

Travaux dirigés de probabilités

Rémi Peyre et Frédéric Simon

Printemps 2009

Voici les énoncés des exercices que nous avons donnés en travaux dirigés dans le cadre du cours « Probabilités » de Christophe Sabot pour les élèves de L3 de l'ÉNS Lyon au second trimestre de l'année universitaire 2008/09. Chaque semaine, la séance de TD aborde les thèmes traités en cours le matin même. À l'intérieur de chaque feuille, nous avons tenté de ranger les exercices par difficulté croissante.

Afin de se repérer plus facilement dans les exercices, nous avons inséré quelques repères dont voici la légende :

- Concernant la difficulté des questions :
 - Une étoile (★) indique une question plus difficile que les autres ;
 - Deux étoiles (★★) indiquent une question très difficile, que ceux qui veulent peuvent faire chez eux.
- Concernant l'intérêt pédagogique des exercices :
 - Un cœur (♥) indique un résultat particulièrement intéressant du point de vue de la culture mathématique ;
 - Un carreau (◇) indique un résultat qu'il est particulièrement utile de connaître quand on fait des probabilités.

À vos neurones !

les moniteurs

Feuille 1 : Théorie de la mesure

19 janvier 2009

Exercice 1 (Tribus)

1. Soit $\mathcal{F}$ une tribu de Ω et $B \subset \Omega$, pas nécessairement mesurable. Montrer que la classe de parties de B définie par $\{A \cap B, A \in \mathcal{F}\}$ est une tribu de B , que l'on appelle *tribu trace* de $\mathcal{F}$ sur B .

2. Montrer que toute intersection dénombrable de tribus (sur un même ensemble) est une tribu. Peut-on enlever « dénombrable » ?

3. (★) Montrer que la réunion croissante d'une suite de tribus n'est pas forcément une tribu. (On pourra chercher un contre-exemple avec $\Omega = \mathbb{N}$).

Exercice 2 (Théorème d'Ulam)

☛ Dans tout cet exercice, le recours à l'axiome du choix est autorisé.

1. Soit $(X, \mathcal{F})$ un ensemble mesurable non vide. Montrer qu'il existe toujours au moins une mesure de probabilité sur $(X, \mathcal{F})$.

Dans la suite de l'exercice, on admet l'existence d'un ensemble totalement ordonné ω_1 , non dénombrable, tel que pour tout $x \in \omega_1$ l'ensemble $\{y \in \omega_1; y < x\}$ est dénombrable.

2. Montrer qu'il existe une fonction $f : \omega_1 \times \omega_1 \longrightarrow \mathbb{N}$ telle que $x < y < z \Rightarrow f(x, z) \neq f(y, z)$. À votre avis, peut-on construire une telle fonction explicitement ?

3. À partir de la fonction f , construire des ensembles F_x^n , pour $(x, n) \in \omega_1 \times \mathbb{N}$, tels que le tableau

$$\begin{array}{cccc} F_x^0 & F_y^0 & F_z^0 & \dots \\ F_x^1 & F_y^1 & F_z^1 & \dots \\ F_x^2 & F_y^2 & F_z^2 & \dots \\ \dots & \dots & \dots & \dots \end{array} \quad (1)$$

dont les lignes sont indexées par $\mathbb{N}$ et les colonnes par ω_1 , vérifie les propriétés suivantes :

- Chaque ligne est constitué d'ensemble deux à deux disjoints ;
- Pour tout $x \in \omega_1$, la réunion des ensembles dans la colonne d'indice x est $\{y; y > x\}$.

Nous allons maintenant démontrer qu'il n'existe aucune mesure de probabilité sur $(\omega_1, \mathfrak{P}(\omega_1))$ qui soit diffuse, c.à.d. qui donne une masse nulle à chaque singleton. Pour cela, nous allons nous donner une probabilité μ sur $(\omega_1, \mathfrak{P}(\omega_1))$ et montrer qu'elle a

nécessairement un atome, c.à.d. un point de masse non nulle.

4. ($\diamond, \star$) Pour $(\varepsilon_i)_{i \in I}$ une famille non dénombrable de réels strictement positifs, montrer qu'il existe nécessairement $J \subset I$ dénombrable tel que $\sum_{i \in J} \varepsilon_i = +\infty$.

5. En déduire que dans chaque ligne du tableau (1), il n'y a qu'un nombre dénombrable de F_x^n dont la mesure par μ est non nulle.

6. Montrer qu'il existe alors une colonne du tableau (1) dans laquelle tous les F_x^n sont de μ -mesure nulle.

Soit maintenant x_0 l'indice d'une colonne vérifiant la propriété de la question 6.

7. Montrer que $\mu(\{x \in \omega_1 ; x \leq x_0\}) = 1$.

8. En déduire l'existence d'un atome pour μ .

L'hypothèse du continu affirme que $[0, 1[$ est en bijection avec ω_1 . K. Gödel a démontré en 1938 que cette hypothèse ne risquait pas de créer de contradiction avec le reste de la théorie des ensembles.

9. ($\heartsuit$) Montrer que, si on accepte l'hypothèse du continu, alors il n'existe aucune mesure sur $([0, 1[, \mathfrak{P}([0, 1[))$ qui étende la mesure de Lebesgue.

Exercice 3 (Nombres de Liouville)

Un réel x est qualifié de nombre de Liouville si, pour tout $n \geq 1$, il est possible de trouver deux entiers p et q avec $q \geq 2$ tels que

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^n}. \quad (1)$$

J. Liouville a démontré en 1844 qu'un nombre de Liouville irrationnel est nécessairement transcendant (c.à.d. que ce n'est la racine d'aucun polynôme à coefficients entiers). On notera L l'ensemble des nombres de Liouville.

1. ($\heartsuit, \star$) Montrer le théorème de Liouville.

2. Expliquer pourquoi L est un G_δ -dense de $\mathbb{R}$, c.à.d. une intersection dénombrable d'ouverts denses.

3. En déduire l'existence d'une infinité non dénombrable de nombres de Liouville irrationnels.

4. ($\star$) Donner explicitement un exemple de nombre de Liouville irrationnel.

5. Démontrer qu'en fait, L a la puissance du continu, c.à.d. qu'on peut y injecter $\mathbb{R}$.

6. Pourquoi L est-il borélien ?

7. Soit U_n l'ensemble des points de $[0, 1[$ vérifiant (1) pour une valeur de n fixée. Pour $n \geq 3$, majorer la mesure de Lebesgue de U_n .

8. En déduire que la mesure de Lebesgue de $L \cap [0, 1[$ est nulle. Expliquer pourquoi on peut même conclure que L tout entier est de mesure nulle.

9. Soit $\overline{\mathbb{Q}} \cap \mathbb{R}$ l'ensemble des nombres algébriques réels. Pourquoi $\overline{\mathbb{Q}} \cap \mathbb{R}$ est-il borélien

et de mesure de Lebesgue nulle ?

10. En déduire l'existence d'un nombre transcendant qui n'est pas de Liouville.

11. (★) En utilisant l'exercice 6, définir explicitement un nombre transcendant qui n'est pas de Liouville.

Exercice 4 (Limite projective de mesures)

Soit $\Omega =]0, 1]$ muni de sa tribu borélienne $\mathcal{F}$. Sur Ω , on définit les tribus $\mathcal{F}_n = \sigma\{]0, 1/i], i \leq n\}$ pour tout $n \geq 1$.

1. Construire des mesures de probabilités P_n sur $(\Omega, \mathcal{F}_n)$ compatibles, i.e. $P_{n+1}(A) = P_n(A)$ pour tout $A \in \mathcal{F}_n$, telles qu'il n'existe aucune mesure de probabilité sur $(\Omega, \mathcal{F})$ dont les restrictions aux $\mathcal{F}_n$ soient les P_n .

Exercice 5 (Complétion des mesures)

Soit $(\Omega, \mathcal{F}, \mu)$ un espace probabilisé. $A \subset \Omega$ est dit négligeable pour μ s'il existe $B \supset A$ tel que $B \in \mathcal{F}$ et $\mu(B) = 0$. La tribu complétée de $\mathcal{F}$ pour μ est alors la tribu engendrée par les éléments de $\mathcal{F}$ et les ensembles négligeables. Nous la noterons $\overline{\mathcal{F}}^\mu$.

1. Pour $A \subset \Omega$, montrer que $A \in \overline{\mathcal{F}}^\mu$ si et seulement si

$$\sup_{\substack{B \subset A \\ B \in \mathcal{F}}} \mu(B) = \inf_{\substack{B \supset A \\ B \in \mathcal{F}}} \mu(B).$$

2. (◇) Montrer qu'il existe une unique extension de μ à $\overline{\mathcal{F}}^\mu$, et la décrire.

Exercice 6 (Régularité automatique)

Soit $d \geq 1$ un entier ; la tribu borélienne de $\mathbb{R}^d$ est notée $\mathcal{B}$.

1. Montrer que tout ouvert de $\mathbb{R}^d$ peut s'écrire comme une union dénombrable de compacts.

2. (◇,★) Soit $\mathbb{P}$ une probabilité sur $(\mathbb{R}^d, \mathcal{B})$. Montrer que pour tout $B \in \mathcal{B}$, pour tout $\varepsilon > 0$, on peut trouver un compact K et un ouvert U tels que $K \subset B \subset U$ et $\mathbb{P}(U) - \varepsilon \leq \mathbb{P}(B) \leq \mathbb{P}(K) + \varepsilon$.

3. Est-il vrai que tout borélien est réunion dénombrable de compacts, resp. intersection dénombrable d'ouverts ?

Exercice 7 (La mesure de Lebesgue et la topologie)

On travaille sur $\mathbb{R}$ muni de la mesure de Lebesgue, laquelle est notée λ .

1. Un ouvert de $\mathbb{R}$ de mesure finie est-il forcément borné ?
2. Un borélien de mesure strictement positive est-il forcément d'intérieur non vide ?
3. Un ouvert dense de $]0, 1[$ a-t-il forcément une mesure 1 ?
4. (★) Construire un borélien A de $\mathbb{R}$ tel que pour tout intervalle ouvert borné non vide I on ait $0 < \lambda(A \cap I) < \lambda(I)$.

Soit $(d_n)_{n \geq 0}$ une suite d'éléments de $]0, 1[$ et posons $K_0 = [0, 1]$. On définit la suite $(K_n)_{n \geq 0}$ de la manière suivante : connaissant K_n , qui est une réunion d'intervalles fermés disjoints, on définit K_{n+1} en retirant dans chacun des intervalles de K_n un intervalle ouvert centré au même point, de longueur d_n fois celle de l'intervalle. On pose $K = \bigcap_n K_n$.

5. Que vaut $\lambda(K)$?
6. (♥) À l'aide de la question 5, construire deux compacts de $[0, 1]$ homéomorphes, l'un de mesure nulle, l'autre de mesure strictement positive.

Exercice 8 (Construction d'un ensemble non borélien)

☞ Cet exercice est donné à titre de complément culturel. Sa résolution utilise massivement la théorie des ordinaux ; il est donc inutile d'essayer de faire l'exercice si vous ne connaissez pas cette théorie.

☞ Dans tout cet exercice, les objets construits doivent être parfaitement explicites.

On note ω_1 le plus petit ordinal non dénombrable, et c un ensemble ayant la puissance du continu, par exemple $c = \{0, 1\}^{\mathbb{N}}$.

1. Rappeler pourquoi ω_1 est égal à l'ensemble des ordinaux dénombrables. Montrer que tout ordinal dénombrable est isomorphe à une partie de $\mathbb{Q}$ pour l'ordre naturel de ce dernier, et en déduire une surjection explicite $c \twoheadrightarrow \omega_1$.

On définit par récurrence ordinale les ensembles $(\mathcal{B}_\alpha)_{\alpha \in \omega_1}$, qui sont des ensembles de parties de $\mathbb{R}$, par :

- $\mathcal{B}_0$ est l'ensemble des ouverts de $\mathbb{R}$;
 - $\mathcal{B}_{\alpha+1}$ est l'ensemble des complémentaires des éléments de $\mathcal{B}_\alpha$;
 - Pour λ un ordinal-limite, $\mathcal{B}_\lambda$ est l'ensemble des réunions d'ensembles appartenant à un des $\mathcal{B}_\alpha$ pour $\alpha < \lambda$.
2. (◇) Montrer qu'une partie de $\mathbb{R}$ est borélienne si et seulement si elle appartient à $\mathcal{B}_\alpha$ pour un $\alpha \in \omega_1$.

