

Densités

Marc Lorenzi

5 mai 2024

1 Probabilités

1.1 Quelques rappels

Nous allons dans cet article travailler exclusivement dans l'univers $\mathbb{N}^*$ des entiers naturels non nuls. Une *tribu* (sur $\mathbb{N}^*$) est un ensemble non vide $\mathcal{T}$ de parties de $\mathbb{N}^*$ stable par passage au complémentaire et par union dénombrable. On montre facilement qu'une tribu est aussi stable par union finie, par différence et par intersection finie ou dénombrable, et contient $\emptyset$ et $\mathbb{N}^*$. Par exemple, l'ensemble $\mathcal{P}(\mathbb{N}^*)$ des parties de $\mathbb{N}^*$ est une tribu. Une *probabilité* est une fonction $\mathbb{P} : \mathcal{T} \longrightarrow [0, 1]$, où $\mathcal{T}$ est une tribu, vérifiant

- $\mathbb{P}(\mathbb{N}^*) = 1$.
- Pour toute famille dénombrable $(A_i)_{i \in I}$ d'éléments disjoints de $\mathcal{T}$, la famille $(\mathbb{P}(A_i))_{i \in I}$ est sommable et

$$\mathbb{P}\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mathbb{P}(A_i)$$

Le triplet $(\mathbb{N}^*, \mathcal{T}, \mathbb{P})$ est un *espace probabilisé*.

1.2 Une tentative

Lorsque nous choisissons un espace probabilisé $(\mathbb{N}^*, \mathcal{T}, \mathbb{P})$, nous aimerions que certaines parties simples de $\mathbb{N}^*$ appartiennent à la tribu $\mathcal{T}$ et aient une certaine probabilité. Par exemple, soit $a \in \mathbb{N}^*$. Si je tire un entier au hasard, quelle est la probabilité qu'il soit multiple de a ? La réponse qui vient à l'esprit est $1/a$. Posons plus précisément la question. Nous aimerions que pour tout $a \in \mathbb{N}^*$, $a\mathbb{N}^*$ appartienne à $\mathcal{T}$ et $\mathbb{P}(a\mathbb{N}^*) = 1/a$. En fait, un tel espace probabilisé n'existe pas.

Lemme 1. *Soit $\mathcal{T}$ une tribu sur $\mathbb{N}^*$ contenant les ensembles $a\mathbb{N}^*$ où $a \in \mathbb{N}^*$. Alors, $\mathcal{T} = \mathcal{P}(\mathbb{N}^*)$.*

Démonstration. Soit $n \in \mathbb{N}^*$. Écrivons

$$n = \prod_{p \in \mathcal{P}} p^{\alpha_p}$$

où $\mathcal{P}$ est l'ensemble des nombres premiers et les α_p sont des entiers presque tous nuls. Par l'existence et l'unicité de la décomposition en produit de nombres premiers, n est l'unique entier tel que pour tout $p \in \mathcal{P}$, p^{α_p} divise n et p^{α_p+1} ne divise pas n . On a donc

$$\{n\} = \bigcap_{p \in \mathcal{P}} (p^{\alpha_p} \mathbb{N}^* \setminus p^{\alpha_p+1} \mathbb{N}^*)$$

Une tribu est stable par différence et par intersection finie ou dénombrable, donc $\{n\} \in \mathcal{T}$. De là, comme une tribu est stable par union finie ou dénombrable, $\mathcal{T}$ contient toutes les parties de $\mathbb{N}^*$. $\square$

Lemme 2. Soit $(\mathbb{N}^*, \mathcal{T}, \mathbb{P})$ un espace probabilisé tel que pour tout $a \in \mathbb{N}^*$, $a\mathbb{N}^* \in \mathcal{T}$ et $\mathbb{P}(a\mathbb{N}^*) = 1/a$. Alors, pour tout $n \in \mathbb{N}^*$, $\mathbb{P}(\{n\}) = 0$.

Démonstration. Soient $a, b \in \mathbb{N}^*$ premiers entre-eux. On a

$$a\mathbb{N}^* \cap b\mathbb{N}^* = ab\mathbb{N}^*$$

et donc

$$\mathbb{P}(a\mathbb{N}^* \cap b\mathbb{N}^*) = \mathbb{P}(ab\mathbb{N}^*) = \frac{1}{ab} = \mathbb{P}(a\mathbb{N}^*)\mathbb{P}(b\mathbb{N}^*)$$

Les événements $a\mathbb{N}^*$ et $b\mathbb{N}^*$ sont donc indépendants. De là, les événements $\mathbb{N}_a = \mathbb{N}^* \setminus a\mathbb{N}^*$ et $\mathbb{N}_b = \mathbb{N}^* \setminus b\mathbb{N}^*$ sont aussi indépendants et donc

$$\mathbb{P}(\mathbb{N}_a \cap \mathbb{N}_b) = \mathbb{P}(\mathbb{N}_a)\mathbb{P}(\mathbb{N}_b) = \left(1 - \frac{1}{a}\right) \left(1 - \frac{1}{b}\right)$$

Une récurrence immédiate montre que ce résultat s'étend à un nombre quelconque d'entiers premiers entre eux deux à deux.

Soit $n \in \mathbb{N}^*$. Soit $N \in \mathbb{N}^*$ tel que $n < N$. Pour tout entier premier $p > n$, n n'est pas multiple de p et donc $n \in \mathbb{N}_p$. De là,

$$\{n\} \subseteq \bigcap_{n < p \leq N} \mathbb{N}_p$$

Des nombres premiers distincts sont premiers entre eux deux à deux, donc

$$\mathbb{P}(\{n\}) \leq \prod_{n < p \leq N} \mathbb{P}(\mathbb{N}_p) = \prod_{n < p \leq N} \left(1 - \frac{1}{p}\right)$$

Par un résultat classique, le produit ci-dessus tend vers 0 lorsque N tend vers l'infini et donc $\mathbb{P}(\{n\}) = 0$. $\square$

Proposition 3. Il n'existe pas d'espace probabilisé $(\mathbb{N}^*, \mathcal{T}, \mathbb{P})$ tel que pour tout $a \in \mathbb{N}^*$, $a\mathbb{N}^* \in \mathcal{T}$ et $\mathbb{P}(a\mathbb{N}^*) = 1/a$.

Démonstration. Supposons le contraire. Par les deux lemmes précédents, les singletons appartiennent à $\mathcal{T}$ et sont de probabilité nulle. On a

$$\mathbb{N}^* = \bigcup_{n=1}^{\infty} \{n\}$$

Cette union est disjointe, donc

$$\mathbb{P}(\mathbb{N}^*) = \sum_{n=1}^{\infty} \mathbb{P}(\{n\}) = 0$$

Contradiction, puisque $\mathbb{P}(\mathbb{N}^*) = 1$. $\square$

Il n'est donc pas possible de définir une probabilité sur $\mathbb{N}^*$ qui réponde à certaines de nos intuitions. Nous sommes pourtant persuadés que notre question de départ est pertinentes et que notre réponse l'est tout autant. Nous allons voir qu'en remplaçant le concept de probabilité par un autre concept, celui de *densité*, nos intuitions seront effectivement les bonnes.

