

Leçon 262 - Modes de convergence d'une suite de variables aléatoires.
Exemples et applications.

Cadre : Toutes les variables aléatoires sont définies sur un espace probabilisé (Ω, A, P) à valeurs dans \mathbb{R}^d pour un $d \geq 1$, muni de la tribu borélienne et de la mesure de Lebesgue.

1. Convergence presque sûre et convergence en probabilité. —

1. Autour de la convergence presque sûre. —

- Def : Une suite de v.a. $(X_n)_n$ converge P-presque sûrement vers une v.a. X si $P(\{w \in \Omega \text{ tq } X_n(w) \not\rightarrow X(w)\}) = 0$.
On le note $X_n \rightarrow_{P-p.s} X$.
- Pro : On a l'équivalence :
 - i) $X_n \rightarrow_{P-p.s} X$.
 - ii) $P(\{w \in \Omega \text{ tq } \sup_{k \geq n} (\|X_k(w) - X(w)\|) \not\rightarrow 0\}) = 0$
 - iii) $P(\{w \in \Omega \text{ tq } \limsup_n (\|X_n(w) - X(w)\|) \not\rightarrow 0\}) = 0$.
 - iv) $\forall \varepsilon > 0, P(\bigcup_n \bigcap_{m \geq n} \{|X_n - X_m| < \varepsilon\}) = 1$.
- Pro : La limite P-presque sûre est unique modulo l'égalité P-presque partout.
- Pro : Si $X_n \rightarrow_{P-p.s} X$ et $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$ est continue, alors $f(X_n) \rightarrow_{P-p.s} f(X)$.
- Pro : Soit (E_n) une suite de v.a. iid de loi de Bernoulli de paramètre $\frac{1}{2}$. Alors la suite $X_n := \sum_{k=1}^n \frac{E_k}{2^k}$ converge P-p.s vers une v.a. de loi $U_{[0,1]}$.
- Lemme de Borel-Cantelli : Soit $(A_n)_n$ une suite d'évènements de Ω .
 - i) Si $\sum_n P(A_n) < +\infty$, alors $P(\limsup_n (A_n)) = 0$.
 - ii) Si les A_n sont indépendants et si $\sum_n P(A_n) = +\infty$, alors $P(\limsup_n (A_n)) = 1$.
- Cor : Soient $(X_n)_n$ et X des variables aléatoires. On a :
 - i) Si $\forall \varepsilon > 0, \sum_n P(\|X_n - X\| \geq \varepsilon) < +\infty$, alors $X_n \rightarrow_{P-p.s} X$.
 - ii) Si les X_n sont indépendantes, alors $X_n \rightarrow_{P-p.s} 0$ si et seulement si $\sum_n P(\|X_n\| \geq \varepsilon) < +\infty$ pour tout $\varepsilon > 0$.
- Pro : Soit $(X_n)_n$ une suite de v.a. iid de loi $Exp(1)$ donnée par la densité $f(x) = \chi_{\mathbb{R}_+}(x)e^{-x}$.
Alors $\frac{\max_{1 \leq i \leq n} (X_i)}{\ln(n)} \rightarrow_{P-p.s} 1$.

2. Convergence en probabilité. —

- Def : On dit que $(X_n)_n$ converge en probabilité P vers X si $\forall \varepsilon > 0, P(\|X_n - X\| > \varepsilon) \rightarrow_n 0$, et on le note $X_n \rightarrow_P X$.
- Pro : La limite en probabilité est unique modulo égalité P-presque partout.
- Pro : Si $(X_n)_n$ converge P-p.s vers X , alors $(X_n)_n$ en probabilité vers X .
- Pro : Soit $X_n \rightarrow_P X$ et $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$ continue. Alors $f(X_n) \rightarrow_P f(X)$.
- Thm : Si une suite $(X_n)_n$ converge en probabilité vers une v.a. X , alors il existe une sous-suite de $(X_n)_n$ qui converge P-p.s vers X .
- Contre-ex : Pour X_n des v.a. indépendantes à valeurs dans $\{0, 1\}$, avec $P(X_n = 1) = \frac{1}{n}$, X_n converge en probabilité vers la v.a. constante 0, mais ne converge pas P-p.s vers 0 car les évènements $P(X_n = 1)$ sont indépendants et de somme infinie.

