ENS LYON
L3 Probabilités

Informatique

Examen partiel du 21 mars 2018
Correction

Exercice 1 Question de cours

Cf cours

Exercice 2 Graphe aléatoire bipartite

1. Le nombre d'arêtes de $H_{2n,p}$ suit la loi $B(n^2, p)$.
2. Soit N le nombre de sommets isolés. Si A_i est l'événement « le sommet i est isolé », on a par linéarité de l'espérance, on a $\mathbf{E}[N] = \sum \mathbf{P}(A_i) = 2n(1-p)^n$.
3. (a) Si $c > 1$, on a $\mathbf{E}[N] = 2n \exp(n \log(1 - \frac{c \log n}{n})) \to 0$ et donc $\mathbf{P}(N \geqslant 1) \leqslant \mathbf{E}N \to 0$.
   (b) Si $c < 1$, on calcule
$$\mathbf{E}[N^2] = \sum_{i,j=1}^{2n} \mathbf{P}(A_i \cap A_j) = 2n(1-p)^n + 2n(n-1)(1-p)^{2n} + 2n^2(1-p)^{2n-1}$$
   d'où il vient que $\mathbf{E}[N^2]/\mathbf{E}[N]^2$ tend vers 1. On utilise l'inégalité de Tchebychev pour conclure que
$$\mathbf{P}(N = 0) = \mathbf{P}(N - \mathbf{E}[N] \leqslant \mathbf{E}[N]) \leqslant \frac{\mathbf{Var}(N)}{\mathbf{E}[N]^2} = \frac{\mathbf{E}[N^2]}{\mathbf{E}[N]^2} - 1 \to 0.$$
4. Le degré d_i du sommet i suit la loi $B(n, 1/2)$. Par l'inégalité de Chernoff I, on a donc
$$\mathbf{P}(d_i \geqslant \frac{n}{2} + a) \leqslant \exp(-2a^2/n).$$
   Ainsi, par la borne de l'union,
$$\mathbf{P}(\max_i d_i \geqslant \frac{n}{2} + a) \leqslant 2n \exp(-2a^2/n).$$
   Cette quantité tend vers 0 si $a = C\sqrt{n \log n}$ avec $2C^2 > 1$.

Exercice 3 Interrupteurs

Partie I :

1. On écrit
$$1 = \mathbf{E}[X] = \mathbf{E}[X\mathbf{1}_{X<1/4}] + \mathbf{E}[X\mathbf{1}_{X\geqslant 1/4}] \leqslant \frac{1}{4} + \mathbf{E}[X\mathbf{1}_{X\geqslant 1/4}].$$
   Par l'inégalité de Cauchy–Schwarz, $\mathbf{E}[X\mathbf{1}_{X\geqslant 1/4}] \leqslant \sqrt{\mathbf{E}[X^2]\mathbf{P}(X \geqslant 1/4)} \leqslant \sqrt{3}\sqrt{\mathbf{P}(X \geqslant 1/4)}$. On obtient la minoration voulue pour $\gamma = 3/16$.
2. On a $\mathbf{E}[Y^2] = \mathbf{Var}(Y) = \sum_i \mathbf{Var}(X_i) = n$ (par inédependance). On a ensuite
$$\mathbf{E}[Y^4] = \sum_{i,j,k,l=1}^{n} \mathbf{E}[X_i X_j X_k X_l].$$
   L'indépendance des X_i et le fait que $\mathbf{E}[X_i] = 0$ implique $\mathbf{E}[X_i X_j X_k X_l] = 0$ dès qu'un indice apparaît une unique fois parmi $\{i, j, k, l\}$. Les seuls termes non nuls sont veux où $i = j = k = l$ ou $i = j \neq k = l$ ou $i = k \neq j = l$ ou $i = l \neq j = k$. On a donc
$$\mathbf{E}[Y^4] = n + 3n(n-1) = 3n^2 - 2n.$$
   On applique la question précédente à $X = Y^2/n$, d'où $\mathbf{P}(Y^2/n \geqslant 1/4) = \mathbf{P}(|X_1 + \cdots + X_n| \geqslant \frac{\sqrt{n}}{2}) \geqslant \gamma$. Enfin,
$$\mathbf{E}[|X_1 + \cdots + X_n|] \geqslant \frac{\sqrt{n}}{2}\mathbf{P}\left(|X_1 + \cdots + X_n| \geqslant \frac{\sqrt{n}}{2}\right) \geqslant \frac{\gamma\sqrt{n}}{2}.$$

Partie II :

1. Soit $(a_{ij})_{1\leqslant i,j\leqslant n}$ des v.a. i.i.d. de loi uniforme sur $\{-1,1\}$. Quel que soit le choix de b et c, on a

$$\mathbf{P}(F(a,b,c)\geqslant t)\leqslant \exp(-t^2/2n^2)$$

par l'inégalité de Chernoff (en effet, $F(a,b,c)$ est la somme de n^2 v.a. de loi uniforme sur $\{-1,1\}$). Par la borne de l'union,

$$\mathbf{P}(\max_{b,c} F(a,b,c)\geqslant t)\leqslant 4^n\exp(-t^2/2n^2).$$

Lorsque $t>\sqrt{2n^3\log 4}$, cette probabilité est <1 et donc $\mathbf{P}(\max_{b,c}F(a,b,c)<t)>0$: il existe donc un choix de a tel que $\max_{b,c}F(a,b,c)<t$, d'où $V(n)=O(n^{3/2})$.

