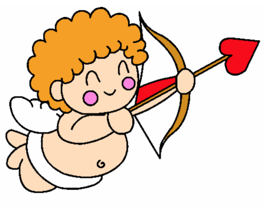
Mathématiques Appliquées - F. Gaunard
http://frederic.gaunard.com
ENC Bessières, Paris 17e.



Devoir Maison n°4

Solution

Exercice 1

Une urne contient 2 boules : une blanche et une rouge. On effectue dans cette urne une succession de tirages en respectant la règle suivante : si la boule tirée est blanche, on l'enlève du jeu et on ajoute une boule rouge dans l'urne. Si la boule tirée est rouge, on l'enlève du jeu et on ajoute une boule blanche dans l'urne.

Pour tout $i \in \mathbb{N}$, on note N_i le nombre de boules blanches contenues dans l'urne après le tirage i.

On introduit alors, pour tout $k \in \mathbb{N}^*$, les évènements R_k (resp. B_k) "la boule piochée au $k-$ième tirage est rouge (resp. blanche)".

(1) Si on remplace une boule par une boule de l'autre couleur, il y a alors deux boules de la même couleur et aucune autre ce qui entraîne ensuite un retour à la situation initiale. Le nombre de boules blanches varie donc entre 0 et 2. On a donc, pour tout $i \in \mathbb{N}$, $N_i(\Omega) = \{0, 1, 2\}$.

(2) Soit $(n_1, n_2, ..., n_{i+1}) \in \{0, 1, 2\}^{i+1}$,

$$P\left(N_{i+1} = n_{i+1} \mid \bigcap_{k=1}^{i} [N_k = n_k]\right) = P_{[N_i = n_i]}(N_{i+1} = n_{i+1}),$$

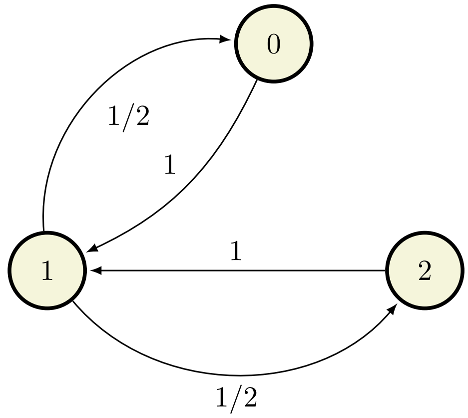
car le nombre de boules blanches après $i+1$ tirages ne dépend que du nombre de boules blanches avant ce tirage (c'est à dire après i tirages) et de la pioche qui suit. Cette probabilité conditionnelle est alors égale à

$$P_{[N_i=j]}(N_{i+1} = k) = \begin{cases} P_{[N_i=j]}(R_{i+1}), & \text{si } j = 0 \text{ et } k = 1 \\ P_{[N_i=j]}(B_{i+1}), & \text{si } j = 2 \text{ et } k = 1 \\ P_{[N_i=j]}(B_{i+1}), & \text{si } j = 1 \text{ et } k = 2 \\ P_{[N_i=j]}(R_{i+1}), & \text{si } j = 1 \text{ et } k = 0 \\ 0, & \text{sinon} \end{cases} = \begin{cases} 1, & \text{si } j = 0 \text{ et } k = 1 \\ 1, & \text{si } j = 2 \text{ et } k = 1 \\ 1/2, & \text{si } j = 1 \text{ et } k = 2 \\ 1/2, & \text{si } j = 1 \text{ et } k = 0 \\ 0, & \text{sinon} \end{cases}$$

La matrice de transition associée est donc donnée par

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}.$$

On peut représenter le graphe de transition comme suit :



(3) Sans trop de difficulté, vu que l'évolution de la chaîne est assez simple.

```
def simul_N(i) :
    N=1
    for k in range(i):
    if N==1 :
        if rd.random() <= 1/2 :
            N=2
        else :
            N=0
   else :
       N=1
   return N
```

