République Algériennne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Mohamed Khider, Biskra

Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie

Départementde Mathématiques

Mémoire présenté pour obtenir le diplôme de

Master en "**Mathematiques Appliquées**"

Option : Probabilités

Par GUELLAI Rima

Titre :

Lois dans $\mathbb{R}^d$, convergence des lois

Devant le Jury :

Dr.	Labed Boubakeur	U. Biskra	Président
Dr.	Zaghdoudi Kadhem	U. Biskra	Encadreur
Dr.	Bougherara Saliha	U. Biskra	Examinatrice

Soutenu Publiquement le 27/06/2022

Dédicace

A mon père **Guellai Abdel Karim** d.être pour moi le meilleur du père

Que Dieu le protège et prolonge sa vie.

A ma mère **Triche.L** qui m'a soutenu et encouragé durant ces années d'études.

Qu' elle trouve ici le témoignage de ma profond reconnaissance.

A mes sœurs, je leur souhaite de réussir dans leur vie et a chacun ma famille

Remerciements

E tiens tout d'abord à remercier **ALLAH** qui m'a donné la force et la patience d'accomplir ce modeste travail.

En second lieu, je tiens à remercier mon encadrant **Dr Zaghdoudi Kadhem**, pour l'orientation et les conseils qu'elle m'a prodigués pendant la période du travail.

Mes vifs remerciements vont également aux mombres du jury pour l'intérêt qu'ils ont porté à ma recherche en acceptant d'examiner mon travail et de l'enrichir par leurs proposition.

Mes remerciements s'étendent également à tous mes enseignants durant les années d'études et à ma famille surtout mon père, ma mère, mes sœurs et mes amis qui par leurs prières et leurs encouragements, j'ai pu surmonter tous les obstacles.

Enfin, je tiens à remercier toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

Notations et symbols

$V.a$	Variable aléatoire.
$E(.)$	Espérance mathématique.
$V(.)$	Variance mathématique.
$Cov(X,Y)$	Covariance mathématique du couple (X,Y).
$D(o,r)$	Disque de rayon r et centre o.
$D^{\circ}(o,r)$	Disque ouvert.
Ω	L'espace fondamental.
$\mathcal{F}$	Tribu définis sur l'espace fondamental Ω.

Table des matières

Introduction

La théorie des probabilités est une théorie mathématique ayant pour but de décrire et d'interpréter les phénomènes dans lesquels le hasard intervient. Cette théorie permet de comprendre comment les procédures inférentielles sont développées et utilisées et quelles en sont les limites. Dans un problème de probabilité, on suppose connues les propriétés d'une population et on répond à des questions concernant un échantillon prélevé dans cette population.

Cette mémoire est consacré aux diverses notions de convergence d'une suite de variables aléatoires, principalement la convergence en loi, mais aussi la convergence en probabilité, la convergence presque-sûre et la convergence en moyenne quadratique sont énoncées.

La convergence en loi des variables aléatoires est un type de convergence est sans doute le plus délicat à comprendre, en partie parce qu'il s'agit d'une convergence de mesures (ce sont les lois des variables aléatoires qui convergent et non les variables elles-mêmes).

Cette note est composée de trois chapitres :

Dans le Chapitre 1 : La première partie est un rappel de base des variables aléatoires, avec un rappel des deux grandes familles de variables aléatoires que sont les variables discrètes et les variables strictement continues, et la seconde partie de ce chapitre est une introduction aux vecteurs aléatoires et à leurs types (vecteurs discrets et vecteurs strictement continus). Le mode de calcul de la fonction caractéristique est également mentionné.

Dans le Chapitre 2 : Le deuxième chapitre s'attache à donner un exemple pour chaque type de vecteur aléatoire avec quelques caractéristiques de chaque type. Le premier exemple est Le vecteur absolument continu gaussien, quant au deuxième exemple le vecteur discrète multinomial.

Dans le Chapitre 3 : La première partie commence donc par donner la notion de convergence presque sûre. On introduit ensuite la convergence en probabilité.

On introduit dans une deuxième partie la convergence L_p et équi-intégrabilité héritée de la théorie de la mesure, et on étudie notamment les propriétés des espaces L_p.

La dernière partie est consacrée à la convergence en loi, et une relation entre la convergence des fonctions caractéristiques et convergence en loi.

Et la dernière relation, nous mentionnons la relation entre les types de convergence et quelques exemples de convergence des lois.

....

Chapitre 1

Fonction caractéristique d'une loi dans $\mathbb{R}^d$

Dans ce chapitre, certaines propriétés générales des variables aléatoires et des vecteurs aléatoires seront définies et mentionnées, ainsi que la façon de calculer la fonction caractéristique.

1.1 Variables aléatoires

Remarquons tout d'abord qu'en dépit de sa dénomination, une variable aléatoire n'est pas une variable au sens habituel.

C'est une application d'un espace probabilisé $(\Omega, \mathcal{F}, P)$ dans un espace mesurable (E, ξ), ou' ξ désigne une tribu définie sur E. Une variable aléatoire, que l'on désignera toujours par la majuscule d'une des dernières lettres de l'alphabet, permet de munir (E, ξ) d'une probabilité image. La seule condition exigée pur réalisé cet objectif est de supposer que la variable aléatoire soit mesurable.

Définition 1.1.1 *Toute application mesurable X d'un espace probabilisé $(\Omega, \mathcal{F}, P)$ dans un espace (E, ξ) définit une variable aléatoire. X vérifie donc la propriété de mesurabilité*

$$\forall B \in \xi, \qquad X^{-1}(B) \in \mathcal{F}.$$

On construit sur (E, ξ) une probabilité dite probabilité image de la probabilité P par la variable aléatoire X, grâce ou théorème,

Théorème 1.1.1 *La mesure P_X définie sur (E, ξ) par*

$$\forall B \in \xi, \quad P_X(B) = P(X^{-1}(B)).$$

Et la probabilité image de P par X définie sur (E, ξ)

Preuve. vérifions les deux axiomes définissant une probabilité

1. $$P_X(E) = P(X^{-1}(E)) = P(\Omega) = 1$$

2. $$P_X(\cup B_i) = P(X^{-1}(\cup B_i)) = \sum_i P(X^{-1}(B_i)) = \sum_i P_X(B_i).$$

   Si l'espace E est égal à $\mathbb{R}$, ou le munit de sa tribu borélienne et les variables aléatoires définies sur E sont dites réelles.

■

Variables aléatoires discrètes

Définition 1.1.2 *Une v.a réelle X définie sur $(\Omega, \mathcal{F}, P)$ est discrète si $P(\Omega)$ est un ensemble au plus dénombrable de $\mathbb{R}$*

$$X(\Omega) = \{x_i\}_{i \subset I \subseteq \mathbb{Z}}$$
$$\forall i P_X(x_i) = P(X^{-1}(x_i)) = p_i.$$

On pourra écrire : $P_X = \sum_i p_i \delta_{x_i}$ ou δ_{x_i} est la mesure de support $\{x_i\}$. Rappelons que tout borélien B, est égal a' 1 si et à 0 sinon.

Variables aléatoires continue à densité

Définition 1.1.3 *une v.a X est dite absolument continue ou continue, s'il existe une fonction f_X presque surement positive dite densité de probabilité (ou densité) de X d'intégrale sur $\mathbb{R}$ égale à l'unité et telle que pour tout borélien B de $\mathbb{R}$,*

$$P_X = \int_B f_X(x)dx,$$

le resultat précédent s'exprime sous une forme plus intuitive

$$\forall x \in \mathbb{R}, \quad P(X \in [x, x + \triangle x]) = f_X(x)\triangle x + o(\triangle x).$$

Rappel Une fonction vérifie presque sûrement une propriété, si elle la vérifie partout sauf éventuellement sur un ensemble de mesure nulle.