3. (★★) Construire une surjection explicite $c \twoheadrightarrow \mathcal{B}(\mathbb{R})$.
4. (◇) S'en servir pour construire une partie explicite de $\mathbb{R}$ qui n'est pas borélienne.
5. Dans quelle question a-t-on utilisé l'axiome du choix ? En remplaçant ω_1 par ω'_1 défini comme le plus petit ordinal stable par réunion dénombrable, refaire l'exercice sans recourir à cet axiome.

Feuille 2 : Variables aléatoires

26 janvier 2009

Exercice 1 (Une distraction du savant Cosinus)

N voyageurs ($N \geq 2$) s'apprêtent à monter dans un wagon contenant N places numérotées, chacun étant muni d'un billet avec réservation. Le premier passager n'est autre que le savant Cosinus, qui, distrait comme il l'est toujours, ne regarde pas le numéro de sa réservation et prend une place au hasard. Les passagers suivants, quand ils entrent dans le train, s'asseyent alors à leur place réservée si celle-ci est encore libre, et sinon prennent une place au hasard parmi celles qui restent.

1. Calculer par récurrence sur N la probabilité que le dernier voyageur à entrer finisse assis à la place qu'il avait réservée.
2. (★) Retrouver le résultat de la question 1 par un argument direct.

Exercice 2 (Recensement des écureuils)

On veut estimer le nombre N d'écureuils dans une forêt. Pour cela on en capture k , on leur met une petite marque sur la patte et on les relâche. Une semaine après (on suppose qu'aucun écureuil n'est mort ou né dans l'intervalle) on en capture ℓ et on compte ceux d'entr'eux qui portent la marque.

1. Calculer la probabilité qu'on trouve m écureuils marqués en fonction de N , k et ℓ .
☞ La loi de m est appelée loi hypergéométrique de paramètres N , k et ℓ .
2. (♥) Calculer la valeur de N pour laquelle la probabilité d'observer une valeur m donnée est maximale. En déduire une estimation du nombre d'écureuils dans la forêt (le nombre d'écureuils évalué par cette méthode est appelé *estimateur du maximum de vraisemblance*).

Exercice 3 (Inégalité de la moyenne)

Soient $x_1, \dots, x_n$ des entiers strictement positifs. La moyenne arithmétique des x_i est $(x_1 + \dots + x_n)/n$, et leur moyenne géométrique est $(x_1 x_2 \dots x_n)^{1/n}$.

1. (◇) Montrer que la moyenne géométrique des x_i est toujours inférieure ou égale à leur moyenne arithmétique.
2. Quels sont les cas d'égalité de la question 1 ?

Pour $\alpha > 0$, on définit la α -moyenne des x_i comme

$$M_\alpha((x_i)_{1 \leq i \leq n}) = \sqrt[\alpha]{\frac{x_1^\alpha + \cdots + x_n^\alpha}{n}}. \quad (1)$$

3. Montrer que $M_\alpha((x_i)_i)$ est une fonction croissante de α , et déterminer les cas d'égalité.

4. Identifier la limite de $M_\alpha((x_i)_i)$ quand $\alpha \rightarrow 0$.

Dorénavant on appelle $M_0((x_i)_i)$ la limite trouvée à la question 4, et on étend la formule (1) aux $\alpha < 0$.

5. ($\heartsuit$) Montrer que $M_\alpha((x_i)_i)$ est une fonction croissante de α sur $\mathbb{R}$ tout entier.

6. Calculer les limites de $M_\alpha((x_i)_i)$ quand $\alpha \rightarrow +\infty$, resp. $\alpha \rightarrow -\infty$.

Exercice 4 (Densité de la mesure image)

Soit X une variable aléatoire réelle dont la loi μ possède une densité m par rapport à la mesure de Lebesgue. On considère une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$; le but de cet exercice est de calculer, si elle existe, la densité de la loi de $f(X)$.

1. On suppose f de classe $\mathcal{C}^1$ avec une dérivée strictement positive. Montrer que $f(X)$ a une densité par rapport à la mesure de Lebesgue et la calculer.

2. Même question en supposant que f est de classe $\mathcal{C}^1$ avec une dérivée strictement négative.

3. Même question avec $f(x) = |x|$.

4. ($\diamond$) Même question en supposant que $\mathbb{R}$ est recouvert par un ensemble dénombrable d'intervalles sur l'intérieur desquels f est $\mathcal{C}^1$ et de dérivée non nulle.

La loi du $\chi^2(1)$ est la loi de X^2 quand X est une variable aléatoire de loi gaussienne centrée réduite, i.e. de densité $m(x) = e^{-x^2/2}/\sqrt{2\pi}$.

5. Montrer que la loi du $\chi^2(1)$ est à densité et calculer cette densité.

6. ($\diamond, \star\star$) Construire une fonction f de classe $\mathcal{C}^1$ strictement croissante et une variable aléatoire X à densité telles que $f(X)$ ne soit pas à densité.

Exercice 5 (Simulation d'une variable aléatoire réelle)

Soit μ une loi de probabilité sur $\mathbb{R}$ et F sa fonction de répartition (avec la convention de continuité à droite). Pour $p \in]0, 1[$, on note

$$F^{-1}(p) = \inf\{x \in \mathbb{R}; F(x) \geq p\}.$$

1. Si U est une variable aléatoire uniforme sur $]0, 1[$, montrer que $F^{-1}(U)$ a pour loi μ .

2. Supposons ici que μ est diffuse, càd. que les singletons sont de mesure nulle. Si X est une variable aléatoire de loi μ , montrer que $F(X)$ suit la loi uniforme sur $]0, 1[$.

3. Imaginons que l'on sache simuler la loi uniforme sur $]0, 1[$. Comment simuler une loi exponentielle de paramètre λ ? Même question avec la loi de Cauchy γ , définie par $d\gamma(x) = 1/\pi(x^2 + 1)dx$.

Soit $a \geq 1$ un entier. On considère une variable aléatoire X_1 de loi uniforme sur $]0, 1[$, et on définit par récurrence les X_n , $n \geq 2$, par $X_n = aX_{n-1} - \lfloor aX_{n-1} \rfloor$, i.e. X_n est la partie fractionnaire de aX_{n-1} .

4. Montrer que tous les X_n ont la même loi. En déduire que, pour tout $k \geq 2$, la loi du vecteur aléatoire $(X_{m+1}, X_{m+2}, \dots, X_{m+k})$ ne dépend pas de m .

5. (★) Soient $I_1, \dots, I_k$ des intervalles de $]0, 1[$. Montrer que, quand $a \rightarrow \infty$,

$$\mathbb{P} \left[\forall n \in \{1, \dots, k\} \quad X_n \in I_n \right] \rightarrow \prod_{n=1}^k |I_n|,$$

où $|I|$ note la longueur de l'intervalle I .

6. (♡, ★) Expliquer du mieux possible le fonctionnement du générateur pseudo-aléatoire suivant (utilisé entre autres par le langage C) :

$$\begin{aligned} \text{SEED} &= (1103515245 * \text{SEED} + 12345) \text{ MOD } 2^{**31} \\ X &= \text{SEED} / 2^{**31}, \end{aligned}$$

où SEED est une valeur qui n'est jamais modifiée en-dehors des appels au générateur aléatoire, $**$ note l'exponentiation et X est censé être une variable aléatoire uniforme sur $[0, 1[$ dont chaque tirage est indépendant.

Exercice 6 (Mesures et moments)

La loi semi-circulaire est la loi de probabilité σ sur $\mathbb{R}$ définie par $d\sigma(x) = \mathbb{1}_{|x| \leq 2} \cdot (\sqrt{4 - x^2}/2\pi)dx$.

1. Calculer tous les moments de σ , càd. tous les $\int x^k d\sigma(x)$ pour $k \in \mathbb{N}^*$.

2. Montrer que toute mesure de probabilité sur $\mathbb{R}$ ayant les mêmes moments que σ est nécessairement à support compact, i.e. qu'il existe un compact dans lequel toute la masse de cette mesure est concentrée.

3. (◇) Montrer qu'une mesure de probabilité sur $\mathbb{R}$ à support compact (donc en particulier σ) est caractérisée par la suite de ses moments.

Soit μ la mesure signée sur $\mathbb{R}$ définie par $d\mu(x) = \mathbb{1}_{x \geq 0} \sin(\sqrt{3}x^{1/3})e^{-x^{1/3}}dx$.

4. Montrer que la valeur absolue de μ a tous ses moments finis, i.e. que pour tout $k \in \mathbb{N}$ on a $\int x^k |d\mu|(x) < \infty$.

5. (★) Calculer tous les moments de μ , i.e. tous les $\int x^k d\mu(x)$ pour $k \in \mathbb{N}$.

6. (♡) En déduire l'existence de deux mesures de probabilité distinctes sur $\mathbb{R}$ dont tous les moments sont finis et coïncident.

Exercice 7 (Événements extrêmes)

Pour $N \geq 0$, on considère N variables aléatoires indépendantes $X_1^N, \dots, X_N^N$ de lois uniformes sur $[0, 1]$. Pour $p \in [0, 1]$, on note $S^N(p)$ le nombre de i tels que $X_i^N \in [0, p]$.

1. Quelle est la loi de $S^N(p)$?
2. Pour $q \geq p$, déterminer la loi de $S^N(q) - S^N(p)$.
3. ($\diamond$) Soient $n_1, \dots, n_k$ des entiers de somme N et $p_1, \dots, p_k$ des réels positifs de somme 1. Calculer

$$\mathbb{P} [\forall j \in \{0, \dots, k\} \quad S^N(p_1 + \dots + p_j) = n_1 + \dots + n_j] .$$

☛ La loi de $(S^N(p_1), S^N(p_1 + p_2) - S^N(p_1), \dots, N - S^N(1 - p_k))$ est appelée loi k -nomiale de paramètres N et $(p_1, \dots, p_k)$.

4. ($\diamond$) Pour $s \geq 0$ un entier, $x \geq 0$ un réel, calculer

$$\pi_s(x) = \lim_{N \rightarrow \infty} \mathbb{P}[S^N(x/N) = s].$$

Vérifier que $\sum_{s \in \mathbb{N}} \pi_s(x) = 1$, et identifier la loi qui attribue une probabilité $\pi_s(x)$ à l'entier s .

On renumérote les X_i^N en les triant par ordre croissant : $X_{(1)}^N \leq X_{(2)}^N \leq \dots \leq X_{(N)}^N$ avec $\{X_{(i)}^N\}_{1 \leq i \leq N} = \{X_i^N\}_{1 \leq i \leq N}$.

5. Montrer que la loi de $NX_{(1)}^N$ converge, quand $N \rightarrow \infty$, vers une loi de probabilité limite sur $\mathbb{R}_+$, et identifier cette loi. (La notion de convergence des lois utilisée ici sera la convergence simple des fonctions de répartition).

6. ($\star$) Pour $s \geq 2$ un entier, montrer que la loi de $NX_{(s)}^N$ converge, quand $N \rightarrow \infty$, vers une loi de probabilité limite sur $\mathbb{R}_+$. Calculer la densité de cette loi par rapport à la mesure de Lebesgue.

7. ($\star\star$) Sauriez-vous décrire simplement la loi limite de $NX_{(s)}^N$ à partir d'une loi classique ?

Feuille 3 : Indépendance

2 février 2009

Exercice 1 (Regroupement par paquets)

Soient $\mathcal{B}_1, \dots, \mathcal{B}_n$ des tribus indépendantes. Soient $1 \leq i_1 < i_2 < \dots < i_k = n$ et notons $\mathcal{D}_1 = \sigma(\mathcal{B}_1, \dots, \mathcal{B}_{i_1})$, $\mathcal{D}_2 = \sigma(\mathcal{B}_{i_1+1}, \dots, \mathcal{B}_{i_2})$, $\dots$, $\mathcal{D}_k = \sigma(\mathcal{B}_{i_{k-1}+1}, \dots, \mathcal{B}_n)$. L'objectif de cet exercice est de montrer que $\mathcal{D}_1, \dots, \mathcal{D}_k$ sont des tribus indépendantes.

1. Montrer qu'il suffit de prouver le cas $k = 2$.

En vertu de la question 1, on suppose maintenant $k = 2$.

