Divisors

Dimitris Koukoulopoulos

Université de Montréal

CNTA XVI
11 June 2024

The multiplicative structure of random integers

Question

Choose $n \in [1, x]$ uniformly at random. What is the distribution of its divisors?

Problem

The events $\{d_1|n\}$ and $\{d_2|n\}$ could have strong dependencies due to common prime factors of d_1 and d_2.

Examples:

- If $4|n$, then we automatically have $2|n$
- If we know that $2|n$, then the probability that $6|n$ is $1/3$ and not $1/6$ (but the probability that $5|n$ remains $1/5$).

Easier question

What is the distribution of the set of prime factors $\{p|n\}$ of a randomly chosen n?

Warm-up: scale calibration

Prime factors

$$\mathbb{E}_{n\leqslant x}\left[\sum_{p|n,\ p\in[y,z]} 1\right] = \sum_{p\in[y,z]} \mathbb{P}_{n\leqslant x}(p|n) \sim \sum_{p\in[y,z]} \frac{1}{p} \sim \log\log z - \log\log y$$

Divisors

$$\mathbb{E}_{n\leqslant x}\left[\sum_{d|n,\ d\in[y,z]} 1\right] \sim \sum_{d\in[y,z]} \frac{1}{d} \sim \log z - \log y$$

Early days of probabilistic number theory

Theorem (Hardy-Ramanujan (1917))

Most integers $n \leqslant x$ have about $\log\log x$ prime factors

Theorem (Erdős–Kac (1940))

If $\omega(n) = \#\{p|n\}$ and we fix $a < b$, then

$$\mathbb{P}_{n\leqslant x}\left(\frac{\omega(n) - \log\log x}{\sqrt{\log\log x}} \in [a, b]\right) \sim \frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}\mathrm{d}t.$$

- For $n \leqslant x$, we have $\omega(n) = \sum_{p\leqslant x} \mathbb{1}_{p|n}$
- Kubilius: model the RVs $(\mathbb{1}_{p|n})_{p\leqslant x}$ by independent Bernoulli's $(B_p)_{p\leqslant x}$ with $\mathbb{P}(B_p = 1) = \frac{1}{p}$.

The distribution of intermediate prime factors

Prime factors form a Poisson Process

Let $I_1, \ldots, I_k$ be disjoint subintervals of $[1, x]$. Then

$$\mathbb{P}_{n \leqslant x}\Big(\#\{p|n,\ p \in I_j\} = m_j\ \forall j\Big) \approx \prod_{j=1}^{k} e^{-\lambda_j} \frac{\lambda_j^{m_j}}{m_j!}$$

with $\lambda_j = \sum_{p \in I_j} \frac{1}{p} \sim \log\log b_j - \log\log a_j \quad$ if $I_j = [a_j, b_j]$.

Prime factors form a Brownian motion when normalized (Billingsley)

$$\rho_N : [0,1] \to \mathbb{R}, \qquad \rho_N(t) := \frac{\#\big\{p|n : \log\log p \leqslant t \log\log x\big\} - t\log\log x}{\sqrt{t \log\log x}}$$

Then ρ_N converges in distribution to the *Brownian motion* in $[0,1]$.

Just one piece of a puzzle. . .

- Every integer n can be written uniquely as a product of primes. We then have

$$\mathbb{E}_{n\leqslant x}\left[\sum_{p|n,\ p\in[y,z]} 1\right] = \sum_{p\in[y,z]} \mathbb{P}_{n\leqslant x}(p|n) \sim \sum_{p\in[y,z]} \frac{1}{p} \sim \log\log z - \log\log y$$

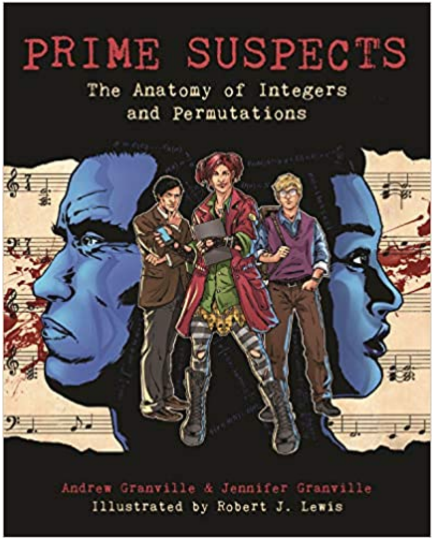
- Every monic polynomial $A \in \mathbb{F}_q[T]$ can be written uniquely as a product of monic irreducibles. We then have

$$\mathbb{E}_{\deg(A)=m}\left[\sum_{P|A,\ \deg(P)\in[k,\ell]} 1\right] = \sum_{P:\ \deg(P)\in[k,\ell]} \frac{1}{q^{\deg(P)}} \sim \log k - \log \ell$$

