La méthode probabiliste — méthode du premier moment

Exercice 1 **Usines de montres**

Deux usines A et B produisent des montres. L'usine A produit une montre défectueuse sur 100 et B une sur 200 (on supposera les états des montres indépendants). Un détaillant reçoit une caisse de n montres de chacune des usines, sans pouvoir connaître l'origine des caisses. Il ouvre une caisse au hasard, prend une montre au hasard et vérifie l'état de cette montre, qui fonctionne.

Quelle est alors la probabilité qu'une seconde montre prise au hasard dans la même caisse fonctionne aussi ?

Exercice 2 **Points fixes d'une permutation**

Quelles sont l'espérance et la variance du nombre de points fixes d'une permutation sur $\{1, \ldots, n\}$ pour une distribution uniforme des permutations ?

Exercice 3 **Analyse du tri rapide**

On rappelle l'algorithme du tri rapide :

Données : Une liste S de n entiers distinct
Résultat : La liste triée des éléments de S
début
  si *S a 0 ou 1 élément* **alors** retourner S;
  sinon
    Choisir un élément x (pivot) de S et diviser les autres éléments en deux sous-listes
      — S_1, liste des éléments de S qui sont $< x$;
      — S_2, liste des éléments de S qui sont $> x$;
    Tri_Rapide($S1$) ; Tri_rapide(S_2);
    Retourner la liste S_1, x, S_2.
  fin
fin

Algorithme 1 : Tri_Rapide

1. Donner un exemple de liste qui nécessite $\Omega(n^2)$ comparaisons pour la trier avec cet algorithme.

On veut montrer que si les pivots sont choisis indépendamment et uniformément, alors l'espérance du nombre de comparaisons est de $2n \ln n + O(n)$. On note $y_1 < y_2 < \cdots < y_n$ les éléments de la liste.

2. Quelle est la probabilité que deux éléments y_i at y_j soient comparés ?
3. En déduire le résultat.
4. Que se passe-t-il si on choisit toujours le premier élément de la liste comme pivot ? Quelle est la différence avec le choix d'un pivot aléatoire ?

Exercice 4 **Recherche de la médiane**

On considère l'algorithme suivant :

1. Expliquer pourquoi cet algorithme termine en $O(n)$ et pourquoi cet algorithme renvoie soit ÉCHEC soit la bonne réponse.

On veut maintenant connaître la probabilité que la réponse soit ÉCHEC. Un ÉCHEC appartient à au moins un de ces trois événements :
- $\mathcal{E}_1 : Y_1 = |\{r \in R \mid r \leq m\}| < \frac{1}{2}n^{3/4} - \sqrt{n}$;
- $\mathcal{E}_2 : Y_2 = |\{r \in R \mid r \geq m\}| < \frac{1}{2}n^{3/4} - \sqrt{n}$;

Données : Un ensemble S de n entiers
Résultat : La médiane de S, noté m
début
  Choisir un multi-ensemble R de $\lceil n^{3/4} \rceil$ éléments de S choisis uniformément et indépendamment;
  Trier R;
  Soit d le $\lfloor \frac{1}{2} n^{3/4} - \sqrt{(n)} \rfloor$-ième plus petit élément de R;
  Soit u le $\lceil \frac{1}{2} n^{3/4} + \sqrt{(n)} \rceil$-ième plus petit élément de R;
  Comparer tous les éléments de S à d et u. Soient $C = \{x \in S \mid d \leq x \leq u\}$, $\ell_d = |\{x \in S \mid x < d\}|$ et $\ell_u = |\{x \in S \mid x > u\}|$;
  si $\ell_d > n/2$ *ou* $\ell_u > n/2$ **alors** ÉCHEC;
  ;
  sinon si $|C| > 4n^{3/4}$ **alors** ÉCHEC;
  ;
  sinon
    Trier C;
    Retourner le $\lfloor n/2 \rfloor - \ell_d + 1$-ième élément de C trié.
  fin
fin

Algorithme 2 : Calcul de la médiane randomisé

— $\mathcal{E}_3$: $m \in C$ et $|C| > 4n^{3/4}$.

