Un Théorème Limite Centrale empirique dans L1 pour des suites de variables aléatoires stationnaires.

Sophie Dede

Un Théorème limite central empirique dans $\mathbf{L}^1$ pour des suites de variables aléatoires stationnaires.

Sophie Dede

LPMA, UPMC Université Paris 6, Case courrier 188,
4, Place Jussieu, 75252 Paris Cedex 05, France.
E-mail:sophie.dede at upmc.fr

Résumé. Dans ce papier, nous obtenons des résultats asymptotiques pour la distance $\mathbf{L}^1$ de Wasserstein entre la fonction de répartition et la fonction de répartition empirique d'une suite de variables aléatoires stationnaires. Ensuite, nous donnons en exemples, des applications aux systèmes dynamiques et aux processus linéaires causaux. Pour prouver notre résultat principal, nous donnons un Théorème limite central pour des suites ergodiques stationnaires de variables aléatoires à valeurs dans $\mathbf{L}^1$. Les conditions obtenues sont de type projectives. Un des principaux outils utilisés est l'approximation par des différences de martingales.

Abstract. In this paper, we derive asymptotic results for the $\mathbf{L}^1$-Wasserstein distance between the distribution function and the corresponding empirical distribution function of a stationary sequence. Next, we give some applications to dynamical systems and causal linear processes. To prove our main result, we give a Central Limit Theorem for ergodic stationary sequences of random variables with values in $\mathbf{L}^1$. The conditions obtained are expressed in terms of projective-type conditions. The main tools are martingale approximations.

Mots clés: statistique des processus-séries temporelles, Modèles semi et non paramétriques

1 Introduction

La distance de Kantorovich ou la distance $\mathbf{L}^1$ de Wasserstein entre deux mesures de probabilité P_1 et P_2 sur $\mathbb{R}$, et d'espérance finie, est définie par

$$d_1(P_1, P_2) := \inf \Big\{ \int |x-y| \, d\nu(x,y) : \ \nu \in \mathcal{P}(\mathbb{R}^2) \text{ de marginales } P_1, \, P_2 \Big\},$$

où $\mathcal{P}(\mathbb{R}^2)$ est l'espace des mesures de probabilité sur $\mathbb{R}^2$.

Soit Λ_1 l'espace des fonctions 1-Lipschitziennes. Il est connu que d_1 peut s'écrire également,

$$d_1(P_1, P_2) = \int |F_2(t) - F_1(t)|\, dt = \sup_{f \in \Lambda_1} \Big| \int f dP_1 - \int f dP_2 \Big|,$$

avec F_1 (respectivement F_2) la fonction de répartition associée à P_1 (respectivement à P_2).

Soit $(X_i)_{i\in\mathbb{Z}}$ une suite de variables aléatoires réelles stationnaires. Nous nous intéressons, ici, au Théorème limite central pour la distance $\mathbf{L}^1$ de Wasserstein, définie par,

$$\int_{\mathbb{R}} |F_n(t) - F_X(t)|\, dt, \tag{1.1}$$

avec F_X la fonction de répartition commune aux variables aléatoires (X_i), et F_n la fonction de répartition empirique correspondante.

Dans la littérature, de nombreux travaux sur la distance de Kantorovich ou $\mathbf{L}^1$ de Wasserstein ont déjà été faits, en particulier pour une suite de variables aléatoires i.i.d $X = (X_i)_{i\in\mathbb{Z}}$ (cf par exemple del Barrio, Giné et Matrán (1999)). Rappelons que si X a pour fonction de répartition F_X, alors la condition

$$\int_{-\infty}^{\infty} \sqrt{F_X(t)(1 - F_X(t))}\, dt < \infty,$$

est équivalente à

$$\Lambda_{2,1}(X) := \int_0^{\infty} \sqrt{\mathbb{P}(|X| > t)}\, dt < \infty.$$

Dans leur Théorème 2.1, del Barrio, Giné et Matrán (1999) montrent que si $(X_i)_{i\in\mathbb{Z}}$ est une suite de variables aléatoires i.i.d, alors le processus $\sqrt{n}(F_n - F_X)$ converge en loi dans $\mathbf{L}^1$ vers le processus $\{B(F(t)), t \in \mathbb{R}\}$, où B est un pont brownien, si et seulement si $\Lambda_{2,1}(X) < \infty$. Notre résultat principal généralise le Théorème 2.1 de del Barrio, Giné et Matrán (1999) au cas des suites de variables aléatoires stationnaires, satisfaisant à des conditions de dépendance appropriées.

