TP1 – Quelques exemples de martingales

On rappelle que pour exploiter tout le potentiel de Python dans un contexte mathématique, il est nécessaire d'installer ou d'importer un certain nombre de modules à chaque session. Pour cette séance, on pourra par exemple charger les modules suivants :

```
import math
import numpy as np
from matplotlib import pyplot as plt
import numpy.random as rd
from random import *
from scipy.stats import *
```

Exercice 1. *Modèle de Wright–Fischer*
On fixe un entier $N \geq 1$. On définit une suite de variables aléatoires $(X_n)_{n\geq 0}$ à valeurs dans $\{0, 1, \ldots, N\}$ comme suit : $X_0 = k \in \{0, \ldots N\}$ fixé et déterministe et pour $n \geq 0$ on choisit X_{n+1} selon une loi binomiale $\mathcal{B}(N, \frac{X_n}{N})$.

1. Générer et tracer une trajectoire de la suite $(X_n)_{n\geq 0}$.
2. Illustrer le fait que lorsque n tend vers l'infini, (X_n) converge presque sûrement vers une variable limite X_∞ à valeurs dans le $\{0, N\}$.
3. On choisit $N = 100$, $n = 1000$, $m = 1000$, simuler m fois la variable X_n et déterminer la proportion de fois ou $X_n = 0$. Répéter l'opération en faisant variant le point de départ $k \in \{0, \ldots N\}$, que remarquez-vous ?
4. On fixe $0 < \alpha < 1$, et on choisit maintenant X_{n+1} selon une loi binomiale de paramètres $\mathcal{B}\left(N, \alpha\frac{X_n}{N} + (1-\alpha)\left(1 - \frac{X_n}{N}\right)\right)$. Que remarquez-vous sur les trajectoires ?

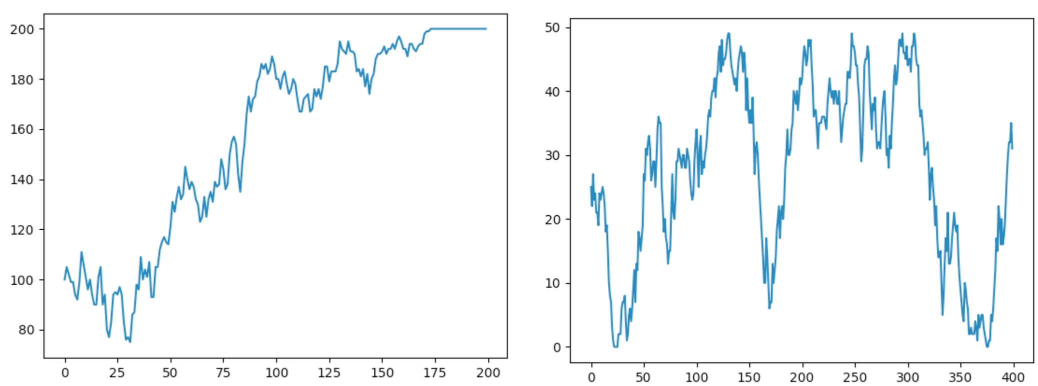

Figure 1 – Deux trajectoires du modèle de Wright-Fisher, sans et avec mutation.

Exercice 2. *Urne de Polya*
On dispose d'un stock infini de balles de 2 couleurs, disons rouges et bleues. Dans une urne, on place initialement 1 boule rouge et 1 balle bleu. On tire une balle au hasard dans l'urne, si elle est rouge on la remet dans l'urne avec une nouvelle balle rouge, si elle est bleue on la remet dans l'urne avec une nouvelle balle bleue. Après n étapes, il y a ainsi $n+2$ balles au total dans l'urne, on note N_n le nombre de balle rouges dans l'urne après n étapes et $X_n = \frac{N_n}{n+2}$ la proportion de balles rouges dans l'urne. On a ainsi $N_0 = 1$, $X_0 = 1/2$ et

$$N_{n+1} = \begin{cases} N_n + 1 & \text{avec probabilité } X_n, \\ N_n & \text{avec probabilité } 1 - X_n. \end{cases}, \qquad X_{n+1} = \frac{N_{n+1}}{n+3}.$$

1. Générer et tracer une trajectoire de la suite $(X_n)_{n\geq 0}$.
2. Illustrer le fait que lorsque n tend vers l'infini, (X_n) converge presque sûrement vers une variable limite X_∞ à valeurs dans le $[0, 1]$.
3. On choisit $n = 500$, $m = 1000$, simuler m fois la variable X_n et tracer la fonction de répartition empirique des données. Superposer la fonction de répartition d'une variable uniforme sur $[0, 1]$. Que remarquez-vous ?

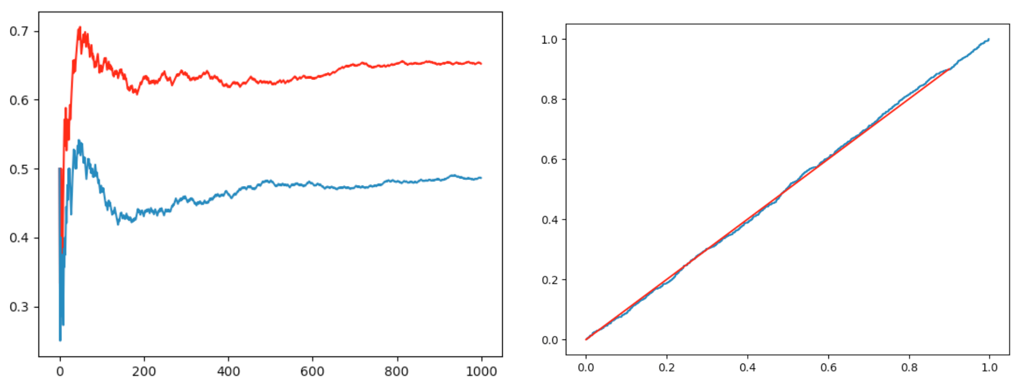

FIGURE 2 – Deux trajectoires du modèle de Polya et la fonction répartition empirique.