(4) Commençons par déterminer l'ensemble des vecteurs propres à gauche $X = \begin{pmatrix} x & y & z \end{pmatrix}$ tels que $XA = X$. On a

$$XA = X \iff \left\{\begin{array}{rcl} 2x - y &=& 0 \\ x - y + z &=& 0 \\ y - 2z &=& 0 \end{array}\right. \iff \left\{\begin{array}{rcl} x &=& z \\ y &=& 2z \end{array}\right. \iff X = z\begin{pmatrix} 1 & 2 & 1 \end{pmatrix}, \quad z \in \mathbb{R}.$$

L'état stable est un vecteur stochastique; la somme de ses composantes doit faire 1. Ainsi, l'état stable de la chaîne est

$$\Pi = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

(5) Ce calcul rapide donne

$$A^2 = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix}, \quad A^3 = A.$$

On voit alors qu'il est nécessaire de distinguer la parité de i. Une récurrence, vraiment immédiate pour une fois, donne alors, pour tout $i \in \mathbb{N}^*$,

$$A^i = \left\{\begin{array}{ll} \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix}, & \text{si } i \text{ est pair} \\ \\ \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}, & \text{si } i \text{ est impair} \end{array}\right.$$

(6) Notons U_i le $i-$ème état probabiliste de la chaîne, c'est à dire, pour $i \in \mathbb{N}$,

$$U_i = \begin{pmatrix} P(N_i = 0) & P(N_i = 1) & P(N_i = 2) \end{pmatrix}.$$

On a notamment $U_0 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$. Par la formule des probabilités totales, on peut obtenir (on a un s.c.e avec $\{[N_i = j] : j \in [\![0, 2]\!]$ seulement pour $i \geq 2$ mais la formule s'étend facilement pour $i \geq 0$)

$$U_{i+1} = U_i A,$$

ce qui donne, par récurrence toute aussi immédiate, pour tout $i \in \mathbb{N}$,

$$U_i = U_0 A^i.$$

La loi de N_i dépend alors de la parité de i. On peut écrire

$$P(N_i = 0) = P(N_i = 2) = \left\{ \begin{array}{rl} 1/2, & \text{si } i \text{ est impair} \\ 0, & \text{si } i \text{ est pair} \end{array} \right. \quad \text{et} \quad P(N_i = 1) = \left\{ \begin{array}{rl} 1, & \text{si } i \text{ est pair} \\ 0, & \text{si } i \text{ est impair} \end{array} \right.$$

(7) Il est clair qu'une telle chaîne n'est pas convergente, et ce bien que l'état stable existe !

Exercice 2

Soit X une variable aléatoire.
On appelle *médiane* de X tout réel m qui vérifie les deux conditions :

$$P(X \leq m) \geq \frac{1}{2} \text{ et } P(X \geq m) \geq \frac{1}{2}.$$

On suppose que X suit la loi exponentielle de paramètre $\lambda > 0$.

(1) (a) Comme $F_X(x) = 0$ si $x < 0$, il est clair qu'un tel m est nécessairement positif. De l'expression $F_X(x) = 1 - e^{-\lambda x}$ si $x \geq 0$, on a

$$\begin{aligned} m \text{ médiane de } X \quad & \Longleftrightarrow \quad F_X(m) \geq 1/2 \text{ et } 1 - F_X(m) \geq 1/2 \\ & \Longleftrightarrow \quad 1 - e^{-m\lambda} \geq 1/2 \text{ et } e^{-m\lambda} \geq 1/2 \\ & \Longleftrightarrow \quad e^{-m\lambda} = 1/2 \\ & \Longleftrightarrow \quad m = \frac{\ln(2)}{\lambda}. \end{aligned}$$

Ceci donne bien que X admet une unique médiane m qui vaut $\dfrac{\ln(2)}{\lambda}$.

