

# Préparation à l'Agrégation Externe Université de Lorraine

## Méthodes

Clément Dell'Aiera

Quelques adresses utiles :

- le site de l'agrégation de mathématiques <http://agreg.org>, vous y trouverez des textes pour vous entraîner, et surtout les comptes rendus du jury. Aussi, la liste des logiciels acceptés à l'agreg : **Python, Scilab, Octave, Sage, Maxima, Xcas, R**. Tous sont libres et gratuits.
- la page de la préparation de Rennes <https://perso.univ-rennes1.fr/florent.malrieu/agreg-probas.html>

## 1 Théorèmes limites dans le cas IID

Les deux théorèmes suivant sont à maîtriser impérativement.

**Théorème 1** (LGN et TCL). Soient  $X_1, \dots, X_n$  des variables aléatoires indépendantes identiquement distribuées, centrées et réduites, on note  $\hat{\theta}_n = \frac{1}{n} \sum X_j$ . Alors

$$\begin{aligned}\hat{\theta}_n &\rightarrow_{ps} 0, \\ \sqrt{n}\hat{\theta}_n &\rightarrow_{\mathcal{L}} 0.\end{aligned}$$

Pour l'illustrer avec Scilab, on peut :

- tirer aléatoirement des variables aléatoires iid
- calculer la moyenne empirique et montrer qu'elle converge vers l'espérance
- répéter l'opération pour obtenir un échantillon de moyennes empiriques
- tracer la fonction de répartition empirique, et montrer qu'elle converge en norme sup vers la fonction de répartition d'une normale.
- montrer que cela n'arrive pas avec certaines lois bien choisies. Par exemple la loi de Cauchy.

Un bon contre-exemple est la loi de Cauchy, de densité

$$P(dx) = \frac{1}{\pi} \frac{1}{1+x^2} dx.$$

On voit qu'elle ne satisfait pas les hypothèses : elle n'a pas de moment d'ordre 1, puisqu'équivalente à  $\frac{1}{\pi x^2}$  en  $+\infty$ .

Pour la simuler, on peut utiliser le lemme suivant.

**Lemme 1.** Soit  $X$  une variable aléatoire de fonction de répartition  $F$  et  $U$  suivant une loi uniforme sur l'intervalle  $[0, 1]$ . Alors  $F^{-1}(U)$  et  $X$  ont mêmes lois.

**Preuve 1.** On rappelle que les fonctions de répartition sont caractérisées par les propriétés suivantes :

- $F$  est càd-làg (continue à droite limite à gauche),
- $F$  est croissante,
- $\lim_{-\infty} F = 0$  et  $\lim_{\infty} F = 1$ .

$F$  n'admet pas forcément d'inverse, mais une inverse généralisée définie par

$$F^{-1}(y) = \inf\{x : F(x) \leq y\}.$$

Alors :

$$P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x).$$

□

Pour, la loi de Cauchy, un simple calcul donne  $F(x) = \frac{1}{\pi} \arctan(x) + \frac{1}{2}$ , et  $F^{-1}(u) = \tan(\pi(y - \frac{1}{2}))$ .

Le code suivant illustre cela, avec une fonction courte pour afficher des fonctions de répartition empiriques.

```
N=500
m=10

Loi normale, exponentielle, et de Cauchy
```

```
Z=grand(1,N,'nor',0,1)
W= grand(1,N,'exp',2)
U=rand(1,N)
Cauchy=30*tan(% pi*(U-0.5))
```

### **Loi des grands nombres**

```
function esperance=esperance(X)
p,q
=size(X)
scf()
plot2d([1 :q],cumsum(X)./[1 :q],2)
xlabel("Loi des grands nombres" )
endfunction
```

```
esperance(Z)
plot2d([0 :N],0*[0 :N])
```

### **FDR empirique**

```
// Theoreme Central Limite
function plotFDR=plotFDR(X)
[p,q]=size(X) ;
scf()
plot2d(-sort(-X),(1/q)*[1 :q],2)
xlabel("Fonction de repartition empirique")
endfunction
```

```
h=0.1 ;
t=[-5 :h :5] ;
```

```
plotFDR(Z)
ynor=cdfnor("PQ",t,0*t,0*t+1)
plot2d(t,ynor)
```

```
plotFDR(W)
t=[0 :h :5] ;
ypoi=cdfpoi("PQ",t,0*t+2)
plot2d(t,ypoi)
```

```
plotFDR(Cauchy)
```

## 2 Chaînes de Markov

On rappelle qu'une chaîne de Markov est l'analogue probabiliste d'une suite récurrente d'ordre 1. C'est la donnée d'un couple  $(E, P)$  où  $E$  est un espace d'état, et  $P : E \times E \rightarrow [0, 1]$  un noyau de transition.

