République Algérienne 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

DÉPARTEMENT DE MATHÉMATIQUES

saleh/Desktop/Nouveau dossier (4)/UMKB$_{l}ogo.wmf$

Cours présenté par Dr : Yakhlef Samia

Module : Probabilités Avancées

3ème Année

2019

Table des matières

Introduction

Chapitre 1

Variables aléatoires

Définition 1.0.1 *Soient* $(\Omega, \mathcal{A})$ *et* $(E, \mathcal{B})$ *deux espaces probabilisables. Une application* f *de* Ω *vers* E *est dite mesurable (ou* $\mathcal{A}$*-mesurable si la confusion est possible) si*

$$\forall B \in \mathcal{B}, \quad f^{-1}(B) \in \mathcal{A}.$$

On a déjà vu que $f^{-1}(\mathcal{B})$ *est une tribu. Dire que la fonction* f *est mesurable revient donc à dire que* $f^{-1}(\mathcal{B}) \in \mathcal{A}$*. Ainsi, pour tout événement* B*, l'ensemble :*

$$f^{-1}(B) = \{\omega \in \Omega / f(\omega) \in B\}$$

est un événement de la tribu initiale. On utilise parfois la notation $f^{-1}(B) = [f \in B]$ *.*

Notons en particulier que toute fonction continue est mesurable. De même, pour tout événement A de la tribu $\mathcal{A}$, la fonction 1_A est mesurable de $(\Omega, \mathcal{A})$ dans $(E, \mathcal{B})$.

Proposition 1.0.1 – *Si* f *et* g *sont deux fonctions mesurables de alors les fonctions* $f+g$ *et* fg *sont encore mesurables.*$(\Omega, \mathcal{A})$ *vers* $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ *.*

– *Si* f *et* g *sont deux fonctions mesurables de* $(\Omega, \mathcal{A})$ *vers* $(\Omega', \mathcal{A}')$ *et de* $(\Omega', \mathcal{A}')$ *vers* $(\Omega'', \mathcal{A}'')$ *respectivement, la fonction* $g \circ f$ *est mesurable de* $(\Omega, \mathcal{A})$ *vers* $(\Omega'', \mathcal{A}'')$*.*

– *Si* (f_n) $(\Omega, \mathcal{A})$ *vers* $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ *alors les fonctions*

$$\sup_n f_n, \quad \inf_n f_n, \quad \lim\sup_n f_n \ \textit{et} \ \lim\inf_n f_n$$

sont mesurables, à condition qu'elles ne prennent pas de valeurs infinies.

Définition 1.0.2 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé. Une application mesurable* X *de* $(\Omega, \mathcal{A}, P)$ *vers* $(E, \mathcal{B})$ *est appelée variable aléatoire.*

Tous les résultats sur les fonctions mesurables restent donc vrais pour les variables aléatoires.

Notation 1.0.1 *Si* $(E, \mathcal{B}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, *l'application* X *est dite variable aléatoire réelle (v.a.r.) ou unidimensionnelle ou univariée.*
Si $(E, \mathcal{B}) = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$, *l'application* X *est dite vecteur aléatoire ou variable aléatoire multidimensionnelle ou multivariée.*
Si $(E, \mathcal{B})$ *est tout, ou une partie, de* $(\mathbb{Z}, \mathcal{B}_{\mathbb{Z}})$, *l'application* X *est dite v.a.discrète.*

1.1 Loi de probabilité d'une variable aléatoire

Définition 1.1.1 *Soit* X *une variable aléatoire de* $(\Omega, \mathcal{A}, P)$ *vers* $(E, \mathcal{B})$. *Définissons une application* P_X *de* $\mathcal{B}$ *vers* $[0, 1]$ *par :*

$$\forall B \in \mathcal{B}, \ P_X(B) = P\left(X^{-1}(B)\right) = P\left(\{\omega \in \Omega / X(\omega) \in B\}\right).$$

La définition précédente a bien un sens puisque l'on a $X^{-1}(B) \in \mathcal{A}$, par mesurabilité de X. On peut donc prendre la probabilité de cet évènement

Définition 1.1.2 P_X *est appelée probabilité image de* P *par* X *ou encore loi de probabilité de la variable aléatoire* X. *On note* $P_X(B) = P(X \in B)$.

Ainsi, tout événement lié à X est connu dès que l'on connaît la loi P_X de X : On oubliera donc souvent dans la suite le détail de l'application $\omega \to X(\omega)$ et on ne se préoccupera pas de ce qu'est exactement $(\Omega, \mathcal{A}, P)$: On raisonnera uniquement sur $(E, \mathcal{B})$ et P_X.
Dans la suite on va considérer uniquement les variables aléatoires réelles.

1.1.1 Fonction de répartition

Définition 1.1.3 *On appelle fonction de répartition (f.d.r.) de la v.a.r. X, la fonction F_X définie sur $\mathbb{R}$ par*

$$F_X(x) = P_X(]-\infty, x]) = P(X \leq x).$$

Proposition 1.1.1 *La fonction de répartition F_X d'une v.a.r. X satisfait les propriétés suivantes :*

1. $0 \leq F_X(x) \leq 1$ *pour tout x dans $\mathbb{R}$;*
2. *La fonction F_X est croissante (au sens large) ;*
3. *La fonction F_X est continue à droite ;*
4. *On a*

$$\lim_{x \to -\infty} F_X(x) = 0 \ \text{et} \ \lim_{x \to +\infty} F_X(x) = 1$$

Preuve. La propriété 1 est évidente puisque la probabilité de n'importe quel événement est toujours positive et inférieure à 1.
Pour établir 2 considérons x et $x\prime$ deux réels tels que $x \leq x\prime$. On a bien sûr l'inclusion

$$]-\infty, x] \subset]-\infty, x']$$

et donc

$$P(]-\infty, x]) \leq P(]-\infty, x']).$$

Pour 3 considérons une suite (h_n) de réels décroissante vers 0. Pour tout x dans $\mathbb{R}$, on a :

$$P_X\left(\left]x, x+h_n\right]\right) = F_X\left(x+h_n\right) - F_X(x)$$

Or la suite d'intervalles $\left(\left]x, x+h_n\right]\right)_n$ est décroissante avec n. Ainsi il vient

$$\lim_{n\to+\infty} P_X\left(\left]x, x+h_n\right]\right) = P_X\left(\bigcap_n \left]x, x+h_n\right]\right) = P_X(\emptyset) = 0$$

alors

$$\lim_{n\to+\infty} F_X\left(x+h_n\right) = F_X(x)$$

et la fonction F_X est donc bien continue à droite.

Établir le 4, considérons la suite d'intervalles $\left(\left]-\infty, n\right]\right)_n$ décroissante vers $\emptyset$ quand n tend vers $+\infty$. On a :

$$\lim_{x\to-\infty} F_X(x) = \lim_{n\to+\infty} F_X(-n) = \lim_{n\to+\infty} P_X\left(\left]-\infty, -n\right]\right) = P_X\left(\bigcap_n \left]-\infty, -n\right]\right) = P_X(\emptyset) = 0$$

on a aussi

$$\lim_{x\to+\infty} F_X(x) = \lim_{n\to+\infty} F_X(n) = \lim_{n\to+\infty} P_X\left(\left]-\infty, n\right]\right) = P_X\left(\bigcup_n \left]-\infty, n\right]\right) = P_X\left(\left]-\infty, +\infty\right[\right) = 1$$

■

Théorème 1.1.1 *Toute fonction F définie sur $\mathbb{R}$ et vérifiant les propriétés 1, 2, 3 et 4 est une fonction de répartition d'une v.a.r.*

Proposition 1.1.2 *Le saut $p_0 = F_X(x_0) - F_X\left(x_0^-\right)$ de la fonction de répartition F_X au point x_0 est égal à $P(X = x_0)$*

Preuve. Soit (h_n) une suite de réels strictement positifs, décroissante vers 0. On a

$$\begin{aligned}
F_X(x_0) - F_X\left(x_0^-\right) &= F_X(x_0) - \lim_{n\to+\infty} F_X(x_0 - h_n) \\
&= \lim_{n\to+\infty} \left(F_X(x_0) - F_X(x_0 - h_n)\right) \\
&= \lim_{n\to+\infty} \left(P(X \leq x_0) - P(X \leq x_0 - h_n)\right) \\
&= \lim_{n\to+\infty} \left(P(x_0 - h_n \leq X \leq x_0)\right) \\
&= P\left(\bigcap_n X \in \left]x_0 - h_n, x_0\right[\right) \\
&= P(X = x_0)
\end{aligned}$$

■

1.2 Variables aléatoires discrètes

Soit X une v.a de Ω dans $F = X(\Omega)$.

Dans ce cas, F est dénombrable et $F = X(\Omega) = \{x_1, ..., x_i, ...\}$. Dans ce cas particulier, Ω et F sont munis des tribus $\mathcal{P}(\Omega)$ et $\mathcal{P}(F)$

Définition 1.2.1 *Toute application de Ω dans F est une Variable aléatoire discrète (limage réciproque par X d'un ensemble de F est un ensemble de Ω).*

Nous voulons caractériser la loi X

Définition 1.2.2 *La loi de X est une probabilité P_X sur F telle que pour tout B C F*

$$P_X(B) = P(X \in B) \tag{1.1}$$

Remarque 1.2.1 *Une probabilité sur un ensemble F dénombrable est caractérisée par les probabilités de ses singletons, il suffit donc de nous limiter à $B = \{x_i\}$*

Proposition 1.2.1 *La loi* $\mathbb{P}_X$ *de* X *est caractérisée par* $\left\{\left(x_i, P_i^X\right), x_i \in F\right\}$, *avec*

$$P_i^X = P\left(X = x_i\right) = P_X\left(\{x_i\}\right) \tag{1.2}$$

Remarque 1.2.2 *La loi* P_X *de* X *est caractérisée par (1.2) des que* $F = X\left(\Omega\right)$ *est dénombrable, même si* Ω *ne l'est pas, on a* $\forall B \in P\left(F\right)$

$$\begin{aligned} P_X\left(B\right) &= P_X\left(\bigcup_{i, x_i \in B} \{x_i\}\right) \\ &= \sum_{i, x_i \in B} P_X\left(\{x_i\}\right) \\ &= \sum_{i, x_i \in B} P\left(X = x_i\right) = P\left(\bigcup_{i, x_i \in B} \{X = x_i\}\right) = P\left(X \in B\right). \end{aligned}$$

1.2.1 Variables aléatoires discrètes usuelles

★ $\Omega = \{\text{pile}, \text{face}\}$, $p \in \left]0, 1\right[$. On pose :$X = 1$ si pile, $X = 0$ si face

$$P_X\left(1\right) = P\left(X = 1\right) = p,\ P_X\left(0\right) = P\left(X = 0\right) = 1 - p$$

La loi de X est appelée **loi de Bernoulli** de paramètre P.

