Espérance conditionnelle pour les vecteurs gaussiens

- $(X_1, \dots, X_d) \in \mathbb{R}^d$ vecteur aléatoire est gaussien si toute combinaison linéaire des X_i est gaussienne : $\forall \lambda_1 \dots \lambda_d$, $\sum_{i=1}^{d} \lambda_i X_i$ suit une loi gaussienne
- Si $(X_1, \dots X_d)$ est gaussien, $(X_i) \in L^2$. Pour tout Y aléatoire dans $\mathbb{R}^d$ et L^2, $\mathbb{E}(Y \mid (X_1, \dots X_d))$ est la projection orthogonale de Y sur le sous-espace vectoriel des variables aléatoires L^2 et mesurables $/ \sigma(X_1, \dots, X_d)$
- ce sous-espace vectoriel est l'ensemble des variables aléatoires de la forme $h(X_1, \dots X_d)$ où h est mesurable.

En fait, on va voir que cette projection orthogonale est la projection orthogonale sur l'espace vectoriel engendré par $(X_1, \dots, X_d)$ $\to$ espace de dimension finie

---

<u>Prop</u> $(X_1, \dots X_m, Y_1, \dots Y_n)$ vecteur gaussien $\in \mathbb{R}^{m+n}$.

Alors, $(X_1, \dots X_m)$ et $(Y_1 \dots Y_n)$ sont des vecteurs gaussiens et ils sont indépendants si et seulement si $\forall i \in [1, m]$, $\forall j \in [1, n]$

$\mathrm{cov}(X_i, Y_j) = 0$

<u>Remarque</u> : indépendance $\Rightarrow$ covariance nulle (toujours vrai)

covariance nulle + vecteur gaussien $\Rightarrow$ indépendance

mais covariance nulle $\not\Rightarrow$ indépendance

<u>dém</u> : $\mathbb{E}\left[\exp\left(i\langle\xi, X_1 \dots X_m, Y_1 \dots Y_n\rangle\right)\right] = \exp\left(\frac{1}{2}\, {}^t\xi\, Q\, \xi\right)$

où Q est la matrice de la forme quadratique associée au vecteur gaussien $Q_{i,j} = \mathrm{cov}(Z_i Z_j)$ si Z est un vecteur gaussien

covariance nulle $\Rightarrow$ ${}^t\xi\, Q\, \xi = \sum_{j,k=1}^{m} \xi_j \xi_k \,\mathrm{cov}(X_i, X_k) + \sum_{j,k=1}^{n} \xi_j \xi_k \,\mathrm{cov}(Y_j, Y_k)$

$$\mathbb{E}\left[\exp\left(i\langle\xi, X_1 \dots X_m, Y_1 \dots Y_n\rangle\right)\right] = \mathbb{E}\left[\exp\left(i\sum_{j=1}^{m} \xi_j X_j\right)\right] \mathbb{E}\left[\exp\left(i\sum_{j=1}^{n} \xi_{m+j} Y_j\right)\right]$$

$$= \mathbb{E}\left[\exp\left(i\langle\xi^{(1)}, (X_1 \dots X_m)\rangle\right)\right] \mathbb{E}\left[\exp\left(i\langle\xi^{(2)}, (Y_1, \dots Y_n)\rangle\right)\right]$$

$\xi^{(1)} = (\xi_1 \dots \xi_m)$ $\quad \xi^{(2)} = (\xi_{m+1} \dots \xi_{m+n})$

Par la propriété des transformées de Fourier des variables aléatoires, on en déduit que $(X_1 \dots X_m)$ indépendant de $(Y_1 \dots Y_n)$

<u>Th</u> (1) Soit $(X, Y_1, Y_2 \dots Y_n)$ vecteur gaussien centré : ($\mathbb{E}X = 0$, $\forall i$ $\mathbb{E}(Y_i) = 0$)

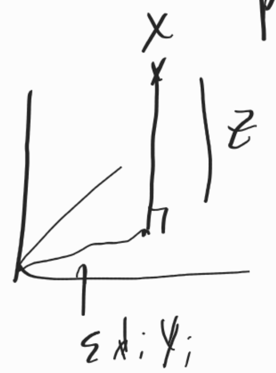
alors $\mathbb{E}(X \mid Y_1 \dots Y_n)$ est la projection orthogonale sur l'espace vectoriel engendré par $Y_1 \dots Y_n$.

