Devoir maison 1

A rendre pour le 12 décembre 2017.

`Exercice 1` `Cas d'égalité dans l'inégalité de Markov`

1. Pour une valeur de $a > 0$ fixée, donner un exemple de distribution où l'égalité de Markov est vérifiée.

2. Existe-t-il une distribution pour laquelle l'inégalité est une égalité pour deux valeurs a et b distinctes ?

`Exercice 2` `Schéma de Flajolet-Martin`

Le but de cet algorithme est de calculer une approximation du nombre d'éléments distincts dans un multi-ensemble en ne faisant qu'une passe sur les nombres. On suppose donc S un multi-ensemble de N éléments.

On note F le nombre d'éléments distincts dans S et on cherche à calculer $\tilde{F}$ une approximation de F.

Pour ce faire, on suppose disposer d'une fonction de hachage parfaite h : pour tout $k \in S$, $h(k) \sim Unif[0, 2^w - 1]$, où $w = \lceil \log N \rceil$. En particulier, pour $k_1 \neq k_2$, $h(k_1)$ et $h(k_2)$ sont indépendants.

On adopte le schéma suivant :
- Soit z_k le nombre de 0 en tête de la représentation binaire de $h(k)$.
- $Z = \max_{k \in S} z_k$
- $\tilde{F} = 2^Z$.

Par exemple, si $w = 5$ et $h(k) = 6 = (00110)_2$, $z_k = 2$.

On veut montrer

$$\forall c \geq 3,\ P(\frac{1}{c} \leq \frac{\tilde{F}}{F} \leq c) \geq 1 - \frac{3}{c}.$$

1. Montrer que $\forall r \in [0, w]$, $P(z_k \geq r) = 2^{-r}$.

Pour $r \in [0, w]$ et $k \in S$, on définit la v.a. $X_k(r) = 1_{z_k \geq r}$ et $X(r) = \sum_{k \text{ distincts de } S} X_k(r)$.

2. Calculer l'espérance et la variance de $X_k(r)$ et de $X(r)$.

Soient r_1 le plus petit entier r tel que $2^r > cF$ et r_2 le plus petit entier r tel que $2^r \geq F/c$.

3. Montrer que le schéma est correct si $X(r_1) = 0$ et $X(r_2) \neq 0$.

4. Montrer que $P(X(r_1) \geq 1) \leq \frac{1}{c}$.

5. Montrer que $P(X(r_2) = 0) \leq \frac{2}{c}$.

6. Conclure.

`Exercice 3` `Composantes connexes de taille 2`

On se place dans le cadre des graphes aléatoires Erdös-Rényi, dans l'espace $\mathcal{G}(n, p)$, où p pourra varier avec n. On note $C(n)$ la variable aléatoire du nombre de composantes connexes de taille 2 d'un graphe de $\mathcal{G}(n, p)$.

1. Quelle est l'espérance du nombre de composantes connexes de taille 2 dans $\mathcal{G}(n, p)$.

2. En déduire que si $p(n) \gg \frac{\ln n}{n}$ (c'est-à-dire, $\ln n/n = o(p(n))$), alors $\lim_{n\to\infty} \mathbf{P}(C(n) \neq 0) = 0$.

3. Calculer $\mathbf{E}[C(n)^2]$. En déduire que $\mathrm{Var}(C(n)) \leq \mathbf{E}[C(n)] + \frac{1-(1-p)^4}{(1-p)^4}\mathbf{E}[C(n)]^2$.

4. En déduire que si $1/n^2 \ll p(n) \leq 1/n$, alors $\lim_{n\to\infty} \mathbf{P}(C(n) = 0) = 0$.

`Exercice 4` `Intersection d'arêtes dans un hypergraphe`

Un hypergraphe est un ensemble de sous-ensembles. Les ensembles sont appelés les arêtes de l'hypergraphe et les éléments de chacune des arêtes sont appelés les sommets de l'hypergraphe.

Une arête dans laquelle tous les sommets sont de même couleur est appelé une arête monochromatique. Soit H un hypergraphe r-uniforme de nombre chromatique $k+1$ (i.e. le nombre de couleurs minimal tel qu'aucune arête ne soit monochromatique). Montrer qu'il existe au moins une arête de H qui en intersecte $\frac{k^{r-1}}{4}$ autres.

`Exercice 5` `Transverses latins`

Soit $(\Omega, \mathbb{P})$ un espace de probabilité et soient $A_1, ..., A_m$ des évènements de Ω. Un graphe de dépendance asymétrique $G = (S, A)$ pour les évènements $A_1, ..., A_m$ est un graphe tel que si $S \subset [1, m], i \notin S$ et $\forall j \in S, \{i, j\} \notin S$ alors $\mathbb{P}(A_i | \cap_{j\in S} \overline{A_j}) \leq \mathbb{P}(A_i)$.

