Séminaire de probabilités (Strasbourg)

Luca Pratelli

Une caractérisation de la convergence dans L^1. Application aux quasimartingales

Séminaire de probabilités (Strasbourg), tome 26 (1992), p. 61-69

<http://www.numdam.org/item?id=SPS_1992__26__61_0>

UNE CARACTERISATION DE LA CONVERGENCE DANS L^1. APPLICATION AUX QUASIMARTINGALES

Luca Pratelli
Dipartimento di Matematica, Università di Pisa
Via Buonarroti 2, I-56127 Pisa

RÉSUMÉ. – Étant données des variables aléatoires réelles intégrables X_n, X, on étudie une condition de convergence de (X_n) vers X, qui est moins restrictive que la convergence faible dans L^1. On donne ensuite de très faibles conditions (ne portant que sur les lois des X_n et sur la loi de X) à ajouter à la condition précédente pour la transformer en la convergence forte dans L^1.

Comme application des résultats obtenus, on expose une démonstration directe de la convergence dans L^1 pour une martingale uniformément intégrable.

1. QUELQUES LEMMES SUR L'INTÉGRABILITÉ UNIFORME

Un critère classique, qui remonte à la Vallée Poussin (voir [1, p.38]), affirme que, pour qu'une suite X_n de variables aléatoires réelles soit uniformément intégrable, il faut et il suffit qu'il existe une application G de $\mathbb{R}_+$ dans $\mathbb{R}_+$, telle que l'on ait $\lim_{t\to\infty} G(t)/t = \infty$ et

(1.1) $$\sup_n E[G\circ|X_n|] < \infty.$$

La partie, de ce critère, concernant la nécessité de la condition peut être ainsi renforcée:

(1.2) LEMME. *Soit* (X_n) *une suite de variables aléatoires réelles uniformément intégrables. Il existe alors un homéomorphisme strictement convexe* F *de* $\mathbb{R}_+$ *sur* $\mathbb{R}_+$, *admettant une dérivée non bornée, tel que la suite* $(F\circ|X_n|)$ *soit uniformément intégrable.*

DÉMONSTRATION. On sait (voir [1, p.39]) qu'il existe une fonction réelle g définie dans $\mathbb{R}_+$, nulle en 0, croissante, non bornée, constante sur chaque intervalle de la forme $[k,k+1[$ (avec k entier), telle qu'en posant

$$G(x) = \int_0^x g(t)dt \quad \text{pour } x \geq 0,$$

la relation (1.1) soit vérifiée. Il est clair qu'on peut construire un homéomorphisme f de $\mathbb{R}_+$ sur $\mathbb{R}_+$ qui soit majoré par g+1. En outre, quitte à remplacer f par log(1+f), on peut supposer que f(x)/g(x) converge vers 0 lorsque x tend vers l'infini. Posons

$$F(x) = \int_0^x f(t)dt \qquad \text{pour } x \geq 0.$$

La fonction F est alors un homéomorphisme strictement convexe de $\mathbb{R}_+$ sur $\mathbb{R}_+$ qui répond à la question. En effet, étant donné $\varepsilon > 0$, on peut trouver une constante réelle positive c telle que l'on ait $\{ F > c \} \subset \{ F \leq \varepsilon G \}$. On a alors, pour tout n,

$$\int_{\{F\circ|X_n|>c\}} F\circ|X_n| \, dP \leq \varepsilon \, E[G\circ|X_n|] \leq \varepsilon \, M,$$

où M désigne le premier membre de (1.1). Puisque ε M est arbitrairement petit, cela prouve que la suite $(F\circ|X_n|)$ est uniformément intégrable.

Voilà un autre critère d'intégrabilité uniforme qui nous sera utile dans la suite:

(1.3) LEMME. *Étant donnée une suite* (X_n) *de variables aléatoires réelles intégrables, posons, pour tout nombre réel* t, $p(t) = \limsup_n E[(X_n-t)^+]$.

