X2016 – MAP 311

PC 1 – mardi 18 avril 2017 – Espaces probabilisés

Corrigé des questions non abordées en PC

Igor Kortchemski – igor.kortchemski@cmap.polytechnique.fr

Corrigé des exercices non traités sur http://www.normalesup.org/~kortchem/MAP311 *un peu après la PC.*

2 Lemmes de Borel–Cantelli

Exercice 5. **(Limite supérieure d'ensembles)** Soit $(A_n)_{n\geq 1}$ une suite d'événements d'un espace probabilisable $(\Omega, \mathcal{A})$. On pose

$$\limsup_{n\to\infty} A_n := \bigcap_{p\geq 1}\left(\bigcup_{n\geq p} A_n\right).$$

(1) Que représente l'événement $\limsup_{n\to\infty} A_n$?

(2) Si $A = \mathbb{R}$, donner $\limsup_{n\to\infty} A_n$ dans les trois cas suivants :

(a) $A_n = [-1/n, 3+1/n]$

(b) $A_n = [-2-(-1)^n, 2+(-1)^n[$

(c) $A_n = p_n\mathbb{N}$, où $(p_n)_{n\geq 1}$ est la suite des nombres premiers et $p_n\mathbb{N}$ est l'ensemble des multiples de p_n,

(3) Comparer $\limsup_{n\to\infty}(A_n \cup B_n)$ et $\limsup_{n\to\infty}(A_n \cap B_n)$ avec $\limsup_{n\to\infty} A_n$ et $\limsup_{n\to\infty} B_n$.

(4) Que représente l'événement $\liminf_{n\to\infty} A_n := \bigcup_{p\geq 1}\left(\bigcap_{n\geq p} A_n\right)$?

Corrigé :

(2) (b) On a $A_n = [-3, 3[$.

(4) Il s'agit de l'événement où tous les A_n sont réalisés à partir d'un certain rang.

□

Rappel des lemmes de Borel-Cantelli.

∗ **Premier lemme de Borel-Cantelli.** Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilité et $(A_n)_{n\geq 1}$ une suite d'événements. On suppose que $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$. Alors $\mathbb{P}(\limsup_{n\to\infty} A_n) = 0$. En d'autres termes, de manière équivalente :

- presque sûrement, A_n n'est réalisé qu'un nombre fini de fois ;
- presque sûrement, A_n n'est pas réalisé à partir d'un certain rang ;
- avec probabilité 0, A_n est réalisé une infinité de fois.

∗ **Deuxième lemme de Borel-Cantelli.** Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilité et $(A_n)_{n\geq 1}$ une suite d'événements **indépendants**. On suppose que $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$. Alors $\mathbb{P}(\limsup_{n\to\infty} A_n) = 1$. En d'autres termes, de manière équivalente :

- presque sûrement, A_n est réalisé une infinité de fois.
- avec probabilité 0, A_n n'est pas réalisé à partir d'un certain rang.

Exercice 6. **(Marche aléatoire vue de l'origine)** Soit $(Y_n)_{n\geq 1}$ une suite de variables aléatoires de Bernoulli indépendantes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ telles que $\mathbb{P}(Y_n = 1) = p$ et $\mathbb{P}(Y_n = -1) = 1-p$ avec $0 < p < 1$ et $p \neq 1/2$. On considère la marche aléatoire $Z_n = Y_1 + Y_2 + \cdots + Y_n$ (avec $Z_0 = 0$). On note $A_n = \{Z_n = 0\}$.

Pour des questions, demande d'explications etc., n'hésitez pas à m'envoyer un mail.

(1) Que représente l'événement $\limsup_{n\to\infty} A_n$?

(2) Montrer que $\mathbb{P}(\limsup_{n\to\infty} A_n) = 0$.

(3) Lorsque $p = 1/2$, on peut montrer en utilisant la formule de Stirling que

$$\mathbb{P}(Z_{2n} = 0) \underset{n\to\infty}{\sim} \frac{1}{\sqrt{\pi n}}.$$

Peut-on en déduire directement que $\mathbb{P}(\limsup_{n\to\infty} A_n) = 1$?

