Qu'est-ce qu'un nombre aléatoire ?

Marc Lorenzi

23 mai 2025

1 Introduction

Le module `random` de Python contient un certain nombre de fonctions renvoyant des « nombres aléatoires ». Voici ce que dit la documentation du langage à propos de la fonction `randint`.

`random.randint(a, b)`

> Return a random integer N such that $a \leq N \leq b$.

Le nombre 12 est-il un nombre aléatoire entre 3 et 15 ? Y a-t-il des nombres qui sont aléatoires et d'autres qui ne le sont pas ? Cette question n'a évidemment aucun sens. Il serait peut-être plus judicieux de parler de suites de nombres aléatoires ou, plus exactement, de suites aléatoires de nombres.

Soit $(\Omega, \mathfrak{T}, \mathbb{P})$ un espace probabilisé. On suppose cet espace assez riche pour qu'il existe une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes définies sur Ω suivant toutes la loi uniforme $\mathcal{U}(\llbracket a, b \rrbracket)$. Donnons-nous enfin $\omega \in \Omega$. Pour tout $n \geq 1$, soit $x_n = X_n(\omega)$.

Définition 1. La suite $(x_n)_{n\geq 1}$ est une *suite aléatoire* d'entiers de $\llbracket a, b \rrbracket$.

2 Construction

Comment construire une telle suite $(X_n)_{n\geq 1}$? Nous allons discuter d'un cas particulier, celui où $a = 0$ et $b = 1$. Dans ce cas, les X_n suivent la loi de Bernoulli $\mathcal{B}(1/2)$. Nous allons construire les X_n à l'aide du *développement dyadique* des réels de l'intervalle $[0, 1[$. En considérant des développements p-adiques, où p est un entier supérieur ou égal à 2, nous pourrions traiter le cas $a = 0$, $b = p - 1$ puis, par translation, traiter le cas général.

Prenons pour univers l'ensemble $\Omega = [0, 1[$. Munissons Ω de la tribu $\mathfrak{T}$ des boréliens. Rappelons que $\mathcal{T}$ est la plus petite tribu sur Ω qui contient les intervalles. Munissons ensuite $(\Omega, \mathfrak{T})$ de la *mesure de Lebesgue* $\mathbb{P}$. Il nous suffira de savoir que pour tout intervalle I inclus dans Ω, $\mathbb{P}(I)$ est la longueur de I.

Tout réel $x \in \Omega$ s'écrit de façon unique

$$x = \sum_{n=1}^{\infty} \frac{a_n}{2^n}$$

où les a_n appartiennent à $\{0,1\}$ et la suite $(a_n)_{n\geq 1}$ n'est pas stationnaire vers 1. Pour tout $n \in \mathbb{N}^*$, posons

$$X_n(x) = a_n$$

Nous définissons ainsi une fonction $X_n : [0,1[\longrightarrow \{0,1\}$. Pour tout $x \in [0,1[$, $X_n(x)$ est le n^e bit du développement dyadique de x.

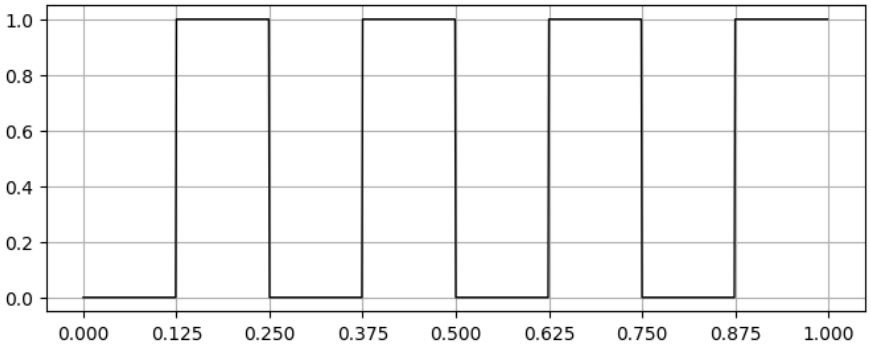

FIGURE 1 – La fonction X_3

Proposition 1. *Pour tout $n \geq 1$, X_n est une variable aléatoire qui suit la loi de Bernoulli $\mathcal{B}(1/2)$.*