1.3 Probabilités sur $\mathbb{N}^*$

Nous allons nous intéresser dans ce qui suit uniquement à des espaces probabilisés $(\mathbb{N}^*, \mathcal{T}, \mathbb{P})$ où $\mathcal{T} = \mathcal{P}(\mathbb{N}^*)$. Une probabilité $\mathbb{P} : \mathcal{P}(\mathbb{N}^*) \rightarrow [0, 1]$ est entièrement caractérisée par ses valeurs sur les singletons. Définir une telle probabilité revient à se donner une suite $\mu = (\mu_n)_{n \in \mathbb{N}^*}$ de réels positifs telle que la série $\sum_{n \geq 1} \mu_n$ soit convergente, de somme $S > 0$, et de poser pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}_\mu(\{n\}) = \frac{\mu_n}{S}$$

On a alors pour tout ensemble $E \subseteq \mathbb{N}^*$,

$$\mathbb{P}_\mu(E) = \frac{1}{S} \sum_{k \in E} \mu_k = \frac{1}{S} \sum_{k=1}^{\infty} \mu_k \mathbb{1}_E(k)$$

où $\mathbb{1}_E$ est la *fonction indicatrice* de E . Écrivons les choses un peu différemment. Pour tout $n \in \mathbb{N}^*$, posons

$$S_n(E, \mu) = \sum_{k=1}^n \mu_k \mathbb{1}_E(k)$$

En particulier, $S_n(\mathbb{N}^*, \mu) = \sum_{k=1}^n \mu_k$ est non nulle si n est assez grand (car cette somme tend vers S). Posons également

$$\Delta_n(E, \mu) = \frac{S_n(E, \mu)}{S_n(\mathbb{N}^*, \mu)}$$

On a alors

$$\mathbb{P}_\mu(E) = \lim_{n \rightarrow \infty} \Delta_n(E, \mu)$$

Quel est l'intérêt d'écrire $\mathbb{P}_\mu$ sous cette forme ? C'est que nous pouvons maintenant envisager de considérer des suites μ dont la série associée est *divergente*.

2 Probabilités et Densités

2.1 Densité

Soit μ une suite de réels positifs telle que

$$\sum_{n=1}^{\infty} \mu_n = +\infty$$

Définition 1. Soit $E \subseteq \mathbb{N}^*$. L'ensemble E possède une *densité* relativement à μ si la suite de terme général $\Delta_n(E, \mu)$ est convergente. La densité de E relativement à μ est alors

$$\mathbb{D}_\mu(E) = \lim_{n \rightarrow \infty} \Delta_n(E, \mu)$$

Remarquons que $\Delta_n(E, \mu) \in [0, 1]$. La densité de l'ensemble E , si elle existe, est donc un réel appartenant à $[0, 1]$. Nous noterons $\mathcal{D}_\mu$ l'ensemble des parties de $\mathbb{N}^*$ admettant une densité relativement à μ . On a donc

$$\mathbb{D}_\mu : \mathcal{D}_\mu \rightarrow [0, 1]$$

Deux suites μ nous intéresseront plus particulièrement :

- La suite $\text{NAT} = (1)_{n \geq 1}$. On parle alors de la *densité naturelle* de E . Dans ce cas,

$$\Delta_n(E, \text{NAT}) = \frac{1}{n} \sum_{k=1}^n \mathbb{1}_E(k) = \frac{|E \cap \llbracket 1, n \rrbracket|}{n}$$

- La suite $\text{LOG} = (1/n)_{n \geq 1}$. On parle alors de la *densité logarithmique* de E . On a dans ce cas

$$\Delta_n(E, \text{LOG}) = \frac{1}{H_n} \sum_{k=1}^n \frac{\mathbb{1}_E(k)}{k}$$

où H_n est le *nième nombre harmonique*. Comme $H_n \sim \ln n$, il revient au même de poser

$$\Delta_n(E, \text{LOG}) = \frac{1}{\ln n} \sum_{k=1}^n \frac{\mathbb{1}_E(k)}{k}$$

2.2 Un calcul simple de densité

Soit $a \in \mathbb{N}^*$. Considérons l'ensemble $E = a\mathbb{N}^*$. Calculons la densité naturelle et la densité logarithmique de E . Pour tous $n, k \in \mathbb{N}^*$, $ka \leq n$ si et seulement si $k \leq \lfloor n/a \rfloor$. Ainsi,

$$E \cap \llbracket 1, n \rrbracket = \{ka : k \in \llbracket 1, \lfloor n/a \rfloor \rrbracket\}$$

On a donc

$$\Delta_n(E, \text{NAT}) = \frac{1}{n} \left\lfloor \frac{n}{a} \right\rfloor$$

Remarquons l'encadrement

$$\frac{1}{a} \leq \Delta_n(E, \text{NAT}) < \frac{1}{a} + \frac{1}{n}$$

Nous en déduisons que lorsque n tend vers l'infini, $\Delta_n(E, \text{NAT})$ tend vers $1/a$. Ainsi, $a\mathbb{N}^* \in \mathcal{D}_{\text{NAT}}$ et $\mathbb{D}_{\text{NAT}}(a\mathbb{N}^*) = \frac{1}{a}$.

Passons à la densité logarithmique. Pour tout $n \geq 2$,

$$\Delta_n(E, \text{LOG}) = \frac{1}{H_n} \sum_{k=1}^{\lfloor n/a \rfloor} \frac{1}{ka} = \frac{1}{a} \frac{H_{\lfloor n/a \rfloor}}{H_n}$$

Or,

$$H_{\lfloor n/a \rfloor} \sim \ln \left\lfloor \frac{n}{a} \right\rfloor \sim \ln n$$

donc $\Delta_n(E, \text{LOG})$ tend vers $1/a$ lorsque n tend vers l'infini. Ainsi, $a\mathbb{N}^* \in \mathcal{D}_{\text{LOG}}$ et $\mathbb{D}_{\text{LOG}}(a\mathbb{N}^*) = \frac{1}{a}$.

2.3 Propriétés des densités

Les densités possèdent un certain nombre de propriétés analogues aux probabilités. Dans ce qui suit, μ est une suite de réels positifs telle que la série $\sum_{n \geq 1} \mu_n$ diverge. Pour tout entier n assez grand, la somme $\sum_{k=1}^n \mu_k$ est strictement positive, il est donc possible de diviser par cette somme. Pour alléger la rédaction, nous nous permettrons d'écrire « pour tout $n \geq 1$ », alors qu'il conviendrait de rajouter « assez grand ».

Proposition 4. [ENSEMBLES FINIS]

Pour toute partie finie E de $\mathbb{N}^*$, $E \in \mathcal{D}_\mu$ et $\mathbb{D}_\mu(E) = 0$.