1.1.1 Fonction de répartition

Définition 1.1.4 *Si X est une v.a la fonction de répartition de X est la fonction $F_X : \mathbb{R} \to [0,1]$ définie par*

$$F_X(t) = P(X \leq t) = P_X(]-\infty, t]), \qquad \forall t \in \mathbb{R}.$$

Propriété 1.1.1

1. *La fonction F_X est continue à droite et $\lim_{t \to -\infty} F_X(t) = 0$, $\lim_{t \to +\infty} F_X(t) = 1$. Inversement, si on se donne une fonction F ayant ces propriétés, alors il existe un unique mesure de probabilité μ telle que*

$$\mu(]-\infty, t]) = F_X(t), \qquad \forall t \in \mathbb{R}.$$

2. *Si X est une v.a discrète*

$$F_X(t) = \sum_{x \in]-\infty, t]} P(X = x).$$

3. *Si X est une v.a continue*

$$F_X(t) = \int_{x \in]-\infty, t]} f(x)dx.$$

4. *soit a,b $\in \mathbb{R}$ avec $a < b$ alors*

$$P(a < X \leq b) = F_X(b) - F_X(a).$$

Exemple 1.1.1 *Si X est une v.a binomiale $\beta(n, p)$, on a*

$$P_X(x) = C_n^k p^k (1-p)^{n-k}, \qquad \forall x \in \{0, 1, ..., n\}$$

alors

$$F_X(x) = \sum_{k=0}^{x} C_n^k p^k (1-p)^{n-k},$$

ou x est le plus grand entier inferieur ou égale à X.

1.1.2 Fonction génératrice d'une variale aléatoire

Définition 1.1.5 *A' toute v.a X a' valeurs dans $\mathbb{N}$, ou associe sa fonction génératrice*

$$G_X(z) = E(z^X) = \sum_{n\in\mathbb{N}} P(X = x) z^n, \qquad \forall z \in D(0,1) \subset \mathbb{C}.$$

$G_X(z)$ est la transformée en z (utilisée par les automaticiens) de la suite $\{P(X = n)\}_{n\in\mathbb{N}}$, continue dans son disque de convergence $D(0,1)$, holomorphe dans son intérieur $D^\circ(0,1)$.

La fonction génératrice d'un v.a entière positive caractérise complètement sa loi. D'un point de vue pratique, son utilisation facilite souvent le calcule de l'espérance et de la variance d'une v.a.

Définition 1.1.6

La fonction génératrice G_X d'un v.a entière X caractérise sa loi

$$\forall n \in \mathbb{N}, \quad P_X(n) = \frac{1}{n!} \left. \frac{d^x G_X(z)}{dz^n} \right|_{z=0}.$$

Pour toute v.a X entière telle que $E(X^i)$ assiste

$$\forall i \in \{1, 2, ..., n\} : \; E(X) = \left. \frac{dG_X(z)}{dz} \right|_{z=1}, \qquad E(X^2) = \left. \frac{d^2 G_X(z)}{dz^2} \right|_{z=1} + E(X)$$

$$\forall i \in \{1, 2, ..., n\} : E(X(X-1)...(X-(n-1))) = \left. \frac{d^i G_X(z)}{dz^i} \right|_{z=1}.$$

1.1.3 Fonction caractéristique

Définition 1.1.7 *Soit X une v.a (discrète ou continue) on appelle fonction caractéristique de X l'application définie par*

$$\forall t \in \mathbb{R} \quad \varphi_X(t) = E(\exp itX) = E(\cos tX + i \sin tX),$$

les v.a $\cos tX$ et $\sin tX$ sont mesurables, bornées par 1 donc intégrable alors $\varphi_X(t)$ est bien définie.

1. Si X est une v.a absolument continue alors

$$\forall t \in \mathbb{R} \varphi_X(t) = \int_{-\infty}^{+\infty} \exp(itx) f(x) d_x.$$

2. Si X est une v.a discrète

$$\forall t \in \mathbb{R} \varphi_X(t) = \sum_n p_n \exp(itx).$$

Théorème 1.1.2 *La fonction caractéristique d'un somme de v.a indépendantes $X_1, ..., X_n$ de lois quelconques est égale ou produit des fonctions caractéristique*

$$\forall t \in \mathbb{R}, \qquad \varphi_{\sum_{j=1}^n X_j}(t) = \Pi_{j=1}^n \varphi_{X_j}(t)$$

Preuve. Par définition

$$\varphi_{\sum_{j=1}^n X_j}(t) = E\left(\exp i \sum_j^n X_j t\right)$$

donc :

$$= \Pi_{j=1}^n E(\exp itX_j) = \Pi_{j=1}^n \varphi_{X_j}(t).$$

■

Exemple 1.1.2 *1. X_1 variable de Bernoulli, de loi de probabilité*

$$P_{X_1} = p\delta_1 + (1-p)\delta_0$$

2. X_2 variable

$$\mathcal{N}(0,1)\varphi_{X_2}(t) = \frac{\exp -\frac{t^2}{2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp(-\frac{1}{2}(x-it)^2)dx.$$

(Intégrale complexe que l'on calcule sur contour rectangle standard centré en 0), est égale à

$$\varphi_{X_2}(t) = \exp -\frac{t^2}{2}.$$

1.2 Vecteurs aléatoires

Description des vecteurs aléatoires

Dans ce qui suite on désignera par β_d la tribu des boréliens de $\mathbb{R}^d$.

Définition 1.2.1 *Un vecteur aléatoire réel de dimension n est un vecteur dont les composantes sont des v.a réelles, il est décrit par une application mesurable $X = (X_1, ..., X_d)$ définie sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ à valeurs dans $(\mathbb{R}, \beta_d)$. Pour tout évènement B appartenant à la tribu β_d,*

$$X^{-1}(B) \in \mathcal{F}.$$

Définition 1.2.2 *La loi de probabilité P_X dite loi conjointe de X est définit sur*

$(\mathbb{R}^d, \beta_d)$ *par*

$$\forall B \in \beta_d, \quad P_X(B) = P((X_1, X_2, ..., X_d) \in B).$$

La fonction de répartition $F_X(x_1, x_2, ..., x_d)$ *est définie par*

$$\forall (x_1, x_2, ..., x_d) \in \mathbb{R}^d, \quad F_X(x_1, x_2, ..., x_d) = P(X_1 \leq x_1, X_2 \leq x_2, ..., X_d \leq x_d).$$

1.2.1 Vecteurs discrets

Définition 1.2.3 *Un vecteur aléatoire* X *discret de dimension* d *prend ses values dans un ensemble dénombrable* E *inclus dans* $\mathbb{R}^d$.

