Université de Rennes
Modélisation - Proba/stats

Préparation à l'agrégation
année 2019/2020

# Feuille de travaux pratiques - Python #7

Emeline LUIRARD

Dans ce TP, nous allons illustrer des éléments de la théories de martingales à temps discrêt. Dans une première partie, on met en évidence les propriétés fondamentales des martingales (stationnarité de l'espérance, théorème d'arrêt, convergence). La seconde partie est consacrée à l'étude des arbres de Galton–Watson, un exemple classique d'utilisation des martingales. Les exercices à traiter en priorité sont indiqués en rouge.

## 1 Mise en évidence des propriétés des martingales

On rappelle qu'une suite de variables aléatoires réelles $(M_n)_{n\geq 0}$ est une **sur/sous/martingale par rapport à une filtration $\mathcal{F}_n$** si pour tout $n \geq 0$ :

— $M_n$ est intégrable i.e. $\mathbb{E}[|M_n|] < +\infty$,
— $M_n$ est $\mathcal{F}_n-$mesurable,
— $\mathbb{E}[M_{n+1}|\mathcal{F}_n] \leq M_n$ (sur), $\mathbb{E}[M_{n+1}|\mathcal{F}_n] \geq M_n$ (sous), $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n$ (martingale).

### 1.1 Sur la propriété de martingale

En particulier, selon que la suite $M_n$ est une sur/sous/martingale, la suite des espérances $\mathbb{E}[M_n]$ est décroissante/croissante/constante. Si l'on collecte plusieurs échantillons d'une suite de données, cette propriété fondamentale permet ainsi de "décider" si les données en questions correspondent ou non à des réalisation d'une martingale.

**Exercice 1.** *Constance de l'espérance*
*Les vecteurs de données `X` et `Y` téléchargeables ici et là correspondent chacun à $m = 200$ réalisations deux suites de variables aléatoires $(X_n)_{1\leq n\leq 10}$ et $(Y_n)_{1\leq n\leq 10}$.*

1. *Qui de $(X_n)$ et $(Y_n)$ vous semble être une martingale ?*
2. *La suite qui n'est pas une martingale est en fait une sous-martingale dont le compensateur est déterministe. Estimez ce compensateur.*

### 1.2 Sur le théorème d'arrêt

Étant donnés une martingale $(M_n)$ par rapport à une filtration $\mathcal{F}_n$ et deux temps d'arrêt $S \leq T$ pour cette même filtration, un théorème d'arrêt est un résultat du type

$$M_S = \mathbb{E}[M_T|\mathcal{F}_S].$$

On a besoin d'hypothèse sur la martingale ou sur les temps d'arrêts. Voici les principales hypothèses :

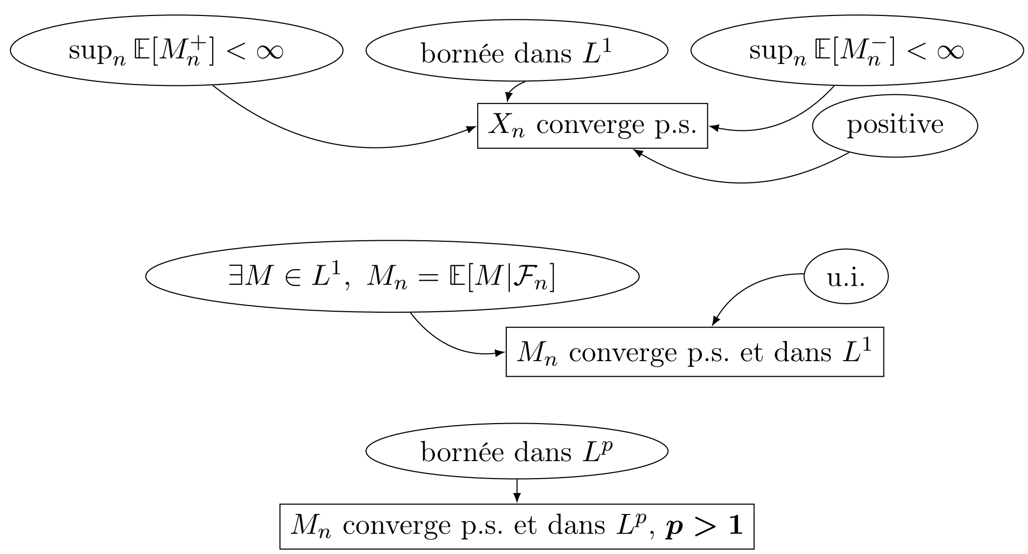
| **Temps d'arrêt** | **Martingale** |
|---|---|
| bornés | $\emptyset$ |
| $\dots$ | $\dots$ |
| $\emptyset$ | u.i. |

**Exercice 2.** *Fixation dans le modèle de Wright–Fisher*
*Soient $k$ et $N$ deux entiers tels que $0 < k < N$. On définit par récurrence une suite $(X_n^N)_{n\geq 0}$ de la façon suivante : $X_0^N := k$ et pour pour tout $n \geq 0$ et $i \in \{0, \dots, N\}$, la loi de $X_{n+1}^N$ sachant que $X_n^N = i$ est la loi binomiale $\mathcal{B}(N, i/N)$. On vérifie aisément que la suite $(X_n^N)_n$ est une martingale bornée par rapport à sa filtration naturelle. On introduit enfin le temps de fixation $T := \inf\{n \geq 0,\ X_n^N = 0 \text{ ou } X_n^N = N\}$.*

