La loi forte des grands nombres L^2

Yves Coudène, Septembre 2006

Voici une preuve de la loi forte des grands nombres, lorsque les variables aléatoires sont dans L^2.

Théorème :
Soit $(\Omega, \mathcal{T}, \mu)$ un espace mesuré, soient $(X_i)_{i\in\mathbf{N}}$ des fonctions intégrables définies de Ω dans $\mathbf{R}$, de carrés intégrables. On suppose que pour tout i, j, $i \neq j$, $\int_\Omega X_i \, d\mu = 0$, $\int_\Omega X_i^2 \, d\mu = 1$ et $\int_\Omega X_i X_j \, d\mu = 0$; alors :

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to+\infty]{} 0 \quad p.s.$$

Preuve :

• Par l'inégalité de Markov,

$$\mu\Big(\Big|\tfrac{1}{m}\sum_{i=1}^{m} X_i\Big| > \varepsilon\Big) \leq \frac{1}{m^2\varepsilon^2}E((\sum X_i)^2)$$

$$E((\textstyle\sum X_i)^2) = E(\displaystyle\sum_{i,j} X_iX_j) = E(\sum X_i^2) + 2E(\sum_{i<j} X_iX_j) = m$$

donc $\mu\Big(\Big|\tfrac{1}{m}\sum_{i=1}^{m} X_i\Big| > \varepsilon\Big) \leq \dfrac{1}{m\varepsilon^2}$

• On veut appliquer le lemme de Borel-Cantelli, mais $\sum \frac{1}{m} = +\infty$. Remplaçons m par m^2 :

$$\mu\Big(\Big|\frac{1}{m^2}\sum_{i=1}^{m^2} X_i\Big| > \varepsilon\Big) \leq \frac{1}{m^2\varepsilon^2}.$$

Comme $\sum \frac{1}{m^2} < +\infty$, le lemme de Borel-Cantelli montre que :

presque partout, $\forall \varepsilon > 0, \exists M \in \mathbf{N}, \forall\, m \geq M,\ \Big|\tfrac{1}{m^2}\displaystyle\sum_{i=1}^{m^2} X_i\Big| < \varepsilon.$

On en déduit, en prenant $\varepsilon = 1/K, K \in \mathbf{N}^*$,

$$\frac{1}{m^2}\sum_{i=1}^{m^2} X_i \xrightarrow[m\to+\infty]{} 0 \qquad p.p.$$

• Soit $n \in \mathbf{N}$. Soit m la partie entière de $\sqrt{n}$. On a : $m^2 \leq n \leq (m+1)^2 - 1$
On utilise à nouveau le lemme de Borel-Cantelli :

$$\mu\Big(\frac{1}{n}\Big|\sum_{i=m^2+1}^{n} X_i\Big| > \varepsilon\Big) \leq \frac{1}{n^2\varepsilon^2}E(\sum X_i^2) \leq \frac{1}{n^2\varepsilon^2}(n-m^2) \leq \frac{2}{n^{3/2}\varepsilon^2}$$

Comme $\sum \frac{1}{n^{3/2}} < +\infty$, on en déduit que :

$$\frac{1}{n}\sum_{i=m^2+1}^{n} X_i \xrightarrow[n\to+\infty]{} 0 \qquad p.p.$$

Remarques :

– L'hypothèse $\int X_i \; d\mu = 0$ est une hypothèse de centrage.

– Si les X_i sont centrées et que la suite des $(X_i)_{i\in\mathbf{N}}$ est stationnaire, on peut remplacer les hypothèses sur les X_i par une condition dite de "corrélations sommables" :

$$\textit{La Série } \sum_i \int X_1 X_i \; d\mu \textit{ est absolument convergente.}$$

– Lorsque $(\Omega, \mathcal{T}, \mu)$ est un espace probabilisé, et que les X_i ne sont pas centrés, on peut remplacer les hypothèses sur les X_i par une hypothèse de "décorrélation" :

$$\int X_i X_j \; d\mu - \int X_i \; d\mu \int X_j \; d\mu = 0.$$

La moyenne des X_i converge alors presque surement, dès que la moyenne des $E(X_i)$ converge, et ces deux moyennes sont égales. C'est l'énoncé classique de la loi des grands nombres pour des variables aléatoires L^2.

– Lorsque $(\Omega, \mathcal{T}, \mu)$ est un espace probabilisé, et que les X_i sont des fonctions bornées par une constante C, on peut simplifier la seconde partie de la preuve en utilisant la majoration : $\Big| \sum_{i=m^2+1}^{n} X_i \Big| \leq C(n - m^2) \leq 2C\sqrt{n}$

Voici un corollaire de la loi forte des grands nombres, qui s'obtient en considérant des fonctions de la forme $\mathbf{1}_A(X_i)$:

Corollaire :
Soient $(\Omega, \mathcal{T}, \mu)$ un espace probabilisé, $(X_i)_{i\in\mathbf{N}}$ une suite de variables aléatoires indépendantes équidistribuées, et $A \in \mathbf{R}$ un borélien. Alors, pour presque tout $\omega \in \Omega$,

$$\frac{1}{n} \; Card \Big\{ i \in \{1, 2, .., n\} \mid X_i(\omega) \in A \Big\} \xrightarrow[n\to+\infty]{} P(X_1 \in A)$$

Remarquons pour terminer, qu'il est possible d'obtenir une vitesse dans la loi des grands nombres lorsque les variables aléatoires sont dans L^2. Il suffit de remplacer ε par $n^{-0,2}\varepsilon$ dans la preuve du théorème pour obtenir :

$$\frac{1}{n^{0,8}} \sum_{i=1}^{n} X_i \longrightarrow 0 \quad p.p.$$

La normalisation optimale est en fait en $\frac{1}{n^{1/2\;+\varepsilon}}$; c'est une application du théorème des trois séries, mais cela ne semble pas pouvoir se déduire de la preuve donnée plus haut.