LOI DES GRANDS NOMBRES. THÉORÈME LIMITE CENTRAL. APPLICATIONS. 250

<u>Cadre</u> : $(X_n)_{n\geq 1}$ suite de variables aléatoires réelles, indépendantes et identiquement distribuées (v.a.r. i.i.d.) définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

<u>Notation</u> : $S_n = \sum_{k=1}^{n} X_k$

<u>Problématique</u> : quel est le comportement de S_n lorsque n tend vers $+\infty$?

I_ Loi des grands nombres :

Pour $X_n \sim \mathcal{B}(p)$, $\frac{S_n}{n}$ est la fréquence empirique des succès. On s'attend à ce que cette fréquence soit proche de p.

I.1°) Loi faible :

<u>Lemme</u> (inégalités)
Soit X v.a.r. Alors pour tout $a>0$: $\mathbb{P}(|X|\geq a) \leq \frac{\mathbb{E}(|X|)}{a}$ (Markov)

$\mathbb{P}(|X-\mathbb{E}(X)|\geq a) \leq \frac{Var(X)}{a^2}$ (Tchebychef)

<u>Thm</u> (loi faible des grands nombres)
Soit $(X_n)_n$ suite de v.a.r. i.i.d. de L^2. Alors $\frac{S_n}{n} \xrightarrow[n\to+\infty]{\mathbb{P}} \mathbb{E}(X_1)$.

Plus précisément,
$$\forall \varepsilon>0,\ \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}(X_1)\right|\geq \varepsilon\right) \leq \frac{Var(X_1)}{n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0$$
fournit une vitesse de convergence.

<u>Appl.</u> (polynômes de Bernstein)
Soit $f : [0,1] \longrightarrow \mathbb{R}$ une fonction continue.
<u>Def</u> : on appelle polynômes de Bernstein les polynômes définis par :
$$\forall n\geq 1,\ B_n(x) \overset{\text{déf.}}{=} \mathbb{E}_{(x)}\left[f\left(\frac{S_n}{n}\right)\right] = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right),$$
où $x\in[0,1]$ et les X_n suivent une loi de Bernoulli de paramètre x.

<u>Thm</u> (de Bernstein)
On a : B_n converge uniformément vers f sur $[0,1]$.
Plus précisément, $\|f - B_n\|_\infty \leq C\,\omega\left(\frac{1}{\sqrt{n}}\right)$ où $C>0$ et ω est le module de continuité de f.

<u>Cor</u> (thm de Weierstrass)
Les polynômes sont denses dans $(\mathcal{C}^0([a,b], \mathbb{R}), \|\cdot\|_\infty)$, $\forall a,b\in\mathbb{R}$ $(a<b)$

I.2°) Estimation des grands écarts, pour les Bernoulli :

Dans ce paragraphe, les X_n suivent des lois de Bernoulli $\mathcal{B}(p)$, $p\in]0,1[$.

<u>Idée</u> : on avait $\mathbb{P}\left(\left|\frac{S_n}{n} - p\right|\geq\varepsilon\right) \leq \frac{p(1-p)}{n\varepsilon^2}$ $\forall\varepsilon>0$, et on établit maintenant que cette vitesse de convergence est exponentielle.

<u>Thm</u> (estimation des grands écarts)
Posons, pour tout $\varepsilon\in]0,1-p[$, $h_+(\varepsilon) = (p+\varepsilon)\ln\frac{p+\varepsilon}{p} + (1-p-\varepsilon)\ln\frac{1-p-\varepsilon}{1-p}$.
Alors $h_+(\varepsilon)>0$ et $\forall n\geq 1$, $\mathbb{P}\left(\frac{S_n}{n}\geq p+\varepsilon\right) \leq e^{-nh_+(\varepsilon)}$.

<u>Cor.</u> : si $0<\varepsilon<p$, notons $h_-(\varepsilon) = h_+(-\varepsilon)$ qui est bien définie et >0.
Alors pour ε tq $0<\varepsilon<\min(p,1-p)$, $\mathbb{P}\left(\left|\frac{S_n}{n}-p\right|\geq\varepsilon\right) \leq e^{-nh_+(\varepsilon)} + e^{-nh_-(\varepsilon)}$.