★ Infinité de lancers indépendants : loi du premier succès

$$\mathbb{P}(\text{premier pile au lancer } k)$$

$$= \mathbb{P}\left(FF...FP\right) = P\left(1 - P\right)^{k-1}$$

Définition 1.2.3 *Une variable* ***géométrique*** *de paramètre* $P \in \left]0, 1\right[$ *est une variable a valeurs dans* $\mathbb{N}^*$ *telle que* $\forall k \in \mathbb{N}^*$,

$$P\left(X=k\right)=P\left(1-P\right)^{k-1}$$

Représentation graphique d'une lois géométriques : Nous représentons sur un graphique les couples $\left(k, P\left(1-p\right)^{k-1}\right)$ pour $k \in \mathbb{N}^*$ et différentes valeurs de p.

– **Variable aléatoire Binomiale de paramètres** n **et** P : Soit S_n nombre de piles obtenus après n lancers indépendants

$$S_n=\sum_{i=1}^{n} X_i \ ; \ S_n\left(\Omega\right)=\{0,...,n\}$$

$$P\left(S_n=k\right)=P\left(\left\{\omega\in\Omega \ / \ \sum_{i=1}^{n} X_i\left(\omega\right)=k\right\}\right)=C_n^k P^k\left(1-p\right)^{n-k}.$$

Définition 1.2.4 *La loi de S_n est appelée loi **Binomiale de paramètre** $n \in \mathbb{N}^*$ et $p\in\left]0,1\right[$ que l'on note $B\left(n,P\right)$*

Exemple 1.2.1 *Pour $n=4$, on à la v.a S_4 suit la loi $B\left(n,P\right)$ et donc on a*

$$P\left(S_4=k\right)=C_4^k\left(\frac{1}{4}\right)^k\left(\frac{3}{4}\right)^{4-k}, \forall k\in\{0,1,2,3,4\}$$

$$P\left(S_4=0\right)=\left(\frac{3}{4}\right)^4 \ , \ P\left(S_4\rangle 1\right)=4\frac{1}{4}\left(\frac{3}{4}\right)^3=\left(\frac{3}{4}\right)^3$$

$$P\left(S_4=2\right)=C_4^2 p^2\left(1-p\right)^2=\frac{4!}{2!2!}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^2=\frac{6}{4^2}\times\frac{3^2}{4^2}$$

– **Variable aléatoire de poisson :** Modélisation des réalisations d'un événement rare. Soit X une variable aléatoire de loi de Binomiale $B\left(n,a_n\right)$ avec $na_n\to\theta\in\mathbb{R}_+^*$ quand $n\to+\infty$

$$P_j\left(a_n,n\right)=:P_X\left(j\right)=\begin{cases} C_n^j\left(a_n\right)^j\left(1-a_n\right)^{n-j} \text{ si } j\leq n \\ 0 \text{ si } j\geq n+1\end{cases}$$

Alors on a : $\lim\limits_{n\to+\infty} P_j(a_n,n) = P_j := e^{-\theta}\frac{\theta^j}{j!}$,$\forall j\in\mathbb{N}$. En effet, on a

$$\begin{aligned}
P_j(a_n,n) &= C_n^j(a_n)^j(1-a_n)^{n-j} \\
&= \frac{n!}{j!\,(n-j)!}(a_n)^j(1-a_n)^{n-j} \\
&= \frac{1}{j!}\underbrace{na_n}_{\substack{\downarrow\\ \theta}}\underbrace{(n-1)\,a_n}_{\substack{\downarrow\\ \theta}}...\underbrace{(n-j+1)a_n}_{\substack{\downarrow\\ \theta}}(1-a_n)^{n-j} \\
&= \frac{1}{j!}\theta^j e^{-\theta},
\end{aligned}$$

on a utiliser

$$(1-a_n)^{n-j} = e^{(n-j)len(1-a_n)} \to \underset{n\to+\infty}{e^{-\theta}}$$

car

$$a_n \underset{n\to+\infty}{\sim} \frac{\theta}{n} \Longrightarrow \ln(1-a_n) \underset{n\to+\infty}{\sim} -a_n.$$

Définition 1.2.5 *On appelle v.a de* ***Poisson*** *de paramètre $\theta>0$ une v.a X a valeurs dans $\mathbb{N}$ telle que*

$$P(X=k) = e^{-\theta}\ \frac{\theta^k}{k!}\ ,\forall k\in\mathbb{N},$$

sa loi est la **loi de Poisson** de paramètre θ notée $\mathcal{P}(\theta)$

Exemple 1.2.2 *En peut prendre comme exemple l'exemple précédent***,** *on peut approcher X (en loi) par une v.a Z de loi de Poisson $\mathcal{P}(2)$ on peut Vérifier que l'approximation est très bonne, soit X une variable aléatoire qui suit la loi Binomiale $B(100,0,02)$ et donc*

$np = 2$, et soit Z une variable aléatoire de loi de Poisson $\mathcal{P}(2)$, on a alors

k	$P(X=k)$	$P(Z=k)$
0	$0,1326$	$0,13533$
1	$0,2706$	$0,27066$
2	$0,2734$	$0,27066$
3	0.1822	$0,1804$
4	$0,09$	$0,0902$

1.2.2 Espérance pour une variable aléatoire sur Ω fini ou dénombrable

Motivation : donner un résumé quantitatif de variable aléatoire X.

- **Moyenne arithmétique** $\dfrac{x_1 + ... + x_n}{n}$
- **Modèle abstrait**

$$M_n = \sum_{\omega \in \Omega} f_n(\{\omega\}) X(\omega) \quad , ou \quad f_n \quad \text{est la fréquence empirique,}$$

On a

$$f_n(\{\omega\}) \underset{n \to +\infty}{\to} P(X = \omega),$$

cette limite sera justifier plus tard (Loi des grands nombres).

Définition 1.2.6 *Si*

$$\sum_{\omega \in \Omega} P(\omega) |X(\omega)| < +\infty$$

On appelle espérance de X le nombre

$$E(X) := \sum_{\omega \in \Omega} X(\omega) P(\omega).$$

On dit que la v.a X est intégrable si

$$E(|X|) =: \sum_{\omega\in\Omega} P(\omega)\,|X(\omega)|\, \langle +\infty$$

Si Ω est fini toutes les v.a sont intégrables.

Soit $\mathbb{L}^1$ l'ensemble des variables aléatoires intégrables. C'est un sous espace vectoriel de l'ensemble des applications de Ω dans F. En effet, il est non vide et stable par combinaison linéaire

Proposition 1.2.2 *1. Si $X(\omega) = a$ pour tout ω , alors*

$$E(X) = a\sum_{\omega\in\Omega} P(\omega) = a$$

2. ***Linéarité** : On a*

$$E(aX + bY) = aE(X) + bE(Y)$$

3. *On a $X \in L^1 \Longleftrightarrow |X| \in L^1$ et on a l'inégalité suivante*

$$|E(X)| \leq E(|X|)\,.$$

4. ***Positivité :** Si $X \geq 0$, alors $E(X) \geq 0$. $X \geq 0$ signifie que pour tout $\omega \in \Omega, X(\omega) \geq 0$.*

5. *Si $X, Y \in L^1$ telles que $X \leq Y$, alors $E(X) \leq E(Y)$.*

La remarque suivante fait le lien entre la probabilité d'un événement et l'espérance de l'indicatrice de cet événement

Preuve. Pour la preuve de 2

$$\begin{aligned} E(aX+bY) &= \sum_{\omega\in\Omega}(aX(\omega)+bY(\omega))P(\omega) \\ &= a\sum_{\omega\in\Omega}X(\omega)P(\omega)+b\sum_{\omega\in\Omega}Y(\omega)P(\omega) \\ &= aE(X)+bE(Y) \end{aligned}$$

Pour 4 on a si $X \leq Y$, alors $E(X) \leq E(Y)$, c'est une conséquence de la positivité et de la linéarité car

$$0 \leq E(Y-X) = E(Y) - E(X).$$

■

Remarque 1.2.3 (fondamentale) *Soit $A \subset \Omega$, rappelons que l'indicatrice de l'ensemble A est la v.a $X = 1_A$ définie par*

$$X(\omega) = \begin{cases} 1 \ si\ \omega \in A \\ 0 \ si\ \omega \notin A \end{cases}$$

C'est donc une v.a de Bernoulli et nous avons

$$E(X) = E(1_A) = P(X=1) = P(A).$$

Calcul de l'espérance

Si $E(|X|) < +\infty$, alors

$$E(X) = \sum_{x_i\in F} x_i P(X=x_i). \tag{1.3}$$

Plus généralement :

Théorème 1.2.1 *Soit $f : F \to \mathbb{R}$, telle que $E(|f(X)|) < +\infty$, alors*

$$E(f(X)) = \sum_{x_i \in F} f(x_i) \mathbb{P}(X = x_i). \tag{1.4}$$

Preuve. Supposons que $E(f(|X|)) < +\infty$ ($i.e$ la série de terme générale $P(\omega) f(X(\omega))$ est absolument convergente)

$$E(f(X)) = \sum_{\omega \in \Omega} P(\omega) \;\; f(X(\omega))$$

$$\Omega = U_{i, x_i \in F} \{\omega, X(\omega) = x_i\}$$

on va faire une somation par pacquer pour cela on a besoin de la c.v absolue de la série on a donc

$$\begin{aligned}
E(f(X)) &= \sum_{x_i \in F} \left(\sum_{\omega, X(\omega) = x_i} P(\omega) \quad f(X(\omega)) \right) \\
&= \sum_{x_i \in F} \left(\sum_{\omega, X(\omega) = x_i} P(\omega) \;\; f(x_i) \right) \\
&= \sum_{x_i \in F} \left(f(x_i) \left(\sum_{\omega, X(\omega) = x_i} P(\omega) \right) \right) \\
&= \sum_{x_i \in F} f(x_i) P \left(\bigcup_{\omega, X(\omega) = x_i} \{\omega\} \right) \\
&= \sum_{x_i \in F} f(x_i) P(X = x_i).
\end{aligned}$$

■

Exemple 1.2.3 *$X \sim P(\theta)$, $\theta \rangle 0$ (loi de Poisson), $\forall k \in \mathbb{N}$, $\mathbb{P}(X = k) = e^{-\theta} \dfrac{\theta^k}{k!}$*

$$\begin{aligned} E(X) &= \sum_{k \in \mathbb{N}} k P(X = k) \\ &= \sum_{k \in \mathbb{N}^*} k e^{-\theta} \frac{\theta^k}{k!} = e^{-\theta} \theta \sum_{k \in \mathbb{N}^*} \frac{\theta^{k-1}}{(k-1)!} \\ &= e^{-\theta} \theta \sum_{\theta \geq 0} \frac{\theta^l}{l!} = e^{-\theta} \theta e^{\theta} = \theta \end{aligned}$$

$$\begin{aligned} E(X(X-1)) &= \sum_{k \in \mathbb{N}} k(k-1) e^{-\theta} \frac{\theta^k}{k!} \\ &= \sum_{k \geq 2} k(k-1) e^{-\theta} \frac{\theta^k}{k!} \\ &= \left(\sum_{k \geq 2} \frac{\theta^{k-2}}{(k-2)!} \right) e^{-\theta} \theta^2 = \left(\sum_{l \geq 0} \frac{e^l}{l!} \right) e^{-\theta} \theta^2 \\ &= e^{\theta} e^{-\theta} \theta^2 = \theta^2 \end{aligned}$$

