RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE

Université Mouloud Mammeri de Tizi Ouzou
Faculté des Sciences
Département de Mathématiques

MÉMOIRE DE MASTER

Filière : Mathématiques
Spécialité : Probabilités & Statistique

Par

Fatokoma Barry

SUR LE MOUVEMENT BROWNIEN ET LE THÉORÈME CENTRAL LIMITE FONCTIONNEL

Soutenu le 04/10/2022 devant le jury :

Dr.	**Abdelghani Hamaz**	UMMTO	Président de jury
Dr.	**Dalila Merabet**	UMMTO	Examinateur
Pr.	**Hamadouche Djamel**	UMMTO	Encadreur

Année Universitaire : 2021/2022

Remerciements

Je tiens à remercier tout d'abord **Mr Hamadouche Djamel** *pour sa disponibilité et en s'assurant la direction et le suivi du présent travail.*

Mes sincères remerciements vont à l'endroit de tous les professeurs qui ont contribué à ma formation.

Je tiens à remercier mon pays d'accueil **Algérie***, mon pays d'origine le* **Mali** *et leurs collaborations bilatérales.*

Enfin, mes sincères remerciements à mes parents et tous ceux qui ont contribué à la réalisation de ce modeste mémoire.

Table des matières

1 Introduction générale

En théorie des probabilités, il existe différentes notions de convergence de variables aléatoires. Certaines de ces notions ne sont pas spécifiques des probabilités, mais de l'analyse en générale, comme la convergence presque sûre, ou encore la convergence L^p. La convergence en loi de suites de variables aléatoires est un concept appartenant plus spécifiquement à la théorie des probabilités, utilisé notamment en statistique et dans l'étude des processus stochastiques.
Le mouvement brownien communément appelé processus de Wiener, est une description mathématique du mouvement aléatoire d'une «grosse» particule qui immerge dans un fluide et qui n'est soumise à aucune interaction, que des chocs avec les «petites» molécules du fluide environnant. Il en résulte un mouvement très irrégulier de la grosse particule, qui a été décrite pour la première fois en 1827 par le botaniste Robert Brown en observant des mouvements des particules à l'intérieur du fluide.
Les deux notions ci-dessus s'entrecroisent sur le théorème central limite fonctionnel, qui concerne la convergence en loi des processus de sommes partielles vers le mouvement brownien.
Pour le travail ci, nous allons parler sur le mouvement brownien. Tout d'abord nous allons faire une brève présentation du mouvement brownien, son historique, les différentes applications, et donner ses propriétés. Dans le chapitre suivant nous nous intéressons à la convergence en loi dans $\mathbb{R}$ et au théorème centrale limite. Dans le troisième chapitre nous allons parler sur les différents espaces de convergence qui sont l'ensemble des fonctions discontinues (Espace $\mathbf{D}$), l'ensemble des fonctions continues (Espace $\mathbf{C}$); l'équitension et la compacité sur ces espaces. Appliquer le théorème de Donsker sur ces espaces fonctionnels. En dernier nous allons parler du théorème central limite fonctionnel et faire une application sta-

tistique.
Concernant le mouvement brownien les propriétés prises en compte sont : propriété de Markov, les trajectoires du mouvement brownien, propriétés de martingale. Nous allons aussi parler des différents types de mouvement brownien qui sont : le mouvement brownien géométrique, le mouvement brownien multidimensionnel.
Sur la convergence en loi et le théorème central limite dans $\mathbb{R}$ nous allons donner la définition et quelques propriétés sur cette convergence, parler sur le théorème de Lévy et sur le théorème central limite des variables aléatoires sur $\mathbb{R}$.
A propos de la convergence étroite sur des espaces fonctionnels et le principe d'invariance de Donsker nous allons introduire la notion de relative compacité, l'equitension et le lien qui existe entre les deux. Etudier l'espace $\mathbf{D}$ et sa compacité ; l'espace $\mathbf{C}$ et l'equitension. Énoncer et donner la démonstration du théorème de Donsker sur les espaces fonctionnels énumérés ci-dessus.
L'application statistique concerne le processus de sommes partielles (lissé ou non), la convergence de ces processus vers le mouvement brownien. Nous allons faire un test d'hypothèse et déterminer la statistique que nous allons utiliser. Faire une application numérique en determinant la région de rejet et la puissance du test.

2 Mouvement Brownien

2.1 Présentation

Le mouvement brownien communément appelé processus de Wiener, est une description mathématique du mouvement aléatoire d'une «grosse» particule qui immerge dans un fluide et qui n'est soumise à aucune interaction, que des chocs avec les «petites» molécules du fluide environnant. Il en résulte un mouvement très irrégulier de la grosse particule, qui a été décrite pour la première fois en 1827 par le botaniste Robert Brown en observant des mouvements des particules à l'intérieur du fluide.
Ensuite Delsaux en 1877 explique les changements incessants de direction de trajectoire par les chocs entre les particules de pollen et les molécules d'eau ; Bachelier en 1900 met en évidence le caractère "Markovien" du mouvement brownien, en vue d'étudier les cours de la bourse ; Einstein en 1905 détermine la densité de transition du mouvement brownien par l'intermédiaire de l'équation de la chaleur et relie ainsi le mouvement brownien et les équations aux dérivées partielles de type parabolique ; Smoluchowski en 1905 décrit le mouvement brownien comme une limite de promenades aléatoires ; N Wiener en 1923 réalise la première étude mathématique rigoureuse et donne une démonstration de l'existence du mouvement brownien ; P. Lévy en 1948 s'intéresse aux propriétés fines des trajectoires du mouvement brownien ; depuis les travaux d'Ito, Watanable, Meyer, Yor LeGall, ...
Ce mouvement permet de décrire avec succès le comportement thermodynamique des gaz (Théorie cinétique des gaz), ainsi que le phénomène de diffusion. Il est aussi très utilisé dans des modèles de mathématiques financières.

2.2 Mouvement brownien - Généralités

Définition 2.1 : Un processus $(B_t)_{t\geq 0}$ est un mouvement brownien standard si :
$(B_t)_{t\geq 0}$ est un processus gaussien,
$(B_t)_{t\geq 0}$ est presque-surement continu,
Pour tous s, t, $E(B_t) = 0$ et $E(B_t B_s) = \min(s,t)$.

Propriétés 2.1 :

1. Si $B_0 = 0$, B est appelé mouvement brownien standard.

2. Si $B_0 = B_1 = 0$, B est appelé pont brownien.

3. Si $P(B_0 = 0) = 1$ (le mouvement est issu de l'origine).

4. $\forall s \leq t$, $B_t - B_s$ est une variable réelle de loi gaussienne centrée, de variance $(t-s)$.

5. $\forall n$, $\forall t_i$, $0 \leq t_0 \leq t_1 \leq ... \leq t_n$, les variables

$$(B_{t_n} - B_{t_{n-1}}, ..., B_{t_1} - B_{t_0}, B_{t_0})$$

sont indépendantes.

6. Pour tout (t,s), la variable $B_{t+s} - B_t$ est indépendante de la tribu du passé avant t, soit $F_t^B = \sigma(B_u, u \leq t)$.

2.2.1 Généralisation

Soit Z un processus défini par $Z_t = a + B_t$; Z est un mouvement brownien issu de a.
On dit que X est un mouvement brownien de drift μ et de coefficient de diffusion σ si :

$$X_t = x + \mu t + \sigma B_t$$

où B est un mouvement brownien. La v.a. X_t est une v.a. gaussienne d'espérance $x + \mu t$, de variance $\sigma^2 t$ et $x = B_0$.
Pour tout (t,s), la v.a $X_{t+s} - X_t$ est indépendante de $F_s^X = \sigma(X_u, u \leq s)$.

2.3 Propriétés

On peut montrer que le mouvement brownien s'obtient comme limite de promenades aléatoires renormalisées.
Soit, sur un espace probabilisé (Ω,F,P), une famille de v.a. $(X_i)_{i\geq 1}$ de Bernoulli indépendantes équidistribuées.

$$P(X_i = 1) = P(X_i = -1) = 1/2, \quad i \geq 1.$$

On associe à cette famille la suite $(S_n, n \geq 0)$ définie par :

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

On dit que la suite (S_n) est une promenade aléatoire.
On a $E(S_n) = 0$, $Var(S_n) = n$.
Remarquons que la suite $(S_m - S_n, m \geq n)$ est indépendante de $(S_0, S_1, ..., S_n)$ et que $S_m - S_n$ est de même loi que S_{m-n}.
On procède alors à une double renormalisation. Soit N fixé

- On ramène l'intervalle de temps [0,N] à [0,1].

- On change l'échelle des valeurs prises par S_n.

Plus précisément, on définit une famille de variables aléatoires indexées par les réels de la forme $k/N, k$ des entiers naturels, par

$$U_{\frac{k}{N}} = \frac{1}{\sqrt{N}} S_k.$$

On a $E(U_{k/N}) = 0$ et $Var(U_{k/N}) = k/N$.
Les propriétés d'indépendance et de stationnarité de la promenade aléatoire restent vérifiées.

- Si $k \geq k'$, $\quad U_{k/N} - U_{k'/N}$ est indépendante de $(U_{p/N}; \quad p \leq k')$.

- Si $k \geq k'$, $U_{\frac{k}{N}} - U_{\frac{k'}{N}}$ a même loi que $U_{\frac{k-k'}{N}}$.

On définit un processus à temps continu $(U_t, t \geq 0)$ à partir de $U_{k/N}$ en imposant à la fonction $t \to U_t$ d'être affine entre k/N et $(k+1)/N$. Pour cela, N étant fixé, on remarque que pour tout t réel, il existe $k(t)$ un entier naturel unique tel que $\frac{k(t)}{N} \leq t \leq \frac{k(t)+1}{N}$ et on pose :

$$U_t^N = U_{\frac{k}{N}} + N(t - \frac{k}{N})(U_{\frac{k+1}{N}} - U_{\frac{k}{N}})$$

où $k = k(t)$.
Pour $t = 1$ on a $U_t^N = \frac{1}{\sqrt{N}} S_N$. Le théorème central limite implique que U_1^N converge en loi vers une variable aléatoire gaussienne centrée réduite.
On montre alors que le processus U^N converge (au sens de la convergence en loi) vers un mouvement brownien B.
En particulier $U_t^N \rightarrow^{Loi} B_t$ et $(U_{t_1}^N, ..., U_{t_n}^N) \rightarrow^{Loi} (B_{t_1}, ..., B_{t_n})$ pour tout k-uple $(t_1, ..., t_k)$.

Proposition 2.1 : Le processus B est un processus gaussien, d'espérance nulle et de covariance $Cov(B_t, B_s) = s \wedge t$.
Le processus $(X_t = x + \mu t + \sigma B_t, t \geq 0)$ est un processus gaussien d'espérance $x + \mu t$ et de covariance

$$E[(X_t - E(X_t))(X_s - E(X_s))] = \sigma^2(s \wedge t).$$

- On note $E_x(f(B_s))$ l'espérance de $f(B_s)$ quand B est un mouvement brownien issu de x. Cette quantité est égale à :

$$E(f(x + B_s)) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{+\infty} f(x+y) \exp(-\frac{y^2}{2t}) dy$$

où B est un mouvement brownien issu de 0.

- On note $P_x(B_s \in A) = P(x + B_s \in A)$ on a $P_x(B_s \in A)$ est la densité de la v.a B_s.

Proposition 2.2 : Si $(B_t \geq 0)$ est un mouvement brownien, alors

i) Le processus $\hat{B}_t = -B_t$ est un mouvement brownien.

ii) Le processus $\tilde{B}_t = \frac{1}{c} B_{c^2 t}$ est un mouvement brownien.

2.3.1 Propriété de Markov

Théorème 2.1 : Pour f borélienne bornée, et pour $u > t$,

$$E(f(B_u)/F_t^B) = E(f(B_u)/\sigma(B_t)).$$

Pour tous s et t, le processus $(W_t, t \geq 0)$ défini par $W_t =^{def} B_{t+s} - B_s$ est un mouvement brownien indépendant de F_s.
Pour $u > t$, conditionnellement à B_t, la v.a B_u est de loi gaussienne d'espérance B_t et de variance $u - t$. Alors

$$E(\mathbf{1}_{B_u \leq x}/F_t^B) = E(\mathbf{1}_{B_u \leq x}/\sigma(B_t)),$$

pour $u \geq t$.

Propriété de Markov forte Soit T un temps d'arrêt à valeurs finies, on a alors

$$E(f(B_{T+s})/F_T^B) = E(f(B_{T+s})/\sigma(B_T)).$$

En particulier, pour tout temps d'arrêt fini T, le processus $(W_t, t \geq 0)$ défini par $W_t =^{def} B_{t+T} - B_T$ est un mouvement brownien indépendant de F_T.

2.3.2 Trajectoires du mouvement brownien

Les trajectoires du mouvement brownien sont continues et sont p.s. "nulle part différentiables".

Convergence : Soit n fixé et $t_j = \frac{j}{2^n} t$ pour j variant de 0 à 2^n.
Alors $\sum_{j=1}^{2^n} [B(t_j) - B(t_{j-1})]^2 \to t$ quand $n \to \infty$, la convergence ayant lieu en moyenne quadratique et presque sûrement.

Proposition 2.3 : Soit σ une subdivision de l'intervalle $[0, t]$ caractérisée
par $0 = t_0 \leq t_1 \leq ... \leq t_n = t$. Soit V_t la variation de la trajectoire du mouvement brownien sur $[0, t]$ définie par

$$V_t(\omega) = sup_\sigma \sum_i |B_{t_i+1}(\omega) - B_{t_i}(\omega)|.$$

Alors $V_t(\omega) = \infty$ p.s.

2.3.3 Propriétés de martingale

Proposition 2.4 : Le processus B est une martingale. Le processus $(B_t^2 - t, t \geq 0)$ est une martingale.
Réciproquement, si X est un processus continu tel que X et $(X_t^2 - t, t \geq 0)$ sont des martingales, X est un mouvement brownien.
Soient B_1 et B_2 deux mouvements browniens indépendants. Le produit $B_1 B_2$ est une martingale.

Définition 2.2 : On dit que B est un (F_t)-mouvement Brownien si B et $(B_t^2 - t, t \geq 0)$ sont des (F_t)-martingales.