Démonstration. Soit E une partie finie de $\mathbb{N}^*$. Il existe un entier $N \geq 1$ tel que $E \subseteq \llbracket 1, N \rrbracket$. Pour tout $n \geq N$,

$$S_n(E, \mu) = S_N(E, \mu)$$

ne dépend pas de n . On a donc

$$\Delta_n(E, \mu) = \frac{S_N(E, \mu)}{\sum_{k=1}^n \mu_k}$$

Comme la série $\sum_{k \geq 1} \mu_k$ diverge vers $+\infty$, cette quantité tend vers 0 lorsque n tend vers l'infini. $\square$

Proposition 5. [DENSITÉ DE L'UNIVERS]

$\mathbb{N}^* \in \mathcal{D}_\mu$ et $\mathbb{D}_\mu(\mathbb{N}^*) = 1$.

Démonstration. Pour tout $n \in \mathbb{N}^*$, $\Delta_n(\mathbb{N}^*, \mu) = 1$. $\square$

Proposition 6. [CROISSANCE]

Soient $A, B \in \mathcal{D}_\mu$ tels que $A \subseteq B$. On a alors $\mathbb{D}_\mu(A) \leq \mathbb{D}_\mu(B)$.

Démonstration. Pour tout $n \geq 1$, $\Delta_n(A, \mu) \leq \Delta_n(B, \mu)$ d'où le résultat en passant à la limite. $\square$

Proposition 7. [UNION DISJOINTE]

Soient $A, B \in \mathcal{D}_\mu$ deux ensembles disjoints. Alors, $A \cup B \in \mathcal{D}_\mu$ et

$$\mathbb{D}_\mu(A \cup B) = \mathbb{D}_\mu(A) + \mathbb{D}_\mu(B)$$

Démonstration. Comme A et B sont disjoints,

$$\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B$$

De là, facilement, pour tout $n \geq 1$,

$$\Delta_n(A \cup B, \mu) = \Delta_n(A, \mu) + \Delta_n(B, \mu)$$

d'où le résultat en faisant tendre n vers l'infini. $\square$

Par une récurrence immédiate, $\mathcal{D}_\mu$ est stable par union finie d'ensembles disjoints. En revanche, il n'est pas stable par union infinie dénombrable d'ensembles disjoints. En effet, les singletons appartiennent à $\mathcal{D}_\mu$ et sont disjoints deux à deux. Nous verrons bientôt qu'il existe des parties E de $\mathbb{N}^*$ qui n'appartiennent pas à $\mathcal{D}_\mu$. Un tel ensemble E est pourtant une union disjointe dénombrable de singletons.

Notation. Pour toute partie A de $\mathbb{N}^*$, notons $\bar{A} = \mathbb{N}^* \setminus A$.

Proposition 8. Pour tout ensemble $A \in \mathcal{D}_\mu$, $\bar{A} \in \mathcal{D}_\mu$ et

$$\mathbb{D}_\mu(\bar{A}) = 1 - \mathbb{D}_\mu(A)$$

Démonstration. Soit $A \in \mathcal{D}_\mu$. On a

$$\mathbb{1}_{\bar{A}} = \mathbb{1}_{\mathbb{N}^*} - \mathbb{1}_A$$

Pour tout $n \geq 1$,

$$\begin{aligned} S_n(\overline{A}, \mu) &= \sum_{k=1}^n \mu_k(1 - \mathbb{1}_A(k)) \\ &= S_n(\mathbb{N}^*, \mu) - S_n(A, \mu) \end{aligned}$$

De là,

$$\Delta_n(\overline{A}, \mu) = 1 - \Delta_n(A, \mu)$$

et cette quantité tend vers $1 - \mathbb{D}_\mu(A)$ lorsque n tend vers l'infini. $\square$

L'ensemble $\mathcal{D}_\mu$ est-il stable par réunion ou intersection de deux ensembles non nécessairement disjoints ? La réponse est non. Nous allons voir un contre-exemple dans le cas de la densité naturelle.

Proposition 9. $\mathcal{D}_{\text{NAT}}$ n'est pas stable par intersection finie.

Démonstration. Soit $A = 2\mathbb{N}^*$ l'ensemble des entiers pairs. Soit

$$E = \bigcup_{k=0}^{\infty} \llbracket 2^{2k}, 2^{2k+1} - 1 \rrbracket$$

Posons $B = (A \cap E) \cup (\overline{A} \cap \overline{E})$. L'ensemble A appartient à $\mathcal{D}_{\text{NAT}}$ et $\mathbb{D}_{\text{NAT}}(A) = 1/2$. On peut montrer (un peu péniblement) qu'il en est de même pour B . En revanche, $A \cap B = A \cap E$. Nous montrerons dans le paragraphe suivant que E n'admet pas de densité naturelle. Des calculs analogues à ceux que nous ferons un peu plus loin permettent de montrer que $A \cap E$ n'appartient pas non plus à $\mathcal{D}_{\text{NAT}}$. $\square$

Proposition 10. $\mathcal{D}_{\text{NAT}}$ n'est pas stable par réunion finie.

Démonstration. Reprenons les ensembles A et B de la preuve précédente. Par la proposition 8, $\overline{A}$ et $\overline{B}$ appartiennent à $\mathcal{D}_{\text{NAT}}$. En revanche, toujours par la proposition 8,

$$\overline{A} \cup \overline{B} = \overline{A \cap B}$$

n'appartient pas à $\mathcal{D}_{\text{NAT}}$, sinon son complémentaire $A \cap B$ serait aussi dans $\mathcal{D}_{\text{NAT}}$. $\square$

2.4 Ensembles n'ayant pas de densité naturelle

Nous allons voir qu'il existe des parties de $\mathbb{N}^*$ qui n'admettent pas de densité. À titre d'échauffement, commençons par la densité naturelle.

Considérons l'ensemble

$$E = \bigcup_{q=0}^{\infty} \llbracket 2^{2q}, 2^{2q+1} - 1 \rrbracket$$

E est l'ensemble des entiers dont le nombre de chiffres en base 2 est impair.

Proposition 11. $E \notin \mathcal{D}_{\text{NAT}}$.

Démonstration. Soit $n \in \mathbb{N}^*$. Il existe un unique entier $m \in \mathbb{N}$ tel que

$$2^m \leq n < 2^{m+1}$$

Supposons pour commencer que $m = 2p + 1$ est impair. On a

$$E \cap \llbracket 1, n \rrbracket = \bigcup_{q=0}^p \llbracket 2^{2q}, 2^{2q+1} - 1 \rrbracket$$

La réunion est disjointe donc

$$\begin{aligned} |E \cap \llbracket 1, n \rrbracket| &= \sum_{q=0}^p |\llbracket 2^{2q}, 2^{2q+1} - 1 \rrbracket| \\ &= \sum_{q=0}^p 4^q \\ &= \frac{1}{3}(2^{2p+2} - 1) \end{aligned}$$

Prenons $n = 2^{2p+1}$. On a

$$\Delta_n(E, \text{NAT}) = \frac{2}{3} \left(1 - \frac{1}{2^{2p+2}} \right)$$

qui tend vers $2/3$ lorsque p tend vers l'infini. Si l'ensemble E possède une densité naturelle, alors $\mathbb{D}_{\text{NAT}}(E) = 2/3$.

