Feuille du Chapitre 1 : Rappels et Chaîne de Markov (contrôlées)

1. Rappels

1.1. Tribus.

Exercice 1. Rappeler la définition d'une tribu $\mathcal{E}$ sur un ensemble E. Rappeler la définition d'un espace mesurable. Donner des exemples simples d'espaces mesurables. Rappeler la définition d'un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$.

Un sous ensemble $\mathcal{E}$ de l'ensemble des partie s de E est une tribu si

- $E \in \mathcal{E}$,
- $A \in \mathcal{E}$, alors $A^c \in \mathcal{E}$,
- $(A_n)_n$ est une suite d'éléments de $\mathcal{E}$, alors $\bigcup_n A_n \in \mathcal{E}$.

Un ensemble mesurable désigne un couple $(E, \mathcal{E})$, où E est un ensemble et $\mathcal{E}$ est une tribu.

Si E est un ensemble, alors $\{\emptyset, E\}$ est une tribu, $\mathcal{P}(E)$ l'ensemble des parties de E est aussi une tribu. Si $A \subset E$ est un sous ensemble de E, alors $\{\emptyset, A, A^c, E\}$ est une tribu sur E.

On rappelle que la tribu des Boréliens sur $\mathbb{R}$ (ou $\mathbb{R}^d$) est la tribu engendrée par les ouverts de $\mathbb{R}$ (respectivement de $\mathbb{R}^d$). On rappelle également que lorsque l'ensemble E est au plus dénombrable, la tribu "naturelle" sur E est l'ensemble des parties de E.

Une mesure μ sur un espace mesurable $(E, \mathcal{E})$ est une application de $\mathcal{E}$ dans $\mathbb{R} \cup \{+\infty\}$ telle que

- $\mu(\emptyset) = 0$.
- Si $(A_n)_n$ est une suite de $\mathcal{E}$ d'ensembles deux à deux disjoints, alors

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n).$$

Un espace de probabilité est un espace mesurable $(\Omega, \mathcal{F}$, munit d'une mesure $\mathbb{P}$ telle que $\mathbb{P}(\Omega) = 1$.

Exercice 2. Soit $(E, \mathcal{E})$ un espace mesurable et soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilités. Rappeler la définition d'une variable aléatoire $X : \Omega \to E$.

Un variable aléatoire de $(\Omega, \mathcal{F}, \mathbb{P})$ dans $(E, \mathcal{E})$ est une fonction $X : \Omega \to E$ telle que pour tout $B \in \mathcal{E}$, $X^{-1}(B) := \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{E}$.

Exercice 3. Soit $(E, \mathcal{E})$ un espace mesurable et soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilités. Soit X une variable aléatoire à valeur dans E. Montrer que

$$\sigma(X) = \{X^{-1}(B) : B \in \mathcal{E}\}$$

est une sous-tribu de $(\Omega, \mathcal{F}, \mathbb{P})$. Montrer que si $h : E \to E$ est une fonction mesurable, alors $h(X)$ est mesurable par rapport à $\sigma(X)$.

On a $X^{-1}(E) = \Omega$, alors $\Omega \in \sigma(X)$. Se plus, on rappel que pour tout ensemble $\mathcal{I}$ et toute collection $(E_i)_{i\in\mathcal{I}}$ de sous ensemble de E indexée par $\mathcal{I}$, $X^{-1}\left(\bigcup_{i\in\mathcal{I}} E_i\right) = \bigcup_{i\in\mathcal{I}} X^{-1}(E_i)$ et $X^{-1}(A^c) = \left(X^{-1}(A)\right)^c$ pour tout $A \subset E$. Ceci est vrai en particulier quand $\mathcal{I} = \mathbb{N}$ et quand $A \in \mathcal{E}$, ce qui garantit les trois points de la définition d'une tribu.

Comme h est mesurable, pour tout $B \in \mathcal{E}$, $h^{-1}(B) \in \mathcal{E}$. En particulier

$$h(X)^{-1}(B) = \{\omega \in \Omega : h(X(\omega)) \in B\} = \{\omega \in \Omega : X(\omega) \in h^{-1}(B)\} = X^{-1}(h^{-1}(B)).$$

Comme X est une variable aléatoire, et que h est mesurable, si $B \in \mathcal{E}$, $h^{-1}(B) \in \mathcal{E}$ et $h(X)^{-1}(B) = X^{-1}(h^{-1}(B)) \in \mathcal{F}$, et $h(X)$ est une variable aléatoire.

