Chapitre 2: Lois des grands nombres

Prof. Mohamed El Merouani

Université Abdelmalek Essaâdi
Faculté des Sciences de Tétouan
Département de Mathématiques
https://elmerouani.jimdofree.com/calcul-des-probabilités

2020/2021

Plan :

- Introduction
- Définitions
- Loi faible des grands nombres
- Loi forte des grands nombres
- Exercices

Introduction :

- Soit $(X_n)_{n\geq 1}$ une suite de v.a.r. centrées. On cherche des conditions suffisantes pour que la suite de v.a. $(\frac{1}{n}\sum_{k=1}^{n} X_n)_{n\geq 1}$ converge vers 0 au sens de l'un des types de convergence introduits au chapitre précèdent.
- Seules les convergences en probabilité et presque sûre ont fait l'objet d'une étude systématique.
- On parle respectivement de loi faible et de loi forte des grands nombres.

Définitions :

- La suite $(X_n)_{n\geq 1}$ satisfait la loi faible des grands nombres si la suite du terme général $\frac{1}{n}\sum_{k=1}^{n} X_n$ converge vers 0 en probabilité.
- La suite $(X_n)_{n\geq 1}$ satisfait la loi forte des grands nombres si la suite du terme général $\frac{1}{n}\sum_{k=1}^{n} X_n$ converge vers 0 presque sûre.
- On pose, pour tout $n \geq 1$:

$$S_n = \sum_{k=1}^{n} X_k, \qquad Y_n = \frac{S_n}{n}$$

Rappels :

- Rappelons que si X est de carré intégrable ($E(X^2) < +\infty$), sa variance est définie par $Var(X) = E[(X - E(X))^2]$.
- Les X_k sont dites deux à deux non-corrélées si $Cov(X_i, X_j) = 0$, pour tout i, j distincts.
- Ceci se produit en particulier lorsque les X_k sont deux à deux indépendantes.
- Pour des X_k deux à deux non-corrélées, on déduit l'égalité $Var(S_n) = \sum_{k=1}^{n} Var(X_k)$, que l'on écrit : $\sigma^2(S_n) = \sum_{k=1}^{n} \sigma^2(X_k)$.

Loi faible des grands nombres :

Théorème 1 :

Soit $(X_n)_{n\geq 1}$ une suite de v.a. de carrés intégrables, deux à deux non-corrélées. Pour tout $n \geq 1$, posons $E(X_n) = \mu_n$, $Var(X_n) = \sigma_n^2$ et supposons que la suite de terme général $\frac{1}{n}\sum_{k=1}^{n} \mu_k \underset{n\to+\infty}{\longrightarrow} \mu$ et que $\frac{1}{n^2}\sum_{k=1}^{n} \sigma_k^2 \underset{n\to+\infty}{\longrightarrow} 0$. Alors $\frac{S_n}{n} \xrightarrow{P} \mu$.

Preuve :

On a : $E(\frac{S_n}{n}) = \frac{1}{n}\sum_{k=1}^{n} \mu_k$

Les v.a. X_n étant deux à deux non-corrélées, on a :

$Var(\frac{S_n}{n}) = \frac{1}{n^2}\sum_{k=1}^{n} \sigma_k^2$

L'inégalité triangulaire conduit à l'inégalité :

$\left|\frac{S_n}{n} - \mu\right| \leq \left|\frac{S_n}{n} - \frac{1}{n}\sum_{k=1}^{n} \mu_k\right| + \left|\frac{1}{n}\sum_{k=1}^{n} \mu_k - \mu\right|$

Mais, pour tout $\varepsilon > 0$, il existe $N(\varepsilon) \in \mathbb{N}^*$ tel que pour tout $n \geq N(\varepsilon)$ on ait $\left|\frac{1}{n}\sum_{k=1}^{n} \mu_k - \mu\right| \leq \frac{\varepsilon}{2}$

