TD 09 – Méthode probabiliste (corrigé)

Exercice 1. *TCL*

Rappel 1 : Étant donnée une variable aléatoire discrète X à valeurs entières, on appelle *fonction génératrice de* X la fonction $G_X(z) := \mathbf{E}\left[z^X\right]$.

1. Soient X et Y deux variables aléatoires discrètes à valeur dans $\mathbb{N}$.
   Si X et Y sont indépendantes, que peut-on dire de G_{X+Y} ?

   ☞ Si X, Y sont indépendantes et à valeur dans $\mathbb{N}$, on a $G_{X+Y} = G_X G_Y$.

Rappel 2 : Loi de Poisson : Soit X qui suit une loi de Poisson de paramètre λ, alors

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}. \tag{1}$$

2. Montrer qu'une somme de n variables indépendantes de Poisson de paramètre 1 suit une loi de Poisson de paramètre n

   ☞ On calcule la fonction génératrice de $X \sim \mathcal{P}(\lambda)$

$$G_X(z) = \sum_k \frac{z^k \lambda^k}{k!} e^{-\lambda} = e^{\lambda z - \lambda}. \tag{2}$$

   Et si $X = \sum_{i=1}^n X_i$ avec $(X_i)_i$ iid $X_i \sim \mathcal{P}(\lambda)$. On a

$$G_{X_i}(z) = e^{z-1}. \tag{3}$$

$$G_X(z) = \prod_{i=1}^n G_{X_i}(z) = e^{n(z-1)}. \tag{4}$$

   donc $X \sim \mathcal{P}(n)$.

3. Calculer à l'aide du théorème central limite

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^n \frac{n^k}{k!}. \tag{5}$$

   ☞ On a

$$e^{-n} \sum_{k=0}^n \frac{n^k}{k!} = \sum_{k=0}^n \mathbb{P}(X = k) \tag{6}$$

$$= \mathbb{P}(X \leq n) \tag{7}$$

   Par le TCL, on a que

$$\frac{X - n}{\sqrt{n}} \to_{\mathcal{L}} Z \sim \mathcal{N}(0,1) \tag{8}$$

   Or

$$\mathbb{P}(X \leq n) = \mathbb{P}\left(\frac{X-n}{\sqrt{n}} \leq 0\right) \tag{9}$$

   Et on dit que $Z_n \to_{\mathcal{L}} Z$ si $\forall t$

$$\lim_n \mathbb{P}(Z_n \leq t) = \mathbb{P}(Z \leq t) \tag{10}$$

   Donc pour $Z_n := \frac{X-n}{\sqrt{n}}$ et $Z \sim \mathcal{N}(0,1)$, on a

$$\lim_n e^{-n} \sum_{k=0}^n \frac{n^k}{k!} = \lim_n \mathbb{P}(Z_n \leq 0) \tag{11}$$

$$= \lim \mathbb{P}(Z \leq 0) \tag{12}$$

$$= \frac{1}{2}. \tag{13}$$

Exercice 2. *Conditions de convergence*

On se donne X_i une suite infinie de bits aléatoires non biaisés.

1. Montrer que presque sûrement tout mot fini apparaît dans la suite X_i.

☞ Soit w un mot fini. On définit A_j (pour $j \geq 1$) l'événement "$X_{kj+1} \cdots X_{k(j+1)} = w$". On a $\mathbf{P}\{A_j\} = \frac{1}{2^k}$. Par indépendance des X_i (et lemme de groupement par paquets), les événements A_j sont indépendants. On a alors

$$\begin{aligned}\mathbf{P}\{w \text{ n'apparaît pas dans la suite } X_i\} &\leq \mathbf{P}\left\{\cap_{1\leq j\leq\infty}\overline{A_j}\right\} \\ &= \lim_{N\to\infty} \mathbf{P}\left\{\cap_{1\leq j\leq N}\overline{A_j}\right\} \\ &= \lim_{N\to\infty}\left(\frac{2^k-1}{2^k}\right)^N \\ &= 0\end{aligned}$$

Donc w apparaît dans la suite X_i avec probabilité 1.

On note maintenant Y_w l'événement "le mot w apparaît dans la suite X_i". On a vu que pour tout w fini, $\mathbf{P}\{Y_w\} = 1$. De plus, il y a un nombre dénombrable de mots finis, i.e. un nombre dénombrables d'événements Y_w. On en déduit que $\mathbf{P}\{\cap_w Y_w\} = 1$ (en passant par le complémentaire c'est plus propre : $\mathbf{P}\left\{\cup_w \overline{Y_w}\right\} \leq \sum_w \mathbf{P}\left\{\overline{Y_w}\right\} = 0$).

