Processus aléatoires
ENS Paris, 2018-2019

Thomas Budzinski
Bureau V2
thomas.budzinski@ens.fr

TD 1 : Vers le mouvement brownien
Corrigé

Jeudi 13 Septembre

1 Problèmes de mesurabilité

Soit $T > 0$. On note $\mathcal{C}([0,T])$ l'espace des fonctions continues de $[0,T]$ dans $\mathbb{R}$. On note $\mathcal{F}$ la plus petite tribu sur $\mathcal{C}([0,T])$ qui rend mesurable les applications coordonnées $x \to x(t)$ pour tout $t \in [0,T]$.

Exercice 1 On munit $\mathcal{C}([0,T])$ de la norme uniforme. Montrer que $\mathcal{F}$ coïncide avec la tribu borélienne sur $\mathcal{C}([0,T])$. Est-ce toujours vrai en remplaçant $\mathcal{C}([0,T])$ par l'espace $L^\infty([0,T])$ des fonctions mesurables bornées de $[0,T]$ dans $\mathbb{R}$ (toujours muni de la norme uniforme) ?

<u>*Solution de l'exercice 1*</u> On note $\mathcal{B}(\mathcal{C}([0,T]))$ la tribu borélienne issue de la norme uniforme. Les applications coordonnées sont continues pour la norme uniforme, donc mesurables, donc $\mathcal{F} \subset \mathcal{B}(\mathcal{C}([0,T])$ par définition de $\mathcal{F}$.

D'autre part, on veut montrer que tous les ouverts de $\mathcal{C}([0,T])$ sont dans $\mathcal{F}$. Comme $\mathcal{C}([0,T])$ est séparable, tout ouvert est une union dénombrable de boules, donc il suffit de montrer que les boules sont dans $\mathcal{F}$. Il suffit donc de montrer que pour tout $x_0 \in \mathcal{C}([0,T])$ fixé, l'application $x \to \|x_0 - x\|_\infty$ est mesurable pour $\mathcal{F}$. Or, c'est vrai car

$$d(x,x_0) = \sup_{t\in[0,T]} |x(t) - x_0(t)| = \sup_{t\in[0,T]\cap\mathbb{Q}} |x(t) - x_0(t)|,$$

et un sup dénombrable de fonctions mesurables est mesurable.

En remplaçant les fonctions continues par les fonctions bornées, on a toujours $\mathcal{F} \subset \mathcal{B}\left(L^\infty([0,T])\right)$, mais l'autre sens ne marche plus, notamment car les valeurs d'une fonction sur les rationnels ne déterminent plus la fonction. Pour montrer que $\mathcal{F} \neq \mathcal{B}\left(L^\infty([0,T])\right)$, on note $\mathcal{F}_{den}$ l'ensemble des $A \in \mathcal{B}\left(L^\infty([0,T])\right)$ qui ne dépendent que d'un nombre dénombrable de valeurs. Plus précisément, ce sont les A pour lesquels il existe une suite (t_n) telle que, si $x(t_n) = y(t_n)$ pour tout n, alors $x \in A \Leftrightarrow y \in A$. On vérifie que $\mathcal{F}_{den}$ est une tribu, et qu'elle rend mesurable les applications coordonnées. On a donc $\mathcal{F} \subset \mathcal{F}_{den}$. Or, il est facile de vérifier que $\mathcal{F}_{den} \neq \mathcal{B}\left(L^\infty([0,T])\right)$. Par exemple, les ensembles $\{x | \sup_{t\in[0,T]} |x(t)| \leq 1\}$ ou $\{x \text{ continue}\}$ ne sont pas dans $\mathcal{F}_{den}$.

Exercice 2 Soient X et Y deux variables aléatoires à valeurs dans l'espace $\mathcal{C}([0,T])$ muni de la même tribu que ci-dessus. On suppose que pour tout $k \geq 1$ et tous $t_1, \dots, t_k \in [0,T]$, les vecteurs

$$(X_{t_1}, X_{t_2}, \dots, X_{t_k}) \qquad \text{et} \qquad (Y_{t_1}, Y_{t_2}, \dots, Y_{t_k})$$

ont la même loi. Montrer que X et Y ont la même loi.

