Probabilité et processus stochastiques

pour les physiciens et les curieux.

Cours n° 2

13 mars 2009

Michel Bauer

IPhT et LPTEns

Cours de l'IPhT — L'orme des Merisiers

Aujourd'hui :

- Indépendance, conditionnement
  - → Espérances conditionnelles
- Processus stochastiques
  - → Filtrations
  - → Temps d'arrêt
  - → Martingales
  - → Applications (ruine, etc)

Indépendance, conditionnement

Dans l'interprétation des probabilités comme fréquences

* Expériences indépendantes $\leftrightarrow$ Probabilités produits

$(\Omega_1, \mathcal{F}_1, p_1)$ $(\Omega_2, \mathcal{F}_2, p_2)$ $(\Omega, \mathcal{F}, p) = (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, p_1 \otimes p_2)$

$A_1 \in \mathcal{F}_1$ $A_2 \in \mathcal{F}_2$ $p(A_1 \times A_2) = p_1(A_1)\, p_2(A_2)$ se

reinterprète comme $A = A_1 \times \Omega_2$ $A' = \Omega_1 \times A_2$ $A \cap A' = A_1 \times A_2$

$$p(A \cap A') = p(A)\, p(A')$$

** Conditionnement $\equiv$ dans une suite d'expériences indépendantes, ne garder que celle dont le résultat vérifie une (des) condition(s) $\leftrightarrow$ Les fréquences correspondantes définissent une nouvelle loi de probabilité.

$(\Omega, \mathcal{F}, p)$ $A \in \mathcal{F}$ $p(A) \neq 0$ $(\Omega, \mathcal{F}, p_A)$

pour $B \in \mathcal{F}$ $p_A(B) = \dfrac{p(A \cap B)}{p(A)}$ probabilité conditionnelle sachant A

On vérifie facilement que P_A est bien une probabilité sur Ω

Intuitivement, si la fréquence d'apparition de B est la même que la condition A soit imposée ou pas, on dit que l'apparition de B est indépendante de la condition A (petite asymétrie)

Ceci se traduit par $p(A \cap B) = p(A)\,p(B)$

Au vu de ces deux interprétation intuitives :

- $A, B \subset (\Omega, \mathcal{F}, p)$ sont dit indépendants si $p(A \cap B) = p(A)\,p(B)$
- $(A_i)_{i \in I}$ famille d'éléments de $\mathcal{F}$ est dit une famille d'évènements indépendants si $\forall J$ fini, $J \subset I$ $\quad p\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} p(A_i)$
- $(\mathcal{F}_i)_{i \in I}$ famille de sous-tribus de $\mathcal{F}$ est dit une famille de tribus indépendantes si toute famille $(A_i)_{i \in I}$ telle que $A_i \in \mathcal{F}_i$ est une famille d'évènement independants

⊖ Les fonctions mesurables $(X_i)_{i\in I}$ $X_i : (\Omega, \mathcal{F}, p) \to (E_i, \mathcal{E}_i)$ sont indépendantes si les tribus $\mathcal{F}_i \equiv \sigma(X_i) \equiv$ plus petite tribu sur Ω rendant X_i mesurable sont indépendantes.

⋮

On atteint ainsi une assez grande généralité. Revenons sur terre.

Lemme : si $X, Y : (\Omega, \mathcal{F}, p) \to (\mathbb{R}, \mathcal{B})$ sont des variables aléatoires intégrables indépendantes et que XY est intégrable,

$$\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$$

Esquisse de preuve :

* X et Y sont simples : $X = \sum_{i\in I} x_i \mathbb{1}_{A_i}$ $Y = \sum_{j\in J} y_j \mathbb{1}_{B_j}$

I, J finis x_i distincts y_j distincts.

Alors $\sigma(X) = \sigma(\{A_i\})$ $\sigma(Y) = \sigma(\{B_j\})$

$\mathbb{1}_{A_i} \mathbb{1}_{B_j} = \mathbb{1}_{A_i \cap B_j}$ $\quad \mathbb{E}(\mathbb{1}_{A_i \cap B_j}) = p(A_i \cap B_j) = p(A_i)\, p(B_j)$ car

$\sigma(X)$ et $\sigma(Y)$ sont indépendants d'où $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$

** si $\varphi : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ mesurable, $\sigma(\varphi \circ X) \subset \sigma(X)$ par définition de la mesurabilité.

si X et Y sont indépendants $(X_+, Y_+), (X_+, Y_-), (X_-, Y_+), (X_-, Y_-)$ sont indépendants.

*** $n = 1, 2, \ldots$ $\varphi_n : \mathbb{R} \to \mathbb{R}$

$$x \to \begin{cases} 0 & x < 0 \\ \frac{k}{n} & x \in \left[\frac{k}{n}, \frac{k+1}{n}\right[\quad k = 0, 1, \ldots n^2 - 1 \\ n & x > n \end{cases}$$

est mesurable

$\varphi_n \circ X_+$ est simple $\quad 0 \le \varphi_n \circ X_+ \le X_+$

$\varphi_n \circ X_+ \nearrow X_+$ (convergence simple) donc si X_+ est intégrable, on a par exemple la convergence dominée.

idem pour $\varphi_n \circ Y_+$ et $(\varphi_n \circ X_+) \times (\varphi_n \circ Y_+)$ et $+ \to \pm$ ok

$$\mathbb{E}(X_+ Y_+) = \lim_{n\to\infty} \mathbb{E}((\varphi_n \circ X_+)(\varphi_n \circ Y_+)) = \lim_{n\to\infty} \mathbb{E}(\varphi_n \circ X_+)\,\mathbb{E}(\varphi_n \circ Y_+) \quad \text{cv}$$

les φ_n sont simples et $\sigma(\varphi_n \circ X_+) \subset \sigma(X_+) \subset \sigma(X)$ (idem Y)

Puis aussi pour $X_+ Y_-$, etc d'où finalement le résultat.

Remarque : d'après ce qui précède, si X et Y sont indépendants et

$\varphi, \psi : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ sont mesurables,

$\mathbb{E}(\varphi(X)\,\psi(Y)) = \mathbb{E}(\varphi(X))\,\mathbb{E}(\psi(Y))$ pour peu que les $\mathbb{E}$sperances aient un sens.

Espérances conditionnelles (crucial pour la suite)

Soit $(\Omega, \mathcal{F}, p)$ un espace probabilisé et $\mathcal{G} \subset \mathcal{F}$ une tribu

Soit $X : \Omega \to (\mathbb{R}, \mathcal{B})$ une variable aléatoire intégrable

On dit que la variable aléatoire $Y : \Omega \to (\mathbb{R}, \mathcal{B})$ est une

espérance conditionnelle de X sachant $\mathcal{G}$ si

- Y est $\mathcal{G}$-mesurable et intégrable
- pour $A \in \mathcal{G}$ $\quad \mathbb{E}(\mathbb{1}_A X) = \mathbb{E}(\mathbb{1}_A Y)$ ie $\int_A X\,dp = \int_A Y\,dp$ pour $A \in \mathcal{G}$.

L'existence d'une espérance conditionnelle de X sachant $\mathcal{G}$ n'est pas évidente. La preuve de l'existence utilise forcement un ou plusieurs piliers de la théorie des probabilités.

Intuitivement ; l'espérance conditionnelle est l'abstraction ultime de nombreuses idées physiques, et en particulier du groupe de renormalisation (plus précisement, de la décimation) :

$\mathcal{F}$ décrit des fluctuation d'un système jusqu'à une certaine échelle, $\mathcal{G} \subset \mathcal{F}$ donc $\mathcal{G}$ décrit les fluctuations jusqu'à une échelle plus grande que celle de $\mathcal{F}$. Si X s'interprète comme

l'action d'un système à l'échelle de $\mathcal{F}$, Y est l'action effective à l'échelle de $\mathcal{G}$: pour les observables qui ne fluctuent pas à une échelle plus petite que celle de $\mathcal{G}$, Y décrit la même information que X.