(b) On commence par obtenir l'expression de $M(x)$ en distinguant deux cas :

- Si $x < 0$, alors, comme $X(\Omega) = \mathbb{R}_+$, $|X - x| = X - x$ et par linéarité de l'espérance

$$M(x) = E(X) - x = \frac{1}{\lambda} - x.$$

- Si $x \geq 0$, alors

$$\begin{aligned} M(x) &= \int_0^{+\infty} \lambda |t - x| e^{-\lambda t} \mathrm{d}t \qquad \text{(théorème de transfert, sous réserve de convergence)} \\ &= \int_0^x \lambda(x - t) e^{-\lambda t} \mathrm{d}t + \int_x^{+\infty} \lambda(t - x) e^{-\lambda t} \mathrm{d}t \\ &= \frac{2}{\lambda} e^{-\lambda x} + x - \frac{1}{\lambda} \end{aligned}$$

où on a naturellement calculé les intégrales avec une IPP (que l'on fait sur l'intégrale partielle à droite avant de passer à la limite justifiant ainsi qu'elle était convergente) et permettant d'obtenir la formule simplifiée ci-dessus.

Au final, on a $M(x) = \left\{ \begin{array}{rl} \dfrac{1}{\lambda} - x, & \text{si } x < 0 \\ \dfrac{2}{\lambda} e^{-\lambda x} - \dfrac{1}{\lambda} + x, & \text{si } x \geq 0 \end{array} \right.$.

Une étude rapide permet de montrer que M est minimale en $x = \dfrac{\ln(2)}{\lambda}$.

(2) On suppose que le paramètre λ est inconnu. Soit α un réel vérifiant $0 < \alpha < 1$.
Pour n entier de $\mathbb{N}^*$, soit $(X_1, X_2, \ldots, X_n)$ un n-échantillon de variables aléatoires indépendantes et de même loi que X. On pose pour tout $n \in \mathbb{N}^*$: $Z_n = \min(X_1, X_2, \ldots, X_n)$.

(a) C'est encore une question classique de la loi du min. À savoir faire et refaire. Soit $x \geq 0$.

$$\begin{aligned} F_{Z_n}(x) &= P(Z_n \leq x) = 1 - P(Z_n > x) \\ &= 1 - P\left(\bigcap_{k=1}^{n} [X_k > x]\right) \\ &= 1 - \prod_{k=1}^{n} P(X_k > x) \qquad \text{par indépendance des } X_k \\ &= 1 - \left(e^{-\lambda x}\right)^n. \end{aligned}$$

Si $x < 0$, il est clair (toutes les variables X_k prennent des valeurs positives) que $F_{Z_n}(x) = 0$. Ainsi

$$F_{Z_n}(x) = \begin{cases} 0, & \text{si } x < 0 \\ 1 - e^{-n\lambda x}, & \text{si } x \geq 0 \end{cases}$$

et on reconnait une loi exponentielle $Z_n \hookrightarrow \mathcal{E}(n\lambda)$.

(b) On utilise la loi précédente

$$\begin{aligned} F_{Z_n}\left(\frac{c}{\lambda}\right) = \frac{\alpha}{2} &\iff 1 - e^{-nc} = \frac{\alpha}{2} \\ &\iff c = -\frac{1}{n}\ln\left(1 - \frac{\alpha}{2}\right) \end{aligned}$$

$$\begin{aligned} 1 - F_{Z_n}\left(\frac{d}{\lambda}\right) = \frac{\alpha}{2} &\iff e^{-nd} = \frac{\alpha}{2} \\ &\iff d = -\frac{1}{n}\ln\left(\frac{\alpha}{2}\right). \end{aligned}$$