- Pro : (Critère de Cauchy de convergence en probabilité) $(X_n)_n$ converge en probabilité vers X ssi $\forall \varepsilon > 0$, la suite double de terme général $P(\|X_n - X_m\| > \varepsilon)$ tend vers 0.
- Pro : (Métrisabilité de la convergence en probabilité) Pour $d(X, Y) = E[\sup(\|X - Y\|, 1)]$, d est une distance sur l'ensemble des v.a. $\Omega \rightarrow \mathbb{R}^d$.
On a de plus : $X_n \rightarrow_P X$ ssi $d(X_n, X) \rightarrow_n 0$.

3. Loi des grands nombres et applications. —

- Loi faible des grands nombres : Soit $(X_n)_n$ une suite de v.a. avec un moment d'ordre 2, telles que $Cov(X_i, X_j) = 0 \forall i \neq j$.
Si $\frac{1}{n} \sum_{k \leq n} E[X_k] \rightarrow m$ et $\frac{1}{n^2} \sum_{k \leq n} \sigma_{X_k}^2 \rightarrow 0$, alors $\frac{X_1 + \dots + X_n}{n}$ converge en probabilité vers m .
- Ex : Pour $(X_n)_n$ suite de v.a. indépendantes de loi $\frac{1}{2n \ln(n)}(\delta_n + \delta_{-n}) + (1 - \frac{1}{2n \ln(n)})\delta_0$, la suite $\frac{X_1 + \dots + X_n}{n}$ ne converge pas P-p.s vers 0 par Borel-Cantelli, mais converge en probabilité vers 0.
- Théorème de Khintchine (Loi faible des grands nombres) : Soit $(X_n)_n$ une suite de v.a. réelles indépendantes et de même loi, d'espérance finie. Alors la suite de v.a. $\frac{X_1 + \dots + X_n}{n}$ converge en probabilité vers $E[X_1]$.
- App : Pour X_1 de loi de Bernoulli de paramètre p , $\frac{X_1 + \dots + X_n}{n}$ qui est de loi $Bin(n, \frac{p}{n})$ converge en probabilités vers la v.a. constante p .
- Théorème de Kolmogorov-Khintchine : Soit $(X_n)_n$ une suite de v.a. réelles indépendantes et de même loi. On a l'équivalence :
 - i) $\exists c \in \mathbb{R}$ tq $\frac{X_1 + \dots + X_n}{n}$ converge en probabilité vers c .
 - ii) X_1 est intégrable.
- Loi forte des grands nombres : Soient (X_n) des v.a. réelles indépendantes admettant un moment d'ordre 2, telles que $E[X_n] \rightarrow m$ et $\sum_{k \geq 0} frac{\sigma_{X_k}^2}{k^2} < +\infty$.
Alors la suite de v.a. $\frac{X_1 + \dots + X_n}{n}$ converge P-p.s vers m .
- Rem : La moyenne empirique $\overline{X}_n := \frac{X_1 + \dots + X_n}{n}$ permet ainsi d'estimer $E[X_1]$ de façon consistante.
- Dev : Approximation d'intégrale par la méthode de Monte-Carlo : Soit $f : [0, 1]^d \rightarrow \mathbb{R}$ intégrable. On note $I = \int_{[0, 1]^d} f(x) d\lambda(x)$.
Pour $X_{n,k}$ des v.a. réelles iid de loi $U_{[0, 1]}$, on construit $Y_n := (X_{n,1}, \dots, X_{n,d})$ qui est une famille de v.a. iid de loi uniforme sur $[0, 1]^d$. On note $S_n = (f(Y_1) + \dots + f(Y_n))/n$.
Alors, pour tout $0 < \varepsilon \leq (\frac{\|f\|_2}{\|f\|_\infty})^2$ assez petit, on a : $P(I - S_n > \varepsilon) \leq 2 \exp(-n(\frac{\varepsilon \|f\|_\infty}{\|f\|_2})^2)$.
L'approximation de l'intégrale de f par le barycentre de n évaluations données par des v.a. uniformes a ainsi une probabilité de ne pas être ε -proche de l'intégrale de f qui décroît exponentiellement en n .

2. Convergence en norme $L^p, p \geq 1$. —

- Def : Pour tout $1 \leq p < +\infty$, on définit $L^p(\Omega, A, P)$ comme le quotient de l'espace vectoriel des v.a. réelles X telles que X^p est intégrable par le sous-ev des v.a. réelles

nulles P-presque partout.

On définit l'application $\|X\|_p := (E[|X|^p])^{\frac{1}{p}}$.

