C. R. Acad. Sci. Paris, t. 333, Série I, p. 673–676, 2001
Statistique/*Statitics*
(Probabilités/*Probability Theory*)

Principe local d'invariance pour des variables aléatoires i.i.d.

Jean-Christophe BRETON, Youri DAVYDOV

Laboratoire de statistique et probabilités, UFR de mathématiques, M2, Université Lille-1, 59655 Villeneuve-d'Ascq cedex, France
Courriel : breton@jacta.univ-lille1.fr; davydov@jacta.univ-lille1.fr

(Reçu le 5 juin 2001, accepté le 22 août 2001)

Résumé. Pour une suite de variables aléatoires indépendantes identiquement distribuées, les processus normalisés constants par morceaux associés convergent faiblement vers le processus de Wiener. On renforce pour les distributions fonctionnelles la convergence pour la distance en variation pour une classe de fonctionnelles. © 2001 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Local invariance principle for i.i.d. random variables

Abstract. *For a sequence of independent identically distributed random variables, the normalized step-processes associated are weakly convergent to the Wiener process. We strengthen for the functional distributions the convergence for the variation distance for a class of functionals.* © 2001 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

1. Résultat

1.1. *Introduction.* – Soit $(\xi_n,\ n \in \mathbb{N})$ une suite de variables aléatoires indépendantes identiquement distribuées, centrées et de variance 1, définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. On note F la fonction de répartition. On associe le processus constant par morceaux :

$$S_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \xi_i, \quad t \in [0,1]. \tag{1}$$

D'après le théorème de Donsker–Prokhorov, en notant P_n, P les lois de S_n et du processus de Wiener W, en désignant par $\Rightarrow$ la convergence faible dans l'espace de Skorokhod $\mathbb{D}$ des fonctions *cadlag*, on a pour chaque fonctionnelle f, P-presque partout continue :

$$P_n f^{-1} \Longrightarrow P f^{-1}. \tag{2}$$

On propose un principe local en renforçant la convergence (2) en une convergence forte (i.e. pour la distance en variation des mesures signées). Si les lois $P_n f^{-1}$, $P f^{-1}$ sont absolument continues par rapport

Note présentée par Paul DEHEUVELS.

S0764-4442(01)02112-7/FLA
© 2001 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS. Tous droits réservés

à la mesure de Lebesgue λ, la convergence en variation est équivalente à la convergence des densités $\mathrm{d}P_n f^{-1}/\mathrm{d}\lambda$ vers $\mathrm{d}P f^{-1}/\mathrm{d}\lambda$ pour la métrique de $\mathrm{L}^1(\mathbb{R})$. La difficulté principale réside dans la singularité pour chaque n des mesures P_n et P.

1.2. *Théorème. –* Un premier résultat en ce sens est donné dans [4], théorème 20.1 : en imposant à la loi commune des ξ_i d'avoir une densité p absolument continue d'information de Fisher $I_p = \int_{\mathbb{R}} \frac{p'^2}{p}\,\mathrm{d}\lambda$ finie, la convergence forte est obtenue pour f dans une classe de fonctionnelles $\mathcal{M}_P$ décrite dans [4], § 19. La condition $I_p < \infty$ demande que p soit à variation bornée. On propose d'affaiblir cette condition en exigeant un moment supérieur à 2 pour ξ_i et en considérant f dans une classe de fonctionnelles restreinte mais encore assez large. Pour cela, notons $H_P := \big\{\ell \mid \ell(t) = \int_0^t \ell'(s)\,\mathrm{d}s,\ \ell' \in \mathrm{L}^2([0,1])\big\}$ l'espace de Cameron–Martin qui coïncide avec l'ensemble des directions admissibles pour P, on considère :

DÉFINITION 1. – $\mathcal{M}_P^{(1)}$ est l'ensemble des fonctionnelles f localement lipschitziennes telles que pour P-presque chaque x, il existe un voisinage $V(x)$ de x et $\ell \in H_P$ tels que :
– la dérivée $\mathrm{D}_\ell f(x)$ de f en x selon ℓ existe et est non nulle ;
– en notant $S_y = \{h \mid \|h\| = 1$ et $\mathrm{D}_h f(y)$ existe$\}$ et $A = \bigcup_{y \in V(x)} \{y\} \times S_y$, on a $(y,h) \in A \to \mathrm{D}_h f(y)$ bornée et continue.

