Counting the Primes Using Entropy

Ioannis Kontoyiannis
Athens U of Economics & Business

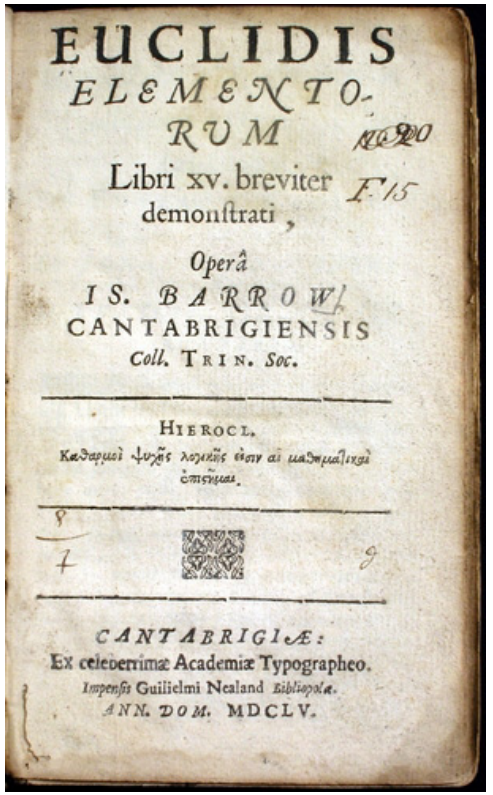
ITW 2008, Porto, Portugal

May 2008

The Primes: Pre-history

Theorem 0. [$<$ 300 BC]

There are infinitely many prime numbers


EUCLIDIS
ELEMENTO-
RUM
Libri xv. breviter demonſtrati,
Operâ
IS. BARROW
CANTABRIGIENSIS
Coll. TRIN. Soc.
HIEROCL.
CANTABRIGIÆ:
Ex celeberrimæ Academiæ Typographeo.
Impenſis Guilielmi Nealand Bibliopolæ.
ANN. DOM. MDCLV.

Let $\pi(n) = \#\{\text{primes } p \ : \ p \leq n\}$

Theorem 0 in modern notation:

$$\pi(n) \to \infty \qquad \text{as } n \to \infty$$

An Information-Theoretic Proof of Theorem 0

Proof. [*Following an idea of Chaitin* (1979)]

Let $N \sim U\{1, 2, \ldots, n\}$ and write it uniquely as

$$N \;=\; p_1^{X_1} \cdot p_2^{X_2} \cdot \cdots \cdot p_{\pi(n)}^{X_{\pi(n)}}$$

This *defines* the RVs X_i. Since $p_i^{X_i}$ divides N

$$2^{X_i} \le p_i^{X_i} \le N \le n \quad \Rightarrow \quad X_i \le \log n$$

Hence:

$$\begin{aligned} \log n &= H(N) \\ &= H(X_1, X_2, \ldots, X_{\pi(n)}) \\ &\le H(X_1) + H(X_2) + \cdots + H(X_{\pi(n)}) \\ &\le \pi(n) \log(\log n + 1) \end{aligned}$$

Therefore:

$$\boldsymbol{\pi(n) \ge \frac{\log n}{\log(\log n + 1)} \to \infty \quad \text{as } n \to \infty}$$

□

How quickly does $\pi(n)$ grow?

We saw $$\pi(n) \geq \frac{\log n}{\log(\log n + 1)}$$

Gauss in 1792 conjectured what has come to be known as the **Prime Number Theorem** (**PNT**)

$$\boldsymbol{\pi(n) \sim \frac{n}{\log_e n}} \quad \text{as } \boldsymbol{n \to \infty}$$

or, equivalently,

$$\boldsymbol{\pi(n) \sim \mathrm{Li}(n) \triangleq \int_2^n \frac{dt}{\log_e t} \sim \frac{n}{\log_e n}}$$

Incidently, the Riemann Hypothesis is *equivalent* to:

$$\boldsymbol{\pi(n) = \mathrm{Li}(n) + O(n^{1/2+\epsilon})}$$

A Second Information-Theoretic Proof of Theorem 0

Proof. [*Following an idea of Hardy & Wright* (1938)]

Let $N \sim U\{1, 2, \ldots, n\}$ and write it uniquely as

$$N \;=\; M^2 \cdot p_1^{Y_1} \cdot p_2^{Y_2} \cdot \cdots \cdot p_{\pi(n)}^{Y_{\pi(n)}}$$

where M is the largest integer s.t. M^2 divides N and the RVs Y_i are binary. Since M^2 divides N, $1 \leq M \leq \sqrt{n}$