Corrigé :

(1) $\limsup_{n\to\infty} A_n$ représente l'ensemble des issus pour lesquelles le trajectoires de Z_n qui repassent une infinité de fois par 0.

2) D'après le (premier) lemme de Borel-Cantelli, il suffit de montrer que $\sum \mathbb{P}(A_n)$ converge. Or $\mathbb{P}(A_n)$ est nul si n est impair et

$$\mathbb{P}(A_{2n}) = \mathbb{P}(Z_{2n} = 0) = \binom{2n}{n} p^n (1-p)^n.$$

Or $\mathbb{P}(A_{2n+2})/\mathbb{P}(A_{2n}) = \frac{(2n+1)(2n+1)}{(n+1)^2} p(1-p) \to 4p(1-p)$ quand $n \to \infty$. Or $4p(1-p) < 1$ puisque $p \neq 1/2$, d'où la convergence de la série.

Remarque. Plus simplement, on peut majorer $\binom{2n}{n}$ par 2^{2n}, de sorte que $\mathbb{P}(A_{2n}) \leq (4p(1-p))^n$ et on conclut de manière similaire.

(3) D'après l'équivalent de l'énoncé, la série de terme général $\mathbb{P}(A_n)$ diverge. Cependant, on ne peut pas appliquer le (deuxième) lemme de Borel Cantelli directement, car les événements A_n ne sont pas indépendants.

Remarque. Lorsque $p = 1/2$, il est cependant vrai que $\mathbb{P}(\limsup_{n\to\infty} A_n) = 1$. Pour le démontrer, on peut procéder comme suit. Notons τ_n le n-ième passage de la marche aléatoire en 0 (τ_n est une variable aléatoire qui vaut $+\infty$ si la marche passe strictement moins de n fois en 0, et $\tau_n < \infty$ si et seulement si la marche repasse n fois par 0), et notons X le nombre total de passages en 0 par la marche aléatoire. Alors

$$X = 1 + \sum_{n\geq 1} \mathbb{1}_{Z_{2n}=0} = 1 + \sum_{n\geq 1} \mathbb{1}_{\tau_n<\infty}.$$

En passant à l'espérance et en utilisant la linéarité de l'espérance pour des variables aléatoires positives, on obtient

$$\infty = \sum_{n\geq 1} \mathbb{P}(Z_{2n} = 0) = \sum_{n\geq 1} \mathbb{P}(\tau_n < \infty).$$

Or on remarque que pour tout $n \geq 2$, $\mathbb{P}(\tau_n < \infty | \tau_{n-1} < \infty) = \mathbb{P}(\tau_1 < \infty)$. En effet, si la marche est revenue en 0 pour la n-ième fois, la probabilité qu'elle repassera en 0 est précisément la probabilité qu'une marche issue de 0 repasse en 0 (car elle y « redémarre » ; une justification précise sera donnée dans le cours de MAP 432 en utilisant la propriété de Markov forte). Donc $\mathbb{P}(\tau_n < \infty) = \mathbb{P}(\tau_n < \infty, \tau_{n-1} < \infty) = \mathbb{P}(\tau_{n-1} < \infty)\mathbb{P}(\tau_1 < \infty)$. Par récurrence, on en déduit que $\mathbb{P}(\tau_n < \infty) = \mathbb{P}(\tau_1 < \infty)^n$, de sorte que

$$\infty = \sum_{n\geq 1} \mathbb{P}(\tau_1 < \infty)^n = \frac{1}{1 - \mathbb{P}(\tau_1 < \infty)}.$$

Donc $\mathbb{P}(\tau_1 < \infty) = 1$ et donc $\mathbb{P}(\tau_n < \infty) = 1$ pour tout $n \geq 1$. Or, par propriété de limite décroissante,

$$\mathbb{P}(\text{la marche passe une infinité de fois en 0}) = \lim_{n\to\infty} \mathbb{P}(\text{la marche repasse au moins } n \text{ fois en 0}) = 1.$$

□

4 Plus appliqué

Exercice 8. (**Un petit calcul**) Combien de fois faut-il, en moyenne, lancer une pièce en l'air pour observer une suite d'un nombre impair de piles, suivi d'un face ?