Proposition 2.5 : Pour tout λ réel, le processus $\exp(\lambda B_t - \frac{1}{2}\lambda^2 t), t \geq 0$ est une martingale.
Réciproquement, si X est un processus continu tel que, pour tout λ réel, $(\exp(\lambda X_t - \frac{1}{2}\lambda^2 t), t \geq 0)$ est une martingale, le processus X est un mouvement brownien

2.4 Mouvement brownien géométrique

Définition 2.3 Soit B un mouvement brownien, b et σ deux constantes. Le processus

$$X_t = x \exp\{(b - \frac{1}{2}\sigma^2)t + \sigma B_t\}, \qquad x = B_0$$

est appelé mouvement Brownien géométrique.
Ce processus est aussi appelé processus “log-normal”. En effet, dans ce cas

$$\ln X_t = (b - \frac{1}{2}\sigma^2)t + \sigma B_t + \ln x$$

et la variable qui est à droite suit une loi normale.

Propriétés 2.2 Le processus $X_t \exp^{-bt}$ est une martingale.
En notant G une v.a. de loi $N(0,1)$,

$$E(f(X_t)/F_s) = E(f(X_t)/X_s) = E(f(x \exp\{(b - \frac{1}{2}\sigma^2)(t-s) + \sigma(B_t - B_s)\}))$$

$$= E(f(x \exp\{(b - \frac{1}{2}\sigma^2)(t-s) + \sigma G\sqrt{t-s}\}))$$

$$= \int_{-\infty}^{\infty} E(f(X_s \exp\{(b - \frac{1}{2}\sigma^2)(t-s) + \sigma y\sqrt{t-s}\}))dy.$$

Ce processus est appelé le modèle Black et Scholes, il est utilisé pour modéliser le prix d'un actif financier. Le rendement de l'actif entre deux dates est mesuré par la différence des logarithmes des cours et est donné par la variable gaussienne

$$(b - \frac{1}{2}\sigma^2)(t-s) + \sigma(B_t - B_s).$$

2.5 Temps d'atteinte

Soit $(B_t, t \geq 0)$ un mouvement brownien et a un nombre réel. Soit

$$T_a = \inf\{t \geq 0; B_t = a\}.$$

Alors T_a est un temps d'arrêt fini p.s. tel que $E(T_a) = \infty$ et pour tout $\lambda \geq 0$,

$$E(\exp -\lambda T_a) = \exp(-|a|\sqrt{2\lambda}).$$

Par inversion de la transformation de Laplace, on obtient la densité de T_a qui est, pour $a > 0$

$$\frac{a}{\sqrt{2\pi t}} \exp(-\frac{a^2}{2t}).$$

2.6 Mouvement brownien multidimensionnel

Soit $B_t = (B_t^{(1)}, B_t^{(2)}, ..., B_t^{(n)})$ un processus n-dimensionnel. On dit que B est un mouvement brownien multidimensionnel si les processus $(B^{(i)}, i \leq n)$ sont des mouvements browniens indépendants.
Le processus n-dimensionnel B est un mouvement brownien si et seulement si les processus $B^{(i)}$ et $B^{(i)}B^{(j)} - \delta_{i,j}t$ sont des martingales (avec $\delta_{i,j} = 0$ pour $i \neq j$ et $\delta_{i,i} = 1$).
On dira que les mouvements browniens à valeurs réelles B_1 et B_2 sont corrélés de coefficient de corrélation ρ si $B_1(t)B_2(t) - \rho t$ est une martingale.
On "décorrèle" les mouvements browniens en introduisant le processus B_3 défini par

$$B_3(t) = \frac{1}{\sqrt{1-\rho^2}}(B_2(t) - \rho B_1(t)).$$

Ce processus est un mouvement brownien indépendant de B_1.

3

Convergence en loi et Théorème Central Limite dans $\mathbb{R}$

3.1 Introduction

Dans ce chapitre nous allons définir la convergence en loi sur $\mathbb{R}$, donner les propriétés de cette convergence et parler sur le théorème central limite.

3.2 Définitions et propriétés

Définition 3.1 Soit $(X_n)_{n\in N}$ une suite de variables aléatoires. On dit que $(X_n)_n$ converge vers X en loi (ou en distribution), et on note $X_n \rightarrow_{n\rightarrow\infty}^{(loi)} X$, si la suite de loi de probabilité $(L_{X_n})_{n\in N}$ converge étroitement vers L_X. En d'autres termes, si pour toute fonction continue bornée h, on a :

$$\lim_{n\rightarrow\infty} E[h(X_n)] = E[h(X)].$$

Propriétés

1. Si $(X_n)_n$ est une suite de variables aléatoires discrètes à valeurs dans le même ensemble E et que $\lim_{n\rightarrow\infty} P(X_n = a) = P(X = a)$, pour tout $a \in E$, alors X_n convers vers X en loi.

2. Si $(X_n)_n$ est une suite de variables aléatoires de densités f_n avec $f_n \leq g$, λ-presque partout et que g est intégrable. Si $\lim_{n\rightarrow\infty} f_n = f$ λ-presque partout, alors (X_n) converge en loi vers une variable aléatoire ayant une densité qui est f. Ceci est une conséquence du théorème de convergence dominée.

Théorème de convergence dominée Soit $(f_n)_{n\in N}$ définie sur un ensemble (S,A,μ) mesurable, une suite de fonctions dans $L^1(S,A,\mu)$ telle que $\lim_n f_n(x) = f(x)$ μ-presque partout et il existe une fonction g, positive, intégrable et mesurable avec $\forall n \in N, \mid f_n \mid \leq g$ μ-presque partout. Alors on a $f \in L^1(S,A,\mu)$ et

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu \quad et \quad \lim_{n\to\infty} \int \mid f_n - f \mid d\mu = 0.$$

Exemples

1. Soit X_n une variable aléatoire suivant une loi uniforme sur $\{\frac{1}{n}, \frac{2}{n}, ..., 1\}$. Alors $(X_n)_n$ converge en loi vers la loi uniforme sur $[0,1]$.
En effet, si h est une fonction continue bornée, alors
$E[h(X_n)] = \frac{1}{n}\sum_{i=1}^n h(\frac{i}{n}) \to_{n\to\infty} \int_0^1 h(x)dx$ par définition de l'intégrale de Riemann. Or si X suit une loi uniforme sur $[0,1]$, alors sa fonction de densité est $\mathbf{1}_{[0,1]}$ et donc $E[h(X)] = \int_0^1 h(x)dx$.

2. Soit X_n de loi gaussienne $N(0,\sigma_n^2)$ avec σ_n qui tend vers 0. Alors X_n converge en loi vers la variable aléatoire nulle (dont la loi est la masse de Dirac en 0).
En effet, si h est continue bornée par M, alors par un changement de variables on a

$$E[h(X_n)] = \frac{1}{\sigma_n\sqrt{2\pi}} \int_{\mathbb{R}} h(x)\exp(-\frac{x^2}{2\sigma_n^2})dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} h(x\sigma_n)\exp(-\frac{x^2}{2})dx.$$

On remarque que $\mid h(x\sigma_n)\exp(-\frac{x^2}{2}) \mid \leq M\exp(-\frac{x^2}{2})$ qui est intégrable, donc par le théorème de convergence dominée, on a

$$\lim_{n\to\infty} E[h(X_n)] = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} h(x)\exp(-\frac{x^2}{2})dx$$

et comme h est continue, on obtient

$$\lim_{n\to\infty} E[h(X_n)] = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} h(0)\exp(-\frac{x^2}{2})dx = h(0).$$

Or si $X = 0$ p.s on a aussi $E[h(X)] = h(0)$.

Proposition 3.1 Soit $(X_n)_{n\in N}$ une suite de variables aléatoires. Alors $X_n \to^{(loi)}_{n\to\infty} X$ si et seulement si pour toute fonction continue à support compact g, on a
$\lim_{n\to\infty} E[g(X_n)] = E[g(X)]$.

Démonstration Comme une fonction continue à support compact et bornée, alors le premier sens est trivial. Supposons maintenant que
$\lim_{n\to\infty} E[g(X_n)] = E[g(X)]$ pour toute fonction continue à support compact et montrons que $X_n \to^{(loi)}_{n\to\infty} X$.
Pour cela, soit h une fonction continue bornée par M. Pour rendre h à support compact, on aimerait multiplier par $\mathbf{1}_{[-L,L]}$ sauf que ceci annule la continuité. Cependant, définissons une fonction ψ_L telle que $\psi_L(x) = 1$ si $x \in [-L, L]$ et $\psi_L(x) = 0$ si $x \in [-(L+1), (L+1)]$. Ainsi, on a $h\psi_L(x)$ est continue à support compact et donc :

$$\lim_{n\to\infty} E[h(X_n)\psi_L(X_n)] = E[h(X)\psi_L(X)].$$

D'autre part, par le lemme de convergence monotone, il existe L_0 tel que si
$L \geq L_0$, on a

$$E[1-\psi_L(X)] \leq P(| X |\geq L) \leq \varepsilon.$$

Ainsi, si $L \geq L_0$ et n suffisamment, grand on a aussi $E[1-\psi_L(X_n)] \leq \varepsilon$ puisque $E[1-\psi_L(X_n)]$ tend vers $E[1-\psi_L(X)]$. Maintenant, on regroupe le tout, en écrivant pour tout $L \geq L_0$ et n assez grand

$$\begin{aligned} | E[h(X_n) - h(X)] | &\leq | E[h(X_n)(1-\psi_L(X_n))] | \\ &+ | E[h(X_n)\psi_L(X_n) - h(X)\psi_L(X)] | \\ &+ | E[h(X)(1-\psi_L(X))] | \\ &\leq 2\varepsilon M + | E[h(X_n)\psi_L(X_n) - h(X)\psi_L(X)] | . \end{aligned}$$

En prenant la limite quand n tend vers l'infini, on aura
$\lim_{n\to\infty} E[h(X_n) - h(X)] \leq 2\varepsilon M$. Ceci étant vrai pour tout $\varepsilon > 0$, le résultat en découle.

3.3 Théorème de Lévy

Proposition 3.2 Soit $(X_n)_{n\in N}$ une suite de variables aléatoires réelles. Si $X_n \to^{(loi)}_{n\to\infty} X$ alors $\lim_n \phi_{X_n}(t) = \phi_X(t)$ pour tout $t \in \mathbb{R}$, où ϕ est la fonction caractéristique.

Démonstration Puisque $x \to e^{itx}$ est continue bornée, donc la convergence en loi implique que $\lim_n E[e^{itX_n}] = E[e^{itX}]$. D'où le résultat.
La proposition reste évidemment vraie si on remplace les variables aléatoires réelles par des vecteurs aléatoires de $\mathbb{R}^d$. Pour la réciproque, si on dispose de variables aléatoires réelles et que $\lim_{n\to\infty} \phi_{X_n} = \phi$, il nous faut nous assurer que ϕ est une fonction continue en 0 et c'est l'énoncé du théorème suivant.

Théorème de continuité de Lévy Soit $(X_n)_{n\in N}$ une suite de variables aléatoires réelles. Si $\phi_{X_n} \to_{n\to\infty} \phi$ et ϕ continue en 0, alors il existe une variable aléatoire X telle que $\phi_X = \phi$ et $X_n \to_{n\to\infty}^{(loi)} X$.

Remarques

- En pratique, si on calcule la limite de la suite de fonctions caractéristiques et on tombe sur une fonction caractéristique usuelle (d'une variable aléatoire qu'on connait), alors ce théorème nous permet de conclure que notre suite de variables aléatoires converge en loi vers la variable aléatoire en question. On peut donc appliquer ce théorème lorsque la limite n'est pas une fonction caractéristique qu'on connait, il faut cependant vérifier que la fonction limite est continue en 0 pour pouvoir affirmer que la suite de variables aléatoires converge en loi.

- Le théorème de Lévy s'étend également à $\mathbb{R}^d$ en faisant l'analogue de la preuve ci-dessus. On peut donc garder en tête que si $(X_n)_{n\in N}$ est une suite de vecteurs aléatoires de $\mathbb{R}^d$ et si $\phi_{X_n(t)}$ converge vers $\phi(t)$ pour $t \in \mathbb{R}^d$ et que ϕ est continue en 0, alors (X_n) converge en loi vers un vecteur aléatoire X, dont la fonction caractéristique est donnée par ϕ.

Définition 3.2 (Famille tendue) On dit qu'une suite de variables aléatoire réelles $(X_n)_{n\in N}$ est tendue si $\forall \varepsilon > 0$, il existe un compact K tel que $P(X_n \in K) \geq 1 - \varepsilon$ pour tout $n \in N$.

Remarques

- Si la suite consiste en une seule variable aléatoire (ou un nombre fini) alors elle est tendue.
En effet, soit $\varepsilon > 0$ et notons que comme F_X tend vers 0 en $-\infty$; alors il existe $M_1 < 0$ tel que $F_X(M_1) \leq \frac{\varepsilon}{2}$. D'autre part, comme F_X tend vers 1 en

∞ ; alors il existe $M_2 > 0$ tel que $F_X(M_2) \geq 1-\frac{\varepsilon}{2}$, on a donc $P(X \in [M_1, M_2]) = F_X(M_1) - F_X(M_1^-) \geq 1-\varepsilon$. Comme $[M_1, M_2]$ est un compact, cela montre que la famille constituée de X seule est tendue.

- La réunion de deux familles tendues forme une famille tendue.

- Une sous suite d'une suite tendue est tendue.

- Une suite de vecteurs aléatoires $(X_n)_{n\in N}$ de $\mathbb{R}^d$ est tendue si toutes les suites de coordonnées sont tendues.

Lemme 3.1 Soit $(X_n)_{n\in N}$ une suite de variables aléatoires réelles. Si $X_n \to^{loi} X$, $\phi_{X_n} \to_{n\to\infty} \phi$ et ϕ continue en 0, alors la suite $(X_n)_{n\in N}$ est tendue.

