Densité d'un vecteur gaussien

Théorème: Soit X un vecteur gaussien de dimension d, de moyenne $m = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_d) \end{pmatrix}$, et de matrice de variances/covariances $K = (\mathrm{Cov}(X_i, X_j))_{\substack{1 \leq i \leq d \\ 1 \leq j \leq d}}$.

Alors la loi de X admet pour densité

$$f_X(x) = \frac{1}{(2\pi)^{d/2} \sqrt{\det(K)}} \exp\left\{ -\frac{1}{2} (x-m)^t K^{-1} (x-m) \right\}$$

Cela signifie que pour tout pavé (ou plus généralement, tout borélien) B de $\mathbb{R}^d$,

$$\mathbb{P}(X \in B) = \int_B \frac{dx_1 \cdots dx_d}{(2\pi)^{d/2} \sqrt{\det(K)}} \exp\left\{ -\frac{1}{2} (x-m)^t K^{-1} (x-m) \right\}.$$

Idée de preuve:

Par changement de base, on se ramène au cas où $K = \mathrm{Diag}(\sigma_1^2, \dots, \sigma_d^2)$. Alors les composantes $X_1 \dots X_d$ de X sont indépendantes, chacune de loi gaussienne de paramètres (m_i, σ_i).

On en déduit $f_X(x) = \frac{1}{(2\pi)^{d/2} \sigma_1 \cdots \sigma_d} \exp\left\{ -\frac{1}{2} \sum_{i=1}^{d} \sigma_i^{-2} (x_i - m_i)^2 \right\}$

qui est la formule cherchée dans le cas diagonal.

III Convergence de suites de variables aléatoires.

A. Différentes notions de convergence.

Soit $(X_n)_{n \geqslant 1}$ une suite de v.a.r.

Définition :

1) X_n converge presque sûrement vers X

$$\left(X_n \xrightarrow[n\to\infty]{p.s.} X \right) \iff \left(\mathbb{P}\left(X_n(\omega) \underset{n\to\infty}{\longrightarrow} X(\omega) \right) = 1 \right)$$

2) X_n converge vers X dans L^p $(p \in \mathbb{N}^*)$

$$\left(X_n \xrightarrow[n\to\infty]{L^p} X \right) \iff \left(\mathbb{E}\left[|X_n - X|^p \right] \underset{n\to\infty}{\longrightarrow} 0 \right)$$

3) X_n converge vers X en probabilité

$$\left(X_n \xrightarrow[n\to\infty]{\mathbb{P}} X \right) \iff \left(\forall \varepsilon > 0 ,\ \mathbb{P}(|X_n - X| > \varepsilon) \underset{n\to\infty}{\longrightarrow} 0 \right)$$

4) X_n converge vers X en loi

$$\left(X_n \xrightarrow[n\to\infty]{\mathcal{L}} X \right) \iff \left(\begin{array}{l} \forall f : \mathbb{R} \to \mathbb{R} \text{ continue bornée,} \\ \mathbb{E}(f(X_n)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(f(X)) \end{array} \right)$$

Relations entre ces convergences :

Théorème :

ⓐ $(P.S.) \Longrightarrow (\mathbb{P}) \Rightarrow (\mathcal{L})$

ⓑ $(L^p) \Longrightarrow (\mathbb{P}) \Longrightarrow (\mathcal{L})$, $\forall p \in \mathbb{N}^*$

ⓒ $(L^p) \Rightarrow (L^q)$, $\forall\ p \geq q$

Toutes les autres implications sont fausses.

Remarque : La convergence en loi est d'une nature particulière : elle ne fait pas intervenir la différence $X_n - X$.

Exemple : $X \sim \mathcal{N}(0, 1)$, $X_n = -X$.

Alors X_n et X ont même loi, donc $\mathbb{E}(f(X_n)) = \mathbb{E}(f(X))$ $(\forall f)$, et $X_n \xrightarrow{\mathcal{L}} X$.

Mais $\mathbb{P}(|X_n - X| > \varepsilon) = \mathbb{P}(|X| > \frac{\varepsilon}{2}) \underset{n\to\infty}{\not\longrightarrow} 0$

Donc $X_n \overset{\mathbb{P}}{\not\longrightarrow} X$.

De même $\mathbb{E}(|X_n - X|^p) = 2\,\mathbb{E}(|X|^p) \underset{n\to\infty}{\not\longrightarrow} 0$.