<u>Prop</u> : On a $h_+(\varepsilon) \underset{\varepsilon\to 0}{=} \frac{\varepsilon^2}{2p(1-p)} + O(\varepsilon^3)$.

<u>Ex</u> : Pour 1000 lancers d'une pièce équilibrée, la probabilité pour que le nombre de « pile » soit supérieur ou égal à 540 est plus petite que 0,05.

I.3°) Loi forte :

<u>Idée</u> : on cherche à améliorer cette fois-ci la convergence dans la loi faible.

Lemme (de Borel-Cantelli) :

Si $(A_n)_n$ est une suite d'évènements, alors

(i) si $\sum_{n\geq 1} \mathbb{P}(A_n) < +\infty$, alors $\mathbb{P}(A_n \text{ i.s.}) := \mathbb{P}\left(\bigcap_{n\geq 1} \bigcup_{k\geq n} A_k\right) = 0$

(i) si les (A_n) sont indépendants, on a la réciproque de (i).

Cor. : En considérant $A_n = \{|X_n - X| \geq \varepsilon\}$, cela fournit un moyen de montrer convergence presque sûre de X_n vers X.

Thm (loi forte des grands nombres, cas L^1).
Soit (X_n) suite de v.a.r. i.i.d. tq $\mathbb{E}(|X_1|) < +\infty$. Alors

$$\frac{S_n}{n} \xrightarrow[n\to+\infty]{p.s.} \mathbb{E}(X_1).$$

Rmq : dans le cas Bernoulli,

- ce thm est une application du I.2°) et du lemme de Borel Cantelli
- on peut affiner le résultat $\frac{S_n}{n} - p \longrightarrow 0$ par l'estimation suivante :

Prop : $S_n - np = O(\sqrt{n \ln n})$ (estimation de Hardy-Littlewood)

Thm (loi forte des grands nombres, cas uniformément bornés)
Soit (X_n) suite de v.a.r. indépendantes centrées tq $\sup_{n,\omega} |X_n(\omega)| < +\infty$.
Alors $\frac{S_n}{n} \xrightarrow[n\to+\infty]{p.s.} 0$

Appl (nombre normaux)
Tout nombre réel $x \in [0,1]$ s'écrit de façon unique $x = \sum_{i=1}^{+\infty} \frac{x_i}{2^i}$ où $x_i \in \{0,1\}$ et la suite (x_i) n'est pas stationnaire à 1 à partir d'un certain rang (développement diadique propre).

Déf. : $x \in [0,1]$ est normal en base 2 si tout bloc de chiffre apparaît avec une fréquence asymptotique de $1/2^\ell$, avec ℓ la longueur de ce bloc.

Prop : $x \in [0,1]$ est normal en base 2 ssi $\forall \ell > 0$, $\forall a \in \{0,1\}^\ell$,

$$\frac{1}{n} \operatorname{card}\{i \mid 1 \leq i \leq n,\ (x_i, x_{i+1}, \dots, x_{i+\ell-1}) = a\} \xrightarrow[n\to+\infty]{} 1/2^\ell.$$

Thm : presque tout nombre $x \in [0,1]$ est normal en base 2.

Appl (méthode de Monte-Carlo)
On cherche à faire un calcul approché d'intégrale ; par exemple $I = \int_{[0,1]^d} f(t)\,dt$ où f intégrable à valeurs réelles, $d \in \mathbb{N}^*$.

Prop : soit $(U_n)_{n\geq 1}$ suite de vecteurs de v.a.r. i.i.d. $\mathcal{U}_{[0,1]^d}$. Alors

$$Z_n \overset{\text{déf}}{=} \frac{f(U_1) + \dots + f(U_n)}{n} \xrightarrow[n\to+\infty]{p.s.} \mathbb{E}(U_1) = I.$$

L'intérêt de cette méthode (par rapport aux méthodes déterministes) est qu'elle ne nécessite pas de régularité sur f et qu'elle ne dépend pas de la dimension (une méthode des rectangles avec 2 points dans chaque direction de $[0,1]^d$ nécessite 2^d évaluations de f ... ie $\simeq 10^{30}$ pour $d = 100$).

