

Probability and number theory: some examples of connections

Jean-Marc Deshouillers

Mathématiques Stochastiques, Université Bordeaux 2

January 13, 1997

[summary by Alain Plagne]

Abstract

We illustrate some connections between probability theory and number theory. Examples are taken from multiplicative number theory (probabilistic behaviour of the function ω), additive number theory (Sidon problem, where arithmetic results are proved by using probability theory; partitions; and Erdős-Rényi type models for sums of powers, where situations on which nothing is known, are modelled) and probability (upper bound for the concentration of the sum of independent and identically distributed integer random variables, where probabilistic results are obtained by using methods from additive number theory).

1. Introduction

One can distinguish at least four types of connections between probability theory and additive number theory.

- (i) When probability leads to models for the integers;
- (ii) When probability techniques permit to prove number theory results;
- (iii) When number theory questions lead to probability questions;
- (iv) When number theory methods permit to prove probability results.

Part 2 is devoted to multiplicative number theory. It provides examples for points (i) and (ii). The third part, which is the heart of this talk, is concerned with additive number theory and illustrated with Sidon's problem (point (ii)), partitions (points (i) and (ii)) and sums of s -th powers (points (i) and (iii)). In the fourth part, which deals with concentration functions, point (iv) is illustrated.

2. Multiplicative Number Theory

We begin with a famous result due to Hardy and Ramanujan [12] concerning the function ω , which counts the number of divisors of an integer

$$\omega(n) = |\{p \text{ such that } p|n\}|.$$

Here and in the sequel, p always denotes a prime number.

Theorem 1. *Let $\Psi(x) \rightarrow \infty$. Then the set of integers such that*

$$(1) \quad |\omega(n) - \log \log n| \leq \Psi(n) \sqrt{\log \log n}$$

has density one.

This result means that

$$\lim_{N \rightarrow +\infty} \frac{|\{1 \leq n \leq N : (1) \text{ holds}\}|}{N} = 1.$$

The original proof is quite technical. We quote here an efficient way to get the result which is due to Turán [17].

$$\begin{aligned} N^{-1} \sum_{n \leq N} \omega(n) &= N^{-1} \sum_{n \leq N} \sum_{p|n} 1 = N^{-1} \sum_{p \leq N} \sum_{\substack{n \leq N \\ n=0 \pmod p}} 1 \\ &= N^{-1} \sum_{p \leq N} (N/p + O(1)) = \sum_{p \leq N} 1/p + O(1) = \log \log N + O(1). \end{aligned}$$

See for example [16] for the last equality. In the same way, we get

$$N^{-1} \sum_{n \leq N} \omega(n)^2 = (\log \log N)^2 + O(\log \log N),$$

thus the number of exceptions to (1) up to N is

$$|\{1 \leq n \leq N : |\omega(n) - \log \log n| \geq \Psi(n) \sqrt{\log \log n}\}| \leq \sum_{n \leq N} \frac{(\omega(n) - \log \log n)^2}{\Psi(n)^2 \log \log n},$$

which is $O(N/\Psi(N)^2)$. This proves the result.

Now let us show a probabilistic approach. Let (X_p) be a family of independent random variables defined by

$$X_p = \begin{cases} 1, & \text{with probability } 1/p, \\ 0, & \text{with probability } 1 - 1/p. \end{cases}$$

The mathematical expectation of $\omega = \sum_{p \leq N} X_p$ is

$$E(\omega) = \sum_{p \leq N} E(X_p) = \sum_{p \leq N} 1/p = \log \log N + o(\log \log N),$$

and the variance of ω is

$$V(\omega) = \sum_{p \leq N} V(X_p) = \sum_{p \leq N} (1 - 1/p)/p = \log \log N + o(\log \log N).$$

Chebyshev's inequality yields

$$\Pr \left\{ |\omega - E(\omega)| \geq \Psi \sqrt{V(\omega)} \right\} \leq 1/\Psi^2,$$

that is the result.

This result can be interpreted, with a probabilistic point of view, as a weak law of large numbers. A much more precise result, that is a central limit theorem, has been proved by Erdős and Kac in 1939 [4, 5].