2. Soient D_1 et D_2 des événements de $\mathcal{D}_1$ et $\mathcal{D}_2$, respectivement de la forme $B_1 \cap \dots \cap B_{i_1}$ et $B_{i_1+1} \cap \dots \cap B_n$ pour des $B_m \in \mathcal{B}_m$. Montrer que D_1 et D_2 sont indépendants.

3. Même question que la question 2, sauf que cette fois-ci D_2 est un événement quelconque de $\mathcal{D}_2$.

4. ($\diamond$) Terminer la démonstration du lemme de regroupement par paquets.

Exercice 2 (Loi du 0-1 de Kolmogorov)

Soient $X_1, \dots, X_n, \dots$ des variables aléatoires indépendantes, et $\mathcal{F}$ la tribu qu'elles engendrent. On définit la tribu asymptotique $\mathcal{T}$ associée aux à la suite des X_n comme l'ensemble des événements ne dépendant que du comportement à l'infini de cette suite, i.e. :

$$\mathcal{T} = \bigcap_{m \geq 0} \bigvee_{n \geq m} \sigma(X_n).$$

1. Pour A et B deux événements ne dépendant chacun que d'un nombre fini de X_n , donner une condition suffisante pour que A et B soient indépendants.

2. Pour A un événement ne dépendant que d'un nombre fini de X_n et $B \in \mathcal{T}$, montrer que A et B sont indépendants.

3. Montrer que pour tout $A \in \mathcal{F}$, pour tout $B \in \mathcal{T}$, A et B sont indépendants.

4. ($\diamond$) En déduire que pour tout $A \in \mathcal{T}$, $\mathbb{P}(A) \in \{0, 1\}$.

Exercice 3 (Loi β)

Soient $U_1, \dots, U_n$ n variables aléatoires indépendantes de loi uniforme sur $[0, 1]$.

1. Montrer que presque-sûrement tous les U_i sont distincts.

Grâce à la question 1, presque-sûrement il existe une unique renumérotation des U_i en ordre croissant : on notera $\{U_1, \dots, U_n\} = \{U^{(1)}, \dots, U^{(n)}\}$ avec $U^{(1)} \leq \dots \leq U^{(n)}$, et on notera σ la permutation de $\{1, \dots, n\}$ telle que $U^{(i)} = U_{\sigma(i)}$.

2. Montrer que la permutation σ et le vecteur aléatoire $U^{(\cdot)}$ sont indépendants.

3. Montrer que σ suit la loi uniforme sur $\mathfrak{S}_n$.

4. Déterminer la loi de $U^{(\cdot)}$.

5. (♥) Montrer que les $U^{(i)}$ sont des variables aléatoires absolument continues par rapport à la mesure de Lebesgue et donner leurs densités.

☛ La loi de $U^{(i)}$ est appelée loi β de paramètres i et $n - i + 1$.

Exercice 4 (Loi du χ^2)

1. (★) Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes de loi $\mathcal{N}(0, 1)$. Quelle est la densité de la loi de $X = \sum_{i=1}^n (X_i)^2$?

Exercice 5 (Inégalité de Hoeffding)

Si Z est une variable aléatoire réelle, on appelle log-transformée de Laplace de Z la fonction à valeurs dans $\mathbb{R} \cup \{+\infty\}$ définie pour $\lambda \in \mathbb{R}_+$ par :

$$\Psi_Z(\lambda) = \ln \mathbb{E}[\exp(\lambda Z)].$$

1. (★) Pour Z une variable aléatoire à valeurs dans un intervalle borné, dériver deux fois Ψ_Z et identifier, pour tout λ , cette dérivée seconde à la variance d'une variable aléatoire Z_λ dont on décrira la loi.

2. Si Z est une variable aléatoire à valeurs dans un intervalle borné I , montrer que la variance de Z est majorée par $|I|^2/4$.

3. Soit Z une variable aléatoire centrée à valeurs dans un intervalle borné I . Dédurre des questions 1 et 2 que pour tout $\lambda \in \mathbb{R}_+$,

$$\Psi_Z(\lambda) \leq \frac{|I|^2 \lambda^2}{8}.$$

4. Soient $Z_1, \dots, Z_n$ des variables aléatoires indépendantes, telles que Z_i est à valeurs dans $[a_i, b_i]$; posons $Z = \sum_{i=1}^n Z_i$ et $\tilde{Z} = Z - \mathbb{E}[Z]$. Montrer que

$$\Psi_{\tilde{Z}} \leq \frac{\lambda^2}{8} \sum_{i=1}^n (b_i - a_i)^2.$$

5. En déduire l'*inégalité de Hoeffding* sur la concentration de la mesure : avec les notations de la question 4, on a pour tout $\varepsilon > 0$:

$$P[|\tilde{Z}| \geq \varepsilon] \leq 2 \exp \left(-\frac{2\varepsilon^2}{\sum_{i=1}^n (b_i - a_i)^2} \right).$$

Feuille 4 : Lemme de Borel-Cantelli, Fonctions caractéristiques

9 février 2009

Lemme de Borel-Cantelli

Exercice 1 (Absence de mesure canonique sur $\mathbb{N}^*$)

Le but de cet exercice est de montrer qu'il n'existe pas de mesure de probabilité $\mathbb{P}$ sur $(\mathbb{N}^, \mathfrak{P}(\mathbb{N}^*))$ telle qu'on ait $\mathbb{P}(\text{« être divisible par } n \text{ »}) = 1/n$ pour tout $n \geq 1$. On va raisonner par l'absurde en supposant donnée une telle probabilité pour aboutir à une contradiction.*

1. Montrer que, pour $n_1, \dots, n_k$ deux à deux premiers entre eux, les événements « être divisible par n_i », $i \in \{1, \dots, k\}$, sont indépendants.

2. (♥) Rappeler comment on démontre la *formule d'Euler* :

$$\forall s > 0 \quad \prod_{p \text{ premier}} \frac{1}{1 - 1/p^s} = \sum_{m=1}^{\infty} \frac{1}{m^s}. \quad (1)$$

Déduire de (1) que la somme des inverses des nombres premiers diverge.

3. En considérant les événements « être divisible par p » pour tous les p premiers, établir alors que $\mathbb{P}$ -presque-tout nombre a une infinité de diviseurs premiers, ce qui est absurde.

Exercice 2 (Loi des grands nombres L^4)

Soient $X_1, \dots, X_n, \dots$ des v.a. i.i.d. centrées admettant des moments d'ordre 4. On note $S_n = X_1 + \dots + X_n$.

1. Calculer $\mathbb{E}((S_n)^4)$.

2. En déduire une majoration de $\mathbb{P}(|\frac{S_n}{n} - \mathbb{E}(X_1)| > \varepsilon)$.

3. (♥) Montrer que $\frac{S_n}{n}$ tend vers $\mathbb{E}(X_1)$ presque-sûrement.

Exercice 3 (Nombres-univers)

Soient $x \in [0, 1[$ un réel et $b \geq 2$ un entier. On dit que x est un nombre-univers en base b quand n'importe quelle séquence finie de chiffres b -aires peut être trouvée dans le développement de x en base b .

1. ($\diamond$) Soit x une variable aléatoire uniforme sur $[0, 1[$ et $0, x_1^b x_2^b x_3^b \dots$ son développement en base b . Montrer que les x_i^b sont i.i.d. et de loi uniforme sur $\{0, \dots, b-1\}$.

2. Pour $\ell \geq 1$ un entier, montrer que les nombres $X_k^{b,\ell}$ de développements b -aires respectifs

$$X_k^{b,\ell} = x_{k\ell+1}^b \dots x_{k\ell+\ell}^b$$

sont indépendants et de loi uniforme sur $\{0, \dots, b^\ell - 1\}$.

3. ($\heartsuit$) En déduire que, presque-sûrement, x est un nombre-univers en base b .

4. En déduire l'existence d'un élément de $[0, 1[$ qui est un nombre-univers en toute base.

5. ($\star\star$) Définir explicitement un tel nombre.

Exercice 4 (Records)

Soient $X_1, X_2, \dots$ des v.a. i.i.d. réelles de fonction de répartition F continue. On pose $A_k = \{X_k > \sup_{j < k} X_j\}$ l'évènement qui exprime que le record est battu au temps k .

1. Montrer que $\forall j < k, \mathbb{P}(X_j = X_k) = 0$.

On réordonne alors les X_i par ordre décroissant ce qui nous définit une permutation aléatoire de $\{1, \dots, n\}$, $\pi_n(i) = j$ si X_i est la j -ième plus grande valeur.

2. Montrer que π_n est uniformément distribuée sur les $n!$ possibilités.

3. En déduire que $\mathbb{P}(A_n) = \frac{1}{n}$.

4. Soient $m < n$ et $i_{m+1}, \dots, i_n$ des éléments distincts de $\{1, \dots, n\}$. Calculer $\mathbb{P}(A_m | \pi_n(j) = i_j \text{ pour } m+1 \leq j \leq n)$. En déduire que les A_k sont indépendants.

5. Soit $R_n = \sum_{m=1}^n \mathbb{1}_{A_m}$ le nombre de fois que le record a été battu au temps n . Montrer que $\frac{R_n}{\sum_{m=1}^n \mathbb{P}(A_m)}$ tend vers 1 en probabilité.

6. ($\star$) Montrer que $\frac{R_n}{\ln(n)}$ tend vers 1 presque-sûrement. (On pourra s'inspirer de l'exercice 2.)

Fonctions caractéristiques

☞ Dans tous les exercices suivants, on admettra que deux lois sur $\mathbb{R}^d$ qui ont la même fonction caractéristique coïncident.

Exercice 5 (Calcul de fonctions caractéristiques)

1. Soit X de loi Γ de paramètre (α, β) . Calculer la fonction caractéristique de X . (On rappelle que la densité de la loi Γ de paramètre (α, β) est $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ si $x \geq 0$, 0 si $x < 0$.)

2. (♥) Soient X et Y des variables aléatoires indépendantes, respectivement de loi $\Gamma(\alpha, \beta)$ et de loi $\Gamma(\delta, \beta)$. Calculer la fonction de répartition de $X + Y$. Que remarque-t-on ?

3. Mêmes questions avec X et Y des variables aléatoires binomiales indépendantes de paramètres (n, p) et (m, p) .

Exercice 6 (Critère d'indépendance de variables aléatoires bornées)

1. (♥) Montrer que la fonction caractéristique d'une variable aléatoire bornée est analytique.

2. Soient X et Y deux variables aléatoires réelles bornées. Montrer que pour que X et Y soient indépendantes, il faut et il suffit que $\forall (k, l) \in \mathbb{N}^2$, $\mathbb{E}(X^k Y^l) = \mathbb{E}(X^k) \mathbb{E}(Y^l)$.

Exercice 7 (Lois stables)

☛ Puisque la fonction caractéristique d'une mesure de probabilité sur $\mathbb{R}^d$ peut être vue comme la transformée de Fourier associée à cette mesure, on peut retrouver une probabilité à partir de sa fonction caractéristique par transformée de Fourier inverse. Ainsi, si μ est une mesure de probabilité sur $\mathbb{R}$ de fonction caractéristique $\hat{\mu}(\xi) = \int e^{i\xi x} d\mu(x)$, la densité de μ par rapport à la mesure de Lebesgue peut formellement s'écrire :

$$\frac{d\mu}{dx}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\xi x} \hat{\mu}(\xi) d\xi, \quad (1)$$

l'égalité (1) devenant vraie stricto sensu dès lors que $\hat{\mu}$ est L^1 .

1. (♦) Montrer que la fonction caractéristique d'une mesure de probabilité sur $\mathbb{R}^d$ est toujours continue.

Soit μ une mesure de probabilité sur $\mathbb{R}$ qui n'est pas une masse de Dirac. On suppose que pour tout $n \geq 2$, si $X_1, \dots, X_n$ sont des variables i.i.d. de loi μ , alors, pour un certain $\lambda(n) > 1$, $(X_1 + \dots + X_n)/\lambda(n)$ a encore pour loi μ .

2. (★) Montrer qu'il existe une constante complexe c , de partie réelle strictement positive, telle que la fonction caractéristique de μ vaut :

$$\hat{\mu}(\xi) = \begin{cases} \exp(-c\xi^\alpha) & \text{si } \xi \geq 0; \\ \exp(-\bar{c}|\xi|^\alpha) & \text{si } \xi \leq 0, \end{cases}$$

où $\alpha = \ln n / \ln \lambda(n)$ pour n'importe quel n — remarquez que cela implique que la connaissance d'un seul $\lambda(n)$ détermine tous les autres.