Inégalité de Markov :
Soit X une V.A.R. admettant une espérance.
Alors $\left(\forall \varepsilon > 0,\ \mathbb{P}(|X| > \varepsilon) \leq \frac{\mathbb{E}(|X|)}{\varepsilon}\right)$

Preuve : Soit Y la VAR qui vaut 1 lorsque $|X| > \varepsilon$, et 0 sinon.
On a $|X| = |X|Y + |X|(1-Y) \geq |X|Y \geq \varepsilon Y$.
D'où $\mathbb{E}(|X|) \geq \varepsilon \mathbb{E}(Y) = \varepsilon \mathbb{P}(|X| > \varepsilon)$. $\square$

Inégalité de Bienaymé-Tchebychev.
Soit X une v.a.r. admettant une espérance et une variance. Alors :
$$\forall \varepsilon > 0,\ \mathbb{P}(|X - \mathbb{E}(X)| > \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

Preuve :

D'après Markov, pour tout $\varepsilon > 0$,
$$\mathbb{P}(|X - E(X)| > \varepsilon) = \mathbb{P}(|X - E(X)|^2 > \varepsilon^2)$$
$$\leq \frac{E(|X - E(X)|^2)}{\varepsilon^2} = \frac{\mathbb{V}(X)}{\varepsilon^2}.$$
$\square$

Application :

<u>Preuve de l'implication $(L^p) \Longrightarrow (\mathbb{P})$</u> :

On suppose que $X_n \xrightarrow[n\to\infty]{L^p} X$. Cela signifie que $\mathbb{E}(|X_n - X|^p) \xrightarrow[n\to\infty]{} 0$.

Soit $\varepsilon > 0$. D'après l'inégalité de Markov,

$$\mathbb{P}(|X_n - X| > \varepsilon)$$

$$= \mathbb{P}(|X_n - X|^p > \varepsilon^p)$$

$$\leq \frac{\mathbb{E}(|X_n - X|^p)}{\varepsilon^p} \xrightarrow[n\to\infty]{} 0 \quad . \quad \square$$

<u>Remarque</u> : Dans le cas $p = 2$, on aurait pu appliquer Bienaymé-Tchebycheff. La convergence (L^2) s'appelle aussi convergence en moyenne quadratique.

Critères de convergence :

Proposition : Soit (X_n) une suite de VAR admettant une moyenne et une variance. Soit $a \in \mathbb{R}$.

$$\left(X_n \xrightarrow[n\to\infty]{L^2} a\right) \Longleftrightarrow \left(\mathbb{E}(X_n) \to a \text{ et } \mathbb{V}(X_n) \to 0\right).$$

Preuve : $\mathbb{E}\left((X_n-a)^2\right) = \underbrace{\mathbb{V}(X_n)}_{\geq 0} + \underbrace{\left(\mathbb{E}(X_n)-a\right)^2}_{\geq 0}$. $\square$

Proposition (critères de convergence en loi) :

1) Si X_n à valeurs dans un même ensemble E fini ou dénombrable, alors

$$\left(X_n \xrightarrow{\mathcal{L}} X\right) \Leftrightarrow \left(\forall k \in E,\ \mathbb{P}(X_n = k) \longrightarrow \mathbb{P}(X=k)\right)$$

2) Si X_n V.A.R. quelconques,

$$\left(X_n \xrightarrow{\mathcal{L}} X\right) \Longleftrightarrow \left(F_{X_n}(x) \xrightarrow[n\to\infty]{} F_X(x),\ \text{pour tout point de continuité } x \text{ de } F_X\right)$$

3) S'il existe $\varepsilon > 0$ tel que $M_{X_n}(t)$ existe $< \infty$ pour tout $n \geqslant 1$ et tout $t \in [-\varepsilon, \varepsilon]$, alors

$$X_n \xrightarrow{\mathcal{L}} X \Longleftrightarrow \left(M_{X_n}(t) \xrightarrow[n\to\infty]{} M_X(t),\ \forall t \in [-\varepsilon, \varepsilon]\right).$$

RAPPEL : $\begin{cases} M_X(t) = \mathbb{E}(\exp(tX)) \\ F_X(x) = \mathcal{P}(X \leq x) \end{cases}$

Quelques propriétés utiles :

Proposition : Soit $a \in \mathbb{R}$.

$$X_n \xrightarrow[n\to\infty]{\mathcal{L}} a \iff X_n \xrightarrow[n\to\infty]{\mathbb{P}} a$$

Preuve : On va montrer ($\Rightarrow$). Soit $\varepsilon > 0$. Les réels $a-\varepsilon$ et $a+\varepsilon$ sont points de continuité de $F_a = \mathbb{1}_{\{x \geq a\}}$. Donc, sous l'hypothèse $X_n \xrightarrow{\mathcal{L}} a$, on a $F_{X_n}(a-\varepsilon) \xrightarrow[n\to\infty]{} 0$ et $F_{X_n}(a+\varepsilon) \xrightarrow[n\to\infty]{} 1$.

