}, \tag{3.13}$$

donde $(I_{\chi,\gamma})$ se distribuye en $(\mathbf{S}^1)^{X_T}$ como una medida de probabilidad de Haar de la clausura S_T del subgrupo generado por los elementos $(x^{i\gamma})_{(\chi,\gamma)\in X_T}$ con $x \in \mathbb{R}$.

Luego, para computar N_q de forma explícita, necesitamos saber qué subgrupo $S_{q,T}$ es. Si ese subgrupo fuera igual a $(\mathbf{S}^1)^{X_{q,T}}$, las

$$(I_{\chi,\gamma})$$

serían independientes y uniformemente distribuidas en $\mathbf{S}^1$, y obtendríamos una fórmula para N_q de (3.13) como suma de términos independientes. Sin embargo, este razonamiento peca de demasiado optimista, pues existen relaciones de dependencia entre las distintas γ, que equivalen a restricciones en el subgrupo S_T. La Hipótesis de Simplicidad Generalizada módulo q es el aserto que indica que todas las relaciones deberían satisfacer las restricciones en S_T. Dichas relaciones son que, si un número complejo z es un cero de $L(s,\chi)$, su conjugado también lo es, pues $\overline{L(\overline{s},\chi)} = L(s,\overline{\chi})$. Luego $(\chi,\gamma) \in X_T$ si y solo si $(\overline{\chi},-\gamma) \in X_T$.

Presentamos, ahora, la anteriormente mencionada Hipótesis de Simplicidad Generalizada módulo q.

Definición 3.3.1 (Hipótesis de Simplicidad Generalizada módulo q)**.** *Sea $q \geq 1$ un entero. La Hipótesis de Simplicidad Generalizada se cumple módulo q si la familia de ordenadas no negativas γ de ceros no triviales de todas las L-funciones de Dirichlet módulo q no triviales, contando multiplicidad, es linealmente independiente sobre $\mathbb{Q}$.*

Esta Hipótesis trae consigo las siguientes implicaciones:

- Para un $\gamma \geq 0$ dado, existe a lo más un caracter de Dirichlet primitivo módulo q χ tal que $L(\frac{1}{2}+i\gamma,\chi) = 0$.
- Todos los ceros no triviales tienen multiplicidad 1.
- Se tiene que $L(\frac{1}{2},\chi) \neq 0$ para cualquier caracter no trivial χ.

Cabe destacar que todos los asertos anteriores son, a día de hoy, conjeturas no precisamente sencillas.

Lema 3.3.2. *Suponiendo la Hipótesis de Simplicidad Generalizada módulo q, el subgrupo S_T viene dado por*

$$S_T = \{(z_{\chi,\gamma}) \in (\mathbf{S}^1)^{X_T} \ : \ z_{\overline{\chi},-\gamma} = \overline{z_{\chi,\gamma}} \text{ para todo } (\chi,\gamma) \in X_T\}, \tag{3.14}$$

para todo $T \geq 2$. En particular, si denotamos por X_T^+ al subconjunto de X_T donde $\gamma \geq 0$, la proyección

$$(z_{\chi,\gamma}) \mapsto (z_{\chi,\gamma})_{(\chi,\gamma)\in X_T^+} \tag{3.15}$$

de S_T en $(\mathbf{S}^1)^{X_T^+}$ es sobreyectiva.

Demostración. En primer lugar, vemos que S_T está contenido en el subgrupo $\tilde{S}_T$ de la derecha de (3.14), ya que cada vector $(z_{\chi,\gamma})_{(\chi,\gamma)\in X_T}$ tiene esta propiedad para $x \in \mathbb{R}$, por la relación entre los ceros de las L-funciones de χ y $\overline{\chi}$.

Para ver que S_T no es un subgrupo propio de $\tilde{S}_T$, basta probar el último aserto, ya que un elemento de $\tilde{S}_T$ está determinado por un único valor de la proyección (3.15).

Pero si esa proyección no es sobreyectiva, entonces existe una familia de enteros no nulos $(m_{\chi,\gamma})_{(\chi,\gamma)\in X_T^+}$ tal que

$$\prod_{(\chi,\gamma)\in X_T^+} x^{i\gamma m_{\chi,\gamma}} = 1$$

para todo $x \in \mathbb{R}$, lo que implica que

$$\sum_{\substack{\chi (mod\ ) \\ \chi\ no\ trivial}} \sum_{\gamma\geq 0} \gamma m_{\chi,\gamma} = 0,$$

lo que contradice la Hipótesis de Simplicidad Generalizada módulo q. □

Por esto, si queremos reformular la Hipótesis sin mecionar el módulo q, podemos hacerlo como *las ordenadas no negativas de los ceros de L-funciones de todos los caracteres primitivos de Dirichlet son $\mathbb{Q}$-linealmente independientes*, tal y como se enuncia en [28].

Ahora podemos enunciar la expresión precisa de la función aleatoria N_q, asumiendo la Hipótesis de de Simplicidad Generalizada. Para ello, sea X^+ el conjunto de las duplas (χ,γ), donde χ es un caracter de Dirichlet no trivial módulo q y $\gamma \geq 0$ es una ordenada no negativa de un cero no trivial de $L(s,\chi)$. Sea $(I_{\chi,\gamma})_{(\chi,\gamma)\in X^+}$ una familia de variables aleatorias independientes e identicamente distribuidas en $\mathbf{S}^1$. Definamos también

$$I_{\overline{\chi},-\gamma} =: \overline{I}_{\chi,\gamma}$$

para todas las ordenadas $\gamma \geq 0$ de los ceros de $L(s,\chi)$. Gracias a (3.3), hemos definido las variables aleatorias $I_{\chi,\gamma}$ para todas las ordenadas de los ceros de $L(s,\chi)$.

Teorema 3.3.3 (Rubinstein-Sarnak)**.** *Sea $q \geq 1$ un entero. Además de la Hipótesis Generalizada de Riemann, supongamos cierta la Hipótesis de Simplicidad Generalizada módulo q. Entonces, la distribución de N_q es la distribución de la suma*

$$m_q - \sum_{\substack{\chi (mod\ q) \\ \chi\ no\ trivial}} \left(\sum_{\substack{\gamma \\ L(\frac{1}{2}+i\gamma,\chi)=0}} \frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma} \right) \overline{\chi}, \tag{3.16}$$

donde la serie converge casi seguro y en L^2 como el límite de las sumas parciales

$$\lim_{T\to+\infty} \sum_{\substack{\chi (mod\ q) \\ \chi\ no\ trivial}} \sum_{\substack{|\gamma|\leq T \\ L(\frac{1}{2}+i\gamma,\chi)=0}} \frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma}. \tag{3.17}$$

En ambas fórmulas, para cada caracter de Dirichlet χ módulo q, la suma recorre las ordenadas de los ceros de $L(s,\chi)$.

Antes de probar este Teorema, podemos dar algunas conclusiones que se extraen del mismo, pues pueden ser buenos indicativos de la existencia de sesgo para algunas clases residuo. Bajo los supuestos del Teorema 3.3.3, tenemos que $\mathbb{E}(N_q) = m_q$, pues la convergencia se mantiene en L^2, y $\mathbb{E}(I_{\chi,\gamma}) = 0$, para todo (χ,γ). Usando la expresión de m_q de la demostración del Teorema 3.2.2, deducimos que

$$\frac{1}{\varphi(q)} \sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times} m_q(a) = 0, \quad \frac{1}{\varphi(q)} \sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times} m_q(a)^2 = \sum_{\substack{\chi^2=1 \\ \chi\ no\ trivial}} 1.$$

Como m_q no es constante, parece razonable pensar que no todas las clases residuo módulo q son iguales, en lo que a representar primos se refiere. Esto es equivalente a a la existencia de un $b \neq 1$ tal que $b^2 = 1$, luego es cierto cuando $q \neq 2$, ya que siempre se puede tomar $b = -1$. Este enunciado puede considerarse como una confimación simple y general de la existencia del sesgo de Chebyshev. Es necesario mencionar que $q = 2$ es una excepción, pues todos los primos (salvo el 2) son impares.

Ahora, probaremos el Teorema 3.3.3.

Demostración. En primer lugar, debemos comprobar que la serie (3.16) converge casi seguro y en L^2 en el sentido del límite (3.17). Basta con probar que, para cada valor $N_q(a)$ de la función aleatoria N_q converge casi seguro y en L^2. Para comprobarlo, vemos que para cualquier $T \geq 2$, tenemos que

$$\sum_{\substack{\chi (mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{|\gamma|\leq T\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)} = \sum_{\substack{\chi (mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{0<\gamma\leq T\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \left(\frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)} + \frac{I_{\overline{\chi},-\gamma}}{\frac{1}{2}-i\gamma}\chi(a)\right)$$

$$= 2 \sum_{\substack{\chi (mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{0<\gamma\leq T\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \Re\left(\frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right), \tag{3.18}$$

de acuerdo con la definición (3.3) de $I_{\chi,\gamma}$, para γ negativa (usamos, además, que bajo la Hipótesis de Simplicidad Generalizada, no hay ceros con $\gamma = 0$). El lado derecho de la igualdad (3.18) es una suma parcial de series de variables aleatorias independientes, por lo que podemos aplicar el Teorema B.3.1. De hecho, tenemos

$$\mathbb{E}\left(\Re\left(\frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right)\right) = 0$$

para cualquier par (χ, γ), y

$$\sum_{\substack{\chi (mod\ q)\\ \chi\ no\ trivial}} \sum_{\gamma>0} \mathbb{V}\left(\Re\left(\frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right)\right) \leq \sum_{\substack{\chi (mod\ q)\\ \chi\ no\ trivial}} \sum_{\gamma>0} \mathbb{E}\left(\left|\frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\right|^2\right)$$

$$= \sum_{\substack{\chi (mod\ q)\\ \chi\ no\ trivial}} \sum_{\gamma>0} \frac{1}{\frac{1}{4}+\gamma^2} < +\infty$$

por la segunda parte de la Proposición C.4.6, por lo que la serie converge casi seguro y en L^2, por el Teorema de Kolmogorov, como dijimos antes.

La función aleatoria N_q es el límite cuando $T \to +\infty$ de

$$N_{q,T} = m_q - \lim_{X\to+\infty}\left(\sum_{\substack{\chi (mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{|\gamma|\leq T\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \frac{I_X^{i\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right).$$

Escribimos una vez más

$$m_q - \sum_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{|\gamma|\leq T\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \frac{I_X^{i\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)} = m_q - 2\sum_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{0<\gamma\leq T\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \Re\left(\frac{I_X^{i\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right).$$

Por la Proposición 3.2.5, o el Corolario 3.2.6 y la Hipótesis de Simplicidad Generalizada módulo q, el límite cuando $X \to +\infty$ de estas funciones aleatorias es

$$m_q - 2\sum_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{0<\gamma\leq T\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \Re\left(\frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right)$$

que a su vez, converge a la función aleatoria N_q por definición, lo cual concluye la prueba. □

Observación 3.3.1. El Teorema 3.3.3 es equivalente al cálculo de la función característica de N_q visto como vector aleatorio. De hecho, así es como se presenta el resultado en [28].

Una vez probado el Teorema 3.3.3, daremos dos consecuencias del mismo: la función característica de N_q y una estimación de la probabilidad de que N_q tome valores grandes. Para el primero, debemos definir antes la función de Bessel J_0 en $\mathbb{R}$:

$$J_0(x) := \frac{1}{2\pi}\int_0^{2\pi} e^{ix\cos(t)}\, dt.$$

Debemos decir que J_0 es una función real y par en x.

Corolario 3.3.4. *Sea $q \geq 2$ un entero. Supongamos ciertas tanto la Hipótesis Generalizada de Riemann como la Hipótesis de Simplicidad Generalizada módulo q. La función característica de la distribución de Rubinstein-Sarnak módulo q N_q viene dada por*

$$\mathbb{E}(e^{it\cdot N_q}) = \exp(it\cdot m_q)\prod_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}}\prod_{\substack{\gamma>0\\ L(\frac{1}{2}+i\gamma,\chi)=0}} J_0\left(\frac{2\cdot|t\cdot\overline{\chi}|}{(\frac{1}{4}+\gamma^2)^{1/2}}\right)$$

para $t \in \mathbb{R}^{(\mathbb{Z}/q\mathbb{Z})^\times}$, donde para cada caracter de Dirichlet χ módulo q, el producto recorre las ordenadas positivas de los ceros de $L(s,\chi)$.

Demostración. Empezaremos escribiendo las series que definen N_q de la siguiente forma, al igual que en la prueba del Teorema 3.3.3:

$$m_q - 2\sum_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{0<\gamma\leq T\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \Re\left(\frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi}\right).$$

Como la función característica de un límite en ley es el límite puntual de las funciones características de la sucesión involucrada, obtenemos usando la independencia de las variables aleatorias $(I_{\chi,\gamma})$ la fórmula producto convergente

$$\mathbb{E}(e^{it\cdot N_q}) = \exp(it\cdot m_q)\prod_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}}\prod_{\gamma>0}\mathbb{E}\left(e^{-2it\cdot\Re\left(\frac{I_{\chi,\gamma}}{1/2+i\gamma}\overline{\chi}\right)}\right)$$

$$= \exp(it \cdot m_q) \prod_{\substack{\chi (mod\ q) \\ \chi\ no\ trivial}} \prod_{\gamma>0} \varphi\left(\frac{t \cdot \overline{\chi}}{\frac{1}{2} + i\gamma}\right)$$

donde, para $z \in \mathbb{C}$, definimos

$$\varphi(z) := \mathbb{E}\left(e^{-2i\Re(zI)}\right)$$

para una variable aleatoria I uniformemente distribuida en el círculo unidad. Por la invariancia de la distribución de I bajo rotaciones (la distribución de $ze^{i\theta}I$ es la misma que la de zI para cualquier $\theta \in \mathbb{R}$), aplicado al ángulo θ tal que $ze^{i\theta} = |z|$, se tiene

$$\varphi(z) = \mathbb{E}\left(e^{-2i\Re(|z|I)}\right) = \mathbb{E}\left(e^{-2i|z|\Re(I)}\right) = \frac{1}{2\pi}\int_0^{2\pi} e^{-2i|z|\cos(t)}\, dt = J_0(2|z|).$$

Por lo tanto, obtenemos que

$$\mathbb{E}(e^{it\cdot N_q}) = \exp(it \cdot m_q) \prod_{\substack{\chi (mod\ q) \\ \chi\ no\ trivial}} \prod_{\gamma>0} J_0\left(\frac{2 \cdot |t \cdot \overline{\chi}|}{(\frac{1}{4} + \gamma^2)^{1/2}}\right),$$

como se quería probar. □

Corolario 3.3.5. *Existe una constante $c_q > 0$ tal que, para $A > 0$, se tiene que*

$$\begin{aligned} c_q^{-1} \exp(-\exp(c_q A^{1/2})) &\leq \liminf_{X\to+\infty} \mathbb{P}_X(\|N_{X,q}\| \geq A) \\ &\leq \limsup_{X\to+\infty} \mathbb{P}_X(\|N_{X,q}\| > A) \leq c_q \exp(-\exp(c_q^{-1} A^{1/2})). \end{aligned}$$

Demostración. En primer lugar, veamos N_q como una variable aleatoria con valores en el espacio de Banach complejo de dimensión finita de las funciones complejas en $(\mathbb{Z}/q\mathbb{Z})^\times$. Tenemos la representación en series

$$N_q = m_q - 2 \sum_{\substack{\chi (mod\ q) \\ \chi\ no\ trivial}} \sum_{L(\frac{1}{2}+i\gamma,\chi)=0} \Re\left(\frac{I_{\chi,\gamma}}{\frac{1}{2} + i\gamma}\overline{\chi}\right).$$

Esta serie converge casi seguro, los términos son independientes y las variables aleatorias $I_{\chi,\gamma}$ están acotadas en módulo por 1. Es más,

$$\mathbb{P}(\|N_q\| > A) \leq \mathbb{P}(\|\tilde{N}_q\| > A)$$

donde

$$\tilde{N}_q = m_q - 2 \sum_{\substack{\chi (mod\ q) \\ \chi\ no\ trivial}} \sum_{L(\frac{1}{2}+i\gamma,\chi)=0} \frac{I_{\chi,\gamma}}{\frac{1}{2} + i\gamma}\overline{\chi},$$

ya que $\|N_q\| \leq \|\tilde{N}_q\|$. Por el Corolario C.4.7, las funciones

$$-\frac{2}{\frac{1}{2} + i\gamma}\overline{\chi}$$

satisfacen las cotas descritas en el Corolario B.3.18, es decir,

$$\sum_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{0<\gamma<T\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \left\| \frac{1}{\frac{1}{2}+i\gamma}\overline{\chi} \right\| \gg (\log T)^2$$

y

$$\sum_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{\gamma>T\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \left\| \frac{1}{\frac{1}{2}+i\gamma}\overline{\chi} \right\|^2 \ll \frac{\log T}{T}$$

para $T \geq 1$. Por el Corolario B.3.18 y la convergencia en ley de $N_{X,q}$ a N_q, deducimos la cota superior

$$\limsup_{X\to+\infty} \mathbb{P}_X(\|N_{X,q}\| > A) \leq \mathbb{P}(\|N_q\| > A) \leq c\exp(-\exp(c^{-1}A^{1/2}))$$

para un número real $c > 0$. En el caso de la cota inferior, basta probar que es cierta para $N_q(a)$, donde a es un elemento fijo de $(\mathbb{Z}/q\mathbb{Z})^\times$. Como las series que expresan $N_q(a)$ no son de la forma que requiere el Corolario B.3.18 para la cota inferior, la transformamos un poco. Tenemos

$$\Re\left(\frac{I_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right) = \frac{1}{2\left(\frac{1}{4}+\gamma^2\right)}\Re(I_{\chi,\gamma}\overline{\chi(a)}) + \frac{\gamma}{\frac{1}{4}+\gamma^2}\Im(I_{\chi,\gamma}\overline{\chi(a)}),$$

para cualquier pareja (χ, γ), lo que implica que

$$N_q(a) = m_q(a) + e_q(a) - 2\sum_{\substack{\chi(mod\ q)\\ \chi\ no\ trivial}} \sum_{\substack{\gamma>0\\ L(\frac{1}{2}+i\gamma,\chi)=0}} \frac{\gamma}{\frac{1}{4}+\gamma^2}\Im(I_{\chi,\gamma}\overline{\chi(a)})$$

donde la variable aleatoria $e_q(a)$ (que aparece por la suma de los primeros términos en la expresión anterior) es uniformemente acotada (por la Proposición C.4.6). Ahora podemos aplicar la cota inferior del corolario B.3.18 a las últimas series. Las variables aleatorias $\Im(I_{\chi,\gamma}\overline{\chi(a)})$ son independientes, simétricas y acotadas por 1, y los supuestos del tamaño de los coeficientes vienen dados por el Corolario C.4.7.

□

4. Forma de las sumas exponenciales

> Si la gente no piensa que las matemáticas son simples, es solo porque no se dan cuenta de lo complicada que es la vida.
>
> *John von Neumann*

4.1. Introducción

En este capítulo se buscará dar una introducción a un tipo especial de sumas exponenciales, las sumas de Kloostermann. Estas sumas tienen un peso importante dentro de la Teoría Analítica de Números.

Para ello, empezaremos definiendo qué son estas sumas, además de mostrar algunas de sus propiedades y resultados relacionados. Posteriormente, introduciremos la distribución de Sato-Tate, su relación con las sumas de Kloosterman y enunciaremos y probaremos un teorema límite que involucra a ambos conceptos. Finalmente, daremos dos resultados que son consecuencia de este último teorema.

Como en capítulos anteriores, se precisan de ciertos conocimientos previos en Teoría de la Probabilidad, Teoría Analítica de Números y Análisis Matemático, particularmente Análisis de Fourier. Si no es el caso, no hay de qué preocuparse. Cualquier resultado auxiliar empleado será explicitado en el Anexo correspondiente para consulta y/o curiosidad del lector.

4.2. Sumas de Kloosterman. Definición y propiedades

Empezaremos definiendo el concepto de suma de Kloosterman.