(c) Observons que

$$\begin{aligned} \left[\frac{c}{\lambda} \leq Z_n \leq \frac{d}{\lambda}\right] &= \left[\frac{Z_n}{d} \leq \frac{1}{\lambda} \leq \frac{Z_n}{c}\right] \\ &= \left[\frac{\ln(2)Z_n}{d} \leq m \leq \frac{\ln(2)Z_n}{c}\right] \end{aligned}$$

Donc

$$P\left(m \in \left[\frac{\ln(2)Z_n}{c}; \frac{\ln(2)Z_n}{d}\right]\right) = P\left(\left[\frac{c}{\lambda} \leq Z_n \leq \frac{d}{\lambda}\right]\right) = 1 - \alpha$$

ou encore, l'intervalle $\left[\frac{\ln(2)Z_n}{c}; \frac{\ln(2)Z_n}{d}\right]$ est un intervalle de confiance pour m au niveau $1 - \alpha$.

Exercice 3

Soient $n \in \mathbb{N}^*$ et $p \in]0; 1[$. On considère une variable aléatoire $N \hookrightarrow \mathcal{B}(n, p)$ et une suite $(U_k)_{k \geq 0}$ de variables aléatoires indépendantes de même loi $\mathcal{U}([0; 1])$. On pose alors

$$T_n = \begin{cases} \min(U_1, U_2, ...U_k), & \text{si } [N = k] \text{ est réalisé avec } k \neq 0 \\ U_0, & \text{si } [N = 0] \text{ est réalisé} \end{cases}$$

De sorte à obtenir une convergence en loi, il nous faut la fonction de répartition de T_n. Notons, pour tout $k \in \mathbb{N}^*$, $M_k = \min(U_1, ..., U_k)$. Soit $x \in]0, 1[$. Commençons par observer que, par indépendance des U_i,

$$\begin{aligned} P(M_k \leq x) &= 1 - P(M_k > x) \\ &= 1 - P([U_1 > x] \cap ... \cap [U_k > x]) = 1 - (1 - x)^k. \end{aligned}$$

Par la formule des probabilités totales appliquée au s.c.e $\{[N=k] : N \in [\![0,n]\!]\}$, on a

$$\begin{aligned}
P(T_n \leq x) &= P_{[N=0]}(T_n \leq x)P(N=0) + \sum_{k=1}^{n} P_{[N=k]}(T_n \leq xP(N=k)) \\
&= P(U_0 \leq x)P(N=0) + \sum_{k=1}^{n} P(M_k \leq x)P(N=k) \\
&= x(1-p)^n + \sum_{k=1}^{n} \binom{n}{k} p^k(1-p)^{n-k}(1-(1-x)^k) \\
&= x(1-p)^n + \sum_{k=1}^{n} \binom{n}{k} p^k(1-p)^{n-k} - \sum_{k=1}^{n} \binom{n}{k} (p(1-x))^k(1-p)^{n-k} \\
&= x(1-p)^n + 1 - (1-p)^n + (1-p+p(1-x))^n - (1-p)^n \\
&= 1 + x(1-p)^n - 2(1-p)^n + (1-x)^n
\end{aligned}$$

Comme $x \in]0,1[$,

$$P(T_n \leq x) \underset{n\to+\infty}{\longrightarrow} 1.$$

Comme $P(T_n \leq x) = 0$ si $x \leq 0$ et $P(T_n \leq x) = 1$ si $x \geq 1$. On a , pour tout $x \neq 0$,

$$P(T_n \leq x) \underset{n\to+\infty}{\longrightarrow} \begin{cases} 0, & \text{si } x < 0, \\ 1, & \text{si } x > 0 \end{cases}$$

ce qui est la fonction de répartition de la variable aléatoire certaine égale à 0. On observera qu'on a pas montré la convergence en $x=0$, mais on n'a pas besoin de le faire; la convergence loi nécessite une convergence en tout point où la fonction de répartition de la loi limite est continue, et celle-ci n'est pas continue en 0. On peut alors écrire

$$T_n \xrightarrow{\mathcal{L}} 0.$$

Problème*

Ce problème n'est pas facile. Il s'adresse aux étudiant.e.s qui préparent les épreuves du Top 3.
Soit $N \in \mathbb{N}^*$ un entier fixé. On note, dans tout ce problème, $E = [\![1,N]\!]$.

