Rappeler les modes de convergence, finir la section 2.3.

Théorèmes limites

Dans tout le chapitre, on se place sur un espace de proba $(\Omega, \mathcal{F}, \mathbb{P})$ sur lequel est définit une suite de v.a indépendantes et identiquement distribuées (i.i.d.)

1. Loi des grands nombres (LGN)

1.1 Loi faible des grands nombres

<u>Théorème</u> : Soit (X_n) une suite de va iid intégrables, alors

$$\frac{S_n}{n} := \frac{1}{n} \sum_1^n X_i \xrightarrow[n\infty]{\mathbb{P}} \mathbb{E}[X_1].$$

<u>Preuve</u> : on suppose d'abord que $X_i \in L^2$. Soit $\varepsilon > 0$ alors comme $\mathbb{E}\left(\frac{S_n}{n}\right) = \mathbb{E}[X_1]$ on a

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}[X_1]\right| > \varepsilon\right) \underset{\text{Tchebychev}}{\leq} \frac{\mathrm{var}\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \frac{1}{n^2\varepsilon^2}\,\mathrm{var}(S_n)$$

$$= \frac{1}{n\varepsilon^2}\,\mathrm{var}(X_1) \xrightarrow[n\infty]{} 0.$$

Dans le cas où $X_i \in \mathbb{L}^1$ on utilise le lemme suivant

<u>Lemme</u> : Si $(a_i),(b_i) \in \mathbb{C}^{\mathbb{N}}$ avec $|a_i| \leq 1$, $|b_i| \leq 1$

alors
$$\left| \prod_1^n a_i - \prod_1^n b_i \right| \leq \sum_1^n |a_i - b_i|.$$

<u>Preuve</u> : par récurrence sur n.

. si $n = 1$ OK.

. si OK pour n fixé alors

$$\left| \prod_1^{n+1} a_i - \prod_1^{n+1} b_i \right| \leq \overbrace{|a_{n+1}|}^{\leq 1} \left| \prod_1^n a_i - \prod_1^n b_i \right| + |a_{n+1} - b_{n+1}| \underbrace{\prod_1^n |b_i|}_{\leq 1}$$

$$\underset{\text{réc}}{\overset{\text{Hyp}}{\leq}} \sum_1^n |a_i - b_i| + |a_{n+1} - b_{n+1}| = \sum_1^{n+1} |a_i - b_i|.$$

$\square$

Retour à la preuve de la loi faible des grands nombres.
Quitte à considérer $X_i - \mathbb{E}[X_i]$, on peut supposer que $\mathbb{E}[X_i] = 0$. Il suffit alors de montrer que $\frac{S_n}{n} \xrightarrow{\mathcal{L}} 0$.
Soit $t \in \mathbb{R}$, on a par indépendance

$$\varphi_{\frac{S_n}{n}}(t) = \varphi_{X_1}\left(\frac{t}{n}\right)^n$$ et d'après le lemme ci-dessus

$$\left| \varphi_{\frac{S_n}{n}}(t) - 1 \right| \leq n \left| \varphi_{X_1}\left(\frac{t}{n}\right) - 1 \right| = t \cdot \frac{n}{t} \left| \varphi_{X_1}\left(\frac{t}{n}\right) - 1 \right|$$

Comme $\mathbb{E}[|X_1|] < \infty$, φ_{X_1} est dérivable en zéro avec $\varphi_{X_1}'(0) = 0$ de sorte que

$$\frac{n}{t} \left| \varphi_{X_1}\left(\frac{t}{n}\right) - 1 \right| \underset{n\infty}{\longrightarrow} |\varphi_{X_1}'(0)| = 0$$

Autrement dit $\frac{S_n}{n} \xrightarrow{\mathcal{L}} 0$, par suite $\frac{S_n}{n} \xrightarrow{\mathbb{P}} 0$.

$\square$

1.2 Loi forte des grands nombres

Théorème : Soit (X_n) une suite de va iid intégrable alors

$$\frac{S_n}{n} = \frac{1}{n} \sum_{1}^{n} X_i \xrightarrow[n\infty]{p.s,\ \mathbb{L}^1} \mathbb{E}[X_1].$$

Remarque : • si (X_n) iid vérifie $\frac{S_n}{n} \xrightarrow{p.s} c \in \mathbb{R}$ alors $\mathbb{E}[|X_1|] < \infty$ et $c = \mathbb{E}[X_1]$.