Theorem 2. *For u a real, one has*

$$(2) \quad \frac{1}{N} |\{1 \leq n \leq N : \omega(n) - \log \log n \leq u \sqrt{\log \log n}\}| \xrightarrow{N \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-t^2/2} dt.$$

A natural question is now: can this result be extended to other (strongly) additive functions (that is such that $f(n) = \sum_{p|n} f(p)$)? More precisely, can one compare (and how far?)

$$\frac{1}{N} |\{1 \leq n \leq N : f(n) \leq u\}| \quad \text{with} \quad \Pr \left\{ \sum_{p \leq N} X_p \leq u \right\},$$

where the X_p 's are independent Bernoulli random variables such that $\Pr\{X_p = f(p)\} = 1/p$ and $\Pr\{X_p = 0\} = 1 - 1/p$. By sophisticated sieve arguments, it can be shown that

$$\frac{1}{N} |\{1 \leq n \leq N : \sum_{p|n, p \leq r} f(p) \leq u\}| - \Pr \left\{ \sum_{p \leq r} X_p \leq u \right\} = O \left[x^{-1/15} + \exp \left(-\frac{1}{8} \frac{\log x}{\log r} \log \frac{\log x}{\log r} \right) \right],$$

which is $o(1)$ when $r = x^{\epsilon(x)}$ tends to $+\infty$ but $\epsilon(x) \rightarrow 0$. This model is called the ‘‘Kubilius model’’.

3. Additive Number Theory

3.1. Sidon problem. In 1932, Sidon raised the following question: is it possible to find a sequence \mathcal{A} of non-negative integers such that

- (i) For every positive integer n , there exist a, b in \mathcal{A} such that $n = a + b$;
- (ii) $|\{a, n - a \in \mathcal{A}\}| = o(n^\epsilon)$ for any $\epsilon > 0$.

In other words: is it possible to build a ‘‘thin’’ basis of order 2? In 1954, Erdős [3] answered positively the question, by proving a more precise result.

Theorem 3. *There exists \mathcal{A} and $0 < c_1 < c_2$ such that*

$$(3) \quad c_1 \log n < |\{a, n - a \in \mathcal{A}\}| < c_2 \log n,$$

for $n \geq 2$.

The proof is probabilistic and particularly short but absolutely not constructive: let (Ω, T, P) be a probabilistic space and X_2, X_3, \dots be independent Bernoulli random variables such that $\Pr\{X_n = 1\} = c\sqrt{(\log n)/n}$. For $\omega \in \Omega$, let $\mathcal{A}(\omega) = \{0, 1\} \cup \{n : X_n(\omega) = 1\}$. Then for almost all $\omega \in \Omega$, $\mathcal{A}(\omega)$ satisfies (3). A very detailed proof of this result is given in the book by Halberstam and Roth [10, chap. 3].

3.2. Restricted partition function. This example underlies the probabilistic interpretation of the powerful Ramanujan-Hardy-Littlewood circle method [18].

For a given N , let $q(N)$ be the number of ways to write

$$N = n_1 + n_2 + \dots + n_r,$$

with $0 < n_1 < n_2 < \dots < n_r (\leq N)$. Another way to say this is: let

$$\mathcal{E}_N = \{0, 1\} \times \dots \times \{0, N\},$$

then

$$q(N) = 2^N \frac{|\{x \in \mathcal{E}_N, \sum x_n = N\}|}{|\{x \in \mathcal{E}_N\}|}.$$

Let X_1, \dots, X_N be independent random variables such that X_n takes the values 0 and n with probability 1/2. We have

$$q(N) = 2^N \Pr\{X_1 + \dots + X_N = N\}.$$

Denote E_N and V_N the expectation and the variance of $X_1 + \dots + X_N$. If we assume that a local limit theorem holds, then we can write

$$q(N) = 2^N \frac{1}{\sqrt{2\pi V_N}} \left(\exp\left(-\frac{(N - E_N)^2}{2V_N}\right) + o(1) \right).$$

An easy computation shows that $E_N \sim N^2/4$, $V_N \sim N^3/12$ and that the exponential term is $o(1)$ and so that $q(N) = o(2^N N^{-3/2})$ which is not interesting. This is due to the fact that N is too far from E_N .

This approach is far from being perfect. The reason is quite clear: in the problem of partitions for an integer N , the integers $1, 2, \dots, N$ do not have the same importance: the small values take part much more frequently in a decomposition of N than the large ones. Thus we have to refine the model by weighting the different integers, with a smaller weight for large integers.