Partie 1 : Préliminaires

(1) On a déjà répondu à cette question mille fois. On peut notamment utiliser la concavité de la fonction $x \mapsto \ln(1+x)$ qui est donc sous toutes ses tangentes, notamment celle en 0 d'équation $y = 1+x$ avec point de contact uniquement en $x=0$.

(2) Soient $(p_n)_{n\in E}$ et $(q_n)_{n\in E}$ deux distributions de probabilités. Soit $n \in E$. Observons que

$$\begin{aligned}
p_n\ln(q_n) - p_n\ln(p_n) &= p_n \ln\left(\frac{q_n}{p_n}\right) = p_n \ln\left(1 + \frac{q_n - p_n}{p_n}\right) \\
&\leq p_n \times \frac{p_n - q_n}{p_n} = p_n - q_n
\end{aligned}$$

avec égalité si et seulement si $p_n = q_n$.
En sommant sur $n \in E$, on a donc

$$\begin{aligned}
\sum_{n\in E} p_n\ln(q_n) - \sum_{n\in E} p_n\ln(p_n) &= \sum_{n\in E} (p_n\ln(q_n) - p_n\ln(p_n)) \\
&\leq \sum_{n\in E} (p_n - q_n) = \sum_{n\in E} p_n - \sum_{n\in E} q_n \\
&= 1 - 1 = 0
\end{aligned}$$

Ou encore

$$\sum_{n\in E} p_n\ln(q_n) \leq \sum_{n\in E} p_n\ln(p_n).$$

Pour que l'égalité ait lieu, il est nécessaire et suffisant que la somme de termes positifs $\sum_{n\in E}(p_n\ln(q_n)-p_n\ln(p_n))$ soit nulle ce qui est équivalent à ce que, pour tout $n\in E$, $p_n=q_n$.

Partie 2 : Estimateur du maximum de vraisemblance pour une matrice de transition

On considère une chaine de Markov $(X_k)_{k\geq 0}$ à N états, dont la loi initiale X_0 est connue et dont on cherche à *estimer* les coefficients de la matrice de transition $A=(p_{i,j})_{(i,j)\in E^2}$.

Notamment, on a, pour tout $k\in\mathbb{N}$, $X_k(\Omega)=E$ et, pour tout $(i,j)\in E^2$,

$$p_{i,j}=P_{[X_k=i]}(X_{k+1}=j)=p_{i,j}.$$

On suppose qu'on dispose alors d'une trajectoire de longueur $n+1$ de la chaîne (avec $n\in\mathbb{N}^*$), c'est dire des variables $(X_0,X_1,...,X_n)$.
Attention, contrairement au contexte classique des problèmes d'estimation, il ne s'agit pas d'un $n-$échantillon d'une même variable aléatoire.
Considérant alors une observation arbitraire $(x_1,...,x_n)\in E^n$ des n derniers termes de cette trajectoire et une observation initiale $x_0\in E$, on introduit la **fonction de vraisemblance** (conditionnelle à $X_0=x_0$) comme la fonction L définie sur l'ensemble des matrices stochastiques de taille N par

$$L(A)=\mathrm{P}_{[X_0=x_0]}\left([X_1=x_1]\cap[X_2=x_2]\cap...\cap[X_n=x_n]\right).$$

On observera que la matrice A contient les probabilités de transition de la chaîne et que la fonction L dépend bien des coefficients de A.
On note également, pour $(i,j)\in E^2$,

$$N_n(i,j)=\mathrm{Card}\left(\{k\in[\![1,n-1]\!]:\ X_k=i \text{ et } X_{k+1}=j\}\right)$$

et

$$M_n(i)=\mathrm{Card}\left(\{k\in[\![1,n-1]\!]:\ X_k=i\}\right).$$

(3) La quantité $N_n(i,j)$ représente le nombre de fois où on passe de l'état i vers l'état j dans l'observation $(x_0,...,x_n)$.
La quantité $M_n(i)$ représente le nombre de fois, où dans les n premières observations on était dans l'état i.