Démonstration Soit $u > 0$ on a par Fubini

$$\frac{1}{u}\int_{-u}^{u}(1-\phi_{X_{n}(t)})dt = E[\frac{1}{u}\int_{-u}^{u}(1-e^{itX_n})dt] = 2E[1-\frac{sin(uX_n)}{uX_n}].$$

D'autre part, si $\mid x \mid\geq 2$ alors $\mid \frac{sinx}{x} \mid\leq \frac{1}{2}$; donc

$$P(\mid X_n \mid\geq \frac{2}{u}) = 2(1-\frac{1}{2})E[\mathbf{1}_{\{\mid X_n\mid\geq\frac{2}{u}\}}] \leq 2E[1-\frac{sin(uX_n)}{uX_n}].$$

On déduit alors que pour tout $n \in N$, on a

$$P(\mid X_n \mid\geq \frac{2}{u}) \leq \frac{1}{u}\int_{-u}^{u}\mid 1-\phi_{X_n}(t)\mid dt.$$

Par le lemme de Fatou, on a

$$\limsup_n P(\mid X_n \mid\geq \frac{2}{u}) \leq \limsup_n \frac{1}{u}\int_{-u}^{u}\mid 1-\phi_{X_n}(t)\mid dt \leq \frac{1}{u}\int_{-u}^{u}\mid 1-\phi_X(t)\mid dt$$

Or ϕ est continue en 0, alors

$$\limsup_{u\longrightarrow 0}\limsup_{n\longrightarrow\infty} P(\mid X_n \mid\geq \frac{2}{u}) \leq \limsup_{u\longrightarrow 0}\frac{1}{u}\int_{-u}^{u}\mid 1-\phi_X(t)\mid dt = 0.$$

Ainsi, si $\varepsilon > 0$, on peut trouver $M > 0$ tel que

$$\limsup_{n\longrightarrow\infty} P(\mid X_n \mid\geq M) < \varepsilon$$

et on déduit que la suite est tendue.

Lemme 3.2 : Soit $(X_n)_{n\in N}$ une suite de variables aléatoires réelles, bornées et tendues. Alors il existe une sous-suite $(X_{nk})_{k\in N}$ et une variable aléatoire X telle que

$$X_{nk} \longrightarrow_{k\longrightarrow\infty} X.$$

Démonstration Soit la suite de fonctions de répartition associées. Pour tout q rationnel, la suite $(F_{X_n}(q))_{n\in N}$ est une suite réelle bornée, et par le théorème de Bolzano-Weierstrass, elle admet une sous-suite convergente. Par le procédé diagonal, on forme une sous-suite $(F_{n_k})_{k\in N}$ qui converge en tout point rationnel. On note G la fonction limite obtenue qu'on étend à $\Re$ en posant

$$F(x) = \inf\{G(q) : q \in Q \bigcap (x, \infty)\}.$$

F est croissante et continue à droite. Montrons que F_{n_k} converge vers F en tout point de continuité de F.
Soit $x \in \Re$ un point de continuité de F et $\varepsilon > 0$. Comme F est continue en x, alors il existe $\eta > 0$ tel que $\mid F(x) - F(y) \mid\leq \varepsilon$ pour tout $y \in [x-\eta, x+\eta]$.
Soit $x - \eta \leq r \leq x \leq s \leq x + \eta$ deux rationnels. Alors

$$F_{n_k}(r) \leq F_{n_k}(x) \leq F_{n_k}(s)$$

et en prenant la limite sur k, on obtient

$$F(r) \leq \liminf_k F_{n_k}(x) \leq \limsup_k F_{n_k}(x) \leq F(s).$$

Or $F(r) \geq F(x) - \varepsilon$ et $F(s) \leq F(x) + \varepsilon$, on déduit donc

$$F(x) - \varepsilon \leq \liminf_k F_{n_k}(x) \leq \limsup_k F_{n_k}(x) \leq F(x) + \varepsilon,$$

pour tout $\varepsilon > 0$, ce qui montre la convergence de la sous-suite $(F_{n_k})_{k\in N}$ vers F en tout point de continuité de F.
Il nous reste à montrer que $\lim_{-\infty} F(x) = 0$ et $\lim_{\infty} F(x) = 1$ pour déduire que F est la fonction de répartition d'une variable aléatoire X. On va se servir du fait que la suite est tendue. Soit $\varepsilon > 0$. Il existe $M > 0$ tel que pour tout $k \in N$ on a $F_{X_{n_k}}(t) \leq \varepsilon$ pour tout $t \leq -M$ et $F_{X_{n_k}}(t) \geq 1 - \varepsilon$ pour tout $t \geq M$. En prenant la limite sur $k \in N$, on déduit que $F_X(t) \leq \varepsilon$ pour tout $t \leq -M$ et $F_{X_{n_k}}(t) \geq 1 - \varepsilon$ pour tout $t \geq M$ (avec t point de continuité). En prenant la limite le long des points de continuité, on déduit que $\lim_{-\infty} F(x) \leq \varepsilon$ et $\lim_{\infty} F(x) \geq 1-\varepsilon$, ceci étant vrai pour tout $\varepsilon > 0$, on déduit le résultat.
Finalement, la caractérisation de la convergence en loi à l'aide de la fonction de répartition permet de conclure la preuve.
On peut maintenant entamer la preuve du théorème de Lévy.

Démonstration du théorème de Lévy Par le lemme **2**, on a que la suite $(X_n)_{n\in N}$ est tendue. Soit $(X_{n_l})_{l\in N}$ une sous-suite quelconque de $(X_n)_{n\in N}$. Puisqu'elle est également tendue, on peut en extraire une sous-suite $(X_{n_{l_k}})_{k\in N}$ qui converge en loi vers une certaine variable aléatoire X.
Par la proposition **3.2**, la fonction caractéristique de $X_{n_{l_k}}$ converge vers ϕ_X. Par l'unicité de la limite, $\phi_X = \phi$. Finalement, on a montré que toute sous-suite de $(X_n)_{n\in N}$, admet une sous-suite qui converge en loi vers X. On déduit donc que (X_n) converge vers X.

3.4 Théorème central limite

Etant donné une suite de variables aléatoires réelles indépendante de même loi, dans L^1, nous avons la loi forte des grands nombres qui implique que
$\frac{1}{n}\sum_{j=1}^n X_j \to_{n\to\infty}^{p.s} E[X_1]$.
On aimerait quantifier plus cette convergence et comprendre à quelle vitesse la convergence a lieu. Pour cela, on cherche à étudier la différence $\sum_{j=1}^n (X_j - E(X_j))$. Supposons que les variables sont en plus dans L^2,

$$E[(\sum_{j=1}^n (X_j - E(X_j))^2] = \sum_{j=1}^n E[(X_j - E(X_j))^2],$$

par linéarité. Comme les X_j ont la même loi, elles ont en particulier la même variance qu'on note σ^2. On en déduit que

$$E[(\sum_{j=1}^n (X_j - E(X_j))^2] = n\sigma^2.$$

Ceci suggère que l'ordre de grandeur de $\sum_{j=1}^n (X_j - E(X_j))$ est $\sigma\sqrt{n}$. Pour rendre cela plus précis

$$\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^n (X_j - E(X_j))$$

est de l'ordre de 1.
Prenons un exemple pour comprendre. Supposons que les X_j sont de Rademarcher indépendantes de paramètre $\frac{1}{2}$ i.e.
$P(X_j = -1) = P(X_j = 1) = \frac{1}{2}$. Notons que $E[X_j] = 0$ et $Var[X_j] = \sigma^2 = 1$ et

calculons la fonction caractéristique de $S_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} X_j$.

$$\phi_{S_n}(t) = \prod_{j=1}^{n} \phi_{X_j}(\frac{t}{\sqrt{n}}) = \prod_{j=1}^{n} \frac{e^{\frac{it}{\sqrt{n}}} + e^{-\frac{it}{\sqrt{n}}}}{2} = \cos^n(\frac{t}{\sqrt{n}}) = e^{n \ln \cos(\frac{t}{\sqrt{n}})}$$

Or $\cos(\frac{t}{\sqrt{n}})$ est équivalent à $1 - \frac{t^2}{2n}$ quand $n \to \infty$ et $\ln(1 - \frac{t^2}{2n})$ est équivalent à $-\frac{t^2}{2n}$. Ainsi, on déduit que

$$\phi_{S_n}(t) \to_{n \longrightarrow \infty} e^{-\frac{t^2}{2}},$$

qui n'est autre que la fonction caractéristique de la loi gaussienne.

Théorème 3.1 (Théorème central limite) Soit $(X_n)_{n \in N}$ une suite de variables aléatoires réelles indépendantes toutes de mêmes lois, dans L^2, telles que $E(| X_n |^2) < \infty$. On note $M = E(X_1)$ la moyenne et $\sigma^2 = Var(X_1)$ la variance. Alors on a

$$\frac{X_n + ... + X_n - nM}{\sigma \sqrt{n}} \longrightarrow_{n \to \infty}^{(loi)} N(0,1).$$

Démonstration On peut supposer sans perdre de généralité que les X_j sont centrées i.e. $E(X_j) = 0$, sinon on travaille avec $X_j - E(X_j)$, de même, on peut supposer que $\sigma = 1$, sinon on travaille avec $\frac{1}{\sigma} X$ à la place de X. Soit $S_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} X_j$ et notons que par indépendance des X_j, on peut écrire

$$\phi_{S_n}(t) = \prod_{j=1}^{n} \phi_{X_j}(\frac{t}{\sqrt{n}}) = \phi_{X_1}^{n}(\frac{t}{\sqrt{n}}).$$

Or par le développement de la fonction exponentielle autour de 0, on peut écrire :

$$E(e^{i \frac{t}{\sqrt{n}} X_1}) = 1 + i \frac{t}{\sqrt{n}} E(X_1) - \frac{t^2}{2n} E(X_1^2) + o(\frac{1}{n}) = 1 - \frac{t^2}{2n} + o(\frac{1}{n}).$$

On obtient alors

$$\lim_{n \to \infty} \phi S_n(t) = \lim_{n \to \infty} (1 - \frac{t^2}{2n} + o(\frac{1}{n}))^n = e^{-\frac{t^2}{2}},$$

qui est la fonction caractéristique d'une loi normale centrée réduite. Par le théorème de Lévy, on peut conclure que S_n converge en loi vers $N(0,1)$.

Remarques

- Parfois on énonce le théorème sans normalisation en écrivant

$$\frac{X_1 + ... + X_n - nM}{\sqrt{n}} \longrightarrow_{n\to\infty}^{(loi)} N(0, \sigma^2).$$

- Ainsi on peut traduire la conclusion du théorème central limite en termes des différentes caractérisations de la convergence en loi. On peut donc affirmer que si $X_1, ..., X_n$ sont des variables aléatoires indépendantes de même loi dans $\mathbf{L}^2$, $(E(\mid X_n \mid^2) < \infty)$ de variance σ^2, alors pour tout $a < b$

$$P(\sum_{j=1}^{n} X_j \in [nE(X_1) + a\sqrt{n}, nE(X_1) + b\sqrt{n}]) \longrightarrow_{n\to\infty} \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2\sigma^2}} dx.$$

4

Convergence étroite dans des espaces fonctionnels et Principe d'invariance de Donsker

4.1 Introduction

Dans ce chapitre nous allons définir la relative compacité, donner ses propriétés et étudier cette relative compacité sur les espaces fonctionnels. Ensuite, donner le théorème de Donsker sur les espaces fonctionnels.

4.2 Relative compacité et équitension

Définition 4.1 [relative compacité] Soit Π une famille de mesure de probabilité sur (S, $\mathbb{F}$). On dit que Π est relativement compacte ; si toute suite d'élément de Π contient une sous suite qui converge faiblement ; c'est à dire, pour toute suite $\{P_n\}$ de Π, il existe une sous suite $\{P_{n'}\}$ et une mesure de probabilité Q (définie sur (S, $\mathbb{F}$), mais pas nécessairement un élément de Π) telles que $P_{n'} \Rightarrow Q$ (ici $\Rightarrow$ signifie convergence faible).

Nous devons décider si oui ou non une famille donnée Π est relativement compacte. Par exemple, supposons qu'on ait des mesures de probabilités P_n et P sur $(C, \mathbb{F})$, que la distribution fini dimensionnelles de P_n convergent faiblement vers celles de P. Supposons, cependant, que nous savons aussi que P_n est relativement compacte. Alors toute sous suite $P_{n'}$ contient une future sous suite $P_{n''}$ qui converge faiblement vers une limite Q. La distribution finie dimensionnelle de Q doit être faiblement bornée que celle de $P_{n''}$ et donc doit coïncider avec la distribution finie dimensionnelle de P (pour $(t_1, ..., t_k)$, $P_{n''}\pi^{-1}_{t_1,...,t_k} \Rightarrow Q\pi^{-1}_{t_1,...,t_k}$ et $P_n\pi^{-1}_{t_1,...,t_k} \Rightarrow P\pi^{-1}_{t_1,...,t_k}$, donc