$\exists\, \lambda_1, \dots \lambda_n$, $\mathbb{E}(X \mid Y_1 \dots Y_n) = \sum_{i=1}^{n} \lambda_i Y_i$

(2) plus généralement, pour toute fonction h mesurable $: \mathbb{R} \to \mathbb{R}_+$,

$$\mathbb{E}\left(h(X) \mid Y_1, \dots, Y_n\right) = \int_{\mathbb{R}} h(x) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx$$

avec $\sigma^2 = \mathbb{E}\,(\underbrace{X - \sum \lambda_i Y_i}_{Z})^2$: carré de la distance de X à sa projection orthogonale



$m = \sum \lambda_i Y_i$

dém : ① On suppose que X non mesurable $/ \sigma(Y_1 \dots Y_n)$

Soit $\hat{X} = \sum \lambda_i Y_i$ la projection orthogonale de X sur l'espace vectoriel engendré par les Y_i $\quad (X - \hat{X}) \perp Y_i$

$\forall i, \quad \operatorname{cov}(X - \hat{X}, Y_i) = \mathbb{E}((X-\hat{X})Y_i) = 0$

donc $(X - \hat{X}, Y_1 \dots Y_n)$ est un vecteur gaussien et $(X - \hat{X})$ est indépendant de $(Y_1 \dots Y_n)$

$$\text{donc } \mathbb{E}(X \mid Y_1 \dots Y_n) = \mathbb{E}(X - \hat{X} \mid Y_1 \dots Y_n) + \mathbb{E}(\hat{X} \mid Y_1 \dots Y_n)$$
$$= \mathbb{E}(X - \hat{X} \mid Y_1 \dots Y_n) + \hat{X}$$
$$= \underbrace{\mathbb{E}(X - \hat{X})}_{0} + \hat{X} = \hat{X}$$

② $Z = X - \hat{X}$, $Z \sim \mathcal{N}(0, \sigma^2)$ et Z indépendante de $(Y_1 \cdots Y_n)$

$$\mathbb{E}(h(X) \mid Y_1 \cdots Y_n) = \mathbb{E}(h(\hat{X} + Z) \mid Y_1 \cdots Y_n)$$
$$= \int h(\hat{X} + z) \underbrace{P_Z(dz)}_{\text{loi de } Z}$$

Martingales

En français, le mot "martingale" n'a pas le même sens en mathématiques et dans le langage courant.

Dans le langage courant, on parle de "martingale" pour une méthode de jeu pour quelqu'un qui joue au casino.

En mathématiques, une martingale est la quantité d'argent que possède le joueur.

$$\mathbb{E}(\underbrace{X}_{\text{variable aléatoire}} \mid \underbrace{\mathcal{F}}_{\text{tribu}})$$

définition : $(\Omega, \mathcal{F}, P)$ espace de probabilités. On appelle filtration une suite croissante de sous tribus de $\mathcal{F}$: $(\mathcal{F}_n)$, $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n \cdots \subset \mathcal{F}$

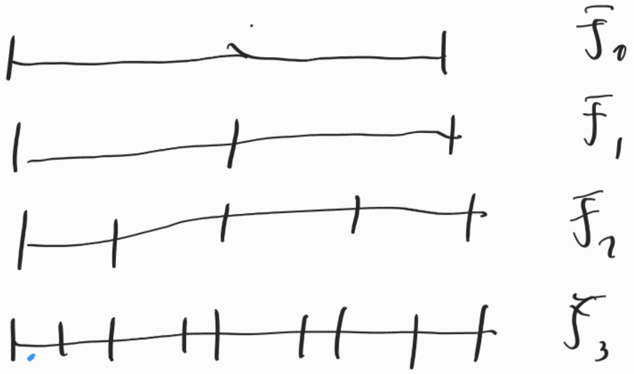
exemples : $X_0, \cdots X_n \cdots$ variables aléatoires, $\mathcal{F}_n = \sigma(X_i, i \leq n)$