Variance

Définition 1.2.7 *Nous dirons qu'une v.a X est de carré intégrable si $E(X^2) < \infty$. L'ensemble des v.a de carré intégrable est noté $\mathbb{L}^2$.*

Supposons que $E(X^2) < \infty$, on appelle ***Variance*** *de X le nombre positif*

$$Var(X) = E\left((X - E(X))^2 \right).$$

on a

$$\begin{aligned} Var(X) &= E\left((X - E(X))^2\right) \\ &= E\left(X^2 - 2XE(X) + (E(X))^2\right) \\ &= \sum_{x_i \in F} P(X = x_i)\, x_i^2 - (E(X))^2 . \end{aligned}$$

$Var(X)$ =“Moyenne des carré-carré des moyennes”. La variance mesure le carré d'une distance a la moyenne

Ecart-Type de X :

$$\sigma(X) = \sqrt{Var(X)}$$

Propriété

Pour tout nombres réels a et b et X une v.a de carré intégrable, il en est de même de $aX + b$ et

$$Var(aX + b) = a^2 Var(X)$$

Définition 1.2.8 *Nous dirons qu'une v.a de carré intégrable X est centrée réduite si* $E(X) = 0$; $Var(X) = 1$

Proposition 1.2.3 *Soit Y une v.a de carrée intégrable. Alors, la variable aléatoire X définie par*

$$X = \frac{Y - E(Y)}{\sigma_Y},$$

est centrée réduite.

Preuve. On a $Var\left(aX+b\right)=a^{2}Var\left(X\right)$

$$
\begin{aligned}
Var\left(aX+b\right) &= E\left(\left(aX+b\right)^{2}\right)-\left(E\left(aX+b\right)\right)^{2} \\
&= E\left(a^{2}X^{2}2abX+b^{2}\right)-\left(aE\left(X\right)+b^{2}\right) \\
&= a^{2}E\left(X^{2}\right)+2abE\left(X\right)+b^{2}-a^{2}\left(E\left(X\right)\right)^{2}-b^{2}-2abE\left(X\right) \\
&= a^{2}\left(E\left(X^{2}\right)-\left(E\left(X\right)\right)^{2}\right) \\
&= a^{2}Var\left(X\right).
\end{aligned}
$$

On a si

$$X=\frac{Y-E\left(Y\right)}{\sigma_{Y}}$$

alors

$$E\left(X\right)=\frac{1}{\sigma_{Y}}\left(E\left(Y\right)-E\left(Y\right)\right)=0$$

et

$$
\begin{aligned}
Var\left(X\right) &= Var\left(\frac{Y-E\left(Y\right)}{\sigma_{Y}}\right) \\
&= \left(\frac{1}{\sigma_{Y}}\right)^{2}Var\left(Y-E\left(Y\right)\right) \\
&= \left(\frac{1}{\sigma_{Y}}\right)^{2}Var\left(Y\right)=\frac{\sigma_{Y}^{2}}{\sigma_{Y}^{2}}=1.
\end{aligned}
$$

■

1.2.3 Fonction génératrice d'une variable aléatoire à valeurs entières

Considérons une variable aléatoire X à valeurs dans $\mathbb{N}$. Jusqu'à présent, nous avons caractérisé la loi d'une telle variable par la donnée (n,P_{n}), où $n\in\mathbb{N}$ et $P_{n}=P\left(X=n\right).$ Nous allons introduire une fonction réelle de variable réelle qui va nous permettre de caractériser simplement la loi de X et d'obtenir ses moments quand ils existent.

Définition 1.2.9 *La fonction génératrice de X est la fonction définie pour $s \in [0,1]$ par*

$$G_X(s) = E\left(s^X\right) = \sum_{n \in \mathbb{N}} P(X = n)\, s^n.$$

C'est la somme d'une série entière dont le rayon de convergence est au moins 1, car $\sum_n P_n = 1$

Nous allons voire que cette fonction caractérise la loi de X.

Fonction génératrice et moments

Proposition 1.2.4 *1. G_X est continue sur $[0,1]$, C^∞ sur $[0,1[$.*

2. $X \in \mathbb{L}^1 \iff G_X$ dérivable (à gauche) en $s = 1$ et $E(X) = G'_X(1)$

3. $X(X-1)\ldots(X-P) \in \mathbb{L}^1 \iff G_X$ est $P+1$ fois dérivable en 1 ,et

$$E\left(X(X-1)\ldots(X-P)\right) = G_X^{(P+1)}(1).$$

Preuve. on a :

1. $G_X^{(s)} = \sum_{n \in \mathbb{N}} S^n P_n$ Somme d'une série entière de rayon de convergence $R \geq 1$, alors (d'après les cours d'analyse) on à
2. G_X est continue sur $[0,1]$ et C^∞ sur $([0,1[)$.

   Montrons que $G_X^{(n)}(0) = P_n\ n! \Longrightarrow P_n = P(X = n) = \frac{G_X^{(n)}(0)}{n!}$.

   Supposons que $X \in \mathbb{L}^1$

$$G_X(s) = E\left(s^X\right) = \sum_n \left(P_n S^n\right)$$

on a donc $\forall s \in [0,1[$

$$\begin{aligned} G'_X(s) &= \sum_n nP_n s^{n-1} \\ &= E\left(Xs^{X-1}\right) \end{aligned}$$

étudions la dérivabilité au point $s=1$

$$\begin{aligned} \frac{G_X(s)-G_X(1)}{s-1} &= \frac{\sum P_n s^n - \sum_n P_n}{s-1} \\ &= \sum_n P_n\left(\frac{s^n-1}{s-1}\right) \\ &= \sum_n P_n\left(1+s+s^2+...+s^{n-1}\right), \end{aligned}$$

puisque on a

$$s^{n-1} = (s-1)\left(1+s+s^2+...+s^{n-1}\right),$$

et comme

$$P_n\left(1+s+s^2+...+s^{n-1}\right) \to nP_n \text{ quand } s \to 1$$

alors

$$\begin{aligned} \lim_{s\to 1}\frac{G_X(s)-G_X(1)}{s-1} &= \sum_n \lim_{s\to 1}\left(P_n\left(1+s+s^2+...+s^{n-1}\right)\right) \\ &= \sum_n nP_n = E(X) \end{aligned}$$

et donc $G'_X(1) = E(X)$.

3. On a pour tout $s \in [0,1[$

$$G_X^{(P+1)}(s) = \left(\sum_n P_n s^n\right)^{(p+1)}$$
$$= \sum_n P_n n(n-1)(n-2)\ldots(n-p)\,S^{n-p+1}?$$

par passage a la limite quand $s \to 1$

$$G_X^{(P+1)}(1) = E\left(X(X-1)\ldots(X-P+1)(X-p)\right).$$

4. En particulier, X est de carré intégrable si et seulement si G_X est deux fois dérivable en 1 et

$$E\left(X(X-1)\right) = G_X''(1)$$

Nous en déduisons que

$$Var(X) = E\left(X(X-1)\right) + E(X) - E(X)^2$$
$$= G_X''(1) + G_X'(1) - \left(G_X'(1)\right)^2.$$

■

Pour calculer les moments d'une variable aléatoire à valeurs entières, il est souvent beaucoup plus rapide d'utiliser les dérivées de la fonction génératrice plutôt qu'un calcul direct

Calcul des moments

- X v.a de Bernoulli on a pour $s \in [0,1]$, $X(\Omega) = \{0,1\}$

$$G_X(s) = \sum_{k \in X(\Omega)} P(X=k)\,s^k$$
$$= (1-p) + sp$$

alors

$$G'_X(1) = p$$

et

$$\begin{aligned} Var(X) &= G''_X(1) + G'_X(1) - (G'_X(1))^2 \\ &= p - p^2 \end{aligned}$$

et donc

$$E(X) = p, Var(X) = p(1-p)$$

- $X \sim B(n,p)$, pour $s \in [0,1]$,

$$\begin{aligned} G_X(s) &= \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} s^k \\ &= (ps + (1-p))^n . \end{aligned}$$

Nous en déduisons

$$E(X) = G'_X(1) = p.n,$$

$$\begin{aligned} Var(X) &= G''_X(1) + G'_X(1) - (G'_X(1))^2 \\ &= n(n-1)p^2 + pn - p^2n^2 \\ &= np((n-1)p + 1 - np) = np(1-p). \end{aligned}$$

- $X \sim \mathcal{P}(\theta)$, pour $s \in [0,1]$,

$$G_X(s) = \sum_{k\in\mathbb{N}} s^k \frac{\theta^k}{k!} e^{-\theta} = e^{s\theta} e^{-\theta} = e^{\theta(s-1)},$$

alors

$$G'_X(s) = \theta e^{\theta(s-1),}$$

et

$$G''_X(s) = \theta^2 e^{\theta(s-1).}$$

Nous en déduisons

$$E(X) = \theta = G'_X(1)$$

$$\begin{aligned} Var(X) &= G''_X(1) + G'_X(1) - (G'_X(1))^2 \\ &= \theta^2 + \theta - \theta^2 = \theta. \end{aligned}$$

1.2.4 Couple de variables aléatoires et variables aléatoires indépendants cas discret

On veut décrire l'évolution aléatoire conjointe de deux variables aléatoires X et Y

Exemple 1.2.4 *$Z = (X, Y)$ décrit le nombre d'années d'études et le nombre de frères et sœurs de l'ainé d'une famille*

$$X(\Omega) = F; Y(\Omega) = G \text{ et } Z(\Omega) = F \times G \text{ est dénombrable.}$$

Définition 1.2.10 *La loi jointe du couple (X, Y) est une probabilité $\mathbb{P}_{(X,Y)}$ sur $F \times G$ caractérisée par la probabilité des singletons. Pour tout $(x, y) \in F \times G$,*

$$\mathbb{P}_{(X,Y)}(\{(x, y)\}) = P(X = x, Y = y)$$

Définition 1.2.11 *La lois marginales Ce sont les lois respectives P_X et P_Y des coordon-*

nées X et Y.