$Q\pi^{-1}_{t_1,\dots,t_k} = P\pi^{-1}_{t_1,\dots,t_k}$). Mais alors, puisque une probabilité de mesure sur C est complement déterminée par sa distribution de Q de dimension finie (la dimension finie forme les ensembles de classes déterminantes), Q et P doivent eux-même coïncider. Donc toute sous suite P_n contient une future sous suite qui converge faiblement vers P.
Notons que, si $\{P_n\}$ converge faiblement vers P, alors elle est relativement compacte.
Supposons que $\{P_n\}$ est relativement compacte et que, pour tout $(t_1,\dots,t_k)$, $P_n\pi^{-1}_{t_1,\dots,t_k}$ convergent faiblement vers la mesure de probabilité $\mu_{t_1,\dots,t_k}$ sur $(\mathbf{R}^k, \mathbb{R}^k)$: Notons que nous ne supposons pas au départ que $\mu_{t_1,\dots,t_k}$ sont des distributions de dimension finie d'une mesure de probabilité sur (C, $\mathbb{F}$). Il s'ensuit toujours que chaque sous suite $\{P_{n'}\}$ contient une future sous suite $\{P_{n''}\}$ qui converge faiblement vers sa limite. Puisque cette limite doit avoir $\mu_{t_1,\dots,t_k}$ comme sa distribution de dimension finie, elle est unique. En conclusion $\{P_n\}$ converge faiblement vers P. Ces résultats fournissent une puissante technique pour prouver la convergence faible dans C et dans d'autres espaces fonctionnels. On prouve d'abord que les lois fini-dimensionnelles converge faiblement et alors la suite en question est relativement compacte. Pour utiliser cette méthode il faut un critère efficace pour la relative compacité.
Considerons en premier $\mathbf{R}$. Soit Π une famille de mesures de probabilites sur $(\mathbf{R}, \mathbb{R})$. Donnons une sous suite $\{P_n\}$ de Π, nous pouvons appliquer une sous suite $\{F_n\}$ correspondant à une distribution de fonctions classique. Il existe une sous suite $\{F_{n'}\}$ et une fonction F telles que

$$F_{n'}(x) \to F(x), \tag{4.1}$$

est valable pour tous ses points de continuité x. La fonction F peut être considérée comme continue à droite, dans ce cas il existe sur $(\mathbf{R}, \mathbb{R})$ une mesure finie μ telle que

$$\mu(a,b] = F(b) - F(a). \tag{4.2}$$

Maintenant il peut arriver que $\mu(\mathbf{R}) < 1$. Par exemple, si P_n correspond à une unité de masse au point n, alors F est identiquement nulle, peut importe que la sous suite $F_{n'}$ soit sélectionnée, donc $\mu(\mathbf{R}) = 0$. Si P_n est une distribution uniforme sur l'intervalle $[-n, n]$, alors $F(x) = \frac{1}{2}$ est seulement possible; encore $\mu(\mathbf{R}) = 0$. Dans cet exemple, la masse "part vers l'infinie" dans un sens évident et intuitif.
Supposons cependant, que la mesure μ déterminée par F dans (4.2) vérifie $\mu(\mathbf{R}) = 1$. Alors μ est une mesure de probabilité et puisque (4.1) tient

aux points de continuité de F, nous avons $P_{n'} \Rightarrow \mu$. Ainsi Π, sera relativement compact si nous nous assurons d'une manière ou d'une autre que chacune de ces mesures μ que nous rencontrons satisfait $\mu(\mathbf{R}) = 1$. Maintenant $\mu(\mathbf{R}) = 1$, si pour tout ε positif, il existe a et b tels que $\mu(a,b] \geq 1-\varepsilon$. Supposons que pour tout ε positif, il existe a et b tels que

$$P_n(a,b] > 1-\varepsilon, \qquad n = 1,2,... \tag{4.3}$$

Puisque (4.3) reste vraie si a diminue et b augmente, on peut amener a et b aux points de continuité de F, auquel cas (4.3) et (4.1) impliquent $\mu(a,b] \geq 1-\varepsilon$.
En conclusion la famille Π est une suite de mesures de probabilités sur $(C,\mathbb{F})$ et est relativement compacte si pour tout ε positif, il existe a et b tels que $P_n(a,b] > 1-\varepsilon$, pour tout P dans Π, une condition qui a pour effet d'empêcher la fuite de la masse évoquée ci-dessus. D'autre part, si cette condition échoue, alors il existe un ε positif tel que pour tout a et b, $P(a,b] > 1-\varepsilon$ pour P dans Π. Choisissons un P_n dans Π avec $P_n(a,b] \leq 1-\varepsilon$. Si la sous suite $\{P_n\}$ converge faiblement pour une mesure de probabilité Q, nous aurions, pour tout x,

$$Q(-x,x) \leq \liminf_{n'} P_{n'}(-x,x) \leq \liminf_{n'} P_{n'}(-n',n'] \leq 1-\varepsilon,$$

ce qui est impossible. Donc la condition est nécessaire et suffisante. Puisque un intervalle $(a,b]$ est un compact fermé, et puisque l'ensemble compact peut être enfermé dans un tel intervalle, la condition peut être remaniée : une famille Π de mesures de probabilité sur $(\mathbf{R},\mathbb{R})$ est relativement compacte si et seulement si, pour tout ε positif il existe un sous compact K tel que $P(K) > 1-\varepsilon$ pour tout P dans Π. Maintenant la condition a un sens dans un espace métrique arbitraire.
Les théorèmes suivants, dus à **Prohorov** vont étendre les résultats précédents à des espaces métriques arbitraires et à des espaces séparables et complets.

Théorème 4.1 [Billingsley] *Si Π est tendue alors elle est relativement compacte.*

Théorème 4.2[Billingsley] *Supposons que S est séparable et complet. Si Π est relativement compacte, alors elle est tendue.*

Démonstration Nous nous tournons d'abord vers la démonstration de la moitié directe, qui est plus utile dans les applications. Nous démontrons successivement le résultat pour $\mathbf{R}^k$, pour $\mathbb{R}^\infty$. Chacune des cas est

traité en le réduisant à celui qui précède.
Le cas $\mathbf{R}^k$. Ici la démonstration est pratiquement la même comme celle déjà donnée pour $\mathbf{R}$. Si $\{P_n\}$ est une suite de Π, alors la fonction $\{F_n\}$ correspondant à une distribution de fonction contient une sous suite $\{F_{n'}\}$ telle que

$$F_{n'}(x) \to F(x); \tag{4.4}$$

en tous points de continuités de F, où F est une fonction continue formulée ci-dessus. Il existe sur $(\mathbf{R}^k, \mathbb{R}^k)$ une mesure μ telle que $\mu(a,b]$ est F différenciable autour des sommes du rectangle de k-dimension $(a,b]$.
Maintenant $P_{n'} \Rightarrow \mu$ résulte si nous prouvons que $\mu(\mathbf{R}^k) = 1$. Etant donné ε, choisissons un ensemble compact K dans $\mathbf{R}^k$ tel que $P_{n'}(K) > 1 - \varepsilon$ pour tout n', ce qui est possible car Π est tendue. Maintenant choisissons a et b tels que $K \subset (a,b]$ tels que les 2^k sommets de $(a,b]$ sont des points de continuité de F (cela est possible car seuls un grand nombre d'hyperplan $(k-1)$-dimensionnels parallèles peuvent avoir une μ-mesure positive). Puisque $P_{n'}(a,b]$ est $f_{n'}$ différenciable autour du sommet de $(a,b]$, (4.4) implique $P_{n'}(a,b] \to \mu(a,b]$. Pour $P_{n'}(a,b] \geq 1-\varepsilon$, $P_{n'}(K) > 1-\varepsilon$, ça suit que $\mu(a,b] \geq 1-\varepsilon$. Puisque ε est arbitraire, $\mu(\mathbf{R}^k) = 1$. Donc Π est relativement compacte.
Pour le cas $\mathbf{R}^\infty$ nous avons besoin du lemme suivant.

Lemme 4.1 *Si Π est une famille tendue sur $(S, \mathbb{F})$ et si h est une cartographie continue de S vers S', alors $\{Ph^{-1} : P \in \Pi\}$ est une famille tendue sur $(S', \mathbb{F}')$.*

Preuve Donnons ε, choisissons dans S un ensemble compact K tel que $P(K) > 1-\varepsilon$ pour tout P dans Π. Si $K' = hK$, alors K' est un compact et $h^{-1}K' \supset K$, donc $Ph^{-1}(K') > 1-\varepsilon$ pour tout P dans Π.
Si Π est une famille tendue sur $(\mathbf{R}^\infty, \mathbb{R}^\infty)$, Par lemme, $\{P\pi_k^{-1} : P \in \Pi\}$ est une famille tendue sur $(\mathbf{R}^k, \mathbb{R}^k)$, pour tout k. On peut sélectionner d'une suite donnée $\{P_n\}$ dans Π, une sous suite $\{P_{n'}\}$ telle que $P_{n'}\pi_k^{-1}$ converge faiblement vers une mesure de probabilité μ_k sur $(\mathbf{R}^k, \mathbb{R}^k)$. On peut choisir une suite $\{P_{n'}\}$, donc $P_{n'}\pi_k^{-1} \Rightarrow \mu_k$ vaut pour tout k simultanément. Puisque les mesures μ_k satisfont évidemment les conditions de cohérence du théorème d'existence de Kolmogorov. Il existe une mesure de probabilité Q sur $(\mathbf{R}^\infty, \mathbb{R}^\infty)$ telle que $Q\pi_k^{-1} = \mu_k$ pour tout k. Alors $P_{n'}\pi_k^{-1} \Rightarrow Q\pi_k^{-1}$ pour tout k, donc les lois fini-dimensionnelles de $P_{n'}$ convergent vers celles de Q, implique que $P_{n'} \Rightarrow Q$. L'équitension implique la relativité compacité dans $\mathbf{R}^\infty$

4.3 L'espace D

Soit $\mathbf{D} = \mathbf{D}[0,1]$ l'espace de fonctions x sur $[0,1]$ qui sont continues à droite et ont une limite à gauche :

i) Pour $0 \leq t < 1$, $x(t+) = \lim_{s \downarrow t} x(s)$ existe, et $x(t+) = x(t)$.

ii) Pour $0 < t \leq 1$, $x(t-) = \lim_{s \uparrow t} x(s)$ existe.
Une fonction x est discontinue en t, si $x(t-)$ est différent de $x(t+)$. La discontinuité d'un élément de $\mathbf{D}$ est de première espèce.
Pour $x \in \mathbf{D}$, et $T_0 \subset [0,1]$, notons que

$$W_x(T_0) = \sup\{|x(s) - x(t)| : s, t \in T_0\}.$$

Le module de continuité de x est défini par

$$W_x(\delta) = \sup_{0 \leq t \leq 1-\delta} W_x[t, t+\delta].$$

Une fonction continue sur $[0,1]$ est uniformément continue. Le lemme suivant donne un résultat correspondant à un élément de $\mathbf{D}$.

Lemme 4.2 Pour tout $x \in \mathbf{D}$ et pour tout ε positif, il existe des points $t_0, t_1, ..., t_r$ tels que

$$0 = t_0 < t_1 < ... < t_r = 1,$$

et

$$W_x[t_{i-1}, t_i) < \varepsilon, \qquad i = 1, 2, .., r.$$

4.3.1 Compacité dans D

Passons maintenant au problème de caractérisation du sous ensemble compact de $\mathbf{D}$. Avec le module $W'_x(\delta)$ défini par

$$W'_x(\delta) = \inf_{\{t_i\}} \max_{0 < i \leq r} W_x[t_{i-1}, t_i).$$

Avant d'énoncer le théorème suivant nous allons définir la topologie de Skorohod.
Soit Λ la classe des applications strictement croissantes et continues sur $[0,1]$. Si $\lambda \in \Lambda$ alors $\lambda(0) = 0$ et $\lambda(1) = 1$. Pour tout x, y dans $\mathbf{D}$ on définit une distance $d(x,y)$ finie ; pour tout $\varepsilon > 0$, il existe un $\lambda \in \Lambda$ tel que

$$\sup_t | \lambda t - t | \leq \varepsilon$$

et

$$\sup_t \mid x(t) - y(\lambda t) \mid \leq \varepsilon.$$

Notons que d_0 est la distance par rapport à $\lambda(0)$ et d est la distance par rapport à $\lambda(1)$.

Théorème 4.3 *A est un ensemble compact fermé dans la topologie de Skorohod si et seulement si*

$$\sup_{x\in A}\sup_t \mid x(t) \mid < \infty \tag{4.5}$$

et

$$\lim_{\delta\to 0}\sup_{x\in A} W'_x(\delta) = 0. \tag{4.6}$$

Notons que, $W'_x(\delta) \leq W_x(2\delta)$, *si* $\delta < \frac{1}{2}$, ( ) est plus faible que

$$\lim_{\delta\to 0}\sup_{x\in A} W_x(\delta) = 0. \tag{4.7}$$

Démonstration Pour prouver la suffisance de ces conditions, il suffit, puisque **D** est d_0-complet, de montrer que A est totalement borné par rapport à d_0. Montrons d'abord que (4.6) et (4.7) impliquent que A est borné par rapport à la métrique d.
Soit un ε positif, choisissons un entier k tel que $\frac{1}{k} < \varepsilon$ et $W'_x(\frac{1}{k}) < \varepsilon$, pour tout x dans A. Prenons H comme un ensemble fini dans l'intervalle $[-\alpha, \alpha]$, où

$$\alpha = \sup_{x\in A}\sup_t \mid x(t) \mid .$$

Soit B l'ensemble fini des y dans D qui prennent sur chacun des intervalles $[\frac{u-1}{k}, \frac{u}{k}]$, une valeur constante de H et satisfont $y(1) \in H$. On va montrer que B est un ε-net par rapport à d.
Choisissons $x \in A$, utilisons l'inéquation $W'_x(\frac{1}{k}) < \varepsilon$ pour tous points t_i, $0 = t_0 < t_1 < ... < t_r = 1$, tels que

$$t_i - t_{i-1} > \frac{1}{k} \tag{4.8}$$

et

$$W'_x[t_{i-1}, t_i) < \varepsilon, \quad 0 < i \leq r. \tag{4.9}$$

Choisissons des entiers u_i, tels que $\frac{u_i}{k} \leq t_i < \frac{(u_i+1)}{k}$, $i = 0, 1, ..., r$. Puisque les u_i doivent être distincts par (4.8), il y a dans Λ un λ qui porte $\frac{u_i}{k}$ dans t_i et est linéaire entre ses points. Choisissons dans B un point y tel que

$$\mid y(\frac{u}{k}) - x(\lambda\frac{u}{k}) \mid < \varepsilon, \qquad 0 \leq u \leq k. \tag{4.10}$$

