DM 1 (due March 9th, before 12h00)

Exercice 1. *Allumettes*

Un fumeur dispose de deux boîtes de $n \geq 1$ allumettes, une dans sa poche gauche et l'autre dans sa poche droite. Lorsqu'il a besoin d'une allumette, il choisit au hasard (avec équiprobabilité) une de ses poches et retire une allumette de la boîte correspondante. Cette expérience continue jusqu'à ce que le fumeur choisisse de prendre une allumette dans une boîte vide.

1. Pour $r \in \{0, \ldots, n\}$, on note u_r la probabilité que, à l'instant où le fumeur réalise qu'une boîte est vide, l'autre boîte contiennent r allumettes. Calculer u_r.

2. Calculer maintenant la probabilité v_r qu'une boîte contienne r allumettes au moment où le fumeur prend la dernière allumette de l'autre boîte.

3. En déduire la probabilité v que la première boîte vidée ne soit pas la boîte reconnue comme vide.

Exercice 2. *Depasser2*

Soit $t_1, t_2, \ldots, t_r$ le résultat d'une sélection (avec remise) de r éléments pris dans l'ensemble $(\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n})$. Soit X la variable aléatoire qui donne le plus petit r satisfaisant

$$t_1 + t_2 + \cdots + t_r > 1$$

On souhaite calculer $\mathbf{E}[X]$. Pour cela, on peut considérer une variante équivalente de l'expérience, qui consiste à choisir avec remise un ensemble d'éléments $t_1, t_2, \ldots, t_r$ dans l'ensemble $(1, 2, \ldots, n)$. X est alors la variable aléatoire du plus petit r qui satisfait

$$t_1 + t_2 + \cdots + t_r > n.$$

1. Montrer que $\mathbf{P}\{X \geq j+1\} = \binom{n}{j}(\frac{1}{n})^j$

2. Montrer que $\mathbf{E}[X] = \sum_{j=0}^{n} \mathbf{P}\{X \geq j+1\}$.

3. En déduire une expression de $\mathbf{E}[X]$ et sa limite quand $n \to \infty$.

Exercice 3. *Princesse*

Une princesse a $n \gg 1$ prétendants parmi lesquels elle doit choisir un époux. Ces prétendants sont numérotés de 1 à n dans l'ordre décroissant de préférence (celui qu'elle préfère porte le numéro 1). Les prétendants lui sont présentés dans un ordre aléatoire de loi uniforme sur $S\ n$ jusqu'à ce qu'elle choisisse celui qui lui semble le meilleur. Le problème est que la princesse ne peut comparer un prétendant qu'avec ceux qu'elle a déjà vus, et de plus elle ne peut choisir un prétendant qu'au moment où il lui est présenté : après c'est trop tard ! (Elle est obligée de choisir le dernier si le processus va jusque là). La princesse souhaite maximiser la probabilité d'épouser le prétendant numéro 1. Pour cela elle applique la stratégie suivante : elle refuse les m premiers prétendants, puis accepte dès qu'on lui présente un prétendant qu'elle préfère aux m premiers.

1. Estimez la valeur optimale de m, puis la probabilité que la princesse épouse le prétendant numéro 1.

Exercice 4. *Rrounding*

Soit U un ensemble à n éléments. On appelle recouvrement de U un ensemble $\mathcal{S} = (S_1, \ldots, S_m)$ 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

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

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

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

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 $(z_i)_{1\leq i\leq m}$ qui minimisent (2). Soient $(X_{i,j})_{1\leq i\leq m, 1\leq j\leq k}$ des variables aléatoires indépendantes vérifiant $\mathsf{P}(X_{i,j} = 1) = z_i$, $\mathsf{P}(X_{i,j} = 0) = 1 - z_i$. 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
$$\mathbf{E}\left[\mathcal{T}\right] \leq k\, \mathrm{OPT}(\mathcal{S}).$$

3. Déterminer une valeur de $c > 0$ telle que, si on pose $k = \lfloor c \log n \rfloor$, on ait
$$\mathsf{P}(\mathcal{T} \text{ est un recouvrement de } U) \geq 1 - \frac{1}{n}.$$

Exercice 5. *Discrepancy*

Let $[m] = \{1, \ldots, m\}$ and $\mathsf{S} = \{S_1, \ldots, S_n\}$ be a family of subsets of $[m]$. Our objective is to color the elements of $[m]$ with two colors, say red and blue, in such a way that all the sets S_i have nearly the same number of red and blue elements. More precisely, a coloring is a map $\chi : [m] \to \{-1, +1\}$. We then define $\chi(S) = \sum_{x\in S} \chi(x)$. Then the discrepancy of the family S with respect to the coloring χ is defined as $\mathrm{disc}(\mathsf{S}, \chi) = \max_{S\in\mathsf{S}} |\chi(S)|$. Our objective is to find a coloring that minimizes the discrepancy so we define $\mathrm{disc}(\mathsf{S}) = \min_\chi \mathrm{disc}(\mathsf{S}, \chi)$.

1. Assume S contains all the subsets of $[m]$. Compute $\mathrm{disc}(\mathsf{S})$.

For each $i \in [m]$ assign $x_i = 1$ with probability $1/2$. Else assign $x_i = -1$

2. What is $\mathbf{E}\left[\chi(S_i)\right]$ where $S_i \in \mathsf{S}$.

3. Define Bernoulli r.v Y_i such that $Y_i = 1$ iff $x_i = 1$ and $Y_i = 0$ iff $x_i = -1$.

4. Find a formula for $\chi(S_i)$ involving the r.v Ys.

5. Show that for any family S of n sets, we have $\mathrm{disc}(\mathsf{S}) \leq 10\sqrt{m \ln(2n)}$.

   You may use that $\sqrt{\dfrac{\ln(2n)}{m}} \leq 1$.