Remarque 1.2.4 *remarquons que*

$$\{X = x\} = U_{Y \in G} \{X = x, Y = y\}$$

et les ensembles $\{X = x\}, \{Y = y\}$ sont disjoints. Nous en déduisons que

$$\begin{aligned} P_X(x) = P(X = x) &= \sum_{Y \in G} P(X = x, Y = y) \\ &= \sum_{Y \in G} P_{(X,Y)}(x, y) \end{aligned}$$

et de même

$$P_Y(y) = P(Y = y) = \sum_{X \in F} P_{(X,Y)}(x, y).$$

1.2.5 Loi conditionnelles

Soit X une v.a discrète à valeurs dans E (dénombrable ou finie) telle que,

$$\forall x \in E : P(X = x) > 0.$$

Définition 1.2.12 *On définie la probabilité conditionnelle sachant $\{X = x\}$ par*

$$P(\cdot/X = x) = \frac{P(\cdot \cap \{X = x\})}{P(X = x)}.$$

Sur l'espace probabilisé $(\Omega, \mathcal{A}, P(\cdot/X = x))$ on peut calculer l'espérance et loi d'une v.a on parle alors d'espérance et de loi conditionnelle sachant $\{X = x\}$.

Définition 1.2.13 *Soit Z une v.a.r positive ou intégrable.*

- *On appelle espérance conditionnelle de Z sachant $\{X = x\}$, l'espérance de Z par rapport à la probabilité $P(\cdot/X = x)$ et on note cette quantité $E(Z/X = x)$,*

– *Autrement dit :*

$$E\left(Z/X=x\right)=\int_{\Omega}ZdP\left(\cdot/X=x\right).$$

– *Notons* Ψ *l'application*

$$\begin{array}{rll}\Psi: & E\rightarrow & \mathbb{R}\\ & x\rightarrow & \Psi\left(x\right)=E\left(Z/X=x\right)\end{array}$$

La v.a.r $\Psi\left(X\right)$, *notée* $E\left(Z/X\right)$, *s'appelle l'espérance conditionnelle de* Z *sachant* X.

Proposition 1.2.5 *Si* Z *est une v.a.r positive ou intégrable, alors* $\forall x\in E$

$$\begin{aligned}E\left(Z/X=x\right)&=\int_{\Omega}ZdP\left(\cdot/X=x\right)\\&=\frac{1}{P\left(X=x\right)}\int_{\{X=x\}}ZdP\\&=\frac{1}{P\left(X=x\right)}E\left(Z\ \mathbf{1}_{\{X=x\}}\right).\end{aligned}$$

Preuve. La formule est vraie si Z est une fonction indicatrice de borélien par linéarité, elle est vraie pour les combinaisons linéaires de fonctions indicatrices de boréliens. Par passage à la limite croissante, elle est vraie pour les v.a positive ■

Proposition 1.2.6 *Si* Z *est une v.a.r positive ou intégrable et si* h *est une fonction de* E *dans* $\mathbb{R}$ *positive ou telle que la v.a* $h\left(X\right)Z$ *est intégrable on a :*

$$E\left(h\left(X\right)Z\right)=\sum_{x\in E}h\left(x\right)E\left(Z/X=x\right)P\left(X=x\right),$$

c'est à dire

$$E\left(h\left(X\right)Z\right)=E\left(h\left(X\right)E\left(Z/X\right)\right),$$

En particulier

$$E(Z) = E(E(Z/X)).$$

Preuve. On peut écrire :

$$\begin{aligned}
E(h(X)Z) &= \sum_{x\in E} E\left(h(X)\, Z\, \mathbf{1}_{\{X=x\}}\right) \\
&= \sum_{x\in E} h(x)\, E\left(Z\, \mathbf{1}_{\{X=x\}}\right) \\
&= \sum_{x\in E} h(x)\, E(Z/X=x)\, P(X=x) \\
&= E(h(X)\, E(Z/X)),
\end{aligned}$$

en prenant la fonction $h = \mathbf{1}$

$$E(Z) = E(E(Z/X))$$

■

Définition 1.2.14 *Soit Y une v.a a valeurs dans $\mathbb{R}^d$, on appelle loi conditionnelle de Y sachant $\{X = x\}$ la loi de Y par rapport a la probabilité $P(\cdot/X = x)$, et on note $P_Y/X = x$ cette loi. La loi $P_{Y/X=x}$ est une probabilité sur $\left(\mathbb{R}^d, B\left(\mathbb{R}^d\right)\right)$ définie pour tout borélien A de $\mathbb{R}^d$ par*

$$P_{Y/X=x}(A) := P(Y \in A/X = x) = \frac{P(Y \in A \cap \{X = x\})}{P(X = x)}$$

et si $\varphi : \mathbb{R}^d \to \mathbb{R}$ un e application mesurable, on a

$$\begin{aligned}
E(\varphi(Y)/X = x) &= \int_{\Omega} \varphi(Y)\, dP(\cdot/X = x) \\
&= \int_{\mathbb{R}^d} \varphi(y)\, dP_{Y/X=x}(y).
\end{aligned}$$

– En fait on a une famille de lois de probabilités, puisque il y'en a une pour chaque valeur de $x \in E$.

– Lorsque ces lois admettent une densités s'appellent les densités conditionnelles de Y sachant que $\{X = x\}$.

– Lorsque les v.a X et Y sont indépendantes il est clair que la loi conditionnelle de Y sachant $\{X = x\}$ est égale à la loi de Y.

Proposition 1.2.7 *Soit Y une v.a à valeur dans $\mathbb{R}^d$ si ϕ est une application mesurable de $E \times \mathbb{R}^d$ dans $\mathbb{R}$ positive où telle que $\phi(X,Y)$ soit intégrable, alors :*

$$E(\phi(X,Y)) = \sum_{x \in E} P(X = x) \int_{\mathbb{R}^d} \phi(x,y)\, dP_{Y/X=x}(y).$$

Preuve. On a

$$\begin{aligned}
E(\phi(X,Y)) &= E(E(\phi(X,Y)/X)) \\
&= \sum_{x \in E} E(\phi(X,Y)/X = x) P(X = x) \\
&= \sum_{x \in E} E(\phi(x,Y)/X = x) P(X = x) \\
&= \sum_{x \in E} \int_{\mathbb{R}^d} \phi(x,y)\, dP_{Y/X=x}(y) P(X = x)
\end{aligned}$$

■

Exemple 1.2.5 *Si X et Y sont deux v.a Indépendante de loi Poisson, l'une de paramètre a $(a > 0)$, l'autre de paramètre b $(b > 0)$.*

Quelle est la loi conditionnelle de X sachant $\{X + Y = n\}$, $n \in \mathbb{N}$?

D'après la définition de la probabilité conditionnelle on a pour tout $n \in \mathbb{N}$

$$P_{X/X+Y=n}(\{k\}) = \frac{P(X = k, X + Y = n)}{P(X + Y = n)},$$

tel que $0 \leqslant k \leqslant n$, *on a d'une part*

$$\begin{aligned} P\left(X=k, X+Y=n\right) &= P\left(X=k, Y=n-k\right) \\ &= e^{-a}\frac{a^k}{k!}e^{-b}\frac{b^{n-k}}{(n-k)!} \\ &= e^{-a-b}\frac{a^k b^{n-k}}{k!\,(n-k)!}, \end{aligned} \tag{1.5}$$

et d'autre part on a aussi

$$\begin{aligned} P\left(X+Y=n\right) &= \sum_{h=0}^{n} P\left(X=k, X+Y=n\right) \\ &= \sum_{h=0}^{n} e^{-(a+b)}a^k \\ &= e^{-(a+b)}\frac{(a+b)^n}{n!}, \end{aligned} \tag{1.6}$$

d'après (1.5) et (1.6)

$$\begin{aligned} P_{X/(X+Y=n)}\left(k\right) &= e^{-(a+b)}\frac{a^k b^{n-k}}{k!\,(n-k)}\frac{n!}{(a+b)^n} \\ &= \frac{n!}{k!\,(n-k)}\left(\frac{a}{a+b}\right)^k\left(\frac{b}{a+b}\right)^{n-k} \end{aligned}$$

$X/(X+Y=n) \backsim B\left(n, \frac{a}{a+b}\right)$, par conséquent $E\left(X/_{X+Y=n}\right) = \left(\frac{a}{a+b}\right)^n$

$$E\left(X/_{X+Y}\right) = \frac{a}{a+b}\left(X+Y\right)$$

1.2.6 Somme de variables aléatoires indépendantes

Proposition 1.2.8 *Soient* X *et* Y *des v.a à valeurs entières, et* $Z = X+Y$. *Alors*

$$P\left(Z=i\right) = \sum_{j\in\mathbb{N}}\mathbb{P}\left((X,Y)=(j,i-j)\right).$$

En particulier, si X et Y sont indépendantes

$$\begin{aligned} P(Z=i) &= \sum_{j\in\mathbb{N}} \mathbb{P}(X=j)\,\mathbb{P}(Y=i-j) \\ &= \sum_{j\in\mathbb{N}} P(X=i-j)\,P(Y=j)\,. \end{aligned}$$

Si X et Y indépendantes, alors

$$G_{X+Y}(s) = G_X(s)\,G_Y(s)\,.$$

1.3 Variables aléatoires réelles

1.3.1 Lois continues

Définition 1.3.1 *On dit qu'une v.a.r. X est de loi continue si sa loi P_X est une mesure de probabilité continue..*

Une v.a.r. continue est donc telle que :

$$P_X(x) = 0$$

Proposition 1.3.1 *Une v.a.r. est continue si et seulement si sa fonction de répartition est continue.*

Définition 1.3.2 *On dit que la loi P_X d'une v.a.r. X admet f_X comme densité s'il existe une telle fonction f_X positive et telle que pour tout $x\in\mathbb{R}$ on ait*

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,du.$$

Une v.a.r. (ou sa loi) qui admet une densité est dite **absolument continue**. Cette définition est équivalente à l'existence d'une fonction f_X positive et telle que

$$\forall B \in \mathcal{B},\ P_X(B) = P(X \in B) = \int_B f_X(u)\, du.$$

où l'intégrale est prise au sens de Lebesgue.

Théorème 1.3.1 *Une fonction f sur $\mathbb{R}$ est une densité de probabilité si et seulement si elle vérifie les trois assertions suivantes :*

i) *f est positive*

ii) *f est mesurable.*

iii) *f est intégrable et*

Proposition 1.3.2 *Si f_X est continue sur un intervalle $[a, b]$, alors F_X est dérivable sur $[a, b]$ et on a $f_X = F'_X$.*

Preuve. On a, pour tout x dans l'intervalle $[a, b]$:

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\, du = \int_{-\infty}^{a} f_X(u)\, du + \int_{a}^{x} f_X(u)\, du,$$

et la proposition découle du résultat classique sur la dérivation de la fonction

$$x \to \int_{a}^{x} f_X(u)\, du.$$

■

Proposition 1.3.3 *Une variable aléatoire réelle absolument continue est continue mais la réciproque est fausse.*

Preuve. Si X est absolument continue, on a alors

$$P_X(x) = \int_{x}^{x} f_X(u)\, du = 0$$

pour tout x dans $\mathbb{R}$ et la variable aléatoire X est bien continue. ■

Principales lois de probabilité sur R absolument continues.

a) Loi uniforme sur $[a, b]$: Une v.a.r. X à valeurs dans $[a, b]$ est dite de loi uniforme sur cet intervalle si elle est absolument continue et admet pour densité

$$f(x) = \frac{1}{b-a} 1_{[a,b]}(x)$$

On note $X \sim U_{[a;b]}$.

