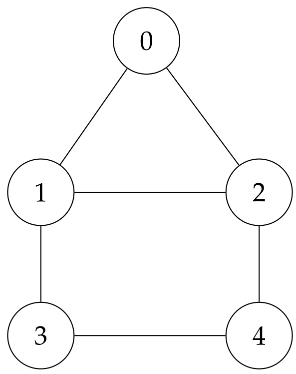
TD 12 – Chaînes de Markov (distributions stationnaires) & Stats (corrigé)

Exercice 1.
On considère une marche aléatoire sur le graphe suivant (c'est à dire qu'à chaque étape, on choisit le sommet suivant uniformément au hasard parmi les voisins du sommet courant).



1. On suppose que la distribution initiale est $\pi_0 = (1, 0, 0, 0, 0)$ (i.e. $X_0 = 0$ avec probabilité 1). Le vecteur de distribution π_n converge-t-il lorsque n tend vers l'infini ? Si oui, déterminer sa limite.

☞ La chaîne de Markov associée à la marche aléatoire sur le graphe est irréductible (le graphe est connexe). Elle est aussi apériodique grâce au triangle formé par les états $0, 1$ et 2. En effet, la période de l'état 0 est le pgcd de 2 (un aller retour $0-1-0$) et 3 (un cycle $0-1-2-0$). Donc l'état 0 est apériodique. Comme la chaîne est irréductible tous les états sont apériodiques. Si on trouve une distribution stationnaire, le théorème de convergence nous assurera une convergence vers cette distribution stationnaire. On observe que la distribution $\pi = (1/6, 1/4, 1/4, 1/6, 1/6)$ est stationnaire (on peut la calculer en résolvant le système d'équations $\pi Q = \pi$ par exemple – *on peut remarquer que les états* 1 *et* 2 *se comportent de la même façon, de même pour* 3 *et* 4*, donc il suffit de considérer un système à trois inconnues.*). Le théorème de convergence nous permet de conclure que $\pi_n \to \pi$ quand n tend vers l'infini.

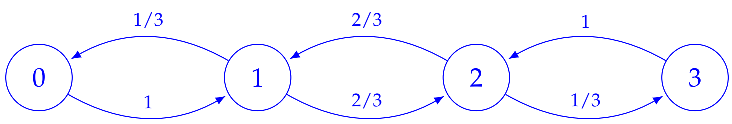
2. Même question si la distribution initiale est $\pi_0 = (0, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$.

☞ La réponse est la même, le théorème de convergence ne dépend pas de l'état initial.

Exercice 2.
On étudie un modèle simple des échanges de molécules de gaz entre deux récipients, et on s'intéresse à l'évolution du nombre de molécules dans le premier récipient. Soit N le nombre total de molécules dans les deux récipients. L'échange est modélisé de la façon suivante : si le premier récipient contient x molécules à l'instant t, alors à l'instant $t+1$ il contient $x+1$ molécules avec probabilité $\frac{N-x}{N}$ et $x-1$ molécules avec probabilité $\frac{x}{N}$.

1. Décrire ce modèle avec une chaîne de Markov, en donnant l'espace des états et la matrice de transition. Pour $N = 3$, dessiner la chaîne de Markov correspondante.

☞



Pour $N \in \mathbb{N}$, on a les formules suivantes : $P_{i,i+1} = \frac{N-i}{N}$, $P_{i,i-1} = \frac{i}{N}$, $P_{i,j} = 0$ sinon.

2. Pour $N \in \mathbb{N}$, donner la distribution stationnaire de la chaîne de Markov.

☞ On utilise la propriété des flux : si on a une 2-partition de l'ensemble des états S en $S_1 \uplus S_2$, le « flux sortant » de S_1 vers S_2 est égal au « flux entrant » de S_2 vers S_1 lorsque l'on est dans un état d'équilibre (*i.e.* lorsque l'on suit une distribution stationnaire). Plus formellement :

$$\sum_{i\in S_1}\sum_{j\in S_2}\pi(i)P_{i,j}=\sum_{i\in S_1}\sum_{j\in S_2}\pi(j)P_{j,i}\,.$$

Ici, en regardant une coupe entre i et $i+1$ pour $0\le i<N$, on sait donc que la distribution stationnaire π satisfait :

$$\pi_i\cdot\tfrac{N-i}{N}=\pi_{i+1}\tfrac{i+1}{N}\qquad\text{i.e.}\qquad \begin{aligned}\pi_{i+1}&=\tfrac{N-i}{i+1}\pi_i\\&=\pi_0\cdot\textstyle\prod_{j=0}^{i}\tfrac{N-j}{j+1}\\&=\tbinom{N}{i+1}\pi_0\end{aligned}$$

Donc pour $0\le i\le N$, on a $\pi_i=\binom{N}{i+1}\pi_0$. Mais on sait aussi que $\sum_i\pi_i=1$, *i.e.* $\pi_0\cdot\sum_{i=0}^{N}\binom{N}{i+1}=1$. On en déduit que $\pi_0=2^{-N}$.

