Correction des exercices du cours: Structures et Algorithmes Aléatoires
cours 5 du 13 novembre 2009.

1. On considère le modèle où des boules sont jetées dans des boites. Montrer que pour une constante c_1, si il y a $c_1\sqrt{n}$ boules alors la probabilité qu'aucune boite ne contienne deux boules est au plus $1/e$. Montrer que pour une constante c_2 (et n suffisament grand), si il y a $c_2\sqrt{n}$ boules, alors la probabilité qu'aucune boite ne contienne deux boules est au moins $1/2$. Optimiser les constantes c_1 et c_2.

   - On se met dans le scenario Poisson avec paramètre $m/n = c_1/\sqrt{n}$. On a alors:

     $$\mathbb{P}(\text{ toutes les boites dans le cas Poisson ont moins de 1 boule}) = e^{-c_1\sqrt{n}}\left(1+\frac{c_1}{\sqrt{n}}\right)^n \to e^{-c_1^2/2}.$$

     On voit donc qu'il faut améliorer le Théorème vu en cours. Ceci est possible en utilisant une monotonicité en m. On considère une fonction f positive ou nulle, telle que $\mathbb{E}[f(X_1^{(m)},\dots,X_n^{(m)})]$ soit monotone décroissante en m

     $$\begin{aligned}
     \mathbb{E}[f(Y_1^{(m)},\dots,Y_n^{(m)})] &= \sum_{k=1}^{\infty}\mathbb{E}[f(Y_1^{(m)},\dots,Y_n^{(m)})|\sum_{i=1}^{n}Y_i^{(m)}=k]\mathbb{P}[\sum_{i=1}^{n}Y_i^{(m)}=k]\\
     &= \sum_{k=1}^{\infty}\mathbb{E}[f(X_1^{(k)},\dots,X_n^{(k)})]\mathbb{P}[\sum_{i=1}^{n}Y_i^{(m)}=k]\\
     &\geq \mathbb{E}[f(X_1^{(m)},\dots,X_n^{(m)})]\sum_{k=1}^{m}\mathbb{P}[\sum_{i=1}^{n}Y_i^{(m)}=k]\\
     &= \mathbb{E}[f(X_1^{(m)},\dots,X_n^{(m)})]\mathbb{P}[\sum_{i=1}^{n}Y_i^{(m)}\leq m]\\
     &\geq \frac{1}{2}\mathbb{E}[f(X_1^{(m)},\dots,X_n^{(m)})],
     \end{aligned}$$

     où on utilise dans la dernière ligne le fait que pour une variable aléatoire de Poisson Z de moyenne μ, on a $\mathbb{P}(Z\leq\mu)\geq 1/2$.
     On en déduit donc que:

     $$\mathbb{P}(\text{ toutes les boites dans le cas réel ont moins de 1 boule}) \leq 2e^{-c_1^2/2}.$$

   - La probabilité pour que la première boite recoive plus de 2 boules est majorée par $\binom{m}{2}1/n^2 \leq c_2/(2n)$. On a donc

     $$\mathbb{P}(\text{ toutes les boites ont moins de 1 boule}) \leq 1 - n\frac{c_2}{2n} = 1-\frac{c_2}{2}.$$

2. On modélise le nombre d'erreurs dans une page des notes de cours par X une variable de Poisson de moyenne μ. Chaque erreur est indépendamment avec probabilité p une erreur de grammaire et avec probabilité $1-p$ une erreur d'orthographe. Quelles sont les lois de Y et Z les variables aléatoires qui comptent respectivement le nombre de fautes de grammaire et d'orthographe par page?

- Traitons le cas de Y:

$$\begin{aligned}
\mathbb{P}(Y=k) &= \sum_{n\geq k} \mathbb{P}(X=n)\binom{n}{k} p^k(1-p)^{n-k} \\
&= e^{-\mu}\sum_{n\geq k}\frac{\mu^n}{n!}\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\
&= \frac{e^{-\mu}(\mu p)^k}{k!}\sum_{n\geq k}\frac{((1-p)\mu)^{n-k}}{(n-k)!} \\
&= \frac{e^{-p\mu}(\mu p)^k}{k!}.
\end{aligned}$$

  Donc Y est une variable aléatoire de Poisson de moyenne μp.

3. Rappel $R(k,k) > n$ ssi il existe un coloriage des arêtes de K_n en rouge et bleu tel qu'il n'existe pas de K_k rouge ni de K_k bleu. Montrer que si

$$e\binom{k}{2}\binom{n-2}{k-2}2^{1-\binom{k}{2}} < 1$$

alors $R(k,k) > n$.

   - On considère un coloriage aléatoire. Pour chaque ensemble S de k sommets de K_n, soit A_S l'événement que le graphe complet sur S est monochromatique. Alors $\mathbb{P}(A_S) = 2^{1-\binom{k}{2}}$. Chaque A_S est mutuellement indépendant des événements A_T sauf si $|S \cap T| \geq 2$. On peut donc appliquer la version du lemme de Lovász symétrique avec $p = 2^{1-\binom{k}{2}}$ et $d < \binom{k}{2}\binom{n-2}{k-2}$, ce qui permet de conclure.