ISFA - L3 Actuariat
Probabilités Avancées
Semestre printemps 2018-2019

Feuille de TD n° 5:

LOI FORTE DES GRANDS NOMBRES

math.univ-lyon1.fr/homes-www/lerouvillois
lerouvillois@math.univ-lyon1.fr
poudevigne@math.univ-lyon1.fr

Exercice 1 : Méthode de Monte-Carlo

Soit f une fonction dans $L^1([0,1]^d)$, et $(U_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. de loi uniforme sur $[0,1]^d$.

1. Montrer que

$$\frac{1}{n}\sum_{i=1}^{n} f(U_i) \xrightarrow[n\to\infty]{p.s} \int_{[0,1]^d} f(x_1,...,x_d)dx_1...dx_d.$$

2. Application : approximation du nombre π.
Soit $(U_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. de loi uniforme sur $[0,1]^2$. On note $U_n = (X_n, Y_n)$. Montrer que

$$\frac{4}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i^2+Y_i^2\leq 1} \xrightarrow[n\to\infty]{p.s} \pi.$$

Exercice 2 : Convergence de la fonction de répartition empirique

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. On note F la fonction de répartition de la loi des X_n. Pour tout $x \in \mathbb{R}$, on pose

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i\leq x}.$$

Montrer que pour tout $x \in \mathbb{R}$,

$$F_n(x) \xrightarrow[n\to\infty]{p.s} F(x).$$

Remarque : Le théorème de Glivenko-Cantelli donne un résultat plus fort : presque sûrement, F_n converge uniformément (en x) vers F.

Exercice 3 : Estimation de paramètres par la méthode des moments

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles i.i.d. On note $\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ la moyenne empirique

1. Montrer que si les X_i ont un moment d'ordre k, alors

$$\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow[n\to\infty]{p.s} \mathbb{E}[X_1^k].$$

2. En déduire que si les X_i sont de carré intégrable, alors la variance empirique

$$V_n := \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

converge presque sûrement vers une limite que l'on précisera.

3. Application :
   (a) On suppose que les X_i sont de loi $\mathcal{N}(\mu,\sigma)$ où μ et σ sont des paramètres inconnus que l'on cherche à estimer. Montrer que

$$(\bar{X}_n, V_n) \xrightarrow[n\to\infty]{p.s} (\mu, \sigma^2).$$

   On dit que ces estimateurs sont consistants.

   (b) On suppose maintenant que les X_i sont de loi uniforme sur $[a,b]$ où a et b sont des paramètres inconnus que l'on chercher à estimer. Monter que

$$(\bar{X}_n - \sqrt{3V_n}, \bar{X}_n + \sqrt{3V_n}) \xrightarrow{p.s} (a,b).$$

Exercice 4 : Loi des grands nombres pour des produits

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendants de même loi et de carré intégrable. Étudier la convergence presque sûre de la suite :

$$S_n = \frac{\sum_{i=1}^{n} X_i X_{i+1}}{n}.$$

Indication : on pourra séparer la somme en deux selon la parité de i

Exercice 5 : Contre-exemple en l'absence de moment d'ordre 1

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a i.i.d suivant une loi de Cauchy centrée de paramètre 1, i.e de densité $f(x) = \dfrac{1}{\pi(1+x^2)}$. On note $\bar{X}_n$ la moyenne empirique.

1. Les hypothèses de la loi forte des grands nombres sont-elles satisfaites ?
2. On admet que la fonction caractéristique d'une loi de Cauchy centrée de paramètre 1 vaut $\phi(\xi) = e^{-|\xi|}$. En déduire la loi de $\bar{X}_n$ pour tout n.
3. Conclure que la suite des $\bar{X}_n$ ne converge ni en probabilité ni presque sûrement vers 0 bien que les X_i soient indépendants et centrés.

Rappel :

- Loi des grands nombres faibles au sens CV en $\mathbb{P}$

Th: Soit $(X_n)_{n\in\mathbb{N}}$, une suite de v.a i.i.d et dans L^1

alors, $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{\mathbb{P}} E[X_1]$

- Loi grands nombres au sens CV p.s

Sous les mêmes hypothèses que le théorème précedent

on a $\bar{X}_n \xrightarrow[n\to\infty]{p.s} E[X_1]$

Exercice 1 : Méthode de Monte-Carlo

Soit f une fonction dans $L^1([0,1]^d)$, et $(U_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. de loi uniforme sur $[0,1]^d$.

