Auteur : Johanne Cohen

Exercices sur les probabilités

https://www.lri.fr/~jcohen/documents/enseignement/td-proba.pdf

Pour s'entrainer.

Exercice 1 Considérons $H(n) = \sum_{i=1}^{n} \frac{1}{i}$. Montrer que $\forall n \geq 1, ln(n) + 1 \geq H(n) \geq ln(n) + 1$

Correction

Soit $n \geq 2$. La fonction$fx \to \frac{1}{x}$ est continue et décroissante sur $]0, +infinity[$. On en déduit que

pour $i \geq 1$, $\frac{1}{i} \geq \int_{k}^{k+1} \frac{1}{t}\,\mathrm{d}t$ et pour $i \geq 2$, $\int_{k-1}^{k} \frac{1}{t}\,\mathrm{d}t \geq \frac{1}{i}$.

— $H(n) = \sum_{i=1}^{n} \frac{1}{i} \geq \sum_{i=1}^{n} \int_{i}^{i+1} \frac{1}{t}\,\mathrm{d}t = \int_{1}^{n+1} \frac{1}{t}\,\mathrm{d}t = ln(n+1)$.

— $H(n) = \sum_{i=1}^{n} \frac{1}{i} = 1 + \sum_{i=2}^{n} \frac{1}{i} \leq 1 + \sum_{i=2}^{n} int_{i-1}^{i} \frac{1}{t}\,\mathrm{d}t = 1 + \int_{1}^{n-1} \frac{1}{t}\,\mathrm{d}t = 1 + ln(n)$

$\square$

Exercice 2 On se donne trois pièces dont une et une seule est biaisée de telle sorte que $Pr(Face) = 2/3$. On lance les pièces et on obtient Face Face Pile. Quelle est la probabilité que la première pièce soit la pièce biaisée ?

Correction

— B : l'évènement tel que les trois lancés sont Face Face Pile dans cet ordre.

— E_i l'évènement tel que le i-ème lancé utilise la pièce biaisée. $Pr(E_i) = 1/3$

— On a $Pr(B|E_1) = Pr(B|E_2) = \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{6}$ et $Pr(B|E_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2}$

— en appliquant la loi de Bayes,

$$Pr(E_1|B) = \frac{Pr(B|E_1)Pr(E_1)}{\sum_{i=1}^{3} Pr(B|E_j)Pr(B|E_j)} = \frac{2}{5}$$

$\square$

Exercice 3 Considérons un module de détection de spam de courrier électronique.

— Le module réussit à identifier les courriers indésirables dans 99% des cas.

— Son taux de faux positifs est toutefois de 2% . C'est à dire le module annonce qu'un message est indésirable alors qu'il ne l'est pas.

Les statistiques officielles indiquent que 10% du courriers électroniques recçus est indésirable.

Quelle est la probabilité qu'un message soit effectivement indésirable lorsque le module indique que c'est le cas ?

Correction

— B : l'évènement « le module détecte un message indésirable »

— X l'évènement « le message est indésirable » biaisée. $Pr(E_i) = 1/3$
— On a $Pr(X) = 1/10$ et $Pr(B|X) = 99/100$ $Pr(B|X^c) = 2/100$
— en appliquant la loi de Bayes,

$$Pr(X|B) = \frac{Pr(B|X)Pr(X)}{Pr(B|X)Pr(X) + Pr(B|X^c)Pr(X^c)} = \frac{1/10 \cdot 99/100}{1/10 \cdot 99/100 + 9/10 \cdot 2/100}$$

□

Exercice 4 Montrer que la loi géométrique est sans mémoire : soit Z une variable aléatoire Z de loi géométrique de paramètres n et p. Alors pour tous $k \geq 0$ et $n \geq 1$,

$$Pr(X = n + k | X > k) = P(X = n)$$

Correction
A faire □

Exercice 5 Soient X et Y des variables aléatoires réelles d'espérance finie et a et b des réels. Alors montrer que $E[aX + bY] = aE[X] + bE[Y]$.

Correction
A FAIRE □

Exercice 6 Calculer l'espérance
— de variable aléatoire Y de loi de Bernoulli de paramètre p.
— de variable aléatoire X de loi binomiale de paramètres n et p.
— de variable aléatoire Z de loi géométrique de paramètres n et p.