L'exercice précédent admet une réciproque, admise dans la proposition suivante :

Proposition. *Soit X, Y deux variables aléatoires de $(\Omega, \mathcal{F}, \mathbb{P})$ à valeur dans $(E, \mathcal{E})$. Alors X est mesurable par rapport à $\sigma(Y)$ si et seulement si il existe une fonction mesurable f de E dans E telle que $X = f(Y)$.*

1.2. Espérance conditionnelle. On rappel d'abord la définition et les propriétés de l'espérance conditionnelle :

PROPOSITION. *Soit $X \in L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ une variable aléatoire et $\mathcal{G} \subset \mathcal{F}$ une sous tribu de $\mathcal{F}$. Il existe une unique (presque surement) variable aléatoire $\mathcal{G}$ mesurable, notée $\mathbb{E}[X|\mathcal{G}]$ et appelée espérance conditionnelle de X sachant $\mathcal{G}$ (c'est à dire que $\sigma(\mathbb{E}[X|\mathcal{G}]) \subset \mathcal{G}$), telle que pour tout $A \in \mathcal{G}$,*

$$\mathbb{E}\big[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A\big] = \mathbb{E}\big[X\mathbb{1}_A\big].$$

De plus, l'espérance conditionnelle vérifie les propriétés suivantes :

- *L'espérance conditionnelle est linéaire, positive, et vérifie les inégalités classiques (Jensen, Hölder,...)*
- $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$,
- *Si Y est $\mathcal{G}$ mesurable, alors, presque surement, $\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}]$,*
- *Si X est indépendante de $\mathcal{G}$, alors $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.*
- *Si $\mathcal{H} \subset \mathcal{G}$ est une autre sous tribu de $\mathcal{F}$, alors $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$*
- *Si f est une fonction mesurable, que X est indépendante de $\mathcal{G}$ et que Y est $\mathcal{G}$ mesurable telles que $f(X,Y) \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, alors $\mathbb{E}[f(X,Y)|\mathcal{G}] = F(Y)$, où pour tout y,*

$$F(y) = \mathbb{E}[f(X,y)].$$

Par ailleurs si $\mathcal{G} = \sigma(Y)$, on note

$$\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)],$$

et par la proprosition précédente, il existe une fonction mesurable h telle que $\mathbb{E}[X|Y] = h(Y)$.

EXERCICE 4. Soit $X \sim \mathcal{N}(0, \sigma^2)$ et Y une variable aléatoire indépendante de X. On définit

$$Z = \exp(-\frac{\sigma^2}{2}Y^2 + XY).$$

(1) Calculer $\mathbb{E}[Z|Y]$.
(2) Montrer que $\mathbb{E}[Z] = 1$

On a

(1) Par les propriétés de l'espérance conditionnelle, on a

$$\mathbb{E}[Z|Y] = e^{-\frac{\sigma^2 Y^2}{2}}\mathbb{E}[e^{XY}|Y] = e^{-\frac{\sigma^2 Y^2}{2}}F(Y),$$

où $F(Y) = \mathbb{E}[e^{Xy}]$. De plus, comme $X \sim \mathcal{N}(0, \sigma^2)$, $F(y) = e^{\frac{\sigma^2 y^2}{2}}$. Ainsi, $\mathbb{E}[Z|Y] = 1$.

