Le théorème de Hardy-Ramanujan

Cortex

April 26, 2026

Présentation Le théorème de Hardy-Ramanujan est un résultat qui affirme que, en moyenne, le nombre de diviseur d'un entier entre 1 et n est $\log\log(n)$. Le théorème va plus loin en affirmant que la variance est "à peine" plus élevée que $\sqrt{\log\log(n)}$.

Remarques préliminaires Le résultat peut sembler insurmontable, mais en réalité, il n'est pas très difficile à prouver, grâce à une utilisation très judicieuse des probabilités. Il faut bien comprendre ce que dit le théorème et ne pas hésiter à l'illustrer graphiquement.

Notations

- Dans la suite, je note $\mathcal{P}$ l'ensemble des nombres premiers et $\sum_{p\leq n} := \sum_{p\in\mathcal{P},\ p\leq n}$.
- Soient $(u_n)_n, (v_n)_n$ deux suites, je note

$$\exists C > 0, \forall n \geq 0,\ \ |u_n| \leq C\,|v_n| \ \ \Leftrightarrow\ \ u_n = \mathcal{O}(v_n)$$

Résultats admis

- On admet le deuxième théorème de Mertens :

$$h(n) := \sum_{p\leq n} \frac{1}{p} = C + \log\log(n) + \mathcal{O}\left(\frac{1}{\log(n)}\right)$$

  Avec $C \in \mathbf{R}$ une constante.

Développement

Pour la suite, pour $n \geq 1$ un entier, notons $\nu(n) := \sum_{p|n} 1$ le nombre de diviseur premier de n. L'objectif est de démontrer que pour tout $\omega : \mathbf{N}_{>0} \to \mathbf{R}$ une suite tendant vers $+\infty$ (que l'on fixe pour la suite) et $X_n \sim U(1,n)$ une loi uniforme, on a

$$\mathbf{P}\left(|\nu(X_n) - \log\log(n)| > \omega(n)\sqrt{\log\log(n)}\right) = o(1)$$

Affirmant qu'à mesure que n devient grand, le part des entiers vérifiant l'inégalité devient infime et par conséquent, que les valeurs de ν se concentre, pour l'essentiel, autour de $\log\log$.

Pour démontrer ce résultat, il faut constater la similitude qu'elle possède avec l'inégalité de Bienaymé-Tchebychev. Si Y désigne une loi de probabilité avec $\mathbf{E}(Y) = \log\log(n)$ et $\mathbf{Var}(Y) = \log\log(n)$ alors on a

$$\mathbf{P}\left(|Y - \log\log(n)| > \omega(n)\sqrt{\log\log(n)}\right) \leq \frac{1}{\omega(n)^2}$$

L'idée est simplement de prouver que $\mathbf{E}(\nu(X_n)), \mathbf{Var}(\nu(X_n)) = \log\log(n) + \mathcal{O}(1)$ en utilisant le résultat admis.

Résultat préliminaire Dans la suite, on fixe $\alpha \in]0,1]$ et $\nu_\alpha(n) := \sum_{p \leq n^\alpha} \mathbf{1}_{p|n}$. L'introduction de α sera justifiée ultérieurement. On démontrera plutôt $\mathbf{E}(\nu_\alpha(X_n)), \mathbf{Var}(\nu_\alpha(X_n)) = \log\log(n) + \mathcal{O}(1)$ en prenant soin de prouver qu'il n'y a pas de différence considérable entre ν et ν_α.

D'une part, pour $n \geq 1$, on a $\nu(n) \geq \nu_\alpha(n)$. D'autre part, le nombre de diviseur premier $p > n^\alpha$ ne peut excéder $\left\lfloor \frac{1}{\alpha} \right\rfloor$. En effet :

$$n^{\alpha(\nu(n) - \nu_\alpha(n))} \leq \prod_{p > n^\alpha,\ p|n} p \leq n$$

Ainsi $0 \leq \nu(n) - \nu_\alpha(n) \leq \left\lfloor \frac{1}{\alpha} \right\rfloor$.

Espérance Pour $d \geq 1$, on a $\mathbf{P}(d|X_n) = \frac{\lfloor \frac{n}{d} \rfloor}{n} = \frac{1}{d} + \mathcal{O}\left(\frac{1}{n}\right)$. De ce fait, il vient

$$\mathbf{E}(\nu_\alpha(X_n)) = \sum_{p \leq n^\alpha} \mathbf{P}(p|X_n) = \sum_{p \leq n^\alpha} \left\{\frac{1}{p} + \mathcal{O}\left(\frac{1}{n}\right)\right\} = h(n^\alpha) + \mathcal{O}(n^{\alpha-1}) = \log\log(n) + \mathcal{O}(1)$$

Variance Pour la variance, on utilise la formule de la variance impliquant les covariances et le fait que pour $p, q \in \mathcal{P}$ distincts, on a $\mathbf{1}_{pq|X_n} = \mathbf{1}_{p|X_n}\mathbf{1}_{q|X_n}$. Ainsi

$$\mathbf{Var}(\nu_\alpha(X_n)) = \sum_{p \leq n^\alpha} \mathbf{P}(p|X_n) - \mathbf{P}(p|X_n)^2 + \sum_{\substack{p,q \leq n^\alpha \\ p \neq q}} \mathbf{Covar}(\mathbf{1}_{p|X_n}, \mathbf{1}_{q|X_n})$$

