

Ramanujan and Probabilistic Number Theory

Gérald Tenenbaum

It has been said that every natural number was a personal friend to Ramanujan. Counting indeed corresponds to a fundamental necessity of human kind: acting on the world. This has incidentally never been more true than in our modernity, in which almost all our activities get numerically coded.

When, at the beginning of the twentieth century, probability theory emerged from the theory of games and, partly due to the influence of actuaries, started to develop as a genuine part of mathematics, who could imagine that the theory of numbers would have anything to do with randomness? Bachelier had investigated models for stock exchange quotes using random variables and had employed stochastic calculus to study their variations. But, venerable arithmetic seemed preserved from these nebulous zones: how could such a crucial basis depend on uncertainty?

Among other advances, Kolmogorov has given a thorough axiomatic construction to probability theory, and the topic has gradually become part of measure theory, hence of mathematics as a whole. It is now straightforward to equip the set of natural integers $\mathbb{N}^* := \{1, 2, 3, \dots\}$ with a probability measure: it suffices to consider a series with non-negative terms

$$\sum_{n \geq 1} p_n = 1$$

and to define the probability measure of an integer sequence $\mathcal{A} \subset \mathbb{N}^*$ by the formula

$$\mathbb{P}(\mathcal{A}) := \sum_{a \in \mathcal{A}} p_a.$$

However, a simple exercise (see, e.g., [19; th. III.1.1]) shows that this setting is incompatible with the natural constraint that the probability of the set of multiples of a given integer m should be equal to $1/m$. The proof only uses the divergence of the series of the reciprocals of the primes, a property already known to Euler. As a consequence, it may be stated that, from the strict framework of probability theory, there is no satisfactory notion of a random integer. By ‘random’ one asks here for a notion according to common sense: a random integer should be even with probability $1/2$, coprime to 6 with probability $5/6$, it should almost never be a prime or a square, and so on.

If mathematics are designed to model the world around us, they are also shaped to support intuition. Consistently, when laying the foundation of probabilistic number theory, Hardy and Ramanujan [5] departed from a purely probabilistic setting and introduced the concept of *normal order* of an arithmetic function.⁽¹⁾ Here is a quote from their epoch making paper:

[..] We may ask what is the *normal order* of [a] function. This phrase requires a little more explanation.

1. That is a function $f : \mathbb{N}^* \rightarrow \mathbb{R}$. One may consider arithmetic functions with complex values or even values in an abstract set, but for purpose of exposition, and in order to stick to the historical viewpoint, this survey will restrict to real valued function.

Suppose that $N(x)$ is the number of numbers, not exceeding x , which possess a certain property P . This property may be a function of n , or of x only, or of both n and x : we shall be concerned only with cases in which it is a function of one variable alone. And suppose further that

$$N(x) \sim x$$

when $x \rightarrow \infty$. Then we shall say, if P is a function of n only, that almost all numbers possess the property, and, if P is a function of x only that *almost all numbers less than x* possess the property. Thus, to take a trivial example, almost all numbers are composite.

The wonderful underlying idea introduced here consists in replacing mere probability by what is nowadays commonly called (*natural*) *density*. If a subset $\mathcal{A} \subset \mathbb{N}^*$ is such that its *frequency* among the first N integers tends to some number δ as $N \rightarrow \infty$, then it is said that \mathcal{A} has *density* δ and one writes $d(\mathcal{A}) = \delta$. Thus, there is a formal link between density and probability: the former is the limit, when it exists, of the latter, computed on the set of the first N integers equipped with the uniform measure.

Natural density is not a probability: it does not satisfy the axiom of countable additivity, which states that the probability of a union of a finite or countably infinite collection of disjoint events is the sum of the corresponding probabilities. Every finite subset of \mathbb{N}^* has density 0 but \mathbb{N}^* itself has density 1. For the same reason, the collection $\mathcal{P}(\mathbb{N}^*)$ of all subsequences of \mathbb{N}^* possessing a density is not a σ -algebra. Finally, it is easy to construct sequences failing to have natural density: we leave to the reader the proof that this is so for the sequence of those integers n with leading digit 1 in the expansion to base 10.

However, density is the powerful tool that gives birth to probabilistic number theory. The set of multiples of any number m has density $1/m$ and, more generally, any intuitive statement on the structure of ‘almost all’ integers turns out to be confirmed in this setting.

The arithmetic counterpart to random variables is the notion of an arithmetic function. These functions usually vary in an intrinsically irregular and erratic way, with the consequence that standard analytical techniques are often useless for describing their behavior. With the idea of normality, Hardy and Ramanujan provide a new path to understanding these objects: aside of the average and extremal orders, easy to define and allowing a cursory classification, they propose a concept capable of reflecting the ‘almost certain’ behavior.