1. Montrer que
$$\frac{1}{n}\sum_{i=1}^{n} f(U_i) \xrightarrow[n\to\infty]{p.s} \int_{[0,1]^d} f(x_1,...,x_d)dx_1...dx_d.$$
2. Application : approximation du nombre π.
Soit $(U_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. de loi uniforme sur $[0,1]^2$. On note $U_n = (X_n, Y_n)$. Montrer que
$$\frac{4}{n}\sum_{i=1}^{n} 1_{X_i^2+Y_i^2\leq 1} \xrightarrow[n\to\infty]{p.s} \pi.$$

Ex 1

$f \in L^1([0,1]^d)$, $(U_n)_{n\in\mathbb{N}}$ v.a i.i.d tq $U_n \sim \mathcal{U}([0,1]^d)$

On applique la LGN forte aux variables $X_i = f(U_i) \in L^1$

On vérifie que les X_i sont i.i.d

finalement: $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{L^2} E[X_1]$

$= E[f(U_1)]$

$= \int_{[0,1]^d} f(x_1,\dots,x_d)\, d\lambda_d(x_1,\dots,x_d)$

2) $U_n = (X_n, Y_n) \in [0,1]^2$

But: $\frac{4}{n}\sum_{i=1}^{n} \mathbb{1}_{X_i^2+Y_i^2\leq 1} \xrightarrow[n\to\infty]{p.s} \pi$

On applique question 1 avec la fonction $f : (x,y) \to 4\,\mathbb{1}_{x^2+y^2\leq 1}$

On obtient $\frac{4}{n} \sum_{i=1}^{n} \mathbb{1}_{x^2+y^2 \leq 1} \xrightarrow[n \to \infty]{p.s} 4 \int_{[0,1]^2} \mathbb{1}_{x^2+y^2 \leq 1} \, d(x,y)$

$x = r\cos(\theta)$

$y = r\sin(\theta)$

$(x,y) \in [0,1]^2 \Longleftrightarrow \begin{cases} 0 \leq r \leq 1 \\ \theta \in [0, \pi/2] \end{cases}$

Changement de variable

$$4 \int_{[0,1]^2} \mathbb{1}_{x^2+y^2 \leq 1} \, d(x,y) = 4 \int_0^{\pi/2} \int_0^1 \mathbb{1}_{r^2 \leq 1} \, r \, dr \, d\theta$$

$$= 4 \frac{\pi}{2} \int_0^1 r \, dr = \frac{4\pi}{2} \times \frac{1}{2} = \pi$$

Exercice 2 : **Convergence de la fonction de répartition empirique**

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. On note F la fonction de répartition de la loi des X_n. Pour tout $x \in \mathbb{R}$, on pose

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i \leq x}.$$

Montrer que pour tout $x \in \mathbb{R}$,

$$F_n(x) \xrightarrow[n\to\infty]{p.s} F(x).$$

Remarque : Le théorème de Glivenko-Cantelli donne un résultat plus fort : presque sûrement, F_n converge uniformément (en x) vers F.

$(X_n)_{n\geqslant 1}$ suite de v.a i.i.d

But : $\forall x \in \mathbb{R}$, $F_n(x) \xrightarrow[n\infty]{p.s} F_{X_1}(x) = \mathbb{P}(X_1 \leq x)$

Rq : Comme les X_n sont i.i.d, elles ont aussi la même fonction de répartition

où $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{X_i \leq x}$

Rappel : LGN (forte) : $(X_n)_{n\geqslant 1}$ v.a i.i.d

Si $E[|X_1|] < \infty$ alors

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\infty]{p.s} E[X_1]$$

Soit $x \in \mathbb{R}$. On pose

$$f(y) = \mathbb{1}_{y \leq x}$$

et on s'interesse aux v.a $Y_i = f(X_i) = \mathbb{1}_{X_i \leq x}$

Comme les X_i sont independants, les $Y_i = f(X_i)$ le sont aussi.

Rappel : • Si $X_1, \dots, X_n$ sont indep et f_i des fonctions mesurables,

Alors, les $f_i(X_i)$ sont aussi independantes

ex : Si X et Y indep, X^2 et Y^3 le sont aussi.

• Si $X_1, \dots, X_n$ independantes.

Si $i = i_1 < i_2 < i_3 \dots < i_p \dots < n$

$p \in \mathbb{N}^*$ fixé

alors les v.a $Y_1 = (X_{i_1}, \dots, X_{i_2 - 1})$

$Y_2 = (X_{i_2}, \dots, X_{i_3 - 1}), \dots, Y_p (X_{i_p}, \dots, X_n)$

sont indep

$i_1 = 1 < i_2 = 5 \Rightarrow Y_1 = (X_1, X_2 \dots X_4)$

ex : X, Y, Z indep $\Rightarrow (X, Y)$ indep de Z

$\Rightarrow X^2 Y^3$ indep de Z

Les X_i ont la même loi, donc par le théoreme de transfert

On montre que les $Y_i = f(X_i)$ ont la même loi

Les $Y_i \in L^1$ car $E[Y_i] = \mathbb{P}(X_i < x) = F_{X_i}(x) < \infty$

On peut donc appliquer la Loi Forte des grands nombres

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} f(X_i) \xrightarrow[n\infty]{P.S} E[f(X_1)]$$

$$= E[\mathbb{1}_{X_1 < x}] = F_{X_1}(x)$$

$$= F(x)$$

Exercice 3 : Estimation de paramètres par la méthode des moments

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles i.i.d. On note $\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ la moyenne empirique