(4) Soit $i\in E$. Lorsqu'on est dans l'état i (pour l'une des n premières observations), on peut donc observer une évolution vers n'importe lequel des états de E à l'observation suivante. En additionnant tous les passages de i vers j, on aura donc le nombres de passages en i, ou encore

$$\sum_{j=1}^{N-1}N_n(i,j)=M_n(i).$$

(5) Par définition de la vraisemblance, la formule des probabilités composées et du fait que (X_k) soit une chaîne de Markov, on a

$$\begin{aligned}
L(A) &= \mathrm{P}_{[X_0=x_0]}\left([X_1=x_1]\cap[X_2=x_2]\cap...\cap[X_n=x_n]\right)\\
&= \mathrm{P}_{[X_0=x_0]}\left([X_1=x_1]\right)P_{[X_0=x_0]\cap[X_1=x_1]}([X_2=x_2])...P_{[X_0=0]\cap...\cap[X_{n-1}=x_{n-1}]}([X_n=x_n])\\
&= P_{[X_0=x_0]}P(X_1=x_1)P_{[X_1=x_1]}(X_2=x_2)...P_{[X_{n-1}=x_{n-1}]}(X_n=x_n)\\
&= \prod_{k=0}^{n-1}P_{[X_k=x_k]}(X_{k+1}=x_{k+1})
\end{aligned}$$

On regroupe alors toutes ces probabilités conditionnelles par couplage (i,j) de changement d'états. C'est à dire, que, pour $(i,j)\in E^2$ fixé, on a

$$\prod_{\substack{0\leq k\leq n-1\\ x_k=i,\ x_{k+1}=j}}P_{[X_k=x_k]}(X_{k+1}=x_{k+1})=p_{i,j}^{N_n(i,j)}$$

Il suit que

$$L(A) = \prod_{k=0}^{n-1} P_{[X_k=x_k]}(X_{k+1} = x_{k+1}) = \prod_{(i,j)\in E^2} \prod_{\substack{0\leq k\leq n-1 \\ x_k=i,\ x_{k+1}=j}} P_{[X_k=x_k]}(X_{k+1} = x_{k+1}) = \prod_{i=1}^{N}\prod_{j=1}^{N} p_{i,j}^{N_n(i,j)}.$$

En prenant le log, on a immédiatement,

$$\ln(L(A)) = \sum_{i=1}^{N}\sum_{j=1}^{N} N_n(i,j)\ln(p_{i,j}).$$

(6) Soit $i \in E$ tel que $M_n(i) \neq 0$ (c'est à dire un état par lequel on passe au moins une fois). Grâce à la question (4), on voit que

$$\left(\frac{N_n(i,j)}{M_n(i)}\right)_{j\in E}$$

est une distribution de probabilité. Comme $(p_{i,j})_{j\in E}$ en est aussi une, la Question (2) donne alors immédiatement

$$\sum_{j=1}^{N} \frac{N_n(i,j)}{M_n(i)} \ln(p_{i,j}) \leq \sum_{j=1}^{N} \frac{N_n(i,j)}{M_n(i)} \ln\left(\frac{N_n(i,j)}{M_n(i)}\right).$$

On note alors, pour tout couple $(i,j) \in E^2$, $\hat{p}_{i,j} = \begin{cases} \dfrac{N_n(i,j)}{M_n(i)}, & \text{si } M_n(i) \neq 0 \\ 0, & \text{sinon} \end{cases}$ et $\hat{A} = (\hat{p}_{i,j})_{(i,j)\in E^2}$.