$\hookrightarrow$ on aurait envie de préciser la vitesse de convergence de cette méthode.

II. THÉORÈME LIMITE CENTRAL :

II 1°) Énoncés et intervalles de confiance :

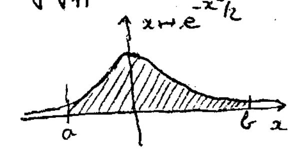
Thm (cas de De Moivre / Laplace) $p \in]0,1[$.
Soit $X_n \sim \mathcal{B}(p)$. Alors S_n, qui est de loi $\mathcal{B}(n,p)$, vérifie :

$$\frac{1}{\sqrt{n}} \left(\frac{S_n - np}{\sqrt{p(1-p)}} \right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

Appl : pour n grand, $\mathcal{B}(n,p)$ peut être approchée par $\mathcal{N}(np,\ n\,p(1-p))$.

Thm (cas général)
Soit $(X_n)_n$ suite de v.a.r. i.i.d, de L^2. Alors $\frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$,
(où $m = \mathbb{E}(X_1)$, $\sigma^2 = \operatorname{Var}(X_1)$). Autrement dit, $\forall a < b$

$$\mathbb{P}\left(a \leq \frac{S_n - nm}{\sigma\sqrt{n}} \leq b\right) \xrightarrow[n\to+\infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\,dx.$$



Ex : Pour 1000 lancers d'une pièce équilibrée, la probabilité pour que le nombre de « piles » soit ≥ 540 est environ 0,006.

Appl : (comme marquée de l'exponentielle).

Prop : $e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} \xrightarrow[n\to+\infty]{} \frac{1}{2}$, via $X_n \sim \mathcal{P}(1)$ où $\mathbb{P}_\mathcal{P} = \sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} \delta_k$, avec $\lambda = 1$

Le théorème limite central (TCL) permet de donner des intervalles de confiance ; notons $\overline{X}_n = \frac{S_n}{n}$.

* Si on cherche à localiser la moyenne m de la loi d'un échantillon $(X_1, \dots, X_n)$: on établit avec probabilité $1-\alpha$ que m se situe dans l'intervalle

$$I(m) = \left[\overline{X}_n - a\frac{\sigma}{\sqrt{n}},\ \overline{X}_n + a\frac{\sigma}{\sqrt{n}}\right]$$ où a est tq $\mathbb{P}(|Z| \leq a) = 1 - \alpha$ avec $Z \sim \mathcal{N}(0,1)$.

Souvent, on prend $\alpha = 0,05$, ce qui donne $a = 1,96$.

* On cherche à localiser la proportion p de succès dans un échantillon $\mathcal{B}(p)$:
  ↳ Tchebycheff donne $I(p) = \left[\bar{X}_n - a\,;\, \bar{X}_n + a\right]$ ($\alpha = \frac{1}{4na^2}$ convient)
  ↳ le TCL est plus précis :
  $$J(p) = \left[\bar{X}_n - \frac{a}{2\sqrt{n}}, \bar{X}_n + \frac{a}{2\sqrt{n}}\right]$$ et comme $\bar{X}_n(1-\bar{X}_n) \to p(1-p)$, on a :
  $$K(p) = \left[\bar{X}_n - a\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}, \bar{X}_n + a\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}\right]$$

Appl (intervalle de confiance pour la méthode de Monte Carlo)
À 95%, on sait que $I \in \left[I_n - \frac{1{,}96\sigma}{\sqrt{n}}, I_n + \frac{1{,}96\sigma}{\sqrt{n}}\right]$, soit encore $|I - I_n| \leq \frac{2\sigma}{\sqrt{n}}$

II. 2°) Vecteurs gaussiens et test du χ^2 :

Déf : Soit X un vecteur colonne aléatoire de $\mathbb{R}^d$ dont les composantes sont de carré intégrable. Le vecteur moyenne de X est donné par $\mathbb{E}(X) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_d))^T$ et la matrice de covariance est
$$Var(X) = (cov(X_i, X_j))_{i,j} \text{ notée } \Gamma,$$
où $cov(X_i, X_j) = \mathbb{E}\left[(X_i - \mathbb{E}(X_i))(X_j - \mathbb{E}(X_j))\right]$.