<u>*Solution de l'exercice 2*</u>
À chercher pour la semaine prochaine !

2 Processus de Poisson

Soit $\lambda > 0$. On rappelle que le *processus de Poisson* $(N_t)_{t\geq 0}$ *d'intensité* λ est défini par

$$N_t = \min\{n \in \mathbb{N} | \xi_0 + \xi_1 + \cdots + \xi_n \geq t\},$$

où $(\xi_i)_{i\geq 0}$ est une suite de variables i.i.d. de loi exponentielle de paramètre λ, c'est à dire de loi $\lambda e^{-\lambda x} \mathbb{1}_{x>0}\, \mathrm{d}x)$. On rappelle également que la loi de Poisson de paramètre λ est définie par $\mathcal{P}_\lambda(k) = e^{-\lambda} \frac{\lambda^k}{k!}$ pour tout $k \in \mathbb{N}$.

Exercice 3 Montrer que pour tous $s, t \geq 0$, les variables N_t et $N_{s+t} - N_t$ sont indépendantes de lois respectives $\mathcal{P}_{\lambda t}$ et $\mathcal{P}_{\lambda s}$.

Montrer que pour tous $0 = t_0 \leq t_1 \leq \cdots \leq t_k$, les variables $N_{t_{i+1}} - N_{t_i}$ pour $0 \leq i \leq k-1$ sont indépendantes de lois respectives $\mathcal{P}_{\lambda(t_{i+1}-t_i)}$.

<u>*Solution de l'exercice 3*</u> Soient $\ell \geq k \geq 0$. On veut calculer $\mathbb{P}\left(N_t = k, N_{s+t} = \ell\right)$. Cet événement équivaut à

$$\{\xi_0 + \cdots + \xi_{k-1} \leq t \leq \xi_0 + \cdots + \xi_k \text{ et } \xi_0 + \cdots + \xi_{\ell-1} \leq s+t \leq \xi_0 + \cdots + \xi_\ell\},$$

donc

$$\begin{aligned}
\mathbb{P}\left(N_t = k, N_{s+t} = \ell\right) &= \lambda^{\ell+1} \int_{\mathbb{R}_+^{\ell+1}} e^{-\lambda(x_0+\cdots+x_\ell)} \mathbb{1}_{\xi_0+\cdots+\xi_{k-1}\leq t\leq \xi_0+\cdots+\xi_k} \\
&\quad \mathbb{1}_{\xi_0+\cdots+\xi_{\ell-1}\leq t\leq \xi_0+\cdots+\xi_\ell}\, \mathrm{d}x_0 \ldots \mathrm{d}x_\ell \\
&= \lambda^{\ell+1} \int_{\mathbb{R}_+^{\ell+1}} e^{-\lambda s_\ell} \mathbb{1}_{s_0\leq\cdots\leq s_{k-1}\leq t\leq s_k\leq\cdots\leq s_{\ell-1}\leq s+t\leq s_\ell}\, \mathrm{d}s_0 \ldots \mathrm{d}s_\ell \\
&= \lambda^\ell \left(\lambda \int_{s+t}^{+\infty} e^{-\lambda s_\ell} \mathrm{d}s_\ell\right) \times \left(\int_{0\leq s_0\leq s_1\leq\cdots\leq s_{k-1}\leq t} \mathrm{d}s_0 \ldots \mathrm{d}s_{k-1}\right) \\
&\quad \times \left(\int_{t\leq s_k\leq s_{k+1}\leq\cdots\leq s_{\ell-1}\leq s+t} \mathrm{d}s_k \ldots \mathrm{d}s_{\ell-1}\right)
\end{aligned}$$

en faisant le changement de variables $s_i = x_0 + x_1 + \cdots + x_i$. Le premier facteur vaut $e^{-\lambda(s+t)}$, le second $\frac{t^k}{k!}$ (c'est un simple calcul par récurrence sur k, en intégrant d'abord sur les $k-1$ plus petites variables, puis sur s_{k-1}) et le troisième vaut $\frac{s^{\ell-k}}{(\ell-k)!}$ (c'est le même calcul que pour le second facteur à changement de variable près). On obtient donc $\mathbb{P}\left(N_t = k_1, N_{s+t} = k_1 + k_2\right) = \frac{(\lambda t)^{k_1}}{k_1!} e^{-\lambda t} \frac{(\lambda s)^{k_2}}{k_2!} e^{-\lambda s}$, d'où le résultat.