☛ Les lois μ ayant une fonction caractéristique de cette forme sont appelées lois α -stables.

Dans les trois questions suivantes, on suppose que μ a tous ses moments polynomiaux finis, i.e. que $\int |x|^k d\mu(x) < \infty$ pour tout $k \geq 0$.

3. Déterminer la valeur de α .

4. Montrer que $\hat{\mu}$ est deux fois dérivable en 0.

5. ($\diamond$) En déduire la loi de μ , à un paramètre réel près. Reconnaissez-vous cette famille de lois ?

Dans les deux questions suivantes, on suppose que $\alpha = 1$ et que μ est symétrique, i.e. que si X a pour loi μ , alors $-X$ aussi.

6. Montrer que $\hat{\mu}$ est symétrique, i.e. que $\hat{\mu}(-\xi) = \hat{\mu}(\xi)$ pour tout ξ .

7. ($\heartsuit$) En déduire la loi de μ , à un paramètre réel près. Reconnaissez-vous cette famille de lois ?

8. ($\star\star$) Notons $u_8 = e^{i\pi/4}$. Pour tout $\xi > 0$, montrer que

$$\Re \left(\int_0^{+\infty} \exp(-u_8 \sqrt{x} + i\xi x) dx \right) = \frac{e^{-1/\xi}}{2\sqrt{\pi}\xi^{3/2}}.$$

Dans les deux questions suivantes, on suppose que $\alpha = 1/2$ et que μ est portée par $\mathbb{R}_+$.

9. ($\star$) Déterminer une densité possible pour μ , à un paramètre réel près.

10. ($\star$) Montrer que μ est nécessairement de la forme trouvée à la question 9.

☛ Cette famille de lois s'appelle les lois de Lévy.

Soient $T_1, T_2, \dots$ des variables aléatoires i.i.d. de loi $\text{Exp}(1)$. Pour $i \geq 1$, on pose :

$$P_i = \left(\sum_{j=1}^i T_j \right)^{-2},$$

puis :

$$\Lambda = \sum_{i=1}^{\infty} P_i.$$

11. Montrer que, presque-sûrement, $\Lambda < \infty$.

12. ($\heartsuit, \star\star$) Montrer que Λ suit une loi de Lévy.

13. ($\star$) Déterminer le paramètre de cette loi de Lévy.

Feuille 5 : Notions de convergence

23 février 2009

Exercice 1 (Comprendre l'énoncé d'un théorème de convergence)

Si $X_1, X_2, \dots$ sont des v.a. i.i.d. sur $\mathbb{R}^d$ de loi μ , on appelle mesure empirique de μ au temps N la mesure $\hat{\mu}^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}$. Le théorème de convergence des mesures empiriques affirme alors, informellement, que « les mesures empiriques convergent μ ».

1. Donner un énoncé rigoureux formalisant la proposition entre guillemets, dans lequel on précisera soigneusement où on utilise quel concept de convergence.

Exercice 2 (Contre-exemples)

1. (♥) Construire des variables aléatoires $X_1, X_2, \dots$ et Y telles que X_k converge en loi vers Y , mais pas en probabilité.

2. (♥) Construire des variables aléatoires $X_1, X_2, \dots$ et Y telles que X_k converge presque-sûrement vers Y , mais pas dans L^1 .

3. (♥) Construire des variables aléatoires $X_1, X_2, \dots$ et Y telles que X_k converge vers Y dans L^p pour tout $p < \infty$, mais pas presque-sûrement.

Exercice 3 (Utilisation de l'inégalité de Kolmogorov)

Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes centrées et admettant des moments d'ordre deux. On note $S_k = \sum_{i=1}^k X_i$.

1. En introduisant les ensembles $E_k = (|S_k| \geq \varepsilon) \cap [\cap_{i=1}^{k-1} (|S_i| < \varepsilon)]$, montrer que pour tout $\varepsilon > 0$,

$$\mathbb{P} \left(\max_{1 \leq k \leq n} \left| \sum_{i=1}^k X_i \right| \geq \varepsilon \right) \leq \frac{1}{\varepsilon^2} \sum_{i=1}^n \sigma_{X_i}^2.$$

Soit X_n une suite de variables aléatoires réelles indépendantes centrées admettant des moments d'ordre deux ; on suppose de plus que $\sum_{i=1}^{\infty} \sigma_{X_i}^2 < \infty$.

2. On pose $A_m = \sup_{k \in \mathbb{N}^*} |S_{m+k} - S_m|$. Montrer que

$$\left\{ A_m > \frac{1}{n} \right\} = \bigcup_{r \in \mathbb{N}^*} \left\{ \sup_{1 \leq k \leq r} |S_{m+k} - S_m| > \frac{1}{n} \right\}.$$

3. En déduire une majoration de $\mathbb{P}(A_m > \frac{1}{n})$.
4. Montrer alors que S_n converge presque-sûrement vers une variable aléatoire.
5. Montrer que S_n converge également dans L^2 .

Exercice 4 (Critères de convergence presque-sûre)

1. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires. Montrer que, s'il existe une série de terme général $\varepsilon_n \geq 0$, convergente, telle que

$$\sum_{n=0}^{\infty} \mathbb{P}(|X_{n+1} - X_n| > \varepsilon_n) < \infty,$$

alors la suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ converge presque-sûrement.

2. Soit $(X_n)_{n \in \mathbb{N}}$ des variables aléatoires. Montrer que si X est une variable aléatoire telle que pour tout $\varepsilon > 0$,

$$\sum_{n=0}^{\infty} \mathbb{P}(|X_n - X| > \varepsilon) < \infty,$$

alors la suite de variables aléatoires X_n converge presque-sûrement vers X .

3. (★) Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires indépendantes qui convergent presque-sûrement vers 0. Montrer qu'alors, on a pour tout $\varepsilon > 0$,

$$\sum_{n=0}^{\infty} \mathbb{P}(|X_n| > \varepsilon) < \infty.$$

Exercice 5 (Convergence étroite et fonctions de répartition)

1. Soient μ et $\mu_1, \mu_2, \dots$ des mesures de probabilité sur $\mathbb{R}^d$ muni de sa tribu borélienne. Montrer que μ_k converge étroitement vers μ si et seulement si, pour toute fonction φ continue à support compact sur $\mathbb{R}^d$, $\int_{\mathbb{R}^d} \varphi d\mu_k \xrightarrow[k \rightarrow \infty]{} \int_{\mathbb{R}^d} \varphi d\mu$.

2. (♥) Soient μ et $\mu_1, \mu_2, \dots$ des mesures de probabilité sur $\mathbb{R}$, et $F, F_1, F_2, \dots$ leurs fonctions de répartition respectives. Montrer que si les F_k convergent simplement vers F , alors les μ_k convergent étroitement vers μ .

3. Mêmes notations, et on suppose de plus que μ est diffuse. Montrer que si les μ_k convergent étroitement vers μ , alors les F_k convergent simplement vers F .

Exercice 6 (Pas de théorème de Cesàro pour la convergence en probabilité)

Soit (X_n) une suite de variables aléatoires réelles indépendantes, la fonction de répartition de X_n étant F_n définie par 0 si $x < 0$ et par $1 - \frac{1}{x+n}$ sinon. On note $S_n = \sum_{k=1}^n X_k$ et $Y_n = \frac{S_n}{n}$.

1. Montrer que X_n converge vers 0 en probabilité.
2. (★) En utilisant $M_n = \max_{1 \leq k \leq n} X_k$, montrer que Y_n ne converge pas vers 0 en probabilité.

Exercice 7 (Convergence en loi et somme)

Soient, sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, deux suites $(X_n)_{n \in \mathbb{N}}$ et $(Y_n)_{n \in \mathbb{N}}$ de variables aléatoires réelles qui convergent en loi respectivement vers les variables aléatoires indépendantes X et Y .

1. ($\diamond$, ★) Si, pour tout $n \in \mathbb{N}$, X_n et Y_n sont indépendantes, démontrer que la suite des variables aléatoires (X_n, Y_n) , $n \in \mathbb{N}$, converge en loi vers (X, Y) .
2. ($\heartsuit$) Soient X et Y deux variables aléatoires réelles indépendantes de même loi de Bernoulli $(\delta_0 + \delta_1)/2$. On pose, pour tout $n \in \mathbb{N}^*$,

$$X_n = X + \frac{1}{n} \quad \text{et} \quad Y_n = (1 - X) - \frac{1}{n}.$$

Étudier la convergence en loi des trois suites $(X_n)_{n \in \mathbb{N}}$, $(Y_n)_{n \in \mathbb{N}}$ et $(X_n + Y_n)_{n \in \mathbb{N}}$. En conclure que la suite des variables aléatoires (X_n, Y_n) , $n \in \mathbb{N}^*$, ne converge pas en loi vers (X, Y) .

Exercice 8 (Topologie sur les mesures)

Soit E un espace métrique muni de sa tribu borélienne, et $C = C_b^0(E)$ l'espace de Banach des fonctions continues bornées sur E munies de la norme du supremum des valeurs absolues. On note C' le dual topologique de C , qu'on munit de sa topologie faible-*, sauf mention explicite du contraire. On note $\mathcal{P}(E)$ l'ensemble des mesures de probabilité sur E .

1. Montrer que $\mathcal{P}(E)$ s'identifie à une partie de C' , et que cette partie est fermée et contenue dans la boule-unité (pour la norme de la topologie forte) de C' .
2. (★) Si E est compact, montrer que C est séparable.
3. ($\diamond$) En déduire que si E est compact, $\mathcal{P}(E)$ muni de la convergence faible est un compact métrisable.
4. Pour $E = [0, 1]$, donner un exemple de métrique qui métrise la topologie faible sur $\mathcal{P}(E)$.

Exercice 9 (Convergence vers une constante)

Soient E un espace métrique, $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. à valeurs dans E et x un point de E ; on note également x la variable aléatoire égale à x p.s..

1. ($\diamond$) Montrer que X_n tend vers x en probabilité si et seulement si X_n tend vers x en loi.

Exercice 10 (Théorème de Skorohod sur $\mathbb{R}$)

Le but de cet exercice est de montrer le théorème suivant : soit (μ_n) une suite de lois de probabilité sur $\mathbb{R}$, qui converge étroitement vers une loi de probabilité μ . On peut alors définir, sur un même espace probabilisé, une suite de variables aléatoires (Y_n) et une variable aléatoire Y , telles que pour chaque n la variable Y_n ait pour loi μ_n et Y ait pour loi μ et telles que l'on ait Y_n qui converge presque-sûrement vers Y . On notera $F_n(t) = \mu_n((-\infty, t])$ et $F(t) = \mu((-\infty, t])$ les fonctions de répartition de μ_n et de μ .

1. ($\diamond$) On considère $[0, 1]$ muni de la tribu borélienne et de la mesure de Lebesgue et les variables aléatoires

$$Y_n(s) = \inf\{t \in \mathbb{R}, F_n(t) \geq s\}.$$

On définit $Y(s)$ de la même manière. Montrer que Y_n a pour loi μ_n et que Y a pour loi μ .

2. ($\heartsuit$) Montrer que $Y_n(s)$ converge vers $Y(s)$ en tout point où Y est continue. En déduire que Y_n converge presque-sûrement vers Y .

3. ($\star$) Montrer que si $\mathbb{P}(X_n \leq x) \rightarrow \mathbb{P}(X \leq x)$ en tout point de continuité de $\mathbb{P}(X \leq \cdot)$, alors pour toute fonction continue bornée g , $\mathbb{E}(g(X_n)) \rightarrow \mathbb{E}(g(X))$.

Exercice 11 (Théorème de Slutsky)

Soient X_n et Y_n deux suites de variables aléatoires réelles convergeant en loi resp. vers une constante x et une v.a. Y .

1. ($\diamond$) Montrer que $X_n + Y_n$ converge en loi vers $Y + x$, et que $X_n Y_n$ converge en loi vers xY (théorème de Slutsky).

2. ($\heartsuit, \star$) Montrer qu'en fait, (X_n, Y_n) converge en loi vers (x, Y) , et en déduire que pour toute fonction continue de $\mathbb{R}^2$ dans $\mathbb{R}$, $f(X_n, Y_n)$ converge en loi vers $f(x, Y)$.

☛ Le théorème de Slutsky est souvent utilisé en statistiques quand on doit calculer des grandeurs faisant intervenir un paramètre du modèle qu'on ne sait qu'estimer asymptotiquement, pour montrer que le résultat obtenu en utilisant l'estimation est asymptotiquement le même que si on avait utilisé le vrai paramètre.