2. Montrer que $P(\mathcal{E}_1) = P(\mathcal{E}_2) \leq \frac{1}{4} n^{-1/4}$.

3. Montrer que $P(\mathcal{E}_3) \leq \frac{1}{2} n^{-1/4}$.

4. Conclure.

`Exercice 5`

Deux événements disjoints de probabilité non nulle peuvent-ils être indépendants ?

`Exercice 6`

1. Donner un exemple d'espace de probabilité avec trois événements A_1, A_2, A_3 indépendants deux à deux mais pas globalement indépendants.

2. Soit $\{A_i\}_{i \in I}$ une famille indépendante d'événements. Est-ce que $\{\tilde{A}_i\}_{i \in I}$ est aussi une famille indépendante, quand pour tout $i \in I$, $\tilde{A}_i \in \{A_i, \bar{A}_i\}$?

`Exercice 7` `Pile ou face : inégalités`

On lance une pièce non biaisée n fois. En utilisant les inégalités de Markov et Chebyshev, donner deux bornes sur la probabilité d'obtenir au moins $3n/4$ fois face.

`Exercice 8` `Problème des secrétaires`

Vous êtes chef d'entreprise et vous avez besoin d'un nouvel assistant. Vous avez un ensemble C de $n \in \mathbb{N}$ candidats à examiner et votre objectif est d'embaucher le candidat qui maximise votre fonction de préférence $f \in \mathcal{F}(C, E)$, où E est un ensemble totalement ordonné et f est une injection. Les candidats arrivent dans un ordre $\preceq$ aléatoire défini sur un espace de probabilité $(\Omega, \mathbb{P})$, i.e. $c_1 \preceq (\omega) c_2$ ssi l'entretien d'embauche du candidat c_1 à lieu avant l'entretien d'embauche du candidat c_2 pour l'instance $\omega \in \Omega$. $\preceq$ suit la distribution uniforme parmi tous les ordres possibles sur C.

Ne sachant pas quel est le niveau de compétence dans le domaine (et donc ne connaissant pas votre fonction de préférence f), vous décidez d'adopter la stratégie suivante : vous vous fixez un entier $m \in \mathbb{N}$, $m < n$ et faites passer les entretiens aux m premiers candidats sans en embaucher aucun. Vous embauchez ensuite le premier candidat qui est meilleur que les m premiers.

1. Soit $i \in [1, n], i > m$, A un ensemble de $i-1$ candidats et $p(A, i)$ le nombre d'ordres totaux sur C tels que :
— l'ensemble des $i-1$ premiers candidats est A
— le meilleur candidat arrive en i-ème position
et $l(A, i)$ le nombre d'ordres totaux sur C tel que :
— l'ensemble des $i-1$ premiers candidats est A
— le meilleur candidat arrive en i-ème position
— le meilleur candidat dans l'ensemble A (i.e. $argmax_{c \in A} f(c)$) arrive dans les m premiers
Exprimez $p(A, i, C)$ et $l(A, i, C)$ en fonction de $p(A \backslash \{argmax_{c \in A} f(c)\}, i-1, C \backslash \{argmax_{c \in A} f(c)\})$

2. Soit $S_{i,n,m}$ l'événement "le candidat arrivant en i-ème position est sélectionné pour n et m fixé" et $M_{i,n,m}$ l'évènement "le candidat arrivant en i-ème position est le meilleur candidat pour n et m fixé". Calculez $\mathbb{P}(S_{i,n,m} | M_{i,n,m})$.

3. Soit $M_{n,m}$ l'évènement "le meilleur candidat est sélectionné pour n et m fixé". Calculez $\mathbb{P}(M_{n,m})$.

4. Montrez que à $(n, m) \in \mathbb{N}^* \times \mathbb{N}^*$ fixés, $n > m$, on a :

$$\frac{n}{m}(\ln(n) - \ln(m)) \leq \mathbb{P}(M_{n,m}) \leq \frac{n}{m}(\ln(n-1) - \ln(m-1))$$

5. Calculez à $x \in \mathbb{R}_+^*$ fixé le $y \in \mathbb{R}_+^*$ qui maximise l'expression $\frac{y}{x}(\ln(x) - \ln(y))$.