2. Fixons $a=(a_{i,j})$ et choisissons (b_i) i.i.d. de loi uniforme sur $\{-1,1\}$. On a alors

$$\max_c F(a,b,c)=\sum_{i=1}^n\left|\sum_{j=1}^n a_{ij}b_j\right|.$$

En utilisant la linéarité de l'espérance, le fait que $(b_j)_j$ et $(a_{ij}b_j)_j$ ont même loi et la question I.2, il vient

$$\mathbf{E}\max_c F(a,b,c)=n\mathbf{E}\left|\sum_{j=1}^n b_j\right|\geqslant\frac{n^{3/2}\gamma}{2}.$$

En particulier, pour tout choix de a, il existe b tel que $\max_c F(a,b,c)\geqslant\frac{n^{3/2}\gamma}{2}$.

Exercice 4 Algorithme probabiliste pour trouver la médiane

1. Si aucun message d'erreur n'est renvoyé, l'algorithme s'exécute en temps $O(n)$; en effet la génération des (Y_i) prend un temps $O(n)$, le tri de F et G prend un temps $O(m\log m)$ pour $m=O(n^{3/4})$, et la détermination de r_d, de r_u et de G nécessite $O(n)$ comparaisons. De plus, l'absence de message d'erreur numéro 2 garantit que la médiane est dans l'intervalle $[d,u]$, donc dans G.

2. (a) Pour l'erreur 1 : comme $\operatorname{card}F=Y_1+\cdots+Y_n$ a la loi $B(n,n^{-1/4})$, on a par l'inégalité de Chernoff II

$$\mathbf{P}(\operatorname{card}F\geqslant 2n^{3/4})\leqslant\exp(-n^{3/4}/3),\quad \mathbf{P}(\operatorname{card}F\leqslant\frac{2}{3}n^{3/4})\leqslant\exp(-n^{3/4}/18).$$

(b) Pour l'erreur 2 : on note E^- l'ensemble des éléments de E inférieurs où égaux à la médiane, et on remarque que $r_d>n/2$ équivaut à $\operatorname{card}(F\cap E^-)<\frac12 n^{3/4}-\sqrt{n}$. La v.a. $\operatorname{card}(F\cap E^-)$ suit la loi $B(\lceil n/2\rceil,n^{-1/4})$ (notons μ sa moyenne) donc par l'inégalité de Chernoff II

$$\mathbf{P}(\operatorname{card}(F\cap E^-)<\frac12 n^{3/4}-\sqrt{n})\leqslant\mathbf{P}(\operatorname{card}(F\cap E^-)\leqslant(1-2n^{-1/4})\mu)\leqslant\exp(-\mu\sqrt{n})\to 0$$

Un argument symétrique traite le cas de $r_u>n/2$ et considérant E^+ l'ensemble des éléments de E supérieurs où égaux à la médiane

(c) Pour l'erreur 3 : si $\operatorname{card}G\geqslant 4n^{3/4}$, alors ou bien $\operatorname{card}(G\cap E^-)\geqslant 2n^{3/4}$ ou bien $\operatorname{card}(G\cap E^+)\geqslant 2n^{3/4}$; ces deux événements ayant même probabilité, il suffit de montrer que $\mathbf{P}(\operatorname{card}(G\cap E^-)\geqslant 2n^{3/4})\to 0$. On remarque que si $\operatorname{card}(G\cap E^-)\geqslant 2n^{3/4}$, alors $r_d\leqslant\frac{n}{2}-2n^{3/4}$ et donc l'ensemble F contient au moins $\frac12 n^{3/4}-\sqrt{n}$ parmi les $\frac{n}{2}-2n^{3/4}$ plus petits éléments de E. La probabilité de ce dernier événement est $\mathbf{P}(X\geqslant(1+\varepsilon)\mathbf{E}X)$, où $X\sim B(\frac{n}{2}-2n^{3/4},n^{-1/4})$ et $\varepsilon=\frac{\sqrt{n}}{n^{3/4}/2-2\sqrt{n}}=O(n^{-1/4})$. Une dernière application de l'inégalité de Chernoff II permet de conclure que la probabilité considérée tend vers 0.