Définition 1.2.4 *La loi* P_X*, dite loi conjointe du vecteur* X*, est définie par*

$$\forall x \in E, \quad P_X(x_1, x_2, ..., x_d) \overset{Déf}{=} P(X_1 = x_1, ..., X_d = x_d).$$

Théorème 1.2.1 *La loi* P_{X_i} *de* X_i*, dite loi marginale de* X_i *est définie par*

$$P_{X_i}(x_i) = P(X_i = x_i) = \sum_{x_1 \in E_1} ... \sum_{x_{i-1} \in E_{i-1}} \sum_{x_{i+1} \in E_{i+1}} ... \sum_{x_n \in E_d} P_X(x_1, ..., x_d).$$

Preuve.

$$P_{X_i}(x_i) = P(X_i = x_i) = P(X_1 \in E_1, ..., X_i = x_i, ..., X_d \in E_d)$$
$$\sum_{x_1 \in E_1} ... \sum_{x_{i-1} \in E_i} \sum_{x_{i+1} \in E_{i+1}} ... \sum_{x_d \in E_d} P_X(x_1, x_2, ..., x_d).$$

■

1.2.2 Vecteurs absolument continus

Définition 1.2.5 *La loi P_X du vecteur absolument continu $X = (X_1, X_2, ..., X_d)$ est définie par sa densité $f_X(x_1, x_2, ..., x_d)$*

$$\forall B \in \beta_d, \quad P_X(B) = \int\limits_B f_X(x_1, x_2, ..., x_d) dx_1 dx_2 ... dx_d,$$

La densité f_X est une fonction positive de $\mathbb{R}^d$ dans $\mathbb{R}$, d'intégrale égale à l'unité.

Théorème 1.2.2 *La loi marginale f_{X_i} de X_i est définie par*

$$f_{X_i}(x_i) = \int\limits_{\mathbb{R}^{d-1}} f_X(x_1, x_2, ..., x_d) dx_1 ... dx_{i-1} dx_{i+1} ... dx_d.$$

1.2.3 Fonction génératrice et caractéristique

Ces deux concepts sont d'une grande efficacité opératoire lorsqu'il s'agit de déterminer les lois de sommes ou de différences de v.a indépendantes, ou des limites de suites de variables ou vecteurs aléatoires.

Fonction génératrice

Définition 1.2.6 *La fonction génératrice d'un vecteur aléatoire $(X_1, ..., X_d)$ à valeurs entières positives est définie par*

$$\forall i \in \mathbb{N}, \quad \forall z_i \in D(0,1), \quad G_{X_1,...,X_d}(z_1, ..., z_d) = E\left(z_1^{X_1}, ..., z_d^{X_d}\right).$$

On démontre facilement des propriétés analogues à celles de la fonction génératrice d'une v.a par exemple $E(X_i)$

$$E(X_i) = \left.\frac{\partial G_{(X_1,...,X_d)}(z_1,...,z_d)}{\partial z_i}\right|_{(z_1=z_2=...=z_d=1)}.$$

Fonction caractéristique

La fonction caractéristique d'une v.a ou d'un vecteur aléatoire quelconque X est la transformation de Fourier de sa mesure de probabilité, qui existe toujours et bénéficie de ses puissantes propriétés.

À l'instar de la fonction génératrice, elle caractérise complètement la loi de probabilité et en contient donc toute l'information.

Définition 1.2.7 *La fonction caractéristique d'un vecteur aléatoire* $X = (X_1, ..., X_d)$ *de loi* P_X *est définie par*

$$\forall t \in \mathbb{R}^d, \quad E(\exp i \langle X, t\rangle) = \int_{\mathbb{R}^d} \exp i \langle X, t\rangle \, dP_X(x_1, ..., x_d) \overset{notée}{=} \varphi_X(t_1, ..., t_d),$$

($\langle X, t\rangle$ *désigne le produit scalaire* $\sum_j t_j X_j$ *et* $j = 0, ..., d$*)*

Remarque 1.2.1 *Si* X *est un vecteur aléatoire de densité* $f_X(x_1, x_2, ..., x_d)$ *sa fonction caractéristique est définie par*

$$\varphi_X(t_1, t_2, ..., t_d) = \int_{\mathbb{R}^d} \exp\left(i \sum_{j=1}^{d} t_j X_j\right) f_X(x_1, x_2, ..., x_d) dx_1 dx_2 ... dx_d.$$

Théorème 1.2.3 *Étant donné deux vecteurs aléatoires* X *et* Y *de même dimen-*

sion et de fonctions caractéristique φ_X *et* φ_Y.

$$\varphi_X = \varphi_Y \iff X \text{ et } Y \text{ ont même loi.}$$

Théorème 1.2.4 *1. Soit la v.a* X *de densité* f_X *et telle que* $\varphi_X \in L^1(\mathbb{R})$, *alour :*

$$f_X(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp(-iXt)\varphi_X(t)dt, \qquad p.p.$$

2. Soit le vecteur X *de densité* f_X *tell que* $\varphi_X \in L^1(\mathbb{R}^n)$ *alors :*

$$f_X(x_1, ..., x_d) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \varphi_X(t_1, t_2, ..., t_d) \exp -i(t_1x_1 + ... + t_dx_d)dt_1, ..., dt_d.$$

1. Si X est à valeurs discrètes x_d

$$P[X = x_d] = \lim_{T \longrightarrow +\infty} \frac{1}{2T} \int_{-T}^{T} \exp(-itx_d)\varphi_X(t)dt.$$

Indépendantes des composantes d'un vecteur aléatoire

Théorème 1.2.5 *Les v.a* $X_1, ..., X_d$ *sont mutuellement indépendantes si pour toute d-aplet de boréliens* $(B_1, B_2, ..., B_d)$ *de* $\mathbb{R}^d$

$$P(X_1 \in B_1, X_2 \in B_2, ..., X_d \in B_d) = P(X_1 \in B_1)...P(X_d \in B_d).$$

Théorème 1.2.6 *Si le vecteur aléatoire $X = (X_1, X_2, ..., X_d)$ est de densité f_X et les v.a X_i de densités de respectives f_{X_i}, alors*

$$\text{(Indépendance des } X_i) \iff f_X(x_1, x_2, ..., x_d) = \Pi_{i=1}^{d} f_{X_i}(x_i).$$

Preuve. $\Longleftarrow$

$$\begin{aligned} P_X(B_1 \times B_2 \times ... \times B_d) &= P(X_1 \in B_1, ..., X_d \in B_d) \\ &= \int\limits_{B_1 \times B_2 \times ... \times B_d} f_X(x)dx \\ &= \int\limits_{B_1 \times B_2 \times ... \times B_d} f_{X_1}(x_1) f_{X_2}(x_2) ... f_{X_d}(x_d) dx_1 dx_2 ... dx_d \\ &= \int\limits_{B_1} f_{X_1}(x_1)dx_1 \int\limits_{B_2} f_{X_2}(x_2)dx_2 ... \int\limits_{B_n} f_{X_d}(x_d)dx_d \\ &= P_{X_1}(B_1)...P_{X_d}(B_d). \end{aligned}$$

$\Longrightarrow$ La réciproque est triviale. ∎

Proposition 1.2.1 *Les variable aléatoires $X_1, X_2, ..., X_d$ sont indépendantes si et seulement si la fonction caractéristique du vecteur $(X_1, X_2, ...X_d)$ est le produit des fonction caractéristique des coordonnées*

$$\varphi_x(t_1, ..., t_d) = \varphi_x(t_1) ... \varphi_{x_d}(t_d).$$

1.2.4 Espérance et la matrice de covariance

Définition 1.2.8 *On définit l'espérance d'un vecteur aléatoire comme le vecteur des espérances existe*

$$E(X) = (E(X_k))_{k=1,...,d}.$$

De même, l'espérance d'une matrice à coefficients aléatoires est la matrice des espérances des coefficients.