Démonstration. Soit $n \in \mathbb{N}^*$. L'événement $\{X_n = 0\}$ est l'ensemble des $x \in \Omega$ dont le n^e bit du développement dyadique est égal à 0. Ces réels sont ceux de la forme

$$x = \sum_{k=1}^{n-1} \frac{a_k}{2^k} + \sum_{k=n+1}^{\infty} \frac{a_k}{2^k} = S_1 + S_2$$

où les a_k appartiennent à $\{0,1\}$. On a

$$S_1 = \frac{1}{2^{n-1}} \sum_{k=1}^{n-1} a_k 2^{n-1-k} = \frac{1}{2^{n-1}} \sum_{k=0}^{n-2} a_{n-1-k} 2^k = \frac{a}{2^{n-1}}$$

où a décrit l'ensemble $[\![0, 2^{n-1} - 1]\!]$. On a également

$$S_2 = \frac{1}{2^n} \sum_{k=n+1}^{\infty} \frac{a_k}{2^{k-n}} = \frac{1}{2^n} \sum_{k=1}^{\infty} \frac{a_{k+n}}{2^k} = \frac{b}{2^n}$$

où b décrit l'ensemble $[0,1[$. Ainsi,

$$\{X_n = 0\} = \frac{1}{2^{n-1}} [\![0, 2^{n-1} - 1]\!] + \frac{1}{2^n}[0,1[= \bigcup_{k=0}^{2^{n-1}-1} \left[\frac{k}{2^{n-1}}, \frac{k}{2^{n-1}} + \frac{1}{2^n}\right[$$

Pour le lecteur un peu perdu dans les notations, voir la figure ci-dessus pour $n = 3$.

L'événement $\{X_n = 0\}$ est la réunion de 2^{n-1} intervalles, c'est donc un borélien. Les intervalles ci-dessus sont disjoints deux à deux et de longueur $1/2^n$, donc

$$\mathbb{P}(X_n = 0) = \sum_{k=0}^{2^{n-1}-1} \frac{1}{2^n} = \frac{1}{2}$$

De là, évidemment,

$$\{X_n = 1\} = \Omega \setminus \{X_n = 0\} \in \mathfrak{T}$$

et

$$\mathbb{P}(X_n = 1) = 1 - \mathbb{P}(X_n = 0) = \frac{1}{2}$$

Ainsi, X_n est une variable aléatoire et $X_n \sim \mathcal{B}(1/2)$. □

Proposition 2. *Les variables aléatoires X_n sont indépendantes.*

Démonstration. Il s'agit de montrer que toute sous-famille finie de $(X_n)_{n\geq 1}$ est une famille de variables aléatoires indépendantes. Soit $N \geq 1$. Soient $a_1, \ldots, a_N \in \{0, 1\}$. Soit $x \in \Omega$. On a $x \in \{X_1 = a_1, \ldots, X_N = a_N\}$ si et seulement si il existe a_{N+1}, a_{N+2}, ... appartenant à $\{0, 1\}$ tels que

$$x = \sum_{k=1}^{N} \frac{a_k}{2^k} + \sum_{k=N+1}^{\infty} \frac{a_k}{2^k} = a + S$$

où, comme nous l'avons déjà vu, S décrit l'intervalle $[0, 1/2^N[$. Ainsi,

$$\{X_1 = a_1, \ldots, X_N = a_N\} = \left[a, a + \frac{1}{2^N}\right[$$

De là,

$$P(X_1 = a_1, \ldots, X_N = a_N) = \frac{1}{2^N} = \prod_{k=1}^{N} \mathbb{P}(X_k = a_k)$$

Les variables aléatoires $X_1, \ldots, X_N$ sont donc indépendantes. □

3 Propriétés des suites aléatoires

Ici encore, concentrons nous sur des variables de Bernoulli de paramètre 1/2.

Que peut-on espérer d'une suite aléatoire de bits ? Possède-t-elle « autant » de 1 que de 0 ? La suite de bits 10101 y apparaît-elle avec la fréquence 1/32 ? Etc. Nous allons voir que la réponse est presque sûrement oui.

3.1 Fréquence d'apparition d'un bit

Soit $a \in \{0, 1\}$. Pour tout $n \in \mathbb{N}^*$, soit $Y_n : \Omega \longrightarrow \{0, 1\}$ définie par $Y_n(\omega) = 1$ si $X_n(\omega) = a$ et $Y_n(\omega) = 0$ sinon. La variable aléatoire Y_n suit une loi de Bernoulli de paramètre

$$\mathbb{P}(Y_n = 1) = \mathbb{P}(X_n = a) = \frac{1}{2}$$

Pour tout $n \geq 1$ et tout $\omega \in \Omega$, la proportion de a dans la suite $(X_1(\omega), \ldots, X_n(\omega))$ est

$$\frac{1}{n} \sum_{k=1}^{n} Y_k(\omega)$$

Par la loi forte des grand nombres, cette proportion tend vers 1/2 presque sûrement.