Pour $p, q \in \mathcal{P}$, $p, q \leq n^\alpha$ et $p \neq q$, on a

$$\mathbf{Covar}(\mathbf{1}_{p|X_n}, \mathbf{1}_{q|X_n}) = \mathbf{P}(pq|X_n) - \mathbf{P}(p|X_n)\mathbf{P}(q|X_n) = \mathcal{O}\left(\frac{1}{n}\right)$$

Ainsi on obtient

$$\mathbf{Var}(\nu_\alpha(X_n)) = h(n^\alpha) + \mathcal{O}(n^{\alpha-1}) - \sum_{p \leq n^\alpha} \frac{1}{p^2} + \mathcal{O}(n^{2\alpha-1}) = \log\log(n) + \mathcal{O}(n^{2\alpha-1}) + \mathcal{O}(1)$$

On constate alors qu'avec $\alpha = 1$, on n'aurait pas eu le résultat attendu. Pour poursuivre, on prend $\alpha < 1/2$.

Conclusion Il ne reste plus qu'à rapporter les résultats précédents dans la formule. En premier lieu, on a

$$\begin{aligned}|\nu(X_n) - \log\log(n)| &\leq |\nu_\alpha(X_n) - \mathbf{E}(\nu_\alpha(X_n))| + |\nu(X_n) - \nu_\alpha(X_n)| + |\log\log(n) - \mathbf{E}(\nu_\alpha(X_n))| \\ &\leq |\nu_\alpha(X_n) - \mathbf{E}(\nu_\alpha(X_n))| + C_\alpha\end{aligned}$$

Où $C_\alpha = |\log\log(n) - \mathbf{E}(\nu_\alpha(X_n))| + \left\lfloor \frac{1}{\alpha} \right\rfloor = \mathcal{O}(1)$, qui n'est pas une variable aléatoire, ce qui a son importance pour appliquer l'inégalité de Bienaymé-Tchebychev. Ainsi, on en déduit :

$$\begin{aligned}&\mathbf{P}\left(|\nu(X_n) - \log\log(n)| > \omega(n)\sqrt{\log\log(n)}\right) \\ \leq &\mathbf{P}\left(|\nu_\alpha(X_n) - \mathbf{E}(\nu_\alpha(X_n))| + C_\alpha > \omega(n)\sqrt{\log\log(n)}\right) \\ \leq &\frac{\mathbf{Var}(\nu_\alpha(X_n))^2}{(\omega(n)\sqrt{\log\log(n)} - C_\alpha)^2} \sim \frac{1}{\omega(n)^2} = o(1)\end{aligned}$$

Remarques a posteriori

- Il existe version plus forte que le théorème de Hardy-Ramanujan : c'est le théorème de Erdös-Kac, qui affirme que pour $n \geq 1$ et $X \sim U(1,n)$, alors pour tout $t \in \mathbf{R}$,

$$\lim_{n\to+\infty} \mathbf{P}\left(\frac{\nu(X) - \log\log(n)}{\sqrt{\log\log(n)}} < t\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{\frac{t^2}{2}} dt$$

- Le résultat $\mathbf{E}(\nu(X_n)) \sim \log\log(n)$ implique immédiatement que

$$\frac{\nu(1) + \cdots + \nu(n)}{n} \sim \log\log(n).$$

  On dit que $\log\log(n)$ est un ordre moyen de $\nu(n)$.

- Dans la littérature, la fonction ν est plutôt noté ω.

- On pourrait, pour faire bonne mesure, préciser à chaque fois en quelles variables dépendent les $\mathcal{O}$.

- Il faut bien comprendre l'avant-dernière inégalité, on utilise le fait que si $A \leq B$ alors $A > C$ implique que $B > C$ et donc $\mathbf{P}(A > C) \leq \mathbf{P}(B > C)$.

Références Il s'agit là d'un développement un peu maison. La référence est *Additive Combinatorics* de Terrence Tao et Van H. Vu aux pages 7-9 dans la version *Cambridge Studies in Advanced Mathematics* (trouvable sur internet).Le livre est très peu répandu.

Recasage Pour ce développement, on peut partir pour 121, 224, 261 (pour l'utilisation d'une inégalité de concentration) et 264.

Il est risqué de le placer dans la 262 (à moins de se tourner vers Erdös-Kac) car la limite illustre une limite de valeur et non de probabilité. D'ailleurs Hardy-Ramanujan peut se réécrire sans probabilité :

$$\frac{1}{n}\#\{|\nu(k) - \log\log(n)| > \omega(n)\sqrt{\log\log(n)},\ k \in \{1,\ldots,n\}\} = o(1)$$

De même pour la 266 car on utilise jamais de l'indépendance entre les variables. Je ne pense pas que l'utilisation des covariances ou l'utilisation de $\mathbf{1}_{pq|X} = \mathbf{1}_{p|X}\mathbf{1}_{q|X}$ justifie un tel recasage.