Définition 1.2.9 *On appelle matrice de covariance du vecteur aléatoire* $X = (X_k)_{k=1...d}$ *la matrice*

$$K_X = (cov[X_h, X_k])_{h,k=1,...,d}.$$

Si ces covariances existent. La manipulation des matrices de covariance est grandement simplifiée par la notation matricielle. Nous écrirons un vecteur X *de* $\mathbb{R}^d$ *comme une matrice à* d *liens et une colonne. Sa transposée, notée* X^t *est une matrice à* 1 *ligne et* d *colonnes. On peut écrire la matrice de comme suit.*

$$K_X = E[(X - E(X))(X - E(X))^t].$$

Chapitre 2

Lois et fonctions caractéristiques usuelles dans $\mathbb{R}^d$

Il existe deux types des les lois usuelles dans $\mathbb{R}^d$, les lois usuelles discrètes et les lois usuelles continues.

2.1 Vecteur gaussien

Définition 2.1.1 *Un vecteur aléatoire* $X = (X_1, X_2, ..., X_d)$ *est gaussien ou normale si toute combinaison linéaire* $\sum_{i=1}^{d} a_i X_i$, $(a_i \in \mathbb{R})$ *de ses composantes est une variable gaussien.*

Théorème 2.1.1 *Les composantes d'un vecteur gaussien son des v.a gaussiens*

Preuve. Pour toute i, on choisit $a_i = 1$ dans la définition précédent et pour tout $j \neq i$ et $a_j = 0$, on conclut à la normalité de X_i grâce à la définition. La définition précédent est fondée sur le théorème de Cramer-Wold qui affirme que la lois d'un vecteur aléatoire $(X_1, X_2, ..., X_d)$ est caractérisée par la connaissance des lois de toutes les combinaisons linéaires possibles de ses composantes. ■

Théorème 2.1.2 *Si les v.a indépendantes* $(X_i)_{i=1,\ldots,d}$ *sont gaussienne, alors le vecteur* $X = (X_1, X_2, ..., X_d)$ *est gaussiens.*

Preuve. $\forall (a_i)_{i=1,\ldots,d} \in \mathbb{R}$, les v.a $a_1X_1, ..., a_dX_d$ sont gaussienne et d'après la définition le vecteur $(X_1, X_2, ..., X_d)$ est gaussiens. ■

Remarque 2.1.1 *Soif si elles sont indépendantes, la normalité de chacune des composantes n'implique pas celle du vecteur, comme le montre l'exemple suivant*

Exemple 2.1.1 *Soit le vecteur* (X, Y) *de dimonsion deux, de densité* $f_{X,Y}$ *égale à* $\frac{1}{\pi}\exp\left(-\frac{x^2+y^2}{2}\right)$ *sur le premier et le troisième quadrant déterminé par un repère orthonormé, et nulle sur les deux autres quadrants. Par construction le vecteur* (X, Y) *n'est pas gaussien, mais ses densités marginales le sont, respectivement égales à* $\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right)$ *et* $\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-y^2}{2}\right)$*.Le produit des densités marginale n'étant pas à la densité conjointe, les composantes sont dépendantes.*

Définition 2.1.2 *Un vecteur gaussien* X^* *est centré réduite, si* $E(X^*) = 0$ *et si* $K_{X^*} = Id$.

Remarque 2.1.2 1. *Les composantes* X_i^* *sont indépendantes et pour tout couple*

$$(i, j), V(X_i) = 1, \mu_i = 0, \qquad cov(X_i, X_j) = 0.$$

2. *La densité* f_X *est définie par*

$$\forall x \in \mathbb{R}^d, \quad f_X(x) = (2\pi)^{-\frac{k}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{d} x_i^2\right).$$

2.1.1 Propriétés des vecteurs gaussiens

1. Soit $X \in \mathbb{R}^d$ un vecteur gaussien et K sa matrice de covariance. Les variables X_d sont indépendantes si et seulement si la matrice K est diagonale.
2. soit $\mu \in \mathbb{R}$ et K un matrice symétrique d'ordre d, définie positive ($\det K \neq 0$).

La loi normale $N(\mu, K)$ admet pour densité.

$$\forall x \in \mathbb{R}^d : \quad f_X(x) = \frac{1}{(\sqrt{2\pi})^d} \frac{1}{\sqrt{\det K}} \exp(-\frac{1}{2}(x-\mu)^t K^{-1}(x-\mu)).$$

Proposition 2.1.1 *L'image linéaire d'un vecteur gaussien est encore un vecteur gaussien, soit $X = (X_1, ..., X_d)$ un vecteur gaussien de dimension d et $A \in M_{(p,d)}(\mathbb{R})$. Alors AX est un vecteur gaussien en dimension p.*

Preuve. En effet si $a \in \mathbb{R}^p$ alors $\langle a, AX\rangle = \langle A^t a, X\rangle$ suit une loi normale car combinaison linéaire des marginales du vecteur gaussien X. ■

2.1.2 Fonction caractéristique du vecteur gaussien

Définition 2.1.3 *Soit X un vecteur gaussien, μ son espérance et K sa matrice de covariance. La fonction caractéristique du vecteur X est définie pour tout $u \in \mathbb{R}^d$ par*

$$\varphi_X(t) = \exp\left[u^t\mu - \frac{1}{2}u^t K u\right].$$

Preuve. Posons $Y = u^t X$. C'est une combinaison linéaire des cordonnées de X. Donc Y suit la lio normale d'espérance $u^t\mu$ et de variance $u^t K u$. La fonction caractéristique de Y évaluée au point τ, vaut donc

$$\varphi_X(\tau) = \exp(i\tau(u^t\tau) - \frac{\tau^2}{2}u^t K u).$$

Pour conclure, il suffit de remarquer que

$$\varphi_X(u_1, u_2, ..., u_d) = \varphi_Y(1).$$

■

Remarque 2.1.3 *La fonction caractéristique du vecteur gaussien X^* centré réduite de dimension d est définie par*

$$\begin{aligned}\varphi_{X^*} &= \exp(-\frac{1}{2}u^t u) \\ &= \exp(-\frac{1}{2}\sum_{j=1}^{d} u_j^2).\end{aligned}$$

2.2 Vecteur multinomial

Le vecteur multinomial est l'extension de dimension d de la variable binomial. soit l'espace fondamental $\Omega = \{A_1, A_2, ..., A_d\}$, ou 'chaque évènement A_i a la probabilité de réalisation p_i, sachant que $\sum_{i=0}^{d} p_i = 1$. On procède à k tirages avec remise dans Ω le vecteur aléatoire $(X_1, X_2, ..., X_d)$ compte les réalisation respectives $(n_1, n_2, ..., n_d)$ des événements $A_1, A_2, ..., A_d$. D'un point de vue pratique, ce modèle probabiliste décrit un sondage avec remise par échantillons de taille d, dans une population partitionnée en d sous-ensembles, décrivant par exemple les classe socioprofessionnelles d'un pays ou d'une région.

Calculons le nombre d'échantillons réalisant l'évènement (n_i fois A_i, n_2 fois A_2,..., n_d fois A_d).Compte tenu de l'indépendance des tirages, on démontre que le nombre de tels échantillons est égale à

$$C_k^{n_1} C_{k-n_1}^{n_2} ... C_{k-(n_1+n_2+...+n_{d-1})}^{n_d} = \frac{k!}{n_1! n_2! ... n_d!}.$$

Or la probabilité de tirer un échantillon de la forme ci-dessus est égale à $p_1^{n_1} ... p_d^{n_k}$. D'ou'

$$P(X_1 = n_1, X_2 = n_2, ..., X_d = n_d) = \frac{k!}{n_1! n_2! ... n_d!} p_1^{n_1} ... p_d^{n_k}.$$

On vérifie facilement que X_i est binomiale de loi $\beta(k, p_i)$ et que

$$P\left(X_1 = n_1, X_2 = n_2, ..., X_d = n_d\right) \neq P(X_1 = n_1) ... P(X_d = n_d),$$

pour tout $(n_1, ..., n_d)$ de somme k, ce qui équivaut à la dépendance des v.a $(X_i)_{i=1,...,d}$.

