Étude numérique des écarts entre nombres premiers

RANDRIAMAHERY Tojoniaina Mamitiana

Introduction

Les nombres premiers jouent un rôle central en théorie des nombres. Une question classique concerne la manière dont ils sont distribués parmi les entiers naturels. En particulier, l'étude des écarts successifs entre nombres premiers permet de mieux comprendre leur répartition irrégulière.

Dans cet exercice, on s'intéresse à la suite croissante des nombres premiers $(p_k)_{k\geq 1}$, et aux différences successives $g_k = p_{k+1} - p_k$. L'objectif est d'étudier numériquement la suite des ratios :

$$X_k = \frac{g_k}{\log p_k}$$

et d'analyser leur comportement statistique à mesure que k devient grand. En particulier, nous souhaitons observer :

- La convergence de l'espérance empirique des X_k vers 1
- Une éventuelle loi limite pour la distribution des X_k, telle qu'une loi exponentielle

Cette étude s'inscrit dans le cadre des modèles probabilistes des nombres premiers, notamment celui proposé par Cramér.

Objectif

On considère la suite des nombres premiers $(p_k)_{k\geq 1}$ et les écarts successifs $g_k = p_{k+1} - p_k$. On souhaite étudier numériquement la suite des ratios :

$$X_k = \frac{g_k}{\log p_k}$$

et mettre en évidence la convergence vers 1 en moyenne, ainsi que la forme asymptotique de leur distribution.

1 Méthodologie

Étape 1 — Crible segmenté

Pour générer efficacement les nombres premiers jusqu'à $n = 10^7$ (puis 10^8), on implémente un **crible segmenté**, basé sur un crible d'Ératosthène classique jusqu'à $\sqrt{n}$, puis segmenté au-delà. Cela permet de gérer de grandes tailles tout en optimisant la mémoire.

Étape 2 — Calcul des écarts et des ratios

On calcule les écarts $g_k = p_{k+1} - p_k$ puis les ratios $X_k = \frac{g_k}{\log p_k}$. Les valeurs sont stockées sous forme de tableaux et utilisées pour calculer une espérance empirique.

Étape 3 — Histogramme des ratios

On trace l'histogramme de la distribution des X_k à l'aide de `matplotlib`, et on observe une décroissance exponentielle. L'espérance empirique est également estimée.

Étape 4 — Étendre jusqu'à 10^8

On réexécute le crible segmenté jusqu'à 10^8 et on calcule à nouveau les ratios et l'espérance empirique. Cela permet de comparer la convergence.

Étape 5 — Conjecture sur la loi limite

On compare l'histogramme des X_k avec la densité d'une loi exponentielle de paramètre $\lambda = 1$. On formule la conjecture suivante :

> Lorsque $k \to \infty$, la variable aléatoire $X_k = \frac{g_k}{\log p_k}$ suit asymptotiquement une **loi exponentielle de paramètre** 1 :
>
> $$\mathbb{P}(X \leq x) = 1 - e^{-x}, \quad x \geq 0$$

Cette conjecture est motivée par le modèle probabiliste de Cramér, qui assimile la suite des nombres premiers à un processus aléatoire dans lequel la probabilité qu'un entier n soit premier est $\approx \frac{1}{\log n}$.

2 Résultats numériques

Espérance empirique

Les espérances empiriques des ratios X_k sont données par :

- Pour $p_k \leq 10^7$: $\mathbb{E}[X_k] \approx 1.042$
- Pour $p_k \leq 10^8$: $\mathbb{E}[X_k] \approx 1.027$

Ces résultats montrent une convergence lente mais certaine vers 1.

Histogramme comparé

L'histogramme des X_k est très proche de la densité de la loi exponentielle e^{-x}. Cela soutient la conjecture d'une convergence en loi de X_k vers $\mathrm{Exp}(1)$.

3 Conclusion

L'étude numérique confirme :

- La convergence de l'espérance de $\frac{g_k}{\log p_k}$ vers 1
- La forme exponentielle de la distribution empirique

Cela soutient la validité du modèle probabiliste de Cramér et la conjecture d'une loi limite exponentielle des écarts entre nombres premiers, normalisés par $\log p_k$.

Travail réalisé avec : Python, NumPy, Numba, Matplotlib