Exercice 3. *Marche aléatoire simple ou ruine du joueur*
Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées de loi $\mu := p\delta_1 + q\delta_{-1}$ où $0 < p, q < 1$ et $p + q = 1$. On pose $S_0 := 0$ et $S_n := X_1 + \ldots + X_n$ pour $n \geq 1$ et on désigne par $(\mathcal{F}_n)_{n\geq 1}$ la filtration naturelle associée. On suppose dans un premier temps que $p \neq q$.

1. Montrer que les suites $W_n := S_n - (2p-1)n$ et $M_n = (q/p)^{S_n}$ issues de 0 et 1 respectivement sont des martingales par rapport $(\mathcal{F}_n)_{n\geq 1}$.
2. Simuler des trajectoires de (W_n) et (M_n) et illustrer leur comportement asymptotique.

On suppose maintenant que $p = q = 1/2$, de sorte la suite $(S_n)_{n\geq 0}$ est une martingale. Soit ϕ la transformée de Laplace de μ i.e. $\phi(\lambda) = \cosh(\lambda) = pe^{\lambda} + qe^{-\lambda}$ pour $\lambda \in \mathbb{R}$.

3. Pour $n \geq 0$, on pose $Y_n := e^{\lambda S_n}/\phi(\lambda)^n$. Montrer que $(Y_n)_{n\geq 0}$ est une martingale.
4. Simuler des trajectoires de la suite (Y_n) et illustrer son comportement asymptotique.

Exercice 4. *Agrégation limitée par diffusion interne*
On considère un modèle de diffusion sur la droite $\mathbb{Z}$. On pose $A_0 = \{0\}$, puis on itère le processus suivant. Étant donné un ensemble $A_n \subset \mathbb{Z}$ contenant 0, on considère une marche aléatoire symétrique $(S_k)_{k\geq 0}$ issue de 0 et arrêtée lorsqu'elle sort de A_n. On définit alors A_{n+1} comme l'ensemble $A_n \cup \{x\}$ où x est le point de sortie de la marche. Par exemple, l'ensemble A_1 est choisi uniformément parmi les deux ensembles $\{-1, 0\}$ et $\{0, 1\}$. Notons $G_n := \min A_n$ et $D_n := \max A_n$, de sorte que A_n est de la forme $A_n = \{G_n, G_n + 1, \dots, D_n - 1, D_n\}$. Puisque le cardinal de A_n est $n + 1$, on a $D_n - G_n = n$. Ainsi, A_n est caractérisé par $X_n := D_n + G_n$. On peut montrer que si $|X_n| = 0$, alors $|X_{n+1}| = 1$ et que si $|X_n| > 0$, alors

$$|X_{n+1}| = \begin{cases} |X_n| - 1 & \text{avec probabilité } \dfrac{1}{2} + \dfrac{|X_n|}{2(n+2)}, \\ \\ |X_n| + 1 & \text{avec probabilité } \dfrac{1}{2} - \dfrac{|X_n|}{2(n+2)}. \end{cases}$$

1. Écrire un programme qui génère une trajectoire de la suite $(|X_m|)_{m=0\dots n}$.
2. Vérifier empiriquement que lorsque n tend vers l'infini, la suite X_n/n tend presque sûrement vers zéro.

Exercice 5. *Thunderbolt*
On peut se demander si les performances sportives seront toujours battues et si oui, à quel rythme. Afin de modéliser cette situation, on considère une suite (X_n) de variables aléatoires indépendantes et de même loi admettant une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^+$. Pour tout entier $n \geq 1$ fixé, on note $(X_{(1)}, \dots, X_{(n)})$ le vecteur composé des variables X_i réordonnées, c'est-à-dire $X_{(1)} < X_{(2)} < \dots < X_{(n)}$, et on note R_n le rang relatif de X_n. Il est clair que R_n est une variable aléatoire à valeurs dans $\{1, 2, \dots, n\}$. Pour $n \geq 1$, on dit qu'il se produit un record à l'instant n si $R_n = 1$. On s'intéresse au comportement asymptotique des suites (Z_n) et (M_n) données par

$$Z_n := \sum_{k=1}^{n} \mathbb{1}_{R_k=1}, \qquad M_n := Z_n - \sum_{k=1}^{n} \frac{1}{k}.$$

La suite (Z_n) compte le nombre de records qui se produisent avant l'instant n.

1. Montrer que les variables aléatoires $R_1, \dots, R_n$ sont indépendantes et que pour $1 \leq k \leq n$:
$$\mathbb{P}(R_n = k) = \frac{1}{n}.$$
2. Montrer que (M_n) est une martingale.
3. Illustrer numériquement les convergences suivantes
$$\lim_{n\to+\infty} \frac{M_n}{\log(n)} = 0 \ \ p.s., \qquad \frac{M_n}{\sqrt{\log(n)}} \xrightarrow{loi} \mathcal{N}(0,1), \qquad \frac{Z_n - \log(n)}{\sqrt{\log(n)}} \xrightarrow{loi} \mathcal{N}(0,1).$$