Vous devez connaître ces définitions, ainsi que

- la définition de la propriété de Markov
- l'équation de Chapman-Kolmogorov
- la définition de mesure invariante  $\nu P = \nu$ , de mesure réversible  $\nu(x)P(x, y) = \nu(y)P(y, x)$ , et savoir montrer que réversible entraîne invariante.

Supposons que l'espace d'états  $E$  est fini. Alors l'espace des probabilités invariantes est un compact convexe non-vidé du simplexe des probabilités sur  $E$ , vue comme un sous-espace de  $\mathbb{R}^{|E|}$ . En particulier, il existe une mesure invariante. Pour l'unicité, il suffit que la chaîne soit irréductible, i.e. que le graphe de transition soit connexe.

**Théorème 2** (Théorème ergodique). Si  $(X_n)$  une trajectoire d'une chaîne de Markov irréductible, d'unique probabilité invariante  $\nu$ . Alors pour toute fonction  $\phi : E \rightarrow \mathbb{R}$ ,

$$\frac{1}{n} \sum_{k=1}^n \phi(X_k) \rightarrow \int_E \phi(x) \nu(dx).$$

Alors

$$\nu(x) = \frac{1}{E[T_x]},$$

où  $T_x = \inf\{n > 1 : X_n = x\}$  est le temps de retour en  $x$ .

De plus, si la chaîne est fortement apériodique, alors on a, pour toute mesure de probabilité initiale  $\mu_0$ , la convergence en loi suivante :

$$\mu_0 P^n \rightarrow \nu,$$

i.e.  $(X_n)_n$  converge en loi vers  $\nu$  quelle que soit sa distribution initiale.

Vous devez penser à la mesure invariante comme la mesure que le système atteint à l'équilibre. Par exemple, dans le modèle de l'urne d'Ehrenfest, on avait  $E = \{0, 1, 2, \dots, m\}$ ,

$$P(k, k-1) = 1 - P(k, k+1) = \frac{k}{m} \quad \text{et} \quad P(0, 1) = P(m, m-1) = 1.$$

Ce modèle est sensé représenter  $m$  particules indistingables dans une pièce, que l'on sépare en deux zones distinctes  $A$  et  $B$ . L'état du système est entièrement déterminé par le nombre  $X_T$  de particules qui se trouvent à l'instant  $T$  dans la zone  $A$  par exemple.

La chaîne est irréductible, mais admet une période : la parité des états visités alterne en fonction du temps. On ne peut donc qu'appliquer la convergence presque-sûre, et pas la convergence en loi, dans le théorème ergodique.

Un simple calcul montre que  $\nu(x) = \frac{1}{2^m} \binom{x}{m}$  est une mesure réversible, donc invariante, qui est unique par irréductibilité. Cela signifie que si l'expérimentateur lance son expérience, quitte la pièce assez longtemps pour que le système atteigne l'équilibre, et qu'il fait une mesure du nombre de particules dans la zone  $A$  lors de son retour, il trouvera  $x$  avec une probabilité proche de  $\nu(x)$  (d'autant plus proche que le temps passe).

Un point inquiétant du modèle est qu'il autorise un retour à l'état 0 ou  $m$ , ce qui signifie que toutes les particules de la pièce sont concentrées dans une partie de la pièce, laissant l'autre vide. Si le jury soulève cette question, vous pouvez alors utiliser le théorème ergodique pour arguer que le temps de retour :

- en  $m$  est de l'ordre de

$$E[T_m] = \frac{1}{\nu(m)} = \frac{1}{2^m}$$

- en  $\frac{m}{2}$  est de l'ordre de

$$E[T_{m/2}] = \frac{1}{\nu(m/2)} \sim \sqrt{\frac{\pi m}{2}}$$

en utilisant la formule de Stirling.

Le nombre  $m$  étant supposé très grand, il faudrait à l'expérimentateur attendre très longtemps pour voir se réaliser une situation pareille. Enfin, si on note  $Z$  une normale centrée réduite :

$$\nu\left(\left[\frac{m}{2} \pm 2\epsilon\sqrt{m}\right]\right) \simeq P(|Z| > \epsilon) = 1 - 2P(Z > \epsilon)$$

et

$$P(Z > \epsilon) = P(e^{tZ} > e^{t\epsilon}) \leq \frac{E[e^{tZ}]}{(t\epsilon)^2}.$$

Ce calcul montre que toute la masse de la mesure invariante est concentrée autour de l'état  $m/2$ , i.e. qu'à l'équilibre, un expérimentateur ne peut mesurer qu'un nombre de particules pratiquement égal dans les deux zones (avec une très forte probabilité).