Alors, pour que les X_n^+ *soient uniformément intégrables, il faut et il suffit que l'on ait* $\lim_{t \to +\infty} p(t) = 0$.

DÉMONSTRATION. L'assertion résulte aussitôt de la double inégalité suivante (valable pour tout couple t,x de nombres réels, avec t positif)

$$(x-t)^+ \leq xI_{]t,\infty[}(x) \leq 2(x - t/2)^+.$$

2. UNE CARACTÉRISATION DE LA CONVERGENCE EN LOI

(2.1) PROPOSITION. *Soit* D *un ensemble partout dense dans* $\mathbb{R}$, *et soient* X_n,X *des variables aléatoires réelles intégrables.*

Les conditions suivantes sont équivalentes:

(a) *La suite* (X_n) *converge en loi vers* X, *et les* X_n^+ *sont uniformément intégrables.*

(b) *Pour tout élément* t *de* D, *on a*

(2.2) $$\lim_n E[(X_n-t)^+] = E[(X-t)^+].$$

DÉMONSTRATION. Il suffit de prouver que (b) entraîne (a). Supposons donc la condition (b) remplie. L'inégalité élémentaire

$$|(x-t)^{+}-(x-s)^{+}| \le |t-s|$$

montre alors que la relation (2.2) a lieu pour tout nombre réel t. Soit f un élément de $\mathscr{C}_c^2(\mathbb{R},\mathbb{R})$. En intégrant par parties, on trouve

$$f(x) = \int_{\mathbb{R}} f''(t)(x-t)^{+}dt.$$

Il en résulte

$$E[f(X)] = \int_{\mathbb{R}} f''(t)E[(X-t)^{+}]dt = \lim_n \int_{\mathbb{R}} f''(t)E[(X_n-t)^{+}]dt = \lim_n E[f(X_n)],$$

et cela montre que (X_n) converge en loi vers X.

Enfin, pour prouver l'intégrabilité uniforme des X_n^{+}, il suffit, grâce au Lemme (1.3), de faire tendre t vers $+\infty$ dans la relation (2.2).

(2.3) REMARQUE. En appliquant la proposition qu'on vient de démontrer aux variables aléatoires $-X_n$, $-X$, on obtient un résultat analogue, où la relation (2.2) est remplacée par la relation

$$(2.4) \qquad \lim_n E[(X_n-t)^{-}] = E[(X-t)^{-}],$$

et où l'intégrabilité uniforme des X_n^{+} est remplacée par celle des X_n^{-}.

3. UNE CARACTÉRISATION DE LA CONVERGENCE EN PROBABILITÉ

(3.1) PROPOSITION. *Étant données, sur l'espace probabilisé* $(\Omega,\mathcal{A},P)$, *les variables aléatoires réelles* X_n, X, *les conditions suivantes sont équivalentes:*

(a) *La suite* (X_n) *converge en probabilité vers* X.

(b) *La suite* (X_n) *converge en loi vers* X *selon toute mesure de probabilité de la forme* $Q = f(X).P$, *avec* f *élément de* $\mathscr{C}_c(\mathbb{R},\mathbb{R})$.

(c) *La suite des vecteurs aléatoires* (X,X_n) *converge en loi vers le vecteur aléatoire* (X,X).

DÉMONSTRATION. L'implication (a) $\Longrightarrow$ (b) résulte aussitôt du fait que la convergence presque sûre (donc aussi la convergence en probabilité) n'est pas détruite par passage à une mesure de probabilité Q absolument continue par rapport à P.

Pour prouver que (b) entraîne (c), remarquons que l'hypothèse (b) signifie que l'on a

$$\lim_n E[f(X)g(X_n)] = E[f(X)g(X)]$$

pour tout couple f,g d'éléments de $\mathscr{C}_c(\mathbb{R},\mathbb{R})$, ce qui équivaut à la convergence vague de la loi du vecteur aléatoire (X,X_n) vers la loi du vecteur aléatoire (X,X).