3. On suppose qu'à l'instant $t=0$, le premier récipient est vide (et le deuxième récipient contient N molécules). On note $T\le 1$ le prochain instant où le premier récipient est vide. Calculer $\mathbf{E}[T]$.

☞ On sait que pour une chaîne irréductible finie, $h_i=\frac{1}{\pi_i}$. En particulier, pour $i=0$, on a $h_0=\frac{1}{\pi_0}=2^n$.

Exercice 3.
Modèle de Poisson
Le modèle de Poisson est l'ensemble $\{\mathcal{P}(\lambda),\lambda>0\}$. On rappelle que $\mathcal{P}(\lambda)$, pour un certain $\lambda>0$, correspond a la loi donnée par

$$\forall k\in\mathbb{N},\quad \mathbb{P}_\lambda(X=k)=\exp(-\lambda)\frac{\lambda^k}{k!}. \tag{1}$$

Vocabulaire :

— Un estimateur $\tilde\lambda_n$ de λ est *convergent* si pour tout $\lambda\in\Lambda$, $\tilde\lambda_n$ converge vers λ en $\mathbb{P}_\lambda$-probabilité, c'est à dire si $\forall\epsilon>0$,

$$\lim_n\mathbb{P}_\lambda(|\tilde\lambda_n-\lambda|\ge\epsilon)=0 \tag{2}$$

— Le *bias* d'estimateur $\tilde\lambda_n$ de λ (tel que $\mathbb{E}_\lambda[\tilde\lambda_n]<\infty$) est la fonction $b(\tilde\lambda_n,\cdot):\Lambda\to\mathbb{R}$ définie comme

$$b(\tilde\lambda_n,\lambda)=\mathbb{E}_\lambda[\tilde\lambda_n]-\lambda \tag{3}$$

Estimateur moment d'ordre 1

1. Calculer $\mathbb{E}_\lambda[X]$ et déduire un estimateur de moment $\tilde\lambda_n^{(1)}$ ☞

$$\mathbb{E}_\lambda[X]=\sum_k k\exp(-\lambda)\frac{\lambda^k}{k!} \tag{4}$$
$$=\lambda \tag{5}$$

Donc on peut prendre $\tilde\lambda_n^{(1)}(X_1,\dots,X_n)=\bar X_n:=\frac{1}{n}\sum_i(X_i)$

2. Quel est le biais de cet estimateur ? ☞ On a $\mathbb{E}_\lambda[\frac{1}{n}\sum_i X_i]=\mathbb{E}_\lambda[X_1]=\lambda$ donc c'est un estimateur non biaisé.
3. Cet estimateur est il consistant ? ☞ Oui, car une loi de Poisson admet un moment d'ordre 2 et on conclut par la loi faible des grands nombres (application de Tchebychev)

Estimateur moment d'ordre 2

1. Calculer un autre estimateur $\tilde\lambda_n^{(2)}$ d'après l'expression de $\mathbb{E}_\lambda[X^2]$. ☞ On trouve $\mathbb{E}_\lambda[X^2]=\lambda(\lambda+1)$ Soit $m(\lambda)=\lambda^2+\lambda$. On a pour tout $s>0$

$$m(\lambda)=s\iff\lambda=\sqrt{\frac14+s}-\frac12=\phi(s) \tag{6}$$

On a que $\tilde\lambda^{(2)}(X_1,\dots,X_n)=\phi(\frac1n\sum_i X_i^2)$ est un estimateur de λ.

