m3pm16div.tex

**Handout: Prime divisor functions; Landau's Poisson extension to PNT; Probabilistic Number Theory**

**Prime Divisor Functions**

Recall the following arithmetic functions: $d(n) :=$ #divisors of $n$; $\omega(n) :=$ # distinct prime divisors $n$; $\Omega(n) :=$ # prime divisors $n$ (counted with multiplicity). So if $n = p_1^{r_1} ... p_k^{r_k}$, we have

$$d(n) = \prod_1^k (1 + r_j), \qquad \omega(n) = k, \qquad \Omega(n) = \sum_1^k r_j.$$

**Theorem**. (i) $\sum_{n \le x} \omega(n) = x \log\log x + C_1 x + O(1/\log x)$,
(ii) $\sum_{n \le x} \Omega(n) = x \log\log x + C_2 x + O(1/\log x)$,
where as above

$$C_1 = \gamma + \sum, \qquad C_2 = C_1 + \sum_p \frac{1}{p(p-1)} = C_1 + S,$$

say.
*Proof.* (i) $\sum_{n \le x} \omega(n)$ is the number of pairs $(p, n)$ with $p|n$ and $n \le x$. For fixed $p$, the number of such pairs is equal to the number of multiples $rp \le x$, i.e. $[x/p]$. So

$$\sum_{n \le x} \omega(n) = \sum_{p \le x} \left( \frac{x}{p} - \left\{ \frac{x}{p} \right\} \right).$$

By the first theorem of II.7,

$$\sum_{p \le x} \frac{x}{p} = x \sum_{p \le x} \frac{1}{p} = x \log\log x + C_1 x + O(x/\log x).$$

Also

$$0 \le \sum_{p \le x} \{x/p\} < \sum_{p \le x} 1 = \pi(x) = O(x/\log x),$$

by Chebyshev's Upper Estimate (III.2). Combining gives (i).
(ii) Similarly. //

*Note.* This theorem gives us:

$$\frac{1}{x} \sum_{n \le x} \omega(n) = \log\log x + C_1 + O(1/\log x);$$

$$\frac{1}{x}\sum_{n\le x}\Omega(n) = \log\log x + C_2 + O(1/\log x).$$

In Probabilistic Number Theory, one thinks of the LHS as the 'mean value' of $\omega, \Omega$. This says that both $\sim \log\log x$. This result, due to Hardy and Ramanujan (1917), corresponds to the Law of Large Numbers (LLN), thinking of $\omega, \Omega$ as random and divisibility by distinct primes as independent events. We prove the Prime Number Theorem (PNT), and one can prove an extension of it, counting $n \le x$ with $k$ prime factors. This corresponds to the relevant Central Limit Theorem (CLT), extending the Law of Large Numbers.

## LANDAU'S POISSON EXTENSION OF PNT: PRIMES PLAY A GAME OF CHANCE

**Theorem (LANDAU 1900**; Handbuch, 1909, 203-211). If $\pi_k(x)$ is the number of $n \le x$ with $k$ distinct prime factors ($k = 1, 2, \ldots$),

$$\pi_k(x) \sim \frac{x}{(k-1)!}.\frac{(\log\log x)^{k-1}}{\log x}.$$

**Lemma** (Handbuch, 203-5). For $F(u, x)$ $(2 \le u \le x)$ s.t.
(i) $F(u, x) \ge 0$;
(ii) for fixed $x > 2$ $F(u, x)/\log u$ decreases in $u$;
(iii) $F(2, x) = o(\int_2^x F(u, x)du/\log u)$ – then

$$\sum_{p\le x} F(p, x) \sim \int_2^x \frac{F(u, x)}{\log u}du.$$

*Proof.* By PNT, $\theta(x) \sim x$, so $\theta(x) = x + x\epsilon(x)$, $\epsilon(x) = o(1)$. So

$$\sum_{p\le x} F(p, x) = \sum_{n=2}^{x} \frac{\theta(n) - \theta(n-1)}{\log n} F(n, x) \qquad \text{(definition of } \theta)$$

$$= \sum_2^x \frac{F(n, x)}{\log n} + \sum_2^{x-1} n\epsilon(n)\Big[\frac{F(n, x)}{\log n} - \frac{F(n+1, x)}{\log(n+1)}\Big] + \frac{F(2, x)}{\log 2} + [x]\epsilon([x])\frac{F([x], x)}{\log[x]}, \qquad (i)$$

by Abel summation. As in the Integral Test (I.4),

$$\sum_2^x \frac{F(n, x)}{\log n} + \frac{F(2, x)}{\log 2} = (1 + o(1))\int_2^x \frac{F(u, x)}{\log u}du.$$

