Bastien Mallein
Bureau V2
`bastien.mallein@ens.fr`

TD 1 – Rappels et probabilités discrètes

28 septembre 2016

Exercice 1 (Paul Valéry)**.** On se donne 100 segments dans le plan, de longueur unité. Chaque segment ayant une direction fixée, montrer qu'il est possible de coller ces segments bout à bout de façon à ce que la distance entre les deux extrémités de la chaîne soit inférieure à 10.

Solution. *« Un homme tirait au sort toutes ses décisions. Il ne lui arriva pas plus de mal qu'aux autres qui réfléchissent. »* Paul Valéry – Choses Tues

Soit $v_1, \ldots v_{100}$ des nombres complexes de module 1, constitués en ajoutant un sens arbitraire aux segments fournis, et $e_1, \ldots e_{100}$ des variables aléatoires i.i.d. vérifiant $\mathbb{P}(e_1 = 1) = \mathbb{P}(e_1 = -1) = 1/2$. On pose $S = \sum_{i=1}^{100} e_i v_i$, et on observe que

$$\mathbb{E}(S.\bar{S}) = \mathbb{E}\left[\|S\|_2^2\right] = 100.$$

On en déduit $\mathbb{P}(\|S\|_2 \leq 10) > 0$, ce qui conclut. □

Exercice 2 (Fonctions Cesàro-continue)**.** On dit qu'une suite (x_n) Cesàro-converge vers x si

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=1}^{n} x_j = x.$$

Une fonction $f : \mathbb{R} \to \mathbb{R}$ est dite Cesàro-continue si pour toute suite (x_n) Cesàro-convergeant vers x, la suite $(f(x_n))$ Cesàro-converge vers $f(x)$.

Déterminer l'ensemble des fonctions mesurables Cesàro-continues.

Solution. Soit $x, y \in \mathbb{R}^2$ et $t \in (0, 1)$. On note (X_n) des variables aléatoires i.i.d. telles que

$$\mathbb{P}(X_1 = x) = 1 - \mathbb{P}(X_1 = y) = t$$

et $S_n = X_1 + \cdots + X_n$. Grâce à la loi des grands nombres, on a

$$\lim_{n\to+\infty} \frac{S_n}{n} = tx + (1-t)y \text{ et } \lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) = (tf(x) + (1-t)f(y)) \text{ p.s.}$$

Par conséquent f préserve les barycentres, ce qui montre que f est affine. □

Exercice 3 (Prise de décision)**.** Camille et Dominique ont deux chèques, et se proposent de les partager de la façon suivante. Camille donne un chèque à Dominique, qui regarde le montant et décide ensuite d'échanger, ou non avec Camille. Proposer à Dominique une stratégie qui lui permette d'avoir le plus gros chèque avec probabilité strictement plus grande que $\frac{1}{2}$.

Solution. Il suffit à Dominique de choisir un nombre X selon une loi exponentielle. Si le montant du chèque proposé est plus petit que X, elle échangera. $\square$

Exercice 4 (Transformée de Verwaat)**.** Soit $n \in \mathbb{N}$, on note l'ensemble des chemins de longueur n

$$\mathcal{C}_n = \left\{(c_0, \dots c_n) \in \mathbb{Z}^{n+1} : c_0 = 0, \forall 1 \leq k \leq n, |c_k - c_{k-1}| = 1\right\}.$$

Pour tout $k \leq n$, on note

$$\Phi^k : c \in \mathcal{C}_n \mapsto (0, c_{k+1} - c_k, c_{k+2} - c_k, \dots, c_n - c_k, c_n - c_k + c_1, c_n - c_k + c_2, \dots, c_n - c_k + c_k).$$

1. Montrer que Φ^k est une bijection de $\mathcal{C}_n$ qui préserve la valeur prise par c_n.
2. Soit S une marche aléatoire au plus proche voisin, en déduire

$$\mathbb{P}(S_0 \geq 0, \dots S_{2n} \geq 0, S_{2n+1} = -1) = \frac{1}{2n+1}\mathbb{P}(S_{2n+1} = -1).$$

3. En conclure

$$\mathbb{P}(S_1 \neq 0, \cdots S_{2n-1} \neq 0, S_{2n} = 0) = \frac{1}{2n-1}\mathbb{P}(S_{2n} = 0).$$

Solution.

1. On observe que $\Phi^{-k} = \Phi^{n-k}$, d'où le caractère bijectif.
2. On note $\mathcal{C}_{2n+1}(-1) = \{c \in \mathcal{C}_{2n+1} : c_{2n+1} = -1\}$. Pour tout $c \in \mathcal{C}_{2n+1}(-1)$, il existe un unique entier k (correspondant au premier temps d'atteinte du minimum de c) tel que $\Phi^k(c)$ reste positif sur $[0, 2n]$. On en conclut,

$$\begin{aligned}
\mathbb{P}(S_{2n+1} = -1) &= \mathbb{P}\left[\exists k \leq 2n : \Phi^k(S)_1 \geq 0, \dots \Phi^k(S)_{2n} \geq 0, \Phi^k(S)_{2n+1} = -1\right] \\
&= \sum_{k=0}^{2n} \mathbb{P}\left[\Phi^k(S)_1 \geq 0, \dots \Phi^k(S)_{2n} \geq 0, \Phi^k(S)_{2n+1} = -1\right] \\
&= (2n+1)\mathbb{P}(S_1 \geq 0, \dots S_{2n} \geq 0, S_{2n+1} = -1).
\end{aligned}$$

3. Connaissant le premier pas, on considère le chemin parcouru par S.

$\square$

Exercice 5 (Distribution de particules)**.** On répartit au hasard r boules dans n cases. Soit $r_1, \cdots r_n$ des entiers tels que $r_1 + \cdots + r_n = r$, où r_j décrit le nombre de boules dans la case j.