Definición 4.2.1 (Suma de Kloosterman)**.** *Sea p un número primo. Para todo par (a, b) de elementos invertibles del cuerpo finito $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, se define la suma (normalizada) de Kloosterman como*

$$Kl(a, b; p) := \frac{1}{\sqrt{p}} \sum_{x \in \mathbb{F}_p^\times} e\left(\frac{ax + b\overline{x}}{p}\right), \tag{4.1}$$

donde denotamos la función $e(z) = e^{2i\pi z}$ y $\overline{x}$ al inverso de x módulo p. Sin el factor normalizante $\frac{1}{\sqrt{p}}$, denotamos a la suma de Kloosterman $K(a, b; p)$.

Por su estructura, podemos observar que las sumas de Kloosterman son sumas finitas, pues recorren las unidades de un cuerpo finito. Una de las propiedades más reseñables de estas sumas es la dada por André Weil, una cota para su módulo:

$$|Kl(a,b;p)| \leq 2. \tag{4.2}$$

A pesar de la simplicidad de la expresión, es un resultado muy importante, pues si intentamos acotar de una forma más intuitiva, el sumatorio de (4.1) recorre $p-1$ raíces de la unidad. Por lo que, al dividir por $\sqrt{p}$, nos queda la cota algo más visualizable $|Kl(a,b;p)| \leq (p-1)/\sqrt{p}$. Esto nos da una idea de cómo cambia el argumento de los sumandos en $\mathbb{C}$, de lo cual no se puede decir que sea precisamente estable, Esto se debe, principalmente, a la alta variabilidad de la función $x \mapsto \overline{x}$ en el conjunto $\{0,\ldots,p-1\}$.

Ahora es cuando entra en juego la Probabilidad. Si modelizamos la suma anterior como una suma de términos independientes uniformemente distribuidos en el círculo unidad

$$S_N = X_1 + \ldots X_N$$

donde los sumandos son (X_n) variables aleatorias independientes y uniformemente distribuidos en el círculo unidad. Por el Teorema Central del Límite, $X_N/\sqrt{N}$ convege en ley a la normal estándar compleja. De esto podemos extraer que el factor normalizante en (4.1) tiene, no solo sentido, sino además razón de ser. Como dato curioso, el valor S_N no es solamente una suma parcial, sino que, visto como hemos descrito anteriormente, es un proceso estocástico muy empleado en varias ramas de la estadística, como el estudio de las series temporales. Este proceso se conoce como paseo aleatorio o caminata al azar, dependiendo del autor del manual que se consulte.

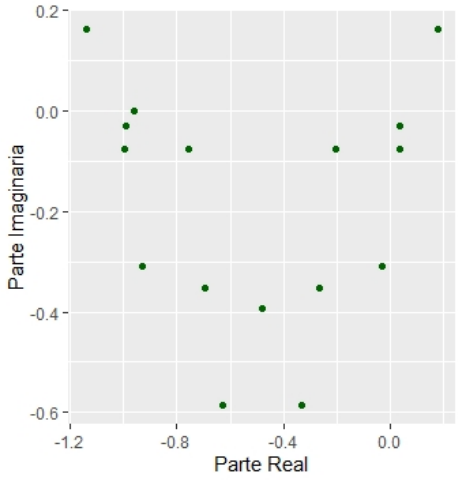


Figura 4.1: Representación de las sumas parciales de $Kl(1,1;17)$

Este hecho sugiere que, desde una perspectiva probabilística, no es mala idea estudiar las sumas parciales de las sumas de Kloosterman y su movimiento en el plano complejo. Para ello, necesitamos imponer un orden en la suma (4.1), el cual se consigue sumando sobre $1 \leq x \leq p-1$. Por lo tanto, consideramos los $p-1$ puntos

$$z_j = \frac{1}{\sqrt{p}} \sum_{1 \leq x \leq j} e\left(\frac{ax+b\overline{x}}{p}\right) \tag{4.3}$$

para $1 \leq x \leq p-1$. Sin embargo, esta nube de puntos no da mucha información de por sí. Para solucionar esto, uniremos los puntos sucesivos mediante segmentos rectos. Podemos ver un ejemplo en la figura 4.2. Si variamos a o b, vemos que las figuras generadas cambian de una forma aparentemente aleatoria, aunque permanece cierta simetría.

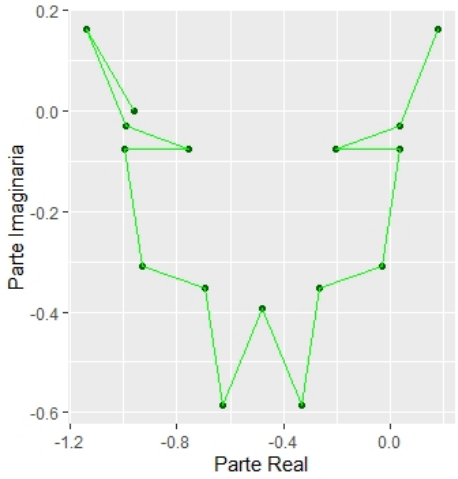


Figura 4.2: Representación del camino generado por las sumas parciales de $Kl(1,1;17)$

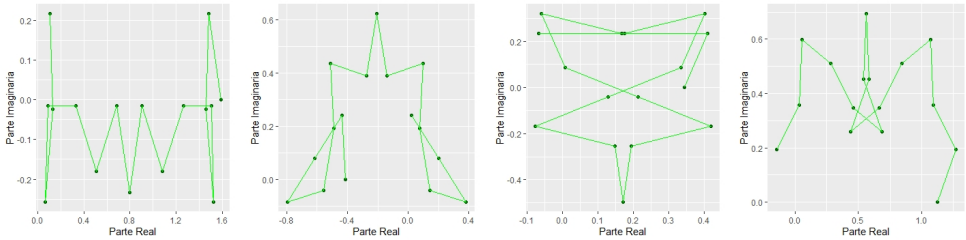


Figura 4.3: Los caminos de las sumas parciales de $Kl(a,1;17)$ con $a = 2,3,4,5$.

De forma natural nos surge la cuestión de si existe un comportamiento estadístico definido para estos caminos de Kloosterman, para $p \to +\infty$, cuando tomamos de manera uniformemente aleatoria una pareja (a,b) de elementos de $\mathbb{F}_p^\times$. En la próxima sección veremos que así es.

4.3. Sumas de Kloosterman. Teorema de distribución

Para poder enunciar este resultado, precisamos de notación auxiliar. Para p primo y (a,b) una pareja de elementos de $\mathbb{F}_p^\times$, denotamos por $K_p(a,b)\colon [0,1] \to \mathbb{C}$ a la función tal que, para $0 \leq j \leq p-2$, su valor en un número real t tal que

$$\frac{j}{p-1} \leq t < \frac{j+1}{p-1}$$

se obtiene interpolando linealmente entre las sumas parciales consecutivas z_j y z_{j+1}, definidas en (4.3). El camino $t \mapsto K_p(a,b)(t)$ es la poligonal descrita anteriormente;

para $t = 0$ tenemos $K_p(a,b)(0) = 0$, y para $K_p(a,b)(1) = Kl(a,b;p)$. Sea $\Omega_p = \mathbb{F}_p^\times \times \mathbb{F}_p^\times$. Podemos ver K_p como una variable aleatoria

$$\Omega_p \to C([0,1]),$$

donde $C([0,1])$ es el espacio de Banach de las funciones continuas $\varphi\colon [0,1] \to \mathbb{C}$ con la norma del supremo $\|\varphi\|_\infty = \sup|\varphi(t)|$. También podemos pensar en la familia de variables aleatorias $(K_p(t))_{t\in[0,1]}$ tales que

$$(a,b) \mapsto K_p(a,b)(t),$$

como un proceso estocástico, donde t tiene el papel de la variable temporal.

Ahora sí, podemos enunciar el teorema principal de esta sección.

Teorema 4.3.1. *Sea $(ST_h)_{h\in\mathbb{Z}}$ una sucesión de variables aleatorias independientes, todas ellas distribuidas según la medida de Sato-Tate*

$$\mu_{ST} = \frac{1}{p}\sqrt{1 - \frac{x^2}{4}}dx$$

en el intervalo $[-2,2]$.

1. *La serie de Fourier aleatoria*

$$K(t) = tST_0 + \sum_{\substack{h\in\mathbb{Z}\\ h\neq 0}} \frac{e(ht)-1}{2i\pi h} ST_h$$

definida para $t \in [0,1]$ converge uniformemente casi seguro, en el sentido de las sumas parciales simétricas

$$K(t) = tST_0 + \lim_{H\to+\infty} \sum_{\substack{h\in\mathbb{Z}\\ 1\leq|h|<H}} \frac{e(ht)-1}{2i\pi h} ST_h.$$

Esta serie de Fourier aleatoria define una variable aleatoria $C([0,1])-$valuada K.

2. *Cuando $p \to +\infty$, las variables aleatorias K_p convergen en ley a K, en el sentido de las variables $C([0,1])-$valuadas.*

En particular, el Teorema 4.3.1 implica que, tomando $t = 1$, las sumas de Kloosterman $Kl(a,b;p) = K_p(a,b)(1)$, vista como variables aleatorias en Ω_p, están distribuidas asintóticamente como $K(1) = ST_0$, es decir, las sumas de Kloosterman están distribuidas según la distribución de Sato-Tate, en el sentido de que, para cualesquiera números reales $-2 \leq \alpha < \beta \leq 2$, tenemos que

$$\frac{1}{(p-1)^2}|\{(a,b) \in \Omega_p \ : \ \alpha < Kl(a,b;p) < \beta\}| = \int_\alpha^\beta d\mu_{ST}(t).$$

Observación 4.3.1. Como curiosidad, este es un famoso resultado de Nick Katz [18]. De hecho, una herramienta clave para la prueba del Teorema 4.3.1 es una extensión de los métodos desarrollados por Katz para probar resultados similares a éste.

Para la prueba del Teorema 4.3.1, seguiremos el siguiente esquema:

1. Probar que la serie de Fourier aleatoria K existe, y es una variable aleatoria $C([0,1]) - valuada$.
2. Probar que una pequeña variante de la sucesión de los coeficientes de Fourier de K_p converge en ley a la sucesión de coeficientes de Fourier de K.
3. Probar que la sucesión $(K_p)_p$ es *ajustada* (ver Definición B.3.24), usando el Criterio de ajuste de Kolmogorov (B.3.25).

Una vez tengamos esto hecho, por la Proposición B.3.26 muestra que una combinación de los pasos (2) y (3) implica que K_p converge a K. Denotaremos por $\mathbb{P}_p$ y $\mathbb{E}_p$ a la probabilidad y la esperanza respecto a la medida uniforme en Ω_p. Antes de empezar la prueba en sí, no está de más ver por qué aparece el límite, y por qué es precisamente esta serie de Fourier.

Lema 4.3.2. *Sea $p \geq 3$ un número primo y $a, b \in \mathbb{F}_p^\times$. Sea $t \in [0,1]$. Entonces tenemos que*

$$\frac{1}{\sqrt{p}} \sum_{1 \leq n \leq (p-1)t} e\left(\frac{an + b\overline{n}}{p}\right) = \sum_{|h|<p/2} \alpha_p(h,t) Kl(a-h, b; p), \tag{4.4}$$

donde

$$\alpha_p(h,t) = \frac{1}{p} \sum_{1 \leq n \leq (p-1)t} e\left(\frac{nh}{p}\right).$$

Demostración. Este es un caso de la fórmula discreta de Plancherel (ver Teorema A.2.4) aplicada a la función característica del intervalo discreto de sumación. Para ver esto, basta insertar las definiciones de $\alpha_p(h,t)$ y $Kl(a-h,b;p)$ en el lado derecho de (4.4). Esto nos muestra que es igual a

$$\begin{aligned}
\sum_{|h|<p/2} \alpha_p(h,t) Kl(a-h,b;p) &= \frac{1}{p^{3/2}} \sum_{|h|<p/2} \sum_{1 \leq n \leq (p-1)t} \sum_{m \in \mathbb{F}_p} e\left(\frac{nh}{p}\right) e\left(\frac{(a-h)m + b\overline{m}}{p}\right) \\
&= \frac{1}{\sqrt{p}} \sum_{1 \leq n \leq (p-1)t} \sum_{m \in \mathbb{F}_p} e\left(\frac{am + b\overline{m}}{p}\right) \frac{1}{p} \sum_{h \in \mathbb{F}_p} e\left(\frac{h(n-m)}{p}\right) \\
&= \frac{1}{\sqrt{p}} \sum_{1 \leq n \leq (p-1)t} e\left(\frac{an + b\overline{n}}{p}\right),
\end{aligned}$$

como dijimos antes, ya que por la ortogonalidad de los caracteres tenemos

$$\frac{1}{p} \sum_{h \in \mathbb{F}_p} e\left(\frac{h(n-m)}{p}\right) = \delta(m,n)$$

para cualesquiera $m, n \in \mathbb{F}_p$, donde $\delta(m,n) = 1$ si $n = m(\mod p)$, y $\delta(m,n) = 0$ en cualquier otro caso. □

Si observamos que $\alpha_p(h,t)$ es una suma de Riemann de la integral

$$\int_0^t e(ht)\, dt = \frac{e(ht) - 1}{2i\pi h}$$

para todo $h \neq 0$, y que $\alpha_p(0,t) \to t$ cuando $p \to +\infty$, podemos observar que el lado derecho de (4.4) se parece a una serie de Fourier del mismo tipo que $K(t)$, con coeficientes dados por una suma de Kloosterman desplazada en vez de ST_h. Ahora presentaremos un teorema de suma importancia, pues contiene la información aritmética importante que nos será de gran utilidad a la hora de probar el Teorema 4.3.1. Sin embargo, su demostración se escapa de nuestros objetivos y herramientas disponibles, por lo que no se incluirá en este trabajo.

Teorema 4.3.3 (Katz, Deligne)**.** *Sea $b \neq 0$ un entero fijado. Para p primo que no divide a b, consideremos la variable aleatoria*

$$S_p \colon a \mapsto (Kl(a-h,b;p))_{h\in\mathbb{Z}}$$

en $\mathbb{F}_p^\times$ con la medida de probabilidad uniforme, tomando valores en el espacio topológico compacto

$$\hat{T} = \prod_{h\in\mathbb{Z}} [-2,2].$$

Entonces S_p converge en ley a la medida de probabilidad producto

$$\bigotimes_{h\in\mathbb{Z}} \mu_{ST}.$$

En otras palabras, la sucesión de variables aleatorias $a \mapsto Kl(a-h,b;p)$ converge en ley a la sucesión $(ST_h)_{h\in\mathbb{Z}}$ de variables aleatorias Sato-Tate idependientes.

Es por este teorema que de (4.4), se extrae que $K_p(t)$ converge en ley a la serie aleatoria

$$tST_0 + \sum_{\substack{h\in\mathbb{Z}\\ h\neq 0}} \frac{e(ht)-1}{2i\pi h} ST_h,$$

que es exactamente $K(t)$. Antes de empezar la prueba del Teorema 4.3.1, daremos un último lema auxiliar.

Lema 4.3.4. *Los coeficientes de Fourier $FT(K_p)$ convergen en ley a $FT(K)$, en el sentido de convergencia de distribución finita.*

Ahora sí, procedemos con la demostración siguiendo el esquema presentado anteriormente.

Demostración. Como hemos dicho anteriormente, seguiremos el esquema de los tres pasos expuesto anteriormente.
Paso 1.-Podemos escribir la serie $K(t)$ como

$$K(t) = tST_0 + \sum_{h\geq 1} \left(\frac{e(ht)-1}{2i\pi h} ST_h - \frac{e(-ht)-1}{2i\pi h} ST_{-h} \right).$$

Renombramos los sumandos como

$$X_h = \frac{e(ht)-1}{2i\pi h} ST_h - \frac{e(-ht)-1}{2i\pi h} ST_{-h}$$

para $h \geq 1$. Estos sumandos son independientes y tienen esperanza 0, pues $\mathbb{E}(ST_h) = 0$. Es más, como ST_h es independiente de ST_{-h}, y tienen varianza 1, tenemos

$$\sum_{h\geq 1} \mathbb{V}(X_h) = \sum_{h\geq 1} \left(\left| \frac{e(ht)-1}{2i\pi h} \right|^2 + \left| \frac{e(-ht)-1}{2i\pi h} \right|^2 \right) \leq \sum_{h\geq 1} \frac{1}{h^2} < +\infty$$

para todo $t \in [0,1]$. Por el Teorema B.3.1, se tiene que para todo $t \in [0,1]$, la serie $K(t)$ converge casi seguro y en L^2 a una variable aleatoria compleja. Para probar la convergencia en $C([0,1])$, usaremos la convergencia de distribuciones finitas junto con el criterio de ajuste de Kolmogorov. Consideremos las sumas parciales

$$K_H(t) = tST_0 + \sum_{1\leq |h|\leq H} \frac{e(ht)-1}{2i\pi h} ST_h$$

para $H \geq 1$. Estas son variables aleatorias $C([0,1]) - valuadas$. La convergencia de $K_H(t)$ a $K(t)$ en L^1 implica, por el Lema B.3.27, que la sucesión $(K_H)_{H\geq 1}$ converge a K en el sentido de distribuciones finitas. Luego, por la Proposición B.3.25, la sucesión converge en el sentido de variables aleatorias $C([0,1]) - valuadas$ si existen constantes $C \geq 0$, $\alpha > 0$ y $\delta > 0$ tal que para cualquier $H \geq 1$, y números reales $0 \leq s < t \leq 1$, tenemos

$$\mathbb{E}(|K_H(t) - K_H(s)|^\alpha) \leq C|t-s|^{1+\delta}. \tag{4.5}$$

Tomemos $\alpha = 4$. Tenemos

$$K_H(t) - K_H(s) = (t-s)ST_0 + \sum_{1\leq |h|\leq H} \frac{e(ht)-e(hs)}{2i\pi h} ST_h.$$

Esta es una suma de variables aleatorias independientes, centradas y acotadas, por lo que por la Proposición B.3.29, es $\sigma_H^2 - subgaussiana$ con

$$\sigma_H^2 = |t-s|^2 + \sum_{1\leq |h|\leq H} \left| \frac{e(ht)-e(hs)}{2i\pi h} \right|^2 \leq |t-s|^2 + \sum_{|h|\neq 0} \left| \frac{e(ht)-e(hs)}{2i\pi h} \right|^2.$$

Por la fórmula de Parseval para series de Fourier ordinarias (ver Teorema A.2.5), tenemos que

$$|t-s|^2 + \sum_{|h|\neq 0} \left| \frac{e(ht)-e(hs)}{2i\pi h} \right|^2 = \int_0^1 |\varphi_{s,t}(x)|^2 \, dx,$$

donde $\varphi_{s,t}$ es la función característica del intervalo $[s,t]$. Por lo tanto, $\sigma_H^2 \leq |t-s|$. Por las propiedades de las variables aleatorias subgaussianas (ver Proposición B.3.30), deducimos que existe una constante $C \geq 0$ tal que

$$\mathbb{E}(|K_H(t) - H_H(s)|^4) \leq C\sigma_H^4 \leq C|t-s|^2,$$

como se expone en (4.5).

