14. Average order of magnitude

Another way to describe a function is to consider its average order, that is, the quantity
$$\frac{1}{n}\sum_{m=1}^{n} f(m).$$
Taking the average tends to smooth out the irregularities and it is quite often possible to approximate the average quite accurately.

The following is an example of how to switch the order of summation, even if it does not often give good estimates.

Proposition 14.1. *If*
$$F(m) = \sum_{d|n} f(d)$$

then
$$\sum_{m=1}^{n} F(m) = \sum_{m=1}^{n} \lfloor\frac{n}{m}\rfloor f(m).$$

Proof. By definition of $F(m)$, we have
$$\sum_{m=1}^{n} F(m) = \sum_{m=1}^{n}\sum_{d|m} f(d).$$

For this order of summation we associate to every value of m, $1 \le m \le n$, all its divisors d. Instead to each value of d, $1 \le d \le n$, one can associate all its multiples kd, $1 \le kd \le n$, so that k takes any value up to $\lfloor n/d \rfloor$. Thus
$$\sum_{m=1}^{n} F(m) = \sum_{d=1}^{n}\sum_{k=1}^{\lfloor n/d \rfloor} f(d)$$
$$= \sum_{d=1}^{n} f(d) \sum_{k=1}^{\lfloor n/d \rfloor} 1$$
$$= \sum_{d=1}^{n} \lfloor\frac{n}{d}\rfloor f(d). \qquad \square$$

If one approximates $\lfloor n/m \rfloor$ by $n/m = O(1)$ one gets
$$\sum_{m=1}^{n} \tau(m) = n\log n + O(n) \quad \text{and} \quad \sum_{m=1}^{n} \sigma(m) = O(n^2).$$

This is not a particularly good estimate and one can do better than this:

Theorem 14.2.

$$\sum_{m=1}^{n} \tau(m) = n \log n + (2\gamma - 1)n + O(n^{1/2}),$$

where γ is Euler's constant.

Proof. By (14.1) we have

$$\sum_{m=1}^{n} \tau(m) = \sum_{m=1}^{n} \lfloor \frac{n}{m} \rfloor.$$

Geometrically, the sum on the RHS is the number of *lattice points* (x, y) (that is, points such that x and y are integers) where x and y are natural numbers, on or below the hyperbola $xy = n$, since if we fix x the number of $1 \leq y \leq n/x$ is precisely $\lfloor n/x \rfloor$.

By symmetry the number of lattice points (x, y) with $x > 0$, $y > 0$ and $xy \leq n$ is equal to twice the number of lattice points $x > 0$, $y > x$ and $xy \leq n$ plus the number of lattice points $x > 0$, $x = y$ and $xy \leq n$.

Hence

$$\sum_{m=1}^{n} \tau(m) = 2 \left(\sum_{x=1}^{\sqrt{n}} \lfloor \frac{n}{x} \rfloor - x \right) + \lfloor \sqrt{n} \rfloor$$

$$= 2n \left(\sum_{x=1}^{\sqrt{n}} \frac{1}{x} \right) + O(\sqrt{n}) - 2 \frac{\lfloor \sqrt{n} \rfloor (\lfloor \sqrt{n} \rfloor + 1)}{2}$$

$$= 2n(\log(\sqrt{n}) + \gamma + O(1/\sqrt{n})) - n + O(\sqrt{n})$$

$$= n \log n + (2\gamma - 1)n + O(\sqrt{n}). \qquad \square$$

One can actually improve the error term $O(\sqrt{n})$. It is known that it can be improved to $O(\sqrt[3]{n})$ but not to $O(\sqrt[4]{n})$. Actually it is conjectured that the error is always smaller than $O(x^{1/4+\epsilon})$ for any $\epsilon > 0$. The problem of finding the best estimate is called the *Dirichlet divisor problem*.

The problem of estimating the number of lattice points inside the circle $x^2 + y^2 \leq n$ is very similar and was considered by Gauss; it is called the *circle problem*. The number of such points is $\pi n + O(n^{1/2})$ and one would like to improve the error as in the Dirichlet divisor problem.

Note the difference between estimating an upper bound for the τ-function versus approximating its average value. Even though it always beats any power of the logarithm, (13.1), on average it is $\log n$.

We now turn to estimating the average value of the φ-function. We will need:

Lemma 14.3. *For $s > 1$*

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}.$$

Proof. Both the series above and the series for the zeta function converge absolutely for $s > 1$, so that we are free to multiply the series together and rearrange the sum in any convenient order.

We have

$$\sum_{m=1}^{\infty} \frac{1}{m^s} \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \sum_{m,n=1}^{\infty} \frac{\mu(n)}{(mn)^s}$$

$$= \sum_{t=1}^{\infty} \frac{1}{t^s} \sum_{d|t} \mu(d)$$

$$= \sum_{t=1}^{\infty} \frac{1}{t^s} M(t)$$

$$= 1. \qquad \square$$

We will also need an upper bound for $\zeta(2)$. In fact

$$\zeta(2) = \frac{\pi^2}{6}.$$

It is easy to see that $\zeta(2) < 2$ as

$$\zeta(2) < 1 + \int_1^{\infty} \frac{dt}{t^2}$$

$$= 2.$$

Theorem 14.4.

$$\sum_{m=1}^{n} \varphi(m) = \frac{3n^2}{\pi^2} + O(n \log n).$$

Proof. Recall that

$$\varphi(m) = m \sum_{d|m} \frac{\mu(d)}{d}.$$

On the other hand

$$\lfloor x \rfloor^2 = x^2 + O(x) \qquad \text{and} \qquad \mu(m) = O(1).$$

3

Therefore

$$\sum_{m=1}^{n} \varphi(m) = \sum_{m=1}^{n} m \sum_{d|m} \frac{\mu(d)}{d}$$

$$= \sum_{d_1 d_2 \leq n} d_2 \mu(d_1)$$

$$= \sum_{d_1=1}^{n} \mu(d_1) \sum_{d_2=1}^{n/d} d_2$$

$$= \sum_{d=1}^{n} \mu(d) \frac{\lfloor n/d \rfloor^2 + \lfloor n/d \rfloor}{2}$$

$$= \frac{1}{2} \sum_{d=1}^{n} \mu(d) \frac{n^2}{d^2} + O\left(\sum_{d=1}^{n} \frac{n}{d}\right)$$

$$= \frac{n^2}{2} \left(\sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} - \sum_{d=n+1}^{\infty} \frac{\mu(d)}{d^2}\right) + O(n \log n)$$

$$= \frac{n^2}{2} \frac{1}{\zeta(2)} + O\left(n^2 \sum_{d=n+1}^{\infty} \frac{1}{d^2}\right) + O(n \log n)$$

$$= \frac{3n^2}{\pi^2} + O(n) + O(n \log n)$$

$$= \frac{3n^2}{\pi^2} + O(n \log n). \qquad \Box$$

Since

$$\sum_{m=1}^{n} m \sim \frac{1}{2} n^2$$

one might say that the average value of $\varphi(n)$ is

$$\frac{6n}{\pi^2} \approx 0.608n.$$