Loi faible des grands nombres :

Preuve : (suite)

Pour tout $n \geq N(\varepsilon)$, on a donc l'inclusion des événements
$\{\left|\frac{S_n}{n} - \frac{1}{n}\sum_{k=1}^{n}\mu_k\right| \leq \frac{\varepsilon}{2}\} \subset \{\left|\frac{S_n}{n} - \mu\right| \leq \varepsilon\}$
ou encore, pour les complémentaires, l'inclusion :
$\{\left|\frac{S_n}{n} - \mu\right| > \varepsilon\} \subset \{\left|\frac{S_n}{n} - \frac{1}{n}\sum_{k=1}^{n}\mu_k\right| > \frac{\varepsilon}{2}\}$
L'inégalité de bienaymé-Tchebychev permet d'écrire :
$P\left(\left|\frac{S_n}{n} - \frac{1}{n}\sum_{k=1}^{n}\mu_k\right| > \frac{\varepsilon}{2}\right) \leq \frac{4}{\varepsilon^2}Var(\frac{S_n}{n}) = \frac{4}{\varepsilon^2}\frac{1}{n^2}\sum_{k=1}^{n}\sigma_k^2$
Il en résulte que pour tout $n \geq N(\varepsilon)$, on a :
$P(\left|\frac{S_n}{n} - \mu\right| > \varepsilon) \leq \frac{4}{\varepsilon^2}\frac{1}{n^2}\sum_{k=1}^{n}\sigma_k^2$
ce qui, en utilisant la seconde hypothèse, démontre le résultat. $\blacksquare$

Loi faible des grands nombres :

Remarque 1 :

En particulier, les hypothèses du théorème précédent sont toutes satisfaites si les v.a. X_n sont indépendantes et de même loi et si X_1 admet un moment d'odre deux.

On étudie un cas particulier du théorème précédent (il lui est toutefois historiquement antérieur), c'est le théorème de Bernoulli.

Théorème de Bernoulli :

Soit une suite de n épreuves de Bernoulli indépendantes et soit X la v.a. égale au nombre de succès au cours de ces n épreuves. Alors $\frac{X}{n} \xrightarrow{P} p$, où p est la probabilité de succès.

Loi faible des grands nombres :

Preuve du théorème de Bernoulli :

On sait que X suit la loi binomiale de paramètres n et p : $X \rightsquigarrow \mathcal{B}(n,p)$.
Donc $E(X) = np$ et $Var(X) = npq$ avec $q = 1-p$.
Soit la v.a. $Y = \frac{X}{n}$, alors $E(Y) = \frac{1}{n}E(X) = \frac{np}{n} = p$ et
$Var(Y) = E\left((\frac{X}{n} - p)^2\right) = \frac{1}{n^2}E((X-np)^2) = \frac{Var(X)}{n^2} = \frac{pq}{n}$
En utilisant l'inégalité de Bienaymé-Tchebychev : $P(|Y-p| \geq \varepsilon) \leq \frac{pq}{n\varepsilon^2}$
Comme nous ne connaissons pas la valeur de p, il sera nécessaire de calculer une valeur extrême supérieure de $\frac{pq}{n\varepsilon^2}$.

$$\frac{d}{dp}\left(\frac{pq}{n\varepsilon^2}\right) = \frac{d}{dp}\left(\frac{p-p^2}{n\varepsilon^2}\right) = \frac{1-2p}{n\varepsilon^2} = 0 \Leftrightarrow p = \frac{1}{2}$$

$$\frac{d^2}{dp^2}\left(\frac{p-p^2}{n\varepsilon^2}\right) = \frac{d}{dp}\left(\frac{1-2p}{n\varepsilon^2}\right) = \frac{-2}{n\varepsilon^2} < 0.$$