Enfin, l'implication (c) $\Longrightarrow$ (a) est évidente, car l'hypothèse (c) entraîne que la suite $(|X_n-X|)$ converge en loi, donc en probabilité, vers la constante $|X-X|=0$.

La proposition qu'on vient de démontrer permet de déduire de (2.1) le corollaire suivant:

(3.2) COROLLAIRE. *Soit* D *un ensemble partout dense dans* $\mathbb{R}$, *et soient* X_n,X *des variables aléatoires réelles intégrables.*

Les conditions suivantes sont équivalentes:

(a) *La suite* (X_n) *converge en probabilité vers* X, *et les* X_n^+ *sont uniformément intégrables.*

(b) *Pour tout élément* t *de* D, *la mesure image* $X((X_n-t)^+.P)$ *converge étroitement vers* $X((X-t)^+.P)$.

DÉMONSTRATION. Il suffit en effet de remarquer que la condition (b) équivaut au fait que, pour tout élément t de D, la suite (X_n) vérifie à la fois la relation (2.2) et la relation suivante:

$$\lim_n E[f(X)(X_n-t)^+] = E[f(X)(X-t)^+] \quad \text{pour } f \in \mathscr{C}_c(\mathbb{R},\mathbb{R}).$$

4. UNE CONDITION MOINS RESTRICTIVE QUE LA CONVERGENCE FAIBLE DANS L^1

Étant données des variables aléatoires réelles intégrables X_n,X, on peut rechercher des conditions minimales à ajouter à la convergence faible de (X_n) vers X dans L^1 pour la transformer en la convergence forte dans L^1.

Nous nous proposons de résoudre ce problème en remplaçant la convergence faible dans L^1 par une condition encore plus faible, à savoir:

(4.1) $$\lim_n E[f(X)X_n] = E[f(X)X] \quad \text{pour } f \in \mathscr{C}_c(\mathbb{R},\mathbb{R}).$$

Cette condition équivaut à la convergence vague de la mesure image $X(X_n.P)$ vers $X(X.P)$. Il s'agit d'une condition effectivement moins restrictive que la convergence faible de (X_n) vers X dans L^1: en effet, elle ne suffit même pas à garantir que la suite des espérances des X_n soit bornée. En outre, même si on la renforce en y ajoutant l'intégrabilité uniforme des X_n (c'est-à-dire la compacité relative faible de l'ensemble des X_n), on n'obtient pas encore une condition équivalente à la convergence faible de (X_n) vers X. On a cependant le résultat suivant:

(4.2) PROPOSITION. *Supposons que les variables aléatoires réelles* X_n, X *vérifient la condition* (4.1) *et que les* X_n *soient uniformément intégrables. On a alors* $\lim_n E[X_n] = E[X]$.

DÉMONSTRATION. Fixons $\varepsilon > 0$. Puisque les X_n sont uniformément intégrables, il existe une constante réelle positive c telle que l'on ait

$$\sup_n \int_{\{|X|>c\}} (|X|+|X_n|)\, dP \le \varepsilon.$$

Choisissons un élément f de $\mathscr{C}_c(\mathbb{R},\mathbb{R})$, avec $I_{[-c,c]} \le f \le 1$. On a alors, pour tout n,

$$| E[(X_n-X)] | \le | E[(X_n-X)f(X)] | + \varepsilon,$$

d'où la conclusion, grâce à (4.1).

Voilà une autre conséquence de la condition (4.1):

(4.3) PROPOSITION. *Supposons que les variables aléatoires réelles intégrables* X_n, X *vérifient la condition* (4.1). *On a alors*

(4.4) $$E[f(X)(X-t)^+] \le \liminf_n E[f(X)(X_n-t)^+],$$

(4.5) $$E[f(X)(X-t)^-] \le \liminf_n E[f(X)(X_n-t)^-]$$

pour tout nombre réel t *et toute fonction* f *positive et semi-continue inférieurement sur* $\mathbb{R}$.