Prenons maintenant $n = 2^{2p+2} - 1$. On a

$$\Delta_n(E, \text{NAT}) = \frac{1}{3}$$

qui tend vers $1/3$ lorsque p tend vers l'infini. Si l'ensemble E possède une densité naturelle, alors $\mathbb{D}_{\text{NAT}}(E) = 1/3$. Contradiction. $\square$

Remarque. Si $m = 2p$ est pair alors

$$E \cap \llbracket 1, n \rrbracket = \bigcup_{q=0}^{p-1} \llbracket 2^{2q}, 2^{2q+1} - 1 \rrbracket \cup \llbracket 2^{2p}, n \rrbracket$$

On a donc

$$\begin{aligned} |E \cap \llbracket 1, n \rrbracket| &= \sum_{q=0}^{p-1} |\llbracket 2^{2q}, 2^{2q+1} - 1 \rrbracket| + |\llbracket 2^{2p}, n \rrbracket| \\ &= \sum_{q=0}^{p-1} 4^q + n - 4^p + 1 \\ &= \frac{1}{3}(2^{2p} - 1) + n - 2^{2p} + 1 \\ &= n - \frac{2}{3}(2^{2p} - 1) \end{aligned}$$

La figure ci-dessous représente $\Delta_n(E, \text{NAT})$ en fonction de n .

Les valeurs $1/3$ et $2/3$ que nous avons trouvées sont remarquables.

Proposition 12.

$$\liminf_{n \rightarrow \infty} \Delta_n(E, \text{NAT}) = \frac{1}{3} \quad \text{et} \quad \limsup_{n \rightarrow \infty} \Delta_n(E, \text{NAT}) = \frac{2}{3}$$

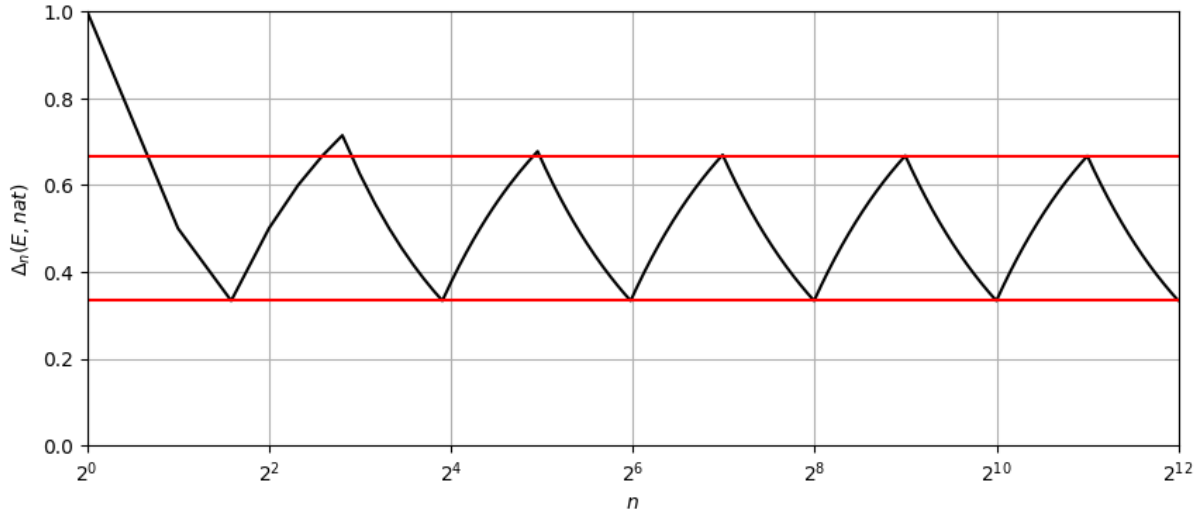


FIGURE 1 – La suite $(\Delta_n(E, \text{NAT}))_{n \geq 1}$

Démonstration. Reprenons les notations de la preuve précédente. Si $m = 2p + 1$ est impair,

$$S_n(E, \text{NAT}) = \frac{1}{3}(2^{2p+2} - 1)$$

Or,

$$2^{2p+1} \leq n \leq 2^{2p+2} - 1$$

donc

$$\frac{1}{3} \leq \Delta_n(E, \text{NAT}) \leq \frac{2}{3} \left(1 - \frac{1}{2^{2p+2}}\right) < \frac{2}{3}$$

Si $m = 2p$ est pair,

$$S_n(E, \text{NAT}) = n - \frac{2}{3}(2^{2p} - 1)$$

donc

$$\Delta_n(E, \text{NAT}) = 1 - \frac{2}{3n}(2^{2p} - 1)$$

Or,

$$2^{2p} \leq n \leq 2^{2p+1} - 1$$

De là,

$$\frac{2}{3n}(2^{2p} - 1) \leq \frac{2(2^{2p} - 1)}{3 \times 2^{2p}} \leq \frac{2}{3}$$

et aussi

$$\frac{2}{3n}(2^{2p} - 1) \geq \frac{2(2^{2p} - 1)}{3(2^{2p+1} - 1)} = \frac{1}{3} - \frac{2}{3(2^{2p+1} - 1)}$$

Ainsi,

$$\frac{1}{3} \leq \Delta_n(E, \text{NAT}) \leq \frac{2}{3} + \frac{2}{3(2^{2p+1} - 1)}$$

Le majorant tend vers $2/3$ lorsque n (et donc p) tend vers l'infini, d'où le résultat. $\square$

2.5 Le cas général

L'ensemble que nous avons étudié dans le paragraphe précédent possède une succession de grands « pleins » et de grands « vides », et c'est la raison pour laquelle il n'admet pas de densité naturelle. Passons au cas général. Le principe sera le même, sauf que nous créerons des pleins et des vides « relativement à la suite μ ».

Proposition 13. $\mathcal{D}_\mu \neq \mathcal{P}(\mathbb{N}^*)$.

Démonstration. Choisissons une suite strictement croissante $(a_n)_{n \in \mathbb{N}}$ d'entiers naturels non nuls vérifiant pour tout $n \in \mathbb{N}$,

$$\sum_{k=1}^{a_{n+1}-1} \mu_k \geq 3 \sum_{k=1}^{a_n-1} \mu_k$$

Une telle suite existe en raison de la divergence de la série $\sum_{n \geq 1} \mu_n$. Remarquons que cette inégalité s'écrit aussi

$$\sum_{k=1}^{a_{n+1}-1} \mu_k \geq 3 \sum_{k=1}^{a_{n+1}-1} \mu_k - 3 \sum_{k=a_n}^{a_{n+1}-1} \mu_k$$

d'où l'on tire

$$\sum_{k=a_n}^{a_{n+1}-1} \mu_k \geq \frac{2}{3} \sum_{k=1}^{a_{n+1}-1} \mu_k$$

Considérons l'ensemble

$$E = \bigcup_{q=0}^{\infty} \llbracket a_{2q}, a_{2q+1} - 1 \rrbracket$$

Nous allons montrer que E ne possède pas de densité naturelle. Pour tout $p \in \mathbb{N}$, prenons $n_p = a_{2p+1} - 1$. Nous avons

$$E \cap \llbracket 1, n_p \rrbracket = \bigcup_{q=0}^p \llbracket a_{2q}, a_{2q+1} - 1 \rrbracket$$

et donc

$$\begin{aligned} S_{n_p}(E, \mu) &= \sum_{q=0}^p \sum_{k=a_{2q}}^{a_{2q+1}-1} \mu_k \\ &\geq \sum_{k=a_{2p}}^{a_{2p+1}-1} \mu_k \\ &\geq \frac{2}{3} \sum_{k=1}^{a_{2p+1}-1} \mu_k \\ &= \frac{2}{3} S_{n_p}(\mathbb{N}^*, \mu) \end{aligned}$$