Définition 2.2.1 *La vecteur aléatoire $X = (X_1, ..., X_d)$ est dit multinomial de paramètres $(p_i)_i$, si pour tout d-uplet $(n_1, n_2, ..., n_n)$ à valeurs dans $\{0, 1, ..., k\}$ et* *tels que* $\sum_{i=1}^{d} n_i = k$

$$P(X_1 = n_1, ..., X_n = n_n) = \frac{k!}{n_1!n_2!...n_n!} p_1^{n_1} p_2^{n_2} ... p_n^{n_k},$$

où les p_i *sont des réels positifs de somme égale à un.*

Définition 2.2.2 *Un vecteur aléatoire* X *à valeurs dans* $\mathbb{N}^d$*, suite la loi multinomiale* $\mathcal{M}(n, p_1, ..., p_n)$*, tell que* $n \in \mathbb{N}$ *et* $\sum_{i=1}^{d} p_i = 1, p_1, ..., p_d \in [0,1]$ *si*

$$\begin{cases} P((X_1, ..., X_d) = (n_1, ..., n_d)) = \frac{n!}{n_1!...n_d!} p_1^{n_1} ... p_d^{n_d} \; pour \; n_1 + n_2 + ... + n_d = n \\ 0 \qquad sinon \end{cases}.$$

2.2.1 Espérance et la matrice de covariance

X_i est de loi binomiale $\beta(kp_i, kp_i(1-p_i))$, alors

$$E(X) = (kp_1, ..., kp_d).$$

Matrice de covariance de X

Posons

$$U_j^i = \begin{cases} 1 & \text{si le } i\text{-ème targe fouit apparaitre } A_j. \\ 0 & \text{sinon.} \end{cases},$$

il est facile de vérifier que $(X_1, X_2, ..., X_d) = \sum_{i=1} \left(U_1^{(i)}, ..., U_d^{(i)}\right)$ ou' les $U_j^{(i)}$ sont des v.a indicatrices des évènements A_j pour chacun des k tirages, les vecteurs $V^{(i)} = (U_j^{(i)}, ..., U_j^{(i)})$ sont indépendants.

D'outre part, $E\left(U_m^{(i)} U_r^{(i)}\right) = 0$ si $m \neq r$, d'où

$$cov\left(U_m^{(i)}U_r^{(i)}\right) = E\left(U_m^{(i)}\right)E\left(U_r^{(i)}\right) = p_m p_r.$$

On en déduit, $cov\left(X_m, X_r\right) = kp_m p_r$

$$cov(X) = k\begin{pmatrix} p_1(1-p_1) & \cdots & p_1 p_d \\ \vdots & \ddots & \vdots \\ -p_k p_d & \cdots & p_d(1-p_d) \end{pmatrix}.$$

2.2.2 Fonction caractéristique du vecteur multinomial.

Définition 2.2.3 *Soit X un vecteur multinomial $\mathcal{M}(n, p_1, ..., p_d), n \in \mathbb{N}$. La fonction caractéristique du vecteur X est définie pour tout $t \in \mathbb{R}^d$ par*

$$\varphi_X(t_1, ..., t_d) = \left(\sum_{j=1}^{d} p_j \exp(it_j p_j)\right)^n.$$

Chapitre 3

Convergence

La convergence de variables aléatoires est une notion essentielle en probabilités et mêne à certains résultats fondamentaux. À titre d'exemple introductif, notons que lors d'un jeu de pile ou face où on lance une pièce dont la fréquence d'apparition de pile est p, la fréquence observée du nombre de pile obtenu après n lancers est de p, pourvu que n . Donc, si p est inconnue, cette observation offre un moyen d'approximer p en comptant les fréquences de pile pour un grand nombre de lancers. En probabilité, plusieurs modes de convergence sont possibles, ils sont introduits dans ce chapitre, dans lequel les suites de variables aléatoires sont supposées construites sur un espace de probabilité $(\Omega, \mathcal{F}, P)$. Pour simplifier on ne considère que des variables aléatoires réelles, mais les énoncés et les résultats restent vrais pour des vecteurs aléatoires à valeurs dans $\mathbb{R}^d$.

3.1 Convergence en probabilité et presque sûre

a) Une suite $(X_n)_{n\in\mathbb{N}}$ de v.a converge presque sûrement vers une v.a X s'il existe un ensemble $C \in F$ de probabilité 1 ($P(C) = 1$) sur le quel la suite converge ponctuellement (ou simplement) vers X. On note

$$X_n \xrightarrow{P_{p.s}} X.$$

b) Une suite $(X_n)_{n\in\mathbb{N}}$ de v.a converge en probabilité vers une v.a X si, pour tout $\varepsilon > 0$, la suite de terme générale $P(|X_n - X| > \varepsilon)$ converge vers 0. On note

$$X_n \xrightarrow{P} X.$$

c) Une suite $(X_n)_{n\in\mathbb{N}}$ de v.a converge presque sûrement (resp. en probabilité) s'il existe une v.a X telle que cette suite converge presque sûrement (resp. en probabilité) vers X.

Remarque 3.1.1

1. Si la suite $(X_n)_{n\in\mathbb{N}}$ de v.a converge presque sûrement (resp. en probabilité),la limite X est $P_{p.s}$ unique. C'est clair pour la première convergence. Pout la seconde, soient X et $X^{'}$ deux limites en probabilité, pour tout $\varepsilon > 0$ et tout $n \in \mathbb{N}$, on a, par l'inégalité triangulaire,

$$\left(\left|X - X^{'}\right| > \varepsilon\right) \subset \left(|X - X_n| > \frac{\varepsilon}{2}\right) \bigcup \left(|X_n - X'| > \frac{\varepsilon}{2}\right).$$

Et par conséquent

$$P\left(\left|X - X^{'}\right| > \varepsilon\right) \leq P\left(|X - X_n| > \frac{\varepsilon}{2}\right) + P\left(|X_n - X'| > \frac{\varepsilon}{2}\right).$$

En passant à la limite, il vient

$$\forall \varepsilon > 0 \qquad P(\left|X - X^{'}\right| > \varepsilon) = 0.$$

On a le résultat en remarquant que

$$(X \neq X') = \bigcup_{n \in \mathbb{N}^*} \left(\left| X - X' \right| > \frac{1}{n} \right).$$

2. La convergence en probabilité $X_n \xrightarrow{P} X$ s'écrit de manière quantifiée :

$$\forall \varepsilon > 0,\ \forall \delta > 0 \qquad \exists N(\varepsilon, \delta) \in \mathbb{N} \quad \text{tel que } n \geq N(\varepsilon, \delta) \Rightarrow P(|X_n - X| > \varepsilon) \leq \delta.$$

Ceci est équivalent à l'assertion

$$\forall \varepsilon > 0 \qquad \exists N(\varepsilon) \in \mathbb{N} \quad \text{tel que } n \geq N(\varepsilon) \Rightarrow P(|X_n - X| > \varepsilon) \leq \varepsilon.$$