Avec plus de deux accroissements, le calcul est similaire mais juste un peu plus long à écrire.

Exercice 4 Soient $N^{(1)}$ et $N^{(2)}$ deux processus de Poisson indépendants d'intensités respectives λ_1 et λ_2. Montrer que $N^{(1)} + N^{(2)}$ est un processus de Poisson d'intensité $\lambda_1 + \lambda_2$.

<u>*Solution de l'exercice 4*</u> D'après l'exercice 2, il suffit de calculer la loi de

$$\left(N_{t_1}^{(1)} + N_{t_1}^{(2)}, \ldots, N_{t_k}^{(1)} + N_{t_k}^{(2)}\right)$$

pour tous $k \geq 0$ et $0 \leq t_1 \leq \cdots \leq t_k$. Plus précisément, d'après l'exercice 3, il suffit de montrer que les variables $N_{t_{i+1}}^{(1)} + N_{t_{i+1}}^{(2)} - N_{t_i}^{(1)} - N_{t_i}^{(2)}$ sont indépendantes, de lois respectives $\mathcal{P}_{(\lambda_1+\lambda_2)(t_{i+1}-t_i)}$. Or, d'après l'exercice 3 et par indépendance de $N^{(1)}$ et $N^{(2)}$ les variables $N_{t_{i+1}}^{(1)} - N_{t_i}^{(1)}$ et $N_{t_{i+1}}^{(2)} - N_{t_i}^{(2)}$ sont toutes indépendantes donc les $N_{t_{i+1}}^{(1)} + N_{t_{i+1}}^{(2)} - N_{t_i}^{(1)} - N_{t_i}^{(2)}$ sont bien indépendantes, et chacune est la somme de deux variables indépendantes de lois $\mathcal{P}_{\lambda_1(t_{i+1}-t_i)}$ et $\mathcal{P}_{\lambda_2(t_{i+1}-t_i)}$. Il suffit donc de vérifier que la somme de deux variables de Poisson X_λ et X_μ de paramètres λ et μ est une variable de Poisson de paramètre $\lambda + \mu$, ce qu'on peut faire par exemple en calculant sa fonction caractéristique :

$$\begin{aligned} \mathbb{E}\left[e^{iu(X_\lambda + X_\mu)}\right] &= \mathbb{E}\left[e^{iuX_\lambda}\right]\mathbb{E}\left[e^{iuX_\mu}\right] \\ &= \left(\sum_{k\geq 0} e^{iuk} e^{-\lambda}\frac{\lambda^k}{k!}\right)\left(\sum_{\ell\geq 0} e^{iu\ell} e^{-\lambda}\frac{\lambda^\ell}{\ell!}\right) \\ &= \exp\left(\lambda e^{iu} - \lambda\right)\exp\left(\mu e^{iu} - \mu\right) \\ &= \exp\left((\lambda+\mu)(e^{iu}-1)\right). \end{aligned}$$

3 Variables gaussiennes et vecteurs gaussiens

On rappelle qu'un vecteur aléatoire X à valeurs dans $\mathbb{R}^d$ est gaussien si pour tout $u = (u_1, \ldots, u_d)$, la variable $u \cdot X$ est gaussienne. En notant $\mu = \mathbb{E}[X]$ la *moyenne* de X et $K = (\mathrm{cov}(X_i, X_j))_{1\leq i,j\leq d}$ sa *matrice de covariance*, on a alors pour tout $u \in \mathbb{R}^d$:

$$\mathbb{E}\left[e^{iu\cdot X}\right] = \exp\left(iu\cdot\mu - \frac{1}{2}{}^t uKu\right).$$

Exercice 5 Soit $X = (X_1, \ldots, X_d)$ un vecteur gaussien centré, i.e. $\mathbb{E}[X] = 0$. Montrer que X_i et X_j sont indépendantes ssi $\mathbb{E}[X_iX_j] = 0$.