6. Montrez que :

$$\lim_{n \to +\infty} \mathbb{P}(E_{n, [\frac{n}{e}]}) = \frac{1}{e}$$

`Exercice 9` `Transformation de listes en arbres`

Soit L une permutation aléatoire de $[1, n]$. On veut transformer L en arbre $\mathcal{A}$ de la manière suivante :

- On part de $\mathcal{A}$ l'arbre vide et du premier élément de la liste
- Lorsque l'on atteint $L[k]$ le k-ème élément de la liste, on l'insère récursivement dans l'arbre $\mathcal{A}$ déjà construit de la manière suivante :
  - si l'arbre est vide, alors on crée un arbre de taille 1 ayant $L[k]$ pour racine
  - sinon, si $L[k]$ est inférieur à la racine de l'arbre, on l'insère récursivement dans le fils gauche, et si $L[k]$ est supérieur à la racine de l'arbre, on l'insère récursivement dans le fils droit.

On suppose que L est distribué selon $\mathcal{U}(\mathfrak{S}_n)$ et on note dans la suite $\mathcal{G}(\mathcal{A})$ et $\mathcal{D}(\mathcal{A})$ pour respectivement les sous-arbres gauche et droit de $\mathcal{A}$ et $\mathcal{N}(\mathcal{A})$ le nombre de listes engendrant l'arbre $\mathcal{A}$ par l'algorithme précédent lorsque l'ensemble de ses noeuds est inclus dans $[1, n]$.

1. Montrez que la liste L donne un arbre $\mathcal{A}$ de racine x ssi

- x est le premier élément de L
- $\exists((g_i)_{i \in [1,k]}, (d_i)_{i \in [1,l]}) \in [1, n]^k \times [1, n]^l$ deux suites croissantes d'indices de $[1, n]$ telle que la liste $[L[g_1], ..., L[g_k]]$ donne le sous-arbre gauche et la liste $[L[d_1], ..., L[d_l]]$ donne le sous-arbre droit.

2. Exprimer $\mathcal{N}(\mathcal{A})$ en fonction de $\mathcal{N}(\mathcal{G}(\mathcal{A}))$ et $\mathcal{N}(\mathcal{D}(\mathcal{A}))$

3. En déduire que la probabilité d'obtenir un arbre $\mathcal{A}$ donné est

$$\mathbb{P}(\mathcal{A}) = \frac{1}{n!} \prod_{x \in [1,n]} \binom{|\mathcal{A}_x| - 1}{|\mathcal{G}(\mathcal{A}_x)|} = \frac{1}{n!} \prod_{x \in [1,n]} \binom{|\mathcal{A}_x| - 1}{|\mathcal{D}(\mathcal{A}_x)|}$$

oú $\mathcal{A}_x$ est le sous-arbre de $\mathcal{A}$ enraciné en x et $|\mathcal{A}|$ est le nombre de sommets de $\mathcal{A}$.

Exercice 10 **Propriétés vraies pour presque tout graphe**

Soit Q une propriété. On dit que Q est vraie pour presque tout graphe dans un espace de probabilités Ω_n constitué des graphes à n sommets sans boucles si

$$P(G \in \Omega_n \text{ tel que } G \text{ satisfait la propriété } Q) \to 1 \text{ quand } n \to \infty.$$

On prend $\Omega_n = \mathcal{G}(n, p)$ et on suppose $0 < p < 1$ fixé.

1. Soit $1 \leq h \leq k$ deux entiers naturels. Montrer que presque tout graphe $G_{n,p} = (V, E)$ de $\mathcal{G}(n, p)$ satisfait la propriété suivante :

pour toute suite de k sommets $x_1, \ldots, x_k$, il existe un sommet x tel que $xx_i \in E$ pour tout $1 \leq i \leq h$ et $xx_i \notin E$ pour tout $h < i \leq k$.

2. En déduire que les propriétés suivantes sont vraies pour presque tout graphe :

— presque tout graphe $G_{n,p}$ a un diamètre de 2 ;
— pour un entier k fixé, un graphe $G_{n,p}$ est de degré minimal au moins k.