Devoir en temps libre

À rendre au plus tard le 07 mai 2024

On s'intéresse dans ce devoir au nombre de diviseurs premiers d'un entier n "typique". En cheminant à travers des résultats de Hardy et Ramanujan (1917), Turàn (1934), puis Erdös et Kac (1940), nous allons voir qu'un entier n possède génériquement $\log\log(n)$ facteurs premiers, avec des fluctuations gaussiennes autour de ce nombre de l'ordre de $\sqrt{\log\log(n)}$.

Notation : dans toute la suite, pour simplifier les expressions, la lettre p désignera un nombre premier, ainsi par $\sum_{p\leq n}$ ou $\prod_{p\leq n}$ on entend la somme ou le produit sur les nombres premiers p plus petits que n.

I. Estimés élémentaires sur les nombres premiers

On fixe un entier n grand et on choisit un nombre N selon la loi uniforme $\mathbb{P}_n$ dans l'ensemble $[|1,n|] = \{1,2,\ldots,n\}$, on note $\mathbb{E}_n$ et var_n la variance et l'espérance sous la loi $\mathbb{P}_n$. Via la décomposition unique en facteurs premiers, on écrit alors

$$N = \prod_{p\leq n} p^{\alpha_{n,p}}$$

où les valuations $p-$adiques $\alpha_{n,p}$ sont des variables aléatoires à valeurs entières.

1. Montrer que lorsque n tend vers l'infini, on a $\theta(n) := \sum_{p\leq n}\log(p) = O(n)$.
   Indice : on pourra remarquer que $\exp(\theta(2n)-\theta(n)) = \prod_{n<p\leq 2n} p \leq \binom{2n}{n} \leq 4^n$.
2. Montrer que pour p premier fixé, la variable aléatoire $\alpha_{n,p}$ vérifie pour $k\geq 0$
$$\mathbb{P}_n\left(\alpha_{n,p}\geq k\right) = \frac{1}{n}\left\lfloor \frac{n}{p^k}\right\rfloor.$$
3. Montrer que pour p premier fixé, on a ainsi l'encadrement
$$\mathbb{E}_n\left[\alpha_{n,p}\right] \in \left[\frac{1}{p}-\frac{1}{n},\, \frac{1}{p-1}\right].$$
4. Déduire de la question précédente et de l'écriture de N en produit de facteurs premiers que lorsque n tend vers l'infini, on a
$$\sum_{p\leq n}\frac{\log(p)}{p} = \log(n) + O(1).$$
5. Déduire de la question précédente que lorsque n tend vers l'infini, on a également
$$\sum_{p\leq n}\frac{1}{p} = \log\log(n) + O(1).$$
   Indice : on pourra par exemple écrire $\dfrac{1}{p} = \dfrac{\log p}{p}\dfrac{1}{\log p} = \dfrac{\log p}{p}\displaystyle\int_1^\infty \mathbb{1}_{t>p}\frac{dt}{t\log^2 t}$.

II. Nombre de facteurs premiers

Dans la suite, étant donné un entier $n \geq 1$, on note $\omega(n)$ son nombre de facteurs premiers

$$\omega(n) := \sum_{p \leq n} \mathbb{1}_{p \,|\, n}.$$

L'objet de cette section est de montrer le théorème suivant.

Théorème (Hardy–Ramanujan, 1917)**.** *Soit $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ une fonction tendant vers l'infini en l'infini et ce de façon arbitrairement lente. Alors, lorsque n tend vers l'infini, on a*

$$\mathit{Card}\left\{x \in [|1, n|],\, |\omega(x) - \log\log x| > \psi(n)\sqrt{\log\log n}\right\} = o(n).$$

Avec les notations de la section précédente, où l'on rappelle que N suit la loi uniforme dans $[|1, n|]$, le théorème ci-dessus s'écrit encore

$$\lim_{n \to +\infty} \mathbb{P}_n\left(|\omega(N) - \log\log N| > \psi(n)\sqrt{\log\log n}\right) = 0.$$

Autrement dit et de façon informelle, avec une probabilité qui tend vers un, un entier typique x possède $\log\log x + O(\sqrt{\log\log x})$ facteurs premiers. La stratégie de preuve proposée ci-après est due à Turàn en 1934 et a signé à cette époque la naissance de la théorie probabiliste des nombres.

