Feuille de TD n° 8:

THÉORÈME CENTRAL LIMITE (SUITE)

math.univ-lyon1.fr/homes-www/lerouvillois
lerouvillois@math.univ-lyon1.fr
poudevigne@math.univ-lyon1.fr

On rappelle le Théorème de Slutsky :

Théorème 1 (Slutsky) *Soit $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ deux suites de vecteurs aléatoires. Si $(X_n)_{n\in\mathbb{N}}$ converge en loi vers un vecteur aléatoire X et que $(Y_n)_{n\in\mathbb{N}}$ converge en loi vers une constante c. Alors $(X_n, Y_n)_{n\in\mathbb{N}}$ converge en loi vers (X, c).*
En particulier, $X_n + Y_n$ converge en loi vers $X + c$, et X_nY_n converge en loi vers cX.

Exercice 1 : Marche aléatoire et mouvement Brownien

Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de v.a i.i.d de carré intégrable, centrée et de variance 1. On définit la marche aléatoire suivante : $S_n := \sum_{i=1}^{n} X_i$, pour tout $n \geq 1$.

1. Vers quelle loi converge $\left(\frac{S_n}{\sqrt{n}}\right)_{n\in\mathbb{N}^*}$?
2. Pour tout $t \in \mathbb{R}_+$, on définit la variable aléatoire $B_t^n := \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}$.
   Vers quelle loi converge $(B_t^n)_{n\in\mathbb{N}^*}$? (on pourra utiliser le Théorème de Slutsky)

Remarque : Le théorème de Donsker montre que la suite de fonction aléatoire $(t \mapsto B_t^n)_{n\in\mathbb{N}^*}$ converge en loi vers le mouvement Brownien.

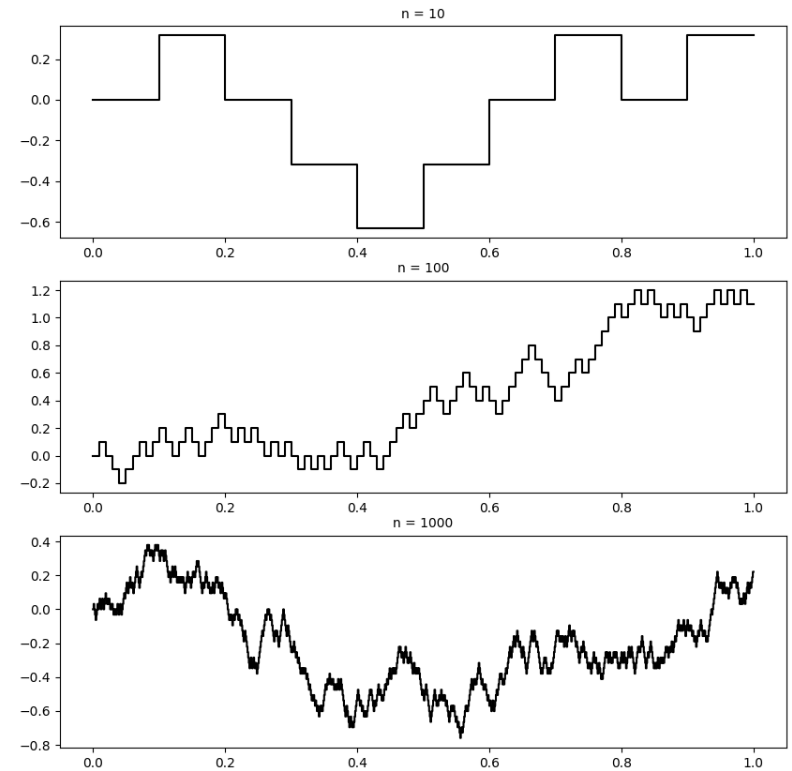


FIGURE 1 – Convergence de la marche aléatoire renormalisée vers le mouvement Brownien. On simule $(S_n)_{n\in\mathbb{N}}$ puis on trace les fonctions $t \mapsto \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}$ pour différentes valeurs de n.

Exercice 1 (Slutsky)

$(X_i)_{i\in\mathbb{N}}$ suite de v.a i.i.d et dans L^2

$E[X_i] = 0$

$V[X_i] = 1$

1) $\frac{S_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$ par théorème centrale limite car les X_i sont i.i.d et dans L^2

2) $t \in \mathbb{R}_+$, $B_t^n = \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}$

$B_t^n = \sqrt{\frac{\lfloor nt \rfloor}{n}} \frac{S_{\lfloor nt \rfloor}}{\sqrt{\lfloor nt \rfloor}}$

Or $nt - 1 < \lfloor nt \rfloor \leq nt$

$\Rightarrow \underset{\substack{\downarrow \\ t}}{t - \frac{1}{n}} < \frac{\lfloor nt \rfloor}{n} \leq t$

Donc $\sqrt{\frac{\lfloor nt \rfloor}{n}} \xrightarrow[n\infty]{} \sqrt{t}$

Par TCL $\frac{S_{\lfloor nt \rfloor}}{\sqrt{\lfloor nt \rfloor}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

et par Slutsky : $B_t^n = \sqrt{\frac{\lfloor nt \rfloor}{n}} \frac{S_{\lfloor nt \rfloor}}{\sqrt{\lfloor nt \rfloor}} \xrightarrow[n\infty]{\mathcal{L}} \sqrt{t}\, \mathcal{N}(0,1)$

<u>Exercice 2</u> : Intervalle de confiance pour une pièce de monnaie.