Feuille 6 : Loi des grands nombres

9 mars 2009

Exercice 1 (Convergence des mesures empiriques)

Soit μ une mesure de probabilité sur $\mathbb{R}^n$. On considère une suite de variables aléatoires i.i.d. $X_1, X_2, \dots$ de loi μ , et pour $N \geq 1$ on note $\hat{\mu}^N$ la mesure (aléatoire) $\frac{1}{N} \sum_{i=1}^N \delta_{X_i}$. Le but de cet exercice est de montrer que, presque-sûrement, $\hat{\mu}^N$ converge faiblement (càd. au sens de la convergence en loi) vers μ .

1. Rappeler comment on peut exprimer la propriété « $\hat{\mu}^N$ converge en faiblement vers μ » en fonction de la donnée de tous les $\hat{\mu}^N(U)$ pour $U \subset \mathbb{R}^n$ ouvert et $N \in \mathbb{N}$.

2. Soit V un ouvert fixé de $\mathbb{R}^n$; montrer que $\hat{\mu}^N(V) \xrightarrow[N \rightarrow \infty]{} \mu(V)$ presque-sûrement.

3. (★) Pourquoi ne peut-on pas en déduire que « presque-sûrement, pour tout ouvert $V \subset \mathbb{R}^n$, $\hat{\mu}^N(V) \rightarrow \mu(V)$ » ? Montrer que par exemple, si μ est la mesure de Lebesgue sur $[0, 1]$, alors presque-sûrement il existe un ouvert (aléatoire) $\tilde{V}$ tel que $\hat{\mu}^N(\tilde{V}) \not\rightarrow \mu(\tilde{V})$.

4. (♡) Pour X un ensemble dénombrable, montrer que l'ensemble des parties finies de X est dénombrable.

5. (♡) Montrer qu'il existe une famille dénombrable $(\hat{V}_i)_{i \in \mathbb{N}}$ d'ouverts de $\mathbb{R}^n$ tel que tout ouvert U de $\mathbb{R}^n$ s'écrive comme réunion de $\hat{V}_i$, càd. qu'il existe $J \subset \mathbb{N}$ tel que $U = \bigcup_{i \in J} \hat{V}_i$.

6. En déduire qu'il existe une famille dénombrable $(V_i)_{i \in \mathbb{N}}$ d'ouverts de $\mathbb{R}^n$ telle que tout ouvert U de $\mathbb{R}^n$ s'écrive comme limite croissante de V_i , càd. qu'il existe une application $i : \mathbb{N} \rightarrow \mathbb{N}$ telle que $U = \lim_{k \nearrow \infty} V_{i(k)}$.

7. (♡, ★) Montrer le théorème de convergence des mesures empiriques.

Exercice 2 (Loi des grands nombres L^1)

Soit μ une mesure de probabilité sur $\mathbb{R}$ ayant un moment d'ordre 1 ; notons $m = \int_{\mathbb{R}} x d\mu(x)$ la moyenne de μ . Soient $X_1, X_2, \dots$ des v.a. i.i.d. de loi μ ; pour $n \geq 1$, posons $S_n = \sum_{i=1}^n X_i$ et $M_n = S_n/n$.

1. Si μ a une variance $\sigma^2 < \infty$, montrer que $\mathbb{E}[(S_n - nm)^2] = n\sigma^2$.

2. Toujours si μ a une variance finie, déduire de la question 1 que M_n converge dans L^1 vers m .

On ne suppose plus que μ a une variance finie. Pour $A \geq 0$ fixé, on pose

$$\tilde{X}_n = \begin{cases} -A & \text{si } X_n \leq -A; \\ X_n & \text{si } -A \leq X_n \leq A; \\ A & \text{si } X_n \geq A, \end{cases}$$

et $h_n = X_n - \tilde{X}_n$.

3. Pourquoi les X_n , resp. les h_n , sont-ils i.i.d. ?

4. Dédurre de la question 2 que

$$\lim_{n \rightarrow \infty} \|M_n - \mathbb{E}[\tilde{X}_1]\|_{L^1} \leq \|h_1\|_{L^1}.$$

5. Montrer que, quand $A \rightarrow \infty$, $\|h_1\|_{L^1} \rightarrow 0$ et $\mathbb{E}[\tilde{X}_1] \rightarrow m$.

6. ($\diamond$) En déduire que M_n converge vers m dans L^1 . Ce résultat s'appelle la *loi des grands nombres* L^1 .

Exercice 3 (Loi des grands nombres L^2)

Soit (X_n) une suite de variables aléatoires indépendantes admettant un moment d'ordre deux. On suppose que :

$$\mathbb{E}(X_n) \rightarrow m \quad \text{et} \quad \sum_{j=1}^{\infty} \frac{1}{j^2} \sigma_{X_j}^2 < \infty.$$

1. Calculer $\|\frac{1}{n} \sum_{j=1}^n (X_j - \mathbb{E}(X_j))\|_2^2$.

2. À l'aide d'une intégration par parties discrète, montrer que $\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{j=1}^n \sigma_{X_j}^2 = 0$.

3. En déduire que $\frac{1}{n} \sum_{k=1}^n X_k$ converge dans L^2 vers m .

Exercice 4 (Réciproques de la loi forte des grands nombres)

☛ Dans tout cet exercice, pour μ une loi de probabilité sur $\mathbb{R}$, on appelle suite des moyennes empiriques de μ la suite de variables aléatoires $\hat{E}_N = \frac{1}{N} \sum_{i=1}^N X_i$, où $X_1, X_2, \dots$ sont i.i.d. selon μ .

1. ($\diamond, \star$) Pour X une variable aléatoire à valeurs positives, montrer que $\mathbb{E}[X] = \int_{x=0}^{\infty} \mathbb{P}(X \geq x) dx$ (Indication : utiliser le théorème de Fubini).

2. En déduire que, si μ est une loi sur $\mathbb{R}$ n'ayant pas de moment d'ordre 1 et $X_1, X_2, \dots$ une suite de v.a. i.i.d. de loi μ , alors p.s. il existe une infinité de $i \in \mathbb{N}^*$ pour lesquels $|X_i| \geq i$ (Indication : utiliser le lemme de Borel–Cantelli).

3. En déduire que, si μ est une loi de probabilité sur $\mathbb{R}$ n'ayant pas de moment d'ordre 1, alors presque-sûrement les moyennes empiriques $\hat{E}_N$ ne convergent vers aucune limite finie.

4. Si X et Y sont deux variables aléatoires réelles indépendantes et que Y n'est pas intégrable, montrer que $X + Y$ n'est pas intégrable non plus (*Indication : utiliser le théorème de Fubini*).

5. En déduire que, si μ n'a pas de moment d'ordre 1, ses moyennes empiriques $\hat{E}_N$ ne sont pas intégrables.

6. (♥) En quoi le résultat de la question 3, resp. de la question 5, est-il une réciproque de la loi forte des grands nombres, resp. de la loi des grands nombres L^1 ?

Considérons la construction consistant à tirer au sort un y de loi uniforme sur $[-1/2, 1/2]$, puis à définir à partir de y la variable aléatoire $1/y\sqrt{|\ln|y||}$; on note μ la loi de cette dernière variable.

7. Montrer que $\mu([-M, M]) = o(M^{-1})$ quand M tend vers l'infini.

8. Montrer que néanmoins μ n'a pas de moment d'ordre 1.

9. Au fait, sauriez-vous montrer que le résultat de la question 7 est toujours vrai dès que μ a un moment d'ordre 1 ? (*Indication : utiliser l'inégalité de Markov*).

10. Montrer que la fonction caractéristique de μ est réelle.

11. (★) Montrer que la fonction caractéristique de μ est dérivable en 0 et de dérivée nulle.

12. En déduire que la loi des $\hat{E}_N$ converge faiblement vers δ_0 (*Indication : utiliser le théorème de Lévy*).

13. (♥) En déduire que les $\hat{E}_N$ convergent en probabilité vers 0.

14. (★) Décrire qualitativement le comportement de la suite des $\hat{E}_N(\omega)$ pour une réalisation ω typique.

Feuille 7 : Théorème-limite central, Variables gaussiennes

30 mars 2009

Exercice 1 (Vecteurs aléatoires et opérateurs affines)

Soient $n, m \geq 0$ et Φ un opérateur affine de $\mathbb{R}^n$ dans $\mathbb{R}^m$, défini par l'expression matricielle :

$$\Phi(x) = Ax + B,$$

pour A une matrice $m \times n$ et B un vecteur-colonne de taille m .

1. ($\diamond$) Soit X un vecteur aléatoire de $\mathbb{R}^n$ intégrable. Exprimer $\mathbb{E}[\Phi(X)]$ en fonction de $\mathbb{E}[X]$.

2. ($\diamond$) Supposons en outre que X est L^2 . Exprimer la matrice de covariance de $\Phi(X)$ en fonction de $\mathbb{E}[X]$ et de la matrice de covariance de X .

Exercice 2 (Contre-exemples sur les lois gaussiennes)

1. Soit X une variable aléatoire $\mathcal{N}(0, 1)$ et soit Z indépendante de X telle que $\mathbb{P}(Z = 1) = \mathbb{P}(Z = -1) = 1/2$. Soit $Y = ZX$. Montrer que Y est de loi $\mathcal{N}(0, 1)$, que X et Y sont décorrélées (càd. que leur covariance est nulle), mais que (X, Y) n'est pas gaussien.

2. Montrer qu'il existe X_1, X_2 des variables aléatoires complexes non corrélées (ce qui signifiera ici que leur covariance *complexe* est nulle) telles que la variable aléatoire (X_1, X_2) (vue cette fois-ci comme un quadruplet *réel*) soit gaussienne mais que X_1 et X_2 ne soient pas indépendantes.

3. ($\star\star$) Montrer qu'il existe trois variables aléatoires gaussiennes X_1, X_2, X_3 deux à deux indépendantes telles que (X_1, X_2, X_3) ne suive pas une loi gaussienne.

Exercice 3 (Densité des lois gaussiennes)

☞ Dans cet exercice, les lois gaussiennes sont toutes implicitement supposées centrées.

1. Pour $n \geq 0$, soient $\xi_1, \dots, \xi_n$ des variables aléatoires réelles i.i.d. de loi $\mathcal{N}(1)$. Montrer que le vecteur aléatoire $X = (\xi_1, \dots, \xi_n)$ est un vecteur gaussien et calculer sa matrice de covariance.

2. Soit M une matrice $n \times n$ inversible. Montrer que MX est un vecteur gaussien et calculer sa covariance.

3. Montrer que MX a une densité sur $\mathbb{R}^n$ et calculer cette densité.

4. (♥) Dédire des questions 2 et 3 que, si C est une matrice symétrique $n \times n$ positive et définie, alors le vecteur gaussien Y de $\mathbb{R}^n$ de matrice de covariance C a une densité par rapport à la mesure de Lebesgue, et calculer cette densité.

5. (♥,★) Si C n'est pas définie, soit r son rang ; montrer que le support de Y (càd. le plus petit fermé contenant Y presque-sûrement) est un sous-espace strict de $\mathbb{R}^n$ de dimension r , et déterminer ce sous-espace.

Exercice 4 (Introduction au mouvement brownien)

Soit μ une loi sur $\mathbb{R}$ ayant un moment d'ordre 2, avec $\int_{\mathbb{R}} x d\mu(x) = 0$ et $\int_{\mathbb{R}} x^2 d\mu(x) = 1$. Soient $X_1, X_2, \dots$ des v.a. i.i.d. de loi μ . Pour $n \in \mathbb{N}$, notons $S_n = \sum_{i=1}^n X_i$, et pour $t \in \mathbb{R}_+$, notons $S_t = S_{[t]}$ (où $[t]$ désigne la partie entière de t). Soient $0 = t_0 < t_1 < t_2 < \dots < t_k$.

1. Pour $n \in \mathbb{N}$, montrer que $S_{nt_1}, S_{nt_2} - S_{nt_1}, \dots, S_{nt_k} - S_{nt_{k-1}}$ sont indépendants.

2. Pour $i \in \{1, \dots, k\}$, montrer que $(S_{nt_i} - S_{nt_{i-1}})/\sqrt{n}$ converge en loi vers $\mathcal{N}(0, t_i - t_{i-1})$.