Notons $t_- = \sup\{t \mid F(t) = 0\}$, $t_+ = \inf\{t \mid F(t) = 1\}$, on a alors :

THÉORÈME 1. – *Soit* $\{\xi_n,\ n \in \mathbb{N}\}$ *une suite de variables aléatoires indépendantes identiquement distribuées centrées de variance* 1, *de fonction de répartition* F. *On suppose qu'il existe* $\gamma > 0$ *tel que* $\xi_1 \in \mathrm{L}^{2+\gamma}(\Omega, \mathcal{F}, \mathbb{P})$, *on suppose que* ξ_1 *admet une densité* p *non nulle p.p. sur* $[t_-, t_+]$. *Alors, pour toute fonctionnelle* $f \in \mathcal{M}_P^{(1)}$, *on a* :

$$P_n f^{-1} \xrightarrow{\mathrm{var}} P f^{-1}. \tag{3}$$

Le théorème 1 s'applique pour des fonctionnelles de type *sup*, $x \mapsto \sup_{t \in [0,1]} \varphi(x(t))$, et *intégrale*, $x \mapsto \int q(x(t))\,\mathrm{d}t$, en supposant φ convexe, q lipschitzienne sur les compacts, de dérivées p.p. non nulles.

2. Schéma de la preuve

2.1. *Idée générale. –* La démonstration du théorème 1 repose sur la méthode dite de *superstructure* qui a été utilisée dans [4] : on applique un découpage de l'espace pour exprimer les distributions fonctionnelles comme mélanges de distributions conditionnelles. La nouveauté principale, qui a permis d'avancer dans ce problème, consiste à représenter les variables aléatoires initiales ξ_n comme des v.a. orthogaussiennes η_n transformées $\xi_n = U(\eta_n)$ puis à considérer les processus $S_n(t)$ en fonction de processus analogues $Z_n(t)$ définis par $(\eta_n)_n$. Cela implique qu'en place des translations admissibles de $S_n(t)$, utilisées dans la preuve de [4], théorème 20.1, on se sert de transformations non linéaires induites par des translations admissibles de $Z_n(t)$. L'analyse du comportement asymptotique des lois conditionnelles correspondant devient plus compliquée et demande des outils supplémentaires.

On considère dans la suite $\mathbb{E}$ le sous-espace de $\mathbb{D}$ fermeture pour la norme uniforme de l'ensemble des fonctions de $\mathbb{D}$ qui ont un nombre fini de sauts en des points rationnels. L'espace $\mathbb{E}$, muni de la norme uniforme $\|\cdot\|$, est un espace de Banach séparable où la convergence faible (2) tient.

La preuve détaillée du théorème 1 est donnée dans [2], on indique dans la suite les idées essentielles : elle consiste en l'étude des hypothèses du résultat général suivant sur la convergence forte de distributions fonctionnelles :

THÉORÈME A ([4], théorème 18.4). – *On considère une suite de mesures de probabilité* $\big\{P_n,\ n \in \overline{\mathbb{N}}\big\}$ *définies sur la* σ*-algèbre borélienne* $\mathcal{B}_X$ *d'un espace métrique complet séparable* (X,d). *On suppose* $P_n \Rightarrow P_\infty$ *et que pour* P_∞*-presque chaque* x, *il existe une boule ouverte* V *centrée en* x, $\varepsilon > 0$ *et une famille* $\big\{G_{n,c},\ n \in \overline{\mathbb{N}},\ c \in (0,\varepsilon]\big\}$ *de transformations mesurables de* X *telles que* :

(i) *pour tout* $c \in (0,\varepsilon)$*, et pour tout* $\delta > 0$*, on a* $\lim_{n\to\infty} P_n\{x \mid |G_{n,c}x - G_{\infty,c}x| \geqslant \delta\} = 0$;

(ii) *pour tout* $c \in (0,\varepsilon)$*, l'application* $G_{\infty,c}$ *est* P_∞*-presque partout continue. De plus, pour toute boule ouverte* S*,* $\rho(S,c) = \sup_{z\in S} \mathrm{d}(z, G_{\infty,c}z) \to 0$ *quand* $c \to 0$;

(iii) $\lim_{c\to 0} \overline{\lim}_{n\to\infty} \|P_n G_{n,c}^{-1} - P_n\| = 0$;

(iv) *pour tout* $\delta \in (0,\varepsilon)$*, avec* $\varphi_{n,z}(c) = f(G_{n,c}z)$*,* $n \in \overline{\mathbb{N}}$*,* $c \in (0,\varepsilon]$*, on a*

$$\int_V \big\|\lambda_{[0,\delta]}\varphi_{n,z}^{-1} - \lambda_{[0,\delta]}\varphi_{\infty,z}^{-1}\big\| \, P_n(\mathrm{d}z) \;\longrightarrow\; 0 \quad \textit{quand } n \to \infty;$$

(v) *pour tout* $\delta \in (0,\varepsilon)$*,* $z \mapsto \lambda_{[0,\delta]}\varphi_{\infty,z}^{-1}$ *de* V *dans l'espace des mesures signées est* P_∞*-p.p. continue.*

Alors $P_n f^{-1} \xrightarrow{\mathrm{var}} P_\infty f^{-1}$.