Puisque un intervalle

$$[\lambda\frac{u}{k},\lambda\frac{u+1}{k})$$

doit être à l'intérieur de l'un des intervalles

$$[\lambda\frac{u_i}{k},\lambda\frac{u_{i+1}}{k})=[t_i,t_{i+1}),$$

la fonction $x(\lambda t)$ ne peut pas, de (4.9), varier de plus que ε lorsque t varie sur un intervalle $[\frac{u}{k},\frac{u+1}{k})$. Puisque y est une constante sur des intervalles de cette forme, (4.10) implique $\mid y(t)-x(\lambda t)\mid< 2\varepsilon$ pour tout t. Par construction,

$$\mid \lambda\frac{u_i}{k}-\frac{u_i}{k}\mid=\mid t_i-\frac{u_i}{k}\mid<\frac{1}{k}<\varepsilon;$$

par linéarité, $\sup_t \mid \lambda t-t\mid<\varepsilon$. Donc $d(x,y)<2\varepsilon$. En conclusion B est un 2ε-net dans le sens de d.
Donnons un η positif, choisissons δ de sorte que $0<\delta<\frac{1}{4}$ et pour que $\delta+W'_x(\delta)<\eta$ tient pour tout $x\in A$. Puis choisissons ε pour que $0<2\varepsilon<\delta^2$. Si B est un ensemble qui vient d'être construit comme un ε-net pour A au sens de la métrique d; alors par le lemme 4.2, B est un η-net pour A au sens de la métrique d_0. Donc A est d_0 totalement borne et donc puisque D est d_0 complet, la fermeture de A est compact.
Cela prouve la suffisance de (4.5) et (4.6). Si est compact, alors il est borné, et (4.5) suit immédiatement (notons que $\sup_t\mid x(t)\mid$ est une distance, au sens soit de d ou de d_0, forment x à la fonction identiquement nulle). Le théorème actuel diffère du théorème Arzela-Ascoli

$$\sup_{x\in A}\mid x(t)\mid<\infty$$

et (4.6) implique (4.5). Il n'est pourtant pas difficile de prouver que (4.5) suit si (4.6) et (4.11) sont vraies pour tout valeur de t.
Il reste à prouver la nécessite de (4.6). Par $\lim_\delta W'_x(\delta)=0$, $W'_x(\delta)$ va vers 0 avec δ pour tout x. Si nous prouvons que $W'_x(\delta)$ est semi-continue supérieur en x pour tout δ, alors la convergence sera uniforme sur un ensemble compact et (3.6) suivra.
Soit x, δ, et ε donnés. Pour la semi continuité supérieure, nous devons trouver un η tel que $d(x,y)<\eta$ implique

$$W'_y(\delta)<W'_x(\delta)+\delta. \tag{4.11}$$

Choisissons d'abord les points t_i vérifiant $0=t_0<t_1<...<t_r=1$, $t_i-t_{i-1}>\delta,\quad i=1,2,...,r$ et

$$W_x[t_i,t_{i-1})<W'_x(\delta)+\frac{\varepsilon}{2},\qquad i=1,2,...,r. \tag{4.12}$$

Maintenant choisissons η suffisamment petit que

$$t_i - t_{i-1} > \delta + 2\eta, \qquad i = 1, 2, ..., r,$$

et

$$\eta < \frac{1}{4}\varepsilon. \tag{4.13}$$

Si $d(x,y) < \eta$, alors, pour un λ dans Λ, nous avons

$$\sup_t \mid \lambda t - t \mid = \sup_t \mid \lambda^{-1} t - t \mid < \eta$$

et

$$\sup_t \mid y(t) - x(\lambda t) \mid \leq \eta. \tag{4.14}$$

Posons $s_i = \lambda^{-1} t_i$. Alors

$$s_i - s_{i-1} > t_i - t_{i-1} - 2\eta > \delta. \tag{4.15}$$

De plus, si s et t appartiennent tous deux à $[s_{i-1}, s_i)$, alors λs et λt les deux se trouvent dans $[t_{i-1}, t_i)$ et en conclusion, par (4.12),(4.13), et (4.14), $\mid y(s) - y(t) \mid < W'_x(\delta) + \varepsilon$. Donc

$$W_y[s_{i-1}, s_i) < W'_x(\delta) + \varepsilon,$$

qui, compte tune de (4.15), implique (4.11). Ce qui complète la démonstration.

4.4 L'espace C

Soit $\mathbf{C} = \mathbf{C}[0,1]$ l'espace des fonctions continues sur $[0,1]$, la distance entre deux éléments $x = x(t)$ et $y = y(t)$ de $\mathbf{C}$ est donnée par

$$\rho(x,y) = \sup_t |x(t) - y(t)|;$$

et c'est facile de vérifier que ρ est une métrique.
L'espace $\mathbf{C}$, est séparable, l'ensemble dénombrable dense est constitué des fonctions polygonales qui sont linéaires sur chaque sous intervalle $[\frac{i-1}{k}, \frac{i}{k}]$, $i = 1, ..., k$, pour un entier k.
Si $\{x_n\}$ est une suite de cauchy de $\mathbf{C}$, alors, pour chaque valeur de t, $\{x_n(t)\}$ est une suite fondamentale sur la droite et à donc une limite $x(t)$. IL est facile de montrer que la convergence $x_n(t) \to x(t)$ est uniforme sur t de

sorte que x la limite de $\{x_n\}$ est dans **C**. Donc **C** est un espace complet.
On définit le module de continuité d'un élément x de **C** par

$$W_x(\delta) = W(x,\delta) = \sup_{|s-t|\leq\delta} | x(s) - x(t) |, \quad 0 < \delta < 1;$$

puisque

$$| W_x(\delta) - W_y(\delta) | \leq 2\rho(x,y).$$

$W_x(\delta)$ est continue en x, pour δ fixé positif. Notons, puisque un élément de **C** est uniformément continu, nous avons

$$\lim_{\delta\to 0} W_x(\delta) = 0, \quad x \in \mathbf{C}.$$

Si A est un sous espace de **C**, alors A est un complet fermé si et seulement si

$$\sup_{x\in A} | x(0) | < \infty;$$

et

$$\lim_{\delta\to 0} \sup_{x\in A} W_x(\delta) = 0.$$

4.4.1 L'équitension dans C

Théorème 4.4 *Soit P_n, P des mesures de probabilités sur $(\mathbf{C}, C)$. Si les lois fini-dimensionnelles de P_n convergent faiblement vers P, et si $\{P_n\}$ est tendue, alors $P_n \Rightarrow P$.*
Si nous devons utiliser ce théorème pour la convergence faible dans C, nous devons donner une caractérisation de l'équitension.
Le module de continuité d'un élément x de C est défini par

$$W_x(\delta) = W(x,\delta) = \sup_{|s-t|\leq\delta} | x(s) - x(t) |, \quad 0 < \delta < 1.$$

Soit $\{P_n\}$ une suite de mesures de probabilités sur $(\mathbf{C}, C)$.

Théorème 4.5 Une suite $\{P_n\}$ est équitendue, si et seulement si, elle vérifie les deux conditions :

i) Pour tout η positif, il existe un a, tel que

$$P_n\{x : |x(0)| > a\} \leq \eta, \quad n \geq 1.$$

ii) Pour tout ε et η positifs, il existe un δ, avec $0 < \delta < 1$, et un entier n_0 tels que

$$P_n\{x : W_x(\delta) \geq \varepsilon\} \leq \eta, \quad n \geq n_0.$$

La condition (i) stipule que $\{P_n \pi_0^{-1}\}$ est tendue. Notons que $W_x(\delta)$ est continue en x, donc mesurable.

Démonstration Supposons $\{P_n\}$ est tendue. Soient ε, $\eta > 0$, on choisit un ensemble compact, tel que $P_n(K) > 1 - \eta$, pour tout n. Par le théorème d'Arzelà-Ascoli, nous avons $K \subset \{x : |x(0)| \leq a\}$ pour tout a assez grand, et $K \subset \{x : W_x(\delta) < \varepsilon\}$, pour δ suffisamment petit, donc (i) et (ii) sont vérifiées (avec $n_0 = 1$ dans(ii)). Cela prouve la nécessite de (i) et (ii).
P est une mesure de probabilité sur $(\mathbf{C}, C)$ qui est tendue, alors par la nécessite de (ii), pour tous ε et η, il existe un δ tel que

$$P\{x : W_x(\delta) \geq \varepsilon\} \leq \eta.$$

Si $\{P_n\}$ vérifie la condition (ii), alors l'inégalité en (ii) est valable pour un nombre fini de n précédant n_0 en diminuant δ si nécessaire. Ainsi nous pouvons supposer, pour prouver la suffisance, que le n_0 dans l'inegalite de (i) est 1.
Admettons que $\{P_n\}$ vérifie (i) et (ii). Etant donné η, choisissons a de sorte que, si

$$A = \{x : |x(0)| \leq a\},$$

alors $P_n(A) \geq 1 - \frac{1}{2}\eta$ pour tout n, et choisissons δ_k de sorte que, si

$$A_k = \{x : W_x(\delta_k) < \frac{1}{k}\},$$

alors $P_n(A_k) \geq 1 - \frac{\eta}{2^{k+1}}$ pour tout n. Si K est une fermeture de $A \bigcap_{k=1}^{\infty} A_k$, alors $P_n(K) \geq 1 - \eta$ pour tout n, et par le théorème de Arselà-Ascoli, K est un compact donc $\{P_n\}$ est équitendue.
Comme le montre la démonstration, nous pouvons exiger que $A = \{x : |x(0)| \leq a\}$, contient tous les $\{P_n\}$. Avec ce changement, le théorème est vrai si $\{P_n\}$ est remplacé par une famille arbitraire π. Dans l'application, nous prouverons souvent $A_k = \{x : W_x(\delta_k) < \frac{1}{k}\}$, avec δ, ε, et η remplacés successivement par $\frac{1}{2}\delta$, 3ε, et 9η.

4.5 Convergence de la marche aléatoire simple vers le mouvement brownien

4.5.1 La marche aléatoire simple

Soit $X_1, X_2, ...$ une suite de variables aléatoires réelles indépendantes et identiquement distribuées, uniformes sur [-1,1], définies sur un espace de probabilité (Ω, F, P). On s'intéresse au comportement de la marche aléatoire $(S_n, n \geq 0)$ associée définie par $S_0 = 0$ et

$$S_n = \sum_{i=1}^{n} X_i = X_1 + X_2 + ... + X_n, \qquad n \geq 1.$$

Par exemple, le théorème de la limite centrale montre que pour tout $a \leq b$, on a

$$P(a < \frac{S_n}{\sqrt{n}} < b) \to_{n\to\infty} \int_a^b \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx$$

ou de façons équivalente

$$\frac{S_n}{\sqrt{n}} \to_{n\to\infty}^{loi} N(0,1).$$

Où N est une variable aléatoire gaussienne centrée, réduite (on notera $N(m, \sigma^2)$ la loi gaussienne de moyenne m et variance σ^2). Ceci détermine la position de S_n lorsque $n \to \infty$. On interprète le résultat précédent en disant que la loi de S_n est proche de $N(m, \sigma^2)$.
Si l'on veut une information plus précise sur la trajectoire de $(S_n, n \geq 0)$, on peut s'intéresser à la position de la marche à des instants $n_1, n_2, n_3, ..., n_k$ rangés par ordre croissant.
On note $n_0 = 0$. Les accroissements

$$S_{n_i} - S_{n_{i-1}} = X_{n_{i-1}+1} + ... + X_{n_i}, \qquad 1 \leq i \leq k,$$

sont indépendants, et ont la même loi que $S_{n_i - n_{i-1}}$. Supposons que $n_1, n_2, n_3, ..., n_k$ sont des suites (indexées par n) et qu'il existe $t_0 = 0 \leq t_1 \leq ..... \leq t_k$ tels que

$$t_i = \lim_{n\to\infty} \frac{n_i}{n}.$$

Alors une nouvelle utilisation de la limite centrale (multidimensionnelle) montre alors que

$$(\frac{S_{n_i} - S_{n_{i-1}}}{\sqrt{n}}, \quad 1 \leq i \leq k) \to_{n\to\infty}^{loi} (N_1, .., N_k),$$

où les variables aleatoires $N_1,...,N_k$ sont indépendantes, gaussiennes, de variances respectives $t_1, t_2 - t_1, ..., t_k - t_{k-1}$. De façon equivanlente, on a

$$(\frac{S_{n_i}}{\sqrt{n}}, \quad 1 \leq i \leq k) \rightarrow^{loi}_{n\rightarrow\infty} (N_1, N_1 + N_2, .., N_1 + ... + N_k). \tag{4.16}$$

Ce qu'on peut interpréter comme la détermination des positions asymptotiques de la marche aléatoire considérée en un nombre fini d'instants $n_1, ..., n_k$.

Peut-on aller plus loin et déterminer le comportement asymptotique de la marche aléatoire dans son intégralité ? Au vu des résultats précédents, il est donc naturel de considérer une fonction aléatoire d'un paramètre réel positif t, correspondant à la marche aléatoire dont le temps est renormalisé par n, et l'espace par $\sqrt{n}$. Notons donc

$$S^{[n]}_t = \frac{S_{[nt]}}{\sqrt{n}}, \qquad t \geq 0.$$

Si l'on en effectue une simulation, on s'aperçoit que, lorsque n augmente, la fonction $S^{[n]}$ semble prendre la forme d'une fonction, de nature aléatoire. Cette fonction limite est appelée mouvement brownien et est un objet le plus fondamental en probabilité. Précisément, un mouvement brownien standard de dimension 1 est une collection $(B_t, t \geq 0)$ de variables aléatoires indexées par $\mathbb{R}_+$, vérifiant les trois propriétés suivantes :

- $B_0 = 0$, p.s.

- La fonction $t \mapsto B_t$ est p.s. continue de $\mathbb{R}_+$ dans $\mathbb{R}$.

- Pour tous $0 = t_0 \leq t_1 \leq .... \leq t_k$, les variables aléatoires $(B_{t_i} - B_{t_{i-1}}, 1 \leq i \leq k)$ sont indépendantes, respectivement de loi $N(0, t_i - t_{i-1})$.

Avec cette definition, il est clair que la loi de $(B_{t_1}, B_{t_2}, ..., B_{t_k})$ est la même que celle de $(N_1, N_1 + N_2, ..., N_1 + ... + N_k)$ avec les mêmes notations que ci-dessous. On reformule donc (3.16) en écrivant que pour tout choix de reels positifs $t_1 \leq t_2 \leq ... \leq t_k$, on a

$$(S^{[n]}_{t_i}, 1 \leq i \leq k) \rightarrow^{loi}_{n\rightarrow\infty} (B_{t_i}, 1 \leq i \leq k). \tag{4.17}$$

Si l'on prend des instants t_i de plus en plus nombreux et resserrés, cette convergence semble bien expliquer que $S^{[n]}$ approche la trajectoire du mouvement brownien B lorsque $n \rightarrow \infty$. Pourtant, les convergences (4.17) ne sont pas entièrement satisfaisantes de ce point de vue. Par exemple, il est naturel de conjecturer que l'on a

$$\sup_{0\le t\le 1} S_t^{[n]} \to_{n\to\infty}^{loi} \sup_{0\le t\le 1} B_t.$$

Mais cette convergence n'est pas une conséquence des convergences en loi (4.17). Disons que, même si l'on choisit des instants t_i très resserrés, la convergence (4.17) n'exclut pas a priori que le processus $S^{[n]}$ puisse prendre des valeurs anormalement élevées entre deux de ces instants. Il nous faut donc développer un convergence en loi pour des variables aléatoires qui sont des fonctions appartenant à un espace muni de la topologie uniforme.