Suppose now the X_i 's are independent random variables taking only the values 0 and n with $\Pr\{X_n = n\} = p_n$. Then

$$q(N) = \frac{\sum_{x \in \mathcal{E}_n, x_1 + \dots + x_N = N} \Pr\{(X_1, \dots, X_N) = x\}}{\prod_{n, x_n = n} p_n \prod_{n, x_n = 0} (1 - p_n)}.$$

If we take $p_n = \exp(-\sigma n)/(1 + \exp(-\sigma n))$, for some σ such that $E(X_1 + \dots + X_N) = N$, that is to say

$$\sigma = \frac{\pi}{2\sqrt{3N}}(1 - 1/8N + O(N^{-2})),$$

we can prove (this is naturally the heart of the matter) that

$$q(N) \sim \frac{1}{4(3N)^{1/4}} \exp(\pi N^{1/2}/\sqrt{3})$$

(see [11, 14]). This argument has been recently developed in [9] in a more general context.

Proof goes as usual, by defining the characteristic function (i.e., the Fourier transform of the image measure)

$$\phi_n(t) = (1 - p_n) + p_n \exp(2\pi i n t).$$

Then

$$\Pr\{X_1 + \dots + X_N = N\} = \int_{\mathbb{R}/\mathbb{Z}} \left(\prod_{n \leq N} \phi_n(t) \right) \exp(-2\pi i N t) dt.$$

The main term (corresponding to major arc) comes from a neighbourhood of 0 on the torus \mathbb{R}/\mathbb{Z} . There “remains” (it is not easy!) to find an upper bound for $|\prod_{n \leq N} \phi_n(t)|$ outside this neighbourhood, that is on the minor arc.

3.3. Probabilistic models for sums of powers. It is well known that sums of two squares have zero density. But for $s \geq 3$, nothing is known about (lower) density of the set of sums of s integral s -th powers, like sums of 3 cubes and of 4 biquadrates.

There are two conjectures. In 1968, Barrucand [1] conjectured that for $s = 3$ and 4 the answer is NO. But, in 1986, Hooley [13] conjectured that the answer is YES for every $s \geq 3$.

In order to guess something in this hard problem, Erdős and Rényi [6], in 1960, considered “pseudo s -th powers”, i.e., random sequences $\mathcal{A}^{(s)}$ defined as in the answer to Sidon’s question, and suggested that the number of representation $r_s(N)$ of an integer N as a sum of s elements from $\mathcal{A}^{(s)}$ should follow (almost surely) a Poisson law. Unfortunately, their proofs contained a gap because of the difficulty of dealing with the quasi-independence of the sets involved. In 1965, Halberstam and Roth [10] overcame the difficulty when $s = 2$, by combinatorial arguments. In 1995, Landreau [15]

proved a correlation inequality, having its own probabilistic interest, which leads to the expected result.

Theorem 4. *Let E_1, \dots, E_N be independent events, and A_1, \dots, A_T be such that each A_t is an intersection of some of the E_n 's. Then*

$$0 \leq \Pr(\cap \bar{A}_t) - \prod \Pr(\bar{A}_t) \leq \sum_{1 \leq t < t' \leq T} (\Pr(A_t \cap A_{t'}) - \Pr(A_t) \Pr(A_{t'})).$$

4. Probability

We illustrate here the possibility of applying number theory ideas to probability theory with the following recent theorem taken from [2], of which we sketch the proof.

Theorem 5. *Let $\frac{\log 4}{\log 3} < \sigma \leq 2$, $n \in \mathbb{N}^*$ and $\epsilon > 0$, X_1, \dots, X_n be i.i.d. integral valued random variables such that*

$$(4) \quad \max_{q \leq 2} \max_{s \bmod q} \sum_{l = s \bmod q} \Pr\{X_1 = l\} \leq 1 - \epsilon$$

and for all $L \geq 2$

$$(5) \quad Q(X_1, L) \leq 1 - L^{-\sigma},$$

where $Q(Y, x)$ denotes $\sup_{h \in \mathbb{R}} \Pr\{h < Y \leq h + x\}$. Then there exists a constant $c = c(\sigma, \epsilon, Q(X_1, 1))$ such that

$$Q(X_1 + \dots + X_n; 1) \leq cn^{-1/\sigma}.$$

The proof, in the same manner as above, uses the characteristic function ϕ of X_1 . If $S_n = X_1 + \dots + X_n$, we have

$$\Pr\{S_n = k\} = \int_0^1 \phi(t)^n \exp(-2\pi ikt) dt,$$

so that $Q(S_n, 1) \leq \int_0^1 |\phi(t)|^n dt$.

We are now reduced to study the large values of ϕ . We first use a lemma which has been introduced in [8] (and which follows from Bochner's theorem).

