Université Denis Diderot Année 2011-2012
Paris VII **M2 modèles aléatoires**

Compléments sur le théorème de Donsker

1 Preuve de Donsker

On note B le mouvement brownien standard.

L'espace des fonctions continues de $[0,1]$ dans $\mathbb{R}$ est noté $\mathcal{C}([0,1],\mathbb{R})$. Avec un léger abus, on note encore $B = (B_t, t \in [0,1])$.

On suppose que les $\{X_i, i \geq 1\}$ sont des variables i.i.d. avec $\mathbb{E}[X_1 = 0]$, $\mathrm{Var}[X_1] = 1$.

On va supposer dans un premier temps (pour simplifier) que X_1 a un support fini, i.e. $\exists K > 0 : \mathbb{P}(|X_1| > K) = 0$.

On note $S_0 := 0$, $S_n = \sum_{i=1}^n X_i$.

Afin de retrouver une fonction continue de $[0,1]$ dans $\mathbb{R}$, on introduit $\mathcal{S}_n$ l'interpolation linéaire de la marche réechelonnée
$\{S_0/\sqrt{n}, S_1/\sqrt{n}, S_2/\sqrt{n}, ..., S_n/\sqrt{n}\}$ sur $[0,1]$. Précisément on construit $\mathcal{S}_n$ de sorte qu'en chaque $t_i = i/n, i \in \{0,...,n\}$, $\mathcal{S}_n(t_i) = S_i/\sqrt{n}$, et sur chaque intervalle $[i/n,(i+1)/n], i = 0,...,n-1$, $\mathcal{S}_n$ est affine.

Théorème 1 (Théorème de Donsker) Quelque soit $\Lambda : \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$ fonctionnelle continue, on a

$$\Lambda(\mathcal{S}_n) \xrightarrow[n\to\infty]{\text{(loi)}} \Lambda(B).$$

Pour prouver le théorème on doit tout d'abord vérifier la convergence des marginales finies dimensionnelles.

Exercice 1.1

1. Soit $t \in [0,1]$ fixé. Quelle est la limite en loi de $S_{[nt]}/\sqrt{n}$? Quelle est la limite en loi de $\mathcal{S}_n(t)$?
2. Soient $0 \leq t_1 < ... < t_p \leq 1$ fixés. Quelle est la limite en loi de
$$(\mathcal{S}_n(t_1), \mathcal{S}_n(t_2) - \mathcal{S}_n(t_1), ..., \mathcal{S}_n(t_p) - \mathcal{S}_n(t_{p-1}))?$$
3. En déduire qu'on a bien, pour $(t_i)_{i=1...p}$ comme ci-dessus ,
$$((\mathcal{S}_n(t_1), \mathcal{S}_n(t_2), ..., \mathcal{S}_n(t_p))) \xrightarrow[n\to\infty]{\text{(loi)}} (B_{t_1}, ..., B_{t_p}).$$

Evidemment, la convergence en loi des marginales finies dimensionnelles n'est pas suffisante pour assurer la convergence en loi d'une suite de processus. Tout au plus, cela permet de garantir que si la suite converge, elle converge vers la bonne limite.

La théorie générale de la convergence des processus permet de conclure si on a la *tension* de la suite, ce qui, d'après le théorème de Prohorov (voir [Billingsley], chapitre 1.6) équivaut dans un espace polonais à sa précompacité : de toute sous-suite, on peut extraire un sous-suite convergente.

Plutôt que de faire appel à cette théorie générale, on va développer un argument ad hoc.

Pour un fonction f, on note $\pi_k f$ l'interpolation linéaire de $\{f(0), f(2^{-k}), ..., f(i2^{-k}), ..., f(1)\}$ sur $[0,1]$.

Exercice 1.2 Fixons $\Lambda : \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$ continue.