While the notion of an almost everywhere constant random variable is rather infertile, that of the normal order of an arithmetic function is incredibly fruitful. Neglecting a set of density zero so as to eliminate scarce, aberrant values, it sheds a new light on variations that otherwise looked desperately chaotic. Order and regularity suddenly emerge, seemingly by magic. Let us quote [5] again:

If [...] $f(n)$ is an arithmetical function of n and $\varphi(n)$ an elementary increasing function, and if, for every positive ε , we have

$$(1 - \varepsilon)\varphi(n) < f(n) < (1 + \varepsilon)\varphi(n)$$

for almost all values of n , we shall say that *the normal order of $f(n)$ is $\varphi(n)$* .

Of course, as Hardy and Ramanujan remark,

[...] it is in no way necessary that a function should possess either a determinate average order or a determinate normal order, or that one should be determinate when the other is, or that, if both are determinate, they should be the same.

Similarly, a given function may have several normal orders, which must all share the same asymptotic behavior. Since the notion is obviously pertinent only for those functions f whose normal order is in a sense simpler than that of f , Hardy and

Ramanujan added the restriction that φ should be elementary and increasing. Today, *elementary* is usually understood as *that can be expressed by means of the symbols of real analysis* while, in the aim of allowing as much flexibility as possible, the prescription of monotonicity has been dropped.

The main normal order result proved in [5] is that, if one defines $\omega(n)$ (resp. $\Omega(n)$) as the total number of prime factors of an integer n counted without (resp. with) multiplicity then both functions have average order and normal order equal to $\log_2 n$. Here and in the sequel, we let \log_k denote the k -fold iterated logarithm to the base e .

Arithmetic is the study of familiar, natural integers, on which two structures are naturally defined: the additive structure, generating all integers from 0 by adding 1 recursively, and the multiplicative structure, generating all integers by multiplying primes together. Basically, one may say that most problems in the field arise from trying to describe mutual interaction between these structures: Fermat's Last Theorem says that perfect powers (multiplicatively defined) to a same exponent $n \geq 3$ cannot add to a perfect n -power, Goldbach's conjecture stipulates that adding two primes generates all even integers, etc. The extraordinary discovery made by Hardy and Ramanujan is that, once the prism of normality is applied *the multiplicative structure⁽²⁾ of an integer only depends on its size*. This is indeed a direct link between the two structures!

The original proof presented in [5] is rather complicated. Through induction the authors show that, for instance in the case of the ω -function, we have, for suitable absolute constants A and B ,

$$(1) \quad \pi_k(x) := |\{n \leq x : \omega(n) = k\}| \leq \frac{Ax(\log_2 x + B)^{k-1}}{(k-1)! \log x} \quad (x \geq 3, k \geq 1).$$

Equipped with such an upper bound, they can easily deduce, by estimating proper tails, that $\omega(n)$ is close to the mean-value $\log_2 x$ for almost all integers $n \leq x$. Actually, given any function $\xi(x) \rightarrow \infty$, they show that

$$(2) \quad |\omega(n) - \log_2 x| \leq \xi(x) \sqrt{\log_2 x}$$

for all but at most $o(x)$ integers $n \leq x$.

At this stage, it should be mentioned that [5] is not the only paper by Ramanujan dealing with the birth of probabilistic number theory. In 1915, he introduced [18] highly composite numbers, which can roughly be described as integers for which the total number of divisors presents a local extremum. And, in the same year 1917, Hardy and Ramanujan published another work [6] in which they provide effective estimates for the densities of subsets of \mathbb{N}^* defined by multiplicative constraints—the probabilistic approach was well anchored in their perspective on arithmetics. We shall not discuss these works here, but stress out the fact that, as is often the case for number theory, new ideas were required to tackle these problems of a new kind: in [18] the technique known as the *benefit method* was introduced,⁽³⁾ and in [6], a novel Tauberian theorem of considerable interest was established in order to estimate the size of an extended class of highly composite numbers.