Sa fonction de répartition est

$$F(x) = \begin{cases} 0 & \text{pour} \quad x \leq a \\ \frac{x-a}{b-a} & \text{pour} \quad a \leq x \leq b \\ 1 & \text{pour} \quad x \geq b. \end{cases}$$

La loi uniforme la plus célèbre est celle dont le support est l'intervalle $[0, 1]$

b) Loi normale $N(\mu, \sigma^2)$: Une v.a.r. X à valeurs dans $\mathbb{R}$ est dite de loi normale de moyenne μ et de variance σ^2 si elle est absolument continue et admet pour densité

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

pour $x \in \mathbb{R}$. La loi $N(0, 1)$ est appelée loi normale centrée réduite.

Notons le résultat suivant

$$\begin{aligned} &\text{si} \quad X \sim N(\mu, \sigma^2) \text{ alors } \frac{X-\mu}{\sigma} \sim N(0,1) \\ &\text{si} \quad X \sim N(0,1) \text{ alors } \mu + \sigma X \sim N(\mu, \sigma^2). \end{aligned}$$

La fonction de répartition de la loi normale n'a pas d'expression explicite mais on l'exprime souvent en fonction de celle de la loi $N(0,1)$, que l'on note souvent φ. On a

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du.$$

Ainsi, si $X \sim N(\mu, \sigma^2)$, alors

$$F_X(x) = \varphi\left(\frac{X-\mu}{\sigma}\right).$$

c) Loi exponentielle : Soit λ un réel strictement positif. Une v.a.r. X à valeurs dans $\mathbb{R}_*^+$ est dite de loi exponentielle de paramètre λ si elle est absolument continue et admet pour densité

$$f(x) = \lambda e^{-\lambda x} 1_{]0,+\infty[}(x).$$

On note $X \sim \mathcal{E}(\lambda)$.

Sa fonction de répartition est :

$$F(x) = 1 - e^{-\lambda x},$$

pour tout x positif.

1.3.2 Changement de variables

Le problème que l'on se propose d'étudier dans cette partie est la détermination de la loi de fonctions d'une v.a.r. dont on connaît la loi.

Soit donc X une v.a.r. de loi P_X et de fonction de répartition F_X. Soit φ une application mesurable de $\mathbb{R}$ vers $\mathbb{R}$. La v.a.r.$Y = \varphi \circ X$ est donc encore une v.a.r. et on cherche à déterminer sa loi.

Une première méthode, convenant autant aux variables discrètes que continues, consiste à déterminer la fonction de répartition F_Y de Y

On a, pour tout y dans $\mathbb{R}$

$$\begin{aligned} F_Y(y) &= P(Y \in]-\infty, y]) = P(Y \in]-\infty, y]) \\ &= P(\varphi \circ X \in]-\infty, y]) = P\left(X \in \varphi^{-1}(]-\infty, y])\right) \\ &= P_X\left(\varphi^{-1}(]-\infty, y])\right). \end{aligned}$$

Voyons un exemple d'application de cette méthode.

Supposons que la v.a.r. X suive une loi $N(0,1)$ et posons $Y = X^2$. On a

$$F_Y(y) = P(Y \leq y) = P\left(X^2 \leq y\right)$$

On constate déjà que l'on a : $F_Y(y) = 0$ si $y \leq 0$. Par ailleurs,

$$\begin{aligned} F_Y(y) &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}), \end{aligned}$$

alors on a

$$\begin{aligned} f_Y(y) &= \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}) \\ &= \frac{1}{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} 1_{\mathbb{R}_*^+}(y). \end{aligned}$$

Une deuxième méthode pour calculer la loi de $\varphi(X) = Y$ est donnée par le théorème suivant suivant et ne convient que pour des variables aléatoires absolument continues.

Théorème 1.3.2 *Soient S et T deux ouverts de $\mathbb{R}$ et X une v.a.r. absolument continue à valeurs dans S et de densité f_X. Soit 'φ une bijection de S vers $T = \operatorname{Im}\varphi$ continûment différentiable ainsi que son inverse (' est dite C^1-difféomorphisme). Alors, la v.a.r. $Y =$*

$\varphi(X)$ est absolument continue, à valeurs dans T et de densité

$$f_Y(y) = f_X\left(\varphi^{-1}(y)\right)\left|\left(\varphi^{-1}\right)'(y)\right| 1_{\operatorname{Im}\varphi}(y).$$

Preuve. On a :

$$F_Y(y) = P_X\left(\varphi^{-1}(]-\infty, y])\right) = \int\limits_{\{x\in S/\ \varphi(x)\leq y\}} f_X(x)\,dx$$

par le changement de variable $u = \varphi(x)$, on obtient si φ est croissante

$$F_Y(y) = \int\limits_{-\infty}^{y} f_X\left(\varphi^{-1}(u)\right)\left(\varphi^{-1}\right)'(u)\,1_T(u)\,du.$$

et donc

$$f_Y(y) = F_Y'(y) = f_X\left(\varphi^{-1}(y)\right)\left(\varphi^{-1}\right)'(y)\,1_{\operatorname{Im}\varphi}(y)$$

on obtient le signe $(-)$ dans le cas φ est décroissante. ■

Exemple 1.3.1 *Appliquons cette formule pour le calcul de la densité de la loi log-normale. Soit X une v.a de loi $N(\mu,\sigma^2)$, on pose $Y = \exp X$ quelle est la loi de Y?*
La fonction '$\varphi = \exp$ est clairement un C^1-difféomorphisme de $\mathbb{R}$ dans $\mathbb{R}_+$ d'inverse $\varphi^{-1} = \ln$' et telle que

$$\left(\varphi^{-1}\right)'(x) = \frac{1}{x}.$$

Ainsi, d'après la formule du changement de variable, on a :

$$\begin{aligned} f_Y(y) &= f_X\left(\varphi^{-1}(y)\right)\left|\left(\varphi^{-1}\right)'(y)\right| 1_{\operatorname{Im}\varphi}(y) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\ln y-\mu}{\sigma}\right)^2}\frac{1}{y}1_{\mathbb{R}_*^+}(y), \end{aligned}$$

qui est la densité de la loi log-normale.

1.3.3 Espérance pour une variable aléatoire continue

Dans cette partie nous voulons généraliser la notion d'espérance introduite dans la section 2, L'idée naturelle est de se ramener à la définition qu'on fait dans la section 2 en approchant X par une suite de variables aléatoires prenant **un nombre fini de valeurs.** On appelle variable aléatoire étagée toute variable aléatoire X qui ne prend qu'un nombre fini de valeurs, disons $a_1, ..., a_p$. D'après (1.3), elle admet donc une espérance donnée par

$$\sum_i^p a_i P\left(X = a_i\right)$$

En va construire l'espérance d'une v.a continue par étape

♦ X **une variable aléatoires positive :** On considère une suite X_n de variables aléatoires positives étagées croissant vers X, par exemple

$$X_n\left(\omega\right) = \begin{cases} k2^{-n} & \text{si } k2^{-n} \leq X\left(\omega\right) \leq \left(k+1\right)2^{-n} \quad 0 \leq k \leq n2^n - 1 \\ n & \text{sinon.} \end{cases}$$

On a

$$X_n\left(\omega\right) \leq X_{n+1}\left(\omega\right) \Rightarrow E\left(X_n\right) \leq E\left(X_{n+1}\right),$$

dans ce cas on définit l'espérance de X

$$E\left(X\right) := \lim_{n\to\infty} E\left(X_n\right)$$

Cette limite existe toujours, elle est positive, mais elle peut être infinie.et elle ne dépend pas de la suite $\left(X_n\right)_n$ approximante choisie.

♦ X **une variable aléatoire de signe quelconque :** Cette variable s'écrit $X = X^+ - X^-$ telle que

$$X^+ = \sup\left\{X, 0\right\} \text{ et } X^- = \sup\left\{-X, 0\right\}$$

et donc

$$|X| = X^+ + X^-,$$

X^+ et X^- sont deux variables aléatoires positives.

Définition 1.3.3 *On dit que la variable aléatoire X est intégrable si les valeurs $E(X^+)$ et $E(X^-)$ sont toutes les deux finies. Dans ce cas, l'espérance mathématique de X est le nombre*

$$E(X) = E(X^+) - E(X^-) \quad \textit{noté aussi} \int_\Omega X(\omega)\, dP(\omega).$$

On appelle L^1 l'ensemble des variables aléatoires intégrables, que l'on pourra aussi noter $L^1(\Omega, \mathcal{A}, P)$ si l'on souhaite préciser l'espace de probabilité sous-jacent.

Puisque $|X| = X^+ + X^-$ la proposition suivante est évidente

Proposition 1.3.4 *Soit X une variable aléatoire. Alors*

$$X \in L^1 \Longleftrightarrow E(|X|) < +\infty.$$

On montre (par passage à la limite) que les propriétés données dans la section 2 restent vraies dans notre cadre général, en effet on a

- **Linéarité :** L^1 est un espace vectoriel, et $\forall X, Y \in L^1, \forall a, b \in \mathbb{R}$,

$$E(aX + bY) = aE(X) + bE(Y)$$

- On a $X \in L^1 \Longleftrightarrow |X| \in L^1$ et on a l'inégalité suivante

$$|E(X)| \leq E(|X|).$$

- **Positivité :**

$$\text{Si } X \geq 0 \text{ , alors } E(X) \geq 0.$$

- Si $X, Y \in L^1$ telles que $X \leq Y$, alors $E(X) \leq E(Y)$.
- Si $X(\omega) = a$ pour tout ω , alors

$$E(X) = a \sum_{\omega \in \Omega} P(\omega) = a.$$

- S'il existe un réel b tel que $|X(\omega)| \leq b$ pour tout ω, alors $X \in L^1$ et $E(X) \leq b$.

1.3.4 Variance et Covariance

Outre l'espace L^1, on définit aussi, comme dans la section 2, l'espace L^2 des variables aléatoires X telles que le carré X^2 soit dans L^1.

Définition 1.3.4 *On dit qu'une variable aléatoire X est de carré intégrable si la variable aléatoire X^2 est intégrable, c'est-à-dire si son espérance est finie.*

Proposition 1.3.5 *L^2 est un sous-espace vectoriel de L^1, et si $X \in L^2$,*

$$|E(X)| \leq E(|X|) \leq \sqrt{E(X^2)}.$$

Définition 1.3.5 *Si $X \in L^2$, on définie la variance de X on la note par $Var(X)$ ou σ^2*

$$Var(X) = E\left((X - E(X))^2\right)$$

On a évidemment encore

$$Var(X) = E(X) - E(X)^2 .$$

Chapitre 2

Vecteurs Aléatoires

2.1 Loi et moment d'un vecteur aléatoire

Description de phénomènes aléatoires qui évoluent dans $\mathbb{R}^n$.