1. Montrer que si les X_i ont un moment d'ordre k, alors
$$\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow[n\to\infty]{p.s} \mathbb{E}[X_1^k].$$
2. En déduire que si les X_i sont de carré intégrable, alors la variance empirique
$$V_n := \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$
converge presque sûrement vers une limite que l'on précisera.

Ex 3 : $(X_n)_{n\geq 1}$, v.a i.i.d

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

1. On suppose que les $X_i \in L^k$, $k \geq 1$

$\quad\quad E[|X_i|^k] < \infty$

But : $\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow[n\infty]{P.S} E[X_1^k]$

Comme les X_i sont i.i.d, il en est de même pour le X_i^k

De plus, les $X_i^k \in L^1$ car $X_i \in L^k$.

Par la loi forte des grands nombres

$$\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow[n\infty]{P.S} E[X_1^k]$$

2\. 2. En déduire que si les X_i sont de carré intégrable, alors la variance empirique

$$V_n := \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

converge presque sûrement vers une limite que l'on précisera.

$V_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ CV p.s ?

On suppose que les $X_i \in L^2$

$V[X] = E[X^2] - E[X]^2$

$$V_n = \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2X_i\bar{X}_n + \bar{X}_n^{\,2}\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}_n \underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i}_{\bar{X}_n} + \underbrace{\frac{1}{n}\sum_{i=1}^{n}\bar{X}_n}_{n(\bar{X}_n)^2}$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}_n^{\,2} + \bar{X}_n^{\,2}$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\bar{X}_n)^2 \xrightarrow{p.s} E[X_1^2] - E[X_1]^2 = V[X_1]$$

$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p.s} E[X_1^2]$, $\quad (\bar{X}_n)^2 \xrightarrow{p.s} E[X_1]^2$

car $X \to X^2$ continue

3. Application :

(a) On suppose que les X_i sont de loi $\mathcal{N}(\mu, \sigma)$ où μ et σ sont des paramètres inconnus que l'on cherche à estimer. Montrer que

$$(\bar{X}_n, V_n) \xrightarrow[n\to\infty]{p.s} (\mu, \sigma^2).$$

On dit que ces estimateurs sont consistants.

(b) On suppose maintenant que les X_i sont de loi uniforme sur $[a, b]$ où a et b sont des paramètres inconnus que l'on chercher à estimer. Monter que

$$(\bar{X}_n - \sqrt{3V_n}, \bar{X}_n + \sqrt{3V_n}) \xrightarrow{p.s} (a, b).$$

a) $X_i \sim \mathcal{N}(\mu, \sigma^2)$ Donc les $X_i \in L^2$

$$(\bar{X}_n, V_n) \xrightarrow[n\infty]{P.S} (\mu, \sigma^2)$$

$$\begin{tikzcd}
\bar{X}_n \arrow[d, "P.S"', "LGN"] & V_n \arrow[d, "P.S"] \\
E[X_1] = \mu & V[X_1] = \sigma^2
\end{tikzcd}$$

car $X_1 \sim \mathcal{N}(\mu, \sigma^2)$

b) $X_i \sim \mathcal{U}([a, b])$

$$(\bar{X}_n - \sqrt{3V_n}, \bar{X}_n + \sqrt{3V_n}) \xrightarrow{P.S} (a, b)$$

On sait que $\bar{X}_n \xrightarrow[n\infty]{P.S} E[X_1] = \dfrac{a+b}{2}$

$$V_n \xrightarrow[n\infty]{P.S} V[X_1] = \frac{(b-a)^2}{12}$$

Comme la fonction $f : x, y \longmapsto x - \sqrt{3y}$ est continue sur $\mathbb{R} \times \mathbb{R}_+$ d'où,

$$\bar{X}_n - \sqrt{3V_n} = f(\bar{X}_n, V_n) \xrightarrow{P.S} \underbrace{E[X_1] - \sqrt{3V[X_1]}}_{\frac{a+b}{2} - \frac{(b-a)}{2} = a}$$

De même, $\bar{X}_n - \sqrt{3V_n} \longrightarrow \frac{a+b}{2} + \frac{(b-a)}{2} = b$