Donc $p = \frac{1}{2}$ est un maximum et par conséquent $\frac{pq}{n\varepsilon^2} \leq \frac{\frac{1}{2}\cdot\frac{1}{2}}{n\varepsilon^2} = \frac{1}{4n\varepsilon^2}$
d'où $P(|Y-p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}$ ou encore $P(|Y-p| < \varepsilon) \geq 1 - \frac{1}{4n\varepsilon^2}$

Par conséquent $\lim_{n\to+\infty} P(|Y-p| < \varepsilon) = 1$ et $\frac{X}{n} \xrightarrow{P} p$. ■

Théorème de Bernoulli :

- Ce théorème veut dire que nous pouvons utiliser $\frac{X}{n}$ pour estimer la probabilité p, puisque c'est une estimation exacte lorsque n est suffisamment grande.
- Dans la pratique, le nombre n doit être fini, par conséquent, on doit déterminer le nombre n de façon que, avec une probabilité suffisamment grande, la fréquence observée $\frac{X}{n}$ soit comprise entre $p - \varepsilon$ et $p + \varepsilon$.
- La fréquence observée des piles dans un jeu de pile ou face, après n tirages, est "proche" de la probabilité (déterministe) p d'obtenir pile, pourvu que n soit grande.
- Donc, si p est inconnue (par exemple, on ne sait pas si la pièce est truquée), nous avons là un moyen de l'approximer.

Loi forte des grands nombres :

Théorème 3 :

Soit $(X_n)_{n\geq 1}$ une suite de v.a. de carrées intégrables, centrées et indépendantes. Pour tout $n \geq 1$, on pose $Var(X_n) = \sigma_n^2 < +\infty$. Si la série $\sum_{n\geq 1} \frac{\sigma_n^2}{n^2}$ est convergente, alors $Y_n = \frac{S_n}{n} \xrightarrow{p.s.} 0$ (*).

Remarque 2 :

Si les X_n ne sont pas centrées, en appliquant (*) aux :
$X_n' = X_n - E(X_n)$, on obtient : $\frac{1}{n}S_n - \frac{1}{n}\sum_{k=1}^{n} E(X_k) \xrightarrow{p.s.} 0$.
Ce résultat est intéressant notamment dans le cas où la suite déterministe $(E(X_n))_{n\geq 1}$ a une limite ℓ quand n tend vers l'infini.
En effet, le théorème de Césaro nous donne alors la convergence vers ℓ de $\frac{1}{n}\sum_{k=1}^{n} E(X_k)$,
d'où la convergence p.s. de $\frac{S_n}{n}$ vers ℓ.
En particulier, si les X_n ont même espérance μ, $\frac{S_n}{n} \xrightarrow{p.s.} \mu$.

Loi forte des grands nombres :

Théorème 4 : (Rajchman)

Soit $(X_n)_{n\geq 1}$ une suite de v.a. de carrées intégrables, centrées et indépendantes. Pour tout $n \geq 1$, on pose $Var(X_n) = \sigma_n^2$. Si $\sup_n \sigma_n^2 < +\infty$, alors :

1. $Y_n \xrightarrow{p.s.} 0$
2. $E(|Y_n|^2) \underset{n\to+\infty}{\longrightarrow} 0$

Preuve :

1. Posons $\sigma^2 = \sup_n \sigma_n^2 < +\infty$;
   alors $\sum_{n\geq 1} \frac{\sigma_n^2}{n^2} \leq \sigma^2 \sum_{n\geq 1} \frac{1}{n^2} < \infty$
   d'où $Y_n \xrightarrow{p.s.} 0$ d'après le théorème 3 précédent.
2. On a : $E(Y_n^2) = Var(Y_n) = \frac{1}{n^2} \sum_{k=1}^{n} \sigma_k^2 \leq \frac{\sigma^2}{n} \underset{n\to+\infty}{\longrightarrow} 0$ ■

Loi forte des grands nombres :

Remarque 3 :

Rajchman a établi les mêmes conclusions en remplaçant "indépendantes" par "deux à deux non-corrélées".