Hence:
$$\begin{aligned} \log n &= H(N) \\ &= H(M, Y_1, Y_2, \ldots, Y_{\pi(n)}) \\ &\leq H(M) + H(Y_1) + H(Y_2) + \cdots + H(Y_{\pi(n)}) \\ &\leq \frac{1}{2}\log n + \pi(n) \end{aligned}$$

Therefore: $\boldsymbol{\pi(n) \geq \frac{1}{2}\log n}$ □

Chebyshev and the PNT

One of the major early players in proof attempts of the PNT was Chebyshev

He didn't prove it, but showed that eventually

$$A\frac{n}{\log_e n} \leq \pi(n) \leq B\frac{n}{\log_e n}$$

His proof was, in part, based on the following:

Theorem 1. [Chebyshev (1852)]

$$C(n) \triangleq \sum_{p \leq n} \frac{\log p}{p} \sim \log n \quad \text{as } n \to \infty$$

Classical Proofs of the PNT

1896

PNT finally proved by Hadamard and independently by de la Vallée-Pousin

Proofs were mathematically heavy, relying on Hadamard's theory of integral functions applied to the Riemann zeta function $\zeta(s)$

1921

G.H. Hardy:

"*if anyone produces an elementary proof of the PNT ... he will show that ... it is time for the books to be cast aside and for the theory to be rewritten*"

1948

Selberg and Erdös announce finding an elementary proof!

In their proof, Chebyshev's Theorem 1 was explicitly used

Information-Theoretic Proof of Theorem 1: Preliminaries

Let $N \sim U\{1, 2, \ldots, n\}$ and write it uniquely as

$$N \;=\; p_1^{X_1} \cdot p_2^{X_2} \cdot \cdots \cdot p_{\pi(n)}^{X_{\pi(n)}}$$

This *defines* the RVs X_i. What is their distribution?

$$\Pr\{X_i \geq k\} \;=\; \Pr\{N \text{ is a multiple of } p_i^k\} \;=\; \frac{1}{n}\left\lfloor \frac{n}{p_i^k} \right\rfloor \;\approx\; \left(\frac{1}{p_i}\right)^k$$

Similarly:

$$\begin{aligned}\Pr\{X_i \geq k \text{ and } X_j \geq \ell\} \;&=\; \Pr\{N \text{ is a multiple of } p_i^k p_j^\ell\} \\ &=\; \frac{1}{n}\left\lfloor \frac{n}{p_i^k p_j^\ell} \right\rfloor \;\approx\; \left(\frac{1}{p_i}\right)^k \left(\frac{1}{p_j}\right)^\ell\end{aligned}$$

Therefore: **The random variables $X_1, X_2, \ldots, X_{\pi(n)}$ are approximately independent Geom$(1/p_i)$**

IT Proof of Theorem 1: Billingsley's Heuristic

$$
\begin{aligned}
\log n &= H(N) \\
&= H(X_1, X_2, \ldots, X_{\pi(n)}) \\
&\overset{??}{\approx} \sum_{i=1}^{\pi(n)} H(X_i) \\
&\overset{??}{\approx} \sum_{i=1}^{\pi(n)} H(\mathsf{Geom}(1/p_i)) \\
&= \sum_{p \le n} \left[\frac{\log p}{p-1} - \log\left(1 - \frac{1}{p}\right) \right] \\
&\approx \sum_{p \le n} \left[\frac{\log p}{p} + \frac{1}{p} \right] \\
&\approx \sum_{p \le n} \frac{\log p}{p} \;=\; C(n)
\end{aligned}
$$

IT Proof of Theorem 1: Lower Bound

To show $\boldsymbol{\liminf_{n\to\infty} \frac{C(n)}{\log n} \geq 1}$

Let $h(\mu) =$ entropy of a Geom RV with mean μ

$$= (\mu+1)\log(\mu+1) - \mu \log \mu$$

Recall $h(\mu)$ is increasing in μ

Note

$$\mu_i \stackrel{\triangle}{=} E(X_i) = \sum_{k\geq 1} \Pr\{X_i \geq k\}$$

$$= \sum_{k\geq 1} \frac{1}{n} \left\lfloor \frac{n}{p_i^k} \right\rfloor$$

$$\leq \sum_{k\geq 1} \frac{1}{p_i^k}$$

$$= \frac{1}{p_i - 1}$$

IT Proof of Theorem 1: Lower Bound

$$
\begin{aligned}
\log n &= H(N) \\
&= H(X_1, X_2, \ldots, X_{\pi(n)}) \\
&\leq \sum_{\boldsymbol{i=1}}^{\boldsymbol{\pi(n)}} \boldsymbol{H(X_i)} \\
&\leq \sum_{\boldsymbol{i=1}}^{\boldsymbol{\pi(n)}} \boldsymbol{h(\mu_i)} \\
&\leq \sum_{\boldsymbol{i=1}}^{\boldsymbol{\pi(n)}} \boldsymbol{h\Big(\frac{1}{p_i - 1}\Big)} \\
&= \sum_{p \leq n} \Big[\frac{\log p}{p-1} - \log\Big(1 - \frac{1}{p}\Big)\Big] \\
&\leq (1 + \epsilon_n) \sum_{p \leq n} \frac{\log p}{p} \\
&= (1 + \epsilon_n) C(n)
\end{aligned}
$$