1. Montrer que $\Lambda(\pi_k B) \xrightarrow[n\to\infty]{\text{(loi)}} \Lambda(B)$.
2. Montrer que l'exercice précédent implique
$$\Lambda(\pi_k \mathcal{S}_n) \xrightarrow[n\to\infty]{\text{(loi)}} \Lambda(\pi_k B).$$
3. En se servant de l'inégalité de Hoeffding, établir l'existence d'un $t_0 > 0$ tel que $\forall t \geq t_0$,
$$\mathbb{P}(\max_{1\leq i\leq m} |S_i| \geq t) \leq \frac{2m}{t^2} \exp\left(-\frac{t^2}{4mK^2}\right).$$
En déduire que
$$\lim_{k\to\infty} \sup_{n\geq 1} \mathbb{P}\left(||\pi_k \mathcal{S}_n - \mathcal{S}_n||_\infty > \delta\right) = 0.$$
4. Conclure que
$$\Lambda(\mathcal{S}_n) \xrightarrow[n\to\infty]{\text{(loi)}} \Lambda(B).$$

Exercice 1.3 1. Que se passe-t-il si on a $\mathrm{Var}(X_1) = \sigma^2$?

2. Adapter la preuve précédente lorsque X_1 a un support quelconque.
Indication :
On pourra introduire les variables *tronquées* $Y_i^{(K)} := X_i^{(K)} \mathbf{1}_{|X_i|\leq K}$, et les variables tronquées et recentrées $X_i^{(K)} := Y_i^{(K)} - \mathbb{E}[Y_i^{(K)}]$.
On définira alors la marche $S^{(K)}$ dont les sauts (tronqués et recentrés) sont i.i.d de loi $X_1^{(K)}$, et son interpolation réechelonée $\mathcal{S}^{(K)}$.

2 Donsker et pont brownien

Définition 2.1 Soit $(B_t, t \geq 0)$ un mouvement brownien standard Le *pont brownien* $\mathcal{P}_t, t \in [0,1]$ est un processus gaussien à trajectoires continues, tel que $\mathcal{P}(0) = \mathcal{P}(1) = 0$, et si $0 \leq s < t \leq 1$, $\mathrm{Cov}(\mathcal{P}_s, \mathcal{P}_t) = s(1-t)$.

Exercice 2.1 Vérifier que le processus $(B_t - tB_1, t \in [0,1])$ est un pont brownien.

Soient $(\xi_i, i \geq 1)$ des variables i.i.d, uniformes sur $[0,1]$. On note
$$\mathcal{P}_n(t) := \frac{1}{\sqrt{n}} \sum_{i=1}^n (\mathbf{1}_{\xi_i \leq t} - t).$$

Théorème 2 (Donsker et pont brownien) Soit $\Lambda : \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$ une fonction continue. Alors
$$\Lambda(\mathcal{P}_n) \xrightarrow[n\to\infty]{\text{(loi)}} \Lambda(\mathcal{P}).$$

Exercice 2.2 Vérifier la convergence des lois finies dimensionnelles.

On admet le reste de la preuve, qui découle de ce que pour tout $\delta > 0$, $\epsilon > 0$ il existe k tel que

$$\limsup_{n\to\infty} \mathbb{P}\left(w_{\mathcal{P}_n}(1/k) > \delta\right) \le \epsilon,$$

où $w_{\mathcal{P}_n}(1/k) := \sup_{s,t\in[0,1],|s-t|<k^{-1}} |\mathcal{P}_n(s) - \mathcal{P}_n(t)|$.

Le critère ci-dessus se vérifie aisément et permet essentiellement d'assurer la tension (et donc la précompacité) de la suite $\mathcal{P}_n$. D'après la convergence des marginales finies dimensionnelles, toute sous-suite convergente ne peut converger que vers le pont brownien, et on conclut.

Pour plus de détails sur les critères de tension pour des suites d'éléments de $\mathcal{C}([0,1],\mathbb{R})$, on pourra consulter le chapitre 3 de [Billingsley].

On peut en fait généraliser le résultat pour des variables $\{\xi_i, i \ge 1\}$ non nécessairement uniformes.

Mais bien entendu, dans le cas général, la loi de la limite change. On obtient un pont brownien changé de temps par l'inverse de la fonction de répartition de ξ_1.

Ce résultat est utilisé en statistiques pour le *test de Kolmogorov-Smirnov.*

Pour plus de détails on pourra consulter le chapitre 7.8 de [Durrett].

3 Plongement de Skorokhod

Le but de la section est de prouver à nouveau le théorème de Donsker. On va le faire dans le cadre de la marche simple [1].

Le but du *plongement* est de retrouver la marche simple dans la trajectoire brownienne.

Commençons par le premier pas X_1 de la marche simple. Il s'agit là d'une variable qui prend la valeur -1 avec probabilité 1/2 et 1 avec probabilité 1/2.