(2)
$$\mathbb{E}[Z] = \mathbb{E}[\mathbb{E}[Z|Y]] = \mathbb{E}[1] = 1.$$

2. Chaînes de Markov

EXERCICE 5. Soit $\mathcal{G} = \sigma(Y)$ où $Y : \Omega \to F$ est une variable aléatoire dans un espace dénombrable F. Soit X une variable aléatoire à valeur dans un espace dénombrable $\mathcal{S}$ et soit $i \in \mathcal{S}$. Montrer que

$$\mathbb{P}(X = i|\mathcal{G}) := \mathbb{E}[\mathbb{1}_{\{X=i\}}|\mathcal{G}] = \sum_{\substack{j \in F \\ \mathbb{P}(Y=j)>0}} \mathbb{P}(X = i|Y = j)\mathbb{1}_{Y=j}.$$

On rappelle que

$$\sigma(Y) = \{Y^{-1}(B) : B \in \mathcal{P}(F)\},$$

où $\mathcal{P}(F)$ désigne l'ensemble des parties de F. De plus, comme F est dénombrable, on a

$$Y^{-1}(B) = \coprod_{j \in B} Y^{-1}(\{j\}),$$

où l'union est disjointe et au plus dénombrable. Ainsi, grâce au propriétés de l'espérance conditionnelle,

$$
\begin{aligned}
\mathbb{E}[\mathbb{E}[\mathbb{1}_{\{X=i\}}|\mathcal{G}]\mathbb{1}\mathbb{1}_{Y^{-1}(B)}] =& \mathbb{E}[\mathbb{1}_{\{X=i\}}\mathbb{1}\mathbb{1}_{Y^{-1}(B)}] \\
=& \sum_{j\in B} \mathbb{E}[\mathbb{1}_{X=i}\mathbb{1}_{Y^{-1}(\{j\})}] \\
=& \sum_{j\in B} \mathbb{P}(X=i, Y=j) \\
=& \sum_{\substack{j\in B\\ \mathbb{P}(Y=j)>0}} \mathbb{P}(X=i|Y=j)\mathbb{P}(Y=j) \\
=& \mathbb{E}\left[\sum_{\substack{j\in F\\ \mathbb{P}(Y=j)>0}} \mathbb{P}(X=i|Y=j)\mathbb{1}_{Y=j}\mathbb{1}_{Y^{-1}(B)}\right].
\end{aligned}
$$

où l'avant dernière égalité provient de la formule des probabilités totales. Notons également que

$$\sum_{\substack{j\in F\\ \mathbb{P}(Y=j)>0}} \mathbb{P}(X=i|Y=j)\mathbb{1}_{Y=j} \in \mathcal{G} = \sigma(Y).$$

Ainsi, grâce à la définition de l'espérance conditionnelle, on vient de montrer que

$$\mathbb{E}[\mathbb{1}_{\{X=i\}}|\mathcal{G}] = \sum_{\substack{j\in F\\ \mathbb{P}(Y=j)>0}} \mathbb{P}(X=i|Y=j)\mathbb{1}_{Y=j} p.s.$$

EXERCICE 6. Prouver la proposition suivante du cours concernant la forme canonique des chaînes de Markov.

PROPOSITION. *Soit $f : \mathbb{N}\times\mathcal{S}\times[0,1]\to\mathcal{S}$ une fonction mesurable.*
Soient $X_0 : \Omega\to\mathcal{S}$ et $(U_n)_{n\in\mathbb{N}\setminus\{0\}}$ une variable aléatoire et une suite de variables aléatoires telles que

(1) Pour tout $n\in\mathbb{N}$, $U_n\sim\mathcal{U}([0,1])$.
(2) $(U_n)_{n\in\mathbb{N}\setminus\{0\}}$ est une suite de variables aléatoires indépendantes.
(3) X_0 et $(U_n)_{n\in\mathbb{N}\setminus\{0\}}$ sous indépendantes.

Alors la suite $(X_n)_{n\in\mathbb{N}}$ définie pour tout $n\in\mathbb{N}$ par

$$X_{n+1} = f(n, X_n, U_{n+1})$$

est une chaîne de Markov de matrices de transition

$$((P(n,i,j))_{i,j\in\mathcal{S}})_{n\in\mathbb{N}} = ((\mathbb{P}(f(n,i,U_1)=j))_{i,j\in\mathcal{S}})_{n\in\mathbb{N}}.$$

On remarque par une récurence imémdiate que pour tout $n\in\mathbb{N}$, X_n est $\sigma(X_0, U_1, \cdots, U_n)$ mesurable. En particulier pour tout $n\in\mathbb{N}$ et tout $k\le n$, X_k est indépendant de U_{n+1}.