DÉMONSTRATION. Pour prouver l'inégalité (4.4), il suffit de remarquer que, pour tout élément g de $\mathscr{C}_c(\mathbb{R},\mathbb{R})$, avec $0 \le g \le fI_{]t,\infty[}$, on a, grâce à l'hypothèse (4.1),

$$\begin{aligned} E[g(X)(X-t)^+] &= E[g(X)(X-t)] = \lim_n E[g(X)(X_n-t)] \\ &\le \liminf_n E[f(X)(X_n-t)^+]. \end{aligned}$$

L'autre inégalité peut être démontrée de manière analogue (ou bien ramenée à la précédente par passage aux variables aléatoires $-X_n, -X$).

(4.6) COROLLAIRE. *Supposons que les variables aléatoires réelles intégrables* X_n, X *vérifient la condition* (4.1). *Alors, pour tout nombre réel* *t tel que l'on ait*

(4.7) $$E[(X-t)^+] = \lim_n E[(X_n-t)^+],$$

la mesure image $X((X_n-t)^+.P)$ *converge étroitement vers* $X((X-t)^+.P)$.

DÉMONSTRATION. L'hypothèse supplémentaire (4.7) permet en effet de renforcer la conclusion de la proposition précédente en affirmant que la relation (4.4) a lieu pour toute fonction f *minorée* et semi-continue inférieurement.

5. LE RÉSULTAT PRINCIPAL: UNE CARACTÉRISATION DE LA CONVERGENCE DANS L^1

Le théorème suivant, qui constitue le résultat principal du présent article, fournit un certain nombre de réponses à la question posée au début du paragraphe précédent.

(5.1) THÉORÈME. *Supposons que les variables aléatoires réelles intégrables* X_n, X *vérifient la condition suivante (moins restrictive que la convergence faible de* (X_n) *vers* X *dans* L^1*):*

$$\lim_n E[f(X)X_n] = E[f(X)X] \quad \textit{pour } f \in \mathscr{C}_c(\mathbb{R},\mathbb{R}).$$

Alors, pour que la suite (X_n) *converge vers* X *au sens fort dans* L^1, *il faut et il suffit qu'elle vérifie l'une quelconque des trois conditions supplémentaires suivantes:*

(a) *La suite* (X_n) *est uniformément intégrable et converge en loi vers* X.

(b) *On a* $\lim_n E[X_n] = E[X]$, *et il existe une fonction réelle* g, *strictement convexe sur* $\mathbb{R}$, *vérifiant la relation*

(5.2) $$\lim_n E[g(X_n)] = E[g(X)] < \infty.$$

(c) *Il existe un ensemble* D, *partout dense dans* $\mathbb{R}$, *tel que l'on ait*

(5.3) $$\lim_n E[|X_n-t|] = E[|X-t|]$$

pour tout élément t *de* D.

DÉMONSTRATION. Il est évident que, si la suite (X_n) converge vers X au sens fort dans L^1, elle vérifie la condition (a).

Montrons que (a) entraîne (b). Remarquons, à cet effet, que si l'on prend F comme dans le Lemme (1.2), la convergence en loi de la suite $(F\circ|X_n|)$ vers $F\circ|X|$ entraîne que l'on a

$$E[F\circ|X|] = \lim_n E[F\circ|X_n|] < \infty,$$

de sorte que la condition (b) est remplie, avec $g(x) = F(|x|)$.

Montrons maintenant que (b) implique (c). Supposons donc la condition (b) remplie. Quitte a retrancher de g une fonction linéaire affine, on pourra supposer inf g = 0. Trois cas sont possibles: ou bien g est strictement croissante, ou bien g est strictement décroissante, ou bien g s'annule en un point c. (Dans ce dernier cas, g est strictement décroissante sur $]-\infty,c]$, strictement croissante sur $[c,\infty[$). En tout cas, g peut se mettre sous la forme