<u>*Solution de l'exercice 5*</u> Le sens direct est immédiat. Pour le sens indirect, d'après le rappel ci-dessus, pour tous u et v, on a

$$\begin{aligned} \mathbb{E}\left[e^{i(uX_i + vX_j)}\right] &= \exp\left(-\frac{1}{2}\left(\mathbb{E}\left[X_i^2\right]u^2 + \mathbb{E}\left[X_j^2\right]v^2 + 2\mathbb{E}\left[X_iX_j\right]uv\right)\right) \\ &= \exp\left(-\frac{1}{2}\left(\mathbb{E}\left[X_i^2\right]u^2 + \mathbb{E}\left[X_j^2\right]v^2\right)\right) \\ &= \mathbb{E}\left[e^{iuX_i}\right]\mathbb{E}\left[e^{ivX_j}\right]. \end{aligned}$$

Comme la fonction caractéristique détermine la loi, le couple (X_i, X_j) a donc la loi d'une paire de gaussiennes indépendantes.

Exercice 6 Soient X, Y et ε trois variables aléatoires indépendantes avec X et Y gaussiennes de loi $\mathcal{N}(0,1)$ et $\mathbb{P}(\varepsilon = -1) = \mathbb{P}(\varepsilon = 1) = \frac{1}{2}$. Lesquels des vecteurs suivants sont gaussiens ?

1. (X, ε)
2. (X, Y)
3. $(X, \varepsilon X)$
4. $(X, \varepsilon Y)$
5. $(\varepsilon|X|, \varepsilon|Y|)$
6. $(X, X+Y)$
7. $(X, X+\varepsilon Y)$
8. $(X, \varepsilon X + Y)$

<u>*Solution de l'exercice 6*</u> Les vecteurs 2, 4, 6 et 7 sont gaussiens. Les vecteurs 2 et 4 le sont car ils ont leurs deux coordonnées indépendantes de loi $\mathcal{N}(0,1)$. Les vecteurs 6 et 7 le sont car ils sont les images respectivement des vecteurs 2 et 4 par une application linéaire.

Le vecteur 1 n'est pas gaussien car sa seconde coordonnée ne l'est pas. Le vecteur 3 ne l'est pas car $\mathbb{P}(X + \varepsilon X = 0) = \frac{1}{2}$ donc la somme de ses deux coordonnées n'est pas gaussienne. Pour montrer que le vecteur 8 ne l'est pas, on peut par exemple calculer la fonction caractéristique de la somme de ses coordonnées :

$$\mathbb{E}\left[e^{iu((1+\varepsilon)X+Y)}\right] = \frac{1}{2}\left(1+e^{-u^2}\right)\frac{e^{-u^2/2}}{\sqrt{2\pi}}.$$

Pour montrer que le vecteur 5 ne l'est pas, on peut soit calculer aussi sa fonction caractéristique, soit remarquer que $(\varepsilon|X|, \varepsilon|Y|)$ a une densité strictement positive sur $\mathbb{R}^+ \times \mathbb{R}^+$ mais une probabilité nulle d'être dans $\mathbb{R}^- \times \mathbb{R}^+$.

Exercice 7 Soit K une matrice symétrique positive. Expliquer comment simuler un vecteur gaussien centré de matrice de covariance K à partir de variables gaussiennes indépendantes.

<u>*Solution de l'exercice 7*</u> Soit d la taille de la matrice K et soit $Z = (Z_1, \ldots, Z_d)$, où les Z_i sont i.i.d. de loi $\mathcal{N}(0,1)$. Notons que Z est un vecteur gaussien de matrice de covariance identité. On cherche à construire notre vecteur de manière linéaire à partir de Z. Plus précisément, soit $A = (a_{i,j})_{1\le i,j\le d}$ une matrice $d \times d$ et soit $X = AZ$. Notons que toute combinaison linéaire des coordonnées de X est une combinaison linéaire des coordonnées de Z, donc est gaussienne, donc X est un vecteur gaussien.