Paso 2.-En primer lugar, denotaremos por $C_0([0,1])$ al subespacio de funciones $f \in C([0,1])$ tal que $f(0) = 0$. Para $f \in C_0([0,1])$, la sucesión $FT(f) = (\tilde{f}(h))_{h\in\mathbb{Z}}$ está definida por $\tilde{f}(0) = f(1)$ y

$$\tilde{f}(h) := \int_0^1 (f(t) - tf(1))e(-ht) \, dt$$

para $h \neq 0$. La aplicación FT es una aplicación lineal continua de $C_0([0,1])$ en $C_0(\mathbb{Z})$, el espacio de Banach de funciones $\mathbb{Z} \to \mathbb{C}$ que tienden a 0 en el infinito.
Empezaremos calculando los coeficientes de Fourier de un camino poligonal. Sean z_0 y z_1 números complejos, y $t_0 < t_1$ números reales. Definimos $\Delta := t_1 - t_0$ y $f \in C([0,1])$ por

$$f(t) = \begin{cases} \frac{1}{\Delta}(z_1(t-t_0) + z_0(t_1-t)) & si \quad t_0 \leq t \leq t_1 \\ \\ 0 & en\ otro\ caso \end{cases}$$

que parametriza el segmento de z_0 en z_1 sobre el intervalo $[t_0, t_1]$. Sea $h \neq 0$. Directamente obtenemos

$$\begin{aligned} \int_0^1 f(t)e(-ht)\,dt = &-\frac{1}{2i\pi h}(z_1 e(-ht_1) - z_0 e(-ht_0)) + \\ &\frac{1}{2i\pi h}(z_1 - z_0)e(-ht_0)\frac{1}{\Delta}\Big(\int_0^\Delta e(-hu)\,du\Big) \\ &= -\frac{1}{2i\pi h}(z_1 e(-ht_1) - z_0 e(-ht_0)) + \\ &\frac{1}{2i\pi h}(z_1 - z_0)\frac{\sin(\pi h\Delta)}{\pi h\Delta}e\Big(-h\Big(t_0 + \frac{\Delta}{2}\Big)\Big). \end{aligned} \tag{4.6}$$

Sea ahora un entero $n \geq 1$, y una familia $(z_0, \ldots, z_n)$ de números complejos. Para $0 \leq j \leq n-1$, sea f_j la función definida como arriba relativa a la pareja de puntos (z_j, z_{j+1}) y al intervalo $\left[\frac{j}{n}, \frac{j+1}{n}\right]$, y definimos

$$f = \sum_{j=0}^{n-1} f_j,$$

por lo que f parametriza el camino poligonal une z_0 a z_1 a z_2 ... a z_n, cada vez con intervalos de igual longitud de $1/n$. Para $h \neq 0$ obtenemos, sumando (4.6), usando la suma telescópica y las relaciones $z_0 = f(0)$ y $z_n = f(1)$ la fórmula

$$\begin{aligned} \int_0^1 f(t)e(-ht)\,dt = &-\frac{1}{2i\pi h}(f(1) - f(0)) + \\ &\frac{1}{2i\pi h}\frac{\sin(\pi h/n)}{\pi h/n}\sum_{j=0}^{n-1}(z_{j+1} - z_j)e\Big(-\frac{h(j+\frac{1}{2})}{n}\Big). \end{aligned} \tag{4.7}$$

Especializaremos esta fórmula para los caminos de Kloosterman. Sea p un primo, $(a,b) \in \Omega_p$, y apliquemos la fórmula para $n = p-1$ y los puntos

$$z_j = \frac{1}{\sqrt{p}}\sum_{1 \leq x \leq j} e\Big(\frac{ax + b\overline{x}}{p}\Big), \quad 0 \leq j \leq p-1.$$

Para $h \neq 0$, el $h-ésimo$ coeficiente de Fourier de $K_p - tK_p(1)$ es la variable aleatoria en Ω_p que lleva la dupla (a,b) en

$$\frac{1}{2i\pi h}\frac{\sin(\pi h/(p-1))}{\pi h/(p-1)}e\Big(-\frac{h}{2(p-1)}\Big)\frac{1}{\sqrt{p}}\sum_{x=1}^{p-1} e\Big(\frac{ax + b\overline{x}}{p}\Big)e\Big(-\frac{hx}{p-1}\Big).$$

Notemos que, para h fijado, tenemos

$$e\Big(-\frac{hx}{p-1}\Big) = e\Big(-\frac{hx}{p}\Big)e\Big(-\frac{hx}{p(p-1)}\Big) = e\Big(-\frac{hx}{p}\Big)(1+O(p^{-1}))$$

para todo p y todo x tal que $1 \leq x \leq p-1$, por lo que

$$\frac{1}{\sqrt{p}}\sum_{x=1}^{p-1} e\Big(\frac{ax+b\overline{x}}{p}\Big)e\Big(-\frac{hx}{p-1}\Big) = Kl(a-h,b;p) + O\Big(\frac{1}{\sqrt{p}}\Big),$$

donde la constante implícita depende de h. Sea

$$\beta_p(h) := \frac{\sin(\pi h/(p-1))}{\pi h/(p-1)} e\Big(-\frac{h}{2(p-1)}\Big).$$

Vemos que $|\beta_p(h)| \leq 1$, luego podemos expresar el $h-$*ésimo* coeficiente de Fourier como

$$\frac{1}{2i\pi h} Kl(a-h,b;p)\beta_p(h) + O(p^{-1/2}),$$

donde la constante implícita depende de h. Cabe destacar que la $0-$*ésima* componente de $FT(K_p)$ es $Kl(a,b;p)$. Dado que $\beta_p(h) \to 1$ cuando $p \to +\infty$ para cualquier h fijado, deducimos del Teorema 4.3.3 y el Lema B.3.31 (aplicado a los vectores de coeficientes de Fourier en $h_1,\ldots,h_m$ para un $m \geq 1$ arbitrario) que $FT(K_p)$ converge en ley a $FT(K)$ en el sentido de distribuciones finitas.

Paso 3.-Ahora debemos comprobar que la sucesión $(K_p)_p$ es ajustada. Según el criterio de Kolmogorov (Proposición B.3.25), es suficiente con encontrar constantes $C \geq 0$, $\alpha > 0$ y $\delta > 0$ tal que, para todo primo $p \geq 3$ y cualesquiera t y s tales que $0 \leq s < t \leq 1$, tenemos que

$$\mathbb{E}_p(|K_p(t) - K_p(s)|^\alpha) \leq C|t-s|^{1+\delta}. \tag{4.8}$$

Denotemos por $\gamma \geq 0$ al número real tal que

$$|t-s| = (p-1)^{-\gamma}.$$

Por lo que γ es mayor cuando t y s son más parecidos. Para probar (4.8), veremos dos casos. En primer lugar, supongamos que $\gamma > 1$. Para ese caso, utilizamos la naturaleza poligonal de los caminos $x \mapsto K_p(x)$, lo que implica que

$$|K_p(t) - K_p(s)| \leq \sqrt{p-1}|t-s| \leq \sqrt{|t-s|}$$

(ya que la *velocidad* del camino es $(p-1)/\sqrt{p} \leq \sqrt{p-1}$). Por lo tanto, para todo $\alpha > 0$ tenemos que

$$\mathbb{E}_p(|K_p(t) - K_p(s)|^\alpha) \leq |t-s|^{\alpha/2}. \tag{4.9}$$

En el siguiente caso ($\gamma \leq 1$), usaremos las sumas parciales discontinuas $\tilde{K}_p(t)$ en vez de $K_p(t)$. Para comprobar que podemos hacer esto, vemos que

$$|\tilde{K}_p(t) - K_p(t)| \leq \frac{1}{\sqrt{p}}$$

para todo primo $p \geq 3$. Por lo tanto, por la desigualdad de Hölder, llegamos para $\alpha \geq 1$ a la relación

$$\begin{aligned}\mathbb{E}_p(|K_p(t) - K_p(s)|^\alpha) &= \mathbb{E}_p(|\tilde{K}_p(t) - \tilde{K}_p(s)|^\alpha) + O(p^{-\alpha/2}) \\ &= \mathbb{E}_p(|\tilde{K}_p(t) - \tilde{K}_p(s)|^\alpha) + O(|t-s|^{-\alpha/2})\end{aligned} \tag{4.10}$$

donde la constante implícita depende solo de α. Tomemos, ahora, $\alpha = 4$. El siguiente cálculo del cuarto momento que viene de las primeras estimaciones de Kloosterman para sus sumas individuales. Tenemos que

$$\tilde{K}_p(t) - \tilde{K}_p(s) = \frac{1}{\sqrt{p}} \sum_{n \in I} e\Big(\frac{an + b\overline{n}}{p}\Big),$$

donde I es el intervalo discreto

$$(p-1)s < n \leq (p-1)t$$

de sumación. La longitud de I es

$$\lfloor (p-1)t \rfloor - \lceil (p-1)s \rceil \leq 2(p-1)|t-s|$$

ya que $(p-1)|t-1| \geq 1$. Expandiendo la cuarta potencia tenemos que

$$\begin{aligned}\mathbb{E}_p(|\tilde{K}_p(t) - \tilde{K}_p(s)|^4) &= \frac{1}{(p-1)^2} \sum_{(a,b) \in \Omega_p} \left| \frac{1}{\sqrt{p}} \sum_{n \in I} e\Big(\frac{an + b\overline{n}}{p}\Big) \right|^4 \\ &= \frac{1}{p^2(p-1)^2} \sum_{a,b} \sum_{n_1,\ldots,n_4 \in I} e\Big(\frac{a(n_1 + n_2 - n_3 - n_4)}{p}\Big) e\Big(\frac{b(\overline{n}_1 + \overline{n}_2 - \overline{n}_3 - \overline{n}_4)}{p}\Big).\end{aligned}$$

Cambiando el orden de los sumatorios conseguimos separar las variables a y b, obteniendo así

$$\frac{1}{p^2(p-1)^2} \sum_{n_1,\ldots,n_4 \in I} \Big(\sum_{a \in \mathbb{F}_p^\times} e\Big(\frac{a(n_1 + n_2 - n_3 - n_4)}{p}\Big)\Big)\Big(\sum_{b \in \mathbb{F}_p^\times} e\Big(\frac{b(\overline{n}_1 + \overline{n}_2 - \overline{n}_3 - \overline{n}_4)}{p}\Big)\Big).$$

Las relaciones de ortogonalidad para caracteres aditivos

$$\frac{1}{p} \sum_{h \in \mathbb{F}_p^\times} e\big(\frac{ah}{p}\big) = \delta(h,0) - \frac{1}{p}$$

(para cualquier $h \in \mathbb{F}_p$) implican que

$$\mathbb{E}_p(|\tilde{K}_p(t) - \tilde{K}_p(s)|^4) = \frac{1}{(p-1)^2} \sum_{\substack{n_1,\ldots,n_4 \in I \\ n_1 + n_2 = n_3 + n_4 \\ \overline{n}_1 + \overline{n}_2 = \overline{n}_3 + \overline{n}_4}} 1 + O(|I|^3 (p-1)^{-3}). \tag{4.11}$$

Fijemos primero n_1 y n_2 en I tales que $n_1 + n_2 \neq 0$. Entonces si n_3, n_4 satisfacen

$$n_1 + n_2 = n_3 + n_4, \quad \overline{n}_1 + \overline{n}_2 = \overline{n}_3 + \overline{n}_4,$$

el valor de n_3+n_4 es fijo, y $\overline{n}_1+\overline{n}_2$ es no nulo, entonces $n_3n_4 = \frac{n_3+n_4}{\overline{n}_1+\overline{n}_2}$ es fijo (en $\mathbb{F}_p^\times$). Por lo tanto, hay a lo más dos duplas (n_3, n_4) que satisfacen las ecuaciones para esos (n_1, n_2) dados. Esto significa que la aportación de esos (n_1, n_2) a (4.11) es $\leq 2|I|^2(p-1)^{-2}$. De forma similar, si $n_1 + n_2 = 0$, la ecuación implica que $n_3 + n_4 = 0$, por lo que la solución solo depende de (n_1, n_3). Luego la aportación en este caso sería $\leq |I|^2(p-1)^{-2}$, y tendríamos

$$\mathbb{E}_p(|\tilde{K}_p(t) - \tilde{K}_p(s)|^4) \ll |I|^2(p-1)^{-2} + |I|^3(p-1)^{-3} \ll |t-s|^2,$$

donde las constantes implícitas son absolutas. Usando (4.10), tenemos que

$$\mathbb{E}_p(|K_p(t) - K_p(s)|^4) \ll |t-s|^2 \tag{4.12}$$

con una constante implícita absoluta. Combinado con (4.9) y $\alpha = 4$ en el rango anterior, esto completa la prueba de ajuste.

Paso Final.- Para acabar la prueba del Teorema 4.3.1, por la Proposición B.3.26, se tiene directamente de los Pasos 2 y 3. □

4.4. Aplicaciones

Gracias al Teorema 4.3.1, podemos obtener algo de información de los caminos de Kloosterman. En esta sección veremos dos ejemplos de aplicaciones del Teorema de distribución, el primero sobre valores grandes de las sumas parciales, y el otro sobre el soporte de las sumas parciales.

Teorema 4.4.1. *Para p primo y $A > 0$, sea $M_p(A)$ y $N_p(A)$ los sucesos*

$$M_p(A) := \left\{(a,b) \in \Omega_p \ : \ \max_{1\leq j\leq p-1} \frac{1}{\sqrt{p}}\left|\sum_{1\leq n\leq j} e\left(\frac{an+b\overline{n}}{p}\right)\right| > A\right\},$$

$$N_p(A) := \left\{(a,b) \in \Omega_p \ : \ \max_{1\leq j\leq p-1} \frac{1}{\sqrt{p}}\left|\sum_{1\leq n\leq j} e\left(\frac{an+b\overline{n}}{p}\right)\right| \leq A\right\}.$$

Existe una constante positiva $c > 0$ tal que, para cualquier $A > 0$ se tiene que

$$c^{-1}\exp(-\exp(cA)) \leq \liminf_{p\to+\infty} \mathbb{P}_p(N_p(A)) \leq \limsup_{p\to+\infty} \mathbb{P}_p(M_p(A)) \leq c\exp(-\exp(c^{-1}A)).$$

Demostración. Las funciones $t \mapsto K_p(a,b)(t)$ describen caminos poligonales en $\mathbb{C}$. Dado que el módulo máximo de un punto del camino se alcanza en uno de los vértices, tenemos que

$$\max_{1\leq j\leq p-1} \frac{1}{\sqrt{p}}\left|\sum_{1\leq n\leq j} e\left(\frac{an+b\overline{n}}{p}\right)\right| = \|K_p(a,b)\|_\infty,$$

por lo que el evento $M_p(A)$ es el mismo que $\{\|K_p\|_\infty > A\}$, y el evento $N_p(A)$ es el mismo que $\{\|K_p\|_\infty \geq A\}$. Por el Teorema 4.3.1 y la composición con la aplicación norma (Proposición B.3.16), las variables aleatorias reales $\|K_p\|_\infty$ convergen en ley a la

variable aleatoria $\|K\|_\infty$, la norma de serie aleatoria de Fourier K. Por las propiedades de la convergencia en ley

$$\mathbb{P}(\|K\|_\infty > A) \leq \liminf_{p\to+\infty} \mathbb{P}_p(N_p(A)) \leq \limsup_{p\to+\infty} \mathbb{P}_p(M_p(A)) \leq \mathbb{P}(\|K\|_\infty \geq A).$$

Por lo tanto, el problema se reduce a una cuestión sobre series aleatorias de Fourier límites.
Consideremos ahora la cota superior. Nos basta probar que existe una constante $c > 0$ tal que

$$\mathbb{P}(\|\Re(K)\|_\infty > A) \leq c\exp(-\exp(c^{-1}A)),$$
$$\mathbb{P}(\|\Im(K)\|_\infty > A) \leq c\exp(-\exp(c^{-1}A)).$$

Dado que la prueba para la parte real y la parte imaginaria son análogas, haremos solo la de la parte real. La variable aleatoria $\Re(K)$ toma valores en el espacio real separable de Banach $C_\mathbb{R}([0,1])$ de funciones reales continuas en $[0,1]$. Es casi seguro la suma de la serie aleatoria de Fourier

$$R := \sum_{h\geq 0} \varphi_h Y_h,$$

donde $\varphi_h \in C_\mathbb{R}([0,1])$ y las variables aleatorias Y_h están definidas como

$$\varphi_0(t) := 2t, \quad Y_0 := \tfrac{1}{2}ST_0,$$

$$\varphi_h(t) := \frac{\sin(2\pi ht)}{8\pi h}, \quad Y_h := \tfrac{1}{4}(ST_h - ST_{-h}) \text{ para } h \geq 1.$$

Podemos observar que las variables aleatorias (Y_h) son independientes y que $|Y_h| \leq 1$ (casi seguro) para todo h. Podemos, por tanto, aplicar la primera cota de la Proposición B.3.32 para probar la cota superior.
Probemos ahora la cota inferior. Nos es suficiente con probar que existe una constante $c > 0$ tal que

$$\mathbb{P}(|\Im(K(1/2))| > A) \geq c^{-1}\exp(-\exp(cA)), \tag{4.13}$$

dado que esto implica que

$$\mathbb{P}(\|K\|_\infty > A) \geq c^{-1}\exp(-\exp(cA)).$$

Tenemos que

$$\Im(K(1/2)) = -\frac{1}{2\pi}\sum_{h\neq 0} \frac{\cos(\pi h) - 1}{h} ST_h = \frac{1}{\pi}\sum_{h\geq 1} \frac{1}{h} ST_h,$$

que es una serie que converge casi seguro en $\mathbb{R}$ con términos independientes, y donde $\frac{1}{h}ST_h$ es simétrica y $|\frac{1}{h}ST_h| \leq 1$ para todo h. Por lo tanto, la cota

$$\mathbb{P}(|\Im(K(1/2))| > A) \geq c^{-1}\exp(-\exp(cA))$$

para algún $c > 0$ se tiene por la segunda parte de la Proposición B.3.32. □

Teorema 4.4.2. *El soporte de la distribución de K es el conjunto de todas las $f \in C_0([0,1])$ tales que*

1. *Tenemos que* $f(1) \in [-2, 2]$.

2. *Para todo* $h \neq 0$, *tenemos que* $\tilde{f}(h) \in i\mathbb{R}$ *y*

$$|\tilde{f}(h)| \leq \frac{1}{\pi|h|}.$$

Demostración. Sea $\mathcal{S}$ el conjunto mencionado en el enunciado. Entonces $\mathcal{S}$ es cerrado en $C([0,1])$, pues es la intersección de cerrados. Por el Teorema 4.3.1, una función muestra $f \in C([0,1])$ del proceso aleatorio K viene dada casi seguro por la serie

$$f(t) = \alpha_0 t + \sum_{h \neq 0} \frac{e(ht) - 1}{2i\pi h} \alpha_h$$

que es uniformemente convergente en el sentido se las sumas parciales simétricas, para algunos números reales α_h tales que $\|\alpha_h\| \leq 2$. Tenemos que $\tilde{f}(0) = f(1) \in [-2, 2]$, y la convergencia uniforme implica que para $h \neq 0$, tenemos

$$\tilde{f}(h) = \frac{\alpha_h}{2i\pi h},$$

por lo que $f \in \mathcal{S}$. Consecuentemente, el soporte de K está contenido en $\mathcal{S}$.
Ahora probaremos la otra inclusión. Por el Lema B.3.33, el soporte de K contiene al conjunto de funciones continuas con expansiones en series simétricas uniformemente convergentes

$$\alpha_0 t + \sum_{h \neq 0} \frac{e(ht) - 1}{2i\pi h} \alpha_h$$

donde $\alpha_h \in [-2, 2]$ para todo $h \in \mathbb{Z}$. En particular, como 0 pertenece al soporte de la medida de Sato-Tate, $\mathcal{S}$ contiene a todas las sumas finitas de este tipo. Sea $f \in \mathcal{S}$ y sea $g(t) = f(t) - tf(1)$. Tenemos que

$$f(t) - tf(1) = \lim_{N \to +\infty} \sum_{|h| \leq N} \hat{g}(h) e(ht) \left(1 - \frac{|h|}{N}\right),$$

en $C_0([0,1])$, por la convergencia uniforme de las medias de Cèsaro de las series de Fourier de funciones periódicas continuas (ver para las medias de Cèsaro, por ejemplo [1] o [31],capítulo 3, Teorema 3.4). Evaluando en 0 y restando llegamos a

$$\begin{aligned} f(t) &= tf(1) + \lim_{N \to +\infty} \sum_{\substack{|h| \leq N \\ h \neq 0}} \tilde{f}(h)(e(ht) - 1)\left(1 - \frac{|h|}{N}\right) \\ &= tf(1) + \lim_{N \to +\infty} \sum_{\substack{|h| \leq N \\ h \neq 0}} \frac{\alpha_h}{2i\pi h}(e(ht) - 1)\left(1 - \frac{|h|}{N}\right) \end{aligned}$$

en $C([0,1])$, donde $\alpha_h = 2i\pi h \tilde{f}(h)$ para $h \neq 0$. Entonces $\alpha_h \in \mathbb{R}$ y $|\alpha_h| \leq 2$, ya que supusimos que $f \in \mathcal{S}$. Por lo tanto, cada función

$$tf(1) + \sum_{\substack{|h| \leq N \\ h \neq 0}} \frac{e(ht) - 1}{2i\pi h} \alpha_h \left(1 - \frac{|h|}{N}\right),$$

pertenece al soporte de K. Como el soporte es cerrado, concluimos que f también pertenece al soporte de K. □

5. Otros temas

> ¿Cómo osamos hablar de leyes del azar? ¿No es, acaso, el azar la antítesis de cualquier ley?
>
> *Bertrand Russell*

5.1. Introducción

Hasta ahora, hemos visto algunos de los resultados más importantes a día de hoy que involucran a la Teoría de la Probabilidad y a la Teoría de Números. El objetivo de este capítulo es dar una breve introducción a la perspectiva probabilística de otros de los problemas de la Teoría de Números, distintos a los mencionados en capítulos anteriores.