Correction
A FAIRE

□

Problème MAX_SAT.

Problème : F formule en forme normale conjonctive (CNF) (par exemple, $F = (x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \neg x_3 \vee x_4) \wedge (\neg x_1 \vee x_2 \vee x_4) \ldots)$.
Problème : quel est le nombre maximal de clauses satisfiables ?
Le problème de décision accocié est NP-complet.

Notons $x_1, \ldots, x_n$ les variables de la formule F. Considérons l'algorithme suivant :

1. chaque variable prend x_i la valeur vrai ou faux suivant la loi de Bernouilli de paramètre de manière indépendante

Exercice 7 Notons k_j le nombre de littéraux de la j-ème clause. Calculer l'espérence de l'évènement E_j « *la j-ème clause est satisfaite* »

Correction
□

Exercice 8 Notons $k = min_{j=1}^{m} k_j$. Considérons la variable aléatoire Y correspondant au nombre de clause satisfaites. Montrer que $E[Y] \geq m(1 - 2^{-k})$.

Correction
A FAIRE

□

Exercice 9 Soit OPT la solution optimale et notons $|OPT|$ le nombre de clauses satisfaites par OPT Montrer que $E[Y] \geq 1/2|OPT|$

Correction
Il suffit de remarque $m \geq |OPT|$ et $(1 - 2^{-k}) \geq 1/2$ □

Remarque : Cet algorithme probabilite est un algorithme d'approximation de rapport 1/2.

Problème de la coupe minimun.

Une coupure d'un graphe est un ensemble d'arêtes telles que les enlever coupe le graphe en deux ou plus composantes connexes. Étant donné un graphe $G = (V, E)$, le problème de la coupure **minimale** consiste à déterminer une coupure de cardinalité minimale du graphe.

Algorithme randomisé On opère $n - 2$ itérations, où n est le nombre de sommets. A chaque itération, on choisit aléatoirement selon une loi uniforme une arête $\{u, v\}$ du graphe, et on la contracte :

- c'est-à-dire, on fusionne u et v en un unique sommet, on élimine toutes les arêtes entre u et v et on garde toutes les autres arêtes.
- Le nouveau graphe obtenu peut avoir des multi-arêtes (plusieurs arêtes entre deux mêmes sommets) mais pas de boucle.

Enfin, les arêtes restantes sont retournées.

Exercice 10 Soit G un graphe à n sommets et $\mathcal{C}$ une coupe minimale de G. Montrer que si $\mathcal{C}$ est de cardinalité k, alors G a au moins $\frac{kn}{2}$ arêtes.

Correction
Soit u un sommet arbitraire de v . Alors si on enlève toutes les arêtes incidentes à un sommet, alors ce sommet devient sa propre composante connexe, et les arêtes incidentes en ce sommet constituent une coupure.

Si chaque sommet est adjacent à au moins k arêtes,
le graphe doit donc avoir au moins $\frac{nk}{2}$ arêtes.

□

Exercice 11 Montrer que une coupe $\mathcal{C}$ sera retournée par notre algorithme si et seulement si aucune des arêtes de $\mathcal{C}$ n'est sélectionée pour être contractée.

Correction
Soit une coupe $\mathcal{C}$ retournée par notre algorithme. Il est évident car à chaque étape l'arête contractée est retirée du graphe.

Supposons que aucune arête de $\mathcal{C}$ n'a été contractée.

1. Toutes arêtes de $\mathcal{C}$ sont retournées. Il n'y a que 2 raisons pour qu'une arête ne soit pas retournée : être contractée ou être transformée en boucle. Si une arête de $\mathcal{C}$ a été transformée en boucle, cela implique que ses 2 extrêmités ont éré fusionnées, ce qui n'est possible que si une autre arête de $\mathcal{C}$ a été contractée.

2. Aucune arête n'appartenant pas à $\mathcal{C}$ n'est retournée. Vrai car l'algorithme ne se termine que lorsqu'il ne reste que 2 sommets (correspondent aux 2 parties séparées par $\mathcal{C}$) : toute arête qui n'appartient pas à $\mathcal{C}$ joint 2 sommets de la même partie et a été éliminée soit par contraction soit comme boucle.