IT Proof of Theorem 1: Upper Bound

We just showed $\liminf_{n\to\infty} \frac{C(n)}{\log n} \geq 1$

To show $\limsup_{n\to\infty} \frac{C(n)}{\log n} \leq 1$

Note

$$\mu_i \stackrel{\triangle}{=} E(X_i) \;=\; \sum_{k\geq 1} \Pr\{X_i \geq k\} \;=\; \sum_{k\geq 1} \frac{1}{n}\left\lfloor \frac{n}{p_i^k} \right\rfloor \;\geq\; \frac{1}{n}\left\lfloor \frac{n}{p_i} \right\rfloor \;\geq\; \frac{1}{p_i} - \frac{1}{n}$$

Elementary Lemma. (Erdös 1930s)

$$\sum_{p\leq n} \log p \;\leq\; 2n$$

Proof. Simple trick involving binomial coefficients. □

IT Proof of Theorem 1: Upper Bound

Since

$$P(N) = \frac{1}{n} \leq \frac{1}{N}$$

we have

$$H(N) \;=\; E\Big[\log \frac{1}{P(N)}\Big] \;\geq\; E[\log N] \;=\; E\Big[\log \prod_{i=1}^{\pi(n)} p_i^{X_i}\Big] \;=\; \sum_{i=1}^{\pi(n)} E(X_i) \log p_i$$

IT Proof of Theorem 1: Upper Bound

Since

$$P(N) = \frac{1}{n} \leq \frac{1}{N}$$

we have

$$H(N) \;=\; E\Big[\log \frac{1}{P(N)}\Big] \;\geq\; E[\log N] \;=\; E\Big[\log \prod_{i=1}^{\pi(n)} p_i^{X_i}\Big] \;=\; \sum_{i=1}^{\pi(n)} E(X_i) \log p_i$$

hence

$$\log n \geq \sum_{p \leq n} \Big(\frac{1}{p} - \frac{1}{n}\Big) \log p = C(n) - \frac{1}{n} \sum_{p \leq n} \log p \geq C(n) - 2$$

i.e.

$$\frac{C(n)}{\log n} \leq 1 + \frac{2}{\log n}$$

□

Final Remarks

Counting-via-entropy gives rough bounds on $\pi(n)$

$$\pi(n) \geq \frac{1}{2} \log n \quad \text{for all } n \geq 2$$

Kolmogorov complexity arguments [Bergman, Tropp, Li & Vitanyi] show

$$\pi(n) \gtrsim \frac{n}{(\log n)^2} \quad \text{as } n \to \infty$$

Entropy-theoretic arguments give optimal results...

$$\sum_{p \leq n} \frac{\log p}{p} \sim \log n \quad \text{as } n \to \infty$$

... as well as finite-n bounds, e.g., for all $n \geq 16$

$$\frac{86}{125} \log n + \frac{47}{20} \leq \sum_{p \leq n} \frac{\log p}{p} \leq \log n + 2$$

$\rightsquigarrow$ **Information-theoretic proof of the PNT...?**

Proof of Erdös's Lemma, paraphrasing Hardy & Wright

Let $\vartheta(n) = \sum_{p \le n} \log p$, take $n \ge 2$ arbitrary, restrict attention to odd n. Every prime $n+1 < p \le 2n+1$ divides the binomial coefficient

$$B := \binom{2n+1}{n} = \frac{(2n+1)!}{n!(n+1)!},$$

since it divides the numerator but not the denominator, and hence the product of all these primes also divides B. In particular, their product must be $\le B$

$$\prod_{n+1<p\le 2n+1} p \le B = \frac{1}{2}\binom{2n+1}{n} + \frac{1}{2}\binom{2n+1}{n+1} \le \frac{1}{2}(1+1)^{2n+1} = 2^{2n},$$

and taking logarithms

$$\vartheta(2n+1) - \vartheta(n+1) = \sum_{n+1<p\le 2n+1} \log p = \log\Big[\prod_{n+1<p\le 2n+1} p\Big] \le 2n$$

Iterating this bound inductively gives the required result