2.2. *Outils utiles.* – On commence par expliciter les outils utiles pour appliquer le théorème A. Comme $f \in \mathcal{M}_P^{(1)}$, pour P-presque chaque x, on dispose d'un voisinage $V_1(x) = \mathrm{B}(x, r_1)$ et d'une direction admissible $\ell \in H_P$ à partir de laquelle on construit les transformations $\{G_{n,c}\}_c$. On trouve alors facilement $V(x) = \mathrm{B}(x,r)$, $P\{\partial V(x)\} = 0$ tels que :

$$\forall (y,h) \in \mathrm{B}(x,r) \times \mathrm{B}\big(\ell/\|\ell\|, r\big) \cap A, \quad \mathrm{D}_h f(y) \geqslant \tfrac{1}{2}\mathrm{D}_\ell f(x) > 0 \quad \text{(quitte à changer } f \text{ en } -f).$$

Soient F la fonction de répartition de ξ_n, $F^{-1}(x) = \inf\{y \mid F(y) \geqslant x\}$ l'inverse généralisé de F et Φ, q les fonction de répartition et densité de $\mathcal{N}(0,1)$. Notons $U = F^{-1} \circ \Phi$, $V = \Phi^{-1} \circ F$, avec $(\eta_n)_n$ une suite orthogaussienne, on peut exprimer $\xi_n = U(\eta_n)$.

Avec $\mathbb{E}_n := \big\{x \in \mathbb{E} \mid x$ constante par morceaux sur $\big[\frac{k}{n}, \frac{k+1}{n}\big)$, $k = 0, \dots, n-1\big\}$, notons Π_n la surjection canonique $\mathbb{E} \to \mathbb{E}_n$ et J_n l'isométrie naturelle de $\mathbb{E}_n \to \mathbb{R}^n$. Avec, pour $\ell \in H_P$ donnée par la définition de $\mathcal{M}_P^{(1)}$, $\bar\ell_n = J_n \circ \Pi_n(\ell) = (\ell_{n,i})_{i\leqslant n}$ où

$$\ell_{n,i} = \sqrt{n}\left(\ell\left(\frac{i}{n}\right) - \ell\left(\frac{i-1}{n}\right)\right), \tag{4}$$

on définit les trois transformations suivantes de $\mathbb{R}^n$:

$$\varphi(\bar x) = \big(V(x_1), \dots, V(x_n)\big), \quad \psi(\bar x) = \big(U(x_1), \dots, U(x_n)\big), \quad G_{n,c}(\bar x) = \bar x + c\bar\ell_n.$$

En notant $a = \int q(x)\,\mathrm{d}U(x)$, on définit alors les transformations utilisées dans le théorème A par

$$G_{n,c} = J_n^{-1} \circ \psi \circ \overline{G}_{n,c} \circ \varphi \circ J_n \circ \Pi_n, \qquad G_{\infty,c}x = x + c\,a\,\ell. \tag{5}$$

Remarque 1. – Le résultat cherché concernant les lois P_n, P, on peut appliquer le théorème de Skorokhod ([1], théorème 6.7) et supposer travailler avec un espace probabilisé pour lequel on a la convergence p.s. $S_n \to W$, on doit alors supposer avoir, pour chaque n, $(\eta_i^n)_{i\leqslant n}$ i.i.d. $\mathcal{N}(0,1)$ et $S_n(\cdot) = \frac{1}{\sqrt{n}}\sum_{i=1}^{[n\cdot]} U(\eta_i^n)$.

2.3. *Étude des hypothèses du théorème* A. – Le point (i) consiste à voir que pour tout $\alpha > 0$,

$$\mathbb{P}\big\{|G_{n,c}S_n - G_{\infty,c}S_n| \geqslant \alpha\big\} \longrightarrow 0. \tag{6}$$

Or,

$$G_{n,c}S_n(\cdot) = \frac{1}{\sqrt{n}}\sum_{i\leqslant[n\cdot]} U\big(\eta_i^n + c\,\ell_{n,i}\big), \qquad G_{\infty,c}S_n(\cdot) = \frac{1}{\sqrt{n}}\sum_{i\leqslant[n\cdot]} U\big(\eta_i^n\big) + c\,a\,\ell(\cdot).$$