De plus, pour tous $1 \le i,j \le d$, on a

$$\begin{aligned}
\mathbb{E}[X_iX_j] &= \mathbb{E}\left[\left(\sum_k a_{i,k}Z_k\right)\left(\sum_\ell a_{j,\ell}Z_\ell\right)\right] \\
&= \sum_{k,\ell} a_{i,k}a_{j,\ell}\,\mathbb{E}\left[Z_kZ_\ell\right] \\
&= \sum_{k,\ell} a_{i,k}a_{j,\ell}\,\mathbb{1}_{k=\ell} \\
&= \sum_k a_{i,k}a_{j,k},
\end{aligned}$$

donc la matrice de covariance de X est A^tA. Il nous suffit donc de trouver A telle que $A^tA = K$. Or, comme K est symétrique positive, il existe une matrice orthogonale Q telle que $K = {}^tQDQ$, où D est une matrice diagonale de coefficients diagonaux $d_1, \ldots, d_n \ge 0$. On peut alors prendre $A = {}^tQ\sqrt{D}Q$, où $\sqrt{D}$ est la matrice diagonale de coefficients diagonaux $\sqrt{d_1}, \ldots, \sqrt{d_n}$.

Exercice 8 (Formule de Wick) Soit $(X_1, \ldots, X_d)$ un vecteur gaussien centré. Montrer que

$$\mathbb{E}\left[\prod_{i=1}^d X_i\right] = \sum_\pi \prod_{\{a,b\}\in\pi} \mathbb{E}\left[X_aX_b\right],$$

où la somme se fait sur toutes les partitions π de $[\![1,d]\!]$ en parties de taille 2.

<u>*Solution de l'exercice 8*</u> Soit Z un vecteur gaussien de matrice de covariance identité. Alors il existe une matrice A telle que X a la même loi que AZ (cf. exercice 7). La formule qu'on veut montrer peut donc se mettre sous la forme

$$\mathbb{E}\left[\prod_{i=1}^d f_i(Z)\right] = \sum_\pi \prod_{\{a,b\}\in\pi} \mathbb{E}\left[f_a(Z)f_b(Z)\right],$$

où les f_i sont des formes linéaires sur $\mathbb{R}^d$. Or, cette dernière formule est linéaire en chacune des f_i, donc il suffit de la prouver dans le cas où chaque f_i est de la forme $z \to z_j$. Autrement dit, il suffit de montrer que pour tous $j_1, \ldots, j_k$, on a

$$\mathbb{E}\left[\prod_{i=1}^d Z_{j_i}\right] = \sum_\pi \prod_{\{a,b\}\in\pi} \mathbb{E}\left[Z_{j_a}Z_{j_b}\right]. \tag{1}$$

Pour tout $j \in [\![1,d]\!]$, on note k_j le nombre de i tels que $j_i = j$. Alors par indépendance des Z_j, le membre de droite vaut

$$\prod_{j=1}^d \mathbb{E}\left[Z^{k_j}\right],$$

où Z est une variable gaussienne centrée de variance 1. Remarquons tout d'abord que si un des k_j est impair, on a $\mathbb{E}[Z^{k_j}] = 0$ donc le membre de gauche dans (1) est nul. D'autre part, si π est une partition de $[\![1, d]\!]$ en paires, une des paires contient un nombre impair de i tel que $j_i = j$. Soit $\{a, b\}$ cette paire. Alors $j_a \neq j_b$, donc $\mathbb{E}[Z_{j_a} Z_{j_b}] = 0$, donc la contribution de π est nulle. Comme c'est vrai pour tout π, le membre de droite est aussi nul. On suppose donc désormais que tous les k_j sont pairs, et on écrit $k_j = 2\ell_j$.

Les moments de Z se calculent par exemple par intégration par partie, et on obtient $\mathbb{E}\left[Z^{2\ell}\right] = \frac{(2\ell)!}{2^\ell \ell!}$ pour tout $\ell \in \mathbb{N}$, donc le membre de gauche de (1) vaut