3. (♥) En déduire que $(S_{nt_1}/\sqrt{n}, \dots, S_{nt_k}/\sqrt{n})$ converge en loi vers un vecteur gaussien dont on déterminera les paramètres (Rappel : Si $\mu_n \rightharpoonup \mu$ et $\nu_n \rightharpoonup \nu$, alors $\mu_n \otimes \nu_n \rightharpoonup \mu \otimes \nu$).

Exercice 5 (Test du χ^2)

On partitionne Ω en k événements A_i de probabilités $p_1, \dots, p_k > 0$. On répète alors n fois l'expérience aléatoire et on compte le nombre de réalisations respectives des A_i : $N_1, \dots, N_k$.

1. Montrer que $(N_1, \dots, N_k)$ suit une loi multinomiale d'effectif n et de paramètres $p_1, \dots, p_k$.

2. Calculer Σ , la matrice de covariance de $(N_1, \dots, N_k)$.

3. La loi limite de $(n^{-1}N_1, \dots, n^{-1}N_k)$ étant dégénérée, on considère le vecteur $X = \frac{1}{\sqrt{n}}(N_1 - np_1, \dots, N_k - np_k)$. Appliquer le théorème central limite et calculer la matrice de covariance Σ de la loi limite de X .

4. (♦) Montrer que $(1, \dots, 1)$ est dans le noyau de Σ . Quelle interprétation probabiliste donner à ce résultat ?

5. (★) Montrer que Σ est de rang $n - 1$. En déduire que la matrice Σ^* obtenue en supprimant la dernière ligne et la dernière colonne de Σ est inversible.

6. On rappelle qu'une variable aléatoire suit une loi du χ^2 à p degrés de liberté quand

elle peut s'écrire comme $\sum_{i=1}^p Z_i^2$ où les Z_i sont indépendants de loi $\mathcal{N}(0, 1)$. Posons $X^* = (X_1, \dots, X_{k-1})$ et soit $D^2 = {}^t X^* (\Sigma^*)^{-1} X^*$; montrer que sa loi limite est une loi du χ^2 à $k - 1$ degrés de liberté.

7. (♥) S'inspirer des questions précédentes pour expliquer comment on peut mettre en place un test (appelé test du χ^2) pour affirmer qu'un dé est vraisemblablement pipé.

Exercice 6 (Moyenne et variance empiriques)

Soient $X_1, \dots, X_n$ des variables aléatoires réelles indépendantes de même loi μ ayant un moment d'ordre deux. On définit les variables aléatoires suivantes, la moyenne et la variance empiriques, par

$$M_n = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{et} \quad \Sigma_n = \frac{1}{n} \sum_{i=1}^n X_i^2 - M_n^2.$$

On note $X = (X_1, \dots, X_n)$.

1. (★) Si μ est la loi gaussienne $\mathcal{N}(m, \sigma^2)$, quelle est la loi de X ? Soit C une matrice orthogonale de dimension n telle que pour tout $j = 1, \dots, n$, on ait $C_{1,j} = \frac{1}{\sqrt{n}}$. Exprimer M_n et Σ_n à l'aide des composantes de CX et en déduire que M_n et Σ_n sont indépendantes.

On s'intéresse maintenant à la réciproque : on suppose que M_n et Σ_n sont indépendantes. On supposera dans un premier temps les X_i centrées. On notera σ la variance de μ , φ la transformée de Fourier de μ , $S_n = nM_n$ et $V_n = n\Sigma_n$.

2. (♥) Calculer $\mathbb{E}[V_n]$ en fonction de σ .

3. (★) Montrer que φ est deux fois différentiable, et calculer $\mathbb{E}[V_n \exp(iuS_n)]$ à l'aide de φ , φ' et φ'' .

4. Montrer que

$$\forall u \in \mathbb{R} \quad \mathbb{E}[V_n \exp(iuS_n)] = (\varphi(u))^n \mathbb{E}[V_n].$$

5. (★) En déduire que φ est solution d'une équation différentielle. Montrer alors que μ suit une loi gaussienne $\mathcal{N}(0, \sigma^2)$.

6. Démontrer que μ est nécessairement gaussienne même si on ne suppose plus les X_i centrées.

Exercice 7 (Un TCL pour des variables sans moment d'ordre 2)

On note μ la loi sur $\mathbb{R}$ à densité $d\mu(x) = \mathbb{1}_{|x| \geq 1} dx / (2x^2)$.

1. Montrer que μ a un moment d'ordre p pour tout $p < 1$, mais qu'elle n'a pas de moment d'ordre 1, ni *a fortiori* d'ordre 2.

2. Expliquer pourquoi la fonction caractéristique de μ , notée $\widehat{\mu}$, est réelle et symétrique.

3. Calculer un équivalent que $1 - \widehat{\mu}(\xi)$ quand ξ tend vers 0 par valeurs positives.

Soient $X_1, X_2, \dots$ des v.a. i.i.d. de loi μ et $\alpha \geq 0$ un réel ; on pose $Y_n^\alpha = n^{-\alpha}(X_1 + \dots + X_n)$. Le but de l'exercice est de déterminer la convergence en loi des Y_n^α pour un α bien choisi.

4. Avec ce formalisme, à quelles valeurs de α correspondent la loi des grands nombres, resp. le théorème-limite central « ordinaire » ?

5. ($\diamond$) Exprimer la fonction caractéristique de Y_n^α en fonction de $\widehat{\mu}$.

6. En déduire que, pour une valeur de α qu'on précisera, les fonctions caractéristiques des $\widehat{Y}_n$ convergent simplement vers une fonction Φ non identiquement égale à 1 dont on donnera l'expression.

7. Pourquoi sait-on à l'avance que Φ est la transformée de Fourier d'une loi de probabilité ν sur $\mathbb{R}$?

8. Calculer la mesure ν dont Φ est fonction caractéristique. Reconnaissez-vous cette loi ?

9. ($\heartsuit$) Résumer les résultats obtenus dans cet exercice sous la forme d'un « théorème-limite central pour μ ».

Feuille 8 : Statistiques

6 avril 2009

Exercice 1 (Maximum de vraisemblance)

On considère le modèle statistique consistant à tirer n variables aléatoires i.i.d. de loi μ_θ , avec $\theta \in \Theta = \mathbb{R} \times \mathbb{R}_+$ et $\mu_{m,\sigma^2} = \mathcal{N}(m, \sigma^2)$.

1. Donner l'estimateur de maximum de vraisemblance pour θ .
2. Soit $\hat{\theta} = (\hat{m}, \hat{\sigma}^2)$ cet estimateur. $\hat{m}$, resp. $\hat{\sigma}^2$, est-il convergent ? sans biais ?

Exercice 2 (Intervalle de confiance non asymptotique pour des gaussiennes)

Soient $X_1, \dots, X_n$ ($2 \leq n \ll \infty$) des variables i.i.d. selon une loi $\mu = \mathcal{N}(m, \sigma^2)$ dont on cherche à évaluer les paramètres. On définit la moyenne et la variance empiriques par

$$\hat{m} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{et} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \hat{m}^2.$$

1. (★) Notant $X = (X_1, \dots, X_n)$, soit C une matrice orthogonale de dimension n telle que pour tout $j = 1, \dots, n$, on ait $C_{1,j} = \frac{1}{\sqrt{n}}$. Exprimer $\hat{m}$ et $\hat{\sigma}^2$ à l'aide des composantes de CX et en déduire que $\hat{m}$ et $\hat{\sigma}^2$ sont indépendantes étant donnée μ . (Remarque : cette question a déjà été posée en VI-1).

2. Donner les lois suivies par $\hat{m}$ et $\hat{\sigma}^2$ en fonction de (m, σ^2) et de n .

3. Quelle propriété remarquable vérifie la loi de la variable aléatoire $\sigma^2/\hat{\sigma}^2$? En déduire comment contruire un intervalle de confiance non asymptotique de niveau α arbitraire pour σ^2 .

4. (★) En considérant la variable $(\hat{m} - m)/\sqrt{\hat{\sigma}^2}$, indiquer comment construire un intervalle de confiance non asymptotique de niveau arbitraire pour m .

5. En déduire un « rectangle de confiance » non asymptotique de niveau $\leq \alpha$ arbitraire pour (m, σ^2) .

La loi de Student à k degrés de libertés ($k \geq 1$) est définie comme la loi d'une variable $T = \frac{Z}{\sqrt{U/k}}$, où $(Z, U) \sim \mathcal{N}(0, 1) \otimes \chi^2(k)$.

6. (★) Déterminer la densité de la loi de Student.

7. (♥) Expliquer l'intérêt de *tabuler* les valeurs des lois de Student (pour $k \ll \infty$), c.à.d. de calculer (l'inverse de) leurs fonctions de répartition et d'éditer les résultats de ces calculs.

8. (♥) Déterminer la limite de loi de Student à k degrés de libertés quand $k \rightarrow \infty$. En déduire un intervalle de confiance asymptotique de niveau arbitraire pour m quand $n \rightarrow \infty$. Reconnaissez-vous cet intervalle ? À quelle famille de lois (plus vaste que les gaussiennes) s'applique-t-il ?

Exercice 3 (Intervalle de confiance non asymptotique dans un cadre non paramétrique)

Soit m la mesure de Lebesgue sur $[0, 1]$ et F sa fonction de répartition. On considère $X_1, \dots, X_n$ i.i.d. selon m , et on note $\hat{F}^n$ la fonction de répartition empirique des X_i .

1. (♥, ★) Pour $\lambda \in \mathbb{R}$, $p \in [0, 1]$, calculer $\mathbb{E}[e^{\lambda(\hat{F}^n(p) - F(p))}]$. En déduire que pour tout $p \in [0, 1]$, pour tout $a \geq 0$,

$$\mathbb{P}\left(|\hat{F}^n(p) - F(p)| \geq \frac{a}{\sqrt{n}}\right) \leq 2e^{-a^2/8}.$$

(Indication : c est un avatar de l'exercice 5 de la feuille 3).

2. L'espace des fonctions càdlàg de $[0, 1]$ dans $\mathbb{R}$ est muni de la norme du supremum $|\cdot|_{\sup}$. Déduire de la question 1 que pour tous $a \geq 0$ et $b \in \mathbb{N}^*$,

$$\mathbb{P}\left(|\hat{F}^n - F|_{\sup} \geq \frac{a}{\sqrt{n}} + \frac{1}{2b}\right) \leq 2be^{-a^2/8}.$$

3. En prenant $a = \frac{2\sqrt{2}}{3}c$, en déduire que pour tout $c \geq 0$,

$$\mathbb{P}\left(|\hat{F}^n - F|_{\sup} \geq \frac{c}{\sqrt{n}}\right) \leq \left[\frac{\sqrt{n}}{4a} + 2\right]e^{-c^2/9}.$$

On s'intéresse maintenant à une mesure de probabilité μ inconnue sur $[0, 1]$, dont on veut déterminer l'espérance à partir de réalisations i.i.d. $X_1, \dots, X_n$ de μ .

4. On suppose que μ est diffuse et de support $[0, 1]$ tout entier (càd. qu'elle ne charge aucun point mais qu'elle charge tout ouvert non vide de $[0, 1]$). Soit G la fonction de répartition de μ et $\hat{G}^n$ sa fonction de répartition empirique. Montrer que pour tout $c \geq 0$,

$$\mathbb{P}\left(|\hat{G}^n - G|_{\sup} \geq \frac{c}{\sqrt{n}}\right) \leq \left[\frac{\sqrt{n}}{4a} + 2\right]e^{-c^2/9}.$$

(Indication : utiliser l'exercice 5 de la feuille 2).

5. On considère l'estimateur $\hat{E} = (X_1 + \dots + X_n)/n$. Comment appelleriez-vous cet estimateur ? Est-il convergent ? A-t-il un biais ?

6. (♦) Montrer que, si ν est une mesure de probabilité sur $[0, 1]$ de fonction de répartition H ,

$$\mathbb{E}[\nu] = 1 - \int_0^1 H(x) dx.$$

7. En déduire une famille d'intervalles de confiance autour de $\hat{E}$ dont on majorera le niveau de façon non asymptotique.

8. Relaxer l'hypothèse selon laquelle μ est diffuse et sans atome.

Exercice 4 (Test de Mann et Whitney)

On dispose d'échantillons indépendants de deux populations différentes, que l'on note $X_1, \dots, X_N$ et $Y_1, \dots, Y_M$, où les X_i et les Y_j sont des réels, et on suppose que les lois les ayant engendrés sont diffuses. Le but du test de Mann–Whitney est de mélanger les deux populations en un unique ensemble de données, et à partir de considérations sur le rang des éléments d'une des deux populations de pouvoir affirmer ou infirmer l'hypothèse que les deux populations suivent la même loi.