Il est trivial que la première assertion implique la seconde. Réciproquement, supposons vraie la seconde assertion et soit $\varepsilon > 0$ et $\delta > 0$. Si $\delta \geq \varepsilon$, il suffit de prendre $N(\varepsilon, \delta) = N(\varepsilon)$, si $\delta < \varepsilon$, on prend $N(\varepsilon, \delta) = N(\delta)$. On a alors, pour tout $n \geq N(\delta)$,

$$P(|X_n - X| > \delta) \leq \delta,$$

et le résultat vient de l'inégalité

$$(|X_n - X| > \varepsilon) \subset (|X_n - X| > \delta).$$

3. Si l'espace des valeurs prises par les v.a est $\mathbb{R}^d$ avec $d \geq 2$, le choix de la norme est indifférent (on le voit facilement en exprimant que toutes les normes sont équivalentes). De plus, pour que $X_n \xrightarrow{P} X$ il faut et il suffit que, pour tout $j = 1, 2, ..., d$, on ait $X_n^j \xrightarrow{P} X^j$, où X_n^j désigne la $j^{éme}$ composante de X_n. La condition nécessaire est triviale, la condition suffisante résulte des

inégalités suivantes (on choisit la norme max)

$$P\left(\max_{1\leq j\leq d}\left|X_n^j - X^j\right| > \varepsilon\right) \leq P\left(\bigcup_{j=1}^{d}\left(\left|X_n^j - X^j\right| > \varepsilon\right)\right) \leq \left(\sum_{j=1}^{d} P\left(\left|X_n^j - X^j\right| > \varepsilon\right)\right).$$

Les conditions de convergence $P_{p.s}$ données ci-dessous sont d'un usage fréquent.

Lemme 3.1.1 *1. Si la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement, elle converge en probabilité et les limites sont $P_{p.s}$ égales.*

2. Si la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X, il existe une sous-suite $(X_{n_j})_{j\in\mathbb{N}}$ qui converge presque sûrement vers X.

3. Pour que la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X il faut et il suffit qu'elle soit de Cauchy pour la convergence en probabilité, c'est à dire que, pour tout $\varepsilon > 0$, la suite double de terme général $P(|X_n - X| > \varepsilon)$ converge vers 0.

Proposition 3.1.1 *Si la série de terme général $P[|X_n - X| > \varepsilon]$ converge, alors (X_n) tend vers X presque sûrement.*

3.2 Convergence L_p et équi-intégrabilité

Si X est une v.a intégrable, le théorème de convergence dominée implique

$$\lim_{a \longrightarrow +\infty} \int_{(|X_i|>a)} |X| \, dP = 0.$$

La notion d'équi-intégrabilité généralise cette propriété à une famille quelconque de v.a en lui donnant un caractère uniforme.

Définition 3.2.1 *La famille de v.a $(X_i)_{i \in I}$, où I est un ensemble quelconque, est équi-intégrabilité si*

$$\lim_{a \longrightarrow +\infty} \sup \int_{(|X_i|>a)} |X_i| \, dP = 0.$$

On donne une condition suffisante d'équi-intégrabilité.

Définition 3.2.2 *Soit $p \geq 1$. Une suite $(X_n)_{n \in \mathbb{N}}$ de v.a admettant un moment d'ordre p converge dans L_p vers une v.a X si*

$$X \in L_p(\Omega, \mathcal{F}, P),$$

et si

$$\lim_{n \longrightarrow +\infty} E\left[|X_n - X|^p\right] = 0.$$

On note

$$X_n \xrightarrow{L_p} X.$$

La suite $(X_n)_{n \in \mathbb{N}}$ de v.a converge dans L_p s'il existe une variable aléatoire $X \in L_p(\Omega, \mathcal{F}, P)$ telle que cette suite converge dans L_p vars X.

Remarque 3.2.1 *Si $p = 1$(resp. $p = 2$) on dit que la suite converge en moyenne (resp. en moyenne quadratique). Si $p \geq 1$, comme il résulte de l'inégalité de Minkowski, l'application $X \longrightarrow [E|X|^p]^{1/p}$ est une semi norme sur L_p.En particulier, si une suite $(X_n)_{n\in\mathbb{N}}$ de v.a converge dans L_p, sa limite est $P_{p.s}$ unique. L'ensemble quotient de $L_p(\Omega, \mathcal{F}, P)$ c'est en vertu de l'inégalité de Minkowski, un espace vectoriel normé, dont la norme est obtenue par passage au quotient de la semi-norme $X \longrightarrow [E|X|^p]^{1/p}$ (en parle de la norme p de X et on note usuellement $\|X\|_p = [E|X|^p]^{1/p}$), dans cet espace de classes de v.a, la limite d'une suite est alors unique. Il est d'usage de noter la mâme façon une v.a et sa classe, on en fera de même pour la semi-norme et la norme associée et on parlera indifféremment de convergence L_p.*

Théorème 3.2.1 *Soient $p \geq 1$ et $(X_n)_{n\in\mathbb{N}}$ une suite de v.a admettant un moment d'ordre p. Les assertion suivantes sont équivalentes*

1. *La suite $(X_n)_{n\in\mathbb{N}}$ converge dans L_p.*
2. *La suite $(X_n)_{n\in\mathbb{N}}$ est de Cauchy dans L_p, c'est à dire que*

$$\lim E|X_n - X|^p = 0.$$

3. *La suite $(|X_n|^p)_{n\in\mathbb{N}}$ est équi-intégrable et il existe $X \in L^p(\Omega, F, P)$ telle que*

$$X_n \xrightarrow{P} X.$$

Convergence en loi

Définition 3.2.3 *Une suite de v.a ou de vecteur aléatoire $(X_n)_{n\in\mathbb{N}}$ de dimension d définis sur (Ω, F, P) converge en loi vers la variable ou le vecteur aléatoire X*

quand n tend vers $+\infty$ (on notes $X_n \xrightarrow{L} X$), si l'une des deux propositions équivalentes suivantes est vérifiée

1. *Pour toute fonction bornée et continue h de $\mathbb{R}^d$ dans $\mathbb{R}$*

$$E(h(X_n))_{n \longrightarrow +\infty} = E(h(X)).$$

2. *la limite ponctuelle de la fonction de répartition F_{X_n} quand n tend vers l'infini est F_X, en tout point ou' F_X est continue.*

2.a *Si $(X_n)_{n\in\mathbb{N}}$ est une suite discrète et X une v.a discrète*

$$(X_n \xrightarrow{L} X) \iff (\forall x_k \quad \lim P_{X_n}(x_k) = P_X(x_k)).$$

2.b *Si les lois des X_n et X sont définies par des densités f_{X_n} et f_X la caractérisation précédent (2) équivaut à la convergence ponctuelle de la suite (f_{X_n}) vers f_X.*

Théorème 3.2.2 (de Lévy) *soient une suite de v.a ou de vecteurs aléatoires $(X_n)_{n\in\mathbb{N}}$ et la suite de leurs fonction caractéristique $(\varphi_{X_n})_n$*

$$(X_n \xrightarrow{L} X) \iff (\forall x, \quad \lim \varphi_{X_n}(x) = \varphi_X(x) \text{ et } \varphi_X \text{ continue en } 0.$$

Remarque 3.2.2 *Une suite de v.a discrètes peut converge en loi vers une v.a absolument continue.*

Illustration : La suite $(X_n)_{n\in\mathbb{N}}$ de v.a discrètes uniformes de lois $P(X_n = \frac{k}{n}) =$

$\dfrac{1}{n+1}$ pour $k = 0, ..., n$ converge en loi vers une v.a X uniforme sur $[0, 1]$.