$$\prod_{j=1}^{d} \frac{(2\ell_j)!}{2^{\ell_j} \ell_j!}.$$

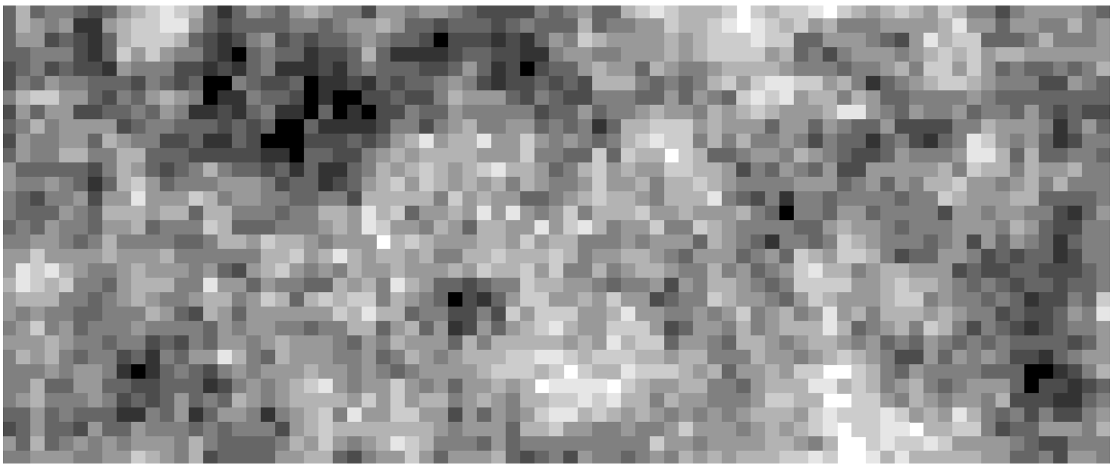
Il reste à étudier le membre de droite de (1). Soit π une partition en paires et $\{a, b\} \in \pi$. Si $j_a \neq j_b$, alors Z_{j_a} et Z_{j_b} sont indépendantes donc $\mathbb{E}[Z_{j_a} Z_{j_b}] = 0$ et la contribution de π s'annule. Si $j_a = j_b$ pour tout $\{a, b\} \in \pi$, alors la contribution de π vaut 1, donc il suffit de compter le nombre de π qui contribuent. Choisir un tel π revient, pour tout j, à partitionner en paires les i tels que $j_i = j$. Or, le nombre de partitions en paires d'un ensemble à 2ℓ éléments vaut $(2\ell - 1) \times (2\ell - 3) \times \cdots \times 3 \times 1 = \frac{(2\ell)!}{2^\ell \ell!}$, donc le nombre de π qui contribuent à la somme vaut

$$\prod_{j=1}^{d} \frac{(2\ell_j)!}{2^{\ell_j} \ell_j!},$$

d'où (1), d'où le résultat.

4 Jolie image

Exercice 9 Que représente la jolie image ci-dessous ?



<u>*Solution de l'exercice 9*</u> Il s'agit d'un champ libre gaussien (Gaussian free field, ou GFF, en anglais), qui est l'analogue naturel d'une marche aléatoire indexée par un graphe quelconque (au lieu de $\mathbb{Z}$). Étant donné un graphe $G = (V, E)$ et un sommet distingué $0 \in V$, le champ libre gaussien $(h_x)_{x \in V}$ sur G est défini comme le vecteur aléatoire tel que $h_0 = 0$, et dont la loi a une densité

$$\frac{1}{Z_G} \exp\left(-\frac{1}{2} \sum_{\{x,y\} \in E} (h_x - h_y)^2 \right)$$

par rapport à la mesure de Lebesgue sur $\mathbb{R}^{V \setminus \{0\}}$. Ici, le graphe G est un tore de dimension 2, de hauteur 32 et largeur 77. Chaque petit carré correspond donc à un sommet et plus h_x est grand, plus le carré correspondant à x est foncé. Pour une image encore plus jolie du même objet, voir par exemple ici :

https://www.math.ucla.edu/~biskup/PIMS/JPGs/GFF.jpg

Si on fixe la dimension et qu'on fait tendre la taille n du tore vers l'infini, cet objet a un comportement un peu surprenant :

- si $d = 1$, alors h ressemble à une marche aléatoire, donc les h_x sont d'ordre $\sqrt{n}$,
- si $d = 2$, la plupart des h_x sont d'ordre $\sqrt{\log n}$, et les plus grands d'ordre $\log n$,
- si $d \geq 3$, la plupart des h_x sont d'ordre $O(1)$, et les plus grands d'ordre $\sqrt{\log n}$.