1. Expliciter la loi de $(r_1, \dots r_n)$ dans chacun des cas suivants.
   (a) On suppose que toutes les configurations de boules sont possibles et équiprobables. Cette distribution s'appelle la « statistique de Maxwell-Boltzmann ».
   (b) On suppose que toutes les configurations possibles pour les nombres de boules dans chaque case sont équiprobables. Cette distribution s'appelle « statistique de Bose-Einstein ».
   (c) On suppose qu'il ne peut y avoir plus d'une boule dans chaque case et que toutes les statistiques de boules autorisées sont équiprobables. Cette distribution s'appelle « statistique de Fermi-Dirac ».
2. Quelle est la probabilité que parmi r personnes, au moins deux personnes aient leur anniversaire le même jour ? Pour quelles valeurs de r cette probabilité est-elle supérieure à $0,5$? à $0,99$?
3. Supposons qu'en Bretagne, il ait eu r averses en une semaine. Quelle est la probabilité qu'il ait plu au moins deux fois le même jour ?

4. Le fichier .tex de ce TD contient 7042 caractères, parmi lesquels r coquilles[1]. Quel modèle est pertinent pour la loi des coquilles ?

Solution. Soit $r_1, \dots r_n$ des entiers positifs ou nuls tels que $r_1 + \cdots + r_n = r$.

1. Pour la statistique de Maxwell-Boltzmann, on a $p(r_1, \dots r_n) = \frac{1}{n^r}\binom{r}{r_1}\binom{r-r_1}{r_2}\cdots\binom{r-r_1-\cdots-r_{n-1}}{r_n}$. Pour la statistique de Bose-Einstein, $p(r_1, \cdots r_n) = \frac{1}{\binom{n+r-1}{r}}$. Pour la statistique de Fermi-Dirac, on a $p(r_1, \cdots r_n) = \frac{1}{\binom{n}{r}}$.
2. On a $\mathbb{P}(\max(r_1, \dots r_n) \geq 2) = 1 - \mathbb{P}(\max(r_1, \dots r_n) \leq 1) = 1 - \frac{364}{365}\frac{363}{365}\cdots\frac{365-r+1}{365}$. En particulier, la probabilité est plus grande que 0,5 dès que $r \geq 23$, plus grande que 0,99 dès que $r \geq 57$.
3. On a $\mathbb{P}(\max(r_1, \dots r_n) \geq 2) = 1 - \mathbb{P}(\max(r_1, \dots r_n) \leq 1) = 1 - \frac{\binom{n}{r}}{\binom{n+r-1}{r}}$.
4. Laissé au lecteur.

□

Exercice 6 (Marche aléatoire sur le cercle)**.** Soit (X_n) une suite de variables aléatoires réelles i.i.d. On note $S_n = X_1 + \cdots + X_n$ mod 1. Déterminer la limite de S_n quand $n \to +\infty$.

Solution. On calcule la fonction caractéristique de $\bar{S}_n$. Pour tout $k \in \mathbb{N}$, on a

$$\mathbb{E}\left[e^{i2k\pi\bar{S}_n}\right] = \mathbb{E}\left[e^{i2k\pi S_n}\right] = \mathbb{E}\left[e^{i2k\pi X_1}\right]^n.$$

Par conséquent, on a

$$\lim_{n\to+\infty} \mathbb{E}\left[e^{i2k\pi\bar{S}_n}\right] = \begin{cases} 0 & \text{si} \quad \left|\mathbb{E}\left[e^{i2k\pi X_1}\right]\right| < 1 \\ 1 & \text{si} \quad \mathbb{E}\left[e^{i2k\pi X_1}\right] = 1 \\ \text{n'existe pas} & \text{sinon.} \end{cases}$$

On suppose pour commencer que pour tout $k \in \mathbb{N}^*$, $\left|\mathbb{E}\left[e^{i2k\pi X_1}\right]\right| < 1$. Dans ce cas, $\bar{S}_n$ converge vers la mesure uniforme sur $[0,1]$, comme on peut s'en convaincre en se ramenant aux fonctions développables en séries de Fourier.