2. Cet estimateur est-il non-biaisé (i.e. $b(\tilde\lambda_n^{(2)},\lambda)=0$) ? ☞ ϕ est strictement concave donc par l'inégalité de Jensen, il vient

$$\mathbb{E}_\lambda[\tilde\lambda_n^{(2)}]=\mathbb{E}_\lambda[\phi(\frac1n\sum_i X_i^2)]<\phi(\mathbb{E}_\lambda\frac1n\sum_i X_i^2)=\phi(\lambda^2+\lambda)=\lambda \tag{7}$$

Donc $\mathbb{E}_\lambda[\tilde\lambda_n^{(2)}]-\lambda<0$, cet estimateur est biaisé.

Estimateur Maximum de Vraisemblance

1. Calculer la vraisemblance du modèle. ☞

$$L(\lambda|x_1,\dots,x_n) = \prod_{i=1}^{n} \mathbb{P}_\lambda(X_i = x_i) = e^{-n\lambda} \frac{\lambda^{\sum_i x_i}}{x_1!\dots x_n!}. \tag{8}$$

2. Calculer l'Estimateur Maximum de Vraisemblance (EMV). ☞ On va calculer

$$\mathrm{argmax}_{\lambda>0} \log L(\lambda|x_1,\dots,x_n) = \mathrm{argmax}_{\lambda>0} - n\lambda + \sum_i x_i \log\lambda \tag{9}$$

Soit $f(\lambda) := -n\lambda + \sum_i x_i \log\lambda$ et on calcule $f'(\lambda) = -n + \sum_i x_i \frac{1}{\lambda}$ et $f'(\lambda) = 0 \implies \lambda = \sum_i x_i/n$? Ainsi, l'EMV pour un modèle Poisson vaut $\hat\lambda_n(X_1,\dots,X_n) = \bar X_n$.

Exercice 4.

Une *k-coloration* d'un graphe est une fonction des sommets du graphe vers un ensemble de k couleurs. On dit qu'elle est *propre* si deux sommets adjacents ne reçoivent jamais la même couleur. Un graphe est dit *k-colorable* s'il possède une k-coloration propre.

Soit G un graphe 3-colorable.

1. Prouver qu'il existe une 2-coloration (non propre) telle qu'aucun triangle n'est monochromatique (un triangle est *monochromatique* si les trois sommets qui le composent reçoivent la meme couleur).

☞ Il existe une coloration propre Rouge, Bleu, Vert. On recolorie les sommets verts en rouge. Chaque triangle contenait déjà un sommet rouge et un sommet bleu avant la recoloration, et c'est toujours le cas après.

On considère maintenant l'algorithme suivant, dont le but est de trouver une telle 2-coloration :

On commence avec une 2-coloration arbitraire.

Tant qu'il y a un triangle monochromatique, on choisit uniformément un sommet parmi les trois sommets qui le composent, et on change sa couleur.

On veut étudier l'espérance du nombre de recolorations avant de s'arrêter.

Comme G est 3-colorable, il existe une coloration propre pour $\{\text{Rouge}, \text{Bleu}, \text{Vert}\}$ (mais on ne la connaît pas). On note R (resp. B, V) l'ensemble des sommets colorés rouge (resp. bleu, vert) dans cette 3-coloration.
Considérons maintenant une 2-coloration arbitraire c de G (en rouge et bleu, disons). Soit $m(c)$ le nombre de sommets de R qui ne sont pas colorés rouge dans c, plus le nombre de sommets de B qui ne sont pas colorés bleu dans c.

2. Que dire si $m(c) = n$ ou $m(c) = 0$?

☞ Dans ces cas, aucun triangle n'est monochromatique et on a terminé.

3. En s'inspirant de l'exemple du cours sur 2-SAT, modéliser l'évolution de $m(c)$ par une chaîne de Markov sur $\{0,\dots,n\}$. Quel·s sont le ou les sommet·s à atteindre pour terminer ?
Pour $j \in \{0,\dots,n\}$, que peut-on dire de l'état j par rapport à l'état $n-j$?

☞ Supposons $m(c) = j \neq 0, n$ et regardons le triangle monochromatique choisi. Sans perte de généralité, on peut supposer qu'il est entièrement rouge. Avec proba $1/3$, on tire le sommet de V et on le recolore (en bleu) ; cela laisse $m(c)$ inchangé. Avec proba $1/3$, on tire le sommet de B et on le recolore ; on passe à $m(c)-1$. Avec proba $1/3$, on tire le sommet de R et on le recolore ; on passe de $m(c)$ à $m(c)+1$.