Empezaremos tratando la equidistribución módulo 1, dando algunos de los resultados asociados a este tema. Proseguiremos con una cuestión muy importante, los famosos *prime gaps*, las distancias entre primos consecutivos. Daremos una breve introducción, también, a la Teoría de Ratner. Continuaremos con algunos resultados sobre las sumas de Rademacher y los grafos de Ramanujan. Seguidamente, describiremos el modelo de Cramér, y daremos algunos de los problemas asociados a éste. Por último, mencionaremos algunas ideas interesantes para la aplicación de la probabilidad.

5.2. Equidistribución módulo 1

A lo largo de este trabajo hemos mencionado que el resultado con el que empezó todo fue el Teorema 1.3.2, desarrollado por Paul Erdös y Mark Kac, el cual nos afirmaba que tomando n como una variable aleatoria unifome discreta en $\{1, \dots, N\}$, la función $\omega(n)$ renormalizada converge en ley a la normal estándar $\mathcal{N}$ cuando $N \to +\infty$. Sin embargo, ya existán resultados que podemos atribuir a la Teoría Probabilística de Números, aquellos referentes a la teoría de la equidistribución módulo 1. Estos resultados se deben especialmente a Weyl [30]. De hecho esta teoría se originó para estudiar las partes fraccionarias de diversas sucesiones $(x_n)_{n\geq 1}$ de números reales, y el hecho de que en muchos casos, estas partes fraccionarias están equidistribuidas en el intervalo $[0, 1]$ con respecto a la medida de Lebesgue.

Podemos expresar esta idea en términos probabiísticos. Si denotamos por $\langle x \rangle$ a la parte fraccionaria de x, es decir, el menor número real tal que $x - \langle x \rangle \in \mathbb{Z}$. Podemos identificar este punto por $e(x) = e^{2i\pi x}$, o por su imagen en $\mathbb{R}/\mathbb{Z}$, según nos convenga. Dada una sucesión $(x_n)_{n\geq 1}$ de números reales, definimos las variables aleatorias S_n en $\Omega_N = \{1, \dots, N\}$ (con la medida de probabilidad uniforme) como

$$S_N(n) = \langle x_n \rangle.$$

Entonces decimos que la sucesión $(x_n)_{n\geq 1}$ está equidistribuida módulo 1 si las variables aleatorias convergen en ley a la distribución uniforme dx en $[0,1]$, con $N \to +\infty$. Weyl probó entre otras cosas, los siguientes resultados.

Teorema 5.2.1. *Se tienen los dos siguientes hechos.*

1. *Sea $P \in \mathbb{R}[X]$ un polinomio de grado $d \geq 1$ con término líder ξX^d, con $\xi \notin \mathbb{Q}$. Entonces la sucesión $(P(n))_{n\geq 1}$ es equidistribuida módulo 1.*
2. *Sea $k \geq 1$ un entero, y sea $\xi = (\xi_1, \ldots, \xi_d) \in (\mathbb{R}/\mathbb{Z})^d$. La clausura del conjunto $\{n\xi \ : \ n \in \mathbb{Z}\} \subset (\mathbb{R}/\mathbb{Z})^d$ es un subgrupo compacto de $(\mathbb{R}/\mathbb{Z})^d$ y las variables aleatorias $T-$valuadas en Ω_N definidas por*

$$K_N(n) := n\xi$$

convergen en ley cuando $N \to +\infty$ a la medida de probabilidad de Haar en T.

Daremos la prueba de la primera parte del teorema. Tanto la prueba de la segunda parte como del lema auxiliar que daremos a continuación se pueden encontrar en [30].

Lema 5.2.2. *Sea $(x_n)_{n\geq 1}$ una sucesión de números reales. Supongamos que para cualquier entero $h \neq 0$, la sucesión $(x_{n+h} - x_n)_{n\geq 1}$ es equidistribuida módulo 1. Entonces (x_n) es equidistribuida módulo 1.*

Demostración. Como dijimos antes, probaremos la primera parte del Teorema 5.2.1. Lo haremos por inducción sobre $d \geq 1$, recordemos que es el grado del polinomio $P \in \mathbb{R}[X]$. Asumimos, sin pérdida de generalidad, que $P(0) = 0$. Si $d = 1$, entonces $P = \xi X$ para algún número real ξ, y $P(n) = n\xi$. Antes supusimos que ξ era irracional, y el resultado se tiene por ser el caso $1-$*dimensional* de la segunda parte del teorema.

Supongamos ahora que $d = deg(P) \geq 2$, y la primera parte se tiene para polinomios de grado menor. Supongamos que $K_N(n) = \langle P(n) \rangle$. Entonces debemos considerar la sucesión auxiliar $K'_N(n) = \langle P(n+h) - P(n) \rangle$, por lo que tenemos el mismo problema pero para el polinomio

$$P(X+h) - P(X) = \xi(X+h)^d - \xi X^d + \cdots = d\xi X^{d-1} + \ldots$$

Ya que estos polinomios tienen grado $d-1$, y el coeficiente líder $\xi \notin \mathbb{Q}$, por la hipótesis de inducción las variables aleatorias K'_N convergen en ley a la medida de Lebesgue. Finalmente, por el Lema 5.2.2, también lo hacen las variables aleatorias K_N. □

Lo fascinante de la equidistribución módulo 1 es, que aunque estemos en situaciones en las que se cumpla, todavía siguen apareciendo problemas, aún por resolver, cuando se intenta estudiar las variaciones existentes. Uno de estos problemas es el estudio de los huecos en una sucesión equidistribuida módulo 1.

Sea (x_n) una sucesión de elementos en $\mathbb{R}/\mathbb{Z}$ equidistribuida módulo 1. Para $N \geq 1$ consideremos los primeros N valores

$$\{x_1, \ldots, x_N\}$$

de la sucesión. El complementario en $\mathbb{R}/\mathbb{Z}$ de este conjunto es la unión disjunta de intervalos (en [0,1), todos menos uno son subintervalos). El número de intervalos es

$\leq N$ (de hecho puede ser $< N$, ya que algunos x_i pueden coincidir). Entonces la pregunta es: ¿cual es la distribución de la longitud de estos huecos? Dicho de otro modo, los intervalos en cuestión son las componentes conexas de $\mathbb{R}/\mathbb{Z} \setminus \{x_1, \ldots, x_N\}$, y nos interesa la medida de Lebesgue de estas componentes conexas. Sea Ω_N el conjunto de los intervalos en $\mathbb{R}/\mathbb{Z} \setminus \{x_1, \ldots, x_N\}$ con la medida de probabilidad uniforme. Definimos las variables aleatorias

$$G_N(I) := N \text{ longitud}(I)$$

para $I \in \Omega_N$. Vemos que el salto medio va a ser aproximadamente de $1/N$, por lo que multiplicar por N lleva a una normalización natural donde la esperanza de G_N es aproximadamente 1. Para el caso de los puntos puramente aleatorios en $\mathbf{S}^1$ independientes, tenemos un resultado probabilístico clásico que nos dice que las variables aleatorias análogas convergen en ley a la distribución exponencial $\mathcal{E}$ en $[0, +\infty)$, es decir, la variable aleatoria tal que

$$\mathbb{P}(a < \mathcal{E} < b) = \int_a^b e^{-x}\, dx$$

para cualesquiera números reales no negativos $a < b$. A esto se le llama también "comportamiento de Poisson". Para cualquier sucesión (determinística, por tanto, aritmética) (x_n) que sea equidistribuida módulo 1, podemos preguntarnos si aparecerá alguna distribución de este estilo. Ya el caso particular de la sucesión $\langle n\xi \rangle >$ para un número irracional fijo ξ, nos lleva a una respuesta bastante llamativa e interesante : el Teorema de las tres longitudes, el cual fue conjeturado por Steinhaus y probado por Sós en [29]. Este teorema dice que hay a lo más tres longitudes distintas entre las partes fraccionales de $\langle n\xi \rangle >$ para $1 \leq n \leq N$, independientemente de N y de $\xi \notin \mathbb{Q}$.

5.3. Espacios entre primos

El Teorema de los Números Primos nos indica que el salto medio entre primos sucesivos de tamaño x es aproximadamente $\log x$. Un problema razonable en vista de la cantidad de conjeturas acerca de la distribución de los números primos es entender la distribución de estos saltos. Una manera de intentarlo es la que presentaremos en esta sección. Para cualquier entero $N \geq 1$, definimos el espacio probabilístico Ω_N como el conjunto de enteros n tales que $1 \leq n \leq N$, con la medida de probabilidad uniforme. Fijemos $\lambda > 0$. Definimos entonces las variables aleatorias

$$G_{\lambda,N}(n) := \pi(n + \lambda \log n) - \pi(n)$$

que mide la cantidad de primos en el intervalo que empieza en n de longitud $\lambda \log n$. Sobre la distribución límite de las variables $G_{\lambda,N}$ existe la siguiente conjetura.

Conjetura 5.3.1. *La sucesión $(G_{\lambda,N})_N$ converge en ley cuando $N \to +\infty$ a una distribución de Poisson con parámetro λ, es decir, para cualquier entero $r \geq 0$, se tiene que*

$$\mathbb{P}_N(G_{\lambda,N} = r) \to e^{-\lambda} \frac{\lambda^r}{r!}.$$

Para curiosidad del lector, esta conjetura apareció por primera vez en los trabajos de Gallagher [10], quien probó que esta conjetura se tendría de una versión adecuada de conjetura de las k-tuplas de Hardy-Littlewood. Parte del interés de la Conjetura 5.3.1 es que la distibución obtenida para los saltos es exactamente lo que se puede esperar para conjuntos puramente aleatorios (consultar Feller [8]).

5.4. Teoría de Ratner

En esta sección presentaremos un resultado asociado a una de las teorías más versátiles y útiles de las Matemáticas, la Teoría de Ratner, llamada así por la mátemática Marina Ratner [27]. Esta teoría se relaciona con la teoría ergódica y los sistemas dinámicos. Si el lector tiene curiosidad por estos temas, puede consultar las obras de Ghys [11], Morris [24] o Einsiedler y Ward [5]. Antes de continuar, si el lector tiene alguna duda sobre los conceptos que aparecerán a continuación, como por ejemplo, espacio cociente o $\mathrm{GL}_2(\mathbb{R})$, puede consultar [9].

Figura 5.1: Marina Ratner, matemática que desarrolló la teoría homónima.

Para mostrar la utilidad de esta teoría en nuestro campo de estudio, presentaremos un resultado de Elkies y McMullen [6]. Para ello, consideremos las sucesiones de las partes fraccionarias de $\sqrt{n}$ para $n \geq 1$ (vistas como elementos de $\mathbb{R}/\mathbb{Z}$). Para $N \geq 1$, definimos el espacio probabilístico Ω_N como el conjunto de componentes conexas de $\mathbb{R}/\mathbb{Z} \setminus \{\langle 1\rangle, \ldots \langle\sqrt{N}\rangle\}$, con la medida de probabilidad uniforme, y definimos sobre Ω_N las variables aleatorias

$$G_N(I) := N \text{ longitud}(I).$$

El resultado mencionado anteriormente, debido a Elkies y McMullen, nos da la distribución límite de G_N cuando $N \to +\infty$, aunque por desgracia, no es una distribución de una familia genérica.

Teorema 5.4.1 (Elkies-McMullen)**.** *Cuando $N \to +\infty$, las variables aleatorias G_N converge en ley a una variable aleatoria en $[0, +\infty)$, con medida de probabilidad $\mu_{EM} = \frac{6}{\pi^2} f(x)dx$, donde f es una función continua, analítica en los intervalos $[0, 1/2]$, $[1/2, 2]$ y $[2, +\infty)$, no es de clase $\mathbf{C}^3$, y cumple $f(x) = 1$ si $0 \leq x \leq 1/2$.*

La restricción de f en los intervalos $[1/2, 2]$ y $[2, +\infty)$ se puede escribir de forma explícita como combinación de funciones elementales. Por ejemplo, para $1/2 \leq x \leq 2$,

sea $r = \frac{1}{2}x^{-1}$ y

$$\psi(r) = \arctan\left(\frac{2r-1}{\sqrt{4r-1}}\right) - \arctan\left(\frac{1}{\sqrt{4r-1}}\right),$$

por lo que tendríamos que

$$f(x) = \frac{2}{3}(4r-1)^{3/2}\psi(r) + (1-6r)\log r + 2r - 1.$$

Para mostrar qué tipo de resultados se asocian a la Teoría de Ratner, daremos un breve esquema de la prueba. La prueba estudia la distribución de los saltos mediante la función L_N definida para $x \in \mathbb{R}/\mathbb{Z}$ tal que $L_N(x)$ nos devuelve la medida del intervalo de salto que contiene a x, siendo el valor de la función 0 si x es de los puntos frontera de los intervalos de los saltos para $\langle 1\rangle, \ldots \langle\sqrt{N}\rangle$. Podemos comprobar, para cualquier $t \in \mathbb{R}$, la medida total en $\mathbb{R}/\mathbb{Z}$ de los puntos que están en un intervalo de salto de longitud menor o igual que t, que es igual a la medida de Lebesgue

$$\mu(\{x \in \mathbb{R}/\mathbb{Z} \ : \ L_N(x) < t\}),$$

viene dada por

$$\int_0^t t \; d(\mathbb{P}_N(G_N < t)) = t\mathbb{P}_N(G_N < t) - \int_0^t \mathbb{P}_N(G_N < t) \; dt.$$

Esto nos dice que basta con entender el comportamiento límite de L_N para así entender la distribución límite de los saltos. Fijemos $t \geq 0$. La idea clave que involucra a la Teoría de Ratner es que si N es el cuadrado de un entero, entonces la probabilidad

$$\mu(\{x \in \mathbb{R}/\mathbb{Z} \ : \ L_N(x) < t\})$$

se puede ver, asintóticamente, cercana a la probabilidad de que cierto retículo afin $\Lambda_{N,x}$ de $\mathbb{R}^2$ interseque al triángulo Δ_t con vértices $(0,0)$,$(1,0)$ y $(0,2t)$ con área t. El retículo es de la forma $\Lambda_{N,x} = g_{N,x} \cdot \mathbb{Z}^2$, para alguna transformación afín $g_{N,z}$. Sea $\mathrm{ASL}_2(\mathbb{R})$ el grupo de transformaciones afines

$$z \mapsto z_0 + g(z)$$

de $\mathbb{R}^2$ cuya parte lineal $g \in \mathrm{GL}_2(\mathbb{R})$ tiene determinante 1, y $\mathrm{ASL}_2(\mathbb{Z})$ el subgrupo de esas transformaciones afines con determinante 1 y coeficientes de ambas partes en $\mathbb{Z}$). Entonces los retículos se pueden ver como elementos del espacio cociente

$$M = \mathrm{ASL}_2(\mathbb{Z})/\mathrm{ASL}_2(\mathbb{R})$$

que parametriza retículos afines $\Lambda \subset \mathbb{R}^2$ con $\mathbb{R}^2/\Lambda$ de área 1. Este espacio admite una medida de probabilidad única $\tilde{\mu}$ que es invariante a la acción por la derecha de $\mathrm{ASL}_2(\mathbb{R})$ por multiplicación. Ahora tenemos, para cada $N \geq 1$, una medida de probabilidad μ_N en M, dla distribución de la variable aleatoria $\mathbb{R}/\mathbb{Z} \to M$ definida por $x \mapsto \Lambda_{N,x}$. La Teoría de Ratner nos proporciona un conjunto de herramientas muy útiles para probar que ciertas medidas en M (o espacios de construcción similar) son iguales a la medida de probabilidad canónica $\tilde{\mu}$, lo que lleva a que la sucesión entera converge en ley a $\tilde{\mu}$. Por lo tanto, se deduce que

$$\mu(\{x \in \mathbb{R}/\mathbb{Z} \ : \ L_N(x) < t\}) \to \tilde{\mu}(\{\Lambda \in M \ : \ M \cap \Delta_t \neq \emptyset\}).$$

Esto nos da, en principio, una forma explícita de la distribución de los saltos, aunque calcularla de forma exacta no es ni mucho menos sencillo.

5.5. Sumas de Rademacher

En esta sección tenemos por objetivo enunciar y probar un resultado relacionado con la distribución de las sumas de Rademacher. Estas sumas tienen la siguiente estructura:

$$\sum_{n\geq 1} \pm x_n, \quad x_1, x_2, \ldots \in \mathbb{R}.$$

El resultado que presentaremos a continuación, junto con los resultados auxiliares y más información adicional, se puede encontrar en el artículo de Montgomery-Smith [23]. Las demostraciones de los resultados auxiliares se pueden encontrar en este mismo artículo, pues debido a que escapan de los objetivos de este trabajo, srán omitidas. Antes de enunciar el resultado, debemos mencionar los objetos que participan en él. En primer lugar, tenemos las variables aleatorias independientes $\varepsilon_1, \varepsilon_2, \ldots$ siguiendo todas ellas una distribución de Bernoulli con función de probabilidad

$$\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = 1/2,$$

para todo $i \in \mathbb{N}$. Es por ello que las sumas a considerar son de la forma $\sum_n \varepsilon_n x_n$, donde $(x_n)_{n\geq 1}$ es una sucesión de números reales tales que $x = (x_n)_{n\geq 1} \in \ell_2$. Definimos, también, la siguiente norma:

$$K(x,t;\ell_1,\ell_2) := K_{1,2}(x,t) = \text{ínf}\{\|x'\|_1 + t\|x''\|_2 \ : \ x', x'' \in \ell_2, \ x' + x'' = x\}.$$

El resultado que presentaremos a continuación se basa en acotar la probabilidad de que la suma sea superior a ciertos valores. Existen otras cotas, como

$$\mathbb{P}\left(\sum_{n\geq 1} \varepsilon_n x_n > t\|x\|_2\right) \leq e^{-\frac{1}{2}t^2}. \tag{5.1}$$

Sin embargo, si $\|x\|_1 < +\infty$, esta expresión no nos da una buena cota inferior, ya que tendríamos otra cota superior

$$\mathbb{P}\left(\sum_{n\geq 1} \varepsilon_n x'_n > \|x'\|_1\right) = 0. \tag{5.2}$$

Ahora sí, podemos enunciar el resultado.

Teorema 5.5.1. *Existe una constante c tal que para todo $x \in \ell_2$ y $t > 0$ se tiene*

$$\mathbb{P}\left(\sum_{n\geq 1} \varepsilon_n x_n > K_{1,2}(x,t)\right) \leq e^{-\frac{1}{2}t^2}$$

y

$$\mathbb{P}\left(\sum_{n\geq 1} \varepsilon_n x_n > c^{-1}K_{1,2}(x,t)\right) \geq c^{-1}e^{-ct^2}.$$

Para probar este teorema, debemos dar antes algunos resultados previos, cuyas demostraciones no son concernientes a los objetivos ni temática de este trabajo.

Definición 5.5.2. *Para $x \in \ell_2$ y $t > 0$, definimos la norma*

$$J(x,t;\ell_\infty,\ell_2) := J_{\infty,2}(x,t) = \text{máx}\{\|x\|_\infty, t\|x\|_2\}.$$

Lema 5.5.3. *Para $t > 0$, los espacios $\big(\ell_2, K_{1,2}(\cdot,t)\big)$ y $\big(\ell_2, J_{\infty,2}(\cdot,t)\big)$ tienen una relación de dualidad mutua, es decir, uno es el dual del otro y viceversa. Es decir, para $x \in \ell_2$ tenemos que*

$$K_{1,2}(x,t) = \sup\left\{\sum_{n\geq 1} x_n y_n \ : \ y \in \ell_2,\ J_{\infty,2}(y,t^{-1}) \leq 1\right\}.$$

Definición 5.5.4. *Para $x \in \ell_2$ y $t \in \mathbb{N}$, definimos la norma*

$$\|x\|_{P(t)} := \sup\left\{\sum_{m=1}^{t}\left(\sum_{n\in B_m}|x_n|^2\right)^{\frac{1}{2}}\right\},$$

donde el supremo se toma sobre todos los conjuntos disjuntos $B_1,\ldots,B_t \subset \mathbb{N}$.

