ENS de Lyon
L3 - Intégration et Probabilités

2nd semestre 2019-2020

TD 4
Caractérisation de lois, indépendance

On admettra le résultat suivant qui sera vu dans le prochain cours : si X et Y sont des variables aléatoires indépendantes, et intégrables, alors XY est intégrable et $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. Cela implique en particulier que $\varphi_{X+Y} = \varphi_X\varphi_Y$, et que $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Exercice 1 *Échauffement*

1. Calculer la loi de la somme de deux variables de Poisson indépendantes.
2. Calculer la loi du minimum de deux lois exponentielles indépendantes.
3. On suppose $\mathbb{P}[A \cap B \cap C] = \mathbb{P}[A]\mathbb{P}[B]\mathbb{P}[C]$. Est-ce que A, B et C sont indépendants ?
4. On suppose que A et ${}^cB$ sont deux événements indépendants. Est-ce que A et B sont indépendants ?
5. Soit F la fonction de répartition de X. On suppose F de classe $\mathscr{C}^1$. Montrer que X est à densité et calculer celle-ci.
6. On suppose F continue. Est-ce que X est à densité ?

Exercice 2 *Lois sur* $\mathbb{N}^*$

La fonction ζ de Riemann est définie pour $s > 1$ par $\zeta(s) = \sum_{n=1}^{+\infty} n^{-s}$. On munit $\mathbb{N}^*$ de la tribu totale $\mathcal{P}(\mathbb{N}^*)$ et de la mesure de probabilité $\mathbb{P}$ définie par $\mathbb{P}(\{x\}) = \frac{1}{\zeta(s)}x^{-s}$, $x \in \mathbb{N}^*$. Pour p premier, on pose $A(n) = n\mathbb{N}^*$ l'événement « être multiple de n ».

1. Justifiez que $\mathbb{P}$ est bien une mesure de probabilité et montrez que pour $p_1, \dots, p_k$ premiers distincts, les évenements $A(p_1), \dots, A(p_k)$ sont indépendants.
2. Montrer la formule d'Euler $\frac{1}{\zeta(s)} = \prod_{p \text{ premier}} \left(1 - \frac{1}{p^s}\right)$.
3. Montrer qu'il n'existe pas de mesure de probabilité $\mathbb{P}$ sur $\mathbb{N}^*$ "bien répartie", au sens que $\mathbb{P}(A(n)) = 1/n$.

Exercice 3 *Convergence*

Vous avez parlé en cours de l'éspérance du nombre de cycles C_n dans une permutation uniforme. Que dire de $\frac{C_n}{\log n}$?

Exercice 4 *Inégalité d'Essen et marches aléatoires.*

On rappelle l'inégalité d'Esséen (*c.f.* TD précédent). Il existe une constante $C > 0$ telle que pour toute variable aléatoire X à valeurs réelles, et pour tout intervalle I de longueur 1, on a

$$\mathbb{P}(X \in I) \leq C \int_{|t|\leq 1} |\mathbb{E}[e^{itX}]| dt.$$

Soient $(X_k)_{k\geq 1}$ une suite de variables aléatoire i.i.d. avec $\mathbb{P}(X_1 = -1) = \mathbb{P}(X_1 = 1) = 1/2$, et soit $(a_k)_{k\geq 1}$ une suite de réels (déterministes). Pour tout entier $n \geq 1$, on définit $S_n = \sum_{1\leq k\leq n} a_k X_k$. On pose

1. Montrer qu'il existe deux constantes $A, B > 0$ universelles telle que pour tout $x \in \mathbb{R}$, $\mathbb{P}(|S_n - x| < 1/2) \leq (A + B\lambda/\delta)/\sqrt{m_n}$, où $m_n = \#\{k \leq n : |a_k| \geq \delta\}$.
   Indication : on pourra utiliser l'estimée de l'intégrale de Wallis : $\int_0^{\pi/2} \cos(u)^n du = \sqrt{\pi/2n} + o(1/\sqrt{n})$.
   Remarque : quand les $|a_k|$ *sont minorés par* 1*, on a en particulier que* $\mathbb{P}(|S_n - x| \leq 1) \leq \frac{C}{\sqrt{n}}$ *(Problème de Littlewood-Offord, résolu par Erdös en* 1945*).*