On obtient donc la chaîne de Markov suivante : pour $j \neq 0, n$, avec proba $1/3$ on passe à $j-1$, avec proba $1/3$ on reste sur j, et avec proba $1/3$ on passe à $j+1$. Pour $j = 0$ ou n, on reste sur l'état courant avec proba 1. Le but est d'atteindre le sommet 0 ou le sommet n. La chaîne est complètement symétrique entre l'état j et l'état $n-j$.

4. Soit h_j l'espérance du nombre de recolorations à effectuer pour terminer lorsqu'on part d'une 2-coloration c pour laquelle $m(c) = j$. Pour $j \in \{1,\dots,n-1\}$, exprimer h_j en fonction de h_{j-1} et h_{j+1}. Déterminer h_0 et h_n.

☞ On a $h_0 = 0$ et $h_n = 0$. Pour $j \in \{1,\dots,n-1\}$, on a :

$$h_j = \frac{1}{3}(1 + h_{j-1} + 1 + h_j + 1 + h_{j+1})$$

autrement dit :

$$h_j = \frac{3}{2} + \frac{1}{2}(h_{j-1} + h_{j+1}).$$

5. Montrer que $h_j = h_{j+1} + f(j)$ pour une certaine fonction f à déterminer, avec $f(0) = -h_1$.

☞ On a $h_0 = h_1 + f(0) = h_1 - h_1 = 0$: ok. Par récurrence, on calcule :

$$h_j = \frac{3}{2} + \frac{1}{2}(h_{j-1} + h_{j+1}) = \frac{3}{2} + \frac{1}{2}(h_j + f(j-1) + h_{j+1})$$

d'où

$$\frac{1}{2}h_j = \frac{3}{2} + \frac{1}{2}f(j-1) + \frac{1}{2}h_{j+1}$$

donc $h_j = h_{j+1} + 3 + f(j-1) = h_{j+1} + f(j)$ avec $f(j) = 3 + f(j-1)$, et donc $f(j) = 3j - h_1$.

6. Prouver que $h_{n/2} = \mathcal{O}(n^2)$. Conclure.

Indication : *On pourra utiliser la relation $h_1 = h_{n-1}$, obtenue par symétrie.*

☞ On a $h_{n-1} = h_n + f(n-1) = 0 + f(n-1)$. Comme $h_1 = h_{n-1}$ et $f(n-1) = 3(n-1) - h_1$, on obtient $h_1 = 3(n-1) - h_1$ d'où $h_1 = 3(n-1)/2$. Donc $h(j) = h_{j+1} + 3j - 3(n-1)/2$. Ainsi,

$$h_j = h_n + \sum_{k=j}^{n-1}(3j - \frac{3}{2}(n-1)) = 0 + 3\sum_{k=j}^{n-1} j - \frac{3(n-1)(n-j)}{2} = 3\frac{(n-1+j)(n-j)}{2} - \frac{3(n-1)(n-j)}{2} = 3\frac{(n-j)}{2}(n-1+j-(n-1)) = \frac{3j(n-j)}{2}$$

d'où

$$h_{n/2} = \frac{3(n/2)^2}{2} = \mathcal{O}(n^2)\,.$$

Comme $h_{n/2}$ est le "milieu" de la chaîne, c'est le pire cas (on peut vérifier que c'est le maximum de h_j) et donc l'espérance du nombre de recolorations est bien quadratique.

Exercice 5.

Une tour se déplace sur un échiquier de 8×8 cases uniformément au hasard (depuis la case de coordonnées (x, y), elle peut se déplacer sur une des cases de l'ensemble $\{(x', y) | x' = x\} \cup \{(x, yi) | y' = y\}$). Au bout de combien de déplacements en moyenne revient-elle à la case de départ ?

☞ On peut modéliser le problème par une chaine de Markov. Il y a 64 états. La chaine est apériodique, irréductible et récurrente. On va calculer la probabilité invariante : sur un graphe G, elle vaut $\pi(x) = \frac{\deg v}{2|E|}$. On en déduit que

$$\pi(x) = \frac{14}{14 \cdot 64} = \frac{1}{64} \tag{10}$$

La tour revient à la case de départ en moyenne en $\frac{1}{\pi}$ déplacements, i.e. 64.

Même question pour un cavalier qui part d'un coin de l'échiquier.

☞ Si v dénote le coin de l'échiquier, on a $\deg(v) = 2$ et on compte que $2|E| = 336$ d'où le temps de retour moyen qui vaut 168.