Soit $n\in\mathbb{N}$, $i_0,\cdots,i_n, j\in\mathcal{S}$. On a

$$
\begin{aligned}
\mathbb{P}(X_{n+1}=j, X_0=i_0,\cdots,X_n=i_n) =& \mathbb{P}(f(n,X_n,U_{n+1})=j, X_0=i_0,\cdots,X_n=i_n) \\
=& \mathbb{P}(f(n,i_n,U_{n+1})=j, X_0=i_0,\cdots,X_n=i_n) \\
=& \mathbb{P}(f(n,i_n,U_{n+1})=j)\mathbb{P}(X_0=i_0,\cdots,X_n=i_n).
\end{aligned}
$$

La premièer égalité provient de la définition de $(X_n)_{n\in\mathbb{N}}$. La seconde égalité vient du fait que l'on a spécifié $X_n=i_n$ et la troisième égalité vient de l'indépendance de $(X_0,\cdots,X_n)$ avec U_{n+1}.

EXERCICE 7. Prouver la proposition suivante qui montrer que toute chaîne de Markov peut s'écrire (en loi) sous la forme d'une chaîne de Markov canonique :

PROPOSITION. *Toute chaine de Markov d'espace d'état $\mathcal{S}$ au plus dénombrable et de matrices de transition $((P(n,i,j))_{i,j\in\mathcal{S}})_{n\in\mathbb{N}}$ admet une forme canonique.*

Indication : soit $n\in\mathbb{N}$ and $i\in\mathcal{S}$, soit $f(n,i,\cdot)$ la fonction quantile de la loi de poids $(P(n,i,j))_{j\in\mathcal{S}}$.

Comme $\mathcal{S}$ est au plus dénombrable, il existe une suite $(j_k)_{k\in\mathbb{N}}$ telle $\mathcal{S} = \coprod_{k\geq 0}\{j_k\}$. Pour $x \in [0,1]$, $n \in \mathbb{N}$ et $i \in \mathcal{S}$, définissons

$$f(n,i,x) = j_0 \mathbb{1}_{[0,P(n,i,j_0))}(x) + \sum_{k\geq 0} j_{k+1}\mathbb{1}_{\left[\sum_{\ell=0}^{k} P(n,i,j_\ell),\sum_{\ell=0}^{k+1} P(n,i,j_\ell)\right)}(x).$$

Soit $U \sim [0,1]$ et $k \in \mathbb{N}$. Alors

$$\begin{aligned}\mathbb{P}(f(n,i,U) = j_k) =& \mathbb{P}\left(U \in \left[\sum_{\ell=0}^{k-1} P(n,i,j_\ell), \sum_{\ell=0}^{k} P(n,i,j_\ell)\right)\right) \\ =& \sum_{\ell=0}^{k} P(n,i,j_\ell) - \sum_{\ell=0}^{k-1} P(n,i,j_\ell) = P(n,i,j_k).\end{aligned}$$

Ici nous avons utilisé la convention $\sum_{\ell=0}^{-1} P(n,i,j_\ell) = 0$. Notons également que f est une fonction mesurable. Sont $(U_n)_{n\in\mathbb{N}\setminus\{0\}}$ une suite de variable aléatoire iid de loi $\mathcal{U}([0,1])$ indépendente de X_0. Soit pour $n \in \mathbb{N}$, $\tilde{X}_{n+1} = f(n,\tilde{X}_n,U_{n+1})$ avec $\tilde{X}_0 = X_0$. Par la Proposition **??**, $(\tilde{X}_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de matrices de transition définie pour $n \in \mathbb{N}$ et $i,j \in \mathcal{S}$ par

$$\tilde{P}(n,i,j) = \mathbb{P}(f(n,i,U_1) = j) = P(n,i,j).$$

Ainsi, par récurrence, pour tout $n \in \mathbb{N}$ et $i_0,\cdots,i_n \in \mathcal{S}$,

$$\mathbb{P}(X_0 = i_0,\cdots,X_n = i_n) = \mathbb{P}(X_0 = i_0)\prod_{k=0}^{n-1} P(k,i_k,i_{,k+1}) = \mathbb{P}(\tilde{X}_0 = i_0,\cdots,\tilde{X}_n = i_n),$$

et $(\tilde{X}_n)_{n\in\mathbb{N}}$ et $(X_n)_{n\in\mathbb{N}}$ ont la même loi, ce qui est l'ennoncée de la proposition.