- Every permutation $\sigma \in S_N$ can be written uniquely as a product of disjoint cycles. We then have

$$\mathbb{E}_{\sigma\in S_N}\left[\sum_{\rho|\sigma,\ \text{length}(\rho)\in[k,\ell]} 1\right] \sim \log k - \log \ell$$

Could these seemingly unrelated anatomies be connected?



Meta-applications to probabilistic Galois theory

Theorem (Bary-Soroker, K., Kozma (2023))

Fix $H \geqslant 35$. If we select f uniformly at random among all monic polynomials of degree n with coefficients in $\{1, 2, \ldots, H\}$, then $\mathrm{Gal}(f) \in \{A_n, S_n\}$ with probability ~ 1 as $n \to \infty$.

In addition, if we select f uniformly at random among all polynomials of degree n with $0, 1$-coefficients, then $\mathrm{Gal}(f) \in \{A_n, S_n\}$ with probability bounded away from 0.

- Breuillard–Varjú (2019) can take $H = 2$ under GRH.
- In 2020, Bary-Soroker and Kozma proved the first statement for any H with at least four distinct prime factors

Rough strategy: reduce f modulo $2, 3, 5, 7$. The reductions should behave approximately like four independent polynomials A_2, A_3, A_5, A_7, with A_p uniformly distributed over monic polynomials of degree n over $\mathbb{F}_p$.

Back to integers: the distribution of large prime factors

The Poisson–Dirichlet distribution

Consider $U_1, U_2, \ldots$ uniform in $[0,1]$ and independent. Then take

$$L_1 = U_1, \quad L_2 = (1-U_1)U_2, \ldots, \quad L_j = (1-U_1)\cdots(1-U_{j-1})U_j, \ldots$$

Let $V_1, V_2, \ldots$ is the sequence $L_1, L_2, \ldots$ ordered decreasingly. Then $\mathbf{V} = (V_1, V_2, \ldots)$ has the Poisson–Dirichlet distribution (of parameter 1).

Large prime factors follow the Poisson–Dirichlet distribution (Billingsley)

Let $n = P_1(n)P_2(n)\cdots$, where $P_1(n) \geqslant P_2(n) \geqslant \cdots$ are primes or ones. Then

$$\mathbb{P}_{n\leqslant x}\Big(P_1(n) \leqslant x^{u_1}, \ldots, P_k(n) \leqslant x^{u_k}\Big) \sim \mathbb{P}\big(V_1 \leqslant u_1, \ldots, V_k \leqslant u_k\big).$$

Example: $\#\{n \leqslant x : P_1(n) \leqslant x^u\} \sim x \cdot \rho(1/u)$, where ρ is the *Dickman–de Bruijn function*.

Arratia's coupling

Theorem (Haddad-K. (2024))

Let $N_x \sim \text{Uniform}(\mathbb{Z} \cap [1, x])$ and let $\mathbf{V} = (V_1, V_2, \ldots)$ follow the Poisson–Dirichlet distr.

There exists a coupling of $\mathbf{V}$ and N_x such that

$$\mathbb{E} \sum_{i \geqslant 1} \left| \frac{\log P_i}{\log x} - V_i \right| = O\left(\frac{1}{\log x} \right),$$

where $N_x = P_1 P_2 \cdots$ with $P_1 \geqslant P_2 \geqslant \cdots$ all primes or ones.

- Arratia proved this in 1998 with $O(\frac{\log \log x}{\log x})$ and conjectured the above (which is optimal).