1. *Écrire un programme qui prend en entrée les entiers $k, N, n$ et qui génère et trace une trajectoire $(X_i^N)_{0\leq i\leq n}$.*
2. *À l'aide de la méthode de Monte-Carlo, illustrer le théorème d'arrêt au temps $T$.*
3. *Estimer les probabilités de sortie $\mathbb{P}(X_T^N = 0)$ et $\mathbb{P}(X_T^N = N)$.*

## 1.3 Convergence des martingales

L'équation de structure $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n$ qui définit la notion de martingale est très rigide et permet, via le lemme des montées de Doob par exemple, de contrôler précisément les fluctuations de la suite $(M_n)$. Ainsi, si $(M_n)_n$ est une martingale, on a :




**Exercice 3.** *Convergence des martingales*
*Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires i.i.d. de loi $\mathcal{N}(0,1)$. On introduit les suites*

$$M_n := \sum_{i=1}^{n} \frac{X_i}{i}, \qquad N_n := \sum_{i=1}^{n} \frac{X_i}{\sqrt{i}}.$$

1. *Montrer que la suite $(M_n)$ est convergente (en quel sens ?). Illustrer cette convergence.*

2. *Quelle est la loi de la limite $M_\infty$ de la suite $(M_n)_n$ ? L'illustrer en superposant un histogramme empirique et la densité limite.*
3. *La suite $(N_n)$ est-elle convergente ?*
4. *Normaliser la suite $(N_n)$ de sorte qu'elle converge en loi vers une limite gaussienne non-triviale. Illustrer cette convergence.*

**Exercice 4.** *Urne de Polyá.*
*Une urne contient initialement $N_0$ boule noires et $B_0$ boules blanches. On tire une boule au hasard (uniformément) dans l'urne, on regarde sa couleur, puis on la replace dans l'urne en ajoutant une boule de la même couleur, de sorte qu'après l'instant 1, il y a maintenant $N_1$ boules noires et $B_1$ boules blanches. On itère ensuite la même procédure. On note $N_k$ et $B_k$ le nombre de boules noires/blanches après le $k-$ième tirage. On vérifie alors aisément que la suite $(X_k)_{k\geq 0}$ définie par $X_k := \frac{B_k}{B_k+N_k}$ est une martingale bornée.*

1. *Écrire un programme qui prend en entrée des entiers positifs $N_0$, $B_0$ et $n$ et qui génère et trace une trajectoire de longueur $n$ de la martingale $(X_k)_{0\leq k\leq n}$.*
2. *Mettre en évidence la convergence de la martingale $(X_k)_{k\geq 0}$.*
3. *On montre que la loi limite est la loi beta $\beta(N_0, B_0)$. Illustrer ce résultat en superposant la densité limite et un histogramme empirique.*
4. *Particulariser au cas $N_0 = B_0 = 1$. Quelle est la loi de $X_k$, la loi limite ?*

## 2 Un exemple classique d'utilisation des martingales

**Exercice 5.** Arbres de Galton–Watson géométriques
*On rappelle que la loi géométrique de paramètre $p \in ]0, 1/2]$ sur $\mathbb{N}$ est la loi $P = \sum_{i\geq 0} p\, q^i \delta_i$ où l'on a posé $q := 1 - p$. Soit $(X_{n,k})_{n\geq 1, k\geq 1}$ une famille de variables aléatoires indépendantes et identiquement distribuées de loi $P$. On définit par récurrence une suite $(Z_n)_{n\geq 0}$ telle que $Z_0 := 1$ et pour $n \geq 0$ :*

$$Z_{n+1} := \sum_{k=1}^{Z_n} X_{n+1,k}.$$

*La suite $Z_n$ représente le nombre d'individus à la génération $n$ d'un arbre de Galton–Watson de loi de reproduction $P$. On adopte ici la convention $\sum_\varnothing = 0$ de sorte que si $Z_n = 0$ alors $Z_{n+1} = 0$. On désigne par $T := \inf\{n \geq 0,\ Z_n = 0\} \in \mathbb{N} \cup \{+\infty\}$ le temps d'extinction de l'arbre. On note $m := \mathbb{E}[Z_1]$ et $\sigma^2 = \mathbb{E}[Z_1^2] - \mathbb{E}[Z_1]^2 \in \mathbb{R}_+$.*

1. *Écrire un programme qui prend en entrée un entier $n \in \mathbb{N}$, le paramètre $p$, et qui génère et trace une trajectoire $(Z_k)_{k=0\ldots n}$.*
2. *On montre aisément que la suite $(Z_n/m^n)$ est en fait une martingale. À $n$ fixé (petit), utiliser la méthode de Monte-Carlo pour vérifier numériquement que l'on a bien $\mathbb{E}(Z_n) = m^n$.*
3. *Lorsque la loi de reproduction est la loi géométrique de paramètre $p$, la probabilité d'extinction $\mathbb{P}(T < +\infty)$ est égale à $p/q$. Retrouver numériquement ce résultat.*
4. *On suppose ici que $m = 1$. Par la méthode de Monte-Carlo, mettre numériquement en évidence les fait suivants :*
   *(a)* $\lim_{n\to+\infty} n\mathbb{P}(Z_n > 0) = 2/\sigma^2$ *;*
   *(b)* $\mathcal{L}(Z_n/n | Z_n > 0)$ *tend lorsque $n$ tend vers l'infini vers la loi exponentielle $\mathcal{E}(2/\sigma^2)$.*

5. *Lorsque* $m > 1$, *mettre en évidence la convergence de la suite* $Y_n := Z_n/m^n$ *vers une variable aléatoire* $Y_\infty$. *Vérifier numériquement que* $\mathbb{E}[Y_\infty] = 1$ *et* $\mathrm{var}(Y_\infty) = \sigma^2/(m^2 - m)$ .