3. Chaînes de Markov contrôlées

3.1. Généralités.

Exercice 8. On suppose que $\mathcal{A} \in \mathcal{B}(\mathbb{R}^d)$ est un Borélien de $\mathbb{R}^d$ et que $\mathcal{S}$ est au plus dénombrable.

Soit

$$f : \mathbb{N}\times\mathcal{S}\times\mathcal{A} \to \mathcal{S}.$$

Montrer que f est mesurable si et seulement si pour tout $(n,i) \in \mathbb{N}\times\mathcal{S}$, $f(n,i,\cdot) : \mathcal{A} \to \mathcal{S}$ est mesurable au sens classique, c'est à dire que pour tout $j \in \mathcal{S}$,

$$\{a \in \mathcal{A} : f(n,i,a) = j\} \in \mathcal{B}(\mathbb{R}^d).$$

Remarquons que $f : \mathbb{N}\times\mathcal{S}\times\mathcal{A} \to \mathcal{S}$ est mesurable si pour tout $B \in \mathcal{P}(\mathcal{S})$, $f^{-1}(B)$ est mesurable pour la tribu sur $\mathbb{N}\times\mathcal{S}\times\mathcal{A}$. Notons que comme $\mathcal{S}$ est au plus dénombrable, la tribu à l'arrivée est naturellement l'ensemble des parties de $\mathcal{S}$, notée $\mathcal{P}(\mathcal{S})$. Si $B \in \mathcal{P}(\mathcal{S})$, il existe une suite $(j_k)_{k\in\mathbb{N}} \in \mathcal{S}^{\mathbb{N}}$ telle que $B = \bigcup_{k\in\mathbb{N}}\{j_k\}$ et $f^{-1}(B) = \bigcup_{k\in\mathbb{N}} f^{-1}(\{j_k\})$

Notons également que pour $j \in \mathcal{S}$,

$$f^{-1}(\{j\}) = \coprod_{\substack{n\in\mathbb{N}\\ i\in\mathcal{S}}}\left(f^{-1}(\{j_k\})\bigcap(\{n\}\times\{i\}\times\mathcal{A})\right)$$

et que

$$\begin{aligned}f^{-1}(\{j\})\bigcap(\{n\}\times\{i\}\times\mathcal{A}) =& \{(m,l,a) \in \mathbb{N}\times\mathcal{S}\times\mathcal{A} : f(m,l,a) = j\}\bigcap(\{n\}\times\{i\}\times\mathcal{A}) \\ =& \{(n,i,a) : a \in \mathcal{A},\ f(n,i,a) = j\}\end{aligned}$$

Comme f est mesurable,

$$f^{-1}(\{j\})\bigcap(\{n\}\times\{i\}\times\mathcal{A}) = \{(n,i,a) : a \in \mathcal{A},\ f(n,i,a) = j\} \in \mathcal{P}(\mathbb{N})\times\mathcal{P}(\mathcal{S})\times(\mathcal{A}\cap\mathcal{B}(\mathbb{R}^d)),$$

ce qui est équivalent à

$$\{a \in \mathcal{A} : f(n,i,a) = j\} \in \mathcal{A}\cap\mathcal{B}(\mathbb{R}^d),$$

et qui prouve le résultat.

Exercice 9 (Forme canonique des chaînes de Markov contrôlées 1). On suppose $\mathcal{A}$ fini. Soit $P : \mathbb{N} \times \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0,1]$ une collection de matrices de transition contrôlées et soit $(f_n : \mathcal{S} \to \mathcal{P}(\mathcal{A}))_{n\in\mathbb{N}}$ une suite de fonction de $\mathcal{S}$ à valeur de $\mathcal{P}(\mathcal{A})$, l'ensemble des mesures de probabilités sur $\mathcal{A}$.