(5.4) $$g(x) = \int \lambda(dt)h(t,x),$$

où λ est une mesure de Radon positive sur $\mathbb{R}$, admettant $\mathbb{R}$ comme support, et où h est une fonction borélienne sur $\mathbb{R}^2$, telle que, pour tout t, la fonction $h(t,\cdot)$ coïncide avec l'une des deux fonctions

$$x \longmapsto (x-t)^+, \qquad x \longmapsto (x-t)^-.$$

En vertu de (4.3), on a, pour tout t,

(5.5) $$E[h(t,X)] \le \liminf_n E[h(t,X_n)].$$

En outre, grâce à l'hypothèse $\lim_n E[X_n]=E[X]$, la relation

(5.6) $$E[h(t,X)] = \lim_n E[h(t,X_n)]$$

est équivalente à chacune des relations (2.2),(2.4).

En utilisant (5.4),(5.5) et (5.2), on obtient

$$\begin{aligned} E[g(X)] &= \int \lambda(dt)\, E[h(t,X)] \\ &\le \int \lambda(dt) \liminf_n E[h(t,X_n)] \\ &\le \liminf_n \int \lambda(dt)\, E[h(t,X_n)] \\ &= \liminf_n E[g(X_n)] = E[g(X)] < \infty. \end{aligned}$$

Cela prouve que la fonction

$$t \longmapsto \liminf_n E[h(t,X_n)] - E[h(t,X)]$$

(positive en vertu de (5.5)) est négligeable pour la mesure λ. L'ensemble où cette fonction est nulle contient donc un ensemble dénombrable D partout dense dans $\mathbb{R}$. Quitte à passer (par un procédé diagonal) à une suite extraite, on peut supposer que la suite (X_n) vérifie, pour tout élément t de D, la relation (5.6), c'est-à-dire les relations équivalentes (2.2),(2.4). Il en résulte que la condition (c) est remplie.

Montrons enfin que la condition (c) implique la convergence forte de (X_n) vers X dans L^1. Remarquons, à cet effet, que, pour tout élément t de D, l'hypothèse (5.3) transforme les inégalités

$$E[(X-t)^+] \le \liminf_n E[(X_n-t)^+], \qquad E[(X-t)^-] \le \liminf_n E[(X_n-t)^-]$$

(fournies par la Proposition (4.3)) en les égalités (2.2),(2.4). Il en résulte, grâce à (4.6), que la condition (b) du Corollaire (3.2) est remplie. Ce corollaire assure alors la convergence en probabilité de (X_n) vers X, ainsi que l'intégrabilité uniforme des X_n^+. Puisque le même raisonnement, appliqué aux variables aléatoires $-X_n$, $-X$, fournit l'intégrabilité uniforme des X_n^-, la démonstration est achevée.

(5.7) COROLLAIRE. *Pour qu'une suite* (X_n) *de variables aléatoires réelles intégrables converge vers une variable aléatoire réelle intégrable* X *au sens fort dans* L^1, *il faut et il suffit qu'elle converge vers* X *en loi et au sens faible dans* L^1.

(5.8) REMARQUE. Chacune des trois conditions (a),(b),(c) qui figurent dans l'énoncé du Théorème (5.1) ne porte que sur les lois des X_n et sur la loi de X. Ce théorème peut être considéré comme un perfectionnement du résultat suivant (cas particulier de [4, Th.3, p.449]): pour que la suite (X_n) converge vers X au sens fort dans L^1, il suffit qu'elle converge vers X au sens faible dans L^1 et qu'il existe une fonction réelle g, strictement convexe sur $\mathbb{R}$, telle que l'espérance de $g(X_n)$ converge vers l'espérance de g(X).

6. APPLICATION AUX QUASIMARTINGALES

On sait que toute quasimartingale (discrète) uniformément intégrable converge dans L^1. Ce résultat peut être démontré en utilisant la décomposition de Rao (voir [2],[3]) et le fait que toute surmartingale positive converge presque sûrement. Nous en exposons ici une démonstration directe, fondée exclusivement sur les résultats du présent article.