Lema 5.5.5. *Si $x \in \ell_2$ y $t^2 \in \mathbb{N}$, entonces*

$$\|x\|_{P(t^2)} \leq K_{1,2}(x,t) \leq \sqrt{2}\|x\|_{P(t^2)}.$$

El siguiente lema se debe a Paley y Zygmund.

Lema 5.5.6. *Sea $x \in \ell_2$, entonces, dado $0 < \lambda < 1$ tenemos que*

$$\mathbb{P}\left(\sum_{n\geq 1}\varepsilon_n x_n > \lambda\|x\|_2\right) \geq \frac{1}{3}(1-\lambda^2)^2.$$

Ahora sí, probaremos el Teorema 5.5.1.

Demostración. En primer lugar, probaremos que

$$\mathbb{P}\left(\sum_{n\geq 1}\varepsilon_n x_n > K_{1,2}(x,t)\right) \leq e^{-\frac{1}{2}t^2}.$$

Dado $\delta > 0$, sea $x', x'' \in \ell_2$ tal que $x' + x'' = x$ y

$$(1+\delta)K_{1,2}(x,t) > \|x'\|_1 + t\|x''\|_2.$$

Entonces tenemos que

$$\begin{aligned}\mathbb{P}\left(\sum_{n\geq 1}\varepsilon_n x_n > (1+\delta)K_{1,2}(x,t)\right) &\leq \mathbb{P}\left(\sum_{n\geq 1}\varepsilon_n x'_n > \|x'\|_1\right) + \mathbb{P}\left(\sum_{n\geq 1}\varepsilon_n x''_n > t\|x''\|_2\right)\\ &\leq 0 + e^{-\frac{1}{2}t^2},\end{aligned}$$

donde la última desigualdad se tiene por (5.1) y (5.2). Tomando $\delta \to 0$, tenemos la cota superior probada.

Ahora debemos probar que, para alguna constante c tenemos que

$$\mathbb{P}\Big(\sum_{n\geq 1}\varepsilon_n x_n > c^{-1}K_{1,2}(x,t)\Big) \geq c^{-1}e^{-ct^2}.$$

En primer lugar, supongamos que $t^2 \in \mathbb{N}$. Dado $\delta > 0$, sean $B_1, \ldots, B_{t^2} \subset \mathbb{N}$ disjuntos tal que $\bigcup_{m=1}^{t^2} B_m = \mathbb{N}$ y

$$\|x\|_{P(t^2)} \leq (1+\delta)\sum_{m=1}^{t^2}\Big(\sum_{n\in B_m}|x_n|^2\Big)^{\frac{1}{2}}.$$

Entonces

$$\begin{aligned}
\mathbb{P}\Big(\sum_{n\geq 1}\varepsilon_n x_n > \frac{1}{2}K_{1,2}(x,t)\Big) &\geq \mathbb{P}\Big(\sum_{n\geq 1}\varepsilon_n x_n > \tfrac{1}{\sqrt{2}}\|x\|_{P(t^2)}\Big)\\
&\geq \mathbb{P}\Big(\sum_{m=1}^{t^2}\sum_{n\in B_m}\varepsilon_n x_n \geq \tfrac{1}{\sqrt{2}}(1+\delta)\sum_{m=1}^{t^2}\Big(\sum_{n\in B_m}|x_n|^2\Big)^{\frac{1}{2}}\Big)\\
&\geq \prod_{m=1}^{t^2}\mathbb{P}\Big(\sum_{n\in B_m}\varepsilon_n x_n \geq \tfrac{1}{\sqrt{2}}(1+\delta)\Big(\sum_{n\in B_m}|x_n|^2\Big)^{\frac{1}{2}}\Big)\\
&\geq \Big(\tfrac{1}{3}\big(1-\tfrac{1}{2}(1+\delta)^2\big)^2\Big)^{t^2},
\end{aligned}$$

donde el último paso se tiene por el Lema 5.5.6. Si tomamos $\delta \to 0$, vemos que

$$\mathbb{P}\Big(\sum_{n\geq 1}\varepsilon_n x_n > \tfrac{1}{2}K_{1,2}(x,t)\Big) \geq \exp(-(\log 12)t^2).$$

Esto prueba el resultado para $t^2 \in \mathbb{N}$. Para $t \geq 1$, vemos que

$$K_{1,2}(x,t) \leq K_{1,2}(x,\lceil t\rceil) \quad y \quad \lceil t\rceil^2 \leq 4t^2,$$

por lo que tenemos el resultado (para $c = 4\log 12$). Para $t < 1$, se deduce directamente del Lema 5.5.6 y la Fórmula de Holmsted A.2.6. □

5.6. Grafos de Ramanujan

El objetivo principal de esta sección es dar una breve introducción a los grafos de Ramanujan, y mostrar el porqué de su importancia. Para ello, mecionaremos algunas ideas presentadas en el artículo de Lubotzky [21]. En primer lugar, debemos definir qué es un grafo de Ramanujan. Sea X un grafo finito conexo $k-regular$ de n vértices, con $k \geq 3$, y sea A su matriz de adyacencia de dimensiones $n \times n$. Si suponemos que es simétrico, entonces todos los autovalores λ de A son reales, y se tiene que $|\lambda| \leq k$, k es un autovalor y $-k$ también lo es si y solo si X es bipartito.

Definición 5.6.1 (Grafo de Ramanujan)**.** *En las condiciones anteriores, se dice que X es un grafo de Ramanujan si todo autovalor λ de A satisface, o bien $|\lambda| = k$, o $|\lambda| \leq 2\sqrt{k-1}$.*

La cota que aparece en la segunda condición, $2\sqrt{k-1}$, no es casual. De hecho, en el artículo de Lubotzky, Phillips y Sarnak [22], se prueba que $2\sqrt{k-1}$ es la mejor cota posible para una familia infinita de grafos $k-regulares$. Esta cota tiene un significado que proviene de la topología algebraica. El recubrimiento universal del grafo X, $\tilde{X}$, descrito anteriormente es T_k, el árbol infinito $k-regular$. Y es un resultado de Kestern el que nos indica que el espectro del operador de adyacencia sobre $L^2(T_k)$ es el intervalo $[-2\sqrt{k-1}, 2\sqrt{k-1}]$. Por lo que, para X, ser un grafo de Ramanujan significa que todos sus autovalores no triviales están en el espectro de su recubrimiento universal $\tilde{X}$.

Otra de las propiedades de estos grafos es que son expansores óptimos desde el punto de vista espectral. Veámoslo, mediante otra definición.

Definición 5.6.2. *Se dice que un grafo X, con las condiciones anteriores, es ε-expansor si $h(X) \geq \varepsilon$, donde $h(X)$ es la constante de Cheeger de X, la cual viene definida como*

$$h(X) := \text{mín}\left\{ \frac{|A(Y, \overline{Y})|}{|Y|} \ : \ Y \subset X, \ |Y| \leq \tfrac{|X|}{2} \right\},$$

donde $A(Y, \overline{Y})$ es el conjunto de aristas entre Y y su complementario $\overline{Y}$.

Si ahora denotamos por $\lambda_1(X) = \text{máx}\{\lambda \ : \ \lambda \neq k, \ \lambda \text{ autovalor de } X\}$. Entonces

$$\frac{h^2|X|}{2k} \leq \lambda_1(X) \leq 2h(X).$$

Si el lector quiere consultar la demostración de este resultado, junto con más resultados de este tipo, puede consultar el artículo [20]. Por lo tanto, los grafos de Ramanujan son expansores. Este tipo de grafos son muy útiles en combinatoria y ciencias de la computación, y por supuesto, en la matemática pura pues son clave en la construcción de muchos algoritmos y redes. En lo que a la probabilidad se refiere, estos grafos son ideales para la construcción de caminos aleatorios contenidos en ellos, pues convergen rápidamente a la distribución uniforme, y en los grafos de Ramanujan convergen de la forma más rápida posible.

Por último, estos grafos también tienen una conexión muy sorprendente con la Teoría de Números. Esto se debe a Ihara, que definió la noción de función zeta de un grafo k-regular X y Sunada, que observó que el grafo X es de Ramanujan si y solo si su función zeta satisface la Hipótesis de Riemann. Esto se puede ver con más detalle en [20].

Si el lector tiene interés en estos grafos, puede consultar los artículos y libros mencionados, pues en ellos se encuentra una amplia variedad de información y resultados sobre ellos.

5.7. Modelo de Cramér. Problemas asociados

En esta sección presentaremos el modelo probabilístico de Cramér para los números primos, junto con algunos de los problemas que presentaba. Estos problemas, junto con la descripción que daremos a continuación del modelo, se pueden encontrar en el artículo de Pintz [25].

Este modelo, propuesto por Cramér a mediados de los años 30, juega, incluso a día de hoy, un papel fundamental a la hora de formular conjeturas sobre los números primos. Por el Teorema de los Números Primos, sabemos que la densidad esperada de los primos menores o iguales que x es de $1/\log x$. El modelo probabilístico de Cramér es una sucesión de variables aleatorias Bernoulli independientes ξ_n, definidas para $n \geq 3$, cuyas funciones de probabilidad son

$$\mathbb{P}(\xi_n = 1) = \frac{1}{\log n}, \quad \mathbb{P}(\xi_n = 0) = 1 - \frac{1}{\log n}.$$

Cramér supuso que para ciertos problemas ξ_n reproduce bien el comportamiento de la función caractrística de los primos, es decir, la función $\chi_{\mathbb{P}}(n)$ que toma el valor 1 cuando n es primo y 0 en caso contrario.

Figura 5.2: Harald Cramér, matemático que desarrolló el modelo que se trata en esta sección.

A modo de aplicación, Cramér usó su modelo para para formular una conjetura sobre saltos grandes entre primos consecutivos. Para cualquier sucesión que tome los valores $\{0, 1\}$ que se corresponda a la de los valores de $(\xi_n)_{n\geq 3}$, podemos asociarle las series P_ν, donde

$$P_{\nu+1} = m \quad si \quad \sum_{n=3}^{m} \xi_n = \nu \quad y \quad \sum_{n=3}^{m-1} \xi_n = \nu - 1.$$

De acuerdo con la heurística de Cramér, los mayores saltos posibles entre primos se comportan de forma similar a los mayores saltos posibles entre P_n. De hecho, probó que, con probabilidad 1, se tiene

$$\limsup_{n\to+\infty} \frac{P_{n+1}-P_n}{(\log P_n)^2} = 1. \tag{5.3}$$

Partiendo de (5.3), Cramér conjeturó la misma relación para los números primos, es decir, si p_n es el n-ésimo número primo, entonces se tiene con probabilidad 1

$$\limsup_{n\to+\infty} \frac{p_{n+1}-p_n}{(\log p_n)^2} = 1. \tag{5.4}$$

En base a los datos empíricos, parece que la conjetura se adecúa bien (aunque no excesivamente bien) a éstos.

Daremos, ahora que tenemos el modelo definido, algunos de los problemas que presenta. En primer lugar, este modelo es demasiado simple como para englobar todas las propiedades que poseen los números primos. Por ejemplo, tenemos, con probabilidad 1, tantos valores pares como impares de los "primos probabilísticos" P_n, es decir

$$|\{P_n \leq x \ : \ 2|P_n\}| \sim |\{P_n \leq x \ : \ 2 \nmid P_n\}| \quad (x \to +\infty).$$

o también, que existen infinitos "primos probabilísticos" P_n consecutivos, es decir

$$\liminf_{n\to+\infty}(P_{n+1}-P_n) = 1.$$

El modelo, además, no es sensible a la propiedad que dado un número natural n, comprobar si es divisible o no por primos pequeños, algo crucial para estudiar la primalidad de n. Esta deficiencia se debe a que las dos propiedades anteriores no las cumplen los números primos, pues todos los primos mayores que dos son impares. Además, asumir que las variables aleatorias ξ_n son independientes es un craso error, pues la segunda propiedad anterior no la cumplen los primos.

Los problemas asociados mencionados anteriormente son sólo algunos de los existentes. Otras deficiencias, como la asociada a los primos gemelos o la descubierta por Maier en 1985, se pueden ver en [25].

5.8. Algunas ideas más

Para acabar, daremos algunos ejemplos más de cómo la Teoría de la Probabilidad puede complementar, con excelentes resultados, a la Teoría de Números.

- Aplicaciones de teoremas límite para medidas de probabilidad aritméticas en otros problemas de la Teoría Analítica de Números.
- Emplear herramientas probabilísticas para modelizar objetos aritméticos, dando ideas, conjeturas o incluso pruebas para problemas relacionados a estos objetos. Un ejemplo de esto es la sección anterior con el modelo de Cramér.

- Usar ideas de la Teoría de números para "desaleatorizar" algunas construcciones o algoritmos. De hecho, hay una gran cantidad de resultados que utilizan la aleatoriedad de ciertos objetos aritméticos para obtener pruebas o propiedades deterministas para objetos matemáticos cuyas propiedades, información o su misma existencia se ha obtenido mediante el uso de herramientas probabilísticas. Un ejemplo perfecto de esto es la construcción de los grafos de Ramanujan, como vimos en secciones anteriores.

Podemos ver, por tanto, que existe una gran variedad de usos de la Teoría de la Probabilidad en la Teoría de Números, tanto los ya descritos como los que están por aparecer. Y es que, parafraseando nuestro título y el de la famosa serie de Marvel, ¿qué pasaría si ... Riemann conoce a Bernoulli?

Anexos

A. Resultados generales

A.1. Álgebra

Definición A.1.1 (Libre de cuadrados)**.** *Se dice que un número entero $n \geq 0$ es libre de cuadrados si no existe primo p tal que $p^2|n$. Es decir, todos los factores primos de n tienen multiplicidad* 1*.*

Proposición A.1.2. *Todo entero positivo N posee, a lo más, un factor primo p tal que $p > \sqrt{N}$.*

Demostración. Sean p_1, p_2 factores primos distintos de N cumpliendo $p_1, p_2 > \sqrt{N}$. Entonces se tendria que

$$N \geq p_1 \cdot p_2 > \sqrt{N}\sqrt{N} = N,$$

lo cual lleva a un absurdo ($N > N$). □

Teorema A.1.3 (Teorema chino del resto)**.** *Para cada entero positivo n con factorización en primos*

$$n = p_1^{r_1} \dots p_k^{r_k}$$

existe un isomorfismo entre los anillos

$$\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/p_1^{r_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p_k^{r_k}\mathbb{Z}.$$

A.2. Análisis matemático

Definición A.2.1. *Supongamos que P es una propiedad de un número entero positivo, y $P(x)$ es la cantidad de números menores o iguales que x que cumplen P. Si se cumple*

$$\lim_{x \to +\infty} \frac{P(x)}{x} = 1,$$

entonces decimos que P se cumple para casi todo x.

Lema A.2.2 (Sumación por partes)**.** *Sea $(a_n)_{n \geq 1}$ una sucesión de números complejos y sea $f : [0, +\infty) \to \mathbb{C}$ una función de clase $\mathbf{C}^1$. Para todo $x \geq 0$ se define*

$$M_a(x) := \sum_{n \geq 1}^{x} a_n.$$

Entonces, para $x \geq 0$ se tiene

$$\sum_{n \geq 1}^{x} a_n f(n) = M_a(x) f(x) - \int_1^x M_a(t) f'(t) dt.$$

Si $M_a(x)f(x)$ tiende a 0 cuando $n \to +\infty$, entonces se tiene

$$\sum_{n\geq 1} a_n f(n) = -\int_1^{+\infty} M_a(t)f'(t)dt,$$

sabiéndose que si la integral o la serie converge absolutamente, ambas lo hacen.

Observación A.2.1. Para una prueba de este resultado, consultar [19], Lema A.1.1.

Definición A.2.3 (Transformada de Mellin)**.** *Sea $\varphi : [0, +\infty) \to \mathbb{C}$ una función continua que decrece más rápido que cualquier polinomio en el infinito. La transformada de Mellin de φ es la función holomorfa $\hat{\varphi}$ definida por la integral*

$$\hat{\varphi}(s) := \int_0^{+\infty} \varphi(s)x^s \frac{dx}{x},$$

para todos los $s \in \mathbb{C}$ para los que la integral tiene sentido, que en nuestro caso, incluye a todos los números complejos que cumplen $\Re(s) > 0$.

Teorema A.2.4 (Identidad de Plancherel)**.** *Sean X_k e Y_k las transformadas discretas de Fourier de x_n e y_n. Entonces se tiene que*

$$\sum_{0\leq n\leq N-1} x_n\overline{y}_n = \frac{1}{N}\sum_{0\leq k\leq N-1} X_k\overline{Y}_k$$

Teorema A.2.5 (Identidad de Parseval)**.** *Dada una función f de cuadrado integrable con coeficientes de Fourier c_n, se cumple que*

$$\|f\|^2 = \langle f, f\rangle_{[0,1]} = \int_0^1 |f(x)|^2\, dx = \sum_{n\in\mathbb{Z}} |c_n|^2.$$

Teorema A.2.6 (Fórmula de Holmsted)**.** *Sea $0 < p_1 < p_2 < +\infty$, y $1/\alpha = 1/p_1 + 1/p_2$. Entonces*

$$K(x, t; \ell_{p_1}, \ell_{p_2}) \sim \left(\sum_{s=0}^{t^\alpha} (x'_s)^{p_1}\right)^{1/p_1} + t\left(\sum_{s=t^\alpha}^{+\infty} (x'_s)^{p_2}\right)^{1/p_2},$$

donde x' es la reordenación decreciente de la sucesión $|x_s|$ en el intervalo $[0, +\infty)$.

Observación A.2.2. Para una prueba de este resultado, consultar [17], Teorema 4.1.

B. Resultados probabilísticos

B.1. Conceptos previos

En primer lugar, presentaremos algunos conceptos fundamentales referentes a la teoría de la probabilidad.

Definición B.1.1 (Espacio Probabilístico)**.** *Se denomina espacio probabilístico a la tripleta* $(\Omega, S, \mathbb{P})$*, donde*

1. Ω *es un conjunto no vacío.*
2. S *es una* $\sigma-$*álgebra que satisface*
   a) $\Omega \in S$.
   b) $A \in S \Rightarrow \bar{A} \in S$.
   c) $A_1, \ldots, A_n \in S \Rightarrow \bigcup_{i=1}^{n} A_i \in S$.
3. $\mathbb{P}$ *es una función de probabilidad, que asigna a cada elemento de* S *un valor del intervalo* $[0, 1]$.