Montrer qu'il existe une fonction $G : \mathcal{P}(\mathcal{A}) \times [0,1] \to \mathcal{A}$ telle que si $(V_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoire iid de loi $\mathcal{U}([0,1])$ indépendante de X_0, et que si $(X_n, A_n)_{n\in\mathbb{N}}$ est définit par la relation de récurrence

$$A_n = G(f_n(X_n), V_n), \quad X_{n+1} \sim P(n, X_n, A_n, \cdot),$$

alors $(X_n, A_n)_{n\in\mathbb{N}}$ est une chaîne de Markov contrôlée de matrices de transitions contrôlées P avec une suite de contrôles $(A_n)_{n\in\mathbb{N}}$ de forme mixte markovienne.

On remarque que si $\mathcal{A} = \{a_1, \cdots, a_{|\mathcal{A}|}\}$, l'ensemble des mesures de probabilité sur $\mathcal{A}$ peut être décrit sous la forme

$$\mathcal{P}(\mathcal{A}) = \left\{ \mu = (p_1, \cdots, p_{|\mathcal{A}|}) \in [0,1]^{|\mathcal{A}|} : \sum_{i=1}^{|\mathcal{A}|} p_i = 1 \right\}.$$

via l'identification $\mu = (p_1, \cdots, p_{|\mathcal{A}|}) \in \mathcal{P}(A)$, $\mu(\{a_k\}) = p_k$. De la même façon que pour la forme canonique des chaîne de Markov (non contrôlées), on écrit

$$G(\mu, u) = a_1 \mathbb{1}_{[0,\mu(a_1))}(u) + \sum_{i=2}^{|A|} a_i \mathbb{1}_{[\mu(a_1)+\cdots+\mu(a_{i-1}), \mu(a_1)+\cdots+\mu(a_i))}(u).$$

Soit $(V_n)_{n\in\mathbb{N}}$ une suite de variable iid de loi $\mathcal{U}([0,1])$

$$A_0 = G(f_0(X_0), U_0).$$

Remarquons aussi que pour $n \in \mathbb{N}$, $i \in \mathcal{S}$ et $a \in \mathcal{A}$, $(P(n,i,a,j))_{j\in\mathcal{S}}$ définit une mesure de probabilité sur $\mathcal{S}$, de la même façon que pour les mesures sur $\mathcal{A}$. Soit $X_1 \sim P(0, X_0, A_0, \cdot)$. On définit alors $A_1 = G(f_1(X_1), V_1)$, de telle sorte que A_1 est bien un contrôle mixte (en fonction de X_1 et V_1). Par récurrence maintenant, on suppose que $X_0, A_0, \cdots, X_n, A_n$ sont construits. On note alors $X_{n+1} \sim P(n, X_n, A_n, \cdot)$, de telle sorte que

$$\begin{aligned}
&(X_0 = i_0, A_0 = a_0, \cdots, X_n = i_n, A_n = a_n, X_{n+1} = j) \\
&\quad = \mathbb{E}[\mathbb{E}[\mathbb{1}_{X_{n+1}=j} | X_0 = i_0, A_0 = a_0, \cdots, X_n = i_n, A_n = a_n] \mathbb{1}_{X_0=i_0, A_0=a_0, \cdots, X_n=i_n, A_n=a_n}] \\
&\quad = \mathbb{E}[P(n, X_n, A_n, j) \mathbb{1}_{X_0=i_0, A_0=a_0, \cdots, X_n=i_n, A_n=a_n}] \\
&\quad = P(n, i_n, a_n, j) \mathbb{P}(X_0 = i_0, A_0 = a_0, \cdots, X_n = i_n, A_n = a_n). \qquad (1)
\end{aligned}$$

On définit alors $X_{n+1} = G(f_{n+1}(X_{n+1}), U_{n+1})$. On a bien définit une chaîne de Markov contrôlée $(X_n, A_n)_{n\in\mathbb{N}}$, de matrices de transition contrôlées P et telle que $(A_n)_{n\in\mathbb{N}}$ est sous forme Markovienne mixte.

