Supplement 3

Asymptotics for the mean values of some arithmetic functions

The following two result have already appeared in this course, we include their proof here for completeness.

Theorem 1: For $x \geq 2$,

$$L(x) := \sum_{n \leq x} \log n = x \log x - x + O(\log x).$$

Theorem 2: For $x \geq 1$,

$$H(x) := \sum_{n \leq x} \frac{1}{n} = \log x + \gamma + O(1/x)$$

where $\gamma := \int_1^\infty \frac{1}{\lfloor t \rfloor} - \frac{1}{t} \, dt \quad \left(= \lim_{x \to \infty} \left(\sum_{n \leq x} \frac{1}{n} - \log x \right) \right).$

Proof of Theorem 1: By the integral test we see that

$$\int_1^N \log t \, dt \leq L(N) \leq \int_1^{N+1} \log t \, dt$$

for any $N \in \mathbb{N}$. Since $\frac{d}{dx}(x \log x - x) = \log x$, the result follows easily. $\square$

Proof of Theorem 2:

$$\sum_{n \leq x} \frac{1}{n} = \log x + \underbrace{\int_1^\infty \frac{1}{\lfloor t \rfloor} - \frac{1}{t} \, dt}_{=: \gamma} - \underbrace{\int_x^\infty \frac{1}{\lfloor t \rfloor} - \frac{1}{t} \, dt}_{\ll \frac{1}{x} \text{ since } \lfloor t \rfloor \geq \frac{t}{2}.} \qquad \square$$

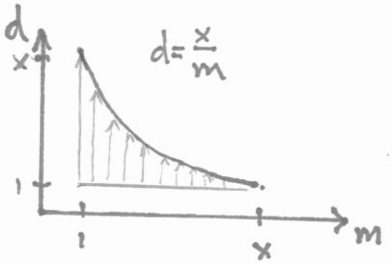
The Dirichlet Hyperbola Method (divisor switching)

Recall that the divisor function $\tau(n)$ counts the number of positive divisors of n, this is clearly a very irregular function as $n \to \infty$. The question of determining the average size of $\tau(n)$ was known as the "divisor problem" and was answered by Dirichlet.

Theorem 3: For $x \geq 1$

$$D(x) := \sum_{n \leq x} \tau(n) = x \log x + (2\gamma - 1)x + O(\sqrt{x}).$$

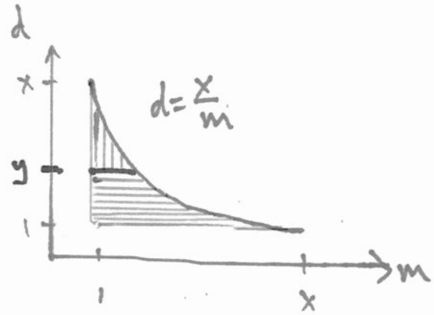
Proof: It is easy to see that



$$\sum_{n \leq x} \tau(n) = \sum_{md \leq x} 1 = \sum_{m \leq x} \sum_{d \leq x/m} 1 \qquad = \frac{x}{m} + O(1)$$

$$\leq x \log x + O(x).$$

To do better than this we employ what is known as the "Dirichlet hyperbola method" (or divisor switching): For a given $1 \leq y \leq x$ (to be determined) we write



$$\sum_{n \leq x} \tau(n) = \sum_{\substack{md \leq x \\ d \leq y}} 1 + \sum_{\substack{md \leq x \\ d > y}} 1.$$

It follows from Theorem 2 that

$$\sum_{\substack{md\le x\\ d\le y}} 1 = \sum_{d\le y}\sum_{m\le x/d} 1 = \sum_{d\le y}\frac{x}{d} + O(y)$$

$$= x\log y + \gamma x + O\left(\frac{x}{y}\right) + O(y).$$

To estimate the second sum we switch the order of m & d,

$$\sum_{\substack{md\le x\\ d>y}} 1 = \sum_{m\le x/y}\ \sum_{y<d\le x/m} 1 = \sum_{m\le x/y}\left(\frac{x}{m}-y\right) + O\left(\frac{x}{y}\right)$$

$$= x\log\left(\frac{x}{y}\right) + \gamma x - x + O(y) + O\left(\frac{x}{y}\right).$$

Combining these two estimates and choosing the optimal value $y = \sqrt{x}$ gives the result. □

Corollary 1 (of Theorems 1 & 3): For $x \geq 1$,

$$\Delta(x) := \sum_{n\le x}(\log n - \tau(n) + 2\gamma) = O(\sqrt{x}).$$

Proof: Immediate.

Theorem 4 (General Hyperbola Method) Let f & g be arithmetic functions with respective summatory functions $F(x) = \sum_{n\le x} f(n)$ & $G(x) = \sum_{n\le x} g(n)$, for any $1\le y\le x$:

$$\sum_{n\le x} f*g(n) := \sum_{md\le x} f(m)g(d) = \sum_{d\le y} g(d)F\left(\frac{x}{d}\right) + \sum_{m\le x/y} f(m)G\left(\frac{x}{m}\right) - F\left(\frac{x}{y}\right)G(y).$$