1. Montrer que presque-sûrement, il n'y a pas d'ex-æquo dans les populations.

2. On pose $Z = (X_1, \dots, X_N, Y_1, \dots, Y_M)$. On considère la statistique d'ordre de ce vecteur et on note R le vecteur des rangs correspondants (R_1 est la place de la valeur X_1 dans la suite ordonnée de façon croissante). Posons

$$U_{Y,X} = \sum_{j=1}^M R_{N+j} - \frac{M(M+1)}{2},$$

montrer que

$$U_{Y,X} = \sum_{i,j=1}^{N,M} \mathbb{1}_{X_i < Y_j}.$$

3. On pose H_0 l'hypothèse selon laquelle les deux échantillons sont issus de la même loi. Montrer que sous cette hypothèse la loi de $U_{Y,X}$ sous H_0 ne dépend pas de la loi commune des individus des deux échantillons.

4. Montrer que sous H_0

$$E[U_{Y,X}] = \frac{NM}{2}.$$

5. Si l'on note $\tilde{\mathbb{1}}_{X_i < Y_j} = \mathbb{1}_{X_i < Y_j} - 1/2$, montrer que, sous H_0 , pour tous $i \neq i', j \neq j'$:

$$E[\tilde{\mathbb{1}}_{X_i < Y_j} \tilde{\mathbb{1}}_{X_i < Y_{j'}}] = \frac{1}{12},$$

$$E[\tilde{\mathbb{1}}_{X_i < Y_j} \tilde{\mathbb{1}}_{X_{i'} < Y_j}] = \frac{1}{12},$$

$$E[(\tilde{\mathbb{1}}_{X_i < Y_j})^2] = \frac{1}{4}.$$

En déduire la variance de $U_{Y,X}$.

6. (★★) Montrer que sous l'hypothèse H_0 , lorsque N et M tendent vers l'infini avec M/N possédant une limite finie non nulle, on a la convergence suivante :

$$\frac{U_{Y,X} - \frac{MN}{2}}{\text{Var}(U_{Y,X})^{1/2}} \xrightarrow{(d)} \mathcal{N}(0, 1).$$

7. (★) Montrer que si par exemple $\mathbb{P}[X > Y] > 1/2$, ce rapport tend presque sûrement vers $+\infty$.

8. En déduire une procédure de test.

Exercice 5 (Test de Neyman–Pearson)

Soit X un espace mesurable et soient μ, ν deux lois de probabilité distinctes sur X . On considère le modèle statistique consistant à tirer une variable aléatoire x , soit selon μ , soit selon ν , de sorte que l'espace des paramètres $\Theta = \{\mu, \nu\}$ ne contient que deux points. On veut tester l'hypothèse que la loi utilisée pour tirer x est μ . Compte tenu de la nature de l'espace des paramètres, l'hypothèse alternative sera donc ici réduite à la seule mesure ν .

On note f et g les densités respectives de μ et ν par rapport à la mesure $\lambda = \mu + \nu$, de sorte que $f + g \equiv 1$. Soit $p \in]0, 1]$; on considère le test consistant à répondre « μ » si $f \geq p$ et « ν » sinon (c'est le test de Neyman–Pearson).

1. Exprimer le niveau de confiance α du test en fonction des données.

2. Soit β la probabilité que le test échoue quand on est dans l'hypothèse alternative. Exprimer β en fonction des données.

3. Montrer qu'on a toujours $\alpha < p/(1 - p)$, resp. $\beta \leq (1 - p)/p$.

☞ β s'appelle l'erreur de deuxième espèce de notre test (par opposition à α qui est l'« erreur de première espèce ». Dans le cas général, β est en fait une fonction de Θ_1 (l'ensemble des paramètres de l'hypothèse alternative) dans $[0, 1]$, alors que α est un nombre (c'est le supremum du risque d'erreur sur Θ_0 , l'ensemble des paramètres sur l'hypothèse nulle). Ici Θ_0 et Θ_1 sont toutes les deux réduites à un point, de sorte qu' α et β jouent en fait des rôles symétriques.

4. (♥) Montrer que le test de Neyman–Pearson est optimal : pour tout test de niveau α ou moins, l'erreur de deuxième espèce est supérieure ou égale au β .

☞ Dans la pratique, il peut être difficile de connaître précisément μ et ν . Dans ce cas-là, plutôt que d'étudier directement les lois de probabilité, on va plutôt étudier leur mesure-image par une certaine statistique de x . De plus, il arrive que même ces mesures-là soient difficiles à déterminer, auquel cas on peut recourir à une approximation. C'est ce que nous nous proposons de faire dans la suite de cet exercice.

On dispose d'une suite S de N bits soi-disant aléatoires (hypothèse nulle), qui a en fait peut-être été écrite par un humain tentant de simuler le hasard (hypothèse alternative). Comme, d'après les psychologues, l'être humain n'a en fait pas une très bonne idée du hasard, on se demande s'il existe un test fiable permettant de trancher entre le hasard véritable et le pseudo-hasard des humains.

Pour s une suite finie de bits de longueur $\leq N$, de longueur $\ell(s)$, on note $\hat{\#}(s)$ le nombre de fois que s apparaît dans S .

On appelle *taux d'alternance de la suite de bits* la proportion $\hat{a}$ de fois où deux chiffres consécutifs sont distincts, i.e.

$$\hat{a} = \frac{\#(01) + \#(10)}{N - 1}.$$

5. Déterminer la loi-limite de $N^{1/2}(\hat{a} - \frac{1}{2})$ sous l'hypothèse nulle.

6. Il paraît que le taux d'alternance des êtres humains tentant de simuler le hasard est en général nettement supérieur à $1/2$. Une expérience réelle sur 22 participants avec $N = 64$ a donné les résultats suivants pour le taux d'alternance sous l'hypothèse Θ_1 : $\mathbb{E}_{\theta_1}[\hat{a}] = 0,538$ et $\text{Var}_{\theta_1}[\hat{a}] = 0,112$. En admettant que la loi de $\hat{a}$ est gaussienne (ce qui est clairement faux en pratique), calculer la réponse du test de Neyman–Pearson en fonction de la valeur de $\hat{a}$ pour $p = 1/2$. Évaluer avec une calculatrice les valeurs de α et β pour ce test.

Soit $k \geq 0$ un (petit) entier. Appelons *statistique de prévisibilité à l'ordre k* le nombre :

$$\hat{p}_k = \frac{1}{N - k} \sum_{\ell(s)=k} (\#(s1) - \#(s0))^2.$$

7. Que mesure $\hat{p}_0$? Exprimer $\hat{a}$ en fonction de $\hat{p}_0$ et $\hat{p}_1$.

8. (★) Montrer que sous l'hypothèse nulle, on a *exactement* $\mathbb{E}[\hat{p}_k] = 1$.

9. (★★) Montrer que, quand $N \rightarrow \infty$, la loi de $\hat{p}_k$ converge vers la loi $\chi^2(2^k)/2^k$.

10. (♥) Quel sens donneriez-vous à $\hat{p}_k$? Pourquoi attend-on d'un mauvais simulateur que, pour k valant quelques unités, $\hat{p}_k$ ait tendance à être de l'ordre de grandeur de N ?

11. L'expérience dont nous parlions précédemment a donné, pour $k = 2$, $\mathbb{E}_{\theta_1}[\hat{p}_2] = 1,628$. En admettant que la loi de $\hat{p}_2$ sous l'hypothèse alternative est proportionnelle à un $\chi^2(4)$ (ce qui est pour le coup à peu près raisonnable), refaire la question 6 avec la statistique de test $\hat{p}_2$.

12. Entre les deux tests évoqués, lequel choisiriez-vous pour discerner un générateur véritablement aléatoire d'un être humain ?

☛ *Remarquez que notre expérience ne nous permettait pas réellement de connaître les lois de $\hat{a}$ et $\hat{p}_k$ sous θ_1 : en fait, on s'est servi d'estimateurs à partir de l'expérience effectuée pour déterminer les paramètres de ces lois. Cette technique est courante en statistiques : quand un des paramètres du modèle n'est pas connu parfaitement, mais peut être estimé avec une bonne précision, alors, si de petites variations de ce paramètre ont une faible influence sur le résultat, on fait souvent comme si ce paramètre était exactement égal à sa valeur estimée.*

Feuille 9 : Espérance conditionnelle

27 avril 2009

Exercice 1 (Statistiques bayésiennes)

L'autre jour, un de mes amis est revenu d'un voyage enchanteresse dans un pays scandinave. Malheureusement, il avait oublié de quel pays il s'agissait. Il y avait donc cinq possibilités qui étaient le Danemark, la Finlande, l'Islande, la Norvège et la Suède. Je lui ai posé deux questions pour le savoir :

1. *La langue du pays ressemblait-elle à l'allemand ?*
2. *Tu m'as dit t'être baigné dans la mer : était-ce la Baltique ?*

Il m'a donné, pour chaque question, une réponse qui était « oui », « non » ou « je ne sais pas ». Je me pose la question de statistiques suivante : comment deviner le pays dans lequel il est allé à partir de ses réponses ?

1. Les réponses attendues pour les questions sont les suivantes :

pays	Q1	Q2
Danemark	oui	non
Finlande	non	oui
Islande	non	non
Norvège	oui	non
Suède	oui	oui

En définissant une distance sur l'espace des couples de réponses, donner un algorithme « naïf » pour deviner le pays visité.

2. Soyons un peu plus fin : en fait, les gens se trompent parfois dans leurs réponses. Admettons que, pour un pays donné, les deux réponses soient indépendantes. La façon dont les gens répondent en fonction du pays est la suivante :

pays	Q1			Q2		
	oui	nsp	non	oui	nsp	non
Danemark	70%	25%	5%	25%	20%	55%
Finlande	10%	40%	50%	50%	35%	15%
Islande	20%	40%	40%	5%	20%	75%
Norvège	60%	25%	15%	10%	25%	65%
Suède	60%	25%	15%	60%	30%	10%

En calculant la probabilité d'obtenir une paire de réponses données en fonction du pays, déterminer l'estimateur du maximum de vraisemblance du pays visité par mon ami.

3. (♥) Encore plus fort : je connais les statistiques de tourisme de chacun des cinq pays. Le Danemark est visité annuellement par 2 000 000 touristes, la Finlande par 2 800 000 touristes, l'Islande 700 000, la Norvège 3 500 000 et la Suède 3 000 000. Construire un modèle probabiliste qui modélise le pays visité par mon ami et les réponses qu'il donne. Calculer alors la loi de probabilité du pays conditionnellement aux réponses données, et en déduire l'*estimateur bayésien* du pays visité par mon ami.

4. Aller jouer quelques parties sur <http://www.akinator.fr> et commenter l'efficacité des méthodes bayésiennes :-)

Exercice 2 (Espérance conditionnelle et indépendance)

Soient $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ et deux sous tribus $\mathcal{A}_1$ et $\mathcal{A}_2$ telles que la tribu engendrée par $\mathcal{A}_1$ et $\sigma(X)$ soit indépendante de $\mathcal{A}_2$. On note $\mathcal{B}$ la tribu engendrée par $\mathcal{A}_1$ et $\mathcal{A}_2$.

1. Montrer que si A est de la forme $A_1 \cap A_2$ avec $A_1 \in \mathcal{A}_1$ et $A_2 \in \mathcal{A}_2$, alors on a $\mathbb{E}(1_A X) = \mathbb{E}(1_A \mathbb{E}(X|\mathcal{A}_1))$.

2. Rappeler quel argument nous permet alors d'affirmer que $\mathbb{E}(X|\mathcal{B}) = \mathbb{E}(X|\mathcal{A}_1)$. Comment interprétez-vous ce résultat ?

3. (◇) En déduire que si $\mathcal{A}_1$ et $\mathcal{A}_2$ sont deux sous-tribus indépendantes et que X est $\mathcal{A}_1$ -mesurable et L^1 , alors $\mathbb{E}(X|\mathcal{A}_2) \equiv \mathbb{E}(X)$.