Commençons par rappeler qu'une suite (X_n) de variables aléatoires réelles intégrables sur $(\Omega,\mathcal{A},P)$, adaptée à une suite croissante $(\mathcal{F}_n)$ de sous-tribus de $\mathcal{A}$, est dite une *quasimartingale* (par rapport à $(\mathcal{F}_n)$) si la suite (D_n) définie par

$$D_n = E[\, X_n - X_{n+1} \mid \mathcal{F}_n]$$

vérifie la relation $\Sigma_n \parallel D_n \parallel_1 < \infty$.

Étant donnée la quasimartingale (X_n), on reconnaît immédiatement que l'on a

$$\left| \int_A (X_n - X_{n+k})dP \right| \le \Sigma_{n\le j<n+k} \parallel D_j \parallel_1$$

pour tout couple n,k d'entiers positifs et tout élément A de $\mathcal{F}_n$.

Il en résulte que, pour tout élément A de l'algèbre $\bigcup_m \mathcal{F}_m$, la suite de nombres réels ($\int_A X_n dP$) est de Cauchy. Par conséquent, si la quasimartingale (X_n) est uniformément intégrable, elle admet une seule valeur d'adhérence pour la topologie faible, de sorte qu'elle converge, pour cette topologie, vers une variable aléatoire intégrable X. Pour prouver que (X_n) converge vers X au sens fort dans L^1, il reste à montrer (voir (5.7)) que (X_n) converge vers X en loi, c'est-à-dire que la condition (b) de la Proposition (2.1) est remplie. Or ceci résulte aussitôt des relations suivantes:

$$E[(X-t)^+] = \lim_n \int_{\{X>t\}}(X_n-t)dP \le \liminf_n E[(X_n-t)^+],$$

$$E[(X_n-t)^+] = \int_{\{X_n>t\}}(X_n-X)dP + \int_{\{X_n>t\}}(X-t)dP \le \Sigma_{j\ge n}\| D_j \|_1 + E[(X-t)^+].$$

La démonstration est donc achevée.

(6.1) REMARQUE. Le résultat qu'on vient de démontrer entraîne à son tour la convergence presque sûre de la quasimartingale (X_n), grâce à l'inégalité suivante (valable pour tout $\varepsilon > 0$ et tout couple n,k d'entiers positifs):

(6.2) $$\varepsilon P\{\sup_{1\le h\le k}|X_n-X_{n+h}|>\varepsilon\} \le \| X_n-X_{n+k} \|_1 + \Sigma_{n<j<n+k}\| D_j \|_1.$$

Pour démontrer cette inégalité, fixons l'entier n et posons

$$T = \inf\{h:\ h\ge 1,\ |X_n-X_{n+h}|>\varepsilon\}, \qquad J_h = I_{\{X_n > X_{n+h}\}} - I_{\{X_n < X_{n+h}\}}.$$

On a alors, pour tout entier h compris entre 1 et k,

$$\varepsilon P\{T=h\} \le \int_{\{T=h\}}|X_n-X_{n+h}|dP = \int_{\{T=h\}}J_h(X_n-X_{n+h})dP$$

$$= \int_{\{T=h\}}J_h(X_n-X_{n+k} - \Sigma_{n+h\le j<n+k} D_j)dP$$

$$\le \int_{\{T=h\}}(|X_n-X_{n+k}| + \Sigma_{n<j<n+k} |D_j|)dP.$$

Pour obtenir l'inégalité (6.2), il ne reste plus qu'à sommer sur h=1,...,k.

BIBLIOGRAPHIE

[1] C. Dellacherie et P.-A. Meyer, *Probabilités et potentiel*, Chap. I-IV, Paris 1975.

[2] C. Dellacherie et P.-A. Meyer, *Probabilités et potentiel*, Chap. V-VIII, Paris 1980.

[3] K.M. Rao, *Quasimartingales*, Math. Scand. 24, 1969, 79-92.

[4] A. Visintin, *Strong convergence results related to strict convexity*, Comm. Partial Differential Equations, 9(5), 1984, 439-466.