Voyons cela en détail si Ω est fini ou dénombrable. Comme nous l'avons vu, il n'est pas vraiment restrictif de supposer que $\mathcal{F} = 2^\Omega$. $\mathcal{G} \subset \mathcal{F}$ est associé à une partition $\mathcal{P}$ de Ω $\quad \Omega = \bigcup_{A \in \mathcal{P}} A$ $\quad \mathcal{G}$ est constitué d'unions arbitraires de $A \in \mathcal{P}$. Toute fonction $X : \Omega \to \mathbb{R}$ est $\mathcal{F}$-mesurable et X intégrable équivaut à $\sum_{\omega \in \Omega} |X(\omega)| \, p(\omega) < +\infty$.

On vérifie que $Y : \Omega \to \mathbb{R}$ est $\mathcal{G}$-mesurable si et seulement si Y est constante sur les $A \in \mathcal{P}$

Si Y est une espérance conditionnelle de X sachant $\mathcal{G}$ on a $\int_A Y\,dp = \int_A X\,dp$ pour tout $A \in \mathcal{G}$ et il en est ainsi si

$\int_A Y\,dp = \int_A X\,dp$ pour tout $A \in \mathcal{P}$. ce qui se traduit par

$\sum_{\omega \in A} Y(\omega)\,p(\omega) = \sum_{\omega \in A} X(\omega)\,p(\omega)$. Mais $Y(\omega)$ est constant sur A,

écrivons $Y(\omega) = \hat{Y}_A$ pour $\omega \in A$. Donc

$$\hat{Y}_A \left(\sum_{\omega \in A} p(\omega) \right) = \sum_{\omega \in A} X(\omega)\,p(\omega) \quad \text{ie} \quad \hat{Y}_A\,p(A) = \mathbb{E}(X \mathbb{1}_A)$$

Deux cas : * $p(A) = 0$ Alors $\hat{Y}_A$ est arbitraire

** $p(A) \neq 0$ Alors $\hat{Y}_A = \dfrac{\mathbb{E}(X \mathbb{1}_A)}{p(A)}$ est déterminé.

Donc si Y est une espérance conditionnelle de X sachant $\mathcal{G}$,

$Y(\omega)$ est déterminé hors d'un ensemble de probabilité nulle :

Si $A \in \mathcal{P}$ et $p(A) \neq 0$ $\quad Y(\omega) = \dfrac{\mathbb{E}(X \mathbb{1}_A)}{p(A)}$ pour $\omega \in A$

Si $A \in \mathcal{P}$ et $p(A) = 0$ $\quad Y(\omega)$ est une constante quelconque sur A.

Réciproquement, on vérifie immédiatement que

$$Y(\omega) = \sum_{\substack{A \in \mathcal{P} \\ p(A) \neq 0}} \mathbb{1}_{\omega \in A} \frac{\mathbb{E}(X \mathbb{1}_A)}{p(A)} + \sum_{\substack{A \in \mathcal{P} \\ p(A) = 0}} \mathbb{1}_{\omega \in A} \hat{Y}_A$$

où $\hat{Y}_A$ est arbitraire est une espérance conditionnelle de X sachant $\mathcal{G}$

<u>Conclusion :</u> si Ω est dénombrable, il existe des espérances conditionnelles de X sachant $\mathcal{G}$. Elle diffèrent sur un ensemble de mesure nulle (s'il y en a !)

Remarques: * Y s'obtient bien en moyennant les fluctuations de X à petite échelle.

* si X^2 est intégrable i.e. $\sum_{\omega \in \Omega} X^2(\omega)\, p(\omega) < +\infty$ on peut obtenir Y par projection : $Z = \sum_{A \in \mathcal{P}} z_A \mathbb{1}_A$ est la variable aléatoire $\mathcal{G}$ mesurable la plus générale, Z^2 est intégrable si $\sum_{A \in \mathcal{P}} z_A^2\, p(A) < +\infty$

$\mathbb{E}((X-Z)^2)$ est minimum si Z est une espérance conditionnelle de X sachant $\mathcal{G}$. (Exercice : vérifier)

En langage mathématique si $X \in L^2(\Omega, \mathcal{F}, p)$ (synonyme de X^2 est intégrable), $L^2(\Omega, \mathcal{G}, p)$ est un sous espace fermé de $L^2(\Omega, \mathcal{G}, p)$ et Y, espérance conditionnelle de X sachant $\mathcal{G}$ est la projection de X sur $L^2(\Omega, \mathcal{G}, p)$.

La théorie des espaces de Hilbert est une des routes vers

Théorème : si $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$ est une variable aléatoire intégrable, et si $\mathcal{G}$ est une sous tribu de $\mathcal{F}$ il existe une ou des espérances conditionnelles de X sachant $\mathcal{G}$. Elles diffèrent éventuellement sur un ensemble de probabilité nulle. On désigne par $\mathbb{E}(X \mid \mathcal{G})$ n'importe laquelle des espérances conditionnelles de X sachant $\mathcal{G}$.

Petit miracle : même si l'existence de $\mathbb{E}(X \mid \mathcal{G})$ demande une preuve non triviale, une fois qu'elle est connue on en n'a plus du tout besoin. Seule la propriété caractéristique

$$\mathbb{E}(1\!\!1_A \, \mathbb{E}(X \mid \mathcal{G})) = \mathbb{E}(1\!\!1_A X) \quad \text{pour } A \in \mathcal{G}$$

sert pour calculer.

Par exemple : * Si X est $\mathcal{G}$ mesurable $\quad \mathbb{E}(X|\mathcal{G}) = X$ (!)

$\mathbb{E}(XY|\mathcal{G}) = X\,\mathbb{E}(Y|\mathcal{G})$ ← Exercice*

* Si $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ sont des tribus (commencer par X v.a. simple)

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H}) = \mathbb{E}(\mathbb{E}(X|\mathcal{H})|\mathcal{G}) = \mathbb{E}(X|\mathcal{H})$$ i.e la plus petite tribu gagne. Preuve :

→ $Y = \mathbb{E}(X|\mathcal{H})$ est $\mathcal{H}$ donc $\mathcal{G}$ mesurable donc $\mathbb{E}(Y) = \mathbb{E}(Y|\mathcal{G})$

→ Si $A \in \mathcal{H}$ $\quad \mathbb{E}(\mathbb{1}_A\,\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H})) = \mathbb{E}(\mathbb{1}_A\,\mathbb{E}(X|\mathcal{G}))$

mais $\mathbb{1}_A \in \mathcal{H} \subset \mathcal{G}$ donc $\mathbb{E}(\mathbb{1}_A\,\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(\mathbb{1}_A X) = \mathbb{E}(\mathbb{1}_A\,\mathbb{E}(X|\mathcal{H}))$

donc par unicité $\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H}) = \mathbb{E}(X|\mathcal{H})$ (avec proba = 1)

* Si $\mathcal{G} = \{\emptyset, \Omega\}$ $\quad \mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$

* $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$ pour $\mathcal{G} \subset \mathcal{F}$

astuce souvent utile pour calculer $\mathbb{E}(X)$

Remarque : en ne parlant pas de probabilités, mais de probabilités relatives, la mécanique statistique "trivialise" les notions de probabilités et espérances conditionnelles. Si Ω est fini ou dénombrable et μ une mesure finie sur Ω (poids de Boltzmann) $Z = \mu(\Omega)$ $p(\omega) = \frac{1}{Z}\mu(\omega)$; si $A \subset \Omega$ $Z_A = \mu(A)$

$$p_A(\omega) = \frac{1}{Z_A} \begin{cases} 0 \text{ si } \omega \notin A \\ \mu(\omega) \text{ si } \omega \in A \end{cases}$$

Si $O : \Omega \to \mathbb{R}$ est telle que $\sum_\omega |O(\omega)|\,\mu(\omega) < \infty$

$$\mathbb{E}(O) = \frac{1}{Z} \sum_\omega O(\omega)\,\mu(\omega) = \langle O \rangle$$

$$\langle O \rangle_A = \frac{1}{\mu(A)} \sum_{\omega \in A} O(\omega)\,\mu(\omega)$$ et si $\mathcal{G}$ est une σ-algèbre

sur Ω $\mathbb{E}(O|\mathcal{G}) = \sum_{A \text{ atome de } \mathcal{G}} \langle O \rangle_A \mathbb{1}_A$ est juste une manière de rassembler les $\langle O \rangle_A$ en un seul objet !