Corrigé : Supposons que x soit la réponse. On considère les premiers lancers de pièces. Si on a F ou PP, le jeu recommence avec la même probabilité, alors que si on a PF, le jeu s'arrête. On a alors

$$x = \frac{1}{2}(x+1) + \frac{1}{4}(x+2) + \frac{2}{4}.$$

La résolution de ce système donne $x = 6$ ou $x = \infty$.

On peut vérifier que $x < \infty$. En effet, si a_n désigne le nombre de manières de lancer une pièce n fois pour qu'un nombre impair de piles suivi d'un face arrive pour la première fois à l'instant n, on a $a_n = a_{n-2} + a_{n-2}$ pour $n \geq 2$. On démontre ensuite aisément par récurrence sur n que $a_n \leq 1.9^n$ pour $n \geq 2$, ce qui implique que

$$x = \sum_{n\geq 2} n \cdot \frac{a_n}{2^n} \leq \sum_{n\geq 2} n \cdot \frac{1.9^n}{2^n} < \infty.$$

Donc $x = 6$.

□

Exercice 9. (**Temps d'attente**) Anne retourne une à une les cartes d'un jeu de 52 cartes bien mélangé, posé face caché sur une table, jusqu'à trouver un as. Combien de cartes aura-t-on vu en moyenne ?

Corrigé : On place toutes les cartes (face visible) en cercle sur la table (de sorte qu'on voit toutes les cartes), et on ajoute un cinquième as fictive entre la première carte du paquet et la dernière carte du paquet. Les 53 cartes sont ainsi partitionnées en 5 intervalles qui ont même loi (chaque intervalle commençant par la carte suivant un as et finissant par un as), de sorte que le nombre moyen de cartes dans le premier intervalle est 53/5. Le nombre de cartes qu'on aura vu étant le nombre de cartes entre le cinquième as (exclus) et le premier as (inclus), on aura donc vu en moyenne 53/5 cartes. □

Exercice 10. (**Piles consécutifs**) Andrea lance une infinité de fois une pièce équilibrée. Pour $n \geq 1$, on note L_n la plus longue séquence de « piles » consécutifs dans les n premiers lancers. Le but de cet exercice est d'étudier le comportement de L_n quand $n \to \infty$. Par exemple :

n	1	2	3	4	5	6	7	8	9	$\cdots$
tirage	F	P	P	F	P	P	P	F	P	$\cdots$
L_n	0	1	2	2	2	2	3	3	3	$\cdots$

Pour modéliser ce problème, Andrea considère $X_1, X_2, \ldots$ des variables aléatoires indépendantes et de même loi telles que $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = 0) = 1/2$. Posons :

$$L_n := \max\{k \geq 1;\ \text{il existe } i \leq n-k \text{ tel que } X_{i+1} = \cdots = X_{i+k} = 1\}.$$

(1) Montrer que pour tout $j \in \{1, \ldots, n\}$ on a $\mathbb{P}(L_n \geq j) \leq \frac{n-j+1}{2^j} \leq \frac{n}{2^j}$.

(2) Montrer que pour tout $j \in \{1, \ldots, n\}$ on a $\mathbb{P}(L_n \leq j) \leq \left(1 - \frac{1}{2^j}\right)^{\lfloor n/j \rfloor}$.

(3) Trouver une suite (a_n) de nombre réels positifs tels que pour tout $\epsilon \in]0,1[$, on a

$$\mathbb{P}\left(1 - \epsilon < \frac{L_n}{a_n} < 1 + \epsilon\right) \underset{n\to\infty}{\longrightarrow} 1.$$

(4) Montrer que pour tout $\epsilon \in]0,1[$, on a $\mathbb{P}\left(\frac{L_n}{a_n} > 1 - \epsilon, \text{ à partir d'un certain rang}\right) = 1$. Montrer que pour tout $\epsilon \in]0,1[$, on a $\mathbb{P}\left(\frac{L_n}{a_n} < 1 + \epsilon, \text{ à partir d'un certain rang}\right) = 1$.