Choose $\epsilon > 0$ arbitrarily small; there exists $U = U(\epsilon)$ with $|\epsilon(u)| < \epsilon$ for $u > U$. So for $x > U + 1$, the sum of the remaining terms on the RHS of (i) is

$$|\sum_{2}^{n-1} n\epsilon(n)\Big[\frac{F(n,x)}{\log n} - \frac{F(n+1,x)}{\log(n+1)}\Big] + [x]\epsilon([x])\frac{F([x],x)}{\log[x]}|$$

$$< O(F(2,x)) + \epsilon \sum_{U}^{n-1} [...] + \epsilon[x]F([x],x)/\log[x]$$

$$= \epsilon \sum_{U}^{x} \frac{F(n,x)}{\log n} + O(F(2,x)) \qquad \text{(by Abel summation again)}$$

$$= \epsilon \int_2^x \frac{F(u,x)}{\log u} du + o(\int_2^x \frac{F(u,x)}{\log u} du).$$

This holds for all $\epsilon > 0$, so LHS $= o(\int_2^x F(u,x)du/\log u)$.
So LHS of (i) is $\sum_{p \le x} F(p,x) = (1 + o(1)) \int_2^x F(u,x)du/\log u$. //

*Proof of the Theorem.* We prove the case $k = 2$ (Handbuch, 205-8):

$$\pi_2(x) \sim x \log\log x/\log x.$$

The general case follows by a similar but more complicated argument (Handbuch, 208-11), or by induction on $k$, an argument due to Wright (HW §22.18, Th. 437, 368-71; J, 140-5).

$$\begin{aligned} \pi_2(x) &:= \#\{n \le x : n \text{ has 2 distinct prime factors}\} \\ &= \frac{1}{2}\#\{(p,q) : p, q \text{ distinct primes}, pq \le x\} \end{aligned}$$

($\frac{1}{2}$ because of $(p,q)$ and $(q,p)$). But $\sum_{p \le x} \pi(x/p)$ is the number of pairs with $p \ne q$, $\pi(\sqrt{x})$ the number of pairs with $p = q$. So by above

$$2\pi(x) = \sum_{p \le x} \pi(x/p) - \pi(\sqrt{x}) = \sum_{p \le x} \pi(x/p) + O(\sqrt{x}/\log x),$$

by PNT or Chebyshev's Upper Estimate. We use the Lemma with

$$F(p,x) := \pi(x/p).$$

For, conditions (i), (ii) are clear. As $\pi(\frac{1}{2}x) \sim \frac{1}{2}x/\log\frac{1}{2}x \sim \frac{1}{2}x/\log x$, (iii) will follow from the relation $(*)$ below:

$$\int_2^x \frac{\pi(x/u)}{\log u}du \sim \frac{2x\log\log x}{\log x}. \tag{$*$}$$

To prove $(*)$:

$$\int_2^x \frac{\pi(x/u)}{\log u}du = \int_2^{x/2} \frac{\pi(x/u)}{\log u}du \qquad (\text{if } u > x/2,\ x/u < 2, \text{ so } \pi(x/u) = 0)$$

$$= x\int_2^{x/2} \frac{\pi(v)}{\log x - \log v}\frac{dv}{v^2} \qquad (v := x/u\text{: } 2 \le u = x/v \le x/2,\ 2 \le v \le x/2).$$

Choose $\epsilon > 0$; for $v \ge V = V(\epsilon)$,

$$\Big|\pi(v) - \frac{v}{\log v}\Big| < \epsilon\frac{v}{\log v}$$

by PNT. So for $x > 2V$,

$$\Big|\int_V^{x/2} \frac{\pi(v)}{\log x - \log v}\frac{dv}{v^2} - \int_V^{x/2} \frac{v/\log v}{\log x - \log v}\frac{dv}{v^2}\Big| < \epsilon\int_V^{x/2} \frac{v/\log v}{\log x - \log v}\frac{dv}{v^2},$$

so $|\int_2^{x/2} ... - \int_2^{x/2} ...| < \epsilon\int_2^{x/2} ... + O(1/\log x)$, as

$$\int_2^V \frac{v/\log v}{\log x - \log v}\frac{dv}{v^2} = O(1/\log x),$$

etc. Since

$$\int_2^{x/2} \frac{v/\log v}{\log x - \log v}\frac{dv}{v^2} = \int_{\log 2}^{\log x - \log 2} \frac{dw}{w(\log x - w)} \qquad (w := \log v)$$

$$= \frac{1}{\log x}\int_{\log 2}^{\log x - \log 2}\Big(\frac{1}{w} + \frac{1}{\log x - w}\Big)dw \qquad (\text{partial fractions})$$

$$= \frac{1}{\log x}(\log(\log x - \log 2) - \log\log 2 - \log\log 2 + \log(\log x - \log 2)) \sim \frac{2\log\log x}{\log x},$$

this gives $(*)$.