Avant de donner l'idée de la preuve, introduisons $\mathbf{L}^1(\mu) = \mathbf{L}^1(\mathbf{T}, \mu)$, où μ est une mesure σ-finie, l'espace de Banach des fonctions réelles μ-intégrables sur $\mathbf{T}$, avec pour norme $\|.\|_{1,\mu}$, définie par $\|x\|_{1,\mu} = \int_{\mathbf{T}} |x(t)|\, \mu(dt)$. Soit $\mathbf{L}^{\infty}(\mu)$ son espace dual.

D'abord, nous donnons le Théorème limite central pour des suites stationnaires ergodiques de différences de martingales dans $\mathbf{L}^1(\mu)$. Puis, par une approximation par des différences de martingales (cf Volný (1993)), nous en déduisons un Théorème limite central pour des suites de variables aléatoires stationnaires ergodiques à valeurs dans $\mathbf{L}^1(\mu)$, et satisfaisant des conditions projectives. Ceci nous permet d'aboutir à des conditions suffisantes pour obtenir des résultats sur le comportement asymptotique de statistiques du type (1.1) pour une importante classe de suites dépendantes. En particulier, les résultats seront appliqués à l'étude de systèmes dynamiques, ainsi qu'à celle des processus linéaires causaux.

2 Théorème limite central pour des suites de variables aléatoires stationnaires dans $\mathbf{L}^1(\mu)$

A partir de maintenant, nous supposerons que la suite stationnaire ergodique $(X_i)_{i\in\mathbb{Z}}$ de variables aléatoires centrées à valeurs dans $\mathbf{L}^1(\mu)$, est donnée par $X_i = X_0 \circ \mathbb{T}^i$, où $\mathbb{T} : \Omega \longrightarrow \Omega$ est une transformation bijective bimesurable conservant la mesure de probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{A})$. Soient $S_n = \sum_{j=1}^n X_j$, les sommes partielles. En considérant $\mathcal{F}_0$, une tribu satisfaisant $\mathcal{F}_0 \subseteq \mathbb{T}^{-1}(\mathcal{F}_0)$, posons $\mathcal{F}_i = \mathbb{T}^{-i}(\mathcal{F}_0)$.

Notation 2.1. Pour tout entier $p \geq 1$ et pour toute variable aléatoire réelle Y, nous notons $\|.\|_p$, la $\mathbf{L}^p$-norme définie par $\|Y\|_p = \mathbb{E}(|Y|^p)^{1/p}$, et $\|.\|_\infty$ caractérise la $\mathbf{L}^\infty$-norme, c'est-à-dire le plus petit réel u tel que $\mathbb{P}(|Y| > u) = 0$.

Voici notre résultat principal:

Théorème 2.2. *Supposons, que pour tout réel t, $\mathbb{E}(X_0(t)|\mathcal{F}_{-\infty}) = 0$, $\mathbb{E}(X_0(t)|\mathcal{F}_\infty) = X_0(t)$ et*

$$\int_{\mathbf{T}} \|X_0(t)\|_2 \, \mu(dt) < \infty. \tag{2.1}$$

Posons $P_0(X(t)) = \mathbb{E}(X(t)|\mathcal{F}_0) - \mathbb{E}(X(t)|\mathcal{F}_{-1})$ et supposons que

$$\sum_{k\in\mathbb{Z}} \int_{\mathbf{T}} \|P_0(X_k(t))\|_2 \, \mu(dt) < \infty. \tag{2.2}$$

Alors

$$n^{-1/2} \sum_{i=1}^n X_0 \circ \mathbb{T}^i \underset{n\to\infty}{\longrightarrow} G \text{ en loi dans } \mathbf{L}^1(\mu), \tag{2.3}$$

où G est une variable aléatoire Gaussienne centrée à valeurs dans $\mathbf{L}^1(\mu)$ avec l'opérateur de covariance suivant: pour tout $f \in \mathbf{L}^\infty(\mu)$,

$$\Phi_G(f, f) = \mathbb{E}\big(\big(f\big(\sum_{k\in\mathbb{Z}} P_0(X_k)\big)\big)^2\big) = \sum_{k\in\mathbb{Z}} \mathrm{Cov}(f(X_0), f(X_k)). \tag{2.4}$$

En conséquence, nous avons

Corollaire 2.3. *Supposons que (2.1) ait lieu. De plus, supposons que*

$$\sum_{n=1}^\infty \frac{1}{\sqrt{n}} \int_{\mathbf{T}} \|\mathbb{E}(X_n(t) \mid \mathcal{F}_0)\|_2 \, \mu(dt) < \infty \,, \tag{2.5}$$

et que

$$\sum_{n=1}^\infty \frac{1}{\sqrt{n}} \int_{\mathbf{T}} \|X_{-n} - \mathbb{E}(X_{-n}(t) \mid \mathcal{F}_0)\|_2 \, \mu(dt) < \infty \,. \tag{2.6}$$

Alors, la conclusion du Théorème 2.2 a lieu.

