AN INFORMATION-THEORETIC PROOF OF THE ERDŐS-KAC THEOREM

Aidan Rocke
aidanrocke@gmail.com

February 24, 2022

ABSTRACT

In this article we show that the Erdős-Kac theorem has an elegant proof via Algorithmic Information Theory.

1 The Algorithmic Probability of a prime number

It is well-known that almost all integers are algorithmically random. Hence, given a prime \( p \in \mathbb{P} \) with binary encoding \( x_p \in \{0, 1\}^\infty \):

\[
K_U(x_p) \sim \log_2 p
\]

so the prime number \( p \) might as well be generated by \( \log_2 p \) tosses of a fair coin.

It follows that the Algorithmic Probability of observing the prime \( p \in \mathbb{P} \) is on the order of:

\[
m(x_p) \sim 2^{\log_2 p} \sim \frac{1}{p}
\]

A more general result, Levin’s Coding theorem, asserts that for any combinatorial object \( X \) with description \( K_U(X) \) [3]:

\[
-\log_2 m(X) \sim K_U(X)
\]

which is independent of the choice of description language \( U \) due to the Invariance theorem [4]:

\[
|K_U(X) - K_{U'}(X)| \leq \text{Cst}
\]

2 The Expected number of Unique Prime Divisors

For any integer \( X \sim U([1, N]) \) we may define its number of Unique Prime Divisors \( \omega(X) = \sum_{p \leq N} X_p \) where \( X_p = 1 \) if \( X \mod p = 0 \) and \( X_p = 0 \) otherwise. Thus, we may calculate the Expectation:

\[
\forall X \sim U([1, N]), \mathbb{E}[\omega(X)] = \sum_{p \leq N} 1 \cdot m(x_p) + 0 \cdot (1 - m(x_p)) = \sum_{p \leq N} m(x_p) \sim \sum_{p \leq N} \frac{1}{p} \sim \ln \ln N
\]

where we used Mertens’ Second theorem \( \sum_{p \leq N} \frac{1}{p} \sim \ln \ln N \).
3 The standard deviation of $\omega(X)$

As the $X_p$ are independent random variables the variance of $\omega(X)$ is linear in $X_p$:

$$\forall X \sim U([1, N]), \text{Var}[\omega(X)] = \sum_{p \leq N} \text{Var}[X_p] = \sum_{p \leq N} \mathbb{E}[X_p^2] - \mathbb{E}[X_p]^2 \sim \sum_{p \leq N} \frac{1}{p} - \frac{1}{p^2} \sim \ln \ln N \quad (6)$$

since $\sum_{p \leq N} \frac{1}{p^2} \leq \frac{\pi^2}{6}$.

4 The Erdős-Kac theorem

In order to prove the Erdős-Kac theorem, it remains to show that $\omega(X) = \sum_{p \leq N} X_p$ satisfies the Lindeberg condition for the Central Limit Theorem:

$$\Lambda_N(\epsilon) = \sum_{p \leq N} \left\langle \left( \frac{X_p}{\sqrt{\text{Var}[\omega(X)]}} \right)^2 : \left| \frac{X_p}{\sqrt{\text{Var}[\omega(X)]}} \right| \geq \epsilon \right\rangle$$

$$\forall \epsilon > 0, \lim_{N \to \infty} \Lambda_N(\epsilon) = 0 \quad (7)$$

where $\langle \alpha : \beta \rangle$ denotes the expectation value of $\alpha$ restricted to outcomes $\beta$.

Given that $\text{Var}[\omega(X)] \sim \sqrt{\ln \ln N}$ and $X_p \in \{0, 1\}$, we find that the Lindeberg condition is satisfied:

$$\forall \alpha \in (0, 1) \forall \epsilon \geq \frac{1}{\sqrt{\ln \ln N}}, \lim_{N \to \infty} \Lambda_N(\epsilon) = 0 \quad (9)$$

Hence, the Central Limit Theorem allows us to deduce that:

$$\forall X \sim U([1, N]), \frac{\omega(X) - \ln \ln N}{\sqrt{\ln \ln N}} \quad (10)$$

converges to the standard normal distribution $\mathcal{N}(0, 1)$ as $N \to \infty$. 

2
References

[1] Hardy, G. H.; Ramanujan, S. (1917), “The normal number of prime factors of a number n”, Quarterly Journal of Mathematics

[2] Marcus Hutter et al. (2007) Algorithmic probability. Scholarpedia, 2(8):2572.

[3] L.A. Levin. Laws of information conservation (non-growth) and aspects of the foundation of probability theory. Problems Information Transmission, 10(3):206-210, 1974. [4] Kolmogorov A.N. Three approaches to the quantitative definition of information. Problems Inform. Transmission, 1(1):1–7, 1965.

[5] Lindeberg, J. W. "Eine neue Herleitung des Exponential-gesetzes in der Wahrscheinlichkeitsrechnung." Math. Zeit. 15, 211-235, 1922.

[6] Erdős, Paul; Kac, Mark (1940). “The Gaussian Law of Errors in the Theory of Additive Number Theoretic Functions”. American Journal of Mathematics.

[7] Rényi, A.; Turán, P. (1958). “On a theorem of Erdős-Kac” . Acta Arithmetica.