Correction des exercices du cours: Structures et Algorithmes Aléatoires

cours 3-4 du 30 octobre et 6 novembre 2009.

1. Identité de Wald: Soit $T, X_1, X_2, \dots$ une suite de variables aléatoires intégrables indépendantes à valeur dans $\mathbb{N}_+$. Les X_i, $i \geq 1$ sont identiquement distribués. On définit $S = X_1 + \cdots + X_T$. Montrer que $\mathbb{E}[S] = \mathbb{E}[X_1]\mathbb{E}[T]$.

   - On a:

$$\begin{aligned}\mathbb{E}[S] &= \sum_{n=0}^{\infty} \mathbb{E}[\mathbf{1}(T=n)X_1 + \cdots + X_n] \\ &= \sum_{n=0}^{\infty} n\mathbb{E}[X_1]\mathbb{P}(T=n) = \mathbb{E}[X_1]\mathbb{E}[T].\end{aligned}$$

2. Calculez, pour le processus de branchement, la moyenne m_n et la variance σ_n^2 du nombre d'individus X_n de la n-ème génération, lorsque $X_0 = 1$ et lorsque $X_0 = k > 1$.

   - On considère le processus de branchement vu en cours $X_{n+1} = \sum_{i=1}^{X_n} Z_i^{(n)}$ où les $Z_i^{(n)}$ sont i.i.d. de même loi que Z. Par l'exercice précédent, on a: $\mathbb{E}[X_{n+1}] = \mathbb{E}[X_n]m$ avec $m = \mathbb{E}[Z]$. On a donc pour X_0 déterministe: $\mathbb{E}[X_n] = m^n X_0$. Pour la variance, on a vu en cours que la fonction génératrice de X_{n+1} est $\varphi_{n+1}(s) = \varphi_n(g_Z(s))$ où g_Z est la fonction génératrice de Z. Soit $u_n = \varphi_n''(1)$ et $v = g_Z''(1)$, on a alors pour X_0 déterministe:

$$u_{n+1} = m^2 u_n + m^n v X_0.$$

   Si $m = 1$ alors $u_n = nvX_0$ et sinon $u_n = X_0 v m^{n-1}\frac{m^n-1}{m-1}$. On a donc

$$Var(X_n) = u_n + \mathbb{E}[X_n] - \mathbb{E}[X_n]^2 = m^n \frac{m^{n+1}-1}{m-1} Var(X_0).$$

3. Soit A une matrice $n \times m$ avec des coefficients dans $\{0,1\}$. Soit

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

   On cherche un vecteur $b \in \{-1,1\}^m$ qui minimize $\|Ab\|_\infty = \max_{i=1,\dots n}|c_i|$. Pour ceci on utilise un algorithme aléatoire très simple: chaque coefficient de b est choisi aléatoirement et indépendamment avec $\mathbb{P}(b_i = 1) = \mathbb{P}(b_i = -1) = 1/2$. Montrer que pour ce choix du vecteur b, on a

$$\mathbb{P}(\|Ab\|_\infty \geq \sqrt{4m\ln n}) \leq \frac{2}{n}.$$

- On considère la i-ème ligne $a^i = (a_{i1}, \ldots, a_{im})$. Soit k le nombre de 1 dans cette ligne. Si $k \leq \sqrt{4m \ln n}$ alors clairement $|a^i b| = |c_i| \leq \sqrt{4m \ln n}$. SI $k > \sqrt{4m \ln n}$ alors les k termes non nuls de la somme $Z_i = \sum_{j=1}^m a_{ij} b_j$ sont des variables aléatoires indépendantes valant $+1$ et -1 avec probabilité $1/2$. En utilisant la borne de Chernoff, on obtient:

$$\mathbb{P}(|Z_i| > \sqrt{4m \ln n}) \leq 2e^{-4m \ln n / 2k} \leq \frac{2}{n^2}.$$

Il suffit ensuite d'appliquer la borne de l'union pour obtenir le résultat.

4. Soit H, un graphe avec v sommets et e arêtes. On appelle $\rho(H) = e/v$ la densité de H. Le graphe H est dit équilibré si pour tout sous-graphe H', on a $\rho(H') \leq \rho(H)$. On considère un graphe équilibré H avec v sommets et e arêtes. Soit $A(G)$ l'événement: H est un sous-graphe (non nécessairement induit) de G. On veut montrer que $p(n) = n^{-v/e}$ est la fonction seuil pour A. Pour tout ensemble S de v éléments, soit A_S l'événement $G|_S$ contient H comme sous-graphe et $X_S = \mathbf{1}(A_S)$. Calculer la moyenne de $X = \sum_{|S|=v} X_S$ et en déduire que si $p << n^{-v/e}$ alors $X = 0$ presque toujours. On considère maintenant le cas $p >> n^{-v/e}$. Montrer que pour des ensembles S, T avec $|S \cap T| = i$, on a: $\mathbb{P}(A_T | A_S) = O(p^{e-(ie/v)})$ et conclure.

- On a par linéarité:

$$\mathbb{E}[X] = \sum_{|S|=v} \mathbb{E}[X_S] = \binom{n}{v} \mathbb{P}(A_S) = \Theta(n^v p^e).$$

Donc si $p << n^{-v/e}$ alors $\mathbb{E}[X] = o(1)$ donc $X = 0$ presque toujours.

- Il y a $O(1)$ copies possibles de H dans T. Chacune a au plus (par la propriété que H est équilibré) ie/v arêtes dont les deux extrémités sont dans S et donc au moins $e - (ie/v)$ autres arêtes. Donc $\mathbb{P}(A_T | A_S) = O(p^{e-(ie/v)})$.
- On note $S \sim T$ ssi $S \neq T$ et S, T ont des arêtes en commun. On a alors

$$\Delta = \sum_{S \sim T} \mathbb{P}(A_S \cap A_T) = \sum_S \mathbb{P}(A_S) \sum_{T \sim S} \mathbb{P}(A_T | A_S)$$

Pour S fixé, on a:

$$\begin{aligned}
\Delta^* &= \sum_{T \sim S} \mathbb{P}(A_T | A_S) \\
&= \sum_{i=2}^{v-1} \sum_{|T \cup S| = i} \mathbb{P}(A_T | A_S) \\
&= \sum_{i=2}^{v-1} O(n^{v-i} p^{e-(ie/v)}) \\
&= \sum_{i=2}^{v-1} o(n^v p^e) = o(\mathbb{E}[X]),
\end{aligned}$$

il est alors facile de conclure comme en cours.