On suppose maintenant qu'il existe $k \in \mathbb{N}^*$ tel que $\left|\mathbb{E}\left[e^{i2k\pi X_1}\right]\right| = 1$. On note

$$T = \inf\left\{k \in \mathbb{N} : \left|\mathbb{E}\left[e^{i2k\pi X_1}\right]\right| = 1\right\}.$$

Si $\mathbb{E}\left[e^{i2T\pi X_1}\right] = e^{i\theta}$, la suite des transformées de Fourier de S_n ne converge pas, donc S_n ne converge pas en loi.

Si $\mathbb{E}\left[e^{i2T\pi X_1}\right] = 1$, cela signifie que $\mathbb{P}(X_1 \in \{n/T, n \in \mathbb{Z}\}) = 1$. On observe alors que

$$\forall k \in \mathbb{Z}, \mathbb{E}\left[e^{2ik\pi X_1}\right] \begin{cases} = 1 & \text{si} \quad T|k \\ \text{est de norme inférieure à 1} & \text{sinon.} \end{cases}$$

En utilisant encore les fonctions développables en séries de Fourier, on montre que

$$S_n \Longrightarrow \text{Unif}(1/T, 2/T, \dots 1\}.$$

□

1. Question subsidiaire : trouver r.

Exercice 7 (Référendum versus Conclave)**.** Soit $n \in \mathbb{N}$, $p \in (1/2, 1]$ et $(X_1, \cdots, X_n)$ des variables aléatoires i.i.d. de loi de Bernoulli de paramètre p.

1. On note $p_n = \mathbb{P}(X_1 + \cdots + X_n > n/2)$. Montrer que p_{2n+1} est croissant avec n et tend vers 1.
2. Soit $a \in \mathbb{R}_+$, et $p = 1/2 + an^{-1/2}$. Calculer la limite de $p_n(a)$.
3. On suppose que $n = 3^m$. On note, pour $1 \leq k \leq 3^m$, $Y_{k,0} = X_k$ et
$$\forall 1 \leq \mu \leq m, \forall k \leq 3^\mu, Y_{k,\mu} = \operatorname{sgn}\left(Y_{3k-2,\mu-1} + Y_{3k-1,\mu-1} + Y_{3k,\mu-1} - 1,5\right).$$
On note $Y_m = Y_{1,m}$. Expliciter $q_m = \mathbb{P}(Y_m = 1)$, ainsi que $\lim_{m\to+\infty} q_m$.
4. Soit $a \in \mathbb{R}_+$ et $\alpha \in (0,1)$ on pose $p = 1/2 + an^{-\alpha}$, pour quelles valeurs de $\alpha \in (0,1)$ la fonction $a \mapsto \lim_{m\to+\infty} q_m(a, \alpha)$ est-elle non-constante ?

Solution.

1. Par simple calcul, on a
$$\begin{aligned} p_{2n+1} &= \sum_{k=n+1}^{2n+1} \left[\binom{2n-1}{k-2} + 2\binom{2n-1}{k-1} + \binom{2n-1}{k}\right] p^k(1-p)^{2n+1-k} \\ &= p_{2n-1} + \binom{2n-1}{n} p^n(1-p)^n(2p-1). \end{aligned}$$

2. On note $S_n = \mathcal{B}(n, 1/2 + an^{-1/2})$ et $R_n = \frac{S_n - n/2}{n^{1/2}}$, pour tout $\xi \in \mathbb{R}$, on a
$$\begin{aligned} \mathbb{E}\left[e^{i\xi R_n}\right] &= \left[e^{i\xi/2n^{1/2}}\left(\tfrac{1}{2} + an^{-1/2}\right) + e^{-i\xi/2n^{1/2}}\left(\tfrac{1}{2} - an^{-1/2}\right)\right]^n \\ &\underset{n\to+\infty}{\longrightarrow} e^{ia\xi - \xi^2/8}. \end{aligned}$$
Par conséquent R_n converge en loi vers une variable aléatoire Gaussienne $\mathcal{N}(a, 1/4)$. Dès lors
$$\lim_{n\to+\infty} p_n(a) = \mathbb{P}(\mathcal{N}(0,1)) \geq -2a).$$

3. Observons que $q_0 = p$ et $q_{m+1} = q_m^3 + 3q_m^2(1 - q_m)$. En étudiant $f(x) = x^3 + 3x^2(1-x)$, on en déduit que pour tout $p > 1/2$,
$$\lim_{m\to+\infty} f^m(p) = 1.$$

4. De plus, on observe que $f'(1/2) = 3/2$, par conséquent, la valeur de α « critique » est
$$\frac{\ln(3/2)}{\ln 3} = 1 - \frac{\ln 2}{\ln 3}.$$

□

Exercice 8 (Easter Egg)**.** Claude s'est inscrit sur un site de rencontres afin de trouver l'être aimé. On lui propose n partenaires successivement. Après chaque proposition, Claude peut soit accepter la proposition et arrêter de chercher, soit décliner et choisir parmi les suivants. Il est interdit de revenir en arrière après avoir dit non. Claude, ayant des goûts très sûrs, sait immédiatement classer la proposition par rapport aux précédents, qui sont faites uniformément au hasard par le site.

Quelle est la stratégie qui permet à Claude de maximiser la probabilité de tomber sur son âme soeur, le meilleur choix possible ?