Une application de la notion d'indépendance :

Une loi de 0,1 : Soit $(\Omega, \mathcal{F}, p)$ un espace probabilisé, et

soit $\mathcal{A} \subset \mathcal{F}$ une algèbre (pas σ-algèbre) tel que $\sigma(\mathcal{A}) = \mathcal{F}$.

Soit $B \in \mathcal{F}$ tel que B est indépendant de $\mathcal{A}$, i.e

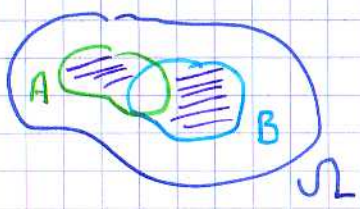
$p(A \cap B) = p(A)\,p(B)$ pour tout $A \in \mathcal{A}$. Alors $p(B) = 0$ ou 1 !

La preuve utilise le lemme d'approximation (cf cours 1)

Si $B \in \mathcal{F}$ et $\varepsilon > 0$, il existe $A \in \mathcal{A}$ tel que

$p(A \Delta B) \leq \varepsilon$ ($A \Delta B \equiv (A \cap B^c) \cup (B \cap A^c)$ partition donc

$p(A \cap B^c) \leq \varepsilon$ et $p(B \cap A^c) \leq \varepsilon$)



$\equiv A \Delta B$

Si B est indépendant de $\mathcal{A}$, B^c aussi et il vient

* $p(A) + p(B \cap A^c) = p(B) + p(A \cap B^c) = p(A \cup B)$

donc $|p(A) - p(B)| \leq p(A \Delta B) \leq \varepsilon$

** $p(B) = p(B, A\ A) + p(B \cap A^c) = p(A)p(B) + p(B \cap A^c)$

↑ B est independant de $\mathcal{A}$

donc $p(B)\ (1 - p(A)) \leq \varepsilon$

En joignant * et ** $p(B)(1-p(B))$ est aussi petit que l'on veut

donc $p(B)\ (1-p(B)) = 0$ ie $p(B) = 0$ ou 1 !

Ceci explique pourquoi on trouve toujours des probabilités 0 ou 1 dans la loi des grandes nombres, la loi du logarithme itéré, etc.

Par exemple dans $(\Omega^\infty, \mathcal{F}^{\otimes\infty}, p^{\otimes\infty})$ cf cours précedent on a

$\mathcal{F}_n$ une suite de σ-algèbres, où $\mathcal{F}_n$ décrit ce qui se passe dans les n premiers copies de Ω si l'on écrit $\Omega^\infty = \Omega^n \times \Omega^{\infty - n}$

Si X est une variable aléatoire sur Ω et si $\omega = (\omega_1, \omega_2, \dots) \in \Omega^\infty$

on définit $Y_n(\omega) \equiv \dfrac{X(\omega_1) + \dots + X(\omega_n)}{n}$

Intuitivement, la convergence d'une suite ne dépend pas de ses n premiers termes donc la propriété $(Y_n(\omega))$ converge vers y est indépendante des n premières valeurs $X(\omega_1), \dots X(\omega_n)$ pourtant

Donc $\{(Y_n) \text{ converge vers } y\}$ est indépendant de tous les $\mathcal{F}_n$ donc

de $\bigcup_n \mathcal{F}_n \equiv \mathcal{A}$. Mais $\sigma(\mathcal{A}) = \mathcal{F}^{\otimes \infty}$ dont nous avons vu

qu'elle permet d'exprimer la convergence. Donc $p(Y_n \to y) = 0$ ou 1 !

2 Ce résultat est trop fauché pour être ultra puissant à lui seul. En particulier, il est rare de savoir grâce à lui que $p(B) = 0$ ou 1 et d'avoir une astuce facile pour montrer que $p(B) > \varepsilon$ pour un ε assez petit, grandissant ainsi $p(B) = 1$

Profitons de cet exemple pour remarquer que les variables aléatoires $X_n(\omega) \equiv X(\omega_n)$ $\omega = (\omega_1, \dots, \omega_n, \dots) \in \Omega^\infty$ sont des variables aléatoires indépendantes sur $(\Omega^\infty, \mathcal{F}^{\otimes\infty}, p^{\otimes\infty})$.

Une part importante de la théorie des probabilités étudie les propriétés de sommes de variables aléatoires indépendantes, notion qui a une fructueuse généralisation, les martingales.

Mais pour en arriver à ces notions, il faut parler de processus aléatoires ----

Processus stochastiques

Champs aléatoire sur $(\Omega, \mathcal{F}, p)$ à valeur dans $(E, \mathcal{E})$ paramétré par S

C'est une application $X : S \times \Omega \to E$ telle que

$$(s, \omega) \mapsto X_s(\omega)$$

pour s fixé $X_s : \Omega \to E$ est une variable aléatoire (i.e. si $A \in \mathcal{E}$, $X_s^{-1}(A) \in \mathcal{F}$)

Un champs aléatoire est donc une collection largement arbitraire de variables aléatoires.

Pour ω fixé $X_\cdot(\omega) : S \to E$ s'appelle un échantillon ou

$$s \mapsto X_s(\omega)$$

une trajectoire ou une configuration du champs.

! Dans la suite, S sera presque toujours un ensemble ordonné ($\mathbb{N}$, $\mathbb{R}^+$, parfois $\mathbb{Z}$ ou $\mathbb{R}$) et on parle alors de **processus aléatoire**, S est alors vu comme un temps qui passe

Si X, Y sont des champs aléatoire $S \times (\Omega, \mathcal{F}, p) \to (E, \mathcal{E})$

- Y est une modification de X si $\forall s \in S \quad p(X_s = Y_s) = 1$
- Y est indiscernable de X si $p(X_s = Y_s \ \forall s \in S) = 1$

Si $X : S \times (\Omega, \mathcal{F}, p) \to (E, \mathcal{E})$ et $X' : S \times (\Omega', \mathcal{F}', p') \to (E, \mathcal{E})$ sont des champs aléatoires

- X et Y ont les mêmes distributions marginales (finies) si

$$p(X_{s_1} \in A_1, \dots, X_{s_k} \in A_k) = p'(X'_{s_1} \in A_1, \dots X'_{s_k} \in A_k) \qquad \forall k \geq 1,$$
$$\forall (s_1, \dots s_k) \in S^k, \quad \forall (A_1, \dots, A_k) \in \mathcal{E}^k$$

2 Remarque philosophique : lors de mesures / expériences réelles, nous avons (au mieux) accès à ces marginales, et $(\Omega, \mathcal{F}, p)$ relève de la modélisation.