Exercice 10 (Forme canonique des chaînes de Markov contrôlées 2). On suppose $\mathcal{A}$ fini. Pour tout $n \in \mathbb{N}$ on se donne $f_n : \mathcal{S} \to \mathcal{P}(\mathcal{A})$, où ici $\mathcal{P}(\mathcal{A})$ désigne l'ensemble des probabilités sur $\mathcal{A}$. $F : \mathbb{N} \times \mathcal{S} \times \mathcal{A} \times [0,1] \to \mathcal{S}$ mesurable. Soit $(U_n)_{n\in\mathbb{N}}$ et $(V_n)_{n\in\mathbb{N}}$ deux suites de variables aléatoires iid de loi $\mathcal{U}([0,1])$ et indépendantes de X_0. Soit $G : \mathcal{P}(A) \times [0,1] \to \mathcal{A}$ la fonction définie à l'exercice précédent.

Montrer que la suite de variables aléatoires $(X_n, A_n)_{n\in\mathbb{N}}$ définie pour tout $n \in \mathbb{N}$ par

$$A_n = G(f_n(X_n), V_n), \quad X_{n+1} = F(n, X_n, A_n, U_n)$$

est une chaîne de Markov contrôlée de matrices de transition contrôlées

$$(P(n,i,a,j) = \mathbb{P}(F(n,i,a,U_1) = j))_{n\in\mathbb{N}, i,j\in\mathcal{S}, a\in\mathcal{A}}$$

pour la filtration $(\mathcal{F}_n = \sigma(X_0, V_0, U_0, \cdots V_n, U_n))_{n\in\mathbb{N}}$ et telle que $(A_n)_{n\in\mathbb{N}}$ s'éxrite sous forme mixte markovienne.

On utilise l'exercice précédent, ce qui permet de conclure.

3.2. Exemples.

Exercice 11 (Approvisionnement de stock). Un centre de livrasion s'approvisionne de la façon suivate :

- S_0 : stock initial
- A l'instant $n \in \mathbb{N}$, S_n représente l'état du stock (un entier dans $\{0, \cdots, N_S\}$) et A_n représente l'approvisionnement (un entier das $\{0, \cdots, N_A\}$).
- Entre n et $n+1$, D_{n+1} représente la demande aléatoire (indépendante des observations passées), de loi homogène (sur $\{0, \cdots, N_D\}$).
- La demande est satisfiate si $S_n + A_n \geq D_{n+1}$, sinon elle est satisfaite dans la mesure du possible.

Écrire la chaîne de Markov contrôlée associé. *Indication :* $S_{n+1} = \max(S_n + A_n - D_{n+1}, 0)$.

Pour $k \in \{0, \cdots, N_D\}$ on note $p_k = \mathbb{P}(D_1 = k)$. Remarquons que pour $n \in \mathbb{N}$, $p_k = \mathbb{P}(D_{n+1} = k)$.

Nous allons désormais construire une matrice de transition contrôlée homogène telle que

$$\mathbb{P}(S_{n+1} = j, X_n = i, A_n = a) = \mathbb{P}((X_n + A_n - D_{n+1})_+ = j, X_n = i, A_n = a) = P(i, a, j)\mathbb{P}(X_n = i, A_n = a).$$

Soit $i \in \{1, \cdots, N_S\}$ et $a \in \{0, \cdots, N_A\}$. Le gérant du magasin ne fera jamais une commande d'approvisionnement qui va dépasser sa capacité de stockage. Ainsi, on définit

$$P(i, a, j) = 0 \quad i + a > N_S, \quad j \in \{0, \cdots, N_S\}.$$

Supponsons maintenant que $0 \leq i + a \leq N_S$. On suppose que $j \in \{1, \cdots, N_S\}$. Si $j > i + a$, quelque soit la demande, il n'est pas possible qu'à la fin de la période de vente, le stock soit plus élevé qu'au début. Ainsi

$$P(i, a, j) = 0 \quad i + a \leq N_S, \quad j > i + a.$$

Supposons désormais que $0 \leq j \leq i + a \leq N_S$. Si $j < i + a - N_D$, il n'est pas possible que les stocks du magasin aient autant baissé, et dans ce cas

$$P(i, a, j) = 0 \quad i + a \leq N_S, \quad j < i + a - N_D.$$

On suppose désormais que $i + a - N_D \leq j \leq i + a$ et que $i + a - N_D \geq 0$. Alors, il n'existe qu'une possilité pour que le stock passe de $i + a$ à j, c'est que la demande soit exactement égale à $i + a - j$. On pose alors