Ainsi, si E possède une densité relativement à μ , alors

$$\mathbb{D}_\mu(E) \geq \frac{2}{3}$$

Prenons maintenant $n_p = a_{2p+2} - 1$. Nous avons toujours

$$E \cap \llbracket 1, n_p \rrbracket = \bigcup_{q=0}^p \llbracket a_{2q}, a_{2q+1} - 1 \rrbracket$$

De là,

$$\begin{aligned}
S_{n_p}(E, \mu) &= \sum_{q=0}^p \sum_{k=a_{2q}}^{a_{2q+1}-1} \mu_k \\
&\leq \sum_{k=1}^{a_{2p+1}-1} \mu_k \\
&= S_{a_{2p+1}-1}(\mathbb{N}^*, \mu) \\
&\leq \frac{1}{3} S_{a_{2p+2}-1}(\mathbb{N}^*, \mu) \\
&= \frac{1}{3} S_{n_p}(\mathbb{N}^*, \mu)
\end{aligned}$$

Ainsi, si E possède une densité relativement à μ , alors

$$\mathbb{D}_\mu(E) \leq \frac{1}{3}$$

Contradiction. L'ensemble E n'a donc pas de densité relativement à μ . $\square$

2.6 Une condition nécessaire et suffisante pour la densité naturelle

Soit E une partie infinie de $\mathbb{N}^*$. Écrivons

$$E = \{a_n : n \in \mathbb{N}^*\}$$

où $(a_n)_{n \in \mathbb{N}^*}$ est une suite strictement croissante d'entiers naturels.

Proposition 14. *Soit $\alpha \in [0, 1]$. On a $E \in \mathcal{D}_{\text{NAT}}$ et $\mathbb{D}_{\text{NAT}}(E) = \alpha$ si et seulement si*

$$(\star) \quad \frac{n}{a_n} \xrightarrow{n \rightarrow \infty} \alpha$$

Démonstration. Supposons $(\star)$. Soit $n \geq a_1$. Par la croissance stricte de la suite, il existe un unique entier $m \in \mathbb{N}^*$ tel que $a_m \leq n < a_{m+1}$. On a

$$E \cap \llbracket 1, n \rrbracket = \{a_1, \dots, a_m\}$$

et donc

$$\frac{|E \cap \llbracket 1, n \rrbracket|}{n} = \frac{m}{n}$$

On en déduit l'encadrement

$$\frac{m}{a_{m+1}} < \frac{|E \cap \llbracket 1, n \rrbracket|}{n} \leq \frac{m}{a_m}$$

Lorsque n tend vers l'infini, il en est de même pour m . Par $(\star)$, le majorant tend donc vers α . Pour le minorant, il suffit de constater que

$$\frac{m}{a_{m+1}} = \frac{m}{m+1} \frac{m+1}{a_{m+1}}$$

Le minorant tend donc aussi vers α . Ainsi, $|E \cap \llbracket 1, n \rrbracket|/n$ tend vers α lorsque n tend vers l'infini.

Inversement, supposons que $E \in \mathcal{D}_{\text{NAT}}$ et $\mathbb{D}_{\text{NAT}}(E) = \alpha$. Pour tout $n \geq 1$,

$$E \cap \llbracket 1, a_n \rrbracket = \{a_1, \dots, a_n\}$$

est de cardinal n . On a donc

$$\frac{n}{a_n} = \frac{|E \cap \llbracket 1, a_n \rrbracket|}{a_n}$$

La suite de terme général $|E \cap \llbracket 1, a_n \rrbracket|/a_n$ est une suite extraite de la suite $(|E \cap \llbracket 1, n \rrbracket|/n)_{n \geq 1}$ qui tend vers α . Elle tend donc aussi vers α . $\square$

Exemples. Cette propriété permet de déterminer facilement la densité de certaines parties de $\mathbb{N}^*$. Considérons d'abord les suites arithmétiques. Soit

$$E = \{na + b : n \in \mathbb{N}^*\}$$

où $a, b \in \mathbb{N}^*$. En posant $a_n = na + b$, nous avons

$$\frac{n}{a_n} = \frac{n}{na + b} \xrightarrow{n \rightarrow \infty} \frac{1}{a}$$

Ainsi, $E \in \mathcal{D}_{\text{NAT}}$ et $\mathbb{D}_{\text{NAT}}(E) = 1/a$. Prenons maintenant une suite géométrique. Soit

$$E = \{ba^n : n \in \mathbb{N}^*\}$$

où $a, b \in \mathbb{N}^*$ et $a \neq 1$. En posant $a_n = ba^n$, nous avons

$$\frac{n}{a_n} = \frac{n}{ba^n} \xrightarrow{n \rightarrow \infty} 0$$

Ainsi, $E \in \mathcal{D}_{\text{NAT}}$ et $\mathbb{D}_{\text{NAT}}(E) = 0$.

Pour tout $n \geq 1$, notons p_n le n ième nombre premier. Une conséquence du théorème des nombres premiers est que lorsque n tend vers l'infini, $p_n \sim n \ln n$. On a donc

$$\frac{n}{p_n} \sim \frac{1}{\ln n} \xrightarrow{n \rightarrow \infty} 0$$

Ainsi, en notant $\mathcal{P}$ l'ensemble des nombres premiers, $\mathcal{P} \in \mathcal{D}_{\text{NAT}}$ et $\mathbb{D}_{\text{NAT}}(\mathcal{P}) = 0$.

2.7 Surjectivité

Définition 2. Une suite $(u_n)_{n \in \mathbb{N}^*}$ de réels de l'intervalle $[0, 1]$ est *uniformément distribuée* si pour tous $a, b \in [0, 1]$ tels que $a < b$,

$$\frac{1}{n} |\{j \in \llbracket 1, n \rrbracket : u_j \in [a, b]\}| \xrightarrow{n \rightarrow \infty} b - a$$

L'existence de suites uniformément distribuées n'est pas évidente. Cela dit, de telles suites existent. Par exemple, si a est un irrationnel et $[x] = x - \lfloor x \rfloor$ dénote la partie fractionnaire du réel x , la suite $([na])_{n \in \mathbb{N}^*}$ est uniformément distribuée. Nous ne le montrerons pas ici.

Proposition 15. Soit $\alpha \in]0, 1]$. Soit $(u_n)_{n \in \mathbb{N}^*}$ une suite uniformément distribuée. Soit

$$E = \{n \in \mathbb{N}^* : u_n \leq \alpha\}$$

L'ensemble E appartient à $\mathcal{D}_{\text{NAT}}$ et $\mathbb{D}_{\text{NAT}}(E) = \alpha$.

