# E4.2: A distribution on $\mathbb{N}$ related to zeta function

## Xinyu Mao

## April 11, 2020

Let $s > 1$, and define $\zeta(s) := \sum_{n \in \mathbb{N}} \frac{1}{n^s}$, as usual. Let $X$ and $Y$ be independent $\mathbb{N}$-valued random variables with

$$\mathcal{P}(X = n) = \mathcal{P}(Y = n) = \frac{n^{-s}}{\zeta(s)}.$$

Define events $E_n := \{X \text{ is divisible by } n\}$. One can easily see that

$$\mathcal{P}(E_n) = \sum_{i \in \mathbb{N}} \mathcal{P}(X = ni) = \sum_{i \in \mathbb{N}} \frac{(ni)^{-s}}{\zeta(s)} = \frac{n^{-s}}{\zeta(s)} \sum_{i \in \mathbb{N}} i^{-s} = n^{-s}. \tag{1}$$

Everything is set up now, and we shall explore some interesting properties of $X$ and $Y$.

**Proposition 1.** $(E_p : p \in \mathbb{P})$ *are independent with $\mathbb{P}$ denoting all prime numbers.*

*Proof.* Let $S := \{p_1, p_2, \ldots, p_m\}$ be a finite subset of $\mathbb{P}$. By definition of $E_n$ and Eq. (1), we immediately have

$$\mathcal{P}\left(\bigcap_{i=1}^m E_i\right) = \mathcal{P}\left(E_{p_1 p_2 \cdots p_m}\right) = (p_1 p_2 \cdots p_m)^{-s} = \prod_{i=1}^m \mathcal{P}(E_i),$$

which implies $(E_p : p \in \mathbb{P})$ are independent. $\qquad\square$

We can use this observation to cast light on **Euler's formula**

$$\frac{1}{\zeta(s)} = \prod_{p \in \mathbb{P}} (1 - \frac{1}{p^s}). \tag{2}$$

The LHS of Eq. (2) is obviously $\mathcal{P}(X = 1)$. The event that $X = 1$ can also be described as $E := \{p \nmid X, \forall p \in \mathbb{P}\} = \bigcap_{p \in \mathbb{P}} E_p^c$. Combining Eq. (1) and Proposition 1, we have $\mathcal{P}(E) = \prod_{p \in \mathbb{P}} \mathcal{P}(E_p^c) = \prod_{p \in \mathbb{P}}(1 - \frac{1}{p^s})$.

Clearly, $(E_{p^2}, p \in \mathbb{P})$ is independent as well(the argument is similar to Proposition 1). Hence,

$$\mathcal{P}(\text{no square other than 1 divides } X) = \mathcal{P}\left(\bigcap_{p \in \mathbb{P}} E_{p^2}^c\right) = \prod_{p \in \mathbb{P}}(1 - \frac{1}{p^{2s}}) = \frac{1}{\zeta(2s)},$$

where the last equality follows from Eq. (2).

Define r.v. $H := \gcd(X, Y)$, and we shall show that

**Proposition 2.** $\mathcal{P}(H = n) = \frac{n^{-2s}}{\zeta(2s)}$.

*Proof.* We need the concept of conditional probability from elementary theory. Note that

$$\mathcal{P}\left(X/n = k | E_n\right) = \frac{(nk)^{-s}}{\sum_{i \geq 1}(ni)^{-s}} = \frac{k^{-s}}{\zeta(s)}.$$

Hence, $(X/n | E_n)$ has the same distribution as $X$. Let $F_n$ be the event that $n$ divides $Y$, then $(Y/n | E_n)$ also has the same distribution. Thus,

$$\begin{aligned}
\mathcal{P}(H = n) &= \mathcal{P}\left(E_n \cap F_n \cap \gcd(\frac{X}{n}, \frac{Y}{n}) = 1\right) \\
&= \mathcal{P}(E_n)\mathcal{P}(F_n)\mathcal{P}\left(\gcd(\frac{X}{n}, \frac{Y}{n}) = 1 | E_n \cap F_n\right) \\
&= n^{-s} \cdot n^{-s} \cdot \mathcal{P}(\gcd(X, Y) = 1).
\end{aligned} \tag{3}$$

Since $(E_p, F_p : p \in \mathbb{P})$ is independent,

$$\mathcal{P}(\gcd(X, Y) = 1) = \mathcal{P}(\bigcap_{p \in \mathbb{P}}(E_p \cap F_p)^c) = \prod_{p \in \mathbb{P}}(1 - \frac{1}{p^{2s}}) = \frac{1}{\zeta(2s)}.$$

Plugging this into Eq. (3), we get the desired result. □

However, we have not rigorously defined conditional probability yet. Can we avoid using conditional probability? The answer is yes, but we need a useful tool, **Möbius function**, which is defined as

$\mu : \mathbb{N} \rightarrow \{-1, 0, 1\}$

$n \mapsto \begin{cases} (-1)^m, \text{if } n = p_1 p_2 \cdots p_m, \text{ where } (p_i)_{i \in [m]} \text{ are distinct prime numbers;} \\ 0, \text{otherwise.} \end{cases}$

Here we list two facts about $\mu$ without giving detailed proof [1].

**Proposition 3.** *(i)* $\sum_{d|n} \mu(d) = [n = 1]$.

*(ii)* $\sum_{n \in \mathbb{N}} \frac{\mu(n)}{n^s} = \prod_{p \in \mathbb{P}} \frac{1}{(1-p^s)} = \frac{1}{\zeta(s)}$.

*Proof.* (i) can be proved by counting the number of $-1$ and $+1$ among all $\mu(d)$. (ii) is simply expansion of the product. $\qquad\square$

Now we are well-equipped to give the coup de grace.

*Another proof of Proposition 2.* Let $G_n$ be the event that $X, Y$ are both divisible by $n$, which is equivalent to $n|H$. Clearly, $\mathcal{P}(G_n) = n^{-2s}$. By Proposition 3 (i),

$$\mathcal{P}(H = n) = \sum_{k=1}^{\infty} \mathcal{P}(H = kn) \sum_{d|k} \mu(d) = \sum_{d=1}^{\infty} \mu(d) \sum_{i=1}^{\infty} \mathcal{P}(H = idn). \qquad (*)$$

Note that $\sum_{i=1}^{\infty} \mathcal{P}(H = in) = \mathcal{P}(G_n)$ for all $n \in \mathbb{N}$. Hence,

$$(*) = \sum_{d=1}^{\infty} \mu(d) \mathcal{P}(G_{dn}) = \sum_{d=1}^{\infty} \mu(d)(dn)^{-2s}$$

$$= n^{-2s} \sum_{d=1}^{\infty} \mu(d) d^{-2s}$$

$$= \frac{n^{-2s}}{\zeta(2s)},$$

where the last equality follows from Proposition 3 (ii). $\qquad\square$

**Remark.** *If we take $s = 1$(this makes no sense) in Proposition 2, we get the probability that $X, Y$ are relatively prime is $\frac{1}{\zeta(2)}$. It is a coincidence that the **arithmetic density** of relatively prime pairs among $\mathbb{N}^2$ is indeed $\frac{1}{\zeta(2)}$! Formally,*

$$\lim_{N \to \infty} \frac{\sum_{i,j \in [N]} [\gcd(i,j) = 1]}{N^2} = \frac{1}{\zeta(2)}.$$

---

[1]This kind of technique is called **Möbius inversion**.