$\square$

Soit $\mathcal{C}$ une coupure minimale de taille k. Notons

— E_i l'événement que l'arête contractée à l'itération i n'est pas dans C.

— $F_i = \bigcap_{j=1}^{i} E_j$ l'événement qu'**aucune** arête de C **n'**a été contractée pendant les premières i itérations.

Exercice 12 Montrer que $\Pr(E_1) = \Pr(F_1) \geq 1 - \frac{2}{n}$.

Correction

$\Pr(E_1) = 1 - \frac{k}{m}$. Comme $m \geq \frac{kn}{2}$, on $\frac{2}{n} \geq \frac{k}{m}$ et $-\frac{k}{m} \geq -\frac{2}{n}$ $\square$

Exercice 13 Montrer que que $\Pr(E_2|F_1) \geq 1 - \frac{2}{n-1}$.

Correction

Il suffit de réaliser que maintenant dans le nouveau graphe, $\mathcal{C}$ est toujours une coupe maximal donc nous avons la relation suivante

$$m - 1 \geq \frac{k(n-1)}{2}$$

Donc $-\frac{k}{m-1} \geq -\frac{2}{n-1}$

$$\Pr(E_2|F_1) = 1 - \frac{k}{m-1} \geq 1 - \frac{2k}{(n-1)k} \geq 1 - \frac{2}{n-1}.$$

$\square$

Exercice 14 Montrer que $\Pr(E_i|F_{i-1}) \geq 1 - \frac{2}{n-i+1}$ et que $\Pr(F_{n-2}) \geq \frac{2}{n(n-1)}$.

Correction

Après la (i-1)-ème itération, Il suffit de réaliser que maintenant dans le nouveau graphe, $\mathcal{C}$ est toujours une coupe maximal donc nous avons la relation suivante

$$m - 1 \geq \frac{k(n-i+1)}{2}$$

$\Pr(E_i|F_{i-1}) \geq 1 - \frac{2k}{(n-i+1)k} = 1 - \frac{2}{n-i+1}$.

$$\begin{aligned}
\Pr(F_{n-2}) &= \Pr(E_{n-2}|F_{n-3})\Pr(E_{n-3}|F_{n-4})\cdots\Pr(E_2|F_1)\Pr(F_1). \\
&\geq \prod_{i=1}^{n-2}(1 - \frac{2}{n-i+1}) \geq \prod_{i=1}^{n-2}(\frac{n-i-1}{n-i+1}) \\
&\geq \frac{n-2}{n}\frac{n-3}{n-1}\frac{n-4}{n-2}\cdots\frac{4}{6}\frac{3}{5}\frac{2}{4}\frac{1}{3} \geq \frac{2}{n(n-1)}.
\end{aligned}$$

$\square$

Exercice 15 En déduire que l'algorithme produit une coupure minimale avec probabilité au moins $\frac{2}{n(n-1)}$.

Correction

Rappellons la définition de F_{n-2}, l'événement qu'**aucune** arête de C **n'**a été contractée pendant les premières $n-2$ itérations.

Cela signifie que l'évènement se réalise alors, l'algorithme retourne une coupe minimum.

$\square$

Exercice 16 Supposons que l'on exécute l'algorithme $n(n-1)\log n$ **fois**, et que l'on produise en sortie la plus petite coupure trouvée dans toutes les itérations. Borner la probabilité que l'on ne produise pas une coupure minimale. *Indication : $1-x \le e^{-x}$ pour $0 \le x \le 1$.*

Correction

- L'algorithme produit une coupure minimale avec probabilité au moins $\frac{2}{n(n-1)}$.
- Puisque l'algorithme est à erreur unilatérale, on peut réduire l'erreur facilement comme précédemment :
  - supposons que l'on exécute l'algorithme $n(n-1)\log n$ **fois**, et que l'on produise en sortie la plus petite coupure trouvée dans toutes les itérations.
  - La probabilité que l'on ne produise pas une coupure minimale est bornée par

$$\left(1-\frac{2}{n(n-1)}\right)^{n(n-1)\log n} \le e^{-2\log n} = \frac{1}{n^2},$$

où l'on a utilisé le fait que $1-x \le e^{-x}$ pour $0 \le x \le 1$.

$\square$