2. En déduire que la presque sûrement tout mot fini apparaît une infinité de fois dans la suite X_i.

☞ Tout mot fini est le sous-mot d'une infinité de mots fini distincts. Donc si une séquence infinie contient tout mot fini, elle contient tout mot fini une infinité de fois (car elle contiendra tous les sur-mots d'un mot fixé, et qu'il y en a une infinité).

Exercice 3. *Un c'est bien, deux c'est mieux*

Deux cent étudiant·es participent à un concours de maths. Le concours comporte 6 questions. Pour chaque question, au moins 120 étudiant·es ont réussi à répondre correctement. Montrer qu'il existe deux étudiant·es qui avaient tout bon à elleux deux (*i.e.* tels que pour chaque question, au moins un·e des étudiant·es a bien répondu).

☞ On considère une distribution uniforme sur les couples d'étudiant·es avec remise. Soient u et v deux entiers entre 1 et 200 tirés uniformément au hasard (et indépendamment). Pour $1 \leq i \leq 6$, on note $u[i] = 1$ si l'étudiant·e u a bien répondu à la question i, et $u[i] = 0$ sinon. On sait que pour tout i,

$$\mathbf{P}\{u[i] = 0\} \leq \frac{80}{200} = \frac{2}{5},$$

où la probabilité est prise sur le choix de u. De plus, on a :

$$\mathbf{P}\{v[i] = 0 \mid u[i] = 0\} \leq \frac{79}{199} < \frac{80}{200}.$$

On en déduit :

$$\mathbf{P}\{u[i] = 0 \text{ et } v[i] = 0\} \leq \frac{4}{25}.$$

D'où, par borne de l'union,

$$\mathbf{P}\{\exists i \text{ t.q. } u[i] = 0 \text{ et } v[i] = 0\} \leq \frac{24}{25} < 1.$$

Donc il existe un choix de u et v (*i.e.* deux étudiant·es) tels que pour toute question, $u[i] = 1$ ou $v[i] = 1$.

Attention : les étudiant·es sont choisi·es aléatoirement mais leurs réponses ne sont pas nécessairement indépendantes ! Une fois les copies corrigées, on a des informations en plus sur la répartition des notes qui font que ni les réponses de deux étudiant·es à la même question, ni les réponses d'un·e étudiant·e à deux questions, ne sont indépendantes.

Exercice 4. *Union d'intervalles*

Soit S une union d'intervalles inclus dans le segment $[0;1]$. On suppose que la longueur totale de S est strictement supérieure à $\frac{1}{2}$. Montrer qu'il existe deux points $x, y \in S$ tels que $|x - y| = 0,1$.

☞ On tire x uniformément au hasard dans $[0;1]$. On définit *(faire un dessin pour expliquer si nécessaire)* :

$$y = \begin{cases} x + 0,1 & \text{si } \lfloor 10x \rfloor \text{ est pair,} \\ x - 0,1 & \text{sinon.}\end{cases}$$

Alors, y est uniformément distribué dans $[0;1]$, mais il n'est pas indépendant de x ; et on a $|x - y| = 0,1$. Par borne de l'union,

$$\mathbf{P}\{x \notin S \text{ ou } y \notin S\} \leq 2(1-t)$$

où t est la longueur totale de S. Or par hypothèse $1 - t < \frac{1}{2}$, donc cette probabilité est strictement inférieure à 1.

On en déduit qu'il existe x et y dans S à distance $0,1$.

Exercice 5. *Un problème complexe*

Soit $a, b \in \mathbb{C}$ et $P = z^2 + az + b$ un polynôme de degré 2 tels que pour tout $z \in \mathbb{C}$, $|z| = 1 \Rightarrow |P(z)| = 1$. Montrer que $a = b = 0$.
Indication : *on pourra considérer* $\mathbf{E}\left[|P(Z)|^2\right]$*, où* Z *est choisi uniformément sur le cercle unité.*
☞ On définit Z une variable aléatoire uniforme sur le cercle unité complexe. En utilisant le fait que $|Z| = 1$, on a :

$$|P(Z)|^2 = (Z^2 + aZ + b)\overline{(Z^2 + aZ + b)} = 1 + |a|^2 + |b|^2 + 2\mathrm{Re}(\bar{a}Z) + 2\mathrm{Re}(\bar{b}Z^2) + 2\mathrm{Re}(a\bar{b}Z)$$

Mais l'espérance de chacune des parties réelles est nulle (c'est une variable centrée en zéro quand Z parcourt le cercle unité, on peut aussi le voir en écrivant la définition de l'espérance et des parties réelles avec les sinus). D'où $\mathbf{E}\left[|P(Z)|^2\right] = 1 + |a|^2 + |b|^2$. Mais comme $|P(z)| = 1$ pour tout z (par hypothèse), on sait aussi que $\mathbf{E}\left[|P(Z)|^2\right] = 1$. D'où $a = b = 0$.

