The Distribution of Prime Numbers and Sums of Singular Series

Vivian Kuperberg

Stanford University

May 22nd, 2020

Patterns in the primes

Motivating Question

Let $\mathcal{H} = \{h_1, h_2, \ldots, h_k\} \subseteq \mathbb{N}$. How often are $n + h_i$ all prime?

$$\sum_{n \leqslant N} \prod_{i=1}^{k} \mathbb{1}_{\mathcal{P}}(n + h_i) = ?$$

$$\sum_{n \leqslant N} \prod_{i=1}^{k} \Lambda(n + h_i) = ?$$

Here we have $\Lambda(n)$ the von Mangoldt function,

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^j \\ 0 & \text{else} \end{cases}$$

First step: the Prime Number Theorem

As a starting point, we have the prime number theorem.

Prime Number Theorem (Hadamard, de la Vallée Poussin)

$$\pi(N) := \sum_{n \leqslant N} \mathbb{1}_{\mathcal{P}}(n) \sim \mathrm{li}(N) \sim \frac{N}{\log N}$$

$$\psi(N) := \sum_{n \leqslant N} \Lambda(n) \sim N$$

Cramér's model

Consider a sequence $Y(n)$, $n \geqslant 3$, of independent random variables with

$$\begin{cases} \mathrm{Prob}(Y(n) = 1) = \frac{1}{\log n} & \text{(``$n$ is prime'')} \\ \mathrm{Prob}(Y(n) = 0) = 1 - \frac{1}{\log n} & \text{(``n is composite'')} \end{cases}$$

The Cramér model expects the sequence of primes to behave like a typical sequence in this random model.

Cramér guess

$$\sum_{n \leqslant N} \prod_{i=1}^{k} \mathbb{1}_{\mathcal{P}}(n + h_i) \sim \frac{N}{(\log N)^k}$$

Patterns relative to q

Instead of asking that $n + h_1, \ldots, n + h_k$ be primes, we can pick a squarefree integer q, and ask that $(n + h_i, q) = 1$ for each i.

"Cramér" guess

Each event $(n + h_i, q) = 1$ is independent and occurring with probability $\frac{\phi(q)}{q}$, so

$$\sum_{n \leqslant q} \prod_{i=1}^{k} \mathbb{1}_{\mathcal{P}(q)}(n + h_i) \sim q \left(\frac{\phi(q)}{q} \right)^k$$

Defining the singular series

How can we do better? For every $p|q$, we need n to avoid the classes $-h_k, \ldots, -h_1 \pmod{p}$.

$$v_{\mathcal{H}}(p) = |\mathcal{H} \pmod{p}| = \#\text{ distinct classes} \pmod{p} \text{ among } \mathcal{H}$$

The number of successful configurations $\pmod{q}$ is

$$\begin{aligned}\prod_{p|q}(p - v_{\mathcal{H}}(p)) &= q\prod_{p|q}(1 - v_{\mathcal{H}}(p)/p) \\ &= \underbrace{q\left(\frac{\phi(q)}{q}\right)^k}_{\text{Cramér guess}} \underbrace{\prod_{p|q}\frac{1 - v_{\mathcal{H}}(p)/p}{(1-1/p)^k}}_{\mathfrak{S}(\mathcal{H};q)}\end{aligned}$$

The Hardy–Littlewood Conjecture

If $q = q_\ell = \prod_{p<\ell} p$ with ℓ large, then "relatively prime to q" is very close to "prime." As $\ell \to \infty$,

$$\mathfrak{S}(\mathcal{H}; q_\ell) \to \prod_p \frac{1 - v_{\mathcal{H}}(p)/p}{(1-1/p)^k} := \mathfrak{S}(\mathcal{H})$$

For large p, $v_{\mathcal{H}}(p) = k$ and $(1-k/p)(1-1/p)^{-k} = 1 + O(p^{-2})$, so the product defining $\mathfrak{S}(\mathcal{H})$ converges.

Hardy–Littlewood Conjecture

$$\sum_{n \leqslant N} \prod_{i=1}^{k} \mathbb{1}_{\mathcal{P}}(n+h_i) \sim \mathfrak{S}(\mathcal{H}) \frac{N}{(\log N)^k}$$

$$\sum_{n \leqslant N} \prod_{i=1}^{k} \Lambda(n+h_i) \sim \mathfrak{S}(\mathcal{H}) N$$

Primes in intervals

Motivating Question

How many primes are there in a small interval?

...in very small intervals

For $\lambda > 0$ and $h = \lambda \log n$,

$$\lim_{N\to\infty} \frac{1}{N} \#\{n \leqslant N : \pi(n+h) - \pi(n) = k\} =?$$

...in medium intervals

Let h be such that h/n is small but $h/\log n$ is large; for example we can take $h = n^{\delta}$, for $\varepsilon \leqslant \delta \leqslant 1 - \varepsilon$. What is the distribution of $\pi(n+h) - \pi(n)$, or of $\psi(n+h) - \psi(n)$?