By the Lemma, $\pi_2(x) \sim x\log\log x/\log x$, proving the case $k = 2$. //

As each $n$ has at least one prime factor, it is better to work with $k+1$ rather than $k$. Writing

$$\lambda := \log \log x$$

(so $\lambda \to \infty$ as $x \to \infty$ – though *extremely slowly*):

$$\frac{1}{x} \pi_{k+1}(x) \sim \frac{(\log \log x)^k}{k! \log x} = \frac{e^{-\lambda} \lambda^k}{k!} \quad (k = 0, 1, 2, \ldots) \qquad (\lambda, x \to \infty).$$

Now $\{e^{-\lambda} \lambda^k / k! : k = 0, 1, 2, \ldots\}$ forms the *Poisson distribution* $P(\lambda)$ of Probability Theory, with parameter $\lambda$ (mean $\lambda$, variance $\lambda$). So:

**Theorem (Landau)**. The *proportion* of primes $\leq x$ with $k+1$ distinct prime factors is asymptotically Poisson distributed with parameter $\lambda := \log \log x$.

The Poisson distribution is "the signature of randomness", in the *discrete* setting (as here). So this suggests that, *in some sense, the primes are randomly distributed* (hence 'Primes play a game of chance' – see below). This is very surprising: in the ordinary sense, nothing could be less random, or more deterministic, or "God-given", than the primes.

Recall the *prime divisor functions* $\omega(n)$ is the number of *distinct* prime divisors of $n$, $\Omega(n)$ is the number of prime divisors of $n$ counted with multiplicity. As before, $\omega$ and $\Omega$ behave similarly here. So we may rephrase Landau's theorem above as saying that both proportions $\omega(n)/n$, $\Omega(n)/n$ are asymptotically Poisson distributed with parameter $\lambda := \log \log n$. Using $X \sim F$ as the usual probabilistic shorthand for "the random variable $X$ has the distribution (function) $F$", we have

**Theorem (Landau's Poisson PNT,** 1900).

$$\omega(n)/n \sim P(\log \log n), \qquad \Omega(n)/n \sim P(\log \log n).$$

With some loss of information (the constants $C_1, C_2$ and the error terms $O(x/\log x)$, we may summarise the results above for comparison. Using $\sim$ now (with a number after it, not a distribution) to denote "is asymptotic to", one has

**Theorem (Hardy and Ramanujan**, 1917).

$$\omega(n)/n \sim \log \log n, \qquad \Omega(n) \sim \log \log n.$$

*Probabilistic Number Theory.* The situation above is sumarised in:
*Kac's Dictum* (Mark KAC (1914-84)): Primes play a game of chance;
*Vaughan's Dictum* (R. C. VAUGHAN (1945-)): It's obvious that the primes are randomly distributed – it's just that we don't know what that means yet[1].

A short calculation (involving the probability generating function) shows that if $X_1, \dots, X_n$ are independent random variables, Poisson $P(1)$ distributed, then their sum is Poisson $P(n)$ (means add; variances add over independent summands; the point is that Poissonianity is preserved). This means that as $n \to \infty$, the *Central Limit Theorem* (CLT) applies: the sum is asymptotically *normally distributed*, or *Gaussian*:

$$P(n) \sim N(n, n)$$

(this statement can be made precise – it is all we need here). Since by Landau's Theorem of III.9 we know that $\omega(n)/n, \Omega(n)/n \sim P(\log \log n)$, this gives

$$\omega(n)/n, \Omega(n)/n \sim N(\log \log n, \log \log n).$$

This is the *Erdös-Kac Central Limit Theorem* of 1939 (Paul ERDÖS (1913-96)).
*Note.* 1. The Erdös-Kac CLT was completed in 1939, during a seminar given by Kac. Erdös was in the audience, and completed the proof by using *sieve methods*, during the talk.
2. This result marks the "official birth" of the subject of Probabilistic Number Theory, though as above the subject really goes back to Landau in 1900 and to Hardy and Ramanujan in 1917. For a textbook account, see e.g. G. TENENBAUM[2]: *Introduction to analytic and probabilistic number theory*, CUP, 1995.
3. Just as the Landau and Erdös-Kac results correspond to the CLT ("Law of Errors"), the Hardy-Ramanujan result corresponds to the LLN ((Weak) Law of Large Numbers – "Law of Averages"). For background, see e.g. my homepage, link to Stochastic Processes (II.7 L12).

[1] My wife's instant response to this:
*(Cecilie) Bingham's Dictum*: Primes play a game of chance – we just don't know the rules yet.

[2] Through my 1986 paper with Tenenbaum, I have my Erdös number of 2 (Erdös had Erdös number 0, his collaborators Erdös number 1, etc.)