Exercice 6. *Lemme local de Lovász*

Soit $k > 6$. On se donne une famille $(A_i)_{i \in I}$ de sous-ensembles d'un ensemble fini F telle que :

1. Pour tout $i \in I$, $|A_i| = k$,
2. Pour tout $x \in F$, $|\{i \in I \mid x \in A_i\}| \leq \frac{2^k}{8k}$.

En utilisant le lemme local de Lovász [1], montrer qu'il existe une partition $F = F_1 \cup F_2$ telle que

$$\forall i \in I, \qquad A_i \cap F_1 \neq \emptyset \qquad \text{et} \qquad A_i \cap F_2 \neq \emptyset\,.$$

Lemme Local de Lovász (rappel) : *Soient n, d des entiers, $0 \leq p \leq 1$ et $A_1, \ldots, A_n$ des événements tels que :*

1. *pour tout $1 \leq i \leq n$, on a $\mathbf{P}\{A_i\} \leq p$,*
2. *les événements $(A_i)_{1 \leq i \leq n}$ admettent un graphe de dépendance de degré $\leq d$,*
3. *on a $4dp \leq 1$.*

Alors on a $\mathbf{P}\left\{\overline{A_1} \cap \ldots \cap \overline{A_n}\right\} > 0$.
☞ On partitionne F au hasard en décidant indépendamment et uniformément si chaque $x \in F$ est dans F_1 ou F_2.
Pour tout $i \in I$, soit E_i l'événement $\{A_i \cap F_1 = \emptyset\} \cup \{A_i \cap F_1 = \emptyset\}$.
Chaque élément $x \in A_i$ appartient à au maximum $\frac{2^k}{k} - 1$ autres A_j par hypothèse 2. Du plus A_i a au maximum k éléments par hypothèse 1, donc finalement E_i est indépendant de tous les autres E_j sauf au plus $\frac{2^k}{8}$ d'entre eux.
On peut alors appliquer le lemme local de Lovász avec :

- $\mathbf{P}\{E_i\} \leq 2 \times \frac{1}{2^k}$ donc $p := \frac{1}{2^{k-1}}$,
- $d := \frac{2^k}{8}$,
- on vérifie qu'on a bien $4dp = 1$.

Et alors $\mathbf{P}\left\{\bigcap_{i \in I} \overline{E_i}\right\} > 0$.
Remarque : le $k > 6$ vient de la condition 2 où on veut $\frac{2^k}{8k} > 1$ (sinon pour tout x, x n'appartient à aucun A_i, donc tous les A_i sont vides, contradiction avec la condition 1).

Exercice 7. *Partition de graphe*

Soit $G = (V, E)$ un graphe non dirigé avec n sommets et m arrêtes.
Montrer qu'il existe une partition de V en deux ensembles disjoints A et B telle que au moins la moitié des arrêtes de G relie un sommet de A et un sommet de B.
☞ On construit une partition $G = A \uplus B$ au hasard, en choisissant pour chaque sommet e s'il appartient à A ou B uniformément et indépendamment. Pour toute arrête $e_i \in E$ de G, on note X_i l'événement « e_i relie A et B ». Alors $\mathbf{P}\{X_i = 1\} = \frac{1}{2}$, donc $\mathbf{E}[X_i] = \frac{1}{2}$.
Notons $M = \sum_{i=1}^{m} X_i$ la v.a. qui compte le nombre d'arrêtes reliant A et B. Son espérance est

$$\mathbf{E}[M] = \sum_{i=1}^{m} \mathbf{E}[X_i] = \frac{m}{2}\,.$$

Mais on a toujours $\mathbf{P}\{M \geq \mathbf{E}[M]\} > 0$ (et $\mathbf{P}\{M \leq \mathbf{E}[M]\} > 0$), donc en particulier $\mathbf{P}\left\{M \geq \frac{m}{2}\right\} > 0$, d'où le résultat.

1. László Lovász (né en 1948), mathématicien hongrois.