1. Montrer que s'il existe un graphe de dépendance asymétrique G pour les évènements $A_1, ..., A_m$ de degré maximal d et un nombre $p \in [0, 1]$ tel que $\forall i \in [1, m], \mathbb{P}(A_i) \leq p$ et $4dp \leq 1$, alors $\mathbb{P}(\cap_{i\in[1,m]}\overline{A_i}) > 0$

Si M est une matrice carrée de dimension n, on appelle une permutation σ un transverse latin ssi les valeurs de $B_{i,\sigma(i)}$ sont toutes deux-à-deux distinctes.

On veut montrer que si $k \leq \frac{n-1}{16}$ est tel que qu'aucun coefficient de B n'apparaît plus de k fois, alors B possède un transverse latin.

2. On se place dans l'espace de probabilité $(\Omega, \mathbb{P})$, où Ω est l'ensemble des permutations de $[1, n]$ et $\mathbb{P}$ est la distribution uniforme sur Ω.

Soit $T = \{(i, j, i', j') | i \neq i', j \neq j', i < i', B_{i,j} = B_{i',j'}\}$, et $\forall (i, j, i', j') \in T, A_{i,j,i',j'}$ l'évènement $[\sigma(i) = j] \cap [\sigma(i') = j']$.

On définit $G = (S, A)$ le graphe ayant pour sommets $S = \{A_t | t \in T\}$ et tel qu'une arête existe entre les sommets $A_{i,j,i',j'}$ et $A_{x,y,x',y'}$ ssi les coefficients $(i, j), (i', j'), (x, y)$ et (x', y') n'occupent pas 4 lignes et colonnes distinctes.

Montrez que G est un graphe de dépendance asymétrique ssi

$$\mathbb{P}(A_{(1,1,2,2)} | \cap_{t\in S} \overline{A_t}) \leq \frac{1}{n(n-1)}.$$

3. Soit $B_{i,j} = \left(\cap_{t\in S}\overline{A_t}\right) \cap [\sigma(1) = i] \cap [\sigma(2) = j]$ et $s_{i,j} = |B_{i,j}|$.

Montrez que $\forall i, j \geq 3, s_{1,2} \leq s_{i,j}$.

4. Montrez que $\forall i \geq 3, s_{1,2} \leq s_{1,i}$.

5. Montrez que $\forall i \geq 2, s_{1,2} \leq s_{i,2}$ et que $\forall i \neq j, s_{i,j} = s_{j,i}$.

6. En déduire que G est un graphe de dépendance asymétrique. Conclure.

`Exercice 6` `Processus de branchement avec paramètre variable`

On considère un processus de branchement. On suppose que l'on a un paramètre λ tiré avant que le premier noeud du processus de branchement commence à se reproduire, distribué selon :

$$f(\lambda) = 1_{\mathbb{R}_+^*}(\lambda) \frac{\left(\frac{q}{p}\right)^{\alpha} \lambda^{\alpha-1}}{\Gamma(\alpha)} e^{-\frac{q}{p}\lambda}$$

et que la distribution du nombre de descendant Z d'un noeud lorsque le paramètre λ est fixé est

$$\mathbb{P}([Z = k]|\lambda) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Montrer que :

$$\mathbb{E}(s^Z) = \left(\frac{q}{1 - ps}\right)^{\alpha}$$

`Exercice 7` `Convergence vers la loi de Poisson`

On pose π une distribution sur $\{-1, 0, 1, \ldots\}$ telle que

$$\pi(k) = \begin{cases} e^{-p} - (1 - p) & \text{si } k = -1 \\ 1 - p & \text{si } k = 0 \\ e^{-p} \frac{p^k}{k!} & \text{sinon.} \end{cases}$$

1. Montrer comment construire des variables aléatoires $X \sim B(p)$ et $Y \sim \mathcal{P}(p)$ telles que $P(X \neq Y) \leq p^2$.

2. En déduire une construction de variables aléatoires $X \sim \text{Bin}(n, p)$ et $Y \sim \mathcal{P}(np)$ telles que $P(X \neq Y) \leq np^2$.

Soient pour tous $k \in \mathbb{N}$, $f(k) = P(\text{Bin}(n, p) = k)$ et $g(k) = P(\mathcal{P}(np) = k)$.

3. Montrer que $\sum_{k \in \mathbb{N}} |f(k) - g(k)| \leq 2np^2$.

4. En déduire que pour tous $X \sim \text{Bin}(n, p)$, $Y \sim \mathcal{P}(np)$ et $A \subseteq \mathbb{N}$,

$$|P(X \in A) - P(Y \in A)| \leq 2np^2.$$