$\mathcal{F}_{n+1} = \sigma(\mathcal{F}_n, X_{n+1})$

$\mathcal{F}_n \subset \mathcal{F}_{n+1}$

- $\Omega = [0,1]$, P mesure uniforme, $\mathcal{F}$ tribu borélienne

$$\mathcal{F}_n = \sigma\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right], 0 \le i \le 2^n - 1 \right)$$



déf : un processus $(X_n)_{n\geq 0}$ est une suite de variables aléatoires

$\uparrow$ à temps discret

un processus $(X_t)_{t\in\mathbb{R}_+}$ à temps continu est une famille de variables aléatoires indexée par $\mathbb{R}_+$

exple : mouvement brownien

déf : un processus (X_n) est adapté à la filtration $(\mathcal{F}_n)$ si $\forall n$, X_n est mesurable par rapport à $\mathcal{F}_n$

déf : Soit $(X_n)_{n\geq 0}$ un processus adapté à une filtration $(\mathcal{F}_n)$. On suppose que $\forall n$, $\mathbb{E}(|X_n|)<\infty$. (X_n) est une martingale par rapport à $\mathcal{F}_n$ si $\forall n$ $\quad \mathbb{E}(X_{n+1} / \mathcal{F}_n) = X_n$

remarque : assez souvent, $\mathcal{F}_n = \sigma(X_i, i \le n)$, $\mathbb{E}(X_{n+1} / X_1 \dots X_n) = X_n$

déf : surmartingale si $\mathbb{E}(X_{n+1} / \mathcal{F}_n) \leq X_n$

sous-martingale $\mathbb{E}(X_{n+1} / \mathcal{F}_n) \geq X_n$

en anglais, martingale, supermartingale, submartingale.

exple : si (Y_i) suite de v.a iid, $\mathbb{E}(Y_i) = 0$

$X_n = Y_1 + \dots + Y_n$, $\mathcal{F}_n = \sigma(X_1 \dots X_n) = \sigma(Y_1 \dots Y_n)$

alors (X_n) est une martingale $/ \mathcal{F}_n$

$$\mathbb{E}(X_{n+1} / \mathcal{F}_n) = \mathbb{E}(X_n + Y_{n+1} / \mathcal{F}_n) = \mathbb{E}(X_n / \mathcal{F}_n) + \mathbb{E}(Y_{n+1} / \mathcal{F}_n)$$

$$= X_n + \mathbb{E}(Y_{n+1} / \mathcal{F}_n) = X_n$$

Y_{n+1} indépendante de $Y_1 \dots Y_n$, donc indépendante de $\mathcal{F}_n$ donc

$$\mathbb{E}(Y_{n+1} / \mathcal{F}_n) = \mathbb{E}(Y_{n+1}) = 0$$

On a un joueur qui joue au casino. au temps n, il joue et son gain est Y_n. La quantité $Y_1 + \dots + Y_n = X_n$ est la quantité d'argent du joueur au temps n

Remarque : $\mathbb{E}(X_{n+1}) = \mathbb{E}(\mathbb{E}(X_{n+1} / \mathcal{F}_n)) = \mathbb{E}(X_n) = \mathbb{E}(X_0)$

En moyenne, on ne gagne rien au casino si le jeu est équitable ($\mathbb{E}\, Y_i = 0$)

Autre exemple : processus de branchement

Idée : on modélise l'évolution d'une population

$(Z_n)_{n\geq 0}$ processus ; $\forall n$, Z_n est le nombre d'individus à la n-ième génération.

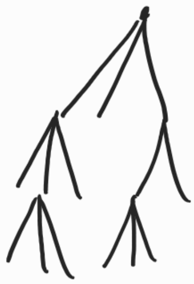
On considère des variables aléatoires $(X_n^{(k)})$ iid, à valeurs dans $\mathbb{N}$, L^1 ; $X_n^{(k)}$ représente le nombre d'enfants du n-ième individu à la génération k.

$\underline{Z_0 = 1}$ (on peut faire plus général)

$Z_1 = X_1^{(0)}$

$Z_2 = X_1^{(1)} + X_2^{(1)} \cdots X_{Z_1}^{(1)}$

$Z_{n+1} = X_1^{(n)} + \dots + X_{Z_n}^{(n)}$



$Z_0 = 1$

$Z_1 = 3$

$Z_2 = 5$

$X_1^{(0)} = 3$

$X_1^{(1)} = 3$, $X_2^{(1)} = 0$, $X_3^{(1)} = 2$

...

