TD 05

Exercice 1. *Fonctions de répartition*

Définitions : Soit une variable aléatoire réelle X de densité de probabilité f_X,
— la **fonction de répartition** associée est :

$$F_X(x) \;=\; \int_{-\infty}^{x} f_X(t)\mathrm{d}t\;,$$

— l'**espérance** de X est définie par

$$\mathbf{E}\left[X\right] \;=\; \int_{-\infty}^{+\infty} t f_X(t)\mathrm{d}t$$

(si l'intégrale est absolument convergente),
— la **variance** de X est définie par

$$\mathbf{Var}\left[X\right] \;=\; \mathbf{E}\left[(X-\mathbf{E}\left[X\right])^2\right] \;=\; \mathbf{E}\left[X^2\right] - \mathbf{E}\left[X\right]^2$$

(si $\mathbf{E}\left[X^2\right]$ existe).

1. Donner la densité, la fonction de répartition, l'espérance et la variance d'une variable aléatoire U suivant la loi uniforme sur $[a,b]$ pour $a<b$.

Soit U une v.a. de loi uniforme sur $[0,2]$. Soit $X := \sqrt{U}$.

2. Calculer la fonction de répartition F_X de X.
3. Calculer la densité f_X de X.
4. Même questions pour $Y := 1/U$.
5. Quelle est l'espérance de Y ?

Exercice 2. *Records*

Soient $U_1,\dots,U_n$ des variables aléatoires indépendantes, uniformément distribuées sur $[0,1]$.
Pour $i \in \{1,\dots,n\}$, on dit que U_i est un **record** si pour tout $j \le i$, on a $U_i \le U_j$.
Calculer l'espérance du nombre de records dans la suite $U_1,\dots,U_n$.

Exercice 3. *Interrupteurs*

1. Montrer qu'il existe une constante $\gamma > 0$ rendant l'énoncé suivant vrai :

   « Si une v.a. positive X vérifie $\mathbf{E}\left[X\right] = 1$ et $\mathbf{E}\left[X^2\right] \le 3$, alors $\mathbf{P}\left\{X \ge 1/4\right\} \ge \gamma$. »

   Indication : définir la variable aléatoire $Y = \mathbf{1}_{X\ge 1/4}$ et se ramener à l'inégalité de Cauchy-Schwarz. $\mathbb{E}(XY) \le \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$

2. Soient $(X_1,\dots,X_n)$ des v.a. i.i.d. vérifiant $\mathbf{P}\left\{X_i = 1\right\} = \mathbf{P}\left\{X_i = -1\right\} = \frac{1}{2}$.
   On pose $Y = \frac{1}{\sqrt{n}}(X_1 + \cdots + X_n)$. Calculer $\mathbf{E}\left[Y^2\right]$ et $\mathbf{E}\left[Y^4\right]$ et en déduire que :

$$\mathbf{E}\left[|X_1 + \cdots + X_n|\right] \ge \frac{\gamma}{2}\sqrt{n}.$$

On considère une grille $n \times n$ d'ampoules ainsi que 3 séries d'interrupteurs : des interrupteurs $a = (a_{ij})_{1 \le i,j \le n}$ associés à chaque ampoule, des interrupteurs $b = (b_i)_{1 \le i \le n}$ associés à chaque ligne et des interrupteurs $c = (c_j)_{1 \le j \le n}$ associés à chaque colonne. Chaque interrupteur prend la valeur -1 ou 1. L'ampoule en position (i,j) est allumée si et seulement si $a_{ij} \times b_i \times c_j = 1$. On considère la quantité

$$\mathbf{F}(a,b,c) = \sum_{i,j=1}^{n} a_{ij} b_i c_j$$

qui est le nombre d'ampoules allumées moins le nombre d'ampoules éteintes.
Deux joueuses jouent au jeu suivant :

1. la joueuse 1 choisit la position des interrupteurs (a_{ij}),
2. puis la joueuse 2 choisit la position des interrupteurs (b_i) et (c_j).

La joueuse 1 veut minimiser $\mathbf{F}(a,b,c)$ et la joueuse 2 veut le maximiser. On considère donc :

$$\mathbf{V}(n) = \min_{a \in \{-1,1\}^{n \times n}} \max_{b,c \in \{-1,1\}^n} \mathbf{F}(a,b,c).$$

3. Montrer que $\mathbf{V}(n) = \mathcal{O}(n^{3/2})$ en considérant le cas où la joueuse 1 joue au hasard.
4. La joueuse 2 applique la stratégie suivante : elle choisit b au hasard, puis ensuite choisit c de façon à allumer le maximum de lampes. Estimer le nombre moyen de lampes allumées par cette stratégie *(indication : utiliser la question 2)* et en déduire que $\mathbf{V}(n) = \Omega(n^{3/2})$.

Exercice 4. *Arrondi*

Soit U un ensemble à n éléments. On appelle recouvrement de U un ensemble $\mathcal{S} = (S_1, \cdots, S_m)$ de parties de U qui vérifie $\bigcup S_i = U$. Étant donné $\mathcal{S}$ un recouvrement de U, on note $\mathrm{OPT}(\mathcal{S})$ le cardinal minimal d'un sous-ensemble de $\mathcal{S}$ qui est encore un recouvrement de U.

1. Expliquer rapidement pourquoi $\mathrm{OPT}(\mathcal{S})$ est la solution du problème d'optimisation suivant :

$$\text{Minimiser } \sum_{i=1}^{m} x_i \text{ sous les contraintes } x_i \in \{0,1\} \text{ et } \sum_{i=1}^{m} x_i \mathbf{1}_{S_i} \ge \mathbf{1}_U \tag{1}$$

On considère le problème suivant qui est une relaxation de (1) :

$$\text{Minimiser } \sum_{i=1}^{m} z_i \text{ sous les contraintes } z_i \in [0,1] \text{ et } \sum_{i=1}^{m} z_i \mathbf{1}_{S_i} \ge \mathbf{1}_U \tag{2}$$

Alors que le problème (1) est NP-difficile, le problème (2) peut être résolu en temps polynomial par les méthodes de programmation linéaire.

2. Soit k un entier, et $(z_i)_{1 \le i \le m}$ qui minimisent (2). Soient $(X_{i,j})_{1 \le i \le m, 1 \le j \le k}$ des variables aléatoires indépendantes vérifiant $\mathbf{P}\left\{X_{i,j} = 1\right\} = z_i$, $\mathbf{P}\left\{X_{i,j} = 0\right\} = 1 - z_i$. On définit un sous-ensemble $\mathcal{T} \subset \mathcal{S}$ par la condition :

$$S_i \in \mathcal{T} \iff \exists j \in \{1, \ldots, k\} \,:\, X_{i,j} = 1\,.$$

Montrer que :

$$\mathbf{E}\left[\#\mathcal{T}\right] \le k \,\mathrm{OPT}(\mathcal{S})\,.$$

3. Déterminer une valeur de $c > 0$ telle que, si on pose $k = \lfloor c \log n \rfloor$, on ait :

$$\mathbf{P}\left\{\mathcal{T} \text{ est un recouvrement de } U\right\} \ge 1 - \frac{1}{n}\,.$$