- Thm : $\|\cdot\|_p$ est une norme sur $L^p(\Omega, A, P)$, et $(L^p(\Omega, A, P), \|\cdot\|_p)$ est complet.
- Cor : Une suite $(X_n)_n$ de $L^q(\Omega, A, P)$ converge en norme L^q ssi elle est de Cauchy pour la norme $\|\cdot\|_q$.
- Pro : Pour tous $q \geq p \geq 1$, on a :
 - i) $L^q(\Omega, A, P) \subset L^p(\Omega, A, P) \subset L^1(\Omega, A, P)$.
 - ii) Si une suite $(X_n)_n$ de $L^q(\Omega, A, P)$ converge vers X en norme $\|\cdot\|_q$, alors elle converge vers X en norme $\|\cdot\|_p$, et converge aussi vers X en probabilité. ($\|\cdot\|_q \Rightarrow \|\cdot\|_p \Rightarrow P$)
- Contre-ex : Pour X_n indépendantes avec $P(X_n = n^c) = \frac{1}{n}$ et $P(X_n = 0) = 1 - \frac{1}{n}$ pour un certain $c > 0$, $X_n \rightarrow_P 0$, mais $\|X_n\|_p = n^{c-\frac{1}{p}}$ diverge si $c > \frac{1}{p}$.
- Rem : Si $(X_n)_n$ converge vers X en norme L^q , alors il existe une extractrice de $(X_n)_n$ qui converge vers X P-p.s.
- Contre-ex : On se place sur $([0, 1], \mathcal{B}([0, 1]), \lambda)$. Pour tout $m \geq 0$, on définit les v.a. $f_m := \chi_{[\frac{k}{2^m}, \frac{k+1}{2^m}]}$ pour $m = 2^n + k, 0 \leq k < 2^n$. Alors $f_m \rightarrow_{\|\cdot\|_p} 0$ mais f_m ne converge pas vers 0 P-p.s.
- Def : Une famille de v.a. réelles $(X_i)_{i \in I}$ est dite uniformément intégrable si : $\lim_{c \rightarrow +\infty} (\sup_{i \in I} (\int_{|X_i| > c} |X_i| dP)) = 0$.
- Ex : Une famille finie est uniformément intégrable. Une famille tq $|X_i| \leq |Y|$, où Y est intégrable, est uniformément intégrable.
- Pro : Une famille $(X_i)_i$ est uniformément intégrable ssi on a :
 - i) $\forall \varepsilon > 0$, il existe $\eta > 0$ tq pour tout $A \in \mathcal{A}$ tq $P(A) < \eta$, on a $\int_A |X_i| dP < \varepsilon \forall i \in I$.
 - ii) $\sup_i (E[|X_i|]) < +\infty$.
- Théorème de Vitali : Soit $(X_n)_n$ est une suite de v.a. réelles qui converge en probabilité vers une v.a. X. La famille $(X_n)_n$ est uniformément intégrable ssi $X \in L^1(\Omega, A, P)$ et $X_n \rightarrow \|\cdot\|_1 X$.

3. Convergence en loi. —

1. Définition et propriétés. —

- Def : Une suite de v.a. réelles X_n converge en loi vers une v.a. réelle X si $E[f(X_n)] \rightarrow E[f(X)]$ pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ continue bornée. On note alors $X_n \rightarrow_{\mathcal{L}} X$.
- Pro : $X_n \rightarrow_{\mathcal{L}} X$ ssi pour F_X, F_{X_n} les fonctions de répartition des v.a., on a : $F_{X_n}(x) \rightarrow F_X(x)$ pour tout x point de continuité de F_X .
- Ex : Pour X_n de loi $N(a_n, \sigma_n^2)$ avec $a_n \rightarrow a$ et $\sigma_n^2 \rightarrow \sigma^2 > 0$, alors X_n converge en loi vers $N(a, \sigma^2)$.
- Pro : Si $X_n \rightarrow_{P-p.s} X$, alors $X_n \rightarrow_{\mathcal{L}} X$. Si $X_n \rightarrow_P X$, alors $X_n \rightarrow_{\mathcal{L}} X$.
- Contre-ex : Pour $X \sim N(0, 1)$, $X_n = (-1)^n X$ est encore de loi $N(0, 1)$, donc converge en loi vers X. Mais X_n ne converge ni P-p.s, ni en probabilité.
- Ex : Si X_n converge en loi vers une v.a. constante c, alors $X_n \rightarrow_P c$.