Il existe une modélisation canonique :

Soit $\tilde{\Omega} = \{ \tilde{\omega} : S \to E \}$: toutes les applications de S dans E

Un cylindre k-dimensionnel ($k \geq 1$) est un ensemble du type

$\{ \tilde{\omega} \in \tilde{\Omega},\ (\tilde{\omega}(s_1), \dots \tilde{\omega}(s_k)) \in A \}$ où $s_1, \dots, s_k \in S$ et $A \in \mathcal{E}^{\otimes k}$

Pour $(s_1, \dots, s_k)$ fixé, les cylindres forment une σ-algèbre $\tilde{\mathcal{F}}_{s_1, \dots s_k}$. Par définition de $\mathcal{E}^{\otimes k}$, $\tilde{\mathcal{F}}_{s_1, \dots s_k}$ est engendrée par les cylindres produits $\{ \tilde{\omega} \in \tilde{\Omega},\ \tilde{\omega}(s_1) \in A_1, \dots, \tilde{\omega}(s_k) \in A_k \}$ où $A_1, \dots, A_k \in \mathcal{E}$.

On note $\tilde{\mathcal{F}}$ la plus petite σ-algèbre sur $\tilde{\Omega}$ contenant tous les $\mathcal{F}_{s_1, \dots, s_k}$.

Alors $X_\cdot : (\Omega, \mathcal{F}) \to (\tilde{\Omega}, \tilde{\mathcal{F}})$ est mesurable :

$$\omega \to \tilde{\omega} \equiv X_\cdot(\omega) : S \to E$$
$$s \to X_s(\omega)$$

Si $A_1, \dots, A_k \in \mathcal{E}$ définit le cylindre $C = \{ \tilde{\omega} \in \tilde{\Omega}\ \ \tilde{\omega}(s_1) \in A_1, \dots \tilde{\omega}(s_k) \in A_k \}$

$X_\cdot^{-1}(C) = \{ \omega \in \Omega \quad X_{s_1}(\omega) \in A_1, \dots, X_{s_k}(\omega) \in A_k \}$. Comme pour s fixé

$\omega \to X_s(\omega)$ est mesurable, $\{\omega, X_{s_1}(\omega) \in A_1\} \dots \{\omega, X_{s_R}(\omega) \in A_R\}$ sont mesurables donc leur intersection finie est aussi dans $\tilde{\mathcal{F}}$.

Note : petite remarque triviale mais bien utile :

$Y : (\Omega, \mathcal{F}) \to E$ $\quad \mathcal{C} \subset \mathcal{P}^E$ et $Y^{-1}(C) \in \tilde{\mathcal{F}}$ pour $C \in \mathcal{C}$

alors $Y : (\Omega, \mathcal{F}) \to (E, \sigma(\mathcal{C}))$ est mesurable. En effet,

$\{A \subset E, \ Y^{-1}(A) \in \tilde{\mathcal{F}}\}$ est une sigma algèbre car les images inverses respectent les opérations ensembliste.

Donc la vérification ci-dessus suffit à montrer que

$X_\bullet : (\Omega, \tilde{\mathcal{F}}) \to (\tilde{\Omega}, \tilde{\tilde{\mathcal{F}}})$ est mesurable.

On peut alors définir la probabilité image

$\tilde{P} = P \circ X_\bullet^{-1}$ sur $(\tilde{\Omega}, \tilde{\tilde{\mathcal{F}}})$ et on vérifie sans peine que

$\tilde{X} : S \times \tilde{\Omega} \to E$

$(s, \tilde{\omega}) \longmapsto \tilde{\omega}(s)$ est un champ aléatoire qui a les mêmes marginales finies que X.

Le théorème de { consistance de Kolmogorov / cohérence

Si pour tout $k \geq 1$ et $s_1, \dots s_k$ on se donne une loi de probabilité $p_{s_1 \dots s_k}$ sur $(E^k, \mathcal{E}^{\otimes k})$ avec les deux propriétés de cohérence

* Si $A_1, \dots A_k \in \mathcal{E}$ et π est une permutation de $[1,k]$

$$p_{s_1, \dots, s_k}(A_1 \times \dots \times A_k) = p_{s_{\pi(1)} \dots s_{\pi(k)}}(A_{\pi(1)} \times \dots \times A_{\pi(k)})$$

** Si $A \in \mathcal{E}^{\otimes k}$

$$p_{s_1, \dots s_{k+1}}(A \times E) = p_{s_1, \dots, s_k}(A)$$

alors (modulo des conditions techniques sur $(E, \mathcal{E})$; $\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)$ est ok par exemple) il existe une unique probabilité $\tilde{p}$ sur $(\tilde{\Omega}, \tilde{\mathcal{F}})$ telle que

$\tilde{p}((\tilde{\omega}(s_1), \dots \tilde{\omega}(s_k)) \in A) = p_{s_1, \dots s_k}(A)$ pour $k \geq 1$ $s_1, \dots, s_k \in S$ et $A \in \mathcal{E}^{\otimes k}$

Remarques : - les conditions * et ** sont clairement nécessaires

- ce théorème est une bonne base pour construire des champs aléatoires, mais en général $\tilde{\Omega}$ est très gros et $\tilde{\mathcal{F}}$ très petit par rapport à ce qu'on voudrait. (cf : construction du Brownien comme processus à trajectoires continues)

Filtrations

Si S est un ensemble ordonné, on interprète $s \in S$ comme le temps qui passe $\{s' \leq s\}$ = passé de s $\{s' \geq s\}$ = futur de s.

Pour incorporer ceci dans la définition d'un processus aléatoire, notion de filtration:

Une filtration sur $(\Omega, \mathcal{F})$ indexée par S est une famille $\mathcal{F}_s$, $s \in S$ de σ-algèbres sur Ω telles que $\begin{cases} \mathcal{F}_s \subset \mathcal{F}_t & s \leq t \\ \mathcal{F}_t \subset \mathcal{F} & t \in S \end{cases}$

La notion de filtration est tout-à fait cruciale :

$\mathcal{F}_t$ est à la fois intuitivement et mathématiquement ce qu'on peut mesurer en faisant des observations jusqu'au temps t.

Un processus aléatoire X sur $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ est dit adapté à la filtration $\{\mathcal{F}_t\}$ si $X_s : \Omega \to (E, \mathcal{E})$ est $\mathcal{F}_s$ mesurable pour tout $s \in S$.

Si X est un processus aléatoire sur $(\Omega, \mathcal{F}, P)$ et si $\mathcal{F}_t^X = \sigma(X_s \; s \leq t)$ alors $\mathcal{F}_t^X$ est une filtration et $X_\cdot$ est adapté à la filtration $\mathcal{F}_\cdot^X$ par construction

Réciproquement, X est adapté à la filtration $\mathcal{F}_t$ si et seulement $\mathcal{F}_t^X \subset \mathcal{F}_t$ pour $t \in S$.

Vocabulaire : $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in S}, P)$ est un espace probabilisé filtré

Si, sur $(\Omega, \mathcal{F}, p)$ on regarde un unique processus $X : S \times \Omega \to E$ la notion de filtration $\mathcal{F}_t$ n'est pas si importante, car on peut toujours travailler avec $\mathcal{F}_t^X$. Mais en général, à partir de X on construit d'autres processus dont la filtration associée peut être plus petite

Temps d'arrêt

Si $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in S},)$ est un espace mesurable filtré, une application $T : \Omega \to S \cup \{\infty\}^*$ est un temps d'arrêt si pour tout $t \in S$

$\{\omega \text{ tq } T \leq t\} \equiv \{\omega \text{ tels que } T(\omega) \leq t\} \in \mathcal{F}_t$ ie est $\mathcal{F}_t$ mesurable.

$T(\omega) \in S \cup \{\infty\}$ s'interprète comme un temps dépendant de Ω. C'est un temps d'arrêt si $\mathcal{F}_t$, la connaissance accumulée jusqu'au temps t,

* ∞ : symbole tel que $t < \infty$ pour tout $t \in S$

permet de decider si $T(\omega) \leq t$.

La notion de temps d'arrêt est elle aussi cruciale dans l'étude des processus stochastiques. Par exemple, si $X : S \times \Omega \to E$ est un processus aléatoire $\omega \to X_{T(\omega)}(\omega)$ est souvent une variable aléatoire interessante.

