概率及中心极限定理在数论上的应用

魏弋凡

2018 年 5 月 11 日

目录

1 前言

我曾被同学问到这样一个问题，

“任取一个的自然数不含平方因子的概率是多少？”

假设不超过 N 的无平方因子数的个数为 $SF(N)$，如果我们把 $\frac{SF(N)}{N}$ 当作是一次又一次实验，挑到无平方因子数的频率，是不是可以直接把其极限（如果存在的话）作为概率呢？

根据 $SF(N)$ 的增长方式（证明见附录），可知上述极限存在且等于 $\frac{6}{\pi^2}$.

但是，这样思考到底合不合理呢？

2 概率数论简介

2.1 自然数集上的概率测度

要想直接使用概率论的工具，就需要建立一个满足 Kolmogorov 公理的概率测度。然而事实上这样的概率测度可能跟我们的直觉有出入，比如这样的概率测度 P 是不存在的（证明可参考 [4]），

$$P(a\mathbb{N}) = \frac{1}{a}, \qquad (a \in \mathbb{N})$$

其中 $a\mathbb{N}$ 表示 a 的所有倍数构成的集合. 但是我们可以这样定义概率测度 [4]，

若 $\sum_{n=1}^{\infty} \lambda_n = 1$，且 $\lambda_n \geq 0$. 对自然数集的某一子集 A，定义

$$P(A) = \sum_{n \in A} \lambda_n.$$

容易验证，这样的 P 将满足概率测度的公理. 常用的概率测度有均匀分布 ν_N，对某一自然数 N,

$$\lambda_n = \begin{cases} \dfrac{1}{N} & , n \leq N, \\ 0 & , n > N, \end{cases} \qquad \nu_N(A) = \sum_{n \in A} \lambda_n = \frac{1}{N} \#\{n : n \in A, n \leq N\}.$$

并且将极限（如果存在的话）$\lim_{n\to\infty} \nu_N(A)$ 称为 A 的自然密度 $d(A)$（上下极限分别对应称为上下自然密度）.

2.2 数论函数的极限分布

许许多多的数论函数有着诡异而且剧烈的变化方式，比如说 Euler 函数 $\phi(n)$，定义为小于 n 且与 n 互素的自然数个数. 见图 1. 事实上，$\phi(n)$ 的变化相当剧烈，以至于 $\frac{\phi(n+1)}{\phi(n)}$ 在整个正实数轴上稠密，见 [1].

又比如说图 2中的 Mertens 函数 $M(x) = \sum_{k=1}^{n} \mu(k)$，其中 $\mu(n)$ 是 Möbius 函数，见 [3]. 虽然每一步都只有很小的变化，但看起来却相当地“随机”. 著名的 Mertens 猜想声称对任意自然数 n，$|M(n)| < \sqrt{n}$. 遗憾的是该猜想并不成立[1].

[1]有趣的是，Riemann 猜想的一个等价命题却是 $\forall \varepsilon > 0$，$M(x) = O\left(x^{\frac{1}{2}+\varepsilon}\right)$

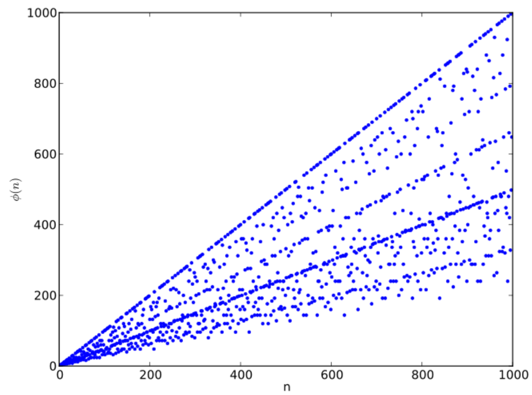

图 1: Euler 函数的前 1000 个值

可见，为众多的数论函数都去找显式公式是很不现实的，不如把他们看成随机变量。联想到概率论中的随机变量的极限分布，我们可以这样来定义数论函数 $f(n)$ 的极限分布，定义

$$F_N(x) = \nu_N\{n : f(n) \leq x\}$$

容易验证这是分布函数，若有某个分布函数 $F(x)$ 使得

$$F_N \xrightarrow{W} F,$$

则称数论函数 $f(n)$ 有极限分布.

2.3 期望

从这个定义上看，我们现在甚至还能讨论有极限分布的数论函数的期望. 以 $\frac{\phi(n)}{n}$ 为例，由 $\phi(n)$ 的定义以及 Möbius 变换有

$$n = \sum_{d|n} \phi(d) \Rightarrow \phi(n) = \sum_{d|n} \mu(d) \frac{n}{d}.$$

因此，

$$\begin{aligned}\sum_{n\leq N} \frac{\phi(n)}{n} = \sum_{n\leq N} \sum_{d|n} \frac{\mu(d)}{d} = \sum_{d\leq N} \frac{\mu(d)}{d} \left\lfloor \frac{N}{d} \right\rfloor \\ = N \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + O\left(\sum_{d>N} \frac{1}{d^2} + \log N \right).\end{aligned}$$