Exemple 2.1.1 $X = (X_1, X_2)$ *couple des deux cordonnées de l'impact de la flèche sur la cible.*

Définition 2.1.1 *On appelle vecteur aléatoire où variable aléatoire à valeurs dans* $\mathbb{R}^d$, *une application mesurable de* $(\Omega, \mathcal{A})$ *dans* $\left(\mathbb{R}^d, \mathbb{B}\left(\mathbb{R}^d\right)\right)$.

2.1.1 Loi d'un vecteur aléatoire

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité, si X est un vecteur aléatoire à valeurs dans $\mathbb{R}^d$, on appelle loi de X, la probabilité $\mathbb{P}_X$ définit par

$$\forall B \in \mathbb{B}\left(\mathbb{R}^d\right), \ \mathbb{P}_X(B) = \mathbb{P}\left(X^{-1}(B)\right) = \mathbb{P}(X \in B)$$

Remarque 2.1.1 *1) On a*

$$B\left(\mathbb{R}^d\right) = \sigma\left(\prod_{i=1}^{n} \left]-\infty, x_i\right], \ (x, ...x_n) \in \mathbb{R}^n\right) = \mathcal{B}(\mathbb{R}) \times ... \times B(\mathbb{R})$$

2) Si $X = (X_{1,...,}X_d)$ un vecteur aléatoire à valeurs dans $\mathbb{R}^d$, chacun de ses composante X_i $(1 \leq i \leq d)$ est une v.a.

Définition 2.1.2 *Si $X = (X_1, ..., X_d)$ est un vecteur à valeurs dans $\mathbb{R}^d$. La loi de la v.a X_i $(1 \leqslant i \leqslant d)$ est appelée la ième loi marginal .*

– La loi P_X de X est la loi joint du d-up let $(X_{1,...,}, X_d)$ probabilité sur $\mathbb{R}^d$.

– La loi P_{X_i} de $X_{_i}$ est la ième loi marginale probabilité sur $\mathbb{R}$.

Remarque 2.1.2 *La question est : qu'elle est le lien entre la collection $(P_{X_i})_{1 \leqslant i \leqslant d}$ et P_X ? comment décrire une probabilité sur $\mathbb{R}^d$?*

La loi du vecteur est caractérisée par sa fonction de répartition

$$\forall (x_1, ..., x_d) \in \mathbb{R}^d. \ F(x_1, ..., x_d) = \mathbb{P}\left(X \in \prod_{i=1}^{d}]-\infty, x_i] \right),$$

c'est une fonction difficile a manipuler.

Definition : La loi de X admet la densité f positive, intégrable sur $\mathbb{R}^n$ et d'intégrale 1 si et seulement si

$$F(x_1, ..., x_d) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_d} f(y_1, ..., y_d) \, d_{y_1} ... d_{y_d}$$

Remarque 2.1.3 *On a pour tout $B \in \mathbb{B}(\mathbb{R})$*

$$\begin{aligned}
\mathbb{P}_{Xi}(B) &= \mathbb{P}(X_i \in B) = \mathbb{P}(X_1 \in \mathbb{R}, ..., X_i \in B, ..., X_d \in \mathbb{R}) \\
&= \mathbb{P}(X \in \mathbb{R} \times ... \times B \times ... \times \mathbb{R}) \\
&= \mathbb{P}_X(\mathbb{R} \times ... \times B \times ... \times \mathbb{R}),
\end{aligned}$$

dans le cas ou f a pour densité f , on obtiens

Proposition 2.1.1 *Si le vecteur $X = (X_1, ..., X_d)$ admet une densité f, alors X_i $(1 \leqslant i \leqslant d)$ a pour densité la fonction f_i donnée par*

$$f_i\,(u) = \int\limits_{\mathbb{R}^{d-1}} f(x_1, ..., x_{i-1}, u, ..., x_{i+1}, ..., x_d) d_{x_1} ... d_{x_{i-1}} d_{x_{i+1}} ... d_{x_n}$$

on appelle cette densité la ième densité marginale.

Preuve. Dans le cas où X a pour densité f :

$$\begin{aligned}
\mathbb{P}_{X_i}(B) &= \mathbb{P}_X(\mathbb{R} \times ... \times B \times ... \times \mathbb{R}) \\
&= \int\limits_{\mathbb{R}\times...\times B\times...\times\mathbb{R}} f(x_1, ..., x_d) d_{x_1} ... d_{x_d} \\
&= \int\limits_B (\int_{\mathbb{R}^{d-1}} f(x_1, ..., x_d) d_{x_1} ... d_{x_{i-1}} d_{x_{i+1}} ... d_{x_n}) d_{x_i} \\
&= \int\limits_B f_i(u) du,
\end{aligned}$$

tel que

$$f_i(u) = \int\limits_{R^{d-1}} f(x_1, ..., x_{i-1}, u, x_{i+1} ..., x_d) d_{x_1} ... d_{x_{i-1}} d_{x_{i+1}} ... d_{x_d}$$

■

Exemple 2.1.2 *On lance une fléchette sur une cible circulaire de rayon 1, le point d'impact est un vecteur aléatoire (X, Y). Si on suppose qu'on lance d'une manière uniforme sur la cible alors la densité de ce vecteur est*

$$f(x, y) = \frac{1}{\pi} \mathbf{1}_{\{x^2+y^2 \leqslant 1\}}\,(x, y)$$

La surface d'un disque de rayoun1 est π

$$
\begin{aligned}
f_X(x) &= \int_{\mathbb{R}} f(x,y)dy = \int_{\mathbb{R}} \frac{1}{\pi}\mathbf{1}_{\{x^2+y^2\leqslant 1\}}(x,y)\,dy \\
&= \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy\right)\frac{1}{\pi}\mathbf{1}_{[-1,1]}(x) \\
&= \frac{2}{\pi}\sqrt{1-x^2}\mathbf{1}_{[-1,1]}(x)
\end{aligned}
$$

X n'est pas une loi uniforme et par symétrie, Y a même loi que X.

Pour calculer la loi du vecteur X, on a la proposition suivante :

Proposition 2.1.2 *Soit X un vecteur aléatoire à valeur dans $\mathbb{R}^d$, une probabilité μ sur $(\mathbb{R}^d, B(\mathbb{R}^d))$ est la loi de X si et seulement si, pour toute fonction φ de $\mathbb{R}^n$ dans $\mathbb{R}$ mesurable positive, on a*

$$
E(\varphi(X)) = \int_{\mathbb{R}^d} \varphi(x)d\mu(x).
$$

Preuve. Condition nécessaire : on a $\varphi : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^d, \mathbb{B}(\mathbb{R}^d), \mu)$ et supposons que μ est la loi de X alors on a

$$
\begin{aligned}
E(\varphi(X)) &= \int_{\Omega} \varphi(X(\omega))dP(\omega) \\
&= \int_{R^d} \varphi(x)d\mu(x),
\end{aligned}
$$

(par le théorème de transfert de mesure).

Condition suffisante : pour $\varphi(X) = \mathbf{1}_B(x)$, $B \in \mathbb{B}(\mathbb{R}^d)$ on a alors d'une part

$$
\int_{R^d} \varphi(x)d\mu(x) = \int_B d\mu(x) = \mu(B),
$$

et d'autre part

$$E(\varphi(x)) = \int\limits_{\Omega} \varphi(X(\omega))d\mathbb{P}(\omega)$$
$$= \int\limits_{\Omega} \mathbf{1}_{X^{-1}(B)}(\omega)d\mathbb{P}(\omega)$$
$$= P(X^{-1}(B)) = \mu(B),$$

alors μ est la loi de X. ■

Remarque 2.1.4 *Nous pouvons, utiliser cette proposition pour trouver la loi d'un vecteur aléatoire image d'un autre vecteur aléatoire de loi connue. Soit X un vecteur aléatoire p.s à valeurs dans un ouvert U de $\mathbb{R}^d$ avec f comme densité. Soit ϕ un C_1 difféomorphisme de U dans un ouvert V de $\mathbb{R}^d$ le vecteur aléatoire $Y = \phi(X)$ est p.s. a valeurs dans V. Pour trouver la loi de Y, nous pouvons écrire pour toute fonction φ mesurable, positive de $\mathbb{R}^d$ dans $\mathbb{R}$*

$$E(\varphi(Y)) = E(\varphi(\phi(X))$$

$$\int\limits_{U} \varphi\left(\phi\left(x\right)\right) f\left(x\right) dx$$

$$\int\limits_{V} \varphi\left(y\right) f\left(\phi^{-1}\left(y\right)\right) \left|J(y)\right| dy$$

où $J(y)$ est le Jacobien de ϕ^{-1} au point y alors la densité de Y est $f(\phi^{-1}(y))\,|J(y)|\,1_V(y)$.

Exemple 2.1.3 $X = (X_1, X_2)$ *un vecteur de densité sur* $\mathbb{R}^2$

$$f(x,y) = \frac{1}{2\pi}\exp(-\frac{x2+y2}{2}),$$

soit $Y = \dfrac{X_1}{X_2}$ *si* $X_2 \neq 0$ *et* $Y = 0 \quad X_2 = 0$

Quelle est la loi de Y ?. Si φ est une fonction mesurable de $\mathbb{R}$ dans $\mathbb{R}$ on a

$$
\begin{aligned}
E(\varphi(Y)) &= E\left(\varphi\left(\frac{X_1}{X_2}\mathbf{1}_{\{X_2\neq 0\}}\right)\right) \\
&= \int\limits_{\mathbb{R}^2} \varphi\left(\frac{X}{X_2}\mathbf{1}_{\{X_2\neq 0\}}\right)\frac{1}{2\pi}exp(-\frac{x_2+x_2}{2})dx_1dx_2 \\
&= \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}} \varphi\left(\frac{X_1}{X_2}\right)\frac{1}{2\pi}exp(-\frac{x_2+x_2}{2})dx_1dx_2,
\end{aligned}
$$

par un changement de variable

$$
\begin{cases} u = \dfrac{x}{y} \\ v = y \end{cases} \Longrightarrow \begin{cases} x = uv \\ y = v \end{cases}
$$

$$
|J| = \begin{vmatrix} v & 0 \\ u & 1 \end{vmatrix},
$$

on a donc

$$
\begin{aligned}
E(\varphi(Y)) &= \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}} \varphi(u)\frac{1}{2\pi}exp\left(\frac{-v^2(u^2+1)}{2}\right)|v|\,dud\theta \\
&= \int\limits_{\mathbb{R}} \varphi(u)\int\limits_{\mathbb{R}} \frac{1}{\pi}exp\left(\frac{-v^2(v^2+1)}{2}\right)vdv)du
\end{aligned}
$$

en effectuions le changement de variable suivants $\dfrac{v^2}{2} = y$ c'est ne bijection sauf sur $R_+ \to R_+$, on obtiens

$$
\begin{aligned}
E(\varphi(Y)) &= \int\limits_{\mathbb{R}} \varphi(u)\frac{1}{\pi}\int\limits_{\mathbb{R}} e^{-y(u^2+1)}dydu \\
&= \int\limits_{\mathbb{R}} \varphi(u)\frac{1}{\pi(1+v^2)}dv
\end{aligned}
$$

La densité de Y est $\dfrac{1}{\pi(1+u^2)}$ c'est la loi de Cauchy.