On definit $\mathcal{F}_T$, tribu associée au temps d'arrêt T par

$$\mathcal{F}_T = \{ A \in \mathcal{F} \text{ tq } A \cap \{T \leq t\} \in \mathcal{F}_t \quad \forall t \in S \}$$

* $\Omega \in \mathcal{F}_T$: définition d'un temps d'arrêt

* Si $A \in \mathcal{F}_T$ alors $A^c \in \mathcal{F}_T$: $A \cup \{T > t\} = \underbrace{A \cap \{T \leq t\}}_{\in \mathcal{F}_t} \cup \underbrace{A \cap \{T > t\} \cup \{T > t\}}_{= \{T > t\} = \{T \leq t\}^c \in \mathcal{F}_t}$

donc $A \cup \{T > t\} \in \mathcal{F}_t$ et

$$A^c \cap \{T \leq t\} = (A \cup \{T > t\})^c \in \mathcal{F}_t$$

* $(A_i)_{i \in I}$ I dénombrable $A_i \in \mathcal{F}_T$ $\forall i \in I$ alors $\bigcup_i A_i \in \mathcal{F}_T$: écrire ...

A nouveau, intuitivement et mathématiquement, $\mathcal{F}_T$ est la connaissance collectée jusqu'au temps T

Martingales : (crucial pour la suite)

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in S}, P)$ un espace probabilisé filtré.

Le processus $(M_t)_{t\in S} : \Omega \to (\mathbb{R}, \mathcal{B})$ est une martingale si

- M_t est intégrable pour $t \in S$ i.e $\mathbb{E}(|M_t|) < +\infty \quad t \in S$
- (M_t) est adapté à $(\mathcal{F}_t)$ ie M_t est $\mathcal{F}_t$ mesurable $t \in S$
- $\mathbb{E}(M_t \mid \mathcal{F}_s) = M_s$ (avec probabilité 1) pour $s \leq t \in S$

" M_t est conservé en moyenne "

Exercice : Si $\mathbb{E}(M) < +\infty$ $M_t = \mathbb{E}(M \mid \mathcal{F}_t)$ est une martingale (dite fermée)

Si dans cette définition on remplace $\mathbb{E}(M_t \mid \mathcal{F}_s) = M_s$ par

$\mathbb{E}(M_t \mid \mathcal{F}_s) \geq M_s$ ou $\mathbb{E}(M_t \mid \mathcal{F}_s) \leq M_s$ on obtient la notion

de sous martingale, ou de super martingale.

Donnons des exemples :

* $S = \mathbb{N}$, X variable aléatoire de moyenne nulle sur $(\Omega, \mathcal{F}, p)$ (ie $\mathbb{E}|X| < +\infty$ et $\mathbb{E}(X) = 0$)

rien à voir

On définit le processus S_n sur $(\Omega^\infty, \mathcal{F}^{\otimes\infty}, p^{\otimes\infty})$ par

$X_n(\omega) = X(\omega_n)$ pour $\omega = (\omega_1, \dots \omega_n, \dots)$ $n \geq 1$ et

$$S_n(\omega) = \begin{cases} 0 & \text{si } n = 0 \\ X_1(\omega) + \dots + X_n(\omega) & \text{si } n \geq 1 \end{cases}$$

Si $\mathcal{F}_n$ est la plus petite tribu rendant la projection

$$\Pi_n : \Omega^\infty \to \Omega^n \quad \text{mesurable}$$
$$\omega = (\omega_1, \dots) \longmapsto (\omega_1 \dots \omega_n)$$

(avec la convention $\mathcal{F}_0 = \{\phi, \Omega^\infty\}$) alors $(\mathcal{F}_n)_{n\in\mathbb{N}}$ est une filtration sur $(\Omega, \mathcal{F}, p)$ qui collecte l'information disponible au temps n.

- $S_n(\omega) = X(\omega_1) + \dots + X(\omega_n)$ est bien $\mathcal{F}_n$ mesurable donc

  $(S_n)_{n\in\mathbb{N}}$ est adapté à $\mathcal{F}_n$

- $|S_n| \leq |X_1| + \dots + |X_n|$ et $\mathbb{E}(|X|) < \infty$ donc

  $\mathbb{E}(|S_n|) < \infty$ et $(S_n)_{n\geq 0}$ est intégrable

- Si $m < n$ $\mathbb{E}(X_n \mid \mathcal{F}_m) = \mathbb{E}(X_n)$ $(= 0$ par hypothèse$)$

Faisons le détail : Si $A = A_1 \times \dots \times A_m \times \Omega^{\infty-m}$ est un cylindre de $\mathcal{F}_m$ et $B = \Omega^{n-1} \times B_n \times \Omega^{\infty-n}$ alors $(B_n \in \mathcal{F})$

$p_\infty(A \cap B) = p(A_1 \times \dots \times A_m \times \Omega^{n-m-1} \times B_n \times \Omega^{\infty-n})$ vaut

$p(A_1) \cdots p(A_m)\, p(B_n) \equiv p_\infty(A) \times p_\infty(B)$. Mais de tels cylindres A engendrent $\tilde{\mathcal{F}}_m$ alors que de tels B engendrent la tribu $\tilde{\mathcal{F}}_{=n}$ du temps exactement n, tribu pour laquelle X_n est par définition mesurable. Donc $\tilde{\mathcal{F}}_m$ et $\tilde{\mathcal{F}}_{=n}$ sont indépendants si $n > m$ et comme $\sigma(X_n) \subset \tilde{\mathcal{F}}_{=n}$ X_n et $\mathbb{1}_A$ sont indépendants pour tout $A \in \tilde{\mathcal{F}}_m$ d'où

$$\mathbb{E}(\mathbb{1}_A X_n) = \mathbb{E}(\mathbb{1}_A)\,\mathbb{E}(X_n)$$

$\mathbb{E}(X_n)$ est un nombre, $\mathcal{F}_0$ donc $\tilde{\mathcal{F}}_m$ mesurable et $\mathbb{E}(\mathbb{1}_A)\,\mathbb{E}(X_n) = \mathbb{E}(\mathbb{1}_A\, \mathbb{E}(X_n))$ d'où

$$\mathbb{E}(X_n \mid \tilde{\mathcal{F}}_m) = \mathbb{E}(X_n) \quad (= 0 \text{ par hypothèse})$$

- En revanche, si $m \geq n$ $\mathbb{E}(X_n \mid \mathcal{F}_m) = X_n$ car X_n est $\mathcal{F}_{=n}$ $(\subset \mathcal{F}_m)$ mesurable.

- On déduit des deux derniers points que $\mathbb{E}(S_n \mid \mathcal{F}_m) = S_m$ $m \leq n$

Donc $(S_n, \tilde{\mathcal{F}}_n)$ est une martingale.