Definición B.1.2 (Variable aleatoria)**.** *Sea* $(\Omega, S, \mathbb{P})$ *un espacio probabilístico. Una función finita, univaluada* X *que lleva* Ω *en* $\mathbb{R}$ *se denomina variable aleatoria si las imágenes inversas bajo* X *de todos los conjuntos borelianos de* $\mathbf{R}$ *son eventos del espacio muestral, es decir*

$$X^{-1}(B) = \{\omega | X(\omega) \in B\} \in S, \quad B \in \mathbb{B},$$

donde $\mathbb{B}$ *es el conjunto de todos los conjuntos borelianos.*

Definición B.1.3 (Esperanza y varianza)**.** *Sea* X *una variable aleatoria discreta con función de probabilidad* $\mathbb{P}[X = x_i]$ *que toma los valores* $\{x_1, \ldots, x_n\}$*. Se define su esperanza como*

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \mathbb{P}[X = x_i].$$

Sea Y *una variable aleatoria continua con función de densidad* $f_Y(y)$ *que toma valores en* $\mathbb{R}$*. Se define su esperanza como*

$$\mathbb{E}[Y] = \int_{\mathbb{R}} y f_Y(y) dy.$$

Para ambos casos, se define la varianza como

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

Definición B.1.4. *Sean* $X_1, \ldots, X_n$ *variables aleatorias definidas sobre el mismo espacio de probabilidad, con funciones de distribución* $F_{X_1}, \ldots, F_{X_n}$ *y función de distribución conjunta* F_X*. Se dice que dichas variables son mutuamente independientes (o, simplemente, independientes) si*

$$F_X(x_1, \ldots, x_n) = F_{X_1}(x_1) \ldots F_{X_n}(x_n), \quad \textit{para todo } x_1, \ldots, x_n \in \mathbb{R}.$$

B.2. Tipos de convergencia

Definición B.2.1 (Convegencia casi segura)**.** *Una sucesión de variables aleatorias* (X_n) *converge con probabilidad* 1 *o de forma casi segura a una variable aleatoria* X *si se tiene que*

$$\mathbb{P}\left(\lim_{n\to+\infty} X_n = X\right) = 1.$$

Esto se expresa como $X_n \xrightarrow{c.s.} X$

Definición B.2.2 (Convegencia en probabilidad)**.** *Una sucesión de variables aleatorias* (X_n) *converge en probabilidad a una variable aleatoria* X *si, para todo* $\varepsilon > 0$ *se tiene que*

$$\lim_{n\to+\infty} \mathbb{P}\big[|X_n - X| \geq \varepsilon\big] = 0.$$

Esto se expresa como $X_n \xrightarrow{P} X$

Definición B.2.3 (Convegencia en ley)**.** *Una sucesión de variables aleatorias* (X_n)*, con funciones de distribución* (F_n)*, converge en ley o distribución a una variable aleatoria* X*, con función de distribución* F*, si se tiene*

$$\lim_{n\to+\infty} F_n(x) = F(x)$$

para todo x *punto de continuidad de* F *y* F_n*. Esto se expresa como* $X_n \xrightarrow{\mathcal{L}} X$

Definición B.2.4 (Convegencia en momentos)**.** *Para* $r \geq 1$ *real, una sucesión de variables aleatorias* (X_n) *converge en* L^r *a una variable aleatoria* X *si los momentos* $\mathbb{E}(|X_n|^r)$ *y* $\mathbb{E}(|X|^r)$ *existen y se tiene*

$$\lim_{n\to+\infty} \mathbb{E}(|X_n - X|^r) = 0$$

En el siguiente esquema, podemos ver las relaciones entre los distintos tipos de convergencia. La flecha, considerando la dirección que toma, indica qué tipo de convergencia implica a otro.

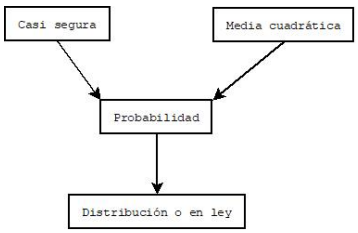


Figura B.1: Representación de las implicaciones entre tipos de convergencia

Para profundizar más en las relaciones entre los tipos de convergencia, consultar [4], capítulo 6.

B.3. Resultados importantes

Teorema B.3.1 (Kolmogorov)**.** *Sea $(X_n)_{n\geq 1}$ una sucesión de variables aleatorias independientes complejas tales que las series*

$$\sum_{n\geq 1} \mathbb{E}[X_n] \tag{B.1}$$

$$\sum_{n\geq 1} \mathbb{V}[X_n] \tag{B.2}$$

convergen. Entonces la serie

$$\sum_{n\geq 1} X_n$$

converge casi seguro, y por tanto, en ley.

Observación B.3.1. Para una prueba de este teorema, consultar [19], páginas 200-202.

Proposición B.3.2. *Sea M un espacio de Banach de dimensión finita. Sean $(X_n)_{n\geq 1}$ y $(X_{n,m})_{n\geq m\geq 1}$ variables aleatorias M-valuadas. Se define $E_{n,m} := X_n - X_{n,m}$. Si se cumple*

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. *Se tiene*
$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{E}(\|E_{n,m}\|) = 0.$$

Entonces las sucesiones (X_n) e (Y_m) convergen en ley, y además, lo hacen a la misma distribución límite.

Observación B.3.2. La prueba de esta Proposición depende de varios resultados auxiliares que no son relevantes para este trabajo. Puede consultarse en [19], páginas 180-181.

Lema B.3.3 (Borel-Cantelli)**.** *Sea $(A_n)_{n\geq 1}$ una colección de sucesos tales que si se cumple*

$$\sum_{n\geq 1} \mathbb{P}(A_n) = +\infty,$$

entonces un elemento del espacio probabilístico base pertenece casi seguro a una infinidad de conjuntos A_n, es decir, el conjunto $A := \cap_{n\geq 1} \cup_{k\geq n} A_k$ cumple $\mathbb{P}(A) = 1$.

Teorema B.3.4 (Método de los momentos)**.** *Sea $m \geq 1$ un entero.*

1. *Sea (μ_n) una sucesión de medidas de probabilidad en $\mathbb{R}^m$ tal que todos los momentos $M_k(\mu_n)$ existen, para $k \geq 1$, y sea μ otra medida de probabilidad en $\mathbb{R}^m$. Supongamos que μ es suave, es decir, es infinitas veces diferenciable. Entonces (μ_n) converge débilmente a μ si para cualquier $m-$tupla k de enteros no negativos se tiene*
$$M_k(\mu_n) \to M_k(\mu)$$
con $n \to +\infty$.

2. *Sea* $(\Omega, \Sigma, \mathbb{P})$ *un espacio de probabilidad. Sea* $(X_n)_{n\geq 1}$ *vectores aleatorios* $\mathbb{R}^m -$ *valuados en* Ω *tales que todos sus momentos* $M_k(X_n)$ *existen, y sea* Y *un vector aleatorio* $\mathbb{R}^m -$ *valuado. Supongamos que* Y *es suave. Entonces* (X_n) *converge en ley a* Y *si para cualquier* $m -$ *tupla* k *de enteros no negativos se tiene*

$$\mathbb{E}(X_{n,1}^{k_1} \dots X_{n,m}^{k_m}) \to \mathbb{E}(Y_1^{k_1} \dots Y_m^{k_m})$$

Observación B.3.3. Para una prueba de este resultado, consultar [2], Teoremas 30.1 y 30.2.

Teorema B.3.5 (Recíproco del método de los momentos)**.** *Sea* $(\Omega, S, \mathbb{P})$ *un espacio probabilístico. Sea* $m \geq 1$ *un entero y sean* $(X_n)_{n\geq 1}$ *vectores aleatorios* $\mathbb{R}^m -$ *valuados en* Ω *tales que todos sus momentos* $M_k(X_n)$ *existen, y tales que existen constantes* $c_k \geq 0$ *cumpliendo*

$$\mathbb{E}(|X_{n,1}|_1^k \dots |X_{n,m}|_m^k) \leq c_k$$

para todo $n \geq 1$*. Supongamos que* X_n *converge en ley a* Y*. Entonces* Y *es suave y para cualquier* $m -$ *tupla* k *de enteros no negativos se tiene*

$$\mathbb{E}(X_{n,1}^{k_1} \dots X_{n,m}^{k_m}) \to \mathbb{E}(Y_1^{k_1} \dots Y_m^{k_m})$$

Observación B.3.4. Para una prueba de este resultado, consultar [2], Teorema 25.12.

Lema B.3.6. *Sea* $m \geq 1$ *un entero. Sea* $(X_n)_{n\geq 1}$ *una sucesión de variables aleatorias variables en algún espacio probabilistico, con valores en* $\mathbb{R}^m$*. Sea* (β_n) *una sucesión de números reales positivos tales que* $\beta_n \to 0$ *cuando* $n \to +\infty$*. Si* (X_n) *converge en ley a una variable aleatoria* $\mathbb{R}^m -$ *valuada* X*, entonces para cualquier sucesión* (Y_n) *de variables aleatorias* $\mathbb{R}^m -$ *valuada tales que* $\|X_n - Y_n\|_\infty \leq \beta_n$ *para todo* $n \geq 1$*, las variables aleatorias* Y_n *convergen a* X*.*

Observación B.3.5. Para una prueba de este resultado, consultar [19], página 183.

Teorema B.3.7 (Teorema central del límite)**.** *Sea* $B \geq 0$ *un número real fijo. Sea* (X_n) *una sucesión de variables aleatorias reales independientes con* $|X_n| \leq B$ *para todo* n*. Sea*

$$\alpha_n := \mathbb{E}(X_n), \quad \beta_n := \mathbb{V}(X_n^2).$$

Sea $\sigma_N \geq 0$ *definido por*

$$\sigma_N^2 := \beta_1 + \dots + \beta_N$$

para $N \geq 1$*. Si* $\sigma_N \to +\infty$ *cuando* $N \to +\infty$*, entonces las variables aleatorias*

$$Y_N = \frac{(X_1 - \alpha_1) + \dots + (X_N - \alpha_N)}{\sigma_N}$$

convergen en ley a la variable aleatoria normal estándar $\mathcal{N}$*.*

Observación B.3.6. Para una prueba de este resultado, consultar [19], páginas 195-196.

Teorema B.3.8 (Teorema de las tres series de Kolmogorov)**.** *Sea* (X_n) *una sucesión de variables aleatorias independientes, y sea* $X_n^{(c)} := X_n I_{[|X_n| \leq c]}$ *la variable aleatoria* X_n *truncada en* c*. Consideremos las tres series siguientes:*

$$\sum_{n\geq 1} \mathbb{P}(|X_n| > c), \quad \sum_{n\geq 1} \mathbb{E}(X_n^{(c)}), \quad \sum_{n\geq 1} \mathbb{V}(X_n^{(c)}).$$

Para que la serie $\sum_{n\geq 1} X_n$ converja casi seguro es necesario que las tres series anteriores converjan para todo positivo c, y suficiente si convergen para algunos c positivos.

Observación B.3.7. Para una prueba de este resultado, consultar [2], Teorema 22.8.

Teorema B.3.9 (Desigualdad de Chebyshev)**.** *Sea X una variable aleatoria y $x \geq 0$ un número real. Se cumple*

$$\mathbb{P}[|X - \mathbb{E}(X)| \geq x] \leq \frac{\mathbb{V}(X)}{x^2}$$

Observación B.3.8. Para una prueba de este resultado, consultar [16], Teorema 1.1.

Proposición B.3.10. *Sea M un espacio métrico completo separable. Sea (X_n) una sucesión de variables aleatorias $M-$valuadas, y μ una medida de probabilidad en M. Entonces, X_n converge en ley a μ si y solo si se cumple*

$$\mathbb{E}(f(X_n)) \to \int_M f(x) d\mu(x)$$

para toda función Lipschitziana acotada $f : M \to \mathbb{C}$.

Observación B.3.9. Para una prueba de este resultado, consultar [19], Proposición B.4.1.

Lema B.3.11. *Sea $r > 0$ un número real. Se tiene que*

$$|\mathbb{E}_T(r^{-it})| \leq \text{mín}\left(1, \frac{1}{T|\log r|}\right).$$

Si $r = n_1/n_2$ para n_1, n_2 enteros positivos distintos, se cumple que

$$\mathbb{E}_T(r^{-it}) \ll \text{mín}\left(1, \frac{\sqrt{n_1 n_2}}{T}\right).$$

Observación B.3.10. Para una prueba de este Lema, consultar [19], Lema 3.2.6.

Proposición B.3.12. *Sea X una variable aleatoria y $a, r \geq 0$. Entonces*

$$\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(|X|^r)}{a^r}$$

Observación B.3.11. Para una prueba de este resultado, consultar [4], aplicar el Teorema 1, sección 3.4 con $h(X) = |X|$.

Teorema B.3.13 (Teorema de De Moivre-Laplace)**.** *Sean $n, k > 0$ enteros con $n > k$, y $p \in (0, 1)$. Entonces se tiene*

$$\frac{\sqrt{2\pi np(1-p)}\binom{n}{k}p^k(1-p)^{n-k}}{e^{-(k-np)^2/2np(1-p)}} \to 1$$

cuando $n \to +\infty$

Teorema B.3.14 (Teorema de Kroneker)**.** *Sea $d \geq 1$ un entero. Sea z un elemento de $\mathbb{R}^d$ y sea T (resp. $\tilde{T}$) la clausura del subgrupo de $(\mathbb{R}/\mathbb{Z})^d$ generado por la clase de z (resp. generado por las clases de los elementos yz con $y \in \mathbb{R}$).*

1. *Cuando $N \to +\infty$, las medidas de probabilidad en $(\mathbb{R}/\mathbb{Z})^d$ definidas por*

$$\frac{1}{N} \sum_{0 \leq n < N} \delta_{nz}$$

convergen en ley a la medida de probabilidad de Haar en T.

2. *Sea λ la medida de Lebesgue en $\mathbb{R}$. Cuando $X \to +\infty$, las medidas de probabilidad μ_x en $(\mathbb{R}/\mathbb{Z})^d$ definidas por*

$$\mu_X(A) := \frac{1}{X} \lambda(\{x \in [0, X] \ : \ xz \in A\})$$

para un subconjunto medible A de $(\mathbb{R}/\mathbb{Z})^d$, convergen en ley a la medida de probabilidad de Haar en $\tilde{T}$.

Observación B.3.12. Para una prueba de este resultado, consultar [19], Teorema B.6.5.

Corolario B.3.15. *Sea M un espacio de Banach separable. Sean (X_n) e (Y_n) sucesiones de variables aleatorias $M-$valuadas.Supongamos que la sucesión (X_n) converge en ley a la variable aleatoria X. Si (Y_n) converge en probabilidad o en L^p para algún $p \geq 1$ (se puede dar $p = +\infty$) a 0, entonces la sucesión $(X_n + Y_n)_n$ converge en ley a X en M.*

Observación B.3.13. Para una prueba de este resultado, consultar [19], Corolario B.4.2.

Proposición B.3.16. *Sea M un espacio métrico. Sea (X_n) una sucesión de variables aleatorias $M-$valuadas tales que X_n converge en ley a X. Para cualquier espacio métrico N y cualquier función continua $\varphi\colon M \to N$, las variables aleatorias $N-$valuada $\varphi \circ X_n$ convergen en ley a $\varphi \circ X$.*

Observación B.3.14. Para una prueba de este resultado, consultar [19], Proposición B.3.2.

Corolario B.3.17. *Sea M un espacio topológico compacto. Sea (X_n) una sucesión de variables aleatorias $M-$valuadas, definidas sobre algunos espacios probabilísticos Ω_n, tales que X_n converge en ley a X. Sea g una función continua en M tal que $g(X_n)$ converge en probabilidad a 0. Entonces, el soporte de X está contenido en el núcleo de g.*

Observación B.3.15. Para una prueba de este resultado, consultar [19], Proposición B.3.4.

Corolario B.3.18. *Sea V un espacio de Banach real o complejo, y $(v_n)_{n \leq 1}$ una sucesión de elementos de V. Sea $(X_n)_{n \leq 1}$ una sucesión de variables aleatorias acotadas en módulo por 1 casi seguro. Supongamos que la serie $\sum_n X_n v_n$ converge casi seguro en V, y sea X su suma. Si se cumple que*

$$\sum_{n \leq N} \|v_n\| \gg (\log N)^2, \quad \sum_{n > N} \|v_n\|^2 \ll \frac{\log N}{N},$$

entonces se tiene que

$$\mathbb{P}(\|X\| > A) \leq c \exp(-\exp(c^{-1}A^{1/2}))$$

y

$$c^{-1} \exp(-\exp(cA^{1/2})) \ll \mathbb{P}(\|X\| > A)$$

para algunos números reales $c, A > 0$.

Proposición B.3.19. *Sea* $\tau \in (1/2, 1)$ *y sea* $U_\tau = \{s \in \mathbb{C} \ : \ \Re(s) > \tau\}$.

1. *El producto de Euler aleatorio definido por*

$$Z(s) := \prod_p (1 - X_p p^{-s})^{-1}$$

converge casi seguro para cualquier $s \in U_\tau$. *Para cualquier subconjunto compacto* $K \subset U_\tau$, *la función aleatoria*

$$Z_K(s) := Z(s) \quad s \in K$$

es una variable aleatoria $H(K) - valuada$.

2. *La serie de Dirichlet aleatoria definida por*

$$\tilde{Z}(s) := \sum_{n \geq 1} X_n n^{-s}$$

converge casi seguro para cualquier $s \in U_\tau$.*Para cualquier subconjunto compacto* $K \subset U_\tau$, *la función aleatoria*

$$\tilde{Z}_K(s) := \tilde{Z}(s) \quad s \in K$$

es una variable aleatoria $H(K) - valuada$.

3. *Se tiene que* $\tilde{Z}_K = Z_K$ *casi seguro.*

Observación B.3.16. Para una prueba de este resultado, consultar [19], Proposición 3.2.9.

Proposición B.3.20. *Sea* $\varphi : [0, +\infty) \to [0, 1]$ *una función suave con soporte compacto tal que* $\varphi(0) = 1$. *Sea* $\hat{\varphi}$ *su transformada de Mellin. Para* $N \geq 1$ *se define la variable aleatoria* $H(D) - valuada$

$$Z_{D,N} := \sum_{n \geq 1} X_n \varphi\left(\frac{n}{N}\right) n^{-s}.$$

Entonces existe $\delta > 0$ *tal que*

$$\mathbb{E}(\|Z_D - Z_{D,N}\|_\infty) \ll N^{-\delta}$$

para $N \geq 1$.

Observación B.3.17. Para una prueba de este resultado, consultar [19], Proposición 3.2.11.

Proposición B.3.21. *Sea $\varphi : [0, +\infty) \to [0, 1]$ una función suave con soporte compacto tal que $\varphi(0) = 1$. Sea $\hat{\varphi}$ su transformada de Mellin (ver Definición A.2.3). Para $N \geq 1$ se define*

$$\zeta_N(s) := \sum_{n\geq 1} \varphi\left(\frac{n}{N}\right) n^{-s},$$

y sea $Z_{N,T}$ la variable aleatoria $H(D) - valuada$

$$t \mapsto (s \mapsto \zeta_N(s + it)).$$

Entonces existe $\delta > 0$ tal que

$$\mathbb{E}_T(\|Z_{D,T} - Z_{N,T}\|_\infty) \ll N^{-\delta} + NT^{-1}$$

para $N \geq 1$ y $T \geq 1$.

Observación B.3.18. Para una prueba de este resultado, consultar [19], Proposición 3.2.12.

Lema B.3.22. *Sea $(a(n))_{n\geq 1}$ una sucesión acotada de números complejos. Para cualesquiera $T \geq 2$ y $\sigma \leq 0$ tenemos*

$$\begin{aligned}\mathbb{E}_T\left(\left|\sum_{n\geq 1} \frac{a(n)}{n^{\sigma+it}}\right|^2\right) &= \sum_{n\geq 1} \frac{|a(n)|^2}{n^{2\sigma}} + O\left(\frac{1}{T} \sum_{\substack{n,m\geq 1\\ m\neq n}} \frac{|a(m)a(n)|}{(mn)^{\sigma-\frac{1}{2}}}\right)\\ &= \mathbb{E}\left(\left|\sum_{n\geq 1} \frac{X_n}{n^\sigma}\right|^2\right) + O\left(\frac{1}{T} \sum_{\substack{n,m\geq 1\\ m\neq n}} \frac{|a(m)a(n)|}{(mn)^{\sigma-\frac{1}{2}}}\right).\end{aligned}$$

Observación B.3.19. Para una prueba de este resultado, consultar [19], Proposición 4.2.6.

Teorema B.3.23 (Criterio de Weyl)**.** *Sea G un grupo abeliano compacto. Una sucesión de variables aleatorias $G - valuadas$ (X_n) converge en ley a una variable aleatoria uniformemente distribuida en G si y solo si, para cualquier caracter no trivial χ de G, se tiene que*

$$\lim_{n\to+\infty} \mathbb{E}(\chi(X_n)) \to 0.$$

Definición B.3.24 (Ajuste)**.** *Sea M un espacio métrico completo separable. Sea $(\mu_i)_{i\in I}$ una sucesión de medidas de probabilidad en M. Se dice que (μ_i) es ajustada si para cualquier $\varepsilon > 0$, existe un subconjunto compacto $K \subset M$ tal que $\mu_i(K) \geq 1 - \varepsilon$ para todo $i \in I$.*

Proposición B.3.25 (Criterio de ajuste de Kolmogorov)**.** *Sea (X_n) una sucesión de variables aleatorias $C([0, 1]) - valuadas$. Si existen números reales $a > 0$, $\delta > 0$ y $C \geq 0$ tales que, para cualesquiera números reales $0 \leq s < t \leq 1$ y $n \geq 1$, tenemos*

$$\mathbb{E}(|X_n(t) - X_n(s)|^\alpha) \leq C|t - s|^{1+\delta},$$

entonces (X_n) es ajustada.