4.6 Théorème de Prokhorov

Soit (E,d) un espace métrique, qu'on supposera séparable et complet. On note $C_b(E,R)$ l'ensemble des fonctions continues bornées de E dans R.

Définition 4.2 Soit $(X_n)_{n\ge 1}$, X une suite de variables aléatoires à valeurs dans E. On dit que (X_n) converge en loi dans E vers X, et on note $X_n \Rightarrow X$ si pour toute fonction $f \in C_b(E,R)$ on a $E[f(X_n)] \to E[f(X)]$.
Le théorème de Prokhorov étant l'outil le plus fondamental pour les théorèmes limite fonctionnels, il donne un critère de "relative compacité" pour la convergence en loi. Pour bien comprendre ce théorème, nous commençons par énoncer le résultat fondamental suivant.

Théorème [Riesz, Banach-Alaoglu] Si E est un compact, alors de toute suite de variables aléatoires à valeurs dans E, on peut extraire une sous-suite qui converge en loi.

Définition 4.3 La suite $(X_n, n\ge 1)$ de variables aléatoires à valeurs dans E, est dite tendue si pour tout $\varepsilon \ge 0$, il existe un compact $K_\varepsilon \subseteq E$ tel que

$$\inf P(X_n \in K_\varepsilon) \ge 1-\varepsilon.$$

Autrement dit, une suite de variables aléatoires est tendue si l'on peut trouver un seul et même compact qui concentre la plupart de la masse des lois de tous les éléments de cette suite.

4.6.1 Théorème [Prokhorov]

Soit $(X_n, n \geq 1)$ une suite tendue de variables aléatoires à valeurs dans E. Alors il existe une sous-suite de $(X_n, n \geq 1)$ qui converge en loi.
Ce théorème est en général utilisé conjointement avec la proposition suivante, qui est une conséquence du fait que la convergence en loi correspond à une topologie métrisable (dite topologie tendue) sur l'ensemble des mesures de probabilités sur E.

Proposition 4.1 Soit $(X_n, n \geq 1)$ une suite de variables aléatoires à valeurs dans E. On a $X_n \Rightarrow X$ si et seulement si de toute suite extraite, on peut reextraire une sous-sous-suite qui converge en loi vers X.
Nous appliquerons cette proposition dans le cas particulier où $E = C([0,1], R)$, muni de la norme uniforme, et que l'on notera désormais C. Nous donnons un critère utile de tension pour les variables aléatoires à valeurs dans C. Dans l'énoncé suivant, pour $x \in C$, la quantité

$$N_\alpha(x) = \sup_{s,t \in [0.1], s \neq t} \frac{| x(s) - x(t) |}{| t - s |^\alpha}$$

est appelée la norme de Hölder d'exposant α de la fonction x.

Théorème 4.2 [critère de Kolmogorov]

- Soit π_k la projection de R^∞ sur R^k définie par $\pi_k(x) = (x_1, .., x_k)$. Pour $k > 1$, soit ψ_k la projection de R^k sur R^{k-1} définie par $\psi_k(x_1, .., x_k) = (x_1, .., x_{k-1})$. Puisque π_k et ψ_k sont continues; elles sont mesurables.

Si P est une mesure de probabilité sur $(R^\infty, \mathbb{R}^\infty)$ définie par une mesure de probabilité μ_k sur $(R^k, \mathfrak{Re}^k)$, on a

$$\mu_k = P\pi_k^{-1}.$$

Puisque

$$\pi_{k-1} = \psi_k \pi_k, \qquad k > 1,$$

les mesures μ_k vérifient alors

$$\mu_{k-1} = \mu_k \psi_k^{-1}, \qquad k > 1.$$

On considère une suite $(X_n)_{n \geq 1}$ de variables aléatoires à valeurs dans C. On suppose que :

- La suite $(X_n(0))_{n\geq 1}$ est tendue,

- Il existe $\beta, p, C \geq 0$ tels que

$$\sup_{n\geq 1} E[\mid X_n(s) - X_n(t) \mid^p] \leq C \mid t - s \mid^{1+\beta}$$

pour tout $s, t \in [0,1]$.
Alors $(X_n)_{n\geq 1}$ est tendue. Plus précisement, pour tout $\alpha \in]0, \frac{\beta}{p}[$ fixé et pour tout $\varepsilon > 0$, il existe $M > 0$ tel que

$$\sup_{n\geq 1} P(N_\alpha(X_n) > M) \leq \varepsilon. \tag{4.18}$$

Pour montrer ce résultat, nous avons recours à un lemme intermédiaire. Notons $D_k = (\frac{i}{2^k} : 0 \leq i \leq 2^k)$ les rationnels dynamiques de niveau k dans $[0,1]$.

Lemme 4.1 : Soit $f : [0,1] \to R$ une fonction. Supposons qu'il existe K, $\alpha \geq 0$ pour lesquels pour tout $k \geq 0$ et pour tout $s, t \in D_k$, consécutifs, on ait
$\mid f(s) - f(t) \mid \leq K \mid t - s \mid^\alpha$. Alors pour tout $s, t \in \bigcup_{n\geq 0} D_n$, on a

$$\mid f(s) - f(t) \mid \leq \frac{2K}{1 - 2^\alpha} \mid t - s \mid^\alpha .$$

En particulier,
si $f \in \mathrm{C}$, on a

$$N_\alpha(f) \leq \frac{2K}{1 - 2^\alpha}.$$

Preuve Soient s, t deux nombres dans $\bigcup_{k\geq 0} D_k$. Alors il existe un unique entier $r \geq 0$ tel que $2^{-r-1} \leq \mid t - s \mid < 2^r$. Sans perdre de généralité, on peut supposer que $s \leq t$. On distingue deux cas.
Dans un premier cas, il existe $i \in \{0, 1, ..., 2^r - 1\}$ tel que $i2^{-r} \leq s$, $t \leq (i+1)2^{-r}$. Dans ce cas, les écritures dynamiques de s, t sont de la forme

$$s = \frac{i}{2^r} + \sum_{k\geq r+1} \frac{i_k(s)}{2^k}, \qquad t = \frac{i}{2^r} + \sum_{k\geq r+1} \frac{i_k(t)}{2^k},$$

où les suites $(i_k(s))_{k\geq r+1}$ et $(i_k(t))_{k\geq r+1}$ sont à valeurs dans $\{0,1\}$, et nulles à partir d'un certaine rang. On a alors, en posant

$$s_l = \frac{i}{2^r} + \sum_{k=r+1}^{r+l} \frac{i_k(s)}{2^k}$$

que s_l et s_{l+1} sont ou bien des nombres dynamiques consécutifs dans D_{r+l+1}, ou bien égaux (ce qu'ils sont pour tout l assez grand, puisqu'alors $s_l = s$) on a donc, par hypothèse

$$|f(s) - f(i2^{-r})| \leq \sum_{i \geq 0} | f(s_l) - f(s_{l+1}) | \leq \sum_{i \geq 0} K 2^{-(l+r+1)\alpha}$$

$$= \frac{K 2^{-(r+1)}}{1 - 2^{-\alpha}} \leq \frac{K}{1 - 2^{-\alpha}} | t - s |^{\alpha} .$$

La même majoration est vraie pour t à la place de s, et on en déduit que

$$| f(s) - f(t) | \leq \frac{2K}{1 - 2^{-\alpha}} | t - s |^{\alpha} .$$

Dans un second cas, il existe $i \in \{1, 2, ..., 2^r - 1\}$ tel que $(i-1)2^r \leq s \leq i2^r \leq t \leq (i-1)2^r$. On écrit alors plutôt

$$s = \frac{i}{2^r} + \sum_{k \geq r+1} \frac{i'_k(s)}{2^k}, \qquad t = \frac{i}{2^r} + \sum_{k \geq r+1} \frac{i_k(t)}{2^k}$$

où de la même manière que précédemment, $(i_k(s))_{k \geq r+1}$ et $(i_k(t))_{k \geq r+1}$ sont des suites de $\{0, 1\}$ nulles à partir d'un certain rang. Le même raisonnement que ci-dessus permet alors de conclure à la même majoration pour $|f(s) - f(t)|$.
Comme l'ensemble $\bigcup_n D_n$ est dense dans $[0, 1]$, si $f \in C$, nous déduisons que la majoration reste valable pour $s, t \in [0, 1]$. Ceci termine la preuve du lemme. Nous prouvons à présent la proposition (2.5). Posons

$$Z_{n,k} = \max_{0 \leq i \leq 2^k - 1} | X_n((i-2)2^{-k}) - X_n(i2^{-k}) | .$$

On a alors, par l'inégalité de Markov,

$$P(Z_{n,k} \geq K 2^{-k\alpha}) \leq P(\max_{0 \leq i \leq 2^k - 1} | X_n((i-2)2^{-k}) - X_n(i2^{-k}) | \geq K 2^{-k\alpha})$$

$$\leq \max_{0 \leq i \leq 2^k - 1} \frac{E[| X_n((i-2)2^{-k}) - X_n(i2^{-k}) |^p]}{K^p 2^{-pk\alpha}}.$$

Par hypothèse, cette dernière quantité est majorée par $CK^{-p} 2^{-(\beta - p\alpha)k}$. Si l'on choisit $\alpha \in]0, \frac{\beta}{p}[$, on déduit que ces majorants sont sommables sur k. Le lemme (2.7) implique que

$$P(N_\alpha(X_n) > \frac{2K}{1 - 2^{-\alpha}}) \leq P(\exists k \geq 0 : Z_{n,k} > K 2^{-k\alpha})$$

$$\leq \frac{C}{K^p} \sum_{k \geq 0} 2^{-(\beta - p\alpha)k} \leq \frac{C'}{K^p},$$

pour une constante $C' \in]0, \infty[$, qui dépend que de α, β, p, C mais non de n. Le dernier majorant converge vers 0 lorsque $K \to \infty$, d'où l'on déduit que tout $\varepsilon > 0$, il existe $M > 0$ tel que (3) est vraie. Ceci, conjointement à la première hypothèse de proposition (2.5), implique la tension de $(X_n)_{n \geq 1}$. En effet, pour $M > 0$, l'ensemble des $f \in C$ tels que $| f(0) | \leq M$ et $N_\alpha(f) \leq M$ est une famille équicontinue, qui est donc relativement compacte.

4.7 Théorème de Donsker dans D

Soit $X_1, X_2, ...$ une suite i.i.d de variables aléatoires. On pose $S_0 = 0$, et pour tout $n \geq 1, S_n = X_1 + X_2 + ... + X_n$. On étend alors S en une fonction continue de $\mathbb{R}_+$ dans $\mathbb{R}$ par la formule

$$S_t^{(n)} = \frac{S_{[nt]}}{\sqrt{n}} \qquad 0 \leq t \leq 1.$$

On voit ainsi $S^{(n)}$ comme une variable aléatoire à valeurs dans $D[0,1]$.

Théorème 4.3 [Donsker] *Soit $X_1, X_2, ...$ une suite i.i.d de variables aléatoires. La suite $(S^{(n)})_{n \geq 1}$ converge en loi dans D vers le mouvement brownien standard unidimensionnel $(B_t)_{0 \leq t \leq 1}$.*

La loi de ce dernier est caractérisé par le fait que pour tout $0 \leq t_0 \leq t_1 \leq ... \leq t_k$, les variables $B_{t_i} - B_{t_{i-1}}, \quad 1 \leq i \leq k$ sont indépendantes, respectivement de loi gaussienne. En fait, le théorème de Donsker est vrai dans le cadre beaucoup plus général où les variables $(X_i, \quad i \geq 1)$ sont indépendantes de même loi, centrées, et de variance finie. Il se généralise également à des marches aléatoires à valeurs dans $\mathbb{R}^d$. La preuve dans le cas particulier de la marche au plus proche voisin, est un peu plus simple, c'est cette version qu'on déterminera ici en supposant que les variables aléatoires $X_1, X_2, ...$ sont i.i.d., centrées, et vérifient $E[X_1^4] < \infty$. La dernière hypothèse n'est pas nécessaire, mais s'en débarrasser nécessite des critères un peu plus fins que celui de Kolmogorov.

Preuve

Etape 1 : Convergence des lois marginales de dimension finie. Nous fixons des instants $0 \leq t_0 \leq t_1 \leq ... \leq t_k \leq 1$. Nous avons déjà vu que

$$(\frac{S_{[nt_i]} - S_{[nt_{i-1}]}}{\sqrt{n}}, \quad 1 \leq i \leq k) \to^{loi}_{n\to\infty} (N_1, N_2, ..., N_k),$$

où les variables aléatoires $N_1, N_2, ..., N_k$ sont indépendantes, respectivement de loi $N(0,(t_i - t_{i-1})I_d)$. Ceci provient d'une application du théorème central limite vectoriel. Par image continue, on en déduit que

$$(\frac{S_{[nt_i]}}{\sqrt{n}} \quad 1 \leq i \leq k) \to^{loi}_{n\to\infty} (N_1, N_1 + N_2, ..., N_1 + ... + N_k).$$

Enfin, remarquons que

$$| S_t^{(n)} - \frac{S_{[nt]}}{\sqrt{n}} | \leq \frac{X_{[nt]}}{\sqrt{n}},$$

qui converge vers 0 en probabilité pour tout t fixé. Pour cette raison, on a la convergence en probabilité : $(S_{t_i}^{(n)} - \frac{S_{[nt_i]}}{\sqrt{n}}, 1 \leq i \leq k) \to^{P}_{n\to\infty} 0$. Par le lemme de Slutzky, on en déduit que :

$$(S_{t_i}^{(n)}, 1 \leq i \leq k) \to^{loi}_{n\to\infty} (\mathbf{B}_{t_i}, 1 \leq i \leq k).$$

Etape 2 : Tension. Nous montrons enfin que la suite $(S^{(n)})_{n\geq 1}$ est tendue dans D. Pour cela, on applique le critère de Kolmogorov. Soit p un entier. Si s et t sont dans un même intervalle de la forme $[\frac{k}{n}, \frac{k+1}{n}]$, on a alors

$$E[| S_t^{(n)} - S_s^{(n)} |^{2p}] = (\frac{| nt - ns |}{\sqrt{n}})^{2p} \leq | t - p |^p .$$

Sinon, en supposant en plus, $s \leq t$, il existe deux entiers k et k' tels que $k - 1 \leq ns \leq k \leq k' \leq nt \leq k' + 1$, et on a alors

$$E[| S_t^{(n)} - S_s^{(n)} |^{2p}] \leq 2^{2p}(| s - \frac{k}{n} |^p + E[| \frac{S_{k'} - S_k}{\sqrt{n}} |^{2p}] + | s - \frac{k'}{n} |^p)$$

$$\leq 2^{2p+1} | t - s |^p + 2^{2p} E[| \frac{S_{k'-k}}{\sqrt{n}} |^{2p}].$$

Or un calcul rapide montre que $E[S_k^{2p}] \leq C_p k^p$ pour une constante $C_p > 0$ et pour tout $k \geq 0$. En effet, on a $E[S_k^{2p}] = \sum E[X_{i_1} ....X_{i_{2p}}]$ où la somme porte sur tous les indices $i_1 ....i_{2p}$ entre 1 et k, et les espérances concernées sont non-nulles si et seulement si on peut apparier ces indices deux par deux. D'où l'on déduit finalement qu'il existe une constante C'_p telle que pour tous $s, t \in [0,1]$, on ait :

$$E[\mid S_t^{(n)} - S_s^{(n)} \mid^{2p}] \leq C'_p \mid t - s \mid^p .$$

En constatant en plus que $S_0^n = 0$, pour tout n, on peut donc appliquer le critère de Kolmogorov, c'est à dire que la suite est tendue.