Comme $\left\|\ell\left(\frac{[n\cdot]}{n}\right)-\ell(\cdot)\right\| \to 0$ quand $n\to\infty$, pour voir (6), il suffit de voir que

$$\max_{k\leqslant n}\left|\frac{1}{\sqrt{n}}\sum_{i\leqslant k} U\left(\eta_i^n + c\,\ell_{n,i}\right) - U\left(\eta_i^n\right) - c\,a\,\ell_{n,i}\right| \xrightarrow{\mathbb{P}} 0.$$

Pour cela on montre que

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left|E\left(U\left(\eta_i^n + c\,\ell_{n,i}\right) - U\left(\eta_i^n\right) - c\,a\,\ell_{n,i}\right)\right| \longrightarrow 0,$$

$$\overline{\lim}_n \max_{k\leqslant n}\mathbb{P}\left\{\frac{1}{\sqrt{n}}\left|\sum_{i=k}^{n} U\left(\eta_i^n + c\,\ell_{n,i}\right) - U\left(\eta_i^n\right) - E\left(U\left(\eta_i^n + c\,\ell_{n,i}\right) - U\left(\eta_i^n\right)\right)\right| > \alpha\right\} = 0.$$

Le point (ii) est clair.

Pour voir (iii), on constate d'abord que $\|P_n G_{n,c}^{-1} - P_n\| \leqslant \|\mathcal{P}_n \overline{G}_{n,c}^{-1} - \mathcal{P}_n\|$ où $\mathcal{P}_n$ est la loi gaussienne standard de dimension n. On conclut alors en utilisant le lemme 20.1 de [4] : il donne

$$\left\|\mathcal{P}_n \overline{G}_{n,c}^{-1} - \mathcal{P}_n\right\| \leqslant c\, I_q^{1/2}\|\ell'\|_2,$$

ce qui garantit (iii).

Pour le point (iv), on passe par l'étude des deux suites ci-dessous :

$$g_n(\omega,c) = f(G_{n,c}S_n), \quad g_\infty(\omega,c) = f(W + c\,a\,\ell); \qquad h_n(\omega,c) = f(G_{\infty,c}S_n), \quad h_\infty = g_\infty.$$

Pour cela, on applique à ces deux suites le résultat suivant qui se justifie avec le théorème 1 de [3] :

PROPOSITION 1. – *Soient pour* $n\in\overline{\mathbb{N}}$, $f_n : (\Omega\times[0,\delta], \mathcal{F}\times\mathcal{B}([0,\delta]), \mathbb{P}\otimes\bar{\lambda}) \to \mathbb{R}$, $\Omega^*\in\mathcal{F}$, $\Omega^*\subset\Omega$ *tels que* :

1. *pour tout* $\omega\in\Omega^*$*, il existe* $N_1(\omega)$*, tel que, pour tout* $n\geqslant N_1(\omega)$, $f_n(\omega,\cdot)$ *est absolument continue* ;
2. $f_n(\omega,0) \xrightarrow{\mathbb{P}} f_\infty(\omega,0)$ *sur* Ω^* ;
3. *pour tout* $\omega\in\Omega^*$, $\frac{\partial}{\partial c} f_\infty(\omega,c) > 0$ *pour* λ*-presque chaque* $c\in(0,\delta)$;
4. $\left\|\frac{\partial}{\partial c} f_n(\omega,c) - \frac{\partial}{\partial c} f_\infty(\omega,c)\right\|_{\mathrm{L}^1} \xrightarrow{\mathbb{P}} 0$ *sur* Ω^* ;

alors $\left\|\lambda_{[0,\delta]} f_n(\omega,\cdot)^{-1} - \lambda_{[0,\delta]} f_\infty(\omega,\cdot)^{-1}\right\| \xrightarrow{\mathbb{P}} 0$ *sur* Ω^*.

Les études de 3, 4 passent par une analyse approfondie du vecteur tangent à la trajectoire. La conclusion de la proposition 1 pour $(g_n)_n$ et $(h_n)_n$ assure alors le point (iv) du théorème A.

Le point (v) s'étudie facilement avec la proposition 1 déjà utilisée précédemment.

Finalement, le théorème A s'applique et donne la convergence (3). □

Références bibliographiques

[1] Billingsley P., Convergence of Probability Measures, 2nd edition, Wiley, 1968.

[2] Breton J.-C., Davydov Y.A., Principe local d'invariance pour des variables aléatoires i.i.d., Publ. IRMA Lille 55 (II) (2001).

[3] Davydov Y.A., On convergence in variation of one-dimensional image measures, J. Math. Sci. 75 (5) (1995) 1903–1909.

[4] Davydov Y.A., Lifshits M.A., Smorodina N.V., Local properties of distributions of stochastic functionals, Amer. Math. Soc. 173 (1998).