Remarque 4 :

Il en résulte que, dans l'énoncé du théorème de Bernoulli, on peut remplacer la convergence en probabilité par la convergence presque sûre (E. Borel).

Loi forte des grands nombres :

Théorème 5 : (Kolmogorov)

Soit $(X_n)_{n\geq 1}$ une suite de v.a. **intégrables**, centrées, indépendantes et identiquement distribuées. On a : $Y_n \xrightarrow{p.s.} 0$

Corollaire 1 :

Soit $(X_n)_{n\geq 1}$ une suite de v.a. **intégrables**, indépendantes et identiquement distribuées. Alors : $Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{p.s.} E(X_1)$

Le théorème précédent est le résultat "optimal" pour les v.a. indépendantes et identiquement distribuées (i.i.d.) en raison de la "réciproque" suivante :

Théorème 6 :

Soit $(X_n)_{n\geq 1}$ une suite de v.a. indépendantes et identiquement distribuées telle que $\frac{S_n}{n}$ converge presque sûrement. Alors $E|X_1| < +\infty$

Loi forte des grands nombres :

Preuve :

Par hypothèse $\frac{S_n}{n}$ converge p.s. vers une certaine v.a. Y.

Notons $\Omega' = \{\omega \in \Omega, \frac{S_n(\omega)}{n} \underset{n\to+\infty}{\longrightarrow} Y(\omega)\}$.

Alors $P(\Omega') = 1$ et pour tout $\omega \in \Omega'$,

$\frac{X_n(\omega)}{n} = \frac{S_n(\omega)-S_{n-1}(\omega)}{n} = \frac{S_n(\omega)}{n} - \frac{n-1}{n}\frac{S_{n-1}(\omega)}{n-1} \underset{n\to+\infty}{\overset{p.s.}{\longrightarrow}} 0$

puisque $\frac{S_n(\omega)}{n}$ et $\frac{S_{n-1}(\omega)}{n-1}$ convergent vers le même réel $Y(\omega)$.

Considérons l'événement

$B = \{\omega \in \Omega, \exists n_0 = n_0(\omega), \forall n \geq n_0, \frac{|X_n(\omega)|}{n} < 1\}$

Puisque $P(\Omega') = 1$ et $\Omega' \subset B$, l'évébement B a pour probabilité 1, d'où en passant au complémentaire

$P\left(\frac{|X_n(\omega)|}{n} \geq 1 \text{ une infinité de fois}\right) = 0.$

Loi forte des grands nombres :

Preuve : (suite)

Il résulte alors du second lemme de Borel-Cantelli que

$$\sum_{n\geq 1} P\left(\frac{|X_n(\omega)|}{n} \geq 1\right) < +\infty.$$

Les X_n ayant même loi, ceci s'écrit aussi $\displaystyle\sum_{n=1}^{+\infty} P\left(|X_1| \geq n\right) < +\infty.$

Pour finir la preuve, on observe que

$$E(|X_1|) = \int_0^{+\infty} P(|X_1| > t)dt = \sum_{n=0}^{+\infty} \int_n^{n+1} P(|X_1| > t)dt$$

$$\leq \sum_{n=0}^{+\infty} P(|X_1| \geq n) = 1 + \sum_{n=1}^{+\infty} P(|X_1| \geq n) < +\infty. \blacksquare$$

Loi forte des grands nombres :

Rappel :
$E(|X_1|) = \int_0^{+\infty} P(|X_1| > t)dt$
En effet, on a le résultat suivant :

Soit X une v.a. à valeurs dans $[0, +\infty[$, absolument continue, de densité f, de fonction de survie $R(x) = 1 - F(x) = P(X > x)$. Alors on a, dans la demi-droite achevée $[0, +\infty]$, $\int_0^{+\infty} R(x)dx = \int_0^{+\infty} xf(x)dx$