$$P(i, a, j) = p_{i+a-j}.$$

De la même façon si $i + a - N_D < 0$ et $j \in \{1, \cdots, i + a\}$,

$$P(i, a, j) = p_{i+a-j}.$$

Enfin, si $i + a - N_D < 0$, pour que le stock passe à zéro, il faut que la demande ai été plus grande que $i + a$. On pose

$$P(i, a, 0) = \sum_{k=i+a}^{N_D} p_k.$$

On remarque que l'on a construit P telle que pour tout $i, j \in \{0, \cdots, N_S\}$, tout $a \in \{0, \cdots, N_A\}$ et tout $n \in \mathbb{N}$,

$$\mathbb{P}(S_{n+1} = j, X_n = i, A_n = a) = P(i, a, j)\mathbb{P}(X_n = i, A_n = a).$$

Exercice 12 (Bandit multi-bras). Un joueur observe les résultats de K machines à sous différentes. A chaque étape, il en choisit une et une seule et la joue, les $K - 1$ autres ne sont pas jouées.

A chaque partie n, on consigne les états des machines à sous, ainsi, un seul de ces états parmis les $K - 1$ est actualisé.

On suppose que la machine $i \in \{1, \cdots, K\}$ peut être dans les états S^i. La machine i passe d'un état $s_1 \in S^i$ à un état $s_2 \in S^i$ avec probabilité $p(i, s_1, s_2)$.

Écrire une matrice de transition contrôlée avec $S^1 \times \cdots \times S^K$ comme espace d'états et $\{1, \cdots, K\}$ comme espace d'actions qui modélise le problème précédent.

La chaîne de Markov contrôlée va décrire la dynamique suivante : à un instant $n \in \mathbb{N}$ les machines sont à l'état $s = (s^1, \cdots, s^K) \in \mathcal{S} := S^1 \times \cdots \times S^K$. Le joueur observe les machine, choisi la machine $i \in \mathcal{A} = \{1, \cdots, K\}$. L'état de cette dernière est actualisé aléatoirement. Soit $i \in \{1, \cdots, K\}$, $s_1^i, s_2^i \in S^i$ et $s^1, \cdots, s^{i-1}, s^{i+1}, \cdots, s^K \in S^1 \times \cdots \times S^{i-1} \times S^{i+1} \times \cdots \times S^K$, et soit $s_1 = (s^1, \cdots, s^{i-1}, s_1^i, s^{i+1}, \cdots, s^K)$ et $s_2 = (s^1, \cdots, s^{i-1}, s_2^i, s^{i+1}, \cdots, s^K)$. On définit

$$P(s_1, i, s_2) = p(i, s_1^i, s_2^i).$$

Si il existe deux indices $j \neq k \in \{1, \cdots, K\}$ telles que $s, \tilde{s} \in \mathcal{S}$ et $s^k \neq \tilde{s}^k$ et $s^j \neq \tilde{s}^j$, on pose

$$P(s, i, \tilde{s}) = 0$$

pour tout $i \in \{1, \cdots, K\}$.

On remarque que P est une matrice de transition contrôlée homogène.

Exercice 13 (Selection d'un film sur Netflix). Un abonné sélectionne un film sur Netflix avec la règle suivante : il visualise les suggestions de façon séquentielle et ne peut choisir que le film dont il est en train de regarder la notice (pas de retour en arrière possible), c'est à dire que soit il choisit le film dont il est en train de regarder la notice, et alors la sélection s'arrête, soit il regarde la notice suivante, mais ne pourra pas revenir en arrière.

On note $\mathcal{S} = \{0, 1\}$ l'espace d'état. L'état est 1 si le film consulté est le meilleurs choix (au regard des goûts de l'abonné) parmi tous ceux consultés jusqu'alors. L'état est 0 si ce n'est pas le cas.

On note $\mathcal{A} = \{0, 1\}$ l'espace d'actions possible pour l'abonné : 1 si on choisit le film dont on est en train de consulter la notice, 0 sinon.

Montrer qu'on obtient une chaîne de Markov contrôlée de matrices de transition

$$P(n, i, a, 1) = \frac{1}{n+1} = 1 - P(n, i, a, 0), \quad i \in \{0, 1\}, a \in \{0, 1\}.$$