Démonstration. Prenons $a = 0$ et $b = \alpha$ dans la définition ci-dessus. La définition de l'équidistribution s'écrit

$$\frac{1}{n} |E \cap \llbracket 1, n \rrbracket| \xrightarrow{n \rightarrow \infty} \alpha$$

d'où le résultat. $\square$

Nous savons déjà qu'il existe des parties de $\mathbb{N}^*$ de densité naturelle nulle (par exemple les ensembles finis). Nous avons donc en corollaire la proposition ci-dessous

Proposition 16. *La fonction $\mathbb{D}_{\text{NAT}} : \mathcal{D}_{\text{NAT}} \longrightarrow [0, 1]$ est surjective.*

Remarque. Qu'en est-il pour la fonction de densité relativement à une suite μ quelconque ? Est-elle surjective ? Je n'ai rien trouvé à ce sujet dans la littérature alors je préfère m'abstenir.

3 Densité logarithmique

3.1 Le théorème de Davenport-Erdős

Nous allons montrer que si une partie de $\mathbb{N}^*$ possède une densité naturelle alors il possède aussi une densité logarithmique *et ces deux densités sont égales*. Il s'agit du théorème de Davenport-Erdős.

Pour tout ensemble $E \subseteq \mathbb{N}^*$ et tout réel $x > 1$, notons

$$\mathcal{N}(E, x) = \sum_{k=1}^{\lfloor x \rfloor} \mathbf{1}_E(k) = S_{\lfloor x \rfloor}(E, \text{NAT})$$

Posons également

$$\mathcal{L}(E, x) = \sum_{k=1}^{\lfloor x \rfloor} \frac{\mathbf{1}_E(k)}{k} S_{\lfloor x \rfloor}(E, \text{LOG})$$

Les fonctions $\mathcal{A}$ et $\mathcal{L}$ sont constantes, égales respectivement à $\mathcal{A}(E, k)$ et $\mathcal{L}(E, k)$, sur tout intervalle $[k, k+1[$ où k est entier.

Lemme 17. *Pour tout ensemble $E \subseteq \mathbb{N}^*$ et tout réel $x > 1$,*

$$\mathcal{L}(E, x) = \frac{\mathcal{N}(E, x)}{x} + \int_1^x \frac{\mathcal{N}(E, t)}{t^2} dt$$

Démonstration. L'égalité à montrer résulte facilement d'une intégration par parties utilisant la théorie des distributions, mais prouvons ce résultat de façon élémentaire.

Soit $E \subseteq \mathbb{N}^*$. On a pour tout $x > 1$,

$$\begin{aligned}
\int_1^x \frac{\mathcal{N}(E, t)}{t^2} dt &= \sum_{k=1}^{\lfloor x \rfloor - 1} \int_k^{k+1} \frac{\mathcal{N}(E, t)}{t^2} dt + \int_{\lfloor x \rfloor}^x \frac{\mathcal{N}(E, t)}{t^2} dt \\
&= \sum_{k=1}^{\lfloor x \rfloor - 1} \mathcal{N}(E, k) \int_k^{k+1} \frac{dt}{t^2} + \mathcal{N}(E, \lfloor x \rfloor) \int_{\lfloor x \rfloor}^x \frac{dt}{t^2} \\
&= \sum_{k=1}^{\lfloor x \rfloor - 1} \left(\frac{\mathcal{N}(E, k)}{k} - \frac{\mathcal{N}(E, k)}{k+1} \right) + \frac{\mathcal{N}(E, \lfloor x \rfloor)}{\lfloor x \rfloor} - \frac{\mathcal{N}(E, x)}{x} \\
&= \sum_{k=1}^{\lfloor x \rfloor - 1} \frac{\mathcal{N}(E, k)}{k} - \sum_{k=1}^{\lfloor x \rfloor - 1} \frac{\mathcal{N}(E, k)}{k+1} + \frac{\mathcal{N}(E, \lfloor x \rfloor)}{\lfloor x \rfloor} - \frac{\mathcal{N}(E, x)}{x} \\
&= \sum_{k=1}^{\lfloor x \rfloor} \frac{\mathcal{N}(E, k)}{k} - \sum_{k=2}^{\lfloor x \rfloor} \frac{\mathcal{N}(E, k-1)}{k} - \frac{\mathcal{N}(E, x)}{x} \\
&= \mathcal{N}(E, 1) + \sum_{k=2}^{\lfloor x \rfloor} \frac{\mathcal{N}(E, k) - \mathcal{N}(E, k-1)}{k} - \frac{\mathcal{N}(E, x)}{x}
\end{aligned}$$

Remarquons que $\mathcal{N}(E, 1) = \mathbb{1}_E(1)$ et que pour tout $k \in \llbracket 2, \lfloor x \rfloor \rrbracket$, $\mathcal{N}(E, k) - \mathcal{N}(E, k-1) = \mathbb{1}_E(k)$. On a donc

$$\begin{aligned}
\int_1^x \frac{\mathcal{N}(E, t)}{t^2} dt &= \mathbb{1}_E(1) + \sum_{k=2}^{\lfloor x \rfloor} \frac{\mathbb{1}_E(k)}{k} - \frac{\mathcal{N}(E, x)}{x} \\
&= \mathcal{L}(E, x) - \frac{\mathcal{N}(E, x)}{x}
\end{aligned}$$

d'où le résultat. $\square$

Proposition 18. [DAVENPORT-ERDŐS]

Soit $E \subseteq \mathbb{N}^*$. Si $E \in \mathcal{D}_{\text{NAT}}$ alors $E \in \mathcal{D}_{\text{LOG}}$ et $\mathbb{D}_{\text{NAT}}(E) = \mathbb{D}_{\text{LOG}}(E)$.

Dit plus pompeusement, $\mathcal{D}_{\text{NAT}} \subseteq \mathcal{D}_{\text{LOG}}$ et $\mathbb{D}_{\text{LOG}}|_{\mathcal{D}_{\text{NAT}}} = \mathbb{D}_{\text{NAT}}$.

Démonstration. Supposons que E a une densité naturelle. Soit $\varepsilon > 0$. Il existe $N \in \mathbb{N}^*$ tel que pour tout $t \geq N$,

$$\mathbb{D}_{\text{NAT}}(E) - \varepsilon \leq \frac{\mathcal{N}(E, t)}{t} \leq \mathbb{D}_{\text{NAT}}(E) + \varepsilon$$

De là, pour tout $x \geq N$,

$$\begin{aligned}
\mathcal{L}(E, x) &= \frac{\mathcal{N}(E, x)}{x} + \int_1^x \frac{\mathcal{N}(E, t)}{t^2} dt \\
&\leq \mathbb{D}_{\text{NAT}}(E) + \varepsilon + \int_1^N \frac{\mathcal{N}(E, t)}{t^2} dt + (\mathbb{D}_{\text{NAT}}(E) + \varepsilon) \int_N^x \frac{dt}{t} \\
&\leq \mathbb{D}_{\text{NAT}}(E) + \varepsilon + \int_1^N \frac{1}{t} dt + (\mathbb{D}_{\text{NAT}}(E) + \varepsilon) \int_N^x \frac{dt}{t} \\
&= \mathbb{D}_{\text{NAT}}(E) + \varepsilon + \ln N + (\mathbb{D}_{\text{NAT}}(E) + \varepsilon)(\ln x - \ln N)
\end{aligned}$$