Exercice 3.1

1. Montrer qu'il existe un temps d'arrêt T_1 du mouvement brownien tel qu'en loi, $B_{T_1} = X_1$.
2. De façon plus générale, montrer que pour tout $a > 0$ il existe T_a tel qu'en loi $B_{T_a} = aX_1$.

Grâce à la propriété de Markov, il est aisé de généraliser le résultat précédent.

Exercice 3.2 Soient $X_i, i \ge 1$ des variables i.i.d de même loi que X_1. On définit $S_0 = 0, S_n = \sum_{i=1}^n X_i$, de sorte que $(S_n, n \ge 0)$ est la marche simple.

Montrer qu'à chaque n on peut associer un temps d'arrêt T_n du mouvement brownien tel que

$$B_{T_n} \stackrel{\text{(loi)}}{=} S_n, \qquad \text{avec}\{T_n - T_{n-1}, n \ge 1\}\text{i.i.d..}$$

Remarquer qu'on pourrait faire de même avec la marche réechelonnée.

---

1. La généralisation de cette preuve repose avant tout sur la généralisation de l'exercice 3.1. On peut ainsi retrouver n'importe quelle variable centrée et de carré intégrable dans la trajectoire brownienne (voir le chapitre 7.6 de [Durrett] pour des précisions)

Exercice 3.3 1. Soit $(W_n(t) := B(nt)/\sqrt{n}, t \geq 0)$. Quelle est la loi du processus W_n ?

2. En se servant de l'exercice 3.2, montrer que

$$\mathcal{L} - \lim_{n\to\infty} S_n/\sqrt{n} = \mathcal{L} - \lim_{n\to\infty} W_n(\mathbb{E}[T_1]).$$

Quelle est la loi limite commune ?

3. Montrer plus généralement que si $0 \leq t_1 < ... < t_p \leq 1$, on a

$$(S_{[nt_1]}/\sqrt{n}, ..., S_{[nt_p]/\sqrt{n}}) \xrightarrow[n\to\infty]{\text{(loi)}} (W(t_1), ..., W(t_p)).$$

On précisera évidemment la loi du vecteur limite.

Pour finir on a besoin d'une construction un peu plus générale, faisant intervenir une suite de marches.

Plus précisément, considérons $X_{n,m}, n \geq 1, m \geq 1$ des variables aléatoires i.i.d de même loi que X_1, $S_{n,m} := X_{n,1} + ... + X_{n,m}$ et $\mathcal{S}_n$ l'interpolation linéaire réechelonnée par $\sqrt{n}$ de $S_{n,m}, m \geq 1$.

Exercice 3.4 Montrer que, de façon similaire à l'exercice 3.2, on peut choisir les temps d'arrêts τ_m^n de sorte que les $\tau_{m+1}^n - \tau_m^n$ sont i.i.d, et pour tout n,

$$(S_{n,m}/\sqrt{n}, 1 \leq m \leq n) \overset{\text{(loi)}}{=} (B(\tau_m^n), 1 \leq m \leq n).$$

Montrer qu'avec ce choix, pour tout $s \in [0, 1]$, $\tau_{[ns]}^n \xrightarrow{\mathbb{P}} s$ lorsque $n \to \infty$.

L'idée pour conclure est celle de la représentation de Skorokhod. On va utiliser des *représentants* des éléments de la suite, c'est-à-dire, pour chaque $n \geq 1$, un processus $\mathbb{S}_n$ a priori très différent, mais *égal en loi* à $\mathcal{S}_n$).

Prenons tout simplement pour représentants

$$(\mathbb{S}_{n,m}/\sqrt{n}, 1 \leq m \leq n) = (B(\tau_m^n), 1 \leq m \leq n),$$

et notons $\mathbb{S}_n$ l'interpolation linéaire sur $[0, 1]$, de ce processus, i.e. tel que $\mathbb{S}_n(0) = 0, \mathbb{S}_n(m/n) = \mathbb{S}_{n,m}/\sqrt{n}, 1 \leq m \leq n$.

Exercice 3.5 Montrer alors que

$$||\mathbb{S}_n - B||_\infty \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

puis que ceci permet effectivement de conclure à la convergence en loi de $\mathcal{S}_n$ vers B.

Références

[Billingsley] Patrick Billingsley, *Convergence of Probability Measures*, Wiley, 1968.

[Durrett] Rick Durrett, *Probability, Theory and Examples*, 3rd edition, Duxbury Press, 1996.