As can be expected, the posterity of the ideas introduced in [5] is extraordinarily vast. As soon as the early thirties, Erdős systematically used the device that the 'anatomy' of a random integer may be described from its size only. Let $\{p_j(n)\}_{j=1}^{\omega(n)}$

2. Through the number of prime factors here, but subsequent works widen significantly the scope.

3. See, e.g., [17] for further information.

denote the increasing sequence of distinct prime factors of an integer n . The final form of the result was announced in 1946⁽⁴⁾: given $\varepsilon > 0$ and $\xi(x) \rightarrow \infty$, almost all $n \leq x$ satisfy

$$\sup_{\xi(x) \leq j \leq \omega(n)} \left| \frac{\log_2 p_j(n) - j}{\sqrt{2j \log_2 j}} \right| \leq 1 + \varepsilon,$$

and the result is optimal in a strong sense: replacing ε by $-\varepsilon$ corresponds to $o(x)$ integers not exceeding x . So, from a statistical point of view, the complete multiplicative structure of a normal integer n can be precisely described: not only the total number of prime factors may be approximated by a function depending only on the size of n , but almost all prime factors may themselves be evaluated by means of their sole rank.

In 1934, Turán provided a new, very simple proof of the Hardy-Ramanujan theorem (2). The main idea was to rely even more on a probabilistic framework and invoke the classical Bienaymé-Chebyshev inequality. It worked wonderfully and thereby naturally lent itself to fruitful generalizations. Systematically exploited by Kubilius and his collaborators in the fifties and sixties, the underlying setting consists in considering divisibility by primes and prime powers as genuine random variables on the set Ω_N of the first N integers, equipped with uniform probability ν_N . For instance, letting $\xi_p(n) = 1$ if p divides n but p^2 does not, and writing \mathbb{E}_N for the expectation with respect to ν_N , we have

$$\mathbb{E}_N(\xi_p) = \frac{1}{N} \left(\left\lfloor \frac{N}{p} \right\rfloor - \left\lfloor \frac{N}{p^2} \right\rfloor \right) = \frac{1 - 1/p}{p} + O\left(\frac{1}{N}\right)$$

and for distinct primes p, q ,

$$\mathbb{E}_N(\xi_p \xi_q) = \frac{(1 - 1/p)(1 - 1/q)}{pq} + O\left(\frac{1}{N}\right) = \mathbb{E}_N(\xi_p) \mathbb{E}_N(\xi_q) + O\left(\frac{1}{N}\right).$$

Thus, the random variables ξ_p and ξ_q are close to being independent provided the involved primes are not too large. If an arithmetic function f ascribes a secondary role to large primes, as is the case for the Ω and ω functions, then probabilistic results devised for independent variables are likely to hold for f —and they actually do in the present case. Further details and complements may be found in [19; ch. III.3].

It is well known that sums of independent random variables obey the strong law of large numbers and that, when suitably normalized, they converge to the Gaussian law. Thus, with hindsight, it could have been expected that a Gaussian law would apply to complement the Hardy-Ramanujan theorem. This was done in 1940 by Paul Erdős and Mark Kac [3]: uniformly for $N \geq 2$ and $y \in \mathbb{R}$, we have

$$(3) \quad \nu_N\{\omega(n) \leq \log_2 N + y\sqrt{\log_2 N}\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-t^2/2} dt + O\left(\frac{1}{\sqrt{\log_2 N}}\right).$$

The sharp form of the error term was first provided by Rényi and Turán [12] in 1958. For a modern proof, see [19; ch. III.4]. The Poisson law already emerging from (1) has been made effective and intensively studied after works of Sathe [13] and Selberg [14] in the fifties: it furnishes an alternative path to (3). Sathe's paper solved a conjecture of Hardy from 1936 (see [7; p. 56]) related to the case when $\omega(n)$ is close to its mean $\log_2 N$. It relied on a complicated induction. Then Selberg gave an elegant analytic approach.

4. See [4; ch. 1] for a detailed proof.

Erdős and Kac later generalized (3) to certain arithmetical functions that are linear combinations of the ξ_p and their variants associated to prime powers. These functions are characterized by the fact that the image of a product of two coprime integers is the sum of the images. Consequently referred to as *additive*, they stand as analogues in number theory of sums of (quasi) independent variables in probability theory. An important part of probabilistic number theory concerns the distribution of additive function: see, e.g., [1] or [19] for a comprehensive discussion of the subject.

The probabilistic flavour of statements proved using the ideas described above led to the construction a formal model of the probability space (Ω_N, ν_N) by an abstract space (Ω, \mathbb{P}) in which the random variables mimicking divisibility by small primes are strictly independent. This abstract space is usually designated as Kubilius' model—see Kubilius [9], Elliott [1], and the author's description [19; ch. III.6]. The total variation distance (which is the strongest possible) between the two spaces is then defined by the formula

$$K(N, u) := \sup_{A \subset \mathbb{R}} |\nu_N(A) - \mathbb{P}(A)|,$$

where $u \geq 1$ indicates that only events depending on primes $\leq N^{1/u}$ are considered. Kubilius [8] proved in 1956 that $K(N, u) \rightarrow 0$ if, and only if, $u \rightarrow \infty$ with N . The best estimate to date [16] is that, given any $\varepsilon > 0$ and a suitable constant C , we have