- Lemme de Slutsky : Pour $(X_n)_n, (Y_n)_n$ des v.a. réelles telles que $X_n \rightarrow_{\mathcal{L}} X$ et $X_n - Y_n \rightarrow_P 0$, on a $Y_n \rightarrow_{\mathcal{L}} X$.
- Def+Pro : On appelle fonction caractéristique d'une v.a. X l'application φ_X donnée par : $\forall t \in \mathbb{R}^d, \varphi_X(t) := \int_{\mathbb{R}^d} \exp(i \langle x, t \rangle) dP_X(x) = E[\exp(i \langle X, t \rangle)]$. On montre que si $\varphi_X = \varphi_Y$, alors $X \sim Y$: φ_X caractérise la loi de X.
- Dev : Théorème de Lévy : X_n converge en loi vers X ssi φ_{X_n} converge simplement vers φ_X .
- App : Théorème Central de la Limite : Soit X_n une suite de v.a. réelles iid ayant un moment d'ordre 2. Alors la suite $\frac{X_1 + \dots + X_n - nE[X_1]}{\sqrt{\text{var}(X_1)n}}$ converge en loi vers une loi normale centrée réduite $N(0, 1)$.
- Application à un intervalle de confiance.
- Thm : Soit $(S_n)_n$ une suite de v.a. de loi $Bin(n, p_n)$. Si $n.p_n$ converge vers un $\lambda > 0$, alors S_n converge en loi vers une v.a. de Poisson de paramètre λ , càd : $P(S_n = m) \rightarrow e^{-\lambda} \frac{\lambda^m}{m!}$.
- Rem : Dans le cas où p est assez petit, on peut approximer une v.a. de loi $Bin(n, p)$ par P_{np} .
- Théorème des évènements rares de Poisson : Pour tout $n \geq 0$, soit $M_n \geq 2$ et soient $A_{n,1}, \dots, A_{n,M_n}$ des évènements de Ω . On définit $S_n := \sum_{j \leq M_n} \chi_{A_{n,j}}$. Si $\sum_{j \leq M_n} P(A_{n,j})$ converge vers un $\lambda > 0$ et si $\max_j (P(A_{n,j}))$ converge vers 0, alors S_n converge en loi vers une v.a. de loi Poisson de paramètre λ .
- Ex : Illustration avec les clients d'une banque.

2. Applications en statistiques. —

- Application du TCL à la détermination d'un intervalle de confiance
- Application permettant d'oublier des variables inconnues pour former d'autres intervalles de confiance
- Pro : (δ -méthode)

Références

Ouvrard (Probas 2) : Contre-ex de convergence P-p.s. Contre-ex P et pas P-p.s. Loi faible des grands nombres, Th de Khintchine, Loi forte des grands nombres, exemple, Méthode de Monte Carlo.(Dev)(incomplet), Th de Glivenko-Cantelli. Th des évènements rares de Poisson, application aux clients d'une banque. Continuous mapping theorem, Th de Slutsky. Applications du TCL, δ -méthode. Barbe, Ledoux : Convergence P-p.s, définitions équivalentes, critère de Cauchy, unicité p.s, exemple, image d'une suite P-p.s convergente par une fonction continue, $\sum_n \frac{E_n}{2^n}$ avec $E_n \sim B(\frac{1}{2})$ indép, Lemme de Borel-Cantelli, Lemme de Borel-Cantelli de convergence P-p.s, exemple. Convergence P, unicité p.s de la limite, P-p.s implique P, P donne P-p.s à extraction près, métrique associée, critère de Cauchy, image d'une suite P-convergente par une fonction continue, exemple. Def de $L^p(\Omega, A, P)$, norme $\|\cdot\|_p$, espace complet, inclusions topologiques, L^p implique P, L^p donne P-p.s à extraction près, exemple, contre-exemple,

uniforme intégrabilité, def équivalente, exemple des familles finies/dominées, unif intégrable et P-convergente ssi L^1 -convergente. Convergence en loi, def équivalente, exemple, P-convergence implique CV en loi, contre-exemple, convergence \mathcal{L} vers une constante, fonction caractéristique.

Zuily, Queffelec : Théorème de Lévy+TCL.(Dev)

Cadre, Vial : Application du TCL à la détermination d'un intervalle de confiance, Application permettant d'oublier des variables inconnues pour former d'autres intervalles de confiance, Intervalle de confiance asymptotique pour un jeu de pile ou face, δ -méthode.

June 2, 2017

Vidal Agniel, École normale supérieure de Rennes