(5) En déduire que $\mathbb{P}\left(\frac{L_n}{a_n} \to 1 \text{ lorsque } n \to \infty\right) = 1$. Autrement dit, presque sûrement, $\frac{L_n}{a_n}$ converge vers 1.

Corrigé :

(1) Pour $j \geq 1$, on a

$$\begin{aligned} \mathbb{P}(L_n \geq j) &= \mathbb{P}\left(\bigcup_{i=0}^{n-j}\{X_{i+1} = \cdots = X_{i+j} = 1\}\right) \\ &\leq \sum_{i=0}^{n-j} \mathbb{P}(\{X_{i+1} = \cdots = X_{i+j} = 1\}) = \sum_{i=0}^{n-j} \frac{1}{2^j} = \frac{n-j+1}{2^j} \leq \frac{n}{2^j}. \end{aligned}$$

(2) Pour $1 \leq i \leq \lfloor n/j \rfloor$, on pose

$$A_i := \{X_{(i-1)j+1} = X_{(i-1)j+2} = \cdots = X_{ij} = 1\}, \qquad 1 \leq i \leq \lfloor n/j \rfloor.$$

Alors $\cup_{i=1}^{\lfloor n/j \rfloor} A_i \subset \{L_n \geq j\}$. Puisque les événements $A_i, 1 \leq i \leq \lfloor n/j \rfloor$ sont indépendants, ceci entraîne que

$$\mathbb{P}(L_n < j) \leq \mathbb{P}\left(\bigcap_{i=1}^{\lfloor n/j \rfloor} A_i^c\right) = \prod_{i=1}^{\lfloor n/j \rfloor} \mathbb{P}(A_i^c) = \mathbb{P}(A_1^c)^{\lfloor n/j \rfloor} = \left(1 - \frac{1}{2^j}\right)^{\lfloor n/j \rfloor},$$

(3) Il eſt naturel de prendre $a_n = \ln(n)/\ln(2)$. En effet, d'après (1), on a alors

$$\mathbb{P}\left(\frac{L_n}{a_n} \geq 1 + \epsilon\right) \leq \frac{n}{2^{(1+\epsilon)\frac{\ln(n)}{\ln(2)}}} = \frac{1}{n^\epsilon} \quad \underset{n\to\infty}{\longrightarrow} \quad 0$$

D'après (2), on a

$$\mathbb{P}\left(\frac{L_n}{a_n} \leq 1 - \epsilon\right) \leq \left(1 - \frac{1}{n^{1-\epsilon}}\right)^{\frac{\ln(2)}{(1-\epsilon)} \cdot \frac{n}{\ln(n)}} \quad \underset{n\to\infty}{\longrightarrow} \quad 0$$

(pour simplifier on a fait comme si $j = a_n(1-\epsilon)$ et n/j étaient entiers, sinon il faut rajouter des ± 1 par ci par là).

Ainsi, en passant au complémentaire,

$$\mathbb{P}\left(1 - \epsilon \leq \frac{L_n}{a_n} \leq 1 + \epsilon\right) = 1 - \mathbb{P}\left(1 - \epsilon > \frac{L_n}{a_n} \text{ ou } \frac{L_n}{a_n} > 1 + \epsilon\right),$$

et

$$\mathbb{P}\left(1 - \epsilon > \frac{L_n}{a_n} \text{ ou } \frac{L_n}{a_n} > 1 + \epsilon\right) \leq \mathbb{P}\left(1 - \epsilon > \frac{L_n}{a_n}\right) + \mathbb{P}\left(\frac{L_n}{a_n} > 1 + \epsilon\right) \quad \underset{n\to\infty}{\longrightarrow} \quad 0$$

d'après la discussion précédente, ce qui permet de conclure que $\mathbb{P}\left(1 - \epsilon \leq \frac{L_n}{a_n} \leq 1 + \epsilon\right) \to 1$.