Déf (vecteur gaussien)
Soit X un vecteur aléatoire de $\mathbb{R}^d$. X est dit vecteur gaussien si toute combinaison linéaire de ses composantes suit une loi normale sur $\mathbb{R}$.

Ex : $\mathcal{N}(0, I_d)$ est appelée loi gaussienne standard sur $\mathbb{R}^d$. Si $Z_1, \dots, Z_d$ sont des v.a.r. i.i.d. de loi $\mathcal{N}(0,1)$, alors $Z = (Z_1, \dots, Z_d)^T \sim \mathcal{N}(0, I_d)$.

Rmq : La loi d'un vecteur gaussien est caractérisée par m et Γ ; on la note $\mathcal{N}(m, \Gamma)$.
On a : $\Phi_X(u) \overset{def}{=} \mathbb{E}\left[\exp(i\langle u, X\rangle)\right] = \exp\left[i\langle u, m\rangle - \frac{1}{2}\langle u, \Gamma u\rangle\right]$, $\forall u \in \mathbb{R}^d$.

Thm (TCL multivarié)
Soit $(X_n)_n$ une suite de vecteurs aléatoires de $\mathbb{R}^d$, indépendants et de même loi, dont les composantes sont de carré intégrable. Alors, on a :
$$\sqrt{n}\left(\frac{S_n}{n} - m\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \Gamma).$$

Déf (loi du χ^2)
La loi du chi-deux $\chi^2(n)$ à n degrés de liberté est la loi de $X_1^2 + \dots + X_n^2$ où les X_i sont des v.a. indépendantes de loi $\mathcal{N}(0,1)$.

Déf : (distance du chi-deux)
La distance du χ^2 entre deux lois de probabilité p et q sur un ensemble fini $\{1, \dots, k\}$ est $D(p,q) = \sum \frac{(p_i - q_i)^2}{p_i}$.

Rmq : ce n'est pas une « vraie distance ».

Thm (de Pearson) :
Soit X une v.a. continue prenant ses valeurs dans k classes $I_1, \dots, I_k$ de $\mathbb{R}$. X a pour loi $p = \{p_1, \dots, p_k\}$ avec $p_i > 0$, $\forall i \in \{1, \dots, k\}$.
Notons $\hat{p} = (\hat{p}_1, \dots, \hat{p}_k)$ la répartition empirique et $D_n = n\,D(p, \hat{p})$. Alors
$$D_n \xrightarrow[n\to+\infty]{\mathcal{L}} \chi^2(k-1).$$

Appl (test du χ^2)
Soit X comme dans le théorème de Pearson, sauf que l'on ne connaît pas sa loi. On dispose d'un échantillon $(X_1, \dots, X_n)$ de même loi que X. On souhaite tester l'hypothèse $\mathcal{H}_0$: « X a pour loi p », $p = (p_1, \dots, p_k)$ contre $\mathcal{H}_1$: « X n'a pas pour loi p » ;
alors : sous $\mathcal{H}_0$, $D_n \xrightarrow[n\to+\infty]{\mathcal{L}} \chi^2(k-1)$ et sous $\mathcal{H}_1$, $D_n \xrightarrow[n\to+\infty]{p.s.} +\infty$.

Références :
* Barbe, Ledoux
* Lesigne (Pile ou Face. Une introduction aux théorèmes limites du Calcul des Probabilités)
* Cottrell &co (exercices de probabilités)
* Zuily, Queffélec
* Bercu, Chafaï