On a aussi la minoration

$$\begin{aligned}\mathcal{L}(E, x) &= \frac{\mathcal{N}(E, x)}{x} + \int_1^x \frac{\mathcal{N}(E, t)}{t^2} dt \\ &\geq (\mathbb{D}_{\text{NAT}}(E) - \varepsilon) \int_N^x \frac{dt}{t} \\ &= (\mathbb{D}_{\text{NAT}}(E) - \varepsilon)(\ln x - \ln N)\end{aligned}$$

Ainsi,

$$(\mathbb{D}_{\text{NAT}}(E) - \varepsilon)(1 - \frac{\ln N}{\ln x}) \leq \frac{\mathcal{L}(E, x)}{\ln x} \leq \frac{\mathbb{D}_{\text{NAT}}(E) + \varepsilon + \ln N}{\ln x} + (\mathbb{D}_{\text{NAT}}(E) + \varepsilon) \left(1 - \frac{\ln N}{\ln x}\right)$$

Le minorant tend vers $\mathbb{D}_{\text{NAT}}(E) - \varepsilon$ lorsque x tend vers l'infini. Il est donc supérieur ou égal à $\mathbb{D}_{\text{NAT}}(E) - 2\varepsilon$ pour tout x assez grand. De même, le majorant est inférieur ou égal à $\mathbb{D}_{\text{NAT}}(E) + 2\varepsilon$ pour tout x assez grand. Ainsi, pour tout x assez grand

$$\mathbb{D}_{\text{NAT}}(E) - 2\varepsilon \leq \frac{\mathcal{L}(E, x)}{\ln x} \leq \mathbb{D}_{\text{NAT}}(E) + 2\varepsilon$$

On en déduit que $\mathcal{L}(E, x)/\ln x$ tend vers $\mathbb{D}_{\text{NAT}}(E)$ lorsque x tend vers $+\infty$. Ainsi, E possède une densité logarithmique, égale à sa densité naturelle $\mathbb{D}_{\text{NAT}}(E)$. $\square$

3.2 Réciproque ?

La réciproque du théorème de Davenport-Erdős est fausse. Il existe des parties de $\mathbb{N}^*$ qui possèdent une densité logarithmique mais pas de densité naturelle. Autrement dit,

Proposition 19. $\mathcal{D}_{\text{NAT}} \subsetneq \mathcal{D}_{\text{LOG}}$.

Démonstration. Soit

$$E = \bigcup_{k=0}^{\infty} \llbracket 2^{2k}, 2^{2k+1} - 1 \rrbracket$$

Nous avons déjà vu que $E \notin \mathcal{D}_{\text{NAT}}$. Montrons que $E \in \mathcal{D}_{\text{LOG}}$. Soit $n \in \mathbb{N}^*$. Il existe un unique entier $m \in \mathbb{N}$ tel que

$$2^m \leq n < 2^{m+1}$$

Supposons d'abord de $m = 2p + 1$ est impair. On a

$$E \cap \llbracket 1, n \rrbracket = \bigcup_{q=0}^p \llbracket 2^{2q}, 2^{2q+1} - 1 \rrbracket$$

et donc

$$\begin{aligned}S_n(E, \text{LOG}) &= \sum_{q=0}^p \sum_{k=2^{2q}}^{2^{2q+1}-1} \frac{1}{k} \\ &= \sum_{q=0}^p (H_{2^{2q+1}-1} - H_{2^{2q}-1})\end{aligned}$$

Rappelons le développement asymptotique

$$H_n = \ln n + \gamma + O\left(\frac{1}{n}\right)$$

où γ est la *constante d'Euler*. On a

$$\begin{aligned} H_{2^{2q+1}-1} - H_{2^{2q}-1} &= \ln(2^{2q+1} - 1) - \ln(2^{2q} - 1) + O\left(\frac{1}{2^{2q}}\right) \\ &= \ln(2^{2q+1}) - \ln(2^{2q}) + O\left(\frac{1}{2^{2q}}\right) \\ &= \ln 2 + O\left(\frac{1}{2^{2q}}\right) \end{aligned}$$

et donc

$$S_n(E, \text{LOG}) = (p+1) \ln 2 + O(1)$$

Remarquons que $2p+1 = \lfloor \lg n \rfloor$ et donc $p \sim \lg n / 2 = \ln n / (2 \ln 2)$. On a donc

$$\Delta_n(E, \text{LOG}) = \frac{1}{2} + O\left(\frac{1}{\ln n}\right)$$

On obtient le même résultat si m est pair. Ainsi, $E \in \mathcal{D}_{\text{LOG}}$ et

$$\mathbb{D}_{\text{LOG}}(E) = \frac{1}{2}$$

□

4 Un exemple intéressant

Soit $b \geq 2$. Pour tout $j \in \llbracket 1, b-1 \rrbracket$, notons $E_{b,j}$ l'ensemble des entiers dont le premier chiffre de l'écriture en base b est égal à j .

Proposition 20. *Pour tout $j \in \llbracket 1, b-1 \rrbracket$, $E_{b,j}$ possède une densité logarithmique et*

$$\mathbb{D}_{\text{LOG}}(E_{b,j}) = \log_b(j+1) - \log_b j$$

Démonstration. Soit $n \in \mathbb{N}^*$. Soient $k \in \llbracket 1, b-1 \rrbracket$ et $p \in \mathbb{N}$ tels que

$$kb^p \leq m < (k+1)b^p$$

Supposons tout d'abord que $k < j$. On a alors

$$E_{b,j} \cap \llbracket 1, n \rrbracket = \bigcup_{q=0}^{p-1} \llbracket jb^q, (j+1)b^q - 1 \rrbracket$$

et donc

$$\begin{aligned} S_n(E_{b,j}, \text{LOG}) &= \sum_{q=0}^{p-1} \sum_{k=jb^q}^{(j+1)b^q-1} \frac{1}{k} \\ &= \sum_{q=0}^{p-1} (H_{(j+1)b^q-1} - H_{jb^q-1}) \end{aligned}$$

On a

$$\begin{aligned} H_{(j+1)b^q-1} - H_{jb^q-1} &= \ln((j+1)b^q - 1) - \ln(jb^q - 1) + O\left(\frac{1}{b^q}\right) \\ &= \ln(j+1) - \ln j + O\left(\frac{1}{b^q}\right) \end{aligned}$$

et donc

$$\Delta_n(E_{b,j}, \text{LOG}) = \frac{p(\ln(j+1) - \ln j)}{\ln n} + O\left(\frac{1}{\ln n}\right)$$

Or, $p = \lfloor \ln n / \ln b \rfloor$, donc

$$\Delta_n(E_{b,j}, \text{LOG}) = \frac{\ln(j+1) - \ln j}{\ln b} + O\left(\frac{1}{\ln n}\right) = \log_b(j+1) - \log_b j + O\left(\frac{1}{\ln n}\right)$$