Remarque : si le temps est $\mathbb{N}$, par la propriété telescopique des espérances conditionnelles itérées, il suffit de vérifier

$$\mathbb{E}(M_{n+1} \mid \tilde{\mathcal{F}}_n) = M_n \text{ pour avoir } \mathbb{E}(M_n \mid \tilde{\mathcal{F}}_m) = M_m \quad m \leq n.$$

Supposons de plus que $\mathbb{E}(X^2) < \infty$. Alors $\mathbb{E}(S_n^2) < \infty$ pour $n \in \mathbb{N}$. [Ceci s'appuie sur Cauchy-Schwarz :

$\mathbb{E}((|X| + \lambda|Y|)^2) \geq 0 \quad \forall \lambda \in \mathbb{R}$ d'où $\mathbb{E}(X^2) + 2\lambda\,\mathbb{E}(|XY|) + \lambda^2\mathbb{E}(Y^2) \geq 0 \; \forall \lambda$

et $\mathbb{E}(|XY|) \leq (\mathbb{E}(|X|^2)\,\mathbb{E}(|Y|^2))^{1/2}$]

S_n^2 est adapté à la filtration $\mathcal{F}_n$, et

$$\mathbb{E}(S_{n+n}^2 \mid \tilde{\mathcal{F}}_n) = \mathbb{E}((S_n + X_{n+1})^2 \mid \tilde{\mathcal{F}}_n)$$

$$(S_n + X_{n+n})^2 = S_n^2 + 2\,S_n X_{n+1} + X_{n+1}^2$$

$\mathbb{E}(S_n^2 \mid \mathcal{F}_n) = S_n^2$ car S_n^2 est $\mathcal{F}_n$ mesurable

$\mathbb{E}(S_n X_{n+1} \mid \mathcal{F}_n) = \ldots = S_n \mathbb{E}(X_{n+1} \mid \mathcal{F}_n) = S_n \cdot 0 = 0$

$\mathbb{E}(X_{n+1}^2 \mid \mathcal{F}_n) = \mathbb{E}(X_{n+1}^2) = \mathbb{E}(X^2)$ car X_{n+1} est indépendante de $\mathcal{F}_n$

Bilan $\mathbb{E}(S_{n+1}^2 \mid \mathcal{F}_n) = S_n^2 + \mathbb{E}(X^2)$ et

$\mathbb{E}(S_n^2 \mid \mathcal{F}_m) = S_m^2 + (n-m)\,\mathbb{E}(X^2)$

donc S_n^2 est une sous-martingale.

Mais mieux encore $R_n \equiv S_n^2 - n\,\mathbb{E}(X^2)$ est une martingale.

Enfin, supposons que $\mathbb{E}(e^{\alpha X}) < \infty$ (par exemple $\alpha \in i\mathbb{R}$, ou X est bornée, ou ...). Alors $\mathbb{E}(e^{\alpha S_n}) = \mathbb{E}\left(\prod_1^n e^{\alpha X_m}\right) = \prod_1^n \mathbb{E}(e^{\alpha X_m}) = \mathbb{E}(e^{\alpha X})^n < \infty$ par indépendance.

Alors $M_n \equiv \dfrac{e^{\alpha S_n}}{\mathbb{E}(e^{\alpha X})^n}$ est une martingale :

M_n est bien adapté à $\mathcal{F}_n$, et pour tout n

$$M_{n+1} = M_n \cdot \frac{e^{\alpha X_{n+1}}}{\mathbb{E}(e^{\alpha X})} \quad \text{où} \quad X_{n+1} \text{ est indépendant de } \mathcal{F}_n$$

donc $$\mathbb{E}(M_{n+1} \mid \mathcal{F}_n) = M_n \frac{\mathbb{E}(e^{\alpha X_{n+1}})}{\mathbb{E}(e^{\alpha X})} = M_n \quad \text{ok.}$$

M_n est un cas particulier de martingale exponentielle.

Pourquoi les martingales sont-elles importantes ?

Les raisons sont nombreuses, donnons deux exemples.

* Les martingales permettent de changer de loi de probabilité pour les processus :

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in S}, P)$ un espace probabilisé filtré et

M_t une martingale (pour $\mathcal{F}_t$) telle que $M_0 = 1$ (avec proba 1)

et $M_t > 0 \quad \forall t \in S$ (avec proba 1).

Alors $\tilde{P}_t = M_t P$ est une loi de probabilité sur $(\Omega, \mathcal{F}_t)$, et

si $s \leq t$ $\tilde{P}_t = \tilde{P}_s$ sur $\mathcal{F}_s$. Donc il est légitime

d'écrire $\tilde{P}$. plutôt que $\tilde{P}_t$, en se souvenant

néanmoins de n'intégrer contre $\tilde{P}$ que des variables

aléatoires mesurable par un $\mathcal{F}_t$ assez grand.

Si Y est $\mathcal{F}_t$ mesurable et disons bornée pour simplifier

$\tilde{\mathbb{E}}(Y) \equiv \mathbb{E}(M_t Y)$. Si $s \geq t$ $\mathbb{E}(M_s Y) = \mathbb{E}(\mathbb{E}(M_s Y \mid \mathcal{F}_t))$

mais Y est $\mathcal{F}_t$ mesurable donc $\mathbb{E}(M_s Y \mid \mathcal{F}_t) = \mathbb{E}(M_s \mid \mathcal{F}_t)\, Y$

qui vaut $M_t Y$ car $M_\cdot$ est une martingale, donc la

définition est cohérente : $\mathbb{E}(M_s Y) = \mathbb{E}(M_t Y)$ si $s \geq t$ et Y $\mathcal{F}_t$ mesurable.

Application (modeste) : $S = \mathbb{N}$, X, S_n, etc comme dans l'exemple ci-dessus, et M_n est la martingale exponentielle. Si Y est $\mathcal{F}_n$-mesurable, $\tilde{\mathbb{E}}(Y) \equiv \mathbb{E}(M_n Y)$

$$M_n = \frac{e^{\alpha S_n}}{\mathbb{E}(e^{\alpha X})^n} \qquad \alpha \in \mathbb{R} \text{ tel que } \mathbb{E}(e^{\alpha X}) < +\infty$$

Regardons la loi de X_n pour $\tilde{\mathbb{P}}$. Si $\varphi : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ est mesurable et disons bornée

$$\tilde{\mathbb{E}}(\varphi(X_n)) = \mathbb{E}(M_n \varphi(X_n)) = \frac{\mathbb{E}(M_{n-1} e^{\alpha X_n} \varphi(X_n))}{\mathbb{E}(e^{\alpha X})}$$

$$= \frac{\mathbb{E}(e^{\alpha X_n} \varphi(X_n))}{\mathbb{E}(e^{\alpha X})} = \frac{\mathbb{E}(e^{\alpha X} \varphi(X))}{\mathbb{E}(e^{\alpha X})}$$

car M_{n-1} et X_n sont indépendantes et $\mathbb{E}(M_{n-1}) = M_0 = 1$

On vérifie facilement que sous $\tilde{p}$ les variables aléatoires X_n sont toujours indépendantes, mais elles on changé de loi, et en particulier ont en général une moyenne non nulle. Sous $\tilde{p}$, le processus S_n a donc en général un biais systématique.

En fait, cette logique peut être renversée.

Soit $A \in \mathcal{F}$ sur $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in S}, p)$ et $p(A) > 0$ on peut définir $\tilde{\mathbb{E}}_A(Y) \equiv \dfrac{\mathbb{E}(Y \mathbb{1}_A)}{\mathbb{E}(\mathbb{1}_A)}$, nouvelle loi de probabilité sur Ω. Si on définit la martingale fermée

$$M_t \equiv \frac{\mathbb{E}(\mathbb{1}_A \mid \mathcal{F}_t)}{\mathbb{E}(\mathbb{1}_A)}$$

et si Y est $\mathcal{F}_t$ mesurable

$$\tilde{\mathbb{E}}_A(Y) = \frac{\mathbb{E}(Y\mathbb{1}_A)}{\mathbb{E}(\mathbb{1}_A)} = \frac{\mathbb{E}(\mathbb{E}(Y\mathbb{1}_A \mid \mathcal{F}_t))}{\mathbb{E}(\mathbb{1}_A)} \overset{\text{car } Y \text{ est } \mathcal{F}_t\text{-mesurable}}{=} \frac{\mathbb{E}(Y\,\mathbb{E}(\mathbb{1}_A \mid \mathcal{F}_t))}{\mathbb{E}(\mathbb{1}_A)} = \mathbb{E}(Y M_t)$$