On possède une pièce de monnaie qui a une probabilité inconnue p de tomber sur face. On aimerait déterminer un intervalle de confiance pour p. Pour cela, on demande à des volontaires de la lancer $n = 100$ fois de façon indépendante et de noter le nombre de fois S_n où une face a été observée

1. On note p la probabilité que la pièce tombe sur face. Quelle est la loi de S_n ?
2. Par quelle loi peut-on approcher $\dfrac{S_n - np}{\sqrt{np(1-p)}}$?
3. Les tables de statistique de la loi normale nous donne que $\displaystyle\int_{-1.96}^{1.96} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \simeq 0.95$. Montrer que $$\mathbb{P}\left(p \in \left[\frac{S_n}{n} - \frac{1.96\sqrt{p(1-p)}}{\sqrt{n}}, \frac{S_n}{n} + \frac{1.96\sqrt{p(1-p)}}{\sqrt{n}}\right]\right) \simeq 0.95.$$
4. En utilisant l'inégalité $p(1-p) \leq 1/4$, valable pour tout $p \in [0,1]$, en déduire l'intervalle de confiance à 95% :
$$\mathbb{P}\left(p \in \left[\frac{S_n}{n} - \frac{1.96}{2\sqrt{n}}, \frac{S_n}{n} + \frac{1.96}{2\sqrt{n}}\right]\right) \gtrsim 0.95.$$
5. <u>Application numérique :</u> On a lancé 100 fois la pièce et observé 41 faces. Déterminer un intervalle de confiance à 95% pour p. Est-il déraisonnable de dire que la pièce est équilibrée ?

<u>Exercice 3</u> : Théorème central limite vectoriel.

Le but est de montrer le TCL vectoriel à partir du TCL unidimensionnel :

Théorème 2 (TCL vectoriel) *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a dans $\mathbb{R}^d$ i.i.d de carré intégrable. On pose μ le vecteur moyen et Γ la matrice de covariance. Alors*
$$\sum_{i=1}^{n} \frac{X_i - \mu}{\sqrt{n}} = \sqrt{n}\,(\bar{X}_n - \mu) \xrightarrow[n\to\infty]{loi} \mathcal{N}(0,\Gamma).$$

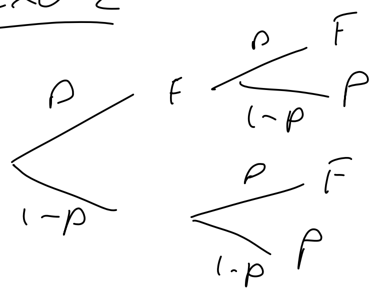
1. En utilisant le théorème de Paul Lévy, montrer qu'une suite de vecteur aléatoire $(Z_n)_{n\in\mathbb{N}}$ converge en loi vers un vecteur aléatoire Z si et seulement si pour tout $\lambda \in \mathbb{R}^d$, $\langle Z_n, \lambda\rangle \xrightarrow[n\to\infty]{loi} \langle Z, \lambda\rangle$.
2. Soit $Z \sim \mathcal{N}(0,\Gamma)$ et $\lambda \in \mathbb{R}^d$. Quelle est la loi de $\langle Z, \lambda\rangle$?
3. Appliquer le TCL à la suite de v.a unidimensionnelle $(\langle X_i, \lambda\rangle)_{i\in\mathbb{N}}$ puis conclure.

<u>Exercice 4</u> : Application du TCL vectoriel : loi multinomiale.

On considère un dé à 6 faces numérotées de 1 à 6. A priori, on ne sait pas si le dé est équilibré. Pour $k \in [\![1,6]\!]$, on note p_k la probabilité de faire k en lançant le dé. On lance n fois le dé de façon indépendante et on note D_i le résultat affiché par le dé au $i^{\text{ème}}$ lancer et $X_i^k = \mathbf{1}_{\{D_i = k\}}$. On note également le nombre de fois où le dé est tombé sur la face k : $S_n^k = \displaystyle\sum_{i=1}^{n} X_i^k$.

1. Quel est le vecteur moyen et la matrice de covariance du vecteur aléatoire $(X_i^1, X_i^2, X_i^3, X_i^4, X_i^5, X_i^6)$?
2. Montrer que $\dfrac{1}{\sqrt{n}}\big((S_n^1, S_n^2, S_n^3, S_n^4, S_n^5, S_n^6) - n(p_1, p_2, p_3, p_4, p_5, p_6)\big)$ converge en loi vers une variable aléatoire dont on précisera la loi.

Exo 2



S_n = # de fois où F a été observé

$$S_n \sim \mathcal{B}(n,p)$$

2) $$\frac{S_n - np}{\sqrt{np(1-p)}} = \frac{S_n - E[S_n]}{\sqrt{V(S_n)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

3) On trouve que

$$\mathbb{P}\left(p \in \left[\frac{S_n}{n} - \frac{1.96\sqrt{p(1-p)}}{\sqrt{n}}, \frac{S_n}{n} + \frac{1.96\sqrt{p(1-p)}}{\sqrt{n}}\right]\right)$$

$$= \mathbb{P}\left(-1.96 \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq 1.96\right)$$

Quand n est très grand, le TCL permet d'approcher la proba précédente

$\approx \mathbb{P}(-1.96 \leq X \leq 1.96)$ où $X \sim \mathcal{N}(0,1)$

$$= \int_{-1.96}^{1.96} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \approx 0.95$$

4) $p \in [0,1]$, on a $p(1-p) \leq \frac{1}{4} = \max_{p\in[0,1]} p(1-p)$

$$\left[\frac{S_n}{n} - \frac{1.96\sqrt{p(1-p)}}{\sqrt{n}}, \frac{S_n}{n} + \frac{1.96}{\sqrt{n}}\sqrt{p(1-p)}\right]$$

$$\subset \left[\frac{S_n}{n} - \frac{1.96}{2\sqrt{n}}, \frac{S_n}{n} + \frac{1.96}{2\sqrt{n}}\right]$$