Des calculs analogues donnent le même résultat si $k = j$ ou $k > j$. Lorsque n tend vers l'infini, $\Delta_n(E_{b,j}, \text{LOG})$ tend vers $\log_b(j+1) - \log_b j$. $\square$

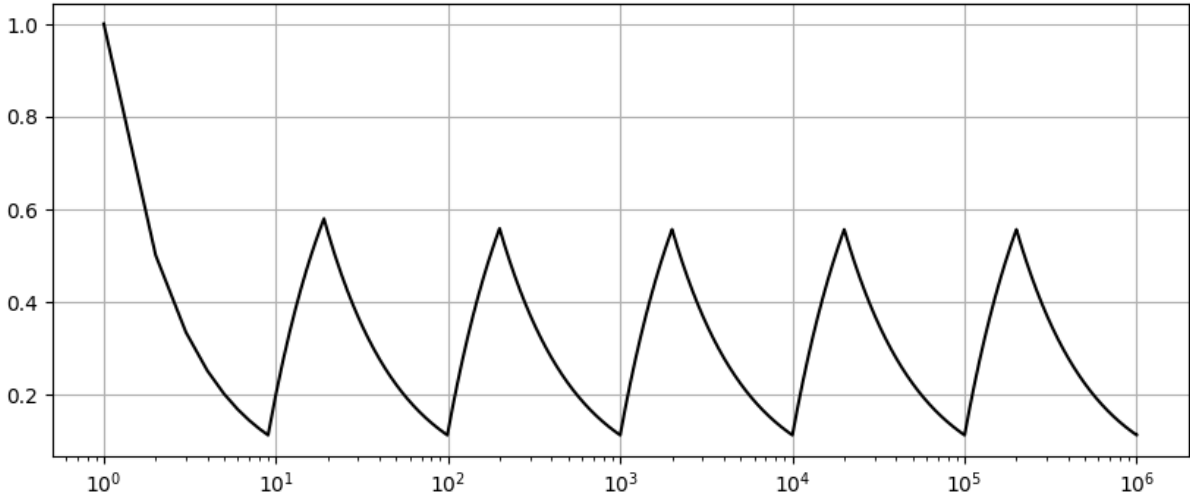


FIGURE 2 – $\Delta_n(E_{10,1}, \text{NAT})$

Jusqu'à quel point ce résultat correspond-il à une réalité ? Livrons-nous à une petite expérience. Le NIST (National Institute of Standards and Technology) fournit la valeur d'environ 350 constantes intervenant en physique. Nous allons tracer, en fonction de $j \in \llbracket 1, 9 \rrbracket$, la proportion de ces constantes dont le premier chiffre (en notation scientifique) est égal à j . En rouge, les proportions en question. En noir, $\log(j+1) - \log j$. La concordance est assez remarquable.

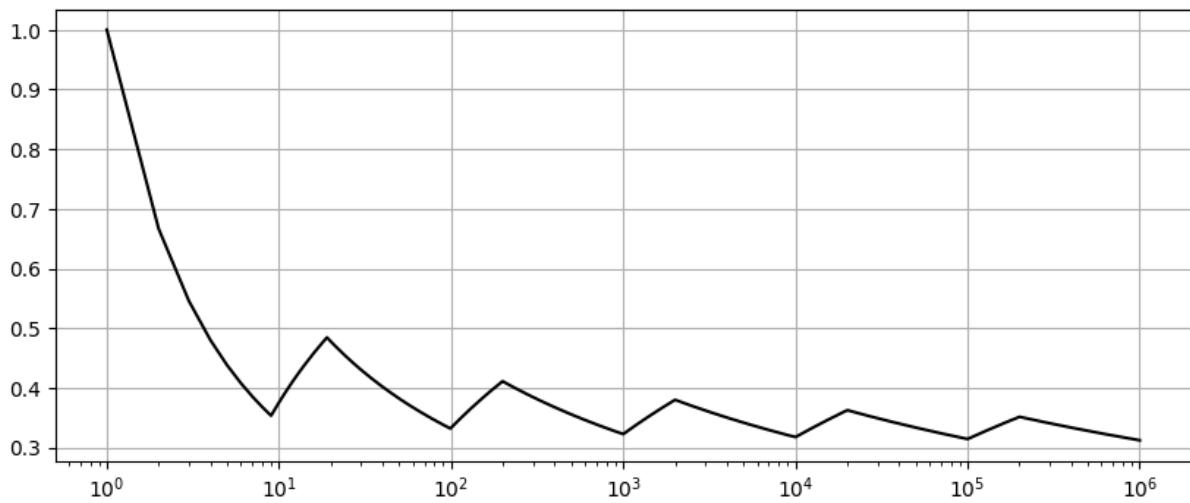


FIGURE 3 – $\Delta_n(E_{10,1}, \text{LOG})$

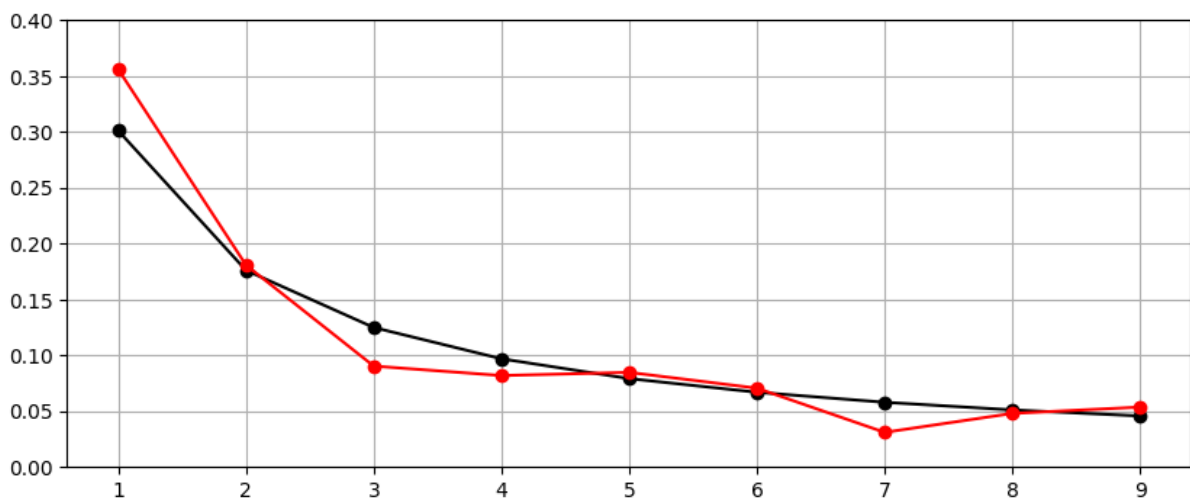


FIGURE 4 – Les constantes de la physique

Bibliographie.

- Gérald Tenenbaum, Introduction to Analytic and Probabilistic Number Theory, Graduate Studies in Mathematics (1995)
- Jörn Steuding, Probabilistic Number Theory (2002).
<http://www.alternatiewiskunde.nl/QED/prob.pdf>
- La page des constantes du NIST :
<https://physics.nist.gov/cuu/Constants/>