Le calcul utilise successivement :

- l'approximation gaussienne de la binomiale,
- la symétrie de la loi normale,
- l'inégalité de Markov.

Voici pour la simulation. Essentiellement, on entre la matrice de transition, et on utilise l'option *markov* de *grand*.

```

//Entree de la matrice de transition de la chaine de Markov associee
au modele d Ehrenfest

function P=ehrenfest(N)
d1=(N-(0 :N-1))/N;
d2=(1 :N)/N;
P=diag(d1,1)+diag(d2,-1);
endfunction;

// Simulation
m=1000
T=100000
x0=m
X=grand(T,'markov',ehrenfest(m),x0)
plot([1 :T],X)

```

On peut facilement voir que le système atteint rapidement l'état d'équilibre sur la figure ???. C'est un fait général : la vitesse de convergence est dominée par le plus grand module  $< 1$  des valeurs propres de la matrice de transition  $P$ . Cela est donné par le théorème de Perron-Frobenius.

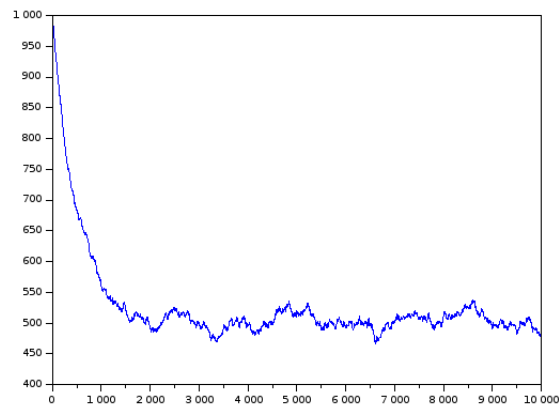


FIGURE 1 – Une trajectoire partant de l'état  $m = 1000$  particules dans la zone  $A$

### 3 Réduction des endomorphismes

Ce numéro présente la réduction des endomorphisme vue sous l'angle des  $k[X]$ -modules. Pour cela, nous nous inspirons fortement du livre de R. Mneimé, *Réduction des endomorphismes*. On rappelle que  $f, g \in \mathcal{L}(E)$  sont équivalents ssi il existe  $T \in GL(E)$  tel que  $Tf = gT$ .

Soit  $k$  un corps,  $E$  un  $k$ -espace vectoriel et  $f \in \mathcal{L}(E)$  un endomorphisme de  $E$ . La loi

$$\forall P \in k[X], \forall v \in E, P.v := P(f)v$$

définit une structure de  $k[X]$ -module sur  $E$ . On notera  $E_f$  le  $k[X]$ -module obtenu. Une remarque importante : soient  $f, g \in \mathcal{L}(E)$ , alors une endomorphisme  $T \in \mathcal{L}(E)$  induit un morphisme de  $k[X]$ -module  $\tilde{T} \in Hom_{k[X]}(E_f, E_g)$  ssi  $Tf = gT$ . On a alors facilement la proposition suivante :

**Proposition 1.** Soient  $f, g \in \mathcal{L}(E)$ . Alors  $E_f$  et  $E_g$  sont isomorphes ssi  $f$  et  $g$  sont équivalents.

Rappelons que  $k[X]$  est un anneau euclidien. L'algorithme du pivot de Gauss assure donc que pour toute matrice  $A \in M_n(k[X])$ , il existe  $Q_0, Q_1 \in GL_n(R)$  et des polynômes  $P_1, \dots, P_r \in k[X]$  vérifiant  $P_j | P_{j+1}$  tels que

$$Q_0 A Q_1 = \begin{pmatrix} P_1 & 0 & & & \\ 0 & \dots & 0 & & \\ & 0 & P_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}$$

avec  $Q_j$  qui sont produit de transvections. Lorsque  $A = \text{mat}(f) - X \cdot \text{id}_E$  la famille de polynômes  $P_1, \dots, P_r$  sont appelés les facteurs invariants de  $f$ .

Dans un tel anneau, l'identité  $A \cdot \text{Com}(A)^T = \text{Com}(A)^T \cdot A = \det(A) I_n$  a toujours un sens, et fournit une preuve simple du théorème de Cayley-Hamilton. En effet, appliquée à la matrice  $A - X I_n \in M_n(k[X])$ , on obtient

$$\chi_A(X) I_n = \text{Com}(A - X I_n)^T \cdot (A - X I_n)$$

mais dans  $E_A$ , on a par définition  $X.e_i = A.e_i = \sum a_{ij}e_j$  donc  $\chi_A(X).e_i = 0$  pour tout  $i$ , i.e.  $\chi_A(A) = 0$ .