4. (◇) Montrer que la réciproque est également vraie : si $\mathcal{A}_1$ et $\mathcal{A}_2$ sont deux sous-tribus telle que pour tout X $\mathcal{A}_1$ -mesurable et L^1 on ait $\mathbb{E}(X|\mathcal{A}) \equiv \mathbb{E}(X)$, alors $\mathcal{A}_1$ et $\mathcal{A}_2$ sont indépendantes.

Exercice 3 (Espérance conditionnelle par rapport à la somme)

Soit X_n une suite de variables aléatoires réelles indépendantes et de même loi μ . Soit $S_n = \sum_{j=1}^n X_j$.

1. Montrer que $\mathbb{E}(X_1|S_n) = \mathbb{E}(X_i|S_n)$ pour tout $1 \leq i \leq n$.

2. En déduire $\mathbb{E}(X_1|S_n)$.

3. Soit $\mathcal{A}_n = \sigma(S_{n+j} | j \in \mathbb{N})$. Utiliser la question 2 de l'exercice 2 pour calculer $\mathbb{E}(X_1|\mathcal{A}_n)$.

Exercice 4 (Calcul d'espérance conditionnelle)

1. Soit Y une variable aléatoire exponentielle de paramètre 1. Calculer $\mathbb{E}(Y|Y \wedge t)$.

Exercice 5 (Une égalité presque sûre)

1. Soit $Y \in L^2(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que $E(Y|X) = f(X)$ et $E(Y^2|X) = f(X)^2$,

pour une certaine fonction f mesurable. Montrer que $Y = f(X)$ $\mathbb{P}$ -p.s..

Exercice 6 (Variables à densité)

Soit (X, Y) un couple de variables aléatoires réelles de densité $f(x, y)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$.

1. Soit g une fonction borélienne bornée. Calculer $\mathbb{E}(g(Y)|X)$.
2. Soit (X, Y) de loi uniforme sur $[0, 1]^2$ et $S = X + Y$. Calculer la densité de X conditionnellement à $S = s$.

Exercice 7 (Loi conditionnelle)

☛ Dans tout cet exercice, les espérances conditionnelles ne sont pas considérées modulo l'égalité presque-sûre. Il faudra implicitement rajouter « pour tout choix d'espérance conditionnelle » (ou parfois « pour un choix ») dans les questions chaque fois que nécessaire.

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et soit $\mathcal{B}$ une sous-tribu de $\mathcal{F}$. Pour $A \in \mathcal{F}$, notons $\mathbb{P}[A|\mathcal{B}] = \mathbb{E}[\mathbb{1}_A|\mathcal{B}]$.

1. On suppose dans un premier temps que $(\Omega, \mathcal{F})$ est isomorphe à $(\mathbb{N}, \mathfrak{P}(\mathbb{N}))$ en tant qu'espace mesurable. Montrer qu'on a p.s. l'égalité suivante entre fonctions sur Ω

$$\sum_{\omega \in \Omega} \mathbb{P}[\{\omega\}|\mathcal{B}] = 1.$$

2. En déduire qu'on peut définir une famille $(\mathbb{P}_\omega)_{\omega \in \Omega}$ de mesures de probabilité sur $(\Omega, \mathcal{F})$, qui soit $\mathcal{B}$ -mesurable (càd. que pour tout $A \in \mathcal{F}$ l'application $\omega \mapsto \mathbb{P}_\omega[A]$ est $\mathcal{B}$ -mesurable), et telle que pour tout $A \in \mathcal{F}$ on ait $\mathbb{P}_\omega[A] = \mathbb{P}[A|\mathcal{B}](\omega)$ (en tant que fonctions de ω). Montrer que cette famille est unique à équivalence presque-sûre près.

3. (♥, ★) Si $(\Omega, \mathcal{F})$ est un espace polonais non dénombrable muni de sa tribu borélienne, montrer qu'il est isomorphe à $\{0, 1\}^{\mathbb{N}}$ muni de sa tribu borélienne (pour la topologie produit).

À partir de maintenant, on suppose que $(\Omega, \mathcal{F})$ est $\{0, 1\}^{\mathbb{N}}$ muni de sa tribu borélienne.

4. (★) Montrer que, si on se donne une application p de l'ensemble $\{0, 1\}^{(\mathbb{N})}$ des suites finies d'éléments de $\{0, 1\}$ dans $[0, 1]$, telle que $p() = 1$ et pour toute suite s , $p(s) = p(s0) + p(s1)$, alors il existe une mesure de probabilité $\mathbb{P}$ sur $\{0, 1\}^{\mathbb{N}}$ telle que $\mathbb{P}$ attribue la masse $p(s)$ à l'ensemble des éléments de $\{0, 1\}^{\mathbb{N}}$ commençant par s (que l'on notera $s\Omega$), et que cette mesure est unique.

5. Montrer que p.s.,

$$\left\{ \mathbb{P}[\Omega|\mathcal{B}] = 1 \quad \text{et} \quad (\forall s \in \{0, 1\}^{(\mathbb{N})}) (\mathbb{P}[s\Omega|\mathcal{B}] = \mathbb{P}[s0\Omega|\mathcal{B}] + \mathbb{P}[s1\Omega|\mathcal{B}]) \right\}.$$

6. (♥) En déduire qu'il existe une unique (à équivalence p.s. près) famille $(\mathbb{P}_\omega)_{\omega \in \Omega}$ de mesures de probabilité sur $(\Omega, \mathcal{F})$, qui soit $\mathcal{B}$ -mesurable (càd. que pour tout $A \in \mathcal{F}$ l'application $\omega \mapsto \mathbb{P}_\omega[A]$ est $\mathcal{B}$ -mesurable), et telle que pour tout $A \in \mathcal{F}$ on ait $\mathbb{P}_\omega[A] = \mathbb{P}[A|\mathcal{B}](\omega)$. Quand $\mathcal{F} = \sigma(X)$ pour une v.a. X , $\mathbb{P}_\omega$ (à noter l'abus de langage puisque la fonction $\omega \mapsto \mathbb{P}_\omega$ n'est définie qu'à équivalence p.s. près) est appelée *loi conditionnelle de $\mathbb{P}$ sachant que $X = X(\omega)$* .

7. (♥) L'espérance sous la loi conditionnelle est-elle la même chose que l'espérance conditionnelle ?

Exercice 8 (Vecteurs gaussiens et espérance conditionnelle)

Soit (X, Y, Z) un vecteur gaussien à valeurs dans $\mathbb{R}^3$. On suppose que $\text{Var}(X) > 0$, $\text{Var}(Y) > 0$ et que $\mathbb{E}[X] = \mathbb{E}[Y] = \mathbb{E}[XY] = 0$.

1. (★) Montrer que $\mathbb{E}[Z | X, Y] = \mathbb{E}[Z | X] + \mathbb{E}[Z | Y] - \mathbb{E}[Z]$ p.s..
2. Montrer que $\mathbb{E}[X | XY] = 0$ p.s., resp. $\mathbb{E}[Y | XY] = 0$ p.s..
3. En déduire $\mathbb{E}[Z | XY]$.

Exercice 9 (Comment gâcher une bonne soirée)

☞ ATTENTION ! L'exercice suivant est réputé pour engendrer des débats houleux et stériles quand on en discute avec des non-probabilistes...

☞ Dans tout cet exercice, on est implicitement censé modéliser les situations de manière réaliste.

Je sais que ma voisine a deux enfants.

1. Supposons que je sache que l'aînée est une fille. Quelle est la probabilité que l'autre enfant soit également une fille ?

2. (♥) Supposons que je sache que ma voisine a au moins une fille. Quelle est la probabilité que ma voisine ait en fait deux filles ?

3. (♥) Supposons que je sache que parmi les enfants de ma voisine, il y a une fille appelée Mathilde. Quelle est la probabilité que l'autre enfant soit également une fille ? (On fera l'hypothèse que la proportion de filles appelées Mathilde est très faible et on négligera les termes d'ordre inférieur).

Feuille 10 : Martingales

4 mai 2009

Exercice 1 (Étude de la convergence d'une martingale)

Soient $X_1, X_2, \dots$ des variables aléatoires indépendantes telles que

pour tout $n \geq 1$, $\mathbb{P}[X_n = 0] = 1 - 1/n$ et $\mathbb{P}[X_n = 1] = \mathbb{P}[X_n = -1] = 1/(2n)$.

On définit par récurrence $M_0 := 0$, et

$$M_n := \begin{cases} X_n & \text{si } M_{n-1} = 0 \\ nM_{n-1}|X_n| & \text{sinon.} \end{cases}$$

On pose $\mathcal{F}_n = \sigma(X_k, 1 \leq k \leq n)$, la tribu engendrée par $X_1, \dots, X_n$ (pour $n \geq 1$), et $\mathcal{F}_0 := \{\emptyset, \Omega\}$.

1. Montrer que pour tout $n \geq 0$, $\mathbb{E}[|M_n|] < \infty$.
2. Montrer que $(M_n)_n$ est une martingale.
3. Soit $A := \{\text{Il existe une infinité de } n \text{ tels que } X_n = 0, X_{n+1} = 1\}$. Montrer que $\mathbb{P}[A] = 1$.
4. Prouver que $\overline{\lim}_n M_n \geq 1$.
5. Calculer $\mathbb{P}[M_n \neq 0 \mid \mathcal{F}_{n-1}]$.
6. Étudier la convergence en probabilité de M_n .
7. Étudier la convergence presque sûre de M_n .
8. En considérant $\mathbb{E}[|M_n| \mid \mathcal{F}_{n-1}]$, établir une relation de récurrence pour $\mathbb{E}[|M_n|]$.
9. Étudier la convergence dans $L^1(\Omega, \mathcal{F}, \mathbb{P})$ de M_n .

Exercice 2 (Variation quadratique)

Soient $X_1, X_2, \dots$ des variables indépendantes, chacune étant L^2 et centrée.

1. Montrer que

$$\left(\sum_{i=1}^n X_i \right)^2 - \sum_{i=1}^n \mathbb{E}[X_i^2]$$

est une martingale pour la filtration naturelle associée aux X_i .

Exercice 3 (Fonction harmonique)

On considère la marche aléatoire biaisée sur $\mathbb{Z}$, c.à.d. une chaîne de Markov sur $\mathbb{Z}$ telle que $\mathbb{P}(X_{n+1} = i + 1 | X_n = i) = p$ et $\mathbb{P}(X_{n+1} = i - 1 | X_n = i) = 1 - p$, pour un $p \in]1/2, 1[$.

1. Déterminer quelles sont les fonctions $f : \mathbb{Z} \rightarrow \mathbb{R}$ telles que $f(X_n)$ soit une martingale.

☛ Ces fonctions sont appelées fonctions harmoniques pour la marche aléatoire biaisée.

Je fais un combat de fleuret contre un ami qui est plus fort que moi : à chaque assaut il a une probabilité $p \in]1/2, 1[$ de faire la touche (il n'y a pas de double touche en fleuret). La règle est que le premier qui a k touches d'avances a gagné. Ayant été chanceux, je me retrouve à mener par ℓ touches d'avance ($0 \leq \ell < k$).

2. Quelle est ma probabilité de remporter le duel ? Application numérique avec $p = 60\%$, $k = 3$ et $\ell = 1$.

Exercice 4 (Inégalité de Doob)

Soit $(M_n)_{n \in \mathbb{N}}$ une martingale qui stationne presque-sûrement ; on note M_∞ la valeur terminale de la martingale. On pose également

$$M_*(\omega) = \sup_{n \in \mathbb{N}} |M_n(\omega)|.$$

Le but de l'exercice est de borner $\mathbb{E}[M_*^2]$ en fonction de $\mathbb{E}[M_\infty^2]$. On supposera ici que M_n est uniformément bornée par une variable L^2 .

1. ($\diamond$) Montrer que $|M_n|$ est une sous-martingale.
2. Montrer que pour tout $A \geq 0$, $\mathbb{E}[\mathbb{1}_{M_* \geq A} |M_\infty|] \geq A \mathbb{P}(M_* \geq A)$.
3. ($\heartsuit$) Montrer que si X est une v.a. positive,

$$\mathbb{E}[X^2] = 2 \int_0^\infty x \mathbb{P}(X \geq x) dx.$$

4. ($\star$) En déduire l'inégalité de Doob :

$$\mathbb{E}[M_*^2] \leq 4 \mathbb{E}[M_\infty^2] \tag{1}$$

(Indication : utiliser l'inégalité de Cauchy-Schwarz).

5. ($\star\star$) Montrer que la constante 4 dans (1) ne peut pas être améliorée.