3.2 Fréquence d'apparition d'uns suite de deux bits

Prenons maintenant deux bits $a, b \in \{0,1\}$. Pour tout $n \in \mathbb{N}^*$, soit $Y_n : \Omega \longrightarrow \{0,1\}$ définie par $Y_n(\omega) = 1$ si $X_n(\omega) = a$ et $X_{n+1}(\omega) = b$, et $Y_n(\omega) = 0$ sinon.

Proposition 3. *Les suites $(Y_{2n})_{n\geq 1}$ et $(Y_{2n+1})_{n\geq 0}$ sont deux suites de variables aléatoires suivant une loi de Bernoulli de paramètre $1/4$.*

Démonstration. Par l'indépendance de X_n et X_{n+1}, la variable aléatoire Y_n suit une loi de Bernoulli de paramètre

$$\mathbb{P}(Y_n = 1) = \mathbb{P}(X_n = a, X_{n+1} = b) = \mathbb{P}(X_n = a)\mathbb{P}(X_{n+1} = b) = \frac{1}{4}$$

Par le lemme des coalitions, les variables aléatoires (X_2, X_3), (X_4, X_5), ... sont indépendantes. Soit $f : \{0,1\}^2 \longrightarrow \{0,1\}$ définie par $f(x,y) = 1$ si $(x,y) = (a,b)$ et $f(x,y) = 0$ sinon. Toujours par le lemme des coalitions, $Y_2 = f(X_2, X_3)$, $Y_3 = f(X_4, X_5)$, ... sont indépendantes.

La même démonstration prouve le résultat pour la suite $(Y_{2n+1})_{n\geq 0}$. □

Par la loi forte des grands nombres,

$$\frac{1}{n}\sum_{k=1}^{n} Y_{2k} \underset{n\to\infty}{\longrightarrow} \frac{1}{4} \quad \text{p.s.}$$

De même,

$$\frac{1}{n}\sum_{k=0}^{n} Y_{2k+1} \underset{n\to\infty}{\longrightarrow} \frac{1}{4} \quad \text{p.s.}$$

Soit $n \in \mathbb{N}^*$. On a

$$\frac{1}{2n}\sum_{k=1}^{2n} Y_k = \frac{1}{2n}\sum_{k=1}^{n} Y_{2k} + \frac{1}{2n}\sum_{k=0}^{n-1} Y_{2k+1}$$

Lorsque n tend vers l'infini, cette quantité tend presque sûrement vers

$$\frac{1}{2}\frac{1}{4} + \frac{1}{2}\frac{1}{4} = \frac{1}{4}$$

De même,

$$\frac{1}{2n+1}\sum_{k=1}^{2n+1} Y_k \underset{n\to\infty}{\longrightarrow} \frac{1}{4} \quad \text{p.s.}$$

Ainsi,

$$\frac{1}{n}\sum_{k=1}^{n} Y_k \underset{n\to\infty}{\longrightarrow} \frac{1}{4} \quad \text{p.s.}$$

Ainsi, la proportion d'apparition de la suite de deux bits ab dans la suite $(X_1(\omega), \ldots, X_n(\omega))$ tend presque sûrement vers $1/4$ lorsque n tend vers l'infini.

3.3 Généralisation

Une démonstration analogue montre que pour tout $p \geq 1$, pour tous, $a_1, \ldots, a_p \in \{0,1\}$, la proportion d'apparition de la suite de p bits $a_1 \ldots a_p$ dans la suite $(X_1(\omega), \ldots, X_n(\omega))$ tend presque sûrement vers $1/2^p$ lorsque n tend vers l'infini.

Pour tout $p \geq 1$ et tous $a_1, \ldots, a_p \in \{0,1\}$, soit $A_{a_1,\ldots,a_p}$ l'ensemble des $\omega \in \Omega$ tels que la proportion d'apparition de la suite de p bits $a_1 \ldots a_p$ dans la suite $(X_1(\omega), \ldots, X_n(\omega))$ tend vers $1/2^p$ lorsque n tend vers l'infini. La discussion précédente montre que $\mathbb{P}(A_{a_1,\ldots,a_p}) = 1$. Soit

$$A = \bigcap_{p \geq 1} \bigcap_{a_1,\ldots,a_p \in \{0,1\}} A_{a_1,\ldots,a_p}$$

L'événement A est une intersection dénombrable d'événements de probabilité 1, donc

$$\mathbb{P}(A) = 1$$

Nous avons donc prouvé le résultat suivant.

Proposition 4. *Pour presque tout $\omega \in \Omega$, on a la propriété suivante :*

Pour tout $p \geq 1$, pour tous, $a_1, \ldots, a_p \in \{0,1\}$, la proportion d'apparition de la suite de p bits $a_1 \ldots a_p$ dans la suite $(X_1(\omega), \ldots, X_n(\omega))$ tend vers $1/2^p$ lorsque n tend vers l'infini.