Proposición B.3.26. *Sea (X_n) una sucesión de variables aleatorias $C_0([0,1])-$valuadas y sea X una variable aleatoria $C_0([0,1])-$valuada. Supongamos que $FT(X_n)$ converge a $FT(X)$ en el sentido de distribución finitas. Entonces (X_n) converge en ley a X en el sentido de las variables aleatorias $C_0([0,1])-$valuadas si y solo si (X_n) es ajustada.*

Observación B.3.20. Para una prueba de este resultado, consultar [19], Proposición B.11.8.

Lema B.3.27. *Sea (X_n) una sucesión de variables aleatorias $C_0([0,1])-$valuadas y sea X una variable aleatoria $C_0([0,1])-$valuada, todas ellas definidas en el mismo espacio probabilístico. Supongamos que, para todo $t \in [0,1]$, las variables aleatorias $(X_n(t))$ convergen a $X(t)$ en L^1. Entonces (X_n) converge a X en el sentido de distribuciones finitas.*

Observación B.3.21. Para una prueba de este resultado, consultar [19], Lema B.11.3.

Definición B.3.28 (Variable aleatoria subgaussiana)**.** *Sea $\sigma > 0$ un número real. Una variable aleatoria real X es σ^2-subgaussiana si tenemos que*

$$\mathbb{E}(\exp(tX)) \leq \exp(\sigma^2 t^2/2)$$

para todo $t \in \mathbb{R}$. Una variable aleatoria compleja Z es σ^2-subgaussiana si $Z = X+iY$, con X e Y variables aleatorias reales σ^2-subgaussiana.

Proposición B.3.29. *Se tienen los dos hechos siguientes.*

- *Sea X una variable aleatoria compleja y $m > 0$ un número real tal que $\mathbb{E}(X) = 0$ y $|X| \leq m$. Entonces X es m^2-subgaussiana.*
- *Sean X_1 y X_2 variables aleatorias independientes tales que X_i es σ_i^2-subgaussiana. Entonces $X_1 + X_2$ es $(\sigma_1^2 + \sigma_2^2)-$subgaussiana.*

Observación B.3.22. Para una prueba de este resultado, consultar [19], Proposición B.8.2.

Proposición B.3.30. *Sea $\sigma > 0$ un número real y sea X una variable aleatoria σ^2-subgaussiana, sea compleja o real. Para cualquier entero $k \geq 0$, existe una constante $c_k \geq 0$ tal que*

$$\mathbb{E}(|X|^k) \leq c_k \sigma^k.$$

Observación B.3.23. Para una prueba de este resultado, consultar [19], Proposición B.8.3.

Lema B.3.31. *Sea $m \geq 1$ un entero. Sean (X_n) e (Y_n) sucesiones de variables aleatorias $\mathbb{R}^m-$valuadas, sea (α_n) una sucesión en $\mathbb{R}^m$ y (β_n) una sucesión de números reales. Supongamos que*

1. *La sucesión (X_n) converge en ley a la variable aleatoria X, y $\|X_n\|$ está acotada por una constante $N \geq 0$, independiente de n.*
2. *Para todo n, tenemos que $\|Y_n\| \leq \beta_n$.*
3. *Tenemos que $\alpha_n \to (1, \ldots, 1)$ y $\beta_n \to 0$ cuando $n \to +\infty$.*

Entonces la sucesión $(\alpha_n \cdot X_n + Y_n)_n$ converge en ley a X en $\mathbb{R}^m$, donde $\cdot$ denota el producto por componentes de vectores.

Observación B.3.24. Para una prueba de este resultado, consultar [19], Corolario B.4.3.

Proposición B.3.32. *Sea V un espacio de Banach separable real o complejo y V' su dual. Sea $(X_n)_{n\geq 1}$ una sucesión de variables aleatorias independientes con $|X_n| \leq 1$ casi seguro, que son reales o complejas según el v¡cuerpo base. Sea $(v_n)_{n\geq 1}$ una sucesión de elementos de V. Supongamos que la serie $\sum X_n v_n$ converge casi seguro en V, y sea X su suma.*

1. *Supongamos que*

$$\sum_{n\leq N} \|v_n\| \ll \log(N), \quad \sum_{n>N} \|v_n\|^2 \ll \frac{1}{N} \tag{B.3}$$

   para todo $N \geq 1$. Entonces existe una constante $c > 0$ tal que para todo $A > 0$ tenemos que

$$\mathbb{P}(\|X\| > A) \leq c \exp(-\exp(c^{-1}A)).$$

2. *Supongamos que V es un espacio de Banach real, que las (X_n) son simétricas, idénticamente distribuidas y reales, y que existe $\lambda \in V'$ con norma 1 tal que*

$$\sum_{n\leq N} |\lambda(v_n)| \gg \log(N) \tag{B.4}$$

   para todo $N \geq 1$. Entonces existe una constante $c' > 0$ tal que para todo $A > 0$, tenemos

$$c'^{-1} \exp(-\exp(c'A)) \leq \mathbb{P}(|\lambda(X)| > A) \leq \mathbb{P}(\|X\| > A).$$

Observación B.3.25. Para una prueba de este resultado, consultar [19], Proposición B.11.13.

Lema B.3.33. *Sea M un espacio topológico compacto. Sea (X_n) una sucesión de variables aleatorias $M-$valuadas, definidas en el mismo espacio probabilístico Ω_n. Supongamos que (X_n) converge en ley a una variable aleatoria X, y sea $N \subset M$ el soporte de la distribución de X.*

1. *Para cualquier $x \in N$ y para cualquier entorno abierto U de x tenemos que*

$$\liminf_{n\to+\infty} \mathbb{P}(X_n \in U) > 0,$$

   y en particular existen un $n \geq 1$ y un $\omega \in \Omega_n$ tal que $X_n(\omega) \in U$.

2. *Para cualquier $x \in M \setminus N$, existe un entorno abierto U de x tal que*

$$\limsup_{n\to+\infty} \mathbb{P}(X_n \in U) = 0.$$

Observación B.3.26. Para una prueba de este resultado, consultar [19], Lema B.3.3.

Proposición B.3.34. *Sea (X_n) una sucesión de variables aleatorias con distribución de Poisson, con parámetro (λ_n). Supongamos que $\lambda_n \to +\infty$ cuando $n \to +\infty$. Entonces*

$$\frac{X_n - \lambda_n}{\sqrt{\lambda_n}}$$

converge en ley a una normal estándar $\mathcal{N}$.

Observación B.3.27. Para una prueba de este resultado, consultar [19], Proposición B.9.1.

Definición B.3.35 (Medida de probabilidad de Haar)**.** *Sea G un grupo compacto. Entonces existe en G una medida de probabilidad de Borel μ_G única, que es invariante bajo traslaciones a la derecha y a la izquierda, es decir, para cualquier función integrable $f\colon G \to \mathbb{C}$ y para cualquier $g \in G$ fijo se tiene*

$$\int_G f(gx)\, d\mu_G(x) = \int_G f(xg)\, d\mu_G(x) = \int_G f(x)\, d\mu_G(x).$$

A esta medida se la conoce como medida de probabilidad de Haar en G.

C. Resultados de la Teoría (Analítica) de Números

C.1. Funciones aritméticas

Empezamos definiendo qué es una función aritmética, y describiendo algunas de sus posibles propiedades.

Definición C.1.1 (Función aritmética)**.** *Una función f compleja definida sobre los enteros positivos se denomina función aritmética.*

Ejemplo C.1.2. Algunas funciones aritméticas importantes son

- La función indicatriz o totient de Euler $\varphi(n) = \left|\{m \leq n | (m,n) = 1\}\right|$
- La función $\omega(n)$ y $\Omega(n)$, que cuentan la cantidad de factores primos de n sin y con multiplicidad, respectivamente.
- Las funciones divisor $\sigma_\alpha(n) = \sum_{d|n} d^\alpha$, definidas para cualquier $\alpha \in \mathbb{C}$.
- La función de Liouville $\lambda(n)$, la cual cumple que $\lambda(1) = 1$ y que si $n = p_1^{\alpha_1} \cdot \ldots \cdot p_m^{\alpha_m}$ entonces
$$\lambda(n) = (-1)^{\alpha_1 + \cdots + \alpha_m}$$
- $\mu(n)$ es la función de Möbius, definida como
$$\mu(n) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } a^2 | n \text{ para algún } a > 1 \\ (-1)^r & \text{si} n \text{ tiene } r \text{ factores primos distintos.} \end{cases}$$

Definición C.1.3 (Función (completamente) multiplicativa)**.** *Sea f una función aritmética. Se dice que f es multiplicativa si f no es idénticamente nula y, para m, n enteros positivos con $(m,n) = 1$ se cumple que*

$$f(mn) = f(m)f(n).$$

Si la anterior igualdad se mantiene para cualesquiera m y n, entonces se dice que f es completamente multiplicativa.

Definición C.1.4 (Función (completamente) aditiva)**.** *Sea f una función aritmética. Se dice que f es aditiva si f no es idénticamente nula y, para m, n enteros positivos con $(m,n) = 1$ se cumple que*

$$f(mn) = f(m) + f(n).$$

Si la anterior igualdad se mantiene para cualesquiera m y n, entonces se dice que f es completamente aditiva.

Teorema C.1.5. *Si $n \geq 1$ entonces*

$$\sum_{d|n} \mu(d) = \left[\frac{1}{n}\right] = \begin{cases} 1 & si \quad n = 1 \\ \\ 0 & si \quad n \neq 1. \end{cases}$$

Demostración. Para $n = 1$, la fórmula es claramente cierta. Supongamos que $n > 1$ y se puede descomponer como $n = p_1^{\alpha_1} \cdot \ldots \cdot p_m^{\alpha_m}$. En la suma $\sum_{d|n} \mu(d)$ los únicos términos no nulos proceden de $d = 1$ y de los divisores de n. Entonces

$$\begin{aligned} \sum_{d|n} \mu(d) &= \mu(1) + \mu(p_1) + \cdots + \mu(p_m) + \mu(p_1p_2) + \cdots + \mu(p_{m-1}p_m) + \cdots + \mu(p_1p_2 \ldots p_m) \\ &= 1 + \binom{m}{1}(-1) + \binom{m}{2}(-1)^2 + \cdots + \binom{m}{m}(-1)^m = (1-1)^m = 0. \end{aligned}$$

□

Teorema C.1.6. *Para $n \geq 1$ se tiene*

$$\varphi(n) = \sum_{d|n} \mu(d)\frac{n}{d}.$$

Demostración. Tenemos que la función totient se puede expresar como

$$\varphi(n) = \sum_{k=1}^{n} \left[\frac{1}{(n,k)}\right]$$

donde k recorre todos los enteros positivos menores o iguales que n. Entonces, por el Teorema C.1.5 sustituyendo n por (n, k), obteniendo

$$\varphi(n) = \sum_{k=1}^{n} \sum_{d|(n,k)} \mu(d) = \sum_{k=1}^{n} \sum_{\substack{d|n \\ d|k}} \mu(d).$$

Para un divisor d de n fijo podemos sumar respecto de todos los k tales que $1 \leq k \leq n$ que son múltiplos de d. Si escribimos $k = qd$, entonces $1 \leq k \leq n$ si y solo si $1 \leq q \leq n/d$. Por lo tanto, la última suma se puede escribir

$$\varphi(n) = \sum_{d|n} \sum_{q=1}^{n/d} \mu(d) = \sum_{d|n} \mu(d) \sum_{q=1}^{n/d} 1 = \sum_{d|n} \mu(d)\frac{n}{d}.$$

□

Proposición C.1.7. *Sea $\varphi(n)$ la función totient de Euler. Se cumple*

1. $\frac{\varphi(n)}{n} = \prod_{p|n}\left(1-\frac{1}{p}\right)$.
2. *$\varphi(n)$ es multiplicativa.*

Demostración. 1. Para $n = 1$ el producto es vacío, ya que no hay primo que divida a 1. Por lo tanto, el valor del producto es 1. Supongamos que $n > 1$ y $p_1, \ldots, p_m$ son los divisores primos de n. Entonces, el producto se puede escribir como

$$\prod_{p|n}\left(1-\frac{1}{p}\right) = \prod_{i=1}^{m}\left(1-\frac{1}{p_i}\right) = 1 - \sum\frac{1}{p_i} + \sum\frac{1}{p_i p_j} - \sum\frac{1}{p_i p_j p_k} + \ldots + \frac{(-1)^m}{p_1 \cdot \ldots \cdot p_m}.$$

Vemos que cada término de la derecha de la igualdad es de la forma $\pm 1/d$, con d divisor de n. El numerador ± 1 es exactamente $\mu(d)$. Dado que $\mu(d) = 0$ si d es divisible por algún cuadrado de algún p_i, entonces la suma de la derecha de la igualdad coincide con $\sum_{d|n}\frac{\mu(d)}{d}$, por lo que por el Teorema C.1.6 queda demostrada la primera parte de la proposición.

2. Sean m, n enteros positivoscon $(m, n) = 1$. Entonces, por la fórmula producto de la función totient

$$\varphi(mn) = \prod_{p|mn}\left(1-\frac{1}{p}\right) = \prod_{p|m}\left(1-\frac{1}{p}\right)\prod_{p|n}\left(1-\frac{1}{p}\right) = \varphi(m)\varphi(n).$$

□

Lema C.1.8. *Para $s \in \mathbb{C}$ tal que $\Re(s) > 1$ se cumple*

$$\sum_{m,n}\frac{\mu(m)\mu(n)}{[m,n]^s} = \sum_{d}\frac{(f \circledast g)(d)}{d^s} = \prod_{p}\left(1-\frac{1}{p^s}\right) = \sum_{n\geq 1}\mu(n)n^{-s} = \frac{1}{\zeta(s)}.$$

C.2. Primos y su distribución

Definición C.2.1 (Función contadora de primos)**.** *Sea $x \geq 0$ un número real. Se define la función $\pi(x)$ como la cantidad de primos menores o iguales que x. Dicho de otro modo*

$$\pi(x) := \big|\{p \text{ primo} \mid p \leq x\}\big|.$$

Definición C.2.2 (Funciones de Chebyshev)**.** *Sea $x \geq 0$ un número real. Se definen las siguientes funciones.*

1. $\psi(x) := \sum_{p^m \leq x}\log p$, *con p primo y $m \geq 1$.*
2. $\vartheta(x) := \sum_{p \leq x}\log p$, *con p primo.*

Teorema C.2.3 (Estimaciones de Chebyshev y Mertens)**.** *Se cumplen los tres enunciados siguientes.*

1. *Existen constantes positivas* c_1 *y* c_2 *tales que*

$$c_1 \frac{x}{\log x} \leq \pi(x) \leq c_2 \frac{x}{\log x}$$

para todo $x \geq 2$.

2. *Para cualquier* $x \geq 3$, *se cumple*

$$\sum_{p \leq x} \frac{1}{p} = \log\log x + O(1).$$

3. *Para cualquier* $x \geq 3$, *se cumple*

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + O(1).$$

Observación C.2.1. Para una prueba de este resultado, consultar [1], Teorema 4.6 para la primera parte, o bien, [14], Teoremas 7, 414, 427 y 425.

Teorema C.2.4. *Para todo* x *se cumple*

$$\psi(x) = \vartheta(x) + O(x^{\frac{1}{2}}(\log x)^2).$$

Observación C.2.2. Para una prueba de este resultado, consultar [14].

Proposición C.2.5. *Sea* $A \geq 0$ *un número real fijo. Para todo* $x \geq 2$, *se tiene que*

$$\prod_{p \leq x} \left(1 + \frac{A}{p}\right) \ll (\log x)^A,$$

$$\prod_{p \leq x} \left(1 - \frac{A}{p}\right)^{-1} \ll (\log x)^A,$$

donde las constantes implícitas sólo dependen de A.

Observación C.2.3. Para una prueba de este resultado, consultar [19], Proposición C.3.6.

Teorema C.2.6 (Teorema de los Números Primos)**.** *Sea* $A > 0$ *un número real arbitrario. Para* $x \geq 2$ *se tiene que*

$$\pi(x) = li(x) + O\left(\frac{x}{(\log x)^A}\right) \tag{C.1}$$

donde $li(x)$ *es la integral logarítmica*

$$li(x) := \int_2^{+\infty} \frac{dt}{\log t}$$

donde la constante implícita depende solo de A. *De forma más general, para* $\alpha \geq 0$ *fijo, se cumple*

$$\sum_{p \leq x} p^\alpha = \int_2^{+\infty} t^\alpha \frac{dt}{\log t} + O\left(\frac{x^{1+\alpha}}{(\log x)^A}\right)$$

donde la constante implícita depende solo de A.

Observación C.2.4. Para una prueba de este resultado, consultar [1].

Teorema C.2.7 (Dirichlet-Hadamard-De la Vallée Poussin)**.** *Para cualquier $q \geq 1$ y $A \geq 1$ fijos, y para cualquier $x \geq 2$ se tiene*

$$\pi(x;q,a) = \frac{1}{\varphi(q)}\frac{x}{\log x} + O\left(\frac{x}{(\log x)^A}\right)$$

$$\sim \frac{1}{\varphi(q)}\pi(x) \sim \frac{1}{\varphi(q)}li(x).$$

C.3. Función $\zeta(s)$ de Riemann. Resultados asociados

Definición C.3.1 (Función $\zeta(s)$ de Riemann)**.** *Sea $s \in \mathbb{C}$ con $\Re(s) > 1$. Se define la función $\zeta(s)$ de Riemann como*

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

Proposición C.3.2 (Fórmula producto)**.** *Con las mismas condiciones de C.3.1, se puede expresar la función anterior como*

$$\zeta(s) = \prod_p (1-p^{-1})^{-1}.$$

Observación C.3.1. Para una prueba de este resultado, consultar [1], aplicando el Teorema 11.7 con $f(n) = 1$.

Proposición C.3.3 (Ecuación funcional)**.** *Para todo $s \in \mathbb{C} \setminus \{1\}$ se cumple*

$$\zeta(1-s) = 2(2\pi)^{-s}\Gamma(s)\cos\left(\frac{\pi s}{2}\right)\zeta(s).$$

Observación C.3.2. Para una prueba de este resultado, consultar [1], Teorema 12.7.

Conjetura C.3.4 (Hipótesis de Riemann)**.** *Sea $S = \{s \in \mathbb{C} \mid 0 \leq \Re(s) \leq 1\}$ la banda cuyos elementos tienen parte real en el intervalo $[0,1]$. Si $s \in S$ y $\zeta(s) = 0$, entonces $\Re(s) = 1/2$.*

Proposición C.3.5. *Sea $T \geq 1$ un número real y sean m, n enteros tales que $1 \leq m, n \leq T$. Sea σ un número real con $\frac{1}{2} \leq \sigma \leq 1$. Entonces se tiene que*

$$\frac{1}{2T}\int_{-T}^{T}\left(\frac{m}{n}\right)^{it}|\zeta(\sigma+it)|^2 dt = \zeta(2\sigma)\left(\frac{(m,n)^2}{mn}\right)^{\sigma}$$
$$+ \frac{1}{2T}\zeta(2-2\sigma)\left(\frac{(m,n)^2}{mn}\right)^{1-\sigma}\int_{-T}^{T}\left(\frac{|t|}{2\pi}\right)^{1-2\sigma} dt + O(\text{mín}(m,n)T^{-\sigma+\varepsilon}).$$

Observación C.3.3. Para una prueba de este resultado, consultar [26].

Figura C.1: Bernhard Riemann, célebre matemático.

Proposición C.3.6 (Factorización de Hadamard)**.** *Sea $s = \sigma + it \in \mathbb{C}$ tal que $\frac{1}{2} \leq \sigma \leq 1$ y $\zeta(s) \neq 0$. Entonces hay $\ll \log(2 + |t|)$ ceros ϱ de ξ tal que $|s - \varrho| \leq 1$, y se cumple*

$$-\frac{\zeta'(s)}{\zeta(s)} = \frac{1}{s} + \frac{1}{s-1} - \sum_{|s-\varrho|<1} \frac{1}{s-\varrho} + O(\log(2 + |t|)),$$

donde la suma recorre los ceros ϱ de $\zeta(s)$ tal que $|s-\varrho| < 1$, contados con multiplicidad, y $\xi(s)$ es la función $\xi(s) := \pi^{-s/2}\Gamma(\frac{s}{2})\zeta(s)$.

Observación C.3.4. Para un esquema de la prueba de este resultado, consultar [19], Proposición C.4.4.

