

DEVOIR MAISON 2

On rappelle les formules usuelles suivantes :

- pour tout $p \in [0, 1]$, $(1 - p) \leq e^{-p}$, et pour tout $p \in [0, 1/2]$, $(1 - p) \geq e^{-p-p^2}$.
- $\binom{n}{k} \leq \frac{n^k}{k!}$ et $k! \geq k^k e^{-k}$

Exercice 1**Points isolés et graphe connexe**

On se place dans le cadre des graphes aléatoires Erdős-Rényi et les notations utilisées sont celles du cours.

Pour chaque sommet, on définit la variable aléatoire :

$$I(x) = \begin{cases} 1 & \text{si } x \text{ est un sommet isolé (sans voisin)} \\ 0 & \text{sinon.} \end{cases}$$

On pose $I = \sum_{x \text{ sommet}} I(x)$ et C la variable aléatoire égale à 1 si le graphe est connexe, 0 sinon.

On dit que la propriété A (qui dépend de n) est satisfaite *avec forte probabilité* si $\lim_{n \rightarrow \infty} \mathbf{P}(A) = 1$. Le but de cet exercice est de prouver le théorème suivant :

Théorème 1. *Si $pn - \ln n \rightarrow \infty$, alors le graphe $G_{n,p}$ est connexe avec forte probabilité et si $pn - \ln n \rightarrow -\infty$, alors le graphe $G_{n,p}$ n'est pas connexe avec forte probabilité.*

On va d'abord s'intéresser aux sommets isolés.

1. Montrer que si $pn - \ln n \rightarrow +\infty$ alors $\lim_{n \rightarrow \infty} \mathbf{P}(I \neq 0) = 0$.

Correction : For each vertex x , $\mathbf{P}(I(x)) = (1 - p)^{n-1}$. So $\mathbf{E}[I] = n(1 - p)^{n-1} \leq ne^{-p(n-1)}$.

Then by the Markov inequality, $\mathbf{P}(I \neq 0) \leq e^p e^{-pn + \ln n}$. So if $pn - \ln n \rightarrow +\infty$ when $n \rightarrow \infty$, then $\mathbf{P}(I \neq 0) \rightarrow 0$.

On suppose maintenant que $pn - \ln n \rightarrow -\infty$.

2. Calculer $\text{Var}(I)$ et montrer que $\text{Var}(I) \leq \mathbf{E}[I] + \mathbf{E}[I]^2 \frac{p}{1-p}$.

Correction :

$$\begin{aligned} \mathbf{E}[I^2] &= \mathbf{E}[(\sum_x I(x))^2] = \sum_x \mathbf{E}[I(x)] + \sum_{x \neq y} \mathbf{E}[I(x)I(y)] \\ &= \mathbf{E}[I] + n(n-1)\mathbf{P}(I(x)=1)\mathbf{P}(I(y)=1 \mid I(x)=1) \\ &= n(1-p)^{n-1} + n(n-1)(1-p)^{2n-3}. \end{aligned}$$

Then,

$$\text{Var}(I) = \mathbf{E}[I] + n(n-1)(1-p)^{2n-3} - n^2(1-p)^{2n-2} \leq \mathbf{E}[I] + \mathbf{E}[I]^2 \frac{p}{1-p}.$$

3. Montrer que $\lim_{n \rightarrow \infty} \mathbf{P}(I = 0) = 0$.

Correction : With the inequality of question 2, $\mathbf{P}(I = 0) \leq \mathbf{E}[I]^{-1} + \frac{p}{1-p}$. But $\mathbf{E}[I] = ne^{(n-1)\ln(1-p)} \geq ne^{(n-1)(-p-p^2)} \geq e^{np^2} e^{\ln n - pn}$. This inequality holds when $p < 1/2$, which is asymptotically true. The second term tends to ∞ , as $pn - \ln n \rightarrow -\infty$, and the first term tends to 1 : for n large enough, $pn < \ln n$, so $pn < \ln n/n$ and $p^2 < (\ln n)^2/n^2$, so $p^2 n < (\ln n)^2/n \rightarrow 0$. Then $\mathbf{E}[I] \rightarrow +\infty$. We also have $\frac{p}{1-p} \rightarrow 0$, so $\lim_{n \rightarrow \infty} \mathbf{P}(I = 0) = 0$.

4. Montrer que dans ce cas, $\lim_{n \rightarrow \infty} \mathbf{P}(C = 1) = 0$.

Correction : If $\lim_{n \rightarrow \infty} \mathbf{P}(I = 0) = 0$, obviously $\lim_{n \rightarrow \infty} \mathbf{P}(C = 1) = 0$, as if there is an isolated vertex, the graph is disconnected.

On s'intéresse maintenant au problème de la connexité au-dessus du seuil ($pn - \ln n \rightarrow \infty$). On s'intéresse plus précisément à la probabilité qu'il n'y ait pas de sommet isolé, alors que le graphe n'est pas connexe, $\mathbf{P}(C = 0, I = 0)$. Soit X_k le nombre total d'arbres couvrants de taille k dans les composantes de taille k .