3 Applications à la fonction de répartition empirique

Soit $Y = (Y_i)_{i\in\mathbb{Z}}$ une suite de variables aléatoires réelles. Nous notons leur fonction de répartition commune par F_Y et par F_n la fonction de répartition empirique correspondante de Y:

$$\forall\ t \in \mathbb{R},\ F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{Y_i\le t}.$$

Soit λ la mesure de Lebesgue sur $\mathbb{R}$. Si $\mathbb{E}(|Y_1|) < \infty$, la variable aléatoire $X_i(.) = \{t \mapsto \mathbf{1}_{Y_i\le t} - F_Y(t), t \in \mathbb{R}\}$ peut être vue comme une variable aléatoire centrée à valeurs dans $\mathbf{L}^1(\lambda)$.

3.1 Suites dépendantes.

Comme nous le verrons dans cette partie, en appliquant le Corollaire 2.3, nous pouvons obtenir des conditions suffisantes pour la convergence dans $\mathbf{L}^1(\lambda)$ du processus $\sqrt{n}(F_n - F_Y)$, dès que la suite Y satisfait quelques conditions de faible dépendance. Posons $\mathcal{F}_0 = \sigma(Y_i, i \le 0)$. Rappelons les définitions des coefficients de dépendance, d'après Dedecker et Prieur (2005): pour tout entier $k \ge 0$,

$$\tilde{\phi}(k) = \sup_{t\in\mathbb{R}} \|\mathbb{P}(Y_k \le t \mid \mathcal{F}_0) - \mathbb{P}(Y_k \le t)\|_\infty\,,$$

et

$$\tilde{\alpha}(k) = \sup_{t\in\mathbb{R}} \|\mathbb{P}(Y_k \le t \mid \mathcal{F}_0) - \mathbb{P}(Y_k \le t)\|_1\,.$$

Quand la suite Y est $\tilde{\phi}$-dépendante, le résultat suivant a lieu:

Proposition 3.1. *Supposons que*

$$\sum_{k\ge1} \sqrt{\frac{\tilde{\phi}(k)}{k}} < \infty \ \textit{et}\ \int_0^\infty \sqrt{\mathbb{P}(|Y| > t)}\, dt < \infty, \tag{3.1}$$

alors $\{t \mapsto \sqrt{n}(F_n(t) - F_Y(t)), t \in \mathbb{R}\}$ *converge dans* $\mathbf{L}^1(\lambda)$*, vers une Gaussienne centrée, de fonction de covariance: pour tous* $f,\ g \in \mathbf{L}^\infty(\lambda)$*,*

$$\Phi_\lambda(f, g) = \int_{\mathbb{R}^2} f(s)g(t)C(s, t)\, dt\, ds \tag{3.2}$$

avec

$$C(s, t) = F_Y(t \wedge s) - F_Y(t)F_Y(s) + 2\sum_{k\ge1}(\mathbb{P}(Y_0 \le t, Y_k \le s) - F_Y(t)F_Y(s)).$$

Remarque 3.2. La Proposition 3.1 est aussi vérifiée en ϕ-mélange (Ibragimov (1962)). Remarquons que ce résultat contient le cas i.i.d, étudié dans del Barrio, Giné et Matrán (1999).

Avant de donner un résultat similaire quand Y est $\tilde{\alpha}$-dépendante, rappelons la définition suivante:

Définition 3.3. Pour toute variable aléatoire positive intégrable Y, nous définissons Q_Y de Y, comme la fonction inverse de queue cadlag $x \to \mathbb{P}(Y > x)$.

Proposition 3.4. *Supposons que*

$$\sum_{k \geq 1} \frac{1}{\sqrt{k}} \int_0^{\tilde{\alpha}(k)} \frac{Q_Y(u)}{\sqrt{u}} \, du < \infty, \tag{3.3}$$

alors la conclusion de la Proposition 3.1 a lieu.