$$K(N, u) \leq Cu^{-u} + CN^{-1+\varepsilon}.$$

It would be impossible to describe the numerous descendants of the ideas originated in the seminal article of Hardy and Ramanujan. Probabilistic ideas are ubiquitous in number theory nowadays. Let us only mention two examples among so many. First, in the opposite direction to Kubilius' model, arithmetic models of probabilistic objects have been constructed: see, e.g., [11] and [10] for the case of Brownian motion. Second, a deep, eighty year old conjecture of Erdős⁽⁵⁾ has recently been confirmed by Field medallist Terence Tao [15], who accomplished this *tour de force* by, in particular, measuring the entropy of multiplicative arithmetic functions—hence shedding light on arithmetical objects by means of probabilistic concepts.

References

- [1] P.D.T.A. Elliott, Probabilistic number theory : mean value theorems, Grundlehren der Math. Wiss. 239; Probabilistic number theory : central limit theorems, *ibid.* 240, Springer-Verlag, New York, Berlin, Heidelberg 1979, 1980.
- [2] P. Erdős, On the distribution function of additive functions, *Ann. of Math.* **47** (1946), 1–20.
- [3] P. Erdős & M. Kac, The Gaussian law of errors in the theory of additive number theoretic functions, *Amer. J. Math.* **62** (1940), 738–742.
- [4] R.R. Hall & G. Tenenbaum, *Divisors*, Cambridge tracts in mathematics 90, Cambridge University Press (1988, paperback ed. 2008).
- [5] G.H. Hardy & S. Ramanujan, The normal number of prime factors of a number n , *Quart. J. Math.* **48** (1917), 76–92.
- [6] G.H. Hardy & S. Ramanujan, Asymptotic formulae for the distribution of integers of various types, *Proc. London Math. Soc.* (2) **16** (1917), 112–123.
- [7] G.H. Hardy, *Ramanujan: Twelve lectures on subjects suggested by his life and work*, Cambridge University Press, Cambridge, England; Macmillan Company, New York, 1940. vii+236 pp.
- [8] J. Kubilius, Probabilistic methods in the theory of numbers. *Uspehi Mat. Nauk* (N.S.) **11** (1956), 2(68), 31–66 ; = Amer. Math. Soc. Translations, vol. 19 (1962), 47–85.
- [9] J. Kubilius, *Probabilistic methods in the theory of numbers*, Amer. Math. Soc. Monographs 11, Providence 1964.
- [10] E. Manstavičius, Natural divisors and the Brownian motion, *J. Théor. Nombres Bordeaux* **8** n° 1 (1996), 159–171.

5. Known as the *discrepancy conjecture*.

- [11] W. Philipp, Arithmetic functions and Brownian motion, in: *Analytic Number Theory*, Proc. Sympos. Pure Math. (St. Louis 1972) 24 (1973), 233–246.
- [12] A. Rényi & P. Turán, On a theorem of Erdős–Kac, *Acta Arith.* **4** (1958), 71–84.
- [13] L.G. Sathe, On a problem of Hardy on the distribution of integers having a given number of prime factors, I, II, *J. Indian Math. Soc.* **17** (1953), 63–141; III, IV, *ibid.* **18** (1954), 27–81.
- [14] A. Selberg, Note on the paper by L.G. Sathe, *J. Indian Math. Soc.* **18** (1954), 83–87.
- [15] T. Tao, The Erdős discrepancy problem, *Discrete Anal.* 2016, Paper No. 1, 29 pp.
- [16] G. Tenenbaum, Crible d’Ératosthène et modèle de Kubilius, in: K. Györy, H. Iwaniec, J. Urbanowicz (eds.), *Number Theory in Progress*, Proceedings of the conference in honor of Andrzej Schinzel, Zakopane, Poland 1997, 1099–1129, Walter de Gruyter, Berlin, New York.
- [17] J.-L. Nicolas, On highly composite numbers, in: *Ramanujan revisited* (Urbana-Champaign, Ill., 1987), 215–244, Academic Press, Boston, MA, 1988.
- [18] S. Ramanujan, Highly composite numbers, *Proc. London Math. Soc.* (2) **14** (1915), 347–409.
- [19] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Graduate Studies in Mathematics 163, Amer. Math. Soc. 2015.

Gérald Tenenbaum
Institut Élie Cartan de Lorraine
Université de Lorraine
BP 70239
54506 Vandœuvre-lès-Nancy Cedex
France
gerald.tenenbaum@univ-lorraine.fr