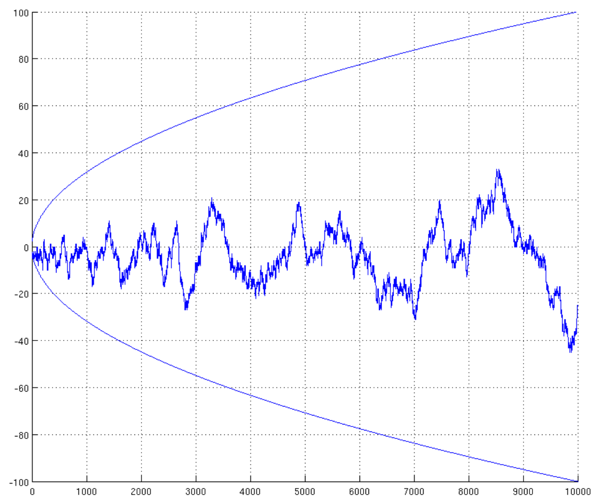

图 2: Mertens 函数的前 10000 个值，曲线为 $\pm\sqrt{n}$

这样我们将得到 $\frac{\phi(n)}{n}$ 的极限分布的期望

$$
\begin{aligned}
E\left(\frac{\phi(n)}{n}\right) &= \lim_{N\to\infty} E_N\left(\frac{\phi(n)}{n}\right) \\
&= \lim_{N\to\infty}\int_{-\infty}^{\infty} x dF_N(x) \\
&= \lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\frac{\phi(n)}{n} \\
&= \sum_{d=1}^{\infty}\frac{\mu(d)}{d^2}.
\end{aligned}
$$

而这个级数绝对收敛，其值为 $\dfrac{6}{\pi^2}$.

2.4 Erdős-Kac 定理

既然期望都可以有，当然特征函数的方法也适用，由 Lévy 连续性定理还可以知道 F 能由 F_N 的特征函数求得. 本定理正好可以用特征函数证出.

定理的对象是所谓的素因子计数函数 $\omega(n)$，表示自然数 n 包含的不同素因子的个数. 该定理声称,

$$\lim_{N\to\infty} \nu_N \left\{ n : \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}} \le x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

正是一个中心极限定理！

该定理的证明首先依赖于一个解析数论上的结果，

$$\sum_{n\le x} z^{\omega(n)} = \lambda(z)x(\log x)^{z-1} + O\left(x(\log x)^{\Re z-1}\right)$$

其中 $z \in \mathbb{C}$，$\lambda(z)$ 是一个整函数.

由此得出关于 $F_N(x)$ 的特征函数 $\phi_N(t)$ 的一个估计，发现其极限恰为

$$e^{-\frac{t^2}{2}},$$

标准正态分布的特征函数，这就证明完毕了。

3 附录

本证明参考了 [2](I.4). 引进 Möbius 函数 $\mu(n)$. 可知

$$SF(N) = \sum_{n\le N} |\mu(n)|,$$

又因为每个自然数 n 都是一个非平方部分 q 和一个平方部分 m^2 的乘积，所以，

$$|\mu(n)| = \begin{cases} 1, & m = 1, \\ 0, & m \neq 1. \end{cases}$$

于是，由 Möbius 函数的一条优美性质可知，

$$|\mu(n)| = \sum_{d|n} \mu(d)$$

注意到 $d|m$ 等价于 $d^2|n$，于是有，

$$\begin{aligned} SF(N) = \sum_{n\le N} |\mu(n)| &= \sum_{n\le N} \sum_{d|m} \mu(d) \\ &= \sum_{n\le N} \sum_{d^2|n} \mu(d) \\ &= \sum_{d^2\le N} \mu(d) \left\lfloor \frac{N}{d^2} \right\rfloor \\ &= N \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + N \cdot O\left(\sum_{d^2>N} \frac{1}{d^2} \right) + O(\sqrt{N}). \end{aligned}$$

又因为（见 [5]），

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \prod_p \left(1 - \frac{1}{p^s}\right) = \frac{1}{\zeta(s)}.$$

于是得到当 $N \to \infty$ 时，

$$SF(N) = \frac{6}{\pi^2} N + O(\sqrt{N}).$$

"*It is evident that the primes are randomly distributed but, unfortunately, we don't know what 'random' means.*"

R. C. Vaughan (摘自 [4] 末尾)

References

[1] A. Schinzel. "Generalisation of a theorem of B.S.K.R. Somayajulu on the Euler's function $\phi(n)$". In: *Ganita* 5 (1954), pp. 123–128.

[2] Gérald Tenenbaum. *Introduction to Analytic and Probabilistic Number Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1995. ISBN: 978-0521412612.

[3] Mathworld. *Möbius function*. URL: `http://mathworld.wolfram.com/MoebiusFunction.html`.

[4] Dr. Jörn Steuding, Dedicated To Prof. Jonas Kubilius, and Jörn Steuding. *Probabilistic Number Theory*.

[5] Wikipedia. *Euler product*. URL: `https://en.wikipedia.org/wiki/Euler_product`.