2. (*) On considère le polynôme aléatoire $P_n(x) = \sum_{i=0}^n X_k x^k$. Montrer que la probabilité que le nombre de racines réelles de P_n dépasse $C\log(n)^4$ tend vers zéro. On pourra d'abord se ramener à $[1/2, 1]$ puis utiliser $\log_2(n)$ fois l'inégalité de Jensen : le nombre de racines d'une fonction analytique sur $D(x,r)$ est majoré par

$$\frac{1}{\log(R/r)} \log\left(\frac{\|f\|_{D(x,R)}}{|f(x)|}\right).$$

Remarque : le bon ordre de croissance du nombre de racines est en réalité $\frac{2}{\pi}\log(n)$ *(Kác 1943, Erdös-Offord, 1956)*

3. Sur ce sujet mais dans $\mathbb{C}$, allez voir les jolies images de `http://math.ucr.edu/home/baez/roots/`.

Exercice 5 *Somme aléatoire*

Dans cet exercice, soit $(Y_i)_i$ une suite i.i.d. et N une variable aléatoire indépendente de la suite $(Y_i)_i$. On pose $X = \sum_{i=1}^N Y_i$.

1. Si Y_1 et N sont intégrables, montrer que $\mathbb{E}[X] = \mathbb{E}[N]\mathbb{E}[Y_1]$.
2. Si Y_i est à valeurs entières, calculer la fonction génératrice de X.
3. Si $Y_i \sim \text{Geom}(p)$, et $N \sim \text{Geom}(a)$. Calculer la loi de X puis donner une interprétation avec une suite de lancers de deux pièces.
4. Avec $Y_i \sim \text{Be}(p)$ et $N \sim \text{Po}(\lambda)$. Montrer également que $X \perp\!\!\!\perp N - X$.

Exercice 6 Maximum de variables à densité

Soient X, Y deux variables aléatoires réelles indépendantes et à densité. Montrer que $\max(X,Y)$ est à densité et calculer celle-ci.

Exercice 7 L'inégalité de Tchebychev est-elle optimale ?

1. Oui. Montrer que si $a > 0$, alors il existe une variable aléatoire réelle X de variance finie non-nulle telle que $\mathbb{P}\Big[|X - \mathbb{E}[X]| \geq a\Big] = \frac{\text{Var}(X)}{a^2}$.
2. Non. Démontrer que si X est une variable aléatoire de carré sommable, alors $a^2\, \mathbb{P}\Big[|X - \mathbb{E}[X]| \geq a\Big] \underset{a\to+\infty}{\longrightarrow} 0$.
3. Presque. Montrer que, pour tout $\varepsilon > 0$, il existe une variable aléatoire réelle X de carré intégrable telle que $a^{2+\varepsilon}\, \mathbb{P}\Big[|X - \mathbb{E}[X]| \geq a\Big] \underset{a\to+\infty}{\longrightarrow} \infty$.

Exercice 8 Une condition forte

Soient X et Y deux variables aléatoires indépendantes, à valeurs dans $[0,1]$.
On suppose que $\mathbb{P}\big[X \leq Y\big] = 1$. Montrer qu'il existe $a \in \mathbb{R}$ tel que $\mathbb{P}\big[X \leq a \leq Y\big] = 1$.

Exercice 9 Lemme de Schwarz-Zippel

Soit K un corps et $P(X_1, \ldots, X_n)$ un polynôme dans $K[X_1, \ldots, X_n]$. On suppose que P est de degré $d \geq 0$, et en particulier, P est non-nul.

Soit S une partie finie de K et $R_1, \ldots, R_n$ des variables aléatoires indépendantes uniformes dans S. Montrer que

$$\mathbb{P}(P(R_1, \ldots, R_n) = 0) \leq \frac{d}{|S|}.$$

On pourra raisonner par récurrence sur n.

Exercice 10 *Modèle du vendeur de journaux*

Chaque jour, vous achetez un certain nombre de produits à un prix a et les vendez à un prix b. Sachant que la demande est aléatoire selon une certaine loi μ, et que ceux que vous n'avez pas vendu à la fin de la journée sont perdus, comment fixer le nombre de produits pour maximiser le gain moyen ?