Cramér predictions in very small intervals

The Cramér model predicts a Poisson distribution in very small intervals. Let

$$\begin{aligned} P_k &= \frac{1}{N}\#\{n \leqslant N : \pi(n+\lambda\log n) - \pi(n) = k\} \\ &\text{“} = \mathrm{Prob}(\pi(n+\lambda\log n) - \pi(n) = k)\text{”} \end{aligned}$$

Cramér guess

$$P_k \sim \binom{\lambda\log n}{k}\frac{1}{(\log n)^k}\left(1-\frac{1}{\log n}\right)^{\lambda\log n - k} \longrightarrow \frac{\lambda^k}{k!}e^{-\lambda}$$

Very small intervals via moments

The rth moment of $\pi(n+h) - \pi(n)$ with $h = \lambda \log n$ is

$$\frac{1}{N}\sum_{n \leqslant N}(\pi(n+h) - \pi(n))^r = \frac{1}{N}\sum_{n \leqslant N}\left(\sum_{\ell=1}^{h} \mathbb{1}_{\mathcal{P}}(n+\ell)\right)^r$$
$$= \sum_{k=1}^{r} \sigma(r,k) \sum_{h_1 < \cdots < h_k \leqslant h} \left(\frac{1}{N}\sum_{n \leqslant N}\prod_{i=1}^{k} \mathbb{1}_{\mathcal{P}}(n+h_i)\right)$$

with $\sigma(r,k)$ the number of ways for k distinct values to occur among the r values of ℓs.

Our original question of counting patterns in the primes has appeared!

Gallagher's theorem

Cramér approach to the inner sum:

$$\sum_{h_1<\cdots<h_k\leqslant h} \frac{1}{N} \sum_{n\leqslant N} \prod_{i=1}^{k} \mathbb{1}_{\mathcal{P}}(n+h_i) \sim \frac{1}{(\log N)^k} \sum_{h_1<\cdots<h_k\leqslant h} 1$$

Hardy–Littlewood approach to the inner sum:

$$\sum_{h_1<\cdots<h_k\leqslant h} \frac{1}{N} \sum_{n\leqslant N} \prod_{i=1}^{k} \mathbb{1}_{\mathcal{P}}(n+h_i) \sim \frac{1}{(\log N)^k} \sum_{h_1<\cdots<h_k\leqslant h} \mathfrak{S}(\{h_1,\ldots,h_k\})$$

Theorem (Gallagher)

$$\sum_{h_1<\cdots<h_k\leqslant h} \mathfrak{S}(\{h_1,\ldots,h_k\}) \sim \sum_{h_1<\cdots<h_k\leqslant h} 1$$

Cramér predictions in medium intervals

We're now considering $h \sim N^{\delta}$, rather than $h = \lambda \log N$, where we saw Poisson behavior with parameter λ. As λ gets quite large, a Poisson distribution with parameter λ approaches a normal distribution with mean λ and variance λ.

Cramér guess

If $h \sim N^{\delta}$, for $n \leqslant N$, $\pi(n+h) - \pi(n)$ has an approximately normal distribution with mean $\sim h/\log N$ and variance $\sim h/\log N$.
Similarly $\psi(n+h) - \psi(n)$ has approximately normal distribution with mean $\sim h$ and variance $\sim h \log N$.

Hardy–Littlewood disagrees with Cramér

The Hardy–Littlewood conjectures tell us that the variance is smaller!

$$\begin{aligned}
&\frac{1}{N}\sum_{n\leqslant N}(\psi(n+h)-\psi(n)-h)^2\\
&\quad=\frac{1}{N}\sum_{n\leqslant N}\left(\sum_{\ell\leqslant h}\Lambda(n+\ell)\right)^2-2\frac{h}{N}\sum_{n\leqslant N}\sum_{\ell\leqslant h}\Lambda(n+\ell)+h^2\\
&\quad\sim\underbrace{\frac{1}{N}\sum_{n\leqslant N}\sum_{\ell\leqslant h}\Lambda(n+\ell)^2}_{\sim h(\log N-1)}+\underbrace{2\sum_{\ell\leqslant h}\mathfrak{S}(\{0,\ell\})}_{\sim h^2-h\log h+Bh}-h^2\\
&\quad\sim h\left(\log\frac{N}{h}+B-1\right)
\end{aligned}$$

The Cramér guess, $h\log N$, is bigger!

Higher moments

Montgomery and Soundararajan computed all moments $\frac{1}{N}\sum_{n\leqslant N}(\psi(n+h)-\psi(n)-h)^r$.
Their key result again concerns sums of singular series:

Theorem (Montgomery–Soundararajan)

$$\sum_{h_1<\cdots<h_k\leqslant h}\mathfrak{S}_0(\mathcal{H})=\begin{cases}(1+o(1))\frac{k!}{2^{k/2}(k/2)!}(-h\log h+Bh)^{k/2} & \text{if } 2|k\\ o((h\log h)^{k/2}) & \text{if } 2\nmid k\end{cases}$$

These are the moments of a normal distribution!