1. Montrer que, pour tout $C > 0$, lorsque n tend vers l'infini, on a

$$\mathbb{P}_n\left(\log\log(n) - \log\log(N) > C)\right] = o(1).$$

2. Montrer que lorsque n tend vers l'infini, on a

$$\mathbb{E}_n\left[\omega(N)\right] = \log\log(n) + O(1).$$

3. De la même façon, montrer que lorsque n tend vers l'infini, on a

$$\mathrm{var}_n(\omega(N)) := \mathbb{E}_n\left[\omega(N)^2\right] - \mathbb{E}_n\left[\omega(N)\right]^2 = O(\log\log(n)).$$

4. En déduire que lorsque n tend vers l'infini

$$\mathbb{P}_n\left(\left|\omega(N) - \mathbb{E}_n\left[\omega(N)\right]\right| > \psi(n)\sqrt{\log\log n}\right) = o(1).$$

5. Conclure.

III. Études des fluctuations

Dans cette dernière partie, on se propose de montrer le théorème ci-dessous qui renforce le théorème de Hardy–Ramanujan, en précisant les fluctuations de la variable aléatoire $\omega(N)$ autour de sa moyenne.

Théorème (Erdös–Kac, 1940)**.** *Lorsque n tend vers l'infini, on a pour tout $t \in \mathbb{R}$*

$$\lim_{n \to +\infty} \mathbb{P}_n \left(\frac{|\omega(N) - \log \log N|}{\sqrt{\log \log N}} > t \right) = \int_t^{+\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}} du.$$

Autrement dit, pour un entier typique x, le nombre de facteurs premiers est $\log \log(x)$ avec des fluctuations gaussiennes de l'ordre de $\sqrt{\log \log(x)}$.

La stratégie de la preuve que nous mettons en oeuvre ci-dessous consiste à comparer la suite des variables $(\mathbb{1}_{p|N})_p$ vues sous la loi $\mathbb{P}_n$, à une suite de variables de Bernoulli indépendantes, pour laquelle le théorème limite central usuel s'applique. Sur un espace de probabilité abstrait $(\Omega, \mathcal{F}, \mathbb{P})$, on considère ainsi une suite de variables indépendantes (X_p), indéxée par les nombre premiers, et telles que

$$\mathbb{P}(X_p = 1) = \frac{1}{p}, \quad \mathbb{P}(X_p = 0) = 1 - \frac{1}{p}, \quad \text{de sorte que} \quad \mathbb{E}[X_p] = \frac{1}{p}, \quad \mathrm{var}(X_p) = \frac{1}{p}\left(1 - \frac{1}{p}\right).$$

On fixe alors $\varepsilon \in]0, 1[$ et on pose

$$S_n := \sum_{p \le n^\varepsilon} X_p, \quad \mu_n := \mathbb{E}[S_n] = \sum_{p \le n^\varepsilon} \frac{1}{p}, \quad \sigma_n^2 = \mathrm{var}(S_n) = \sum_{p \le n^\varepsilon} \frac{1}{p}\left(1 - \frac{1}{p}\right).$$

On définit enfin

$$\widetilde{\omega}(N) := \sum_{p \le n^\varepsilon} \mathbb{1}_{p|N}.$$

1. Justifier que sous $\mathbb{P}_n$, la variable $\omega(N) - \widetilde{\omega}(N)$ est bornée, uniformément en n, de sorte que le théorème d'Erdös–Kac pour $\omega(N)$ est équivalent au même énoncé pour $\widetilde{\omega}(N)$.
2. Montrer que lorsque n tend vers l'infini, on a

$$\mathbb{E}_n\left[\widetilde{\omega}(N)\right] = \log \log(n) + O(1), \qquad \mathrm{var}_n(\widetilde{\omega}(N)) := \log \log(n) + O(1).$$

3. Montrer que la suite S_n vérifie le théorème limite central suivant, pour $t \in \mathbb{R}$

$$\lim_{n \to +\infty} \mathbb{P}\left(\frac{S_n - \mu_n}{\sigma_n} > t \right) = \int_t^{+\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}} du.$$

   Indice : on pourra appliquer le TLC pour des variables indépendantes vérifiant la condition de Lindeberg.

4. Pour $r \ge 1$ entier, montrer que

$$\lim_{n \to +\infty} \left(\mathbb{E}_n\left[\left(\frac{\widetilde{\omega}(N) - \mathbb{E}_n[\widetilde{\omega}(N)]}{\sqrt{\mathrm{var}_n(\widetilde{\omega}(N))}} \right)^r \right] - \mathbb{E}\left[\left(\frac{S_n - \mu_n}{\sigma_n} \right)^r \right] \right) = 0.$$

5. Conclure en utilisant le fait que la loi gaussienne est caractérisée par ses moments et que la convergence en loi vers une gaussienne est équivalente à la convergence des moments.