5. Montrer que $\mathbf{P}(C = 0, I = 0) \leq \sum_{k=2}^{n/2} \mathbf{E}[X_k]$.

Correction : First notice that if there is no isolated vertex, but the graph is not connected, there exists a connected component of size between 2 and $n/2$. Then

$$\mathbf{P}(C = 0, I = 0) \leq \sum_{k=2}^{n/2} \mathbf{P}(X_k \geq 1) \leq \sum_{k=2}^{n/2} \mathbf{E}[X_k].$$

6. Montrer que $\mathbf{E}[X_k] \leq \binom{n}{k} p^{k-1} k^{k-2} (1-p)^{k(n-k)} \leq n(epn)^k \frac{1}{k^2} e^{-pk(n-k)}$. *Indication* : On pourra admettre que le nombre d'arbres couvrants d'un graphe connexe à k sommets est inférieur ou égal à k^{k-2} .

Correction : for each set of k vertices, there can be a spanning tree if they are in the same connected component, which can happen if there are at least $k-1$ edges among them. Moreover, among those vertices, there are at most k^{k-2} spanning trees and those vertices are not connected to any other vertex (otherwise it would not be a spanning tree) :

$$\mathbf{E}[X_k] \leq \binom{n}{k} p^{k-1} k^{k-2} (1-p)^{k(n-k)}.$$

Now, as $(1-u) \leq e^{-u}$, $\binom{n}{k} \leq n^k/k!$ and $k! \geq k^k e^{-k}$,

$$\mathbf{E}[X_k] \leq n^k p^{k-1} k^{k-2} e^{-pk(n-k)} k^{-k} e^k = n(epn)^k \frac{1}{k^2} e^{-pk(n-k)}$$

On s'intéresse au cas où $p = a \ln n/n$ avec $a > 1/2$, ce qui est suffisant pour le cas où $pn - \ln n \rightarrow +\infty$.

7. Montrer que pour k fixé, $\mathbf{E}[X_k] = o(1)$.

Correction : For $k \geq 2$, $\mathbf{E}[X_k] \leq \frac{e^{pk^2}}{k^2} n(epn e^{-pn})^k = \frac{e^{pk^2}}{k^2} n(ea(\ln n)n^{-a})^k \rightarrow 0$ when $n \rightarrow \infty$.

8. En utilisant que $k(n-k) \geq kn/2$, montrer que $\mathbf{E}[X_k] \leq n^{1-k/4}$ pour n suffisamment grand.

Correction : $\mathbf{E}[X_k] \leq n(epn)^k \frac{1}{k^2} e^{-pk(n-k)} \leq n(epn)^k e^{-pnk/2} = n(ea \ln n n^{-a/2})^k$. As $a > 1/2$, there exists N such that for all $n \geq N$, $ea \ln n n^{-a/2} \leq n^{-1/4}$. Then, $\mathbf{E}[X_k] \leq n^{1-k/4}$.

9. Conclure en montrant que $\mathbf{P}(C = 0, I = 0) \xrightarrow{n \rightarrow \infty} 0$ quand $pn - \ln n \rightarrow +\infty$.

Correction : If $p = a \ln n/n$, we now compute $\sum_{k=2}^{n/2} \mathbf{E}[X_k] = o(1) + \sum_{k=5}^{n/2} \mathbf{E}[X_k] \leq o(1) + \sum_{k=5}^{n/2} n^{1-k/4} = o(1) + \sum_{k=1}^{n/2} n^{-k/4} = o(1) + \frac{n^{-1/4} - n^{-n/8}}{1 - n^{-1/4}} = o(1)$. Now, if $p > a \ln n/n$, $\mathbf{E}[X_k]$ decreases, so $\mathbf{P}(C = 0, I = 0) \xrightarrow{n \rightarrow \infty} 0$.

Exercice 2 Codage de Prüfer (exercice bonus ne comptant pas dans le barème)

Soit K_n le graphe complet de taille n , dont les sommets sont numérotés de 1 à n , et soit T un arbre couvrant de K_n (les arbres ne sont pas enracinés, donc une feuille correspond exactement aux sommets de degrés 1). On associe à T le résultat de l'algorithme **Prüfer**(T). Le symbole ε représente le mot (ou la liste vide), et $\text{père}(T, v)$ est le père de v dans l'arbre T .

1. Montrer qu'il y a une bijection entre les arbres couvrants et $\{1, \dots, n\}^{n-2}$.
2. En déduire le nombre d'arbres couvrants de K_n , et donc une borne supérieure du nombre d'arbres couvrants dans un graphe à n sommets.

Algorithme 1 : $\text{Prüfer}(T)$

début $c \leftarrow \varepsilon;$ **tant que** T *a plus de deux sommets* **faire** Soit v la feuille de T avec le plus petit numéro; $c \leftarrow c.\text{père}(T, v);$ Retirer v de T ainsi que son arête incidente à v ;Renvoyer c .