Si Y est $\tilde{\mathcal{F}}_t$- mesurable et $s \le t$ on a

$$\tilde{\mathbb{E}}_A(Y \mid \tilde{\mathcal{F}}_s) = \frac{1}{M_s}\,\mathbb{E}(Y M_t \mid \tilde{\mathcal{F}}_s) \quad \text{car si } B \in \tilde{\mathcal{F}}_s \subset \tilde{\mathcal{F}}_t.$$

$$\tilde{\mathbb{E}}_A\left(1\!\!1_B \frac{1}{M_s}\,\mathbb{E}(Y M_t \mid \tilde{\mathcal{F}}_s)\right) = \mathbb{E}\left(1\!\!1_B\,\mathbb{E}(Y M_t \mid \tilde{\mathcal{F}}_s)\right)$$

définition de $\mathbb{E}(\,)$ $\rightarrow$ $= \mathbb{E}(1\!\!1_B\, Y M_t) = \tilde{\mathbb{E}}_A(1\!\!1_B\, Y)$

$\rightarrow$ Exemple d'illustration : $S = \{0, 1, \dots, 2N\}$ $\quad \Omega = \{-1, 1\}^{2N}$

$\tilde{\mathcal{F}} = 2^{\Omega}$ $\quad \omega = (\omega_1, \dots, \omega_{2N})$ $\quad X_n(\omega) \equiv \omega_n$ $\quad n = 1, \dots, 2N$

$S_0 = 0$ $\quad S_n = \sum_{m=1}^{n} X_m$ $\quad n \ge 1$ $\quad p$ mesure uniforme

$$p(\omega) = \frac{1}{2^{2N}} \quad \text{pour } \omega \in \Omega$$

$\tilde{\mathcal{F}}_0 = \{\phi, \Omega\}$ $\quad \mathcal{F}_n$ information collectée jusqu'au temps n,

ie la plus petite σ-algèbre rendant $X_1, \dots X_n$ mesurables

$A = \{ S_{2N} = 0 \}$: on regarde une marche aléatoire simple symmétrique de $2N$ pas, conditionnée à finir en 0.

Combinatoire élémentaire : $p(A) = \mathbb{E}(\mathbb{1}_A) = \frac{1}{2^{2N}} \binom{2N}{N}$

La martingale associée est $M_n = \mathbb{E}(\mathbb{1}_A \mid \mathcal{F}_n)$ qu'on calcule naïvement. Au temps n la marche est en S_n et $S_{n+m} - S_n$ est une marche aléatoire simple indépendante du passé qui doit finir en $-S_n$ d'où $M_n = \frac{1}{2^{2N-n}} \binom{2N-n}{N - \frac{n+S_n}{2}} \mathbb{1}_{|S_n| \leq 2N-n}$

Exercice : M_n est bien une martingale.

$$\frac{M_{n+1}}{M_n} = \left(1 - \frac{S_n}{2N-n}\right) \frac{\mathbb{1}_{S_n \in [n-2N,\, 2N-n-2]}}{\mathbb{1}_{S_n \in [n-2N,\, 2N-n]}} \mathbb{1}_{X_{n+1}=1}$$

$$+ \left(1 + \frac{S_n}{2N-n}\right) \frac{\mathbb{1}_{S_n \in [n+2-2N,\, 2N-n]}}{\mathbb{1}_{S_n \in [n-2N,\, 2N-n]}} \mathbb{1}_{X_{n+1}=-1}$$

On veut la distribution conditionnelle sous $\tilde{P}_A$ de X_{n+1} sachant $X_1, \dots X_n$ ie $\mathcal{F}_n$

$$\tilde{\mathbb{E}}_A\left(\mathbb{1}_{X_{n+1}=1} \mid \mathcal{F}_n\right) = \mathbb{E}\left(\frac{M_{n+1}}{M_n}\mathbb{1}_{X_{n+1}=1} \mid \mathcal{F}_n\right)$$

$$= \left(\frac{1}{2} - \frac{S_n}{N-\frac{n}{2}}\right) \frac{\mathbb{1}_{S_n \in [n-2N,\, 2N-n-2]}}{\mathbb{1}_{S_n \in [n-2N,\, 2N-n]}}$$

Sous $\tilde{\mathbb{E}}_A$, les variables aléatoires $X_1 \dots X_n$ ne sont plus indépendantes, mais le processus S_n reste Markovien :

connaissant $S_1, \dots S_n$, $S_{n+1} - S_n = \begin{cases} 1 & \text{proba } \frac{1}{2} - \frac{S_n}{N-\frac{n}{2}} \\ -1 & \text{proba } \frac{1}{2} + \frac{S_n}{N-\frac{n}{2}} \end{cases}$

Ceci permet par exemple de simuler numériquement très facilement des marches allant de 0 à 0 en $2N$ pas.

** Martingales et temps d'arrêts : si $S = \mathbb{N}$ ou $\mathbb{R}^+$ et $(M_t)_{t \in S}$ est une martingale pour $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in S}, P)$ avec $\mathcal{F}_0 = \{\phi, \Omega\}$ (σ-algèbre triviale) on a par la propriété de martingale que $\mathbb{E}(M_t) = \mathbb{E}(M_t \mid \mathcal{F}_0) = M_0$. $t \in S$

Sous certaines hypothèses que nous préciserons plus loin, on a aussi $\mathbb{E}(M_T) = M_0$ si T est un temps d'arrêt.

Voyons immédiatement quelque applications à la ruine du joueur en supposant $\mathbb{E}(M_T) = M_0$, T temps d'arrêt et $M_\bullet$ martingales appropriés.

On travaille donc sur l'exemple de la marche aléatoire simple

$S = \mathbb{N}$ $\quad \Omega = \{-1,1\}^{\infty}$ $\quad \mathcal{F}^{\otimes\infty}$ ↓ noté $\mathcal{F}$ $\quad p^{\otimes\infty}$ ↓ noté p $\quad \mathcal{F}_n$ information collectée au temps n

$\omega = (\omega_1, \omega_2, \dots)$ $\quad X_n(\omega) \equiv \omega_n$ $\quad n = 1, 2, \dots$

Les X_n sont des variables aléatoires indépendantes

et $p(X_n = \pm 1) = \frac{1}{2}$ Rq: il n'est pas difficile de traiter le cas avec biais.

$S_0 = 0$ $\quad S_n = X_1 + \dots + X_n$ $\quad n \geq 1$.

On choisit $a, b \in \mathbb{Z}$ $\quad a < 0 < b$ et on définit le temps aléatoire $T = \inf\{n \in \mathbb{N} \text{ tel que } S_n = a \text{ ou } b\}$,

T est le premier instant où la marche touche le bord de $[a,b]$. Ce problème s'appelle le problème de ruine car si A disposant de $-a$ (€) et B disposant de b (€) jouent à pile ou face avec une pièce juste à 1 € le point

et si S_n représente le gain de A, B est ruiné si : S_n atteint b avant a.

- T est bien un temps d'arrêt : en connaissant $S_1, .. S_n$ on peut effectivement décider si la marche a touché a ou b avant le temps n (inclus) ou pas.
- $p(T = \infty) = 0$ i.e. la marche a une probabilité nulle de rester pour toujours dans $]a,b[$ ← s'obtient par des majorations grossières, exercice.