$$\varphi_{X_n}(t) = \sum_{k=0}^{n} \frac{1}{n+1} \exp\left(it\frac{k}{n}\right) = \frac{1 - \exp\left(it\left(\frac{n+1}{n}\right)\right)}{(n+1)\left(1 - \exp it\frac{1}{n}\right)},$$

Continue en 0, quand n tend vers $+\infty$. D'où le résultat par application du théorème de Lévy, puisque

$$\varphi_X(t) = \frac{\exp(it) - 1}{it}.$$

Inversement : Une suite de v.a admettant des densités peut converge vers un v.a discrète. C'est le cas d'un suite de v.a $(X_n)_{n\in\mathbb{N}}$ dont densités $(f_n)_n$ forment un suite d'unités approchées définie en analyse fonctionnelle comme suite de fonction positive d'intégrales égales alunite et telle que pour tout voisinage V de 0 petit soit

$$\lim_{n\longrightarrow+\infty} \int\limits_{C_V} f_n(x)dx \text{ et égale à } 0,$$

Illustration : la suite d'unités approchées égales deux densités gaussiennes $f_{X_n}(x) = \dfrac{n}{\sqrt{2\pi}} \exp\left(-\frac{n^2x^2}{2}\right)$ de moyenne nulle et de variance égale à $\dfrac{1}{n^2}$ converge en loi vers la probabilité discrète δ_0.

Théorème 3.2.3 *Soit $(w_n)_{n\in\mathbb{N}}$ une suite de lois de probabilité sur $\mathbb{R}$. Soit F_n la fonction de répartition de w_n*

$$w_n \forall x \in \mathbb{R}, \quad F(x) = w_n\left(]-\infty, x]\right).$$

Théorème 3.2.4 *Soit w une mesure de probabilité sur $\mathbb{R}$ et F sa fonction de répartition. La suite (w_n) converge faiblement vers w si et seulement si en tout* *point x où F est continue,*

$$\lim_{n \longrightarrow \infty} F_n(x) = F(x).$$

Théorème 3.2.5 *Soient $(X_n)_{n\in\mathbb{N}}$ une suite de vecteurs aléatoires à valeurs dans $\mathbb{R}^k$ et f une fonction continue de $\mathbb{R}^k$ dans $\mathbb{R}$. Si $(X_n)_{n\in\mathbb{N}}$ converge en loi (resp. probabilité) vers X, alors $(f(X_n))_n$ converge en loi (resp. probabilité) vers $f(X)$.*

3.3 Liens entre différents types de convergence

Proposition 3.3.1 *Si (X_n) tend vers X en probabilité, alors il existe une sous-suite qui tend vers X presque sûrement.*

Preuve. Supposons que $(X_n)_{n\in\mathbb{N}}$ tend vers X en probabilité. Définissons la suite d'entier (n_k) par n_k et pour tout $k \geq 0$, n_{k+1} est le premier indice supérieur à n_k tel que

$$P\left[\left|X_{n_{k+1}} - X\right| > \frac{1}{k}\right] < \frac{1}{2^k}.$$

Fixons $\varepsilon > 0$. Il existe k_0 tel que pour tout $k > k_0$, $\frac{1}{K} < \varepsilon$. Alors

$$P\left[\left|X_{n_{k+1}} - X\right| > \varepsilon\right] < \frac{1}{2^k},$$

qui est le terme général d'une série convergente. D'où le résultat par application de la proposition (3.1.1). La convergence en probabilité est un intermédiaire entre la convergence presque sûre et la convergence en loi. ∎

Proposition 3.3.2 *Si (X_n) tend vers X en probabilité, alors (X_n) converge en loi vers la loi de X.*

Preuve. Nous le démontrons dans le cas de v.a réelles en utilisant le critère des fonctions de répartition (théorème (3.3.2)). Soit x un point où la fonction de répartition de X est continue. Nous voulons démontrer que

$$\lim_{n \longrightarrow +\infty} P[X_n \leq x] = P[X \leq x].$$

Fixons $\delta > 0$. Pour tout $\varepsilon > 0$, il existe n_0 tell que pour $n > n_0$

$$P\left[|X_n - X| > \varepsilon\right] < \delta.$$

Décomposons l'événement "$X_n \leq x$"

$$"X_n \leq x" = "X_n \leq x \text{ et } X \leq x + \varepsilon" \text{ ou } "X_n \leq x \text{ et } X > x + \varepsilon",$$

Pour tout $n > n_0$, on a donc

$$P\left[X_n \leq x\right] < P\left[X \leq x + \varepsilon\right] + \delta.$$

On procède de façon symétrique pour la minoration.

$$"X_n > x" = "X_n > x \text{ et } X > x - \varepsilon" \text{ ou } "X_n > x \text{ et } X \leq x + \varepsilon",$$

Donc

$$1 - P\left[X_n \leq x\right] < 1 - P\left[X \leq x - \varepsilon\right] - \delta,$$

Soit

$$P\left[X_n \leq x\right] > P\left[X \leq x - \varepsilon\right] - \delta.$$

Puis que la fonction de répartition de X est continue en x, nous pouvons choisi ε de sorte que

$$P[X \leq x] - 2\delta \leq P[X_n \leq x] \leq P[X \leq x] + 2\delta,$$

D'où le résultat. ■

Proposition 3.3.3 *Si (X_n) converge presque surement vers X alors (X_n) converge vers X en probabilité.*

Preuve.

$$B_n^{\varepsilon} = \text{"} \; |X_n - X| \geq \varepsilon \text{"}.$$

Si (X_n) converge vers X presque surement, alors pour tout $\varepsilon > 0$

$$P[\cap_{n_0 \geq 0} \cup_{n \geq n_0} B_n^{\varepsilon}] = 0,$$

Donc :

$$\lim_{n_0 \longrightarrow \infty} P[\cup_{n \geq n_0} B_n^{\varepsilon}] = 0,$$

D'où le résultat. ■

Proposition 3.3.4 *Si (X_n) converge vers X dans L_p alors $(X_n)_{n \in \mathbb{N}}$ converge vers X en probabilité (et donc aussi en loi vers la loi de X).*

Preuve. C'est une conséquence immédiate de l'inégalité de **Markov** :

$$P[|X_n - X| < \varepsilon] \leq \frac{E[|X_n - X|^p]}{\varepsilon^p}.$$

En particulier, de toute suite qui converge presque sûrement. On utilise une technique analogue pour démontrer que l'espace L_p, muni de la norme $\|.\|$ est un espace vectoriel normé complet. La convergence presque sûre n'entraîne pas forcément la convergence dans L_p (ne serait ce que parce que les X_n peuvent converger presque sûre sans être forcément dans L_p).Le théorème de convergence dominée fournit une condition suffisante. ■

Proposition 3.3.5 *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a dans L_p qui converge presque surement vers une v.a X. Supposons qu'il existe $Y \in L_p$ telle que pour tout $n \in \mathbb{N}$:*

$$P\left[|X_n| \leq Y\right] = 1.$$

Alors $X \in L_p$ et (X_n) converge vers X dans L_p.