2.1.2 Moments d'un vecteur aléatoire

Définition 2.1.3 *Si $X = (X_1, ..., X_d)$ un vecteur aléatoire lorsque $\|X\|$ est intégrable ou lorsque chaque composante est une v.a positive ou nulle, on appelle espérance (ou moyenne d'ordre1) de X, le vecteur de $\mathbb{R}^d$ égal a :*

$$E(X) = (E(X_1), ..., E(X_d)).$$

On peut calculer $E(X_i)$ a laide de la loi de X_i où a laid de la loi de X

$$E(X_i) = \int_{\Omega} X_i(\omega)\, dP(\omega) = \int_{\mathbb{R}} x_i P_{X_i}(x_i) = \int_{\mathbb{R}^d} x_i dP_X(x).$$

Définition 2.1.4 *Si X et Y sont deux v.a réelles de carré intégrable, on appelle covariance de X et Y la quantité*

$$cov(X, Y) = E[(X - E(X))(Y - E(Y))]$$

Proposition 2.1.3 *Toutes les v.a considérées sont de carré intégrable. Alors :*

– (cov est symétrique)

$$Cov(X, Y) = Cov(X, Y)$$

–

$$Cov(X, Y) = Var(X) = E\left(X^2\right) - E(X)^2$$

–

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

– (Cov est bilinéaire) si a_1,$a_2 \in \mathbb{R}$, on a :

$$Cov(a_1X_1 + a_2X_2, Y) = a_1Cov(X_1, Y) + a_2Cov(X_2, Y).$$

Proposition 2.1.4 *Si $X_1, ..., X_d$ sont des v.a.r de carré intégrable, on a la formule*

$$var\left(\sum_{i=1}^{d} X_i\right) = \sum_{i=1}^{d} var(X_i) + 2 \sum_{1\leqslant i<j\leqslant d} cov(X_i, X_j).$$

Preuve. Posons $Y_i = X_i - E\left(X_i\right)$. on a

$$\left(\sum_{i=1}^{d} Y_i\right)^2 = \sum_{i=1}^{d} \left(Y_i\right)^2 + 2 \sum_{1\leqslant i<j\leqslant d} Y_i Y_j,$$

et on prend l'espérence$\left[\sum_{i=1}^{d} Y_i = \sum_{i=1}^{d} X_i - E\left(\sum_{i=1}^{d} X_i\right)\right]$

$$Var\left(\sum_{i=1}^{d} X_i\right) = E\left(\sum_{i=1}^{d} Y_i\right)^2 = \sum_{i=1}^{d} Var\left(X_i\right) + 2 \sum_{1\leqslant i<j\leqslant d} cov(X_i, X_j).$$

■

Définition 2.1.5 *Le coefficient de Corrélation de deux v.a.r X et Y de carrés intégrable et de variance nou nulles est définit par*

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}.$$

Proposition 2.1.5 *Soient X et Y deux v.a.r de carré intégrable. Alors*

$$|Cov(X,Y)| \leqslant \sqrt{Var(X)Var(Y)},$$

et par conséquent

$$|\rho(X,Y)| \leqslant 1,$$

si les Variances de ces v.a.r ne sont pas nulles)

Preuve. Il suffit d'appliquer l'inégalité de Schwartz aux v.a.r $X - E(X)$ et $Y - E(Y)$

$$\begin{aligned} Cov(X,Y) &= E\left(X - E\left(X\right)\right)\left(Y - E\left(Y\right)\right) \\ &\leqslant E((X - E(X))^2)^{\frac{1}{2}} E((Y - E(Y))^2)^{\frac{1}{2}} \\ &= \sqrt{Var(X)Var(Y)}. \end{aligned}$$

■

Définition 2.1.6 *Si $X = (X_1, ..., X_d)$ un vecteur aléatoire à valeurs dans $\mathbb{R}^d$ tel que $E(\|\times\|^2) < \infty$. On appelle matrice de dispersion de X ou (matrice de variance covariance) la matrice $d \times d$ de terme générale*

$$(D_X)_{i,j} = cov(X_i, X_j),$$

$$D_X = E\left[\left(X - E\left(X\right)\right)\left(X - E\left(X\right)^t\right)\right].$$

tel que X et $E(X)$ sont représentes par matrice colonnes.

Proposition 2.1.6 *Si X est un vecteur aléatoire à valeurs dans $\mathbb{R}^d$, tel que $E(\|\times\|^2)\langle\infty$. Si A est la matrice représentant une application linéaire de $\mathbb{R}^d$ dans $\mathbb{R}^d$ et si B est un vecteur de $\mathbb{R}^d$, alors le vecteur aléatoire $Y = AX + B$ a valeurs dans $\mathbb{R}^d$ est tel que $E(\|Y\|^2)\langle\infty$ et on a*

$$E(Y) = A\ E\left(X\right) + B,$$

$$D_Y = AD_XA^t.$$

Preuve. On a

$$\forall_i, X_i \in L^2\left(\Omega, \mathcal{A}, P\right) \Longrightarrow \forall_i, Y_i \in L^2\left(\Omega, \mathcal{F}, P\right),$$

$$E(\|Y\|^2) = \sum_{i=1}^{d} E\left(Y_i^2\right),$$

est finie puisque l'intégrale est linéaire alors on a

$$E(Y) = AE(X) + B.$$

Calculons D_Y :

$$\begin{aligned} D_Y &= E\left[(Y - E(Y))\left(Y - E(Y)^t\right)\right] \\ &= E\left[(AX + B - E(AX + B))(AX + B - E(AX + B))^t\right] \\ &= E\left[A(X - E(X))(X - E(X))^t A^t\right] \\ &= AD_X A^t \end{aligned}$$

■

2.2 Vecteurs aléatoires indépendants

Dans cette section nous allons généraliser la notion d'indépendance de v.a au cas de vecteur aléatoire.

Définition 2.2.1 *X et Y sont indépendants si pour tous boréliens A et B de $\mathbb{R}^n$, on a*

$$P(X \in A, Y \in B) = P(X \in A) \times P(Y \in B).$$

Proposition 2.2.1 *Si X et Y ont des lois à densité, alors X et Y sont indépendantes si et seulement si pour (presque) tous $x, y \in \mathbb{R}^n$*

$$f_{(X,Y)}(x, y) = f_X(x) . f_Y(y) .$$

Preuve. On sait que la loi d'un vecteur est caractériser par sa fonction de répartition $\forall x, y \in \mathbb{R}^n$

$$\begin{aligned} F_{(X,Y)}(x,y) &= P(X \leqslant x, Y \leqslant y) \\ &= P(X \leqslant x)\ P(Y \leqslant y) \\ &= \int_{-\infty}^{x} f_X(u)du \int_{-\infty}^{y} f_Y(v)dv \end{aligned}$$

d'où

$$\int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(u,v)dudv = \int_{-\infty}^{x}\int_{-\infty}^{y} f_X(u)f_Y(v)dudv$$

alors

$$f_{(X,Y)}(u,v) = f_X(u)f_Y(v)$$

réciproquement si on suppose $f_{(X,Y)}(u,v) = f_X(u)f_Y(v)$, alors $\forall x, y \in \mathbb{R}^n$

$$\int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(u,v)dudv = \int_{-\infty}^{x}\int_{-\infty}^{y} f_X(u)f_Y(v)dudv$$

et donc

$$P(X \leqslant x, Y \leqslant y) = P(X \leqslant x)P(Y \leqslant y).$$

■

Théorème 2.2.1 *Soient $X \in \mathbb{R}^m$ et $Y \in \mathbb{R}^n$ des vecteurs aléatoires indépendants, soient g et h mesurables positives ou bornées respectivement sur $\mathbb{R}^m$ et $\mathbb{R}^n$, alors*

$$E(g(X)h(Y)) = E(g(X))E(h(Y)).$$

Preuve. 1) Cas générale : vrais pour $\mathbf{g}$ et $\mathbf{h}$ fonctions indicatrices d'ensembles, donc vrais pour les fonctions étagées, vrais pour les fonctions positives par passage à la limite puis on écrit $\mathbf{g} = \mathbf{g}^+ - \mathbf{g}^-$, $\mathbf{h} = \mathbf{h}^+ - \mathbf{h}^-$

2)Si X et Y ont des lois a densité, la preuve est plus simple

$$\begin{aligned}
E(g(X)h(Y)) &= \int_{\mathbb{R}^m \times \mathbb{R}^n} g(X)h(Y)\ f_{X,Y}(x,y)dxdy \\
&= \int_{\mathbb{R}^m \times \mathbb{R}^n} g(X)h(Y)f_X(x)f_Y(y)dxdy \\
&= \left(\int_{\mathbb{R}^m} g(X)f_X(x)dx\right)\left(\int_{\mathbb{R}^n} h(Y)f_Y(y)dy\right) \\
&= E(g(X))E(h(Y)),
\end{aligned}$$

par le théorème de Fubini car **h** et **g** sont positifs sont bornée. ■

Corollaire 2.2.1 *Si X et Y sont deux variable aléatoires réelles indépendantes, alors*

$$Cov(X, Y) = 0.$$

Preuve. Immédiate d'après le théorème précédent

$$E\left[(X - E(X))(Y - E(Y))\right] = E(X - E(X))E(Y - E(Y)) = 0.$$

■

Remarque 2.2.1 *La réciproque est fausse voir l'exemple des fléchettes.*

on a $Cov(X, Y) = 0$. or X et Y ne sont pas indépendantes.