(4) On remarque que

$$\left(1 - \frac{1}{n^{1-\epsilon}}\right)^{\frac{\ln(2)}{(1-\epsilon)} \cdot \frac{n}{\ln(n)}} \quad \underset{n\to\infty}{\sim} \quad e^{-\frac{\ln(2)}{1-\epsilon} \cdot \frac{n^\epsilon}{\ln(n)}}.$$

Donc $\sum_{n \geq 1} \mathbb{P}\left(\frac{L_n}{a_n} \leq 1 - \epsilon\right) < \infty$. D'après le (premier) lemme de Borel-Cantelli, p.s., pour tout n suffisamment grand, $L_n/a_n > 1 - \epsilon$.

L'autre sens eſt plus délicat, car la série de terme général $1/n^\epsilon$ diverge. Pour remédier à cela, posons $n = n_m = \lfloor m^{2/\epsilon} \rfloor$ de sorte que $\sum_m \mathbb{P}(L_{n_m} \geq (1+\epsilon)a_{n_m}) < \infty$. D'après le (premier) lemme de Borel-Cantelli, p.s., pour tout m suffisamment grand on a $L_{n_m} \leq (1+\epsilon)\ln(n_m)/\ln(2)$. Or, si m eſt suffisamment grand et $n \in [n_{m-1}, n_m[$, on a

$$L_n \leq L_{n_m} < (1+\epsilon)\frac{\ln(n_m)}{\ln(2)} \leq (1+2\epsilon)\frac{\ln(n_{m-1})}{\ln(2)} \leq (1+2\epsilon)\frac{\ln(n)}{\ln(2)},$$

d'où le résultat.

(5) Le but de cette question est essentiellement de pouvoir intervertir le « pour tout $\epsilon > 0$ »et la probabilité $\mathbb{P}$. Pour cela, on remarque que

$$\mathbb{P}\left(\frac{L_n}{a_n} \to 1\right) = \mathbb{P}\left(\text{pour tout } p \geq 1, \text{ à partir d'un certain rang on a } 1 - \frac{1}{p} < \frac{L_n}{a_n} < 1 + \frac{1}{p}\right).$$

Or, à $p \geq 1$ fixé, on a $\mathbb{P}\left(\text{à partir d'un certain rang on a } 1 - \frac{1}{p} < \frac{L_n}{a_n} < 1 + \frac{1}{p}\right) = 1$ (en prenant $\epsilon = 1/p$ dans la question (4)). On conclut en utilisant la question (3) de l'exercice 1.

□

5 Pour aller plus loin

Exercice 11. **(Approximation diophantienne et Borel-Cantelli)** On admet l'existence d'une tribu $\mathcal{A}$ sur $[0,1]$ contenant tous les intervalles inclus dans $[0,1]$ (c'est la tribu borélienne) et d'une mesure de probabilité $\mathbb{P}$ sur $([0,1], \mathcal{A})$ telle que $\mathbb{P}([a,b]) = b - a$ pour tous $0 \leq a \leq b \leq 1$ (c'est la mesure de Lebesgue).

(1) Soit $\epsilon > 0$ fixé. Montrer que

$$\mathbb{P}\left(\left\{x \in [0,1] : \exists \text{ un nombre } \textbf{infini} \text{ de rationnels } p/q \text{ avec } \operatorname{pgcd}(p,q) = 1 \text{ tq } \left|x - \frac{p}{q}\right| \leq \frac{1}{q^{2+\epsilon}}\right\}\right) = 0.$$

Ainsi, « presque tout » x est « mal approchable par des rationnels à l'ordre $2 + \epsilon$ ».

Indication. Pour tout $q \geq 1$, on pourra considérer

$$A_q := [0,1] \cap \bigcup_{p=0}^{q} \left[\frac{p}{q} - \frac{1}{q^{2+\epsilon}}, \frac{p}{q} + \frac{1}{q^{2+\epsilon}}\right].$$

(2) Montrer que

$$\mathbb{P}\left(\left\{x \in [0,1] : \exists \text{ un nombre } \textbf{infini} \text{ de rationnels } p/q \text{ avec } \operatorname{pgcd}(p,q) = 1 \text{ tq } \left|x - \frac{p}{q}\right| \leq \frac{1}{q^{2}}\right\}\right) = 1.$$

Ainsi, « presque tout » x est « bien approchable par des rationnels à l'ordre 2 ».