The DDT theorem

The DDT theorem (Deshouillers–Dress–Tenenbaum (1979))

$$\frac{1}{x}\sum_{n\leqslant x}\frac{\#\{d|n : d\leqslant n^u\}}{\tau(n)} = \frac{2}{\pi}\arcsin\sqrt{u} + O\Big(\frac{1}{\sqrt{\log x}}\Big).$$

A more probabilistic formulation

Recall $N_x \sim \mathrm{Uniform}(\mathbb{Z}\cap[1,x])$. Fix parameters $\alpha_j\in(0,1)$ with $\alpha_1+\cdots+\alpha_k=1$.
Define the random k-factorization $\mathbf{D}_x=(D_{x,1},\ldots,D_{x,k})$ such that

$$\mathbb{P}\Big[D_{x,j}=d_j\ \forall j\ \Big|\ N_x=n\Big] = \prod_{1\leqslant j\leqslant k}\tau_{\alpha_j}(d_j)\quad \text{whenever } d_1\cdots d_k=n.$$

- DDT: $k=2$ and $\alpha_1=\alpha_2=1/2$.
- Sun-Kai Leung (2023): $\mathbb{P}\big[D_{x,j}\leqslant N_x^{u_j}\ \forall j\leqslant k-1\big] = \mathrm{Dirichlet}(\boldsymbol{\alpha};\mathbf{u}) + O\big((\log x)^{-\frac{1}{k}}\big)$.

The Dirichlet law via Arratia's coupling

Theorem (Donnelly–Tavaré (1987))

- Let $\mathbf{V} = (V_1, V_2, \dots)$ be a Poisson-Dirichlet distribution of parameter 1.
- Let $\alpha_j \in (0,1)$ with $\alpha_1 + \cdots + \alpha_k = 1$.
- Let $C_1, C_2, \dots$ be independent RVs s.t. $\mathbb{P}[C_j = \ell] = \alpha_\ell$ for all $\ell = 1, \dots, k$.

Then $(\sum_{i\geqslant 1} V_i \mathbb{1}_{C_i=1}, \dots, \sum_{i\geqslant 1} V_i \mathbb{1}_{C_i=k})$ follows Dirichlet($\boldsymbol{\alpha}$).

Theorem (Haddad–K. (2024))

For $x \geqslant 2$ and $\mathbf{u} \in [0,1]^{k-1}$, we have

$$\mathbb{P}\Big[D_x \leqslant N_x^{u_j}\ \forall j \leqslant k-1\Big] = \text{Dirichlet}(\boldsymbol{\alpha}; \mathbf{u}) + O\Bigg(\sum_{i=1}^{k-1} \frac{1}{(1+u_i \log x)^{1-\alpha_i}(1+(1-u_i)\log x)^{\alpha_i}}\Bigg).$$

A cautionary tale about divisors

Theorem (Tenenbaum 1980)

If $\mathcal{N}$ is any **positive** density set of integers, then there is **no** weak limit for the distributions

$$F_n(u) := \frac{\#\{d|n : \frac{\log d}{\log n} \leqslant u\}}{\#\{d|n\}} \quad \text{as } n \to \infty \text{ over elements of } \mathcal{N}$$

Rough reason: n typically has $\approx (\log n)^{\log 2}$ divisors; these points are neither nearly constant to get a singular measure, nor are there enough of them to cover nicely $[0, \log n]$.

Question

Is the set of $\log d$'s with $d|n$ well-spaced or does it form large clusters?

The Erdős–Hooley function

$$\Delta(n) := \max_{u\in\mathbb{R}} \#\Big\{d|n : \log d \in (u, u+1]\Big\}$$

Conjecture of Erdős (1948), proven by Maier–Tenenbaum (1985)

$\Delta(n) > 1$ for almost all integers n.

Rough reason: For a typical n, there are $\approx (\log n)^{\log 3}$ distinct fractions $\frac{d_1}{d_2}$.

Theorem (Ford–Green–K. (2023))

For almost all n, we have $\Delta(n) \geqslant (\log\log n)^{\eta+o(1)}$ with $\eta \approx 0.35332$.

- Improves on Maier–Tenenbaum (1985, 2009) and La Bretèche–Tenenbaum (2023).
- La Bretèche–Tenenbaum (2023): $\Delta(n) \leqslant (\log\log n)^{c+o(1)}$ with $c \approx 0.6102$.