D'où $\mathbb{P}(|X_n - a| > \varepsilon) = F_{X_n}(a-\varepsilon) + (1 - F_{X_n}(a+\varepsilon)) \xrightarrow[n\to\infty]{} 0$ $\square$

Proposition : Soit $f : \mathbb{R}^2 \to \mathbb{R}$ continue.

$$\begin{cases} X_n \xrightarrow{\mathbb{P}} X \\ Y_n \xrightarrow{\mathbb{P}} Y \end{cases} \Longrightarrow f(X_n, Y_n) \xrightarrow{\mathbb{P}} f(X, Y)$$

Proposition : (Slutsky)

$$\left.\begin{array}{l} X_n \xrightarrow{\mathcal{L}} X \\ Y_n \xrightarrow{\mathcal{L}} a \end{array}\right\} \Longrightarrow \begin{cases} X_n + Y_n \xrightarrow{\mathcal{L}} X + a \\ X_n Y_n \xrightarrow{\mathcal{L}} aX \end{cases}$$

Proposition : Soit $g : \mathbb{R} \to \mathbb{R}$ continue.

(a) $\left(X_n \xrightarrow[n\to\infty]{\mathbb{P}} X\right) \Longrightarrow \left(g(X_n) \xrightarrow[n\to\infty]{\mathbb{P}} g(X)\right)$

(b) $\left(X_n \xrightarrow[n\to\infty]{\mathcal{L}} X\right) \Longrightarrow \left(g(X_n) \xrightarrow[n\to\infty]{\mathcal{L}} g(X)\right)$

B. Loi faible des grands nombres.

Théorème: Soit $(X_n)_{n \geq 1}$ une suite de v.a.r. indépendantes et de même loi (on note i.i.d., pour indépendantes et identiquement distribuées).

On suppose que $\mathbb{E}(|X_n|^2) < \infty$.

Soit $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$.

Alors $\overline{X}_n \xrightarrow[n \to \infty]{\mathbb{P}} m = E(X_1)$.

Preuve: On a

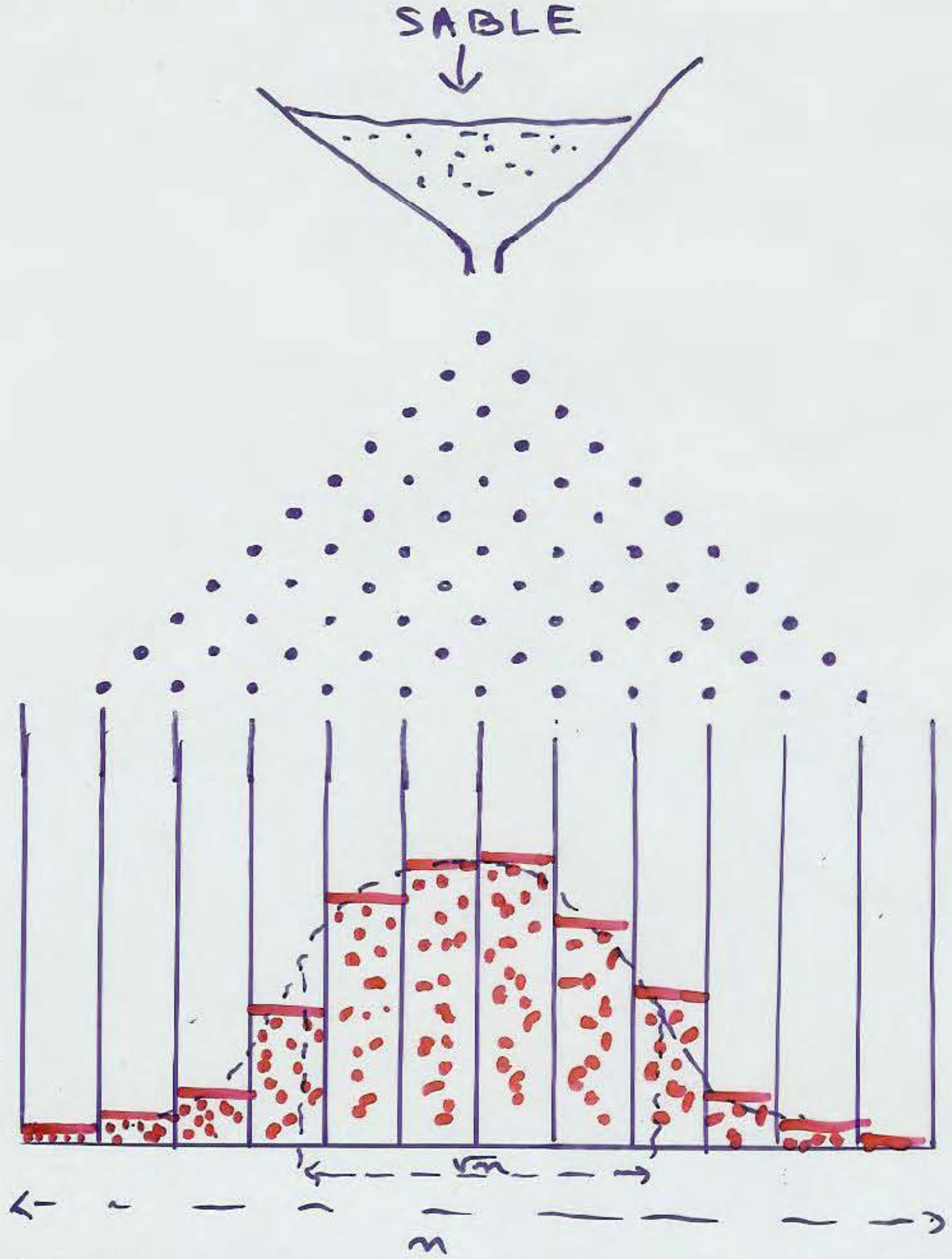
$$\mathbb{V}(\overline{X}_n) = \frac{1}{n^2} \mathbb{V}(X_1 + \cdots + X_n) = \frac{1}{n^2} \sum_{i=1}^{n} \mathbb{V}(X_i) = \frac{1}{n} \mathbb{V}(X_1) \xrightarrow[n \to \infty]{} 0.$$