Lemma 1. *Let $E_\theta = \{t \in \mathbb{R}/\mathbb{Z}, |\phi(t)| \geq \cos \theta\}$, where \mathbb{R}/\mathbb{Z} is once again the torus, we have for $\theta_1, \theta_2 \geq 0$ and $\theta_1 + \theta_2 \leq \pi/2$,*

$$E_{\theta_1} + E_{\theta_2} \subset E_{\theta_1 + \theta_2}.$$

Then we need a result by Freiman [7] on structure of set addition (it has been extended to the torus in [8]).

Theorem 6. *Let \mathcal{A} be a finite subset of \mathbb{Z} such that*

$$|\mathcal{A} + \mathcal{A}| \leq 2|\mathcal{A}| - 1 + b,$$

with $b \leq |\mathcal{A}| - 3$, then there exists an arithmetic progression \mathcal{L} with $|\mathcal{A}| + b$ elements that contains \mathcal{A} .

This enables us to show that either the set of the arguments for which ϕ is large has a small measure, or it has a structure (it is located close to the vertices of a regular polygon), which is not possible in view of (4) and (5).

References

- [1] Barrucand (Pierre). – Sur la distribution empirique des sommes de trois cubes ou de quatre bicarrés. *Comptes-Rendus de l'Académie des Sciences*, vol. 267, 1968, pp. 409–411.
- [2] Deshouillers (J.-M.), Freiman (G. A.), and Yudin (A. A.). – *On Bounds for the Concentration Function, 1*. – Prépublication n° M/95/37, IHES, 1995.
- [3] Erdős (P.). – On a problem of Sidon in additive number theory. *Acta Scientiarum Mathematicarum Szegediensis*, vol. 15, 1954, pp. 255–259.
- [4] Erdős (P.) and Kac (M.). – On the Gaussian law of errors in the theory of additive functions. *Proceedings of the National Academy of Sciences of the USA*, vol. 25, 1939, pp. 206–207.
- [5] Erdős (P.) and Kac (M.). – The Gaussian law of errors in the theory of additive number theoretic functions. *American Journal of Mathematics*, vol. 62, 1940, pp. 738–742.
- [6] Erdős (P.) and Rényi (A.). – Additive properties of random sequences of positive integers. *Acta Arithmetica*, vol. 6, 1960, pp. 83–110.
- [7] Freĭman (G. A.). – *Foundations of a structural theory of set addition*. – American Mathematical Society, Providence, R. I., 1973, *Translations of Mathematical Monographs*, vol. 37. Translated from the Russian.
- [8] Freiman (G. A.), Moskvin (D. A.), and Yudin (A. A.). – Inverse problems of additive number theory and local limit theorem for lattice random variables. In *Number-theoretic studies in the Markov spectrum and in the structural theory of set addition*, pp. 148–162. – 1973.
- [9] Freiman (Gregory A.) and Pitman (Jane). – Partitions into distinct large parts. *Journal of the Australian Mathematical Society*, vol. 57, n° 3, 1994, pp. 386–416.
- [10] Halberstam (Heini) and Roth (Klaus Friedrich). – *Sequences*. – Springer-Verlag, New York-Berlin, 1983, 2nd edition.
- [11] Hardy (G. H.) and Ramanujan (S.). – Asymptotic formulæ in combinatory analysis. *Proceedings of the London Mathematical Society*, vol. 16, 1917, pp. 75–115.
- [12] Hardy (G. H.) and Ramanujan (S.). – The normal number of prime factors of a number n . *Quarterly Journal of Mathematics*, vol. 48, 1917, pp. 76–92.
- [13] Hooley (Christopher). – On some topics connected with Waring's problem. *Journal für die Reine und Angewandte Mathematik*, vol. 369, 1986, pp. 110–153.
- [14] Hua (Loo-Keng). – On the number of partitions of a number into unequal parts. *Transactions of the American Mathematical Society*, vol. 51, 1942, pp. 194–201.
- [15] Landreau (Bernard). – Étude probabiliste des sommes de s puissances s -ièmes. *Compositio Mathematica*, vol. 99, n° 1, 1995, pp. 1–31.
- [16] Tenenbaum (Gérald). – *Introduction to analytic and probabilistic number theory*. – Cambridge University Press, Cambridge, 1995, *Cambridge Studies in Advanced Mathematics*, vol. 46. Translated from the second French edition (1995).
- [17] Turán (P.). – On a theorem of Hardy and Ramanujan. *Journal of the London Mathematical Society*, vol. 9, 1934, pp. 274–276.
- [18] Vaughan (R. C.). – *The Hardy-Littlewood method*. – Cambridge University Press, Cambridge, 1997, 2nd edition, *Cambridge Tracts in Mathematics*, vol. 125.