Corrigé :

(1) Pour tout $q \geq 1$, on note

$$A_q := [0,1] \cap \bigcup_{p=0}^{q} \left[\frac{p}{q} - \frac{1}{q^{2+\epsilon}}, \frac{p}{q} + \frac{1}{q^{2+\epsilon}}\right].$$

Ainsi, $\mathbb{P}\left(A_q\right) \leq 2/q^{1+\epsilon}$. Par conséquent,

$$\sum_{q \geq 1} \mathbb{P}\left(A_q\right) < +\infty.$$

D'après le lemme de Borel-Cantelli, $\mathbb{P}\left(\limsup_{q \to \infty} A_q\right) = 0$, or l'ensemble $\limsup_{q \to \infty} A_q$ contient l'ensemble des réels bien approchables par des rationnels à l'ordre $2 + \epsilon$. Voir http://en.wikipedia.org/wiki/Thue-Siegel-Roth_theorem

(2) Tout d'abord, on démontre le résultat suivant :

Théorème de Dirichlet. Soit α un nombre réel. Pour tout entier $N \geq 2$, il existe deux entiers p et q, avec $0 < q < N$ tels que $|q\alpha - p| < \frac{1}{N}$.
(dans cet énoncé, p et q ne sont pas nécessairement premiers entre eux)

Pour cela, en notant $\{x\}$ la partie fractionnaire de x et $\lfloor x \rfloor$ la partie entière de x, considérons les $N+1$ nombres $0, 1, \{\alpha\}, \{2\alpha\}, \ldots, \{(N-1)\alpha\}$ et les N intervalles $[0, 1/N], [1/N, 2/N], \ldots, [(N-1)/N, 1]$. D'après le principe des tiroirs, il existe deux de ces $N+1$ nombres appartenant au même tiroir.

Si l'un de ces nombres vaut 1 et l'autre $\{m\alpha\}$ avec $0 \leq m \leq N-1$, alors

$$|m\alpha - (\lfloor m\alpha \rfloor + 1)| = |\{m\alpha\} - 1| < \frac{1}{N}.$$

On peut donc prendre $q = m$ et $p = \lfloor m\alpha \rfloor + 1$.

Sinon, on peut trouver $0 \leq \ell < m \leq N-1$ tels que $|\{m\alpha\} - \{\ell\alpha\}| < 1/N$. Donc

$$|(m-\ell)\alpha - (\lfloor m\alpha \rfloor - \lfloor \ell\alpha \rfloor)| = |(m\alpha - \lfloor m\alpha \rfloor) - (\ell\alpha - \lfloor \ell\alpha \rfloor)| = |\{m\alpha\} - \{\ell\alpha\}| < \frac{1}{N},$$

et on peut prendre $q = m - \ell$ et $p = \lfloor m\alpha \rfloor - \lfloor \ell\alpha \rfloor$(on a bien $0 < q < N$).
En particulier, on a

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{qN} < \frac{1}{q^2}.$$

□

Ce résultat implique que si α est irrationnel, il un nombre infini de rationnels p/q avec $\mathrm{pgcd}(p,q) = 1$ tels que $\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^2}$. En effet, supposons que α est irrationnel. D'après le théorème de Dirichlet, il existe une infinité d'entiers p, q tels que

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{qN} < \frac{1}{q^2}.$$

Supposons par l'absurde que les fractions $\frac{p}{q}$ prennent un nombre fini de valeurs sous forme irréductible. Par extraction, il existe alors un couple (p,q) et suite (p_n, q_n) avec $q_n \to \infty$ telle que $\frac{p_n}{q_n} = \frac{p}{q}$ pour tout $n \geq 1$ et

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q_n^2}.$$

En passant à la limite, on trouve $\alpha = \frac{p}{q}$, absurde.

On peut maintenant répondre à la question posée. En effet, on vient de montrer que tout nombre irrationnel est bien approchable par des rationnels à l'ordre 2. Or l'ensemble des nombres irrationnels est de probabilité 1, car son complémentaire est dénombrable et donc de probabilité 0 pour la mesure de Lebesgue, ce qui conclut.

□