3.3.1 Annexe

Les liens entre les différentes convergences sont donnés par le diagramme suivant

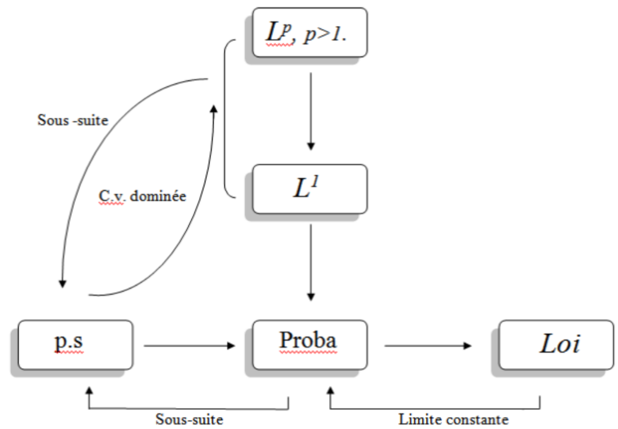

FIG. 3.1 – Liens entre différents types de convergence

3.4 Exemples de convergence en lois

Convergence de la loi hypergéométrique vers la loi binomiale

Soit X_N suivant une loi hypergéométrique $H(N, M, n)$ avec $\frac{M}{N} \longrightarrow p$, $N \longrightarrow \infty$, n restant fixe. Alors X_n tend en loi vers une v.a X suivant une binomiale $B(n, p)$.

Preuve. Pour prouve cela, utilisons les densités par rapport à la mesure de comptage v sur N. Nous avons

$$f_{X_N}(k) = \frac{\begin{pmatrix} M \\ k \end{pmatrix} \begin{pmatrix} N-M \\ n-k \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}, 0 \leq k \leq \min(n, M).$$

Supposons que $M\diagup N \longrightarrow p$, $N \longrightarrow \infty$. Nous avons

$$\begin{aligned}
f_{X_N}(k) &= \frac{M!(M-N)!n!(N-n)!}{k!(M-k)!(n-k)!(N-M-(n-k))!N!} \\
&= \frac{n!}{k!(M-k)!} \times \frac{M!}{(M-k)!} \times \frac{(N-M-(n-k))!}{(N-M-(n-k))!} \times \frac{(M-n)!}{N!} \\
&= \begin{pmatrix} n \\ k \end{pmatrix} \times \left(\frac{M!}{(M-k)}\right) \times \left(\frac{(M-n)!}{N!}\right).
\end{aligned}$$

Mais

$$\begin{aligned}
\left(\frac{M!}{(M-k)}\right) &= M(M-1)...(M-k+2)(M-k+1) \\
&= M^k(1-\frac{k-1}{M})(1-\frac{k-2}{M}) \times ... \times (1-\frac{1}{M}) \\
&= M^k(1+o(1),
\end{aligned}$$

Puis que $M \longrightarrow \infty$ et k est fixe. En suit Puis que, aussi, $N-M = N(1-M\diagup N) \rightsquigarrow N(1-p) \longrightarrow \infty$ et $n-k$ est fixé. Enfin

$$\begin{aligned}
\left(\frac{(M-n)!}{N!}\right) &= \frac{1}{(N-n+1)(N-n+2)...(N-1)N} \\
&= \frac{1}{N^n(1-\frac{n-1}{N})(1-\frac{n-2}{N})...(1-\frac{1}{N})} \\
&= \frac{1}{N^n(1=o(1)}.
\end{aligned}$$

Pour des raisons similaires. Au total, pour tout $0 \leq k \leq n$

$$f_{X_n}(k) = \binom{n}{k}\left(\frac{M}{N}\right)^k\left(\frac{N-M}{N}\right)^k(1+o(1)) \longrightarrow \binom{n}{k}p^k(1-p)^{n-k}.$$

Ainsi pour tout point k du domaine de la densité de la loi binomiale

$$f_X(k) = \binom{n}{k} p^k (1-p)^{n-k},$$

Nous avons

$$\forall (1 \leq k \leq n), f_{X_n}(k) \longrightarrow f_X(k).$$

Nous avons donc la convergence en loi cherchée. ■

Convergence de la loi binomiale vers la loi de Poisson

Soit X_n suivant une loi $\mathcal{B}(n,p)$ avec $p = p_n \longrightarrow 0$, $np_n \longrightarrow \lambda$ et $\lambda > 0$, quand $n \longrightarrow \infty$. Alors X_n tend en loi vers une v.a X suivant une loi de Poisson de paramètre λ.

Preuve. Pour attester cela, utilisons la fonction des moments. Soit X une v.a suivant la loi de Poisson de paramètre $\lambda > 0$. Nous avens $\varphi_{X_n}(t) = (p_n + (1 - p_n)\exp t)^n$ pour $n \geq 1, \varphi_X(t) = \exp(\lambda(\exp t - 1)), t \in \mathbb{R}$ Notons $\lambda_n = np_n \longrightarrow \lambda$.Pour tout t fixée,

$$\varphi_{X_n}(t) = \left(\frac{\lambda_n}{n} + (1 - \frac{\lambda_n}{n})\exp t\right)^n = \left(1 - \frac{\lambda_n(\exp t - 1)}{n}\right)^n \longrightarrow \exp\left(\lambda(\exp t - 1)\right) = \varphi_X(t),$$

Par le résultat classique d'analyse qui affirme que

$$\left(1 + \frac{x_n}{n}\right)^n \longrightarrow \exp x \qquad \text{quand } n \longrightarrow +\infty \qquad \text{pourvu que } x_n \longrightarrow x \in \mathbb{R} \qquad \text{lorsque } n \longrightarrow +\infty.$$

■

Convergence de la loi géométrique vers la loi exponentielle

Soit Y_n pour tout $n \in \mathbb{N}^*$ est une v.a de la loi géométrique de paramètre $\frac{\alpha}{n}$ où $\alpha > 0$, quand $n \longrightarrow \infty$. Alors $X_n = \frac{Y_n}{n}$ tend en loi vers une v.a X suivant une loi de exponentielle de paramètre α.

Preuve. Soit Y une v.a suivant la loi géométrique de paramètre $\alpha > 0$. Nous avons

$$\varphi_{Y_n}(t) = \frac{\alpha/n}{\exp(-it) - 1 + 1/n}.$$

On a : $X_n = \dfrac{y_n}{n}$, alors

$$\varphi_{X_n}(t) = \varphi_{Y_n}(t/n) = \frac{\alpha/n}{\exp(-it/n) - 1 + 1/n},$$

puis que

$$\exp z = 1 + z + z\varepsilon(z), \text{ avec } \lim_{n \longrightarrow 0} \varepsilon(z) = 0,$$

alors si on pose $z = \dfrac{-it}{n}$ donc

$$sin \longrightarrow \infty \Rightarrow z \longrightarrow 0, \text{ i.e } \varphi_{X_n}(t) = \frac{\alpha/n}{-it + \alpha - it\varepsilon(t/n)},$$

et

$$\lim_{n \longrightarrow \infty} \varphi_{X_n}(t) = \frac{\alpha}{\alpha - it}.$$

Est la fonction caractéristique d'un v.a X de la loi exponentielle $X \rightsquigarrow \xi(\alpha)$. ■

Bibliographie

[1] Breton, J-C. (2014). Fondements des Probabilités de $(\Omega, \mathcal{F}, P)$ aux cons equences de la LGN et du TCL. Université de Rennes 1.

[2] Gane, S LO. (2016). convergence vague - Suites de vecteur aléatoire. Saint-Louis, SENEGAL - Calgary.

[3] Kilque, C. (2018). Modes de convergence d'une suite de variables aléatoires. Exemples et applications. Université de Rennes 1.

[4] Ouvrard,J-Y. (2009). Probabilités 2 master agrégation. Gassini Paris.

[5] LABED, S. (2020). Probabilités avancées. Cours de $3^{éme}$ année Licence Mathématiques. Université de Biskra.

[6] Ycart, B, et Genon-Catalot, V. (2004). Vecteurs et suites aléatoires martingales discrètes. Centre de publication Universitaire, Tunis.