(7) Soit $A = (p_{i,j})$ stochastique de $\mathcal{M}_N(\mathbb{R})$. D'après la question précédente, pour tout i tel que $M_n(i) \neq 0$, on a (en multipliant par $M_n(i)$,

$$\sum_{j=1}^{N} N_n(i,j)\ln(p_{i,j}) \leq \sum_{j=1}^{N} N_n(i,j)\ln\left(\frac{N_n(i,j)}{M_n(i)}\right).$$

ce qui implique, en sommant sur i :

$$\sum_{\substack{1\leq i\leq n \\ M_n(i)\neq 0}} \sum_{j=1}^{N} N_n(i,j)\ln(p_{i,j}) \leq \sum_{\substack{1\leq i\leq n \\ M_n(i)\neq 0}} \sum_{j=1}^{N} N_n(i,j)\ln\left(\frac{N_n(i,j)}{M_n(i)}\right).$$

puis en prenant l'exponentielle

$$\prod_{\substack{1\leq i\leq n \\ M_n(i)\neq 0}} \prod_{j=1}^{N} p_{i,j}^{N_n(i,j)} \leq \prod_{\substack{1\leq i\leq n \\ M_n(i)\neq 0}} \prod_{j=1}^{N} \hat{p}_{i,j}^{N_n(i,j)}.$$

SI i est tel que $M_n(i) = 0$, alors $N_n(i,j) = 0$ et donc $p_{i,j}^{N_n(i,j)} = 1$ et $\hat{p}_{i,j}^{N_n(i,j)} = 0^0 = 1$. Ajouter les termes manquants en i du produit revient donc à multiplier par 1, et on peut écrire

$$\prod_{i=1}^{N}\prod_{j=1}^{N} p_{i,j}^{N_n(i,j)} \leq \prod_{i=1}^{N}\prod_{j=1}^{N} \hat{p}_{i,j}^{N_n(i,j)}.$$

ou encore

$$L(A) \leq L(\hat{A}).$$

L'estimateur $\hat{A}$ est alors l'estimateur du maximum de vraisemblance pour la matrice de transition de la chaîne de Markov.

(8) Python.

(a) On commence par remplir les coefficients de $A = (a_{i,j})$ selon les valeurs de $N_n(i,j)$. Pour ce faire, on parcourt la liste des observations et, à chaque tour de boucle, on augmente de 1 le coefficient $N_n(x_k, x_{k+1})$. Attention au décalage entre les indices en Python (qui commencent à 0) et ceux des états des de la chaîne...

```
def EMV_A(N, X) :
    n=len(X)
    A=np.zeros((N,N))
    M=np.zeros(N)
    for k in range(n-1):
        A[X[k]-1][X[k+1]-1]+=1
    for i in range(N) :
        M[i]=np.sum(A[i])
        if M[i] !=0 :
            A[i]=[A[i][j]/M[i] for j in range(N)]
    return A
```

(b) Sur 10000 réalisations de la loi $\mathcal{U}(\llbracket 1,3 \rrbracket)$, il est raisonnable de penser que chacune des valeurs $1, 2$ et 3 est prise autant de fois (conséquence de la loi faible des grands nombres). Ainsi, cela suggère donc une chaîne de Markov où le passage d'un état vers n'importe lequel des autres se fait avec la même probabilité uniforme. La matrice de transition correspondant à cette chaîne serait une matrice 3×3 dont tous les coefficients sont égaux à $1/3$. En exécutant ce programme, on obtient bien

Affichage Python

```
> > > EMV_A(3, rd.randint(1, 4, 10000))
array([[0.33542039, 0.32886106, 0.33571855],
[0.33944394, 0.32630614, 0.33424992],
[0.33185053, 0.32680902, 0.34134045]])
```

c'est à dire une matrice dont tous les coefficients sont proches de $1/3$.