De même, $\mathbb{E}(\overline{X}_n) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = E(X_1) = m$.

Donc $\overline{X}_n$ converge vers m en moyenne quadratique, ce qui entraîne la convergence en probabilité, comme conséquence de Bienaymé-Tchebychef. $\square$

C. Le Théorème central limite

Une petite expérience :



Observation : La répartition des grains de sable dans les cases semble suivre une loi de Gauss de moyenne 0 et de variance $\sigma^2 = (\sqrt{n})^2$.

Théorème: Soit $(X_n)_{n\geq 1}$ une suite de V.A.R. indépendantes et de même loi (i.i.d.), telles que $\mathbb{E}(X_n^2) < \infty$, $\mathbb{E}(X_n) = m$ et $\mathbb{V}(X_n) = \sigma^2$.

Alors

$$Z_n = \frac{\sqrt{n}}{\sigma}(\overline{X}_n - m) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

Preuve: On démontrera le Thm sous l'hypothèse supplémentaire qu'il existe $\varepsilon > 0$ tel que $M_{X_n}(t) = \mathbb{E}(\exp(tX_n)) < \infty$, $\forall t \in [-\varepsilon, \varepsilon]$. On commence par définir les V.A.R. $Y_n = \frac{1}{\sigma}(X_n - m)$, de sorte que $\mathbb{E}(Y_n) = 0$ et $\mathbb{V}(Y_n) = 1$, et $Z_n = \sqrt{n}\,\overline{Y}_n$.

Alors $M_{Z_n}(t) = M_{\frac{Y_1+\cdots+Y_n}{\sqrt{n}}}(t) = M_{Y_1+\cdots+Y_n}\left(\frac{t}{\sqrt{n}}\right)$

$$= \left[M_{Y_1}\left(\frac{t}{\sqrt{n}}\right)\right]^n.$$

Mais pour $\frac{t}{\sqrt{n}} \ll 1$, $M_{Y_1}\left(\frac{t}{\sqrt{n}}\right) = 1 + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)$.

Cela vient de ce que $M_{Y_1}(0) = 1$, $M'_{Y_1}(0) = E(Y_1) = 0$, $M''_{Y_1}(0) = \mathbb{V}(Y_1) = 1$.

Donc $M_{Z_n}(t) = \left(1 + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n$

$$= \exp\left\{n \operatorname{Log}\left(1 + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)\right\}$$

$$= \exp\left(\frac{t^2}{2}(1 + o(1)_{n\to\infty})\right) \xrightarrow[n\to\infty]{} \exp\left(\frac{t^2}{2}\right) = M_{\mathcal{N}(0,1)}(t)$$

□