Conclusion Nous avons montré que la suite $(S^{(n)})_{n \geq 1}$ est tendue, et la convergence des lois marginales de dimension finie vers le mouvement brownien. Toute limite en loi de $S^{(n)}$ le long d'une sous-suite a donc même loi que le mouvement brownien. Donc $S^{(n)} \Rightarrow \mathbf{B}$ dans D.

4.7.1 Lissage polygonal

Après la convergence de $S^{(n)}$ vers un mouvement brownien $\mathbf{B}$ dans l'ensemble de fonction "cadlag" $D[0,1]$, nous allons passer par le lissage de $S^{(n)}$ pour qu'elle soit dans l'ensemble des fonctions continues $C[0,1]$.
Soit $X_1, X_2, ...$ une suite i.i.d de variables aléatoires réelles centrées et réduites. On pose $S_n = \sum_{i=1}^n X_i$, l'espérance de cette somme est $E(S_n) = 0$ et sa variance est $V(S_n) = n \Rightarrow \sigma(S_n) = \sqrt{n}$ et posons

$$\xi_n(t) = \frac{S_{[nt]}}{\sqrt{n}}$$

$\xi_n(t) \to_{n \to \infty}^{loi} \mathbf{B}(t)$ dans $D[0,1]$.
Par le lissage de $\xi_n(t)$ nous retrouvons une autre fonction

$$\zeta_n(t) = \frac{S_{[nt]}}{\sqrt{n}} + (nt - [nt]) X_{[nt]+1}.$$

En appliquant le théorème de Donsker sur $\zeta_n(t)$, on déduit que

$$\zeta_n(t) \to_{n \to \infty}^{loi} \mathbf{B}(t)$$

4.7.2 Théorème de Donsker dans C

Prenons une suite de variables aléatoires $\xi_1, \xi_1, ...$ définie sur $(\Omega, A, \mathbf{P})$. Soit $S_n = X_1 + ... + X_n$ une somme partielle qui définit un élément aléatoire X_n de $\mathbf{C}$ donné par

$$X_n(t, \omega) = \frac{1}{\sigma \sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma \sqrt{n}} \xi_{[nt]+1}(\omega).$$

Le théorème suivant est sur la convergence en loi de X_n, dans C

Théorème 4.4 Supposons que (ξ_n) est une suite de variables aléatoires indépendantes et identiquement distribuées de moyenne nulle et de variance σ^2. La fonction aléatoire X_n définie par

$$X_n(t,\omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega);$$

converge en loi vers B i.e $X_n \longrightarrow^{loi} B$ dans C

Démonstration Nous allons d'abord monter que lois de dimension finie des X_n convergent vers celles de B. Considérons d'abord un seul instant s; nous devons prouver que $X_n(s) \longrightarrow^{loi} B(s)$.
Puisque

$$|X_n(s) - \frac{1}{\sigma\sqrt{n}} S_{[ns]}| \leq \frac{1}{\sigma\sqrt{n}} \xi_{[ns]+1} \longrightarrow^{P} 0;$$

par l'inégalité de Chebyskev, $\frac{1}{\sigma\sqrt{n}} S_{[ns]} \longrightarrow^{loi} B(s)$. Mais c'est une conséquence directe du théorème central limite de Lévy, et $\frac{[ns]}{n} \to s$.
Considérons maintenant deux instants s et t, avec $s < t$, nous devons prouver que

$$(X_n(s), X_n(t)) \longrightarrow^{loi} (B_s, B_t)$$

prouvons que

$$(X_n(s), X_n(t) - X_n(s)) \longrightarrow^{loi} (B_s, B_t - B_s),$$

car

$$|X_n(s) - \frac{1}{\sigma\sqrt{n}} S_{[ns]}| \leq \frac{1}{\sigma\sqrt{n}} \xi_{[ns]+1} \longrightarrow^{P} 0;$$

dans cette relation en remplaçant t par s on aura

$$(\frac{1}{\sigma\sqrt{n}} S_{[ns]}, \frac{1}{\sigma\sqrt{n}} (S_{[nt]} - S_{[ns]})) \longrightarrow^{loi} (B_s, B_t - B_s).$$

Puisque les composants sur la gauche sont indépendants des des composantes de la droite.
Un ensemble de trois moments ou plus, peut être traité de la même manière, et donc les lois de dimension finie convergent correctement. Nous pouvons étendre via le lemme suivant, qui est plus général que celui qui est actuellement réqui.

Lemme 4.3 *Soit $\xi_1, ..., \xi_m$ une suite de variables aléatoires indépendantes de moyenne nulle et de variance σ_i^2 ; posons $S_i = \xi_1 + ... + \xi_i$ et $S_i^2 = \sigma_1^2 + ... + \sigma_i^2$, alors*

$$P\{\max_{i \le m} \mid S_i \mid \ge \lambda S_m\} \le 2P\{|S_m| \ge (\lambda - \sqrt{2}) S_m\}. \tag{4.19}$$

Preuve Prouvons cette inéquation ; notons qu'elle est trivial si $\lambda \le \sqrt{2}$; considérons les ensembles

$$E_i = \{\max_{j<i} \mid S_j \mid < \lambda S_m \le \mid S_j \mid\}.$$

Clairement,

$$\begin{aligned} P\{\max_{i \le m} \mid S_i \mid \ge \lambda S_m\} &\le P\{|S_m| \ge (\lambda - \sqrt{2}) S_m\} \\ &+ \sum_{i=1}^{m} P(E_i \bigcap \{|S_m| < (\lambda - \sqrt{2}) S_m)\}). \end{aligned} \tag{4.20}$$

Puisque $\mid S_i \mid \ge \lambda S_m$ et $\mid S_m \mid < (\lambda - \sqrt{2})$ alors $\mid S_m - S_i \mid \ge \sqrt{2} S_m$, d'après l'inégalité de Chebychev et en réponse à l'indépendance des ξ_i que la somme de (4.20) est au plus

$$\begin{aligned} \sum_{i=1}^{m-1} P(E_i) P\{\mid S_m - S_i \mid \ge \sqrt{2} S_m\} &\le \sum_{i=1}^{m-1} P(E_i) \frac{1}{2S_m^2} \sum_{k=i+1}^{m} \sigma_k^2 \\ &\le \frac{1}{2} \sum_{i=1}^{m-1} P(E_i) \\ &\le \frac{1}{2} P\{\max_{i \le m} \mid S_i \mid \ge \lambda S_m\}. \end{aligned} \tag{4.21}$$

Et maintenant (4.21) et (4.20) combinées donne (4.19).
Appliquons le lemme pour les variables et le théorème nous avons, pour $\lambda > 2\sqrt{2}$,

$$P\{\max_{i \le n} \mid S_i \mid \ge \lambda \sigma \sqrt{n}\} \le 2P\{|S_m| \ge \frac{1}{2} \lambda \sigma \sqrt{n}\}. \tag{4.22}$$

Par le théorème central limite,

$$P\{|S_m| \ge \frac{1}{2} \lambda \sigma \sqrt{n}\} \to P\{|N| \ge \frac{1}{2} \lambda\} < \frac{8}{\lambda^2} E\{|N|^2\}.$$

En conclusion, si ε est positif, nous avons

$$\limsup_{n \to \infty} P\{\max_{i \le} \mid S_i \mid \ge \lambda \sigma \sqrt{n}\} < \frac{\varepsilon}{\lambda^2}$$

pour λ assez grand. L'equitension suit maintenant par le théorème.

Une application comme indiqué plus haut, le théorème de Donsker a une interprétation quantitative : $X_n \to_{loi} B$ disons que, si τ est petit une particule sujette à des déplacements indépendants $\xi_1, \xi_2, \ldots$ à des instants successifs $\tau, 2\tau, \ldots$ apparaitra, vue de loin, effectuer approximativement un mouvement brownien.
Plus important que cette interprétation quantitative est l'utilisation du théorème de Donsker pour prouver les théorèmes limites pour diverses fonction des sommes partielles S_n. L'introduction indique comment utiliser la relation $X_n \to_{loi} B$ pour dériver la limite de $\max_{i\leq n} S_i$; portons maintenant cela en détail.
Puisque $h(x) = \sup_t x(t)$ est une fonction continue sur C, $X_n \to_{loi} B$ implique que

$$\sup_t X_n(t) \to_{loi} \sup_t B_t.$$

La relation évidente

$$\sup_{0\leq t\leq 1} X_n(t) = \max_{i\leq n} \frac{1}{\sigma\sqrt{n}} S_i$$

maintenant implique

$$\max_{i\leq n} \frac{1}{\sigma\sqrt{n}} S_i \to_{loi} \sup_t B_t. \tag{4.23}$$

Ainsi nous aurions la distribution limite de $\max_{i\leq n} S_i$ (correctement normalisée) si nous connaissions la distribution de $\sup_t B_t$. La technique que nous utiliserons pour trouver cette dernière distribution est de calculer la distribution limite de $\max_{i\leq n} S_i$ dans un cas particulier.
Supposons que S_0, S_1,... sont des variables aléatoires d'une marche aléatoire symétrique partant de l'origine. Supposons, que les ξ_n sont indépendants et satisfont

$$P\{\xi_n = 1\} = P\{\xi_n = -1\} = \frac{1}{2}. \tag{4.24}$$

Montrons que, si a est un entier non négatif, alors

$$P\{\max_{0\leq i\leq n} S_i \geq a\} = 2P\{S_n > a\} + P\{S_n = a\}. \tag{4.25}$$

Le cas $a = 0$ est trivial. Supposons que $a > 0$ et prenons $M_i = \max_{0\leq j\leq i} S_j$. Puisque

$$P\{M_n \geq a\} - P\{S_n = a\} = P\{M_n \geq a, S_n < a\} + P\{M_n \geq a, S_n > a\}$$

et

$$P\{M_n \geq a, S_n > a\} = P\{S_n > a\}.$$

(4.25) sera vraie si nous prouvons

$$P\{M_n \geq a, S_n < a\} = P\{M_n \geq a, S_n > a\}. \tag{4.26}$$

De (4.24), on déduit que les 2^n chemins possibles $(S_1, S_2, ..., S_n)$ ont la même probabilité 2^{-n}. En conclusion (4.26) suivra si nous montrons que le nombre de chemins contribuant à l'évènement de gauche est le même que le nombre de chemins contribuant à l'évènement de droite pair, et pour le montrer il suffit de faire correspondre les chemins d'une manière un à un. Etant donné un chemin $(S_1, S_2, ..., S_n)$ contribuant à la main gauche même en (4.26), faites le correspondre au chemin obtenu en réfléchissant par a toute la somme partielle après la première qui atteint la hauteur a. Puisque la correspondance est un à un, (4.26) suit.

Soit α un nombre non négatif arbitraire, et soit $a_n = \alpha n^{\frac{1}{2}}$. Par (4.26), nous avons

$$P\{\max_{i \leq n} \frac{1}{\sqrt{n}} S_i \geq \alpha\} = 2P\{S_n > a_n\} + P\{S_n = a_n\}. \tag{4.27}$$

Par le théorème central limite,

$$P\{S_n \geq a_n\} \to P\{N \geq \alpha\}$$

($\sigma^2 = 1$ dans ce cas, en vue de (4.24)). Puisque le plus grand terme de la distribution Binomiale symétrique tend vers 0, le terme $P\{S_n = a_n\}$ dans (4.27) est négligeable. Donc

$$P\{\max_{i \leq n} \frac{1}{\sqrt{n}} S_i \geq \alpha\} \to 2P\{N \geq \alpha\}, \qquad \alpha \geq 0. \tag{4.28}$$

La combinaisons de (4.28) avec (4.23)(en supposant (4.24)), nous pouvons conclure

$$P\{\sup_t B_t \leq \alpha\} = \frac{2}{\sqrt{2\pi}} \int_0^\alpha e^{-\frac{1}{2}u^2} du, \qquad \alpha \geq 0. \tag{4.29}$$

Bien sur, le côté gauche de (4.29) s'annule si $\alpha < 0$. Nous avons dérive un fait sur le mouvement brownien en combinant le théorème de Donsker avec un calcul impliquant une marche aléatoire, un calcul qui était simple en partie parce qu'il se réduisant à une énumération et en partie parce qu'une marche aléatoire ne peut pas passer au dessus d'un entier positif a sans le traverser.