Remarque 3.5. La Proposition 3.4 est aussi vraie en α-mélange (Rosenblatt (1956)).

3.1.1 Application aux systèmes dynamiques.

Soit T une application de $[0, 1]$ dans $[0, 1]$ préservant la mesure de probabilité μ sur $[0, 1]$. Rappelons que l'opérateur de Perron-Frobenius K de $\mathbf{L}^1(\mu)$ dans $\mathbf{L}^1(\mu)$ est défini par l'égalité suivante: pour tout $h \in \mathbf{L}^1(\mu)$ et $f \in \mathbf{L}^\infty(\mu)$,

$$\int_0^1 (Kh)(x) f(x) \mu(dx) = \int_0^1 h(x)(f \circ T)(x) \mu(dx).$$

Ici, nous nous intéressons à donner des conditions suffisantes pour la convergence dans $\mathbf{L}^1(\lambda)$ de la fonction de répartition empirique associée à F_Y où les variables aléatoires $(Y_i)_{i \in \mathbb{Z}}$ sont définies comme suit: pour une fonction monotone donnée f, posons

$$Y_k = f \circ T^k \,. \tag{3.4}$$

Puisque sous la probabilité $([0, 1], \mu)$, la variable aléatoire $(T, T^2, ..., T^n)$ est distribuée comme $(Z_n, Z_{n-1}, ..., Z_1)$, où $(Z_i)_{i \geq 0}$ est une chaîne de Markov stationnaire de mesure invariante μ et de noyau de transition K (cf Lemme $XI.3$ dans Hennion and Hervé (2001)), la convergence dans $\mathbf{L}^1(\lambda)$ de la fonction de répartition empirique associée à F_Y est déduite de celle de la fonction de répartition empirique associée à $F_{f(Z)}$.

Le cas des applications BV-contractantes (cf par exemple Dedecker et Prieur (2005), pour plus de détails et d'exemples d'applications BV-contractantes).
Nous obtenons le Corollaire suivant:

Corollaire 3.6. *Si T est BV-contractante et f est une fonction monotone de* $]0, 1[$ *dans* $\mathbb{R}$ *satisfaisant* $\int_0^\infty \sqrt{\lambda(|f| > t)} \, dt < \infty$, *alors la conclusion de la Proposition 3.1 a lieu pour la suite* $(Y_k)_{k \in \mathbb{Z}}$ *où* Y_k *est définie par (3.4).*

3.2 Processus linéaires causaux

Maintenant, nous nous intéressons à la suite

$$Y_k = \sum_{j\geq 0} a_j \varepsilon_{k-j}, \tag{3.5}$$

où $(\varepsilon_i)_{i\in\mathbb{Z}}$ est une suite de variables aléatoires réelles i.i.d à valeurs dans $\mathbf{L}^2$ et $\sum_{j\geq 0} |a_j| < +\infty$.

Corollaire 3.7. *Supposons que, ε_0 ait une densité bornée par K et que $|a_0| \neq 0$. De plus, supposons que*

$$\sum_{k\geq 0} \int_0^{(a_k)^2} \frac{Q_{|Y_0|}(u)}{\sqrt{u}}\, du < \infty. \tag{3.6}$$

Alors la conclusion de la Proposition 3.1 a lieu pour une suite $(Y_k)_{k\in\mathbb{Z}}$ où Y_k est définie par (3.5).

Bibliographie

[1] Dede, S. (2008). An empirical central limit theorem in $\mathbf{L}^1$ for stationary sequences. *hal-00347334 ou arXiv 0812.2839.*

[2] Del Barrio, E., Giné, E. et Matrán, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *The Annals of Probability*, 27, No. 2, 1009-1071.

[3] Dedecker, J. et Prieur, C. (2005). New dependence coefficients. Examples and applications to statistics. *Probab. Theory. Relat. Fields*, 132, 203-236.

[4] Hennion, H. et Hervé, L. (2001). Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness. *Lecture Notes in Mathematics*, 1766.

[5] Ibragimov, I.A. (1962). Some limit theorems for stationary processes. *Theory Probab. Appl.*, 7, 349-382.

[6] Rosenblatt, M. (1956). A central limit Theorem and a strong mixing condition. *Proc. Nat. Ac. SC. U.S.A.*, 42, 43-47.

[7] Volný, D. (1993). Approximating martingales and the central limit theorem for strictly stationary processes. *Stochastic Processes and their Applications*, 44, 41-74.