Processus de $\underset{\uparrow}{\underbrace{\text{Galton}}}$ – Watson

Francis Galton : cousin de Darwin

Galton s'intéressait aux noms de famille, transmis par les hommes

Si $Z_n = 0$, alors $Z_{n+1} = 0$, $Z_k = 0$ $k \geq n$ la population s'éteint

(Z_n) processus, $\mathcal{F}_n = \sigma(Z_1 \cdots Z_n)$

$\mathbb{E}(Z_{n+1} | \mathcal{F}_n) = \mathbb{E}(X_1^{(n)} + \dots + X_{Z_n}^{(n)} | \mathcal{F}_n) = \mathbb{E}\left(\sum_{k=1}^{\infty} X_k^{(n)} \mathbb{1}_{k\leq Z_n} | \mathcal{F}_n\right)$

cv monotone pour l'espérance conditionnelle

$= \sum_{k=1}^{\infty} \mathbb{E}(X_k^{(n)} \underbrace{\mathbb{1}_{k\leq Z_n}} | \mathcal{F}_n)$ mesurable $| \mathcal{F}_n$

$$= \sum_{k=1}^{\infty} \mathbb{1}_{(k \leq Z_n)} \mathbb{E}(X_k^{(n)} / \mathcal{F}_n)$$

$\mathcal{F}_n = \sigma(Z_1 \dots Z_n)$ $\forall i$, Z_i est une fonction des $X_k^{(i-1)}$

donc $\sigma(Z_1 \dots Z_n) \subset \sigma(X_k^{(i)}, i \leq n-1)$

Pour tout k, $X_k^{(n)}$ est indépendante des $(X_j^{(i)})$, $i \leq n-1$

de $\sigma(X_j^{(i)}, i \leq n-1)$

$\Rightarrow X_k^{(n)}$ indépendante de $\mathcal{F}_n \Rightarrow \mathbb{E}(X_k^{(n)} / \mathcal{F}_n) = \underbrace{\mathbb{E}(X_k^{(n)})}_{= m}$

m : nombre moyen d'enfants

$$\sum_{k=1}^{\infty} \mathbb{1}_{\{k \leq Z_n\}} \mathbb{E}(X_k / \mathcal{F}_n) = m \sum_{k=1}^{\infty} \mathbb{1}_{(k \leq Z_n)} = m Z_n$$

$$= \mathbb{E}(Z_{n+1} / \mathcal{F}_n)$$

$m \geq 1$ (Z_n) est une sous-martingale

$m \leq 1$ surmartingale

$m = 1$ (Z_n) martingale. masse de Dirac en 1 : p.s, $X_1^{(1)} = 1$

si la loi de $X_1^{(1)}$ n'est δ_1 et si $m=1$ alors p.s, $\exists n$,

$Z_n = 0$ [résultat énoncé par Galton, démonstration fausse

Watson a donné une démonstration juste]

<u>déf</u> : $(H_n)_{n \geq 1}$ processus est prévisible par rapport à une filtration $(\mathcal{F}_n)_{n \geq 0}$

si $\forall n$, H_n borné et $\mathcal{F}_{n-1}$ - mesurable

<u>Prop</u> : (X_n) processus adapté, (H_n) prévisible. On pose $(H.X)_0 = 0$ et $n \geq 1$, $(H.X)_n = H_1(X_1 - X_0) + H_2(X_2 - X_1) \cdots + H_n(X_n - X_{n-1})$

alors (i) si (X_n) martingale, $((H.X)_n)$ est une martingale

(ii) si (X_n) surmartingale et si $H_n \geq 0$ ps pour tout $n \geq 1$, alors $((H.X)_n)$ est une surmartingale

<u>idée</u> : H est la quantité d'argent misée au casino au temps n par un joueur

H prévisible : le joueur décide de la quantité d'argent misée au temps n en sachant ce qui s'est passé jusqu'au temps n-1, mais sans savoir ce qui se passe au temps n. → la quantité d'argent gagnée au temps n = $\underbrace{\text{quantité d'argent misée}}_{H_n} \cdot \underbrace{(X_n - X_{n-1})}_{\mathbb{E}(X_n - X_{n-1}) = 0}$ $\Rightarrow \mathbb{E}(H_n(X_n - X_{n-1})) = 0$

<u>dém</u> : . (H_n) bornées, $(X_n) \in L^1 \Rightarrow \forall n, (H.X)_n \in L^1$

. (H_n) prévisible, (X_n) adapté $\rightarrow (H.X)_n$ adapté

. $$\mathbb{E}((H.X)_{n+1} | \mathcal{F}_n) = \mathbb{E}((H.X)_n + H_{n+1}(X_{n+1} - X_n) | \mathcal{F}_n)$$

$$= (H.X)_n + H_{n+1} \underbrace{\mathbb{E}[(X_{n+1} - X_n) | \mathcal{F}_n]}_{0 \text{ si } (X_n) \text{ martingale}}$$

$$= (H.X)_n$$