Laissons maintenant de côté l'hypothèse (4.24). Si les ξ_n sont indépendants et identiquement distribués et vérifies (4.14), alors (4.23) est valable, et à partir de (4.26) nous pouvons maintenant conclure

$$P\{\frac{1}{\sigma\sqrt{n}} \max_{i \leq n} S_i \leq \alpha\} \to \frac{2}{\sqrt{2\pi}} \int_0^\alpha e^{-\frac{1}{2}u^2} du, \qquad \alpha \geq 0. \tag{4.30}$$

Ainsi nous avons dérivé la distribution limite de $\max_{i\leq n} S_i$ sous les hypothèses du théorème de Lévy.
Cet argument suit un schéma général. Si h est continue sur C ; ou continue sauf aux points formant un ensemble de mesure de Wiener 0 ; alors $X_n \rightarrow^{loi} B$ implique

$$h(X_n) \rightarrow^{loi} h(B). \tag{4.31}$$

(Dans le cas qui vient d'être analysé ; $h(x) = \sup_t x(t)$.) Nous pouvons trouver la distribution limite de $h(X_n)$ si nous connaissons la distribution de $h(B)$, et nous pouvons souvent trouver la distribution de $h(X_n)$ dans un cas particulier, puis en utilisant (4.28) dans l'autre sens.
En conclusion, si les ξ_n sont indépendants et identiquement distribués avec $E\{\xi_n\} = 0$ et $E\{\xi_n^2\} = \sigma^2$, alors la distribution limite de $h(X_n)$ ne dépend d'aucune autre propriété de ξ_n. Pour cette raison, le théorème de Donskers est souvent appelé le Principe d'invariance. Ici nous l'appellerons plutôt le théorème central limite fonctionnel.

Condition nécessaire pour l'équitension Supposons que les ξ_n sont indépendants et identiquement distribués et admettons directement que $X_n \rightarrow^{loi} B$. Nous montrons que, dans cette circonstance, pour chaque ε positif il existe un λ positif tel que

$$P\{\max_{i\leq n} \mid S_i \mid\geq \lambda\sigma\sqrt{n}\} \leq \frac{\varepsilon}{\lambda^2}. \tag{4.32}$$

Dans le cas où $\{\xi_n\}$ est stationnaire, c'est juste l'hypothèse du théorème 4.4, que nous avons vérifié en prouvant le théorème de Donsker.
Soit $Y = \sup_t |B_t|$. Parce que de (4.29) et la symétrique de B sous réflexion par 0, Y a un second moment fini. De (4.31) avec $h(x) = \sup_t |x(t)|$, ça implique que, pour tout λ positif,

$$P\{\max_{i\leq n} \mid S_i \mid\geq \lambda\sigma\sqrt{n}\} \rightarrow P\{Y \geq \lambda\} \leq \frac{1}{\lambda^2}\int_{(Y\geq\lambda)} Y^2 dP. \tag{4.33}$$

Soit $\varepsilon > 0$, choisissons λ assez grand que l'intégrale ici est inférieur à ε ; de (4.33) que (4.32) vaut pour tout n suffisamment grand.
Donc nous voyons, après coup, que (4.32) est la condition à vérifier pour prouver l'équitension du théorème 4.4 de Donsker.
L'argument précédent le montre bien : supposons et que la distribution de dimension finie des X_n converge vers celle de la fonction aléatoire X, où $\sup_t \mid X(t) \mid$ a un second moment fini. Dans ces conditions, si $\{X_n\}$ est tendue, alors l'hypothèse du théorème 4.4 est satisfaite. Dans cette mesure, cette hypothèse est nécessaire.

5

Application statistique

5.1 Introduction

Dans ce chapitre nous allons considérer des processus de sommes partielles, nous étudions la convergence de ces processus vers le mouvement brownien. Nous déterminons ensuite une statistique de test qui pourra nous permettre de calculer la région de rejet et la puissance du test.

5.2 Processus de sommes partielles

Soit $X_1, X_2, \ldots$ une suite de variables aléatoires indépendantes et identiquement distribués, de moyenne m_i et de variance σ^2. On pose $S_0 = 0$ et pour tout $n \geq 1$, $S_n = X_1 + X_2 + \ldots + X_n$. On définit

$$\xi_n(t) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}.$$

Ce processus est un élément de $\mathbf{D}[0,1]$, c'est à dire un processus discontinu sur $[0,1]$ et converge vers le mouvement brownien,

$$\xi_n(t) \rightarrow^{loi}_{\mathbf{D}[0,1]} B(t).$$

Pour la preuve voir chapitre 4.

5.3 Processus de sommes partielles lissé

Soit $\xi_1, \xi_2, \ldots$ une suite de variables aléatoires définie sur (Ω, A, P), soit $S_n = \xi_1 + \xi_2 + \ldots + \xi_n$ une somme partielle qui définit un élément aléatoire

$\tilde{\xi}_n$ de $\mathbf{C}[0,1]$ donné par

$$\tilde{\xi}_n(t,\omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega).$$

D'après le théorème central limite fonctionnel dans le chapitre précédent nous pouvons déduire que $\tilde{\xi}_n$ converge en loi vers le mouvement brownien B, c'est à dire

$$\tilde{\xi}_n(t) \to^{loi}_{\mathbf{C}[0,1]} B(t).$$

Supposons maintenant l'ensemble des fonctions continues sur $[0,1]$ ($\mathbf{C}[0,1]$) muni de la norme uniforme. Puisque cette norme est continue alors

$$\| \tilde{\xi}_n(t) \|_\infty \to^{loi}_{\mathbf{C}[0,1]} \| B(t) \|_\infty .$$

5.4 Test d'hypothèses

Soit $(X_1, X_2, ..., X_n)$ un échantillon de variables aléatoires de moyennes $m_1, m_2, ..., m_n$ respectivement. On veut tester l'hypothèse nulle.
$(H_0) : m_1 = m_2 = ... = m_n$,
contre l'hypothèse alternative
$(H_A) : \exists l < k^* < m^* < n$ tels que
$m_1 = m_{k^*} = m_{m^*+1} = ... = m_n$, $m_{k^*+1} = ... = m_{m^*}$ et $m_{k^*} \neq m_{k^*+1}$.
On note $l^* = m^* - k^*$ la longueur de l'épidémie et on suppose que l^* et $n - l^*$ tendent vers l'infini quand n tend vers l'infini.
On définit le processus de sommes partielles basé sur les X_i par

$$\tilde{\xi}_n(t) = S_{[nt]} + (nt - [nt])\xi_{[nt]+1}, \quad t \in [0,1]. \tag{5.1}$$

$S_0 = 0$ et $S(t) = \sum_{k \le t} X_k$.
Pour une suite de variables aléatoires indépendantes et de même loi, Donsker et Prohorov ont prouvé que si la variance σ^2 est finie, alors $\sigma^{-1} n^{-\frac{1}{2}} \xi_n$ converge en loi dans $C[0,1]$ vers le mouvement brownien B.
Par le théorème de conservation de la convergence en loi par image continue, $g(\sigma^{-1} n^{-\frac{1}{2}} \xi_n)$ converge en loi vers $g(B)$, où g est une fonction continue.
La statistique utilisée ici est

$$T = \| \tilde{\xi}_n(t) \|_\infty$$

et on a

$$T \to_{n\to\infty} \| B(t) \|_\infty,$$
$$\| B \|_\infty \sim N(0,1),$$

quand les variables sont centrées et réduites.

Preuve Soit $F = (C[0,1], \| \, . \, \|_\infty)$ un espace de fonctions continues muni de la norme infinie, $\tilde{\xi}_n$ un élément de F.

$$\| \tilde{\xi}_n \|_\infty = \sup_{t \in [0,1]} | \tilde{\xi}_n(t) | =$$

$$= \sup_{t \in [0,1]} | S_{[nt]} + (nt - [nt]) \xi_{[nt]+1} | .$$

Le théorème de Donsker a une interprétation quantitative : $\tilde{\xi}_n \to_{loi} B$ pour dériver la limite $\max_{i \leq n} \xi_i$; puisque

$$T = \sup_{t \in [0,1]} | \tilde{\xi}_n(t) |$$

est une fonction continue sur $C[0,1]$, $\tilde{\xi}_n \to_{loi} B$ implique que

$$\sup_t \tilde{\xi}_n(t) \to^{loi} \sup_t B(t).$$

La relation évidente

$$sup_t \tilde{\xi}_n(t) = \max_{i \leq n} (S_{[it]} + (it - [it]) \xi_{[it]+1}),$$

maintenant implique

$$\max_{i \leq n} (S_{[it]} + (it - [it]) \xi_{[it]+1}) \to \sup_t B(t).$$

Ainsi nous aurons la distribution limite de $\max_{i \leq n} (S_{[it]} + (it - [it]) \xi_{[it]+1})$ normalisée si nous connaissons la distribution de $\sup_t B(t)$. Or la distribution de $\sup_t B(t)$ est une loi normale $N(0,1)$, si les variables sont centrées et réduites.

5.4.1 Région de rejet

La région de rejet d'un test d'hypothèses est définie par

$$R = \{ \| B(t) \|_\infty > c \},$$

puisque $\| B(t) \|_\infty \sim N(0,1)$, alors la région de rejet s'écrit

$$R = \{ N(0,1) > c \}$$

Détermination de la constante c On utilise la région critique

$$R = \{N(0,1) > c\}.$$

On a

$$P(N(0,1) > c) = \alpha,$$

α est le seuil de confiance et a pour valeur 0.05. Maintenant passons à l'évènement contraire

$$P(N(0,1) > c) = \alpha$$

$$P(N(0,1) \leq c) = 1 - \alpha$$

$$P(N(0,1) \leq c) = 0,95.$$

Avec la table de la loi normale on a $c = 1,65$; donc la région critique s'écrit de la manière suivante

$$R = \{N(0,1) > 1,65\}$$

5.4.2 Puissance du test

La valeur de la puissance est déterminée par $1-\beta$; puisque la valeur de α est fixée la puissance du test est 0,95, mais cette valeur varie en fonction de la moyenne à tester.

5.4.3 Application numérique

Prenons un échantillon de variables aléatoires simulées à partir du langage R.
Le code suivant permet de simuler un échantillon de taille n qui suit une loi normale de moyenne m et de variance σ^2.

```
nor=function(n=20,moy=2,sgm=4) {
n=20
moy=2
sgm=4
for(i in i :n){u = runif(n,0,1)
x= (1/(sgm*sqrt(2*3.14)))*exp((-1/2)*((u-moy)/sgm )^(2))}
print(x)
}
```

Le tableau suivant représente l'échantillon de la simulation.

7.0324	4.3371	0.2038	0.2950
-5.5409	2.4104	3.7675	7.6418
2.7150	4.3899	5.9261	4.9673
8.4838	4.2371	4.2128	5.6453
5.8320	8.5837	-3.1541	5.0164

Nous pouvons déterminer la région de rejet et la puissance de test.

Région de rejet

$$R = \{N(m,\sigma) > c\},$$

donc

$$P\{N(m,\sigma) > c\} = \alpha$$

$$P\{N(0,1) > \frac{c-m}{\sigma/\sqrt{n}}\} = \alpha$$

$$P\{N(0,1) \leq \frac{c-m}{\sigma/\sqrt{n}}\} = 1-\alpha,$$

puisque α est fixé à 0.05, alors par la table de la loi normale on trouve

$$\frac{c-m}{\sigma/\sqrt{n}} = 1.65.$$

Or d'après les données $m = 2.5$, $n = 20$ et $\sigma = 4$; on peut déterminer la valeur de c qui est 3.94. Maintenant la région de rejet s'écrit

$$R = \{N(m,\sigma) > 3.94\}.$$

Puissance de test Comme on a dit avant $1-\beta$ est la puissance de test, mais ici $\beta = 0.33$, alors $1-\beta = 0.67$.

6 Conclusion générale

«Il est facile de manquer le but et difficile de l'atteindre»

Aristote

Au cours de ce mémoire, nous avons prouvé que le processus de sommes partielles converge vers le mouvement brownien, et faire une application statistique sur les tests d'hypothèses en faisant la comparaison des moyennes.
Nous avons pu montrer que le mouvement brownien a des propriétés propres à lui tout en donnant son importance dans différents domaines. Ici en utilisant les différents espaces et leurs propriétés nous avons vu que le processus de sommes partielles converge vers le mouvement brownien, mais en utilisant un autre espace d'étude, nous pouvons voir que ce même processus de sommes partielles va converger vers un pont brownien.
La présentation du mouvement brownien nous a permis de voir la spécificité du mouvement et comprendre que sa convergence ne se fait pas sur $\mathbb{R}$ mais plus tôt sur les espaces fonctionnels. Pour avoir une vue d'ensemble sur cette convergence nous avons fait appel au théorème de Donsker sur les espaces fonctionnels, qui est aussi appelé le théorème central limite fonctionnel.
Sur l'application statistique, nous avons utilisé la loi normale centrée et réduite qui est la limite de notre statistique de test pour déterminer la région de rejet et la puissance du test.

Bibliographie

[1] Billingsley P., *Convergence of probability measures*, J.Willey, New York(1968).

[2] Billingsley P., *Probability and measures*, 2ème édition (1986)

[3] Kerkyacharian G., Roynett B,.*Une démonstration simple des théorèmes de Kolmogorov, Donsker et Ito-Nisio*, C.R. Acad. Sci. Paris Ser. Math. I 312 (1991), 877-882.

[4] Lamperti J.,*On convergence of stochastic processus,*, Trans. Amer. Math. Soc. 104(1962), 430-435.

[5] Burago D., Burago Y., Ivanov A., A *course in metric geometry*, AMS, (2001).

[6] Merabet D., *Le processus empirique sous dépendance*, Thèse de doctorat (2012).

[7] Imeçaoudene K., *Inférence statistique hölderienne pour la détection de la rupture épidémique* Thèse de doctorat (2019).

[8] Graiche F., *Convergence hölderienne des processus empirique et quantiles et applications* Thèse de doctorat (2012).

[9] Ait Ouahra M, Kissami A, Sghir A *Un principe d'invariance de type Donsher dans une classe d'espaces de Besov-Orlicz*, Article, (2012).

Résumé

Il s'agit de décrire d'abord le mouvement brownien et les phénomènes modélisés par ce processus ainsi que ses différentes propriétés. Ensuite établir un théorème central limite fonctionnel pour comprendre l'importance de ce processus limite gaussien qu'est le mouvement brownien et enfin terminer par une application statistique sur un test d'hypothèse où tous les résultats établis auparavant seront appliqués, en particulier la loi gaussienne de ce processus limite mouvement brownien.

Abstract

It is a question of first describing the Brownian motion and the phenomena modeled by this process as well as its different properties. Then establish a functional central limit theorem to understand the importance of this gaussian limit process that is Brownian motion and finally finish with a application on hypothesis test where all the results established before will be applied, in particular the gaussian distribution of this Brownian motion limit process.