Correction des exercices du cours: Structures et Algorithmes Aléatoires
cours 2 du 23 octobre 2009.

1. Montrer que $\mathbb{E}[X^k] \geq \mathbb{E}[X]^k$, pour tout $k \geq 1$ pair.

   - découle de l'inégalité de Jensen.

2. On considère une version déterministe de quicksort où le premier élément de la liste est choisi comme pivot. Quel est le nombre moyen de comparaisons si la liste en entrée est tirée uniformément parmi toutes les possibilitées?

   - La preuve vue en cours fonctionne encore.

3. Vous devez recruter un thésard et vous avez n candidats qui sont triés selon vos préférences. Les candidats se présentent à l'entretien dans un ordre aléatoire, mais à l'issue de l'entretien vous pouvez le classer parmi ceux que vous avez déjà vus. Si à la fin d'un entretien, vous ne proposez pas la thèse, vous perdez toute chance de sélectionner ce candidat comme thésard. Votre but est de proposer la thèse au meilleur candidat. On considère la stratégie suivante: faire passer un entretien à m candidats et tous les rejeter; à partir du $m+1$-ème candidat, choisir le candidat qui est meilleur que tous ceux vus jusqu'à présent. Montrer que la probabilité de choisir le meilleur candidat est:

   $$P(n,m) = \frac{m}{n} \sum_{j=m+1}^{n} \frac{1}{j-1}.$$

   En déduire que

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

   et que $\lim_n \max_m P(n,m) \geq 1/e$.

   - Soit E_j l'événement: le j-ème candidat est le meilleur et est recruté. On a $\mathbb{P}(E_j) = \frac{m}{n(j-1)}$ pour $j \geq m+1$ car il y a une probabilité $1/n$ que le j-ème candidat soit le meilleur et étant donné cet événement, il y a une probabilité $m/(j-1)$ qu'on le choisisse. En effet, soit $X_{ij} = \mathbf{1}$(le i-ème candidat est classé premier parmi les $j-1$ premiers) et $X_j = \sum_{i=1}^m X_{ij}$. Alors $\mathbb{P}(E_i) = \frac{1}{n}\mathbb{P}(X_j = 1) = \frac{1}{n}\mathbb{E}[X_j] = \frac{m}{n(j-1)}$.
   - On obtient ensuite facilement l'inégalité puis on prend $m = n/e$.

4. Soit $v_1, \dots, v_n \in \mathbb{R}^n$ avec $|v_i| \leq 1$. Soit $p_1, \dots, p_n \in [0,1]$ et $w = p_1 v_1 + \dots p_n v_n$. Montrer qu'il existe $\epsilon_1, \dots, \epsilon_n \in \{0,1\}$ tels que $v = \epsilon_1 v_1 + \cdots + \epsilon_n v_n$ et $|w - v| \leq \sqrt{n}/2$.

   - On prend des ϵ_i indépendants avec $\mathbb{P}(\epsilon_i = 1) = p_i = 1 - \mathbb{P}(\epsilon_i = 0)$. Ceci définit un une variable aléatoire $X = |w - v|^2$. On a

   $$X = \sum_{i=1}^{n} \sum_{j=1}^{n} v_i \dot{v}_j (p_i - \epsilon_i)(p_j - \epsilon_j),$$

et donc

$$\mathbb{E}[X] = \sum_{i=1}^{n} |v_i|^2 p_i(1-p_i) \leq \frac{1}{4}\sum_{i=1}^{n} |v_i|^2 \leq \frac{n}{4},$$

et on conclut comme en cours.

5. On rappelle que $R(k,\ell) > n$ signifie qu'il existe un 2-coloriage des arêtes de K_n en rouge et bleu tel qu'il n'existe pas de K_k rouge ni de K_ℓ bleu. Montrer que s'il existe $p \in [0,1]$ tel que

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{t}(1-p)^{\binom{t}{2}} < 1,$$

alors $R(k,t) > n$. En déduire que $R(4,t) \geq \Omega(t^{3/2}/(\ln t)^{3/2})$.

- On considère le coloriage aléatoire où chaque arête est colorié en bleu avec probabilité p et rouge avec probabilité $(1-p)$. Soit X le nombre de K_k rouge plus le nombre de K_t bleu. On a $\mathbb{E}[X] = \binom{n}{k}p^{\binom{k}{2}} + \binom{n}{t}(1-p)^{\binom{t}{2}}$. Donc $\mathbb{E}[X] < 1$ et il exsite un coloriage avec $X = 0$.
- On a

$$\begin{aligned} \binom{n}{4} p^{\binom{4}{2}} + \binom{n}{t}(1-p)^{\binom{t}{2}} &\leq n^4p^6 + \frac{n^t}{t!}e^{-pt(t-1)/2} \\ &\sim n^4p^6 + \frac{1}{\sqrt{2\pi t}}\exp(t(\ln n - \ln t + 1 - p(t-1)/2)). \end{aligned}$$

Pour $n \sim t^{3/2}/(\ln t)^{3/2}$, on choisit $p \sim \ln t/(t+\sqrt{t})$ de telle sorte que $n^4p^6 \to 0$ et

$$\ln n - \ln t + 1 - p(t-1)/2) \sim 1 + \frac{3}{2}\ln t - \ln t - \frac{t-1}{2(t+\sqrt{t})}\ln t \sim -\frac{3}{4}t\ln t$$

donc $\binom{n}{4}p^{\binom{4}{2}} + \binom{n}{t}(1-p)^{\binom{t}{2}} < 1$ pour t suffisament grand.

6. Montrer que $R(k,k) > n - \binom{n}{k}2^{1-\binom{k}{2}}$.

- On considère un coloriage aléatoire des arêtes de K_n. Pour tout ensemble R de k sommets, soit X_R la variable aléatoire indicatrice de l'événement: le graphe induit par R est monochromatique. Soit $X = \sum_R X_R$. On a $\mathbb{E}[X] = \binom{n}{k}2^{1-\binom{k}{2}}$. Donc il existe un coloriage tel que $X \leq m$. On prend un tel coloriage et on retire de K_n un sommet de chaque ensemble à k sommets monochromatique. On retire au plus m sommets et donc il reste s sommetes avec $s \geq n - m$. Le coloriage sur ces s sommets n'a pas d'ensemble de k sommets monochromatique.