C.4. L-funciones de Dirichlet

Definición C.4.1. *Sea $q \geq 1$ un entero. Las L-funciones de Dirichlet módulo q son las series determinadas por los caracteres del grupo de clases residuo invertibles módulo q, es decir, son series de la forma*

$$L(s, \chi) = \sum_{n \geq 1} \frac{\chi(n)}{n^s}.$$

Un caracter módulo q es una función que cumple $\chi\colon (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}^\times$, que se pueden extender a todo $\mathbb{Z}/q\mathbb{Z}$ enviando las clases no invertibles a 0. La función resultante se llama caracter de Dirichlet módulo q.

Conjetura C.4.2 (Hipótesis de Riemann generalizada)**.** *Para cualquier entero $q \geq 1$, para cualquier caracter de Dirichlet módulo q χ y para cualquier cero $\rho = \beta + i\gamma$ de su L-función tal que $0 < \beta \leq 1$, se tiene que $\beta = \frac{1}{2}$*

Proposición C.4.3. *Sea $q \geq 1$ un entero. Para cualquier x e y en $\mathbb{Z}$, tenemos que*

$$\frac{1}{\varphi(q)} \sum_{\chi (mod\ q)} \chi(x)\overline{\chi(y)} = \begin{cases} 1 & si\ x = y \\ 0 & en\ otro\ caso \end{cases}$$

donde la suma recorre todos los caracteres de Dirichlet módulo q.

Observación C.4.1. Para una prueba de este resultado, consultar [19], Proposición C.5.1.

Corolario C.4.4. *Sea $q \geq 1$ un entero y χ un caracter de Dirichlet módulo q no trivial. Asumimos la Hipótesis de Riemann generalizada módulo q para $L(s, \chi)$. Para cualquier $x \geq 2$, se cumple*

$$\int_2^x \left(\sum_{n \leq t} \Lambda(n)\chi(n) \right) dt \ll x^{3/2},$$

donde la constante implícita depende de q.

Observación C.4.2. Para una prueba de este resultado, consultar [19], Corolario C.5.11.

Teorema C.4.5. *Sea $q \geq 1$ un entero y χ un caracter de Dirichlet módulo q no trivial. Para cualquier $x \geq 2$ y $X \geq 2$ tales que $2 \leq x \leq X$ se cumple*

$$\sum_{n \leq x} \Lambda(n)\chi(n) = - \sum_{\substack{L(\beta+i\gamma)=0 \\ |\gamma| \leq X}} \frac{x^{\beta+i\gamma}}{\beta + i\gamma} + O\left(\frac{x(\log qx)^2}{X} \right),$$

donde la suma es sobre los ceros no triviales de $L(s, \chi)$, contados con multiplicidad, y la constante implícita es absoluta.

Observación C.4.3. Para una prueba de este resultado, consultar [19], Teorema C.5.6.

Proposición C.4.6. *Sea χ un caracter de Dirichlet módulo q.*

1. *Para $T \geq 1$, el número $N(T; \chi)$ de ceros ρ de $L(s; \chi)$ tales que*

$$\Re(\rho) > 0, \quad |\Im\rho| \leq T$$

satisface

$$N(T; \chi) = \frac{T}{\pi} \log\left(\frac{qT}{2\pi} \right) - \frac{T}{\pi} + O(\log(q(T+1))).$$

2. *Para cualquier $\varepsilon > 0$, la serie*

$$\sum_{\rho} \rho^{-1-\varepsilon}$$

converge, donde ρ recorre los ceros de $L(s, \chi)$ tales que $\Re(\rho) > 0$.

Observación C.4.4. Para una prueba de este resultado, consultar [19], Proposición C.5.3, primera parte.

Corolario C.4.7. *Sea χ un caracter de Dirichlet módulo q.*

1. *Tenemos que*

$$\sum_{\substack{0<\gamma<T \\ L(\frac{1}{2}+i\gamma,\chi)=0}} \frac{1}{|\frac{1}{2}+i\gamma|} \gg (\log T)^2$$

para T suficientemente grande.

2. *Tenemos que*

$$\sum_{\substack{\gamma>T \\ L(\frac{1}{2}+i\gamma,\chi)=0}} \frac{1}{|\frac{1}{2}+i\gamma|^2} \ll \frac{\log T}{T},$$

para $T \geq 1$.

D. Códigos de R

```
library(numbers)
vec=c()
for (i in 1:10^5) {vec[i] = (omega(i+1) - log(log(10^5)))/
   sqrt(10^5)}
vec=as.factor(vec)
summary(vec)
x = c(-0.923431141731039, -0.283701630724715,
   0.356027880281609 , 0.995757391287933, 1.63548690229426,
   2.27521641330058)
y=c(9700,33760,38844,15855,1816,25)
y = y/sum(y)
rand = rnorm(500)
den = density(rand)
plot(den, lwd = 2, col = "red",main="densidad");lines(x,y)
```

Extracto de código D.1: Código de la gráfica Th. Erdos-Kac $N = 10^5$

```
library(numbers)
vec=c()
for (i in 1:10^6) {vec[i] = (omega(i+1) - log(log(10^6)))/
   sqrt(10^6)}
vec=as.factor(vec)
summary(vec)
x = c(-1.00330925721309, -0.386188945806068,
   0.230931365600957, 0.848051677007983, 1.46517198841501,
   2.08229229982203, 2.69941261122906)
y=c(78734, 288727, 379720, 208034, 42492, 2285, 8)
y = y/sum(y)
rand = rnorm(500)
den = density(rand)
plot(den, lwd = 2, col = "red",main="densidad");lines(x,y)
```

Extracto de código D.2: Código de la gráfica Th. Erdos-Kac $N = 10^6$

```
library(numbers)
vec=c()
for (i in 1:10^7) {vec[i] = (omega(i+1) - log(log(10^7)))/
   sqrt(10^7)}
vec=as.factor(vec)
summary(vec)
x = c(-1.06754965128691, -0.467783315573305,
   0.131983020140295, 0.731749355853896, 1.3315156915675,
   1.9312820272811, 2.5310483629947, 3.1308146987083)
y=c(665134, 2536839, 3642766, 2389433, 691209, 72902, 1716,
   1)
```

```
y = y/sum(y)
rand = rnorm(500)
den = density(rand)
plot(den, lwd = 2, col = "red",main="densidad");lines(x,y)
```

Extracto de código D.3: Código de la gráfica Th. Erdos-Kac $N = 10^7$

```
library(numbers)
vec=c()
for (i in 1:10^8) {vec[i] = (omega(i+1) - log(log(10^8)))/
   sqrt(10^8)}
vec=as.factor(vec)
summary(vec)
x = c(-1.12102936816748, -0.535168585367229,
   0.0506921974330238, 0.636552980233277, 1.22241376303353,
   1.80827454583378, 2.39413532863404, 2.97999611143429)
y=c(5172545, 20364419, 31099435, 22926536, 8236747, 1290853,
   67304, 556)
y = y/sum(y)
rand = rnorm(500)
den = density(rand)
plot(den, lwd = 2, col = "red",main="densidad");lines(x,y)
```

Extracto de código D.4: Código de la gráfica Th. Erdos-Kac $N = 10^8$

```
library(numbers)
vec=c()
for (i in 1:10^5) {vec[i] = (omega(i+1)}
vec=as.factor(vec)
summary(vec)
x = c(1:6)
y=c(9700,33760,38844,15855,1816,25)
y = y/sum(y)
ran = rpois(500,log(log(10^5)))
ran = as.factor(ran)
summary(ran)
xmues = c(0:7)
ymues = c(43, 103, 134, 118, 57, 31, 10, 3)
ymues = ymues/sum(ymues)
plot(xmues,ymues, lwd = 2, col = "red", xlim = c(min(x),max(x
   )), ylim = c(0,max(y)), type = "l");lines(x,y)
```

Extracto de código D.5: Código de la gráfica $\omega(n)$ sin renormalización $N = 10^5$

```
library(numbers)
vec=c()
for (i in 1:10^6) {vec[i] = (omega(i+1)}
vec=as.factor(vec)
summary(vec)
x = c(1:7)
```

```
y=c(78734, 288727, 379720, 208034, 42492, 2285, 8)
y = y/sum(y)
ran = rpois(500,log(log(10^6)))
ran=as.factor(ran)
summary(ran)
xmues = c(0:8)
ymues=c(43, 111, 134, 98, 62, 29, 13, 8 ,2)
ymues=ymues/sum(ymues)
plot(xmues,ymues, lwd = 2, col = "red", xlim = c(min(x),max(x
   )), ylim = c(0,max(y)), type = "l");lines(x,y)
```

Extracto de código D.6: Código de la gráfica $\omega(n)$ sin renormalización $N = 10^6$

```
library(numbers)
vec=c()
for (i in 1:10^7) {vec[i] = (omega(i+1)}
vec=as.factor(vec)
summary(vec)
x = c(1:8)
y=c(665134, 2536839, 3642766, 2389433,  691209,   72902,
   1716,  1)
y = y/sum(y)
ran = rpois(500,log(log(10^7)))
ran=as.factor(ran)
summary(ran)
xmues=c(0:9)
ymues=c(28, 90, 117, 113, 75, 51, 21, 3, 1, 1)
ymues=ymues/sum(ymues)
plot(xmues,ymues, lwd = 2, col = "red", xlim = c(min(x),max(x
   )), ylim = c(0,max(y)), type = "l");lines(x,y)
```

Extracto de código D.7: Código de la gráfica $\omega(n)$ sin renormalización $N = 10^7$

```
library(numbers)
vecphi=c()
for (i in 1:10^5) {vecphi[i] = eulersPhi(i+1)/(i+1)}
ECDF=ecdf(vecphi)
plot(ECDF,col="black",lwd=1)
```

Extracto de código D.8: Código de la gráfica de distribución empírica de $\varphi(n)/n$, $N = 10^5$

```
library(numbers)
vecphi=c()
for (i in 1:10^6) {vecphi[i] = eulersPhi(i+1)/(i+1)}
ECDF=ecdf(vecphi)
plot(ECDF,col="black",lwd=1)
```

Extracto de código D.9: Código de la gráfica de distribución empírica de $\varphi(n)/n$, $N = 10^6$

```
library(numbers)
vecphi=c()
for (i in 1:10^7) {vecphi[i] = eulersPhi(i+1)/(i+1)}
ECDF=ecdf(vecphi)
plot(ECDF,col="black",lwd=1)
```

Extracto de código D.10: Código de la gráfica de distribución empírica de $\varphi(n)/n$, $N = 10^7$

```
library(ggplot2)
x = 1:16
y = c(1,9,6,13,7,3,5,15,2,12,14,10,4,11,8,16)
sumvec = c()
for (j in 1:16){
  sumvec[j] = (1/sqrt(17))*exp(2*1i*pi*(x[j] + y[j])/17)
}
vec = cumsum(sumvec)
datos = as.data.frame(cbind(Re(vec),Im(vec)))
names(datos) = c("Parte_Real", "Parte_Imaginaria")
ggplot(datos,aes(Parte_Real, Parte_Imaginaria)) + geom_point(
   color='darkgreen') +
 xlab("Parte␣Real") + ylab("Parte␣Imaginaria")
```

Extracto de código D.11: Código de las sumas parciales de Kl(1,1;17)

```
library(ggplot2)
x = 1:16
y = c(1,9,6,13,7,3,5,15,2,12,14,10,4,11,8,16)
sumvec = c()
for (j in 1:16){
  sumvec[j] = (1/sqrt(17))*exp(2*1i*pi*(x[j] + y[j])/17)
}
vec = cumsum(sumvec)
datos = as.data.frame(cbind(Re(vec),Im(vec)))
names(datos) = c("Parte␣Real", "Parte␣Imaginaria")
ggplot(datos,aes(Parte_Real, Parte_Imaginaria)) + geom_point(
   color='darkgreen') + geom_path(color='green') +
  xlab("Parte␣Real") + ylab("Parte␣Imaginaria")
```

Extracto de código D.12: Código del camino generado por las sumas parciales de Kl(1,1;17)

```
library(ggplot2)
x = 1:16
y = c(1,9,6,13,7,3,5,15,2,12,14,10,4,11,8,16)
sumvec = c()
for (j in 1:16){
  sumvec[j] = (1/sqrt(17))*exp(2*1i*pi*(2*x[j] + y[j])/17)
```

```
}
vec = cumsum(sumvec)
datos = as.data.frame(cbind(Re(vec),Im(vec)))
names(datos) = c("Parte␣Real", "Parte␣Imaginaria")
ggplot(datos,aes(Parte_Real, Parte_Imaginaria)) + geom_point(
   color='darkgreen') + geom_path(color='green') +
  xlab("Parte␣Real") + ylab("Parte␣Imaginaria")
```

Extracto de código D.13: Código del camino generado por las sumas parciales de Kl(2,1;17)

```
library(ggplot2)
x = 1:16
y = c(1,9,6,13,7,3,5,15,2,12,14,10,4,11,8,16)
sumvec = c()
for (j in 1:16){
  sumvec[j] = (1/sqrt(17))*exp(2*1i*pi*(3*x[j] + y[j])/17)
}
vec = cumsum(sumvec)
datos = as.data.frame(cbind(Re(vec),Im(vec)))
names(datos) = c("Parte␣Real", "Parte␣Imaginaria")
ggplot(datos,aes(Parte_Real, Parte_Imaginaria)) + geom_point(
   color='darkgreen') + geom_path(color='green') +
  xlab("Parte␣Real") + ylab("Parte␣Imaginaria")
```

Extracto de código D.14: Código del camino generado por las sumas parciales de Kl(3,1;17)

```
library(ggplot2)
x = 1:16
y = c(1,9,6,13,7,3,5,15,2,12,14,10,4,11,8,16)
sumvec = c()
for (j in 1:16){
  sumvec[j] = (1/sqrt(17))*exp(2*1i*pi*(4*x[j] + y[j])/17)
}
vec = cumsum(sumvec)
datos = as.data.frame(cbind(Re(vec),Im(vec)))
names(datos) = c("Parte␣Real", "Parte␣Imaginaria")
ggplot(datos,aes(Parte_Real, Parte_Imaginaria)) + geom_point(
   color='darkgreen') + geom_path(color='green') +
  xlab("Parte␣Real") + ylab("Parte␣Imaginaria")
```

Extracto de código D.15: Código del camino generado por las sumas parciales de Kl(4,1;17)

```
library(ggplot2)
x = 1:16
```

```
y = c(1,9,6,13,7,3,5,15,2,12,14,10,4,11,8,16)
sumvec = c()
for (j in 1:16){
  sumvec[j] = (1/sqrt(17))*exp(2*1i*pi*(5*x[j] + y[j])/17)
}
vec = cumsum(sumvec)
datos = as.data.frame(cbind(Re(vec),Im(vec)))
names(datos) = c("Parte␣Real", "Parte␣Imaginaria")
ggplot(datos,aes(Parte_Real, Parte_Imaginaria)) + geom_point(
   color='darkgreen') + geom_path(color='green') +
  xlab("Parte␣Real") + ylab("Parte␣Imaginaria")
```

Extracto de código D.16: Código del camino generado por las sumas parciales de Kl(5,1;17)

E. Códigos de Mathematica

```
ComplexPlot[Zeta[s + I 0], {s, -2 - 2 I, 3 + 3 I},
 PlotLegends -> Automatic]
```

Extracto de código E.1: Código de la imagen módulo zeta Riemann 1

```
ComplexPlot[Zeta[s + I  1000], {s, -2 - 2 I, 3 + 3 I},
 PlotLegends -> Automatic]
```

Extracto de código E.2: Código de la imagen módulo zeta Riemann 2

```
ComplexPlot[Zeta[s + I  5000], {s, -2 - 2 I, 3 + 3 I},
 PlotLegends -> Automatic]
```

Extracto de código E.3: Código de la imagen módulo zeta Riemann 3

```
ComplexPlot[Zeta[s + I  10000], {s, -2 - 2 I, 3 + 3 I},
   PlotLegends -> Automatic]
```

Extracto de código E.4: Código de la imagen módulo zeta Riemann 4

```
Plot[{Re[Zeta[1/2 + I t]], Im[Zeta[1/2 + I t]]}, {t, 0, 100},
    PlotLegends -> "Expressions"]
```

Extracto de código E.5: Código de la imagen zeta Riemann en la línea crítica

Bibliografía

1. Apostol, T. *Introducción a la Teoría Analítica de los Números* (Editorial Reverté, 2018).
2. Billingsley, P. *Probability and measure, Third Edition* (Wiley, 1995).
3. Edwards, H. *Riemann's Zeta Function* (Dover Publications, 2001).
4. Ehsanes Saleh, A. y Rohatgi, V. *An Introduction to Probability and Statistics* (Willey-Interscience, 2015).
5. Einsiedler, M. y Ward, T. *Ergodic theory: with a view towards number theory* (Springer, 2011).
6. Elkies, N. y McMullen, C. Gaps in $\sqrt{n}$ *mod* 1 and ergodic theory. *Duke Math. J.* **123,** 95-139 (2004).
7. Erdös, P. y Kac, M. The Gaussian Law of Errors in the Theory of Additive Number Theoretic Functions. *American Journal of Mathematics* **62,** 738-742 (1940).
8. Feller, W. *An introduction to probability theory and its applications, Vol. 2* (Wiley, 1966).
9. Fraleigh, J. *A First Course in Abstract Algebra, Sixth Edition* (Pearson Education, 2000).
10. Gallagher, P. On the distribution of primes in short intervals. *Mathematika 23,* 4-9 (1976).
11. Ghys, É. en *Séminaire Bourbaki 1991/92, exposés 747* (1992).
12. Granville, A. y Soundararajan, K. Sieving and the Erdös-Kac Theorem. *Equidistribution in number theory, an introduction,* 15-27 (2007).
13. Hardy, G. H. y Ramanujan, S. The normal number of prime factors of a number *n*. *Quart. J.* **48,** 76-92 (1917).
14. Hardy, G. y Wright, E. *An Introduction to the Theory of Numbers, Sixth Edition* (Oxford University Press, 2008).
15. Harper, A. Two new proofs of the Erdös–Kac Theorem, with bound on the rate of convergence, by Stein's method for distributional approximations. *Math. Proc. Camb. Phil. Soc.* **147,** 95-114 (2009).
16. Helfgott, H. Azar y Aritmética. *arXiv.* https://arxiv.org/abs/0909.0922 (2009).
17. Holmstedt, T. Interpolation of quasi-normed spaces. *Math. Scand.* **26,** 177-199 (1970).
18. Katz, N. M. *Gauss Sums, Kloosterman Sums and Monodromy Groups* (Princeton Univ. Press, 1988).
19. Kowalski, E. *An Introduction to Probabilistic Number Theory* (Cambridge University Press, 2021).

20. Lubotzky, A. *Discrete groups, expanding graphs and invariant measures. With an appendix by Jonathan D. Rogawski* (Birkhuser Verlag, 1994).

21. Lubotzky, A. Ramanujan Graphs. *arXiv* (2017).

22. Lubotzky, A., Phillips, R. y Sarnak, P. Ramanujan graphs. *Combinatorica* **8(3),** 261-277 (1988).

23. Montgomery-Smith, S. The Distribution of Rademacher Sums. *Proceedings of the American Mathematical Society* **109,** 517-522 (1990).

24. Morris, D. *Ratner's theorems on unipotent flows* (University of Chicago Press, 2005).

25. Pintz, J. Cramér vs. Cramér. On Cramér's probabilistic model for primes. *Functiones et Approximatio Commentarii Mathematici* **37,** 361-376 (2007).

26. Radziwiłł, M. y Soundararajan, K. Selberg's central limit theorem for $\log|\zeta(\frac{1}{2}+it)|$. *Enseign. Math.* **63,** 1-19 (2017).

27. Ratner, M. On Raghunathan's measure conjecture. *Ann. of Math* **134,** 545-607 (1991).

28. Rubinstein, M. O. y Sarnak, P. Chebyshev's Bias. *Exp. Math.* **3,** 173-197 (1994).

29. Sós, V. On the theory of diophantine approximations. *Acta Math. Acad.Sci. Hungar.* **8,** 461-472 (1957).

30. Weyl, H. Uber die Gleichverteilung von Zahlen mod. Eins. *Math. Ann.* **77** (1914).

31. Zygmund, A. *Trigonometric sums, Third Edition* (Cambridge, 2002).