Donc si p_a est la probabilité de toucher le bord de $[a,b]$ pour la première fois en a i.e. $p_a \equiv p(S_T = a)$ et $p_b \equiv p(S_T = b)$ on a $p_a + p_b = 1$

On va maintenant utiliser les trois martingales que nous

connaissons pour cette situation :

S_n est une martingale

$S_n^2 - n$ est une martingale

$\frac{e^{\alpha S_n}}{(\mathrm{ch}\,\alpha)^n}$ est une martingale. $\forall \alpha$ donc aussi $\frac{\mathrm{ch}\,\alpha\left(S_n - \frac{a+b}{2}\right)}{(\mathrm{ch}\,\alpha)^n}$

- $\mathbb{E}(S_T) = S_0 = 0$ (admis pour l'instant)

  S_T est une variable aléatoire qui vaut $\begin{cases} a & \text{proba } p_a \\ b & \text{——} \quad p_b \end{cases}$

  d'où $\mathbb{E}(S_T) = a p_a + b p_b = S_0 = 0$

  Joint à $p_a + p_b = 1$ il vient $p_a = \frac{b}{b-a}$ $\quad p_b = \frac{-a}{b-a}$

- $\mathbb{E}(S_T^2 - T) = S_0^2 - 0 = 0$ (admis pour l'instant)

  $\mathbb{E}(T) = \mathbb{E}(S_T^2) = a^2 p_a + b^2 p_b = -ab$ temps moyen avant la ruine.

- $\mathbb{E}\left(\dfrac{\operatorname{ch} \alpha \left(S_T - \frac{a+b}{2}\right)}{(\operatorname{ch}\alpha)^T} \right) = \operatorname{ch} \alpha \left(\dfrac{a+b}{2}\right)$ (admis pour l'instant)

$S_T - \frac{a+b}{2}$ vaut $\frac{a-b}{2}$ ou $\frac{b-a}{2}$ donc $\operatorname{ch} \alpha \left(S_T - \frac{a+b}{2}\right) = \operatorname{ch} \alpha \frac{b-a}{2}$

donc $\mathbb{E}\left((\operatorname{ch}\alpha)^{-T} \right) = \dfrac{\operatorname{ch} \alpha \left(\frac{b+a}{2}\right)}{\operatorname{ch} \alpha \left(\frac{b-a}{2}\right)}$

On obtient donc pour un effort minimal la transformée de Laplace de la loi de T :

si l'on définit $\alpha(d)$ par $\operatorname{ch}\alpha = e^d$ il vient

$\mathbb{E}\left(e^{-dT}\right) = \dfrac{\operatorname{ch} \alpha(d) \frac{b+a}{2}}{\operatorname{ch} \alpha(d) \frac{b-a}{2}}$. Le premier pôle $\alpha \frac{b-a}{2} = \pm i \frac{\pi}{2}$

indique que $p(T \geq n)$ décroît comme $\left[\cos\left(\frac{\pi}{b-a}\right)\right]^n$

Il est temps de justifier ces calculs !

On ne peut pas espérer l'impossible : pour la marche aléatoire simple de l'exemple précédent, si T_1 désigne le temps de premier passage en 1, on a $S_{T_1} = 1$ et $\mathbb{E}(1) \neq 0 = S_0$. Cependant :

Théorème ($S = \mathbb{N}$ ou $\mathbb{R}^+$) Si $(M_t, \mathcal{F}_t)_{t \in S}$ est une martingale sur $(\Omega, \mathcal{F}, p)$ et si T est un $\mathcal{F}_t$-temps d'arrêt borné alors $\mathbb{E}(M_T) = M_0$

(T borné $\Leftrightarrow \exists\, t \in S$ tel que $p(T \leq t) = 1$)

Démonstration dans le cas $S = \mathbb{N}$. On suppose $T \leq n$

$$\mathbb{E}(M_T) = \mathbb{E}\left(\sum_{m=0}^{n} M_m \mathbb{1}_{T=m}\right) = \sum_{m=0}^{n} \mathbb{E}(M_m \mathbb{1}_{T=m})$$

$M_m = \mathbb{E}(M_n \mid \mathcal{F}_m)$ car $m \leq n$ d'où

$$\mathbb{E}(M_T) = \sum_{m=0}^{n} \mathbb{E}(\mathbb{E}(M_n \mid \mathcal{F}_m) \mathbb{1}_{T=m}) \quad \text{mais } \mathbb{1}_{T=m} \text{ est } \mathcal{F}_m\text{-mesurable}$$

$$= \sum_{m=0}^{n} \mathbb{E}(M_n \mathbb{1}_{T=m}) = \mathbb{E}\left(M_n \sum_{m=0}^{n} \mathbb{1}_{T=m}\right) = \mathbb{E}(M_n) = M_0$$

C'est en fait le seul théorème vraiment général. Ensuite, c'est du cas par cas. Une approche générale consiste à remplacer T par $T \wedge t = \min(T, t)$ où t est déterministe et à montrer que $\lim_{t \to \infty} \mathbb{E}(M_T - M_{T \wedge t}) = 0$ car le théorème garantit que $\mathbb{E}(M_{T \wedge t}) = M_0$.

Illustrons ceci sur l'exemple de la marche aléatoire avec T le temps de sortie de $[a, b]$, temps non borné donc il faut travailler.

~ Montrons d'abord que $p(T \geq n)$ décroit exponentiellement en n :

à la hache : soit m un entier tel que $m \geq \frac{b-a}{2}$.

Partant de n'importe où dans $[a,b]$ en faisant m pas dans la direction appropriée on touche le bord donc si $k \in \mathbb{N}$ et $n \geq km$ $p(T \geq n) \leq (1-2^{-m})^k$ d'où la décroissance exponentielle et en particulier $p(T=\infty) = 0$.

~ Passons aux choses sérieuses :

* pour $n \leq T$ $S_n \in [a,b]$ et $T \wedge m \leq T$ pour tout m donc $\begin{cases} |S_T - S_{T\wedge m}| = 0 & \text{pour } m \geq T \\ |S_T - S_{T\wedge m}| < b-a & \text{pour } m < T \end{cases}$

$$\mathbb{E}(|S_T - S_{T\wedge m}|) \leq (b-a)\, p(T > m) \to 0 \quad m \to \infty \quad \text{ok !}$$

$$\mathbb{E}(S_T) = 0$$

* de même $|S_T^2 - S_{T\wedge m}^2| = 0 \quad T \leq m$

$|S_T^2 - S_{T\wedge m}^2| < (\max(-a,b))^2 \quad T > m$ d'où

$\mathbb{E}(|S_T^2 - S_{T\wedge m}^2|) < (\max(-a,b))^2 \, p(T > m) \to 0$ qd $m \to \infty$

et $\mathbb{E}(T - T\wedge m) = \sum_{n>m} (n-m)\, p(T=n) \to 0 \quad m \to \infty$

Donc $\mathbb{E}(S_T^2 - T) = 0$

* si $M_n = \dfrac{\operatorname{ch} \alpha (S_n - \frac{b+a}{2})}{(\operatorname{ch}\alpha)^n}$ avec α réel on a

$M_n \leq \operatorname{ch}\alpha (S_n - \frac{(b+a)}{2}) \leq \operatorname{ch}\alpha \frac{(b-a)}{2}$ pour $n \leq T$

donc $|M_T - M_{T\wedge m}| = 0 \quad T \leq m$

$|M_T - M_{T\wedge m}| \leq \operatorname{ch}\alpha \frac{(b-a)}{2} \quad T > m$ et

$$\mathbb{E}\left(|M_T - M_{T\wedge m}|\right) \leq \operatorname{ch}\frac{\alpha(b-a)}{2}\, p(T>m) \to 0 \quad m\to\infty.$$

Donc $\mathbb{E}(M_T) = M_0$

Remarque : En général, pour tout $\omega \in \Omega$ on a

$$\lim_{m\to\infty} T\wedge m\,(\omega) = \lim_{m\to\infty} T(\omega)\wedge m = T(\omega)$$

Si $p(T=\infty) = 0$ alors pour presque

tout ω $\exists n$ tel que $T(\omega)\wedge m = T(\omega)$ $m \geq n$

donc $M_{T\wedge m}$ converge presque surement vers M_T.

On peut essayer d'appliquer des théorème généraux de type convergence dominée. Si $\mathbb{E}(M_T) \neq M_0$, ça signifie que de la probabilité à fuit à l'∞.

Exercice : analyser ce phenomène pour le premier passage en 1.