Hooley's "new technique"

Theorem (Hooley (1979))

$$\frac{1}{x}\sum_{n\leqslant x}\Delta(n) \ll (\log x)^{4/\pi-1} \qquad (4/\pi - 1 < 1)$$

Remark: Hooley was motivated by many applications to problems in Diophantine equations/inequalities, e.g. he deduced $\#\{a^2+b^4+c^4 \leqslant x\} \geqslant x(\log x)^{1-\frac{4}{\pi}-o(1)}$.

Theorem (K.-Tao (2024), Ford-K.-Tao (2024))

$$(\log\log x)^{1+\eta-o(1)} \ll \frac{1}{x}\sum_{n\leqslant x}\Delta(n) \ll (\log\log x)^{11/4}$$

- Improves 2023 u.b. by La Bretèche–Tenenbaum of rough shape $\exp(c\sqrt{\log\log x})$
- Improves 1982 l.b. by Hall–Tenebaum
- La Bretèche–Tenenbaum (2024+): $(\log\log x)^{3/2} \ll \frac{1}{x}\sum_{n\leqslant x}\Delta(n) \ll (\log\log x)^{5/2}$

Ford's work

Theorem (Ford (2008))

$$\mathbb{P}_{n\leqslant x}\Big(\exists d|n,\ d\in[D,2D]\Big) \asymp (\log D)^{-\delta}(\log\log D)^{-3/2} \quad \text{with} \quad \delta = \int_1^{\frac{1}{\log 2}} \log t \,\mathrm{d}t \approx 0.08$$

- If n has $\varrho \log\log x$ prime factors, then it has $\approx (\log x)^{\varrho \log 2}$ divisors d all of whose logarithms $\log d$ lie in $[0, \log x]$.
- To have good chances to "hit" the region $[\log D, \log D + \log 2]$ we need $\varrho \geqslant 1/\log 2$.
- $\mathbb{P}_{n\leqslant x}\big(n \text{ has } \frac{1}{\log 2} \log\log x \text{ prime factors}\big) \asymp (\log x)^{-\delta}(\log\log x)^{-1/2}$.
- If for some scale y, the number of prime factors $\leqslant y$ exceeds

$$\underbrace{(1/\log 2)\cdot \log\log y}_{\text{expected amount}} + \underbrace{C}_{\text{large constant}} ,$$

then the $\log d$'s get "trapped" inside a small region.

Sub-ballistic trajectories

Arguin–Bourgade–Radziwiłł (2023+) proof of Fyodorov–Hiary–Keating conj.

- Theorem (ABR): For a.a. $\tau \in [0, T]$, $$\max_{|t-\tau|\leqslant 1} |\zeta(1/2+it)| \asymp \frac{\log T}{(\log\log T)^{3/4}}.$$
- For *fixed h*, $$\mathbb{P}_{\tau\in[0,T]}\left(|\zeta(1/2+i(\tau+h))| > \frac{\log T}{(\log\log T)^{1/4}}\right) \asymp \frac{1}{\log T}.$$
- Reason for 3/4: If $\exists y$ s.t. $$\left|\prod_{p\leqslant y}\left(1-\frac{1}{p^{1/2+it}}\right)^{-1}\right| > \underbrace{C}_{\text{large constant}} \cdot \underbrace{\frac{\log y}{(\log\log y)^{3/4}}}_{\text{expected amount}},$$

  then "there aren't enough points t" so that for one of them $|\zeta(1/2+it)|$ reaches the value $\frac{\log T}{(\log\log T)^{3/4}}$.

Zeta's cousin

Conjecture (Arguin–Bourgade–K. (2024))

Let $\tau(n;\xi) = \sum_{d|n} d^{i\xi}$. For almost all $n \leqslant x$ with $\varrho \log\log x$ prime factors, we have

$$T(n) := \max_{\xi\in[1,2]} |\tau(n;\xi)| \asymp \frac{(\log x)^{\mu(\varrho)}}{(\log\log x)^{\frac{3}{2\alpha(\varrho)}}}.$$

for certain constants $\mu(\varrho) < \log 2$ and $\alpha(\varrho) > 0$.

- Hall proved $T(n) \leqslant (\log x)^{\mu(1)+o(1)}$ for almost all $n \leqslant x$, where $\mu(1) \approx 0.65238$
- Tenenbaum proved* $T(n) \geqslant (\log x)^{1/2+o(1)}$
- Proof*** of tight u.b. in conjecture, and of weak l.b. with correct exponent of $\log x$.

Thank you for your attention