Higher moments and exponential sums

Montgomery and Soundararajan use a different expression for $\mathfrak{S}(\mathcal{H})$:

$$\mathfrak{S}(\mathcal{H}) = \sum_{\substack{q_1,\ldots,q_k \\ 1\leqslant q_i<\infty}} \left(\prod_{i=1}^{k} \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1,\ldots,a_k \\ 1\leqslant a_i\leqslant q_i \\ (a_i,q_i)=1 \\ \sum a_i/q_i\in\mathbb{Z}}} e\left(\sum_{i=1}^{k} \frac{a_i h_i}{q_i} \right)$$

The blue condition implies that any prime factor of a q_i must divide at least two distinct q_i's. If k is even, one possibility is $q_{2j-1} = q_{2j}$ in pairs, and that the q_i's are otherwise distinct.
They show that terms with $q_{2j-1} = q_{2j}$ contribute the main term, and all other terms contribute to smaller order.
They do this by connecting the sum over singular series to moments of the number of reduced classes mod p, which was previously studied by Montgomery and Vaughan.

Further directions

Question

Montgomery and Soundararajan showed that sums of the singular series $\mathfrak{S}_0$ exhibit square-root cancellation. When k is odd, can this be refined to $O(h^{(k-1)/2}(\log h)^{(k+1)/2})$?

Question

How does this translate to considering regions of number fields?

(K., Rodgers, Roditty-Gershon): Smaller variance occurs in number fields as well, including for regions of different shapes.

Another view of variance: Riemann's explicit formula

Riemann's explicit formula

$$\psi(x) = x - \sum_{\rho} \frac{x^{\rho}}{\rho} + \text{ negligible terms,}$$

with ρ ranging over non-trivial zeroes of the Riemann zeta function.

Then

$$\psi(x+h) - \psi(x) - h = -\sum_{\rho} \frac{(x+h)^{\rho} - x^{\rho}}{\rho} + \text{negligible terms}$$

We assume the Riemann hypothesis. Large zeroes contribute less; we'll consider ρ with $|\rho| < x/h$. For this ρ, $(x+h)^{\rho} - x^{\rho} \sim \rho h x^{\rho - 1}$.

Another view of variance: Riemann's explicit formula

With $\rho = \frac{1}{2} + i\gamma$,

$$
\begin{aligned}
\frac{1}{X}\int_X^{2X} (\psi(x+h) - \psi(x) - h)^2 \mathrm{d}x \\
&\approx \frac{1}{X}\int_X^{2X} \left| -\sum_{\rho} \frac{(x+h)^{\rho} - x^{\rho}}{\rho} \right|^2 \mathrm{d}x \\
&\approx \frac{h^2}{X^2}\int_X^{2X} \left| \sum_{|\gamma| \leqslant X/h} x^{i\gamma} \right|^2 \mathrm{d}x \\
&= \frac{h^2}{X} \sum_{|\gamma_1|, |\gamma_2| \leqslant X/h} X^{i(\gamma_1 - \gamma_2)} \frac{2^{1+i(\gamma_1-\gamma_2)} - 1}{1 + i(\gamma_1 - \gamma_2)}
\end{aligned}
$$

To get an asymptotic, we need to understand the spacings $\gamma_1 - \gamma_2$ of the zeros of ζ along the critical line.

Pair Correlation

Montgomery's Strong Pair Correlation Conjecture

Let

$$F(X, T) := \sum_{0 \leqslant \gamma_1, \gamma_2 \leqslant T} X^{i(\gamma_1 - \gamma_2)} w(\gamma_1 - \gamma_2)$$

with $w(u) := \frac{4}{4+u^2}$. For $T \leqslant X$,

$$F(X, T) = \frac{T}{2\pi} \log T + o(T \log T)$$

This spacing looks like what would happen if the zeroes of ζ behaved like eigenvalues of a random matrix from the Gaussian Unitary Ensemble (GUE), which is the space of Hermitian matrices along with a distribution that is Gaussian with respect to the trace.

Rudnick and Sarnak extended this conjecture to correlations of more than two zeroes. Odlyzko has done large-scale computations of zeroes that support this conjecture.

Pair Correlation and Variance

Theorem (Goldston–Montgomery)

The Strong Pair Correlation Conjecture is equivalent to the following claim. As $X \to \infty$ and for $X^{\varepsilon} \leqslant h \leqslant X^{1-\varepsilon}$,

$$\frac{1}{X}\int_1^X (\psi(x+h) - \psi(x) - h)^2 \mathrm{d}x \sim h \log \frac{X}{h}$$

T. H. Chan extended this to show that with a more precise version of the Pair Correlation Conjecture, we can also include the linear term $(B-1)h$. Higher moments of the distribution of primes have a similar analog in random matrix theory. Rains showed that the eigenvalues of a large power of a random matrix become uniformly distributed. The distribution of the trace then approaches a normal distribution as the power of the random matrix increases.

Thank you!