Preuve :
En effet, $\int_0^{+\infty} R(x)dx = \int_0^{+\infty} \left(\int_x^{+\infty} f(t)dt \right) dx =$
$= \int_0^{+\infty} \left(\int_0^{+\infty} f(t)\mathbb{I}_{\{t>x\}}(t,x)dt \right) dx$
Soit d'après le théorème de Fubini
$\int_0^{+\infty} R(x)dx = \int_0^{+\infty} \left(\int_0^{+\infty} \mathbb{I}_{\{t>x\}}(t,x)dx \right) f(t)dt = = \int_0^{+\infty} tf(t)dt$ ■

Loi forte des grands nombres :

Rappel :
Dans la démonstration précédente, on a fait appel au th. de Fubini pour les intégrales doubles, mais, en fait, il n'y a aucune difficulté ici puisque la fonction à intégrer est positive.

Remarque 5 :

Si $X \geq 0$ et si $E(X) < \infty$, alors $E(X) = \int_0^{+\infty} R(x)dx$

Loi forte des grands nombres :

Complément :

Montrons que si $(X_n)_{n\geq 1}$ est une suite de v.a. indépendantes telle que la suite $(Y_n)_{n\geq 1}$ (avec $Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k$) converge p.s. vers une limite Y, alors Y est p.s. constante.

Pour le voir, remarquons tout d'abord que, pour tout x réel, l'événement $\{Y_k \leq x\}$ est indépendant de $\{Y \leq x\}$. (L'événement $\{Y \leq x\}$ ne peut admettre pour probabilité que 0 ou 1). On a alors $P(\{Y_k \leq x\} \cap \{Y \leq x\}) = P(\{Y_k \leq x\}).P(\{Y \leq x\})$, $\forall x$ réel. En faisant tendre k vers l'infini, on en déduit $P(Y \leq x) = P(Y \leq x)^2$ et donc $P(Y \leq x) = 0$ ou 1 $\forall x$. Comme l'application $x \mapsto P(Y \leq x)$ est une fonction de répartition, elle est nécessairement égale à l'échelon unité. Ainsi Y est p.s. constante. ∎

Exercices :

Exercice 1 :

Soit $(X_n)_{n\geq 1}$ une suite de v.a. indépendantes, de même loi, de carrés intégrables.
On pose $E(X_1) = \mu$, $Var(X_1) = \sigma^2$.
Pour tout entier $n \geq 2$, on définit les v.a. :
$Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ et $Z_n = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - Y_n)^2$.

1. Calculer $E(Z_n)$.
2. Montrer que $Z_n \xrightarrow{p.s.} \sigma^2$, lorsque n tend vers l'infini.

Exercices :

Exercice 2 :

Soit $(X_n)_{n\geq 1}$ une suite de v.a. indépendantes, identiquement distribuées. On suppose que $Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{p.s.} Y$.
Démontrer les propriétés suivantes :

1. $\sum_{n\geq 1} P\left(|X_n| \geq n\right) < +\infty$;
2. Les X_n sont intégrables ;
3. Y =cste, presque sûrement.

Exercices :

Exercice 3 :

Soit $(X_n)_{n\geq 1}$ une suite de v.a.
On pose $S_n = \sum_{k=1}^{n} X_k$.
Montrer que si $\frac{S_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} Y$, alors $\frac{S_n}{n} \xrightarrow{P} 0$,
c'est-à-dire, la suite $(X_n)_{n\geq 1}$ vérifie la loi faible des grands nombres.

Exercice 4 :

Soit $(X_n)_{n\geq 1}$ une suite de v.a. indépendantes de moyennes $E(X_n)$ finies, et soit X une v.a. de variance finie telle que pour tout n les variables $X_1, X_2, \cdots, X_n$ et $X - (X_1 + X_2 + \cdots + X_n)$ soient indépendantes.
Montrer que X_n est de variance finie et que la série de terme général $(X_n - E(X_n))$ converge presque sûrement.