En effet si $\frac{S_n}{n} \to c$ alors $\frac{X_n}{n} = \frac{S_n}{n} - \frac{S_{n-1}}{n-1} \frac{n-1}{n} \xrightarrow{p.s} c - c = 0$

et par Borel-Cantelli, $\forall \varepsilon > 0$ $\sum_n \mathbb{P}\left(\left|\frac{X_n}{n}\right| > \varepsilon\right) < \infty$

Par suite $\mathbb{E}[|X_1|] = \sum \mathbb{P}(|X_1| > n) < \infty$.

• si $X_1 \geq 0$ et $\mathbb{E}[X_1] = +\infty$ alors $\frac{S_n}{n} \xrightarrow{p.s} +\infty$.

Preuve : on donne la preuve dans le cas où $X_i \in \mathbb{L}^2$ de la cv p.s, pour le cas $\mathbb{L}^1$ on renvoie au poly de JC Breton ou à la preuve de l'an prochain avec les martingales inverses.

On suppose donc que $X_1 \in \mathbb{L}^2$ et quitte à considérer $X^+ = \max(X, 0)$ et $X^- = \max(-X, 0)$ on peut supposer que $X_1 \geq 0$. On a alors comme plus haut

$$\mathbb{P}\left(\left| \frac{S_n}{n} - \mathbb{E}[X_1] \right| > \varepsilon \right) \leq \frac{\mathrm{var}(X_1)}{n \varepsilon^2}$$

et donc

$$\sum_n \mathbb{P}\left(\left| \frac{S_{n^2}}{n^2} - \mathbb{E}[X_1] \right| > \varepsilon \right) \leq \sum \frac{\mathrm{var}(X_1)}{n^2 \varepsilon^2} < \infty.$$

D'après Borel-Cantelli, on a donc

$$\frac{S_{n^2}}{n^2} \xrightarrow{p.s} \mathbb{E}[X_1].$$

Ensuite, pour n général on pose $q_n := \lfloor \sqrt{n} \rfloor$ de sorte que $q_n \leq \sqrt{n} \leq q_n + 1$ i.e. $q_n^2 \leq n \leq (q_n+1)^2$

Comme les X_i sont positives, on a alors

$$\underbrace{\frac{q_n^2}{n} \frac{S_{q_n^2}}{q_n^2}}_{\downarrow p.s \; \mathbb{E}[X_1]} \leq \frac{S_n}{n} \leq \underbrace{\frac{S_{(q_n+1)^2}}{(q_n+1)^2} \cdot \frac{(q_n+1)^2}{n}}_{\downarrow p.s \; \mathbb{E}[X_1]}$$

$\square$

Remarque (orale) sur la portée philosophique de la LGN : du hasard peut naître le déterminisme !
Cela légitime la méthode scientifique i.e répéter des expériences dans des conditions similaires pour avoir une idée du résultat théorique.

1.3 Applications de la LGN

Méthode de Monte-Carlo.

Si (X_i) suite iid de loi $\mathbb{P}_X$ et si $g : \mathbb{R} \to \mathbb{R}$ mesurable, tq $g(X_1) \in \mathbb{L}^1$ alors

$$\frac{1}{n} \sum_i^n g(X_i) \xrightarrow[n\infty]{ps, \mathbb{L}^1} \mathbb{E}[g(X_1)] = \int g(x) \, \mathbb{P}_X(dx).$$

Par exemple, si $\mathbb{P}_X$ est a densité f_X, on a alors

$$\frac{1}{n} \sum_i^n g(X_i) \approx \int g(x) f_X(x) \, dx .$$

- Cette méthode peut être utilisée pour approcher des integrales. (voir TP 3)
-

Estimation paramétrique

On considère une élection type référendum ou 2e tour avec 2 issues possibles, disons A ou B.
Le résultat final est une proportion p pour A et $(1-p)$ pour B.
Comment annoncer le résultat / une estimation du résultat à 20h si certains bureaux de votes ferment à 20h.
On choisit un échantillon de votant au hasard. Si cet échantillon est "bien choisi" une personne interrogée aura une proba p de voter pour A, $(1-p)$ pour B.
On note $X_i = 1$ si vote pour A
$\qquad\qquad 0$ si vote pour B

Sur un échantillon de taille n, la proportion pour A sera alors $$\frac{S_n}{n} = \frac{1}{n} \sum_i X_i \xrightarrow[n\infty]{} \mathbb{E}[X_i] = p.$$

Ainsi pour n "assez grand", une estimation du paramètre inconnu p est $\hat{p}_n = \frac{S_n}{n}$.