En effet on a d'une part

$$f_X(x)f_Y(y) = \frac{4}{\pi^2}\sqrt{1-x^2}\sqrt{1-y^2}\mathbf{1}_{[-1,1]}(x)\,\mathbf{1}_{[-1,1]}(y)$$

$$\neq f_{(X,Y)}(x,y) = \frac{1}{\pi}\mathbf{1}_{\left\{x^2+y^2\leqslant 1\right\}}(x,y)$$

et d'autre part

$$Cov(X, Y) = E(XY) - E(X)E(Y) = 0$$

car

$$
\begin{aligned}
E(XY) &= \int_{\mathbb{R}^2} xy\frac{1}{\pi}\mathbf{1}_{\{x^2+y^2\leqslant 1\}}(x,y)dxdy \\
&= \int\limits_{-1}^{1}\int\limits_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} xy\frac{1}{\pi}dydx \\
&= 0
\end{aligned}
$$

et

$$
\begin{aligned}
E(X) &= \frac{2}{\pi}\int_{-1}^{1} x\sqrt{1-x^2}dx \\
&= 0
\end{aligned}
$$

2.2.1 Fonction caractéristique

- Cadre général de vecteurs aléatoires
- Une nouvelle fonction pour caractériser la loi
- Fonction à valeurs complexes
- Mathématiquement : transformée de Fourier

Notation 2.2.1 *Pour $X, Y \in \mathbb{R}^d$, on note*

$$\langle X, Y\rangle = \sum_{j=1}^{d} x_i y_i.$$

Si X est un vecteur aléatoire à valeurs dans $\mathbb{R}^d$ et pour $u \in \mathbb{R}^d$ alors

$$e^{\mathbf{i}\langle u,X\rangle} = \cos\langle u, X\rangle + i\sin\langle u, X\rangle$$

est une v.a bornée (en module) par 1

Définition 2.2.2 *La fonction caractéristique de X est la fonction ϕ_X définie' de $\mathbb{R}^d$ dans*

$\mathbb{C}$ *par :*

$$\phi_X(u) = E\left(e^{\mathbf{i}\langle u,X\rangle}\right) = E\left(e^{\mathbf{i}\sum\limits_{j=1}^{d} ujXj}\right)$$

$$= E\left(\cos\langle u, X\rangle\right) + iE\left(\sin\langle u, X\rangle\right),$$

si X est à valeurs réelles cas ou $(d = 1)$, alors

$$\forall u \in \mathbb{R},\ \phi_X(u) = E\left(e^{\mathbf{i}uX}\right)$$

La fonction ϕ_X ne dépend que de la loi de X

Cas discret

Loi de $\left\{(X_k, P_k), k \in \mathbb{N}\right\}$ alors

$$\forall u \in \mathbb{R}, \phi_X(u) = \sum_{k\in\mathbb{N}} P_k e^{iuX_k}.$$

Supposons que X soit à valeurs dans $\mathbb{N}$

$\phi_X(u) = G_X\left(e^{iu}\right)$, où G_X fonction génératrice de X. En effet

$$\phi_X(u) = \sum_{k\in\mathbb{N}} P_k \left(e^{iu}\right)^k = G_X\left(e^{iu}\right)$$

Cas d'une v.a.r à densité f

$$\forall u \in \mathbb{R}, \phi_X(u) = \int\limits_{\mathbb{R}} e^{\mathbf{i}uX} f_X(x)dx = E\left(e^{\mathbf{i}uX}\right),$$

(calculer l'intégrable d'une fonction complexe (théorème des résidu).

Proposition 2.2.2 *Soit X un vecteur aléatoire de $\mathbb{R}^d$. Alors la fonction ϕ_X est continue de module inférieur a* **1**, *et*

$$\phi_X(0) = 1,\ \phi_X(-u) = \overline{\phi_X(u)}, \forall u \in \mathbb{R}^d$$

Preuve. De la continuité on a

$$\begin{aligned} u_p \to u &\Rightarrow e^{\mathbf{i}\langle Up, X(\omega)\rangle} \to e^{\mathbf{i}\langle U, X(\omega)\rangle},\ \forall \omega \in \Omega \\ &\Rightarrow e^{\mathbf{i}\langle u_p, X\rangle} \to e^{\mathbf{i}\langle u, X\rangle} \quad p.s, \end{aligned}$$

tous ces $v.a.r$ sont bornées en module par **1** par le théorème de convergence dominée

$$\phi_X(u_p) \to \phi_X(u).$$

Par ailleurs,

$$\begin{aligned} \phi_X(-u) &= E(\cos\langle -U, X\rangle) + iE(\sin\langle -U, X\rangle) \\ &= E(\cos\langle U, X\rangle) - iE(\sin\langle U, X\rangle) \\ &= \overline{\phi_X(u)} \end{aligned}$$

■

1) Fonction Caractéristique des variables usuelles

– $X \backsim B(n, p)$

$$\begin{aligned}\phi_X(u) = G_X(e^{iu}) &= \sum_{k=0}^{n} P(X=k)e^{iuk} \\ &= \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} e^{iuk} \\ &= \sum_{k=0}^{n} C_n^k \left(pe^{iu}\right)^k (1-p)^{n-k} \\ &= \left(pe^{iu}(1-p)\right)^n\end{aligned}$$

– $X \backsim P(\lambda)$

$$\begin{aligned}\phi_X(u) = G_X(e^{iu}) &= \sum_{k\geqslant 0} e^{-\lambda}\frac{\lambda^k}{k!}e^{iuk} \\ &= \sum_{k\geqslant 0} e^{-\lambda}\frac{\left(\lambda e^{iu}\right)^k}{k!} \\ &= e^{-\lambda}\sum_{k\geqslant 0} \frac{\left(\lambda e^{iu}\right)^k}{k!} \\ &= e^{-\lambda}e^{\lambda e^{iu}} = e^{\lambda\left(e^{iu}-1\right)}\end{aligned}$$

– $X \backsim U\left([a,b]\right)$

$$\phi_X(u) = \frac{1}{b-a}\int_a^b e^{iux}dx = \frac{e^{iub}-e^{iua}}{iu\left(b-a\right)}$$

– $X \backsim U\left[-a,a\right],\ a>0$

$$\phi_X(u) = \frac{1}{2a}\int_{-a}^{a} e^{iux}dx = \frac{\sin ua}{ua}$$

– X est une variable de loi exponentielle de paramètre $\lambda > 0$

$$\phi_X(u) = \lambda \int_0^{+\infty} e^{-\lambda x} e^{iux} dx = \frac{\lambda}{\lambda - ui}$$

car

$$\underset{x \to +\infty}{Lim} \left| \frac{\lambda}{\lambda - ui} e^{-\lambda x + iux} \right| = 0$$

– X est une variable de loi normale $N(0,1)$

si X est une variable de loi normale $N(0,1)$, sa fonction caractéristique vaut

$$\phi_X(u) = \varrho^{\frac{-U^2}{2}}$$

– $X \backsim N(m, 6^2)$

$X = m + \sigma Y$, où $Y \backsim N(0,1)$.

alors

$$\phi_X(u) = e^{ium - \frac{U^2 6^2}{2}}$$

Remarque 2.2.2 *La méthode générale de calcul dans le cas d'une densité (théorème des résidus analyse complexe)*

Exemple 2.2.1 *Prenons l'exemple d'une v.a qui suit la loi normale centrée réduite* $(i.e\ X \backsim N(0,1))$ *on a*

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \varrho^{\frac{-x^2}{2}}$$

alors

$$\phi_X(u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{iux} e^{\frac{-x^2}{2}} dx.$$

On a pour $s \in \mathbb{R}$

$$
\begin{aligned}
\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{sx} e^{\frac{-x^2}{2}} dx &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varrho^{\frac{s^2}{2}} \varrho^{\frac{-\left(s-x^2\right)}{2}} dx \\
&= e^{\frac{s^2}{2}} X \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{\frac{-\left(s-x^2\right)}{2}} dx \\
&= e^{\frac{s^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-y^2} dy \\
&= e^{\frac{\partial^2}{2}} \\
&= \sum_n \frac{s^{2n}}{2^n n!}, \qquad (2.1)
\end{aligned}
$$

série entière de rayon de convergence ∞, *et pour le terme de droite on a puisque*

$$e^{sx} = \sum_n \frac{s^n x^n}{n!}$$

alors

$$
\begin{aligned}
\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{sx} e^{\frac{-x^2}{2}} dx &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \sum_n \frac{s^n x^n}{n!} dx \\
&= \sum_n \frac{s^n}{n!} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^n e^{-x^2} dx \qquad (2.2)
\end{aligned}
$$

qui est aussi une série entière de rayon ∞

Les deux série (2.1), (2.2) coïncide sur $\mathbb{R}$, *et donc elle coïncide sur* $\mathbb{C}$. *en particulier pour* $s = iu$ *et donc*

$$\phi_X(u) = e^{\dfrac{-u^2}{2}}.$$

Le moment de la variable entière $X \backsim N(0,1)$, *on a d'après (2.1) et (2.2)*

$$\sum_n \frac{s^n}{n!} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^n e^{-x^2} dx = \sum_n \frac{s^{2n}}{2^n n!},$$

si on identifie terme à terme les deux séries on a (Le terme de droite a seulement les terme *paire)*

pour tout n impaire

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} x^n e^{\frac{-x^2}{2}}dx = E\left(X^n\right) = 0,$$

pour tout n pair

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} x^{2k} \varrho^{\frac{-x^2}{2}}dx = \frac{(2k)!}{2^k k!} = E\left(X^{2k}\right) = E\left(X^n\right).$$

En déduit pour $X \backsim N\left(m, \sigma^2\right)$ tel que $X = m + 6y$, où $Y \backsim N(0,1)$.

$$\begin{aligned} E(e^{iuX}) &= E\left(e^{ium}e^{iu\sigma Y}\right) \\ &= e^{ium}E\left(e^{iu\sigma Y}\right) \\ &= e^{ium}\phi_Y\left(u\sigma\right) \\ &= e^{ium}e^{\frac{-u^2\sigma^2}{2}} \\ &= e^{ium-\frac{u^2\sigma^2}{2}}. \end{aligned}$$

2) Propriété fondamentale

Théorème 2.2.2 *La fonction caractéristique ϕ_X caractérise la loi du vecteur aléatoire X (i.e si deux vecteurs aléatoires ont même fonction caractéristique, ils ont même loi).*

Remarque 2.2.3 *Pour la preuve on utilise les propriétés de la transformée de Fourier (preuve technique).*

Proposition 2.2.3 *Soit $X = (X_1, ..., X_n)$ les composants X_i sont indépendantes si et seulement si pour tous $u_1, ..., u_n \in \mathbb{R}$, on a*

$$\phi_X\left(u_1, ..., u_n\right) = \prod_{j=1}^{n} \phi_{X_j}.\left(u_j\right).$$

Preuve. Premièrement supposons que les X_i sont indépendantes, on a donc

$$\begin{aligned}\phi_X(u_1,...,u_n) &= E\left(e^{i(u_1X_1+...+u_nX_n)}\right)\\ &= \prod_{j=1}^{n} E\left(e^{iu_jX_j}\right)\\ &= \prod_{j=1}^{n} \phi_{X_j}.(u_j)\end{aligned}$$

Inversement, supposons que

$$\phi_X(u_1,...,u_n) = \prod_{j=1}^{n} \phi_{X_j}.(u_j).$$

On peut construire des v.a X'_j ,$j \in \{1,...,n\}$ indépendantes, et telles que $\forall j \in \{1,...,n\}$

$$loi\left(X'_j\right) = loi\left(X_j\right)$$

on a donc

$$\begin{aligned}\phi_{X'}(u_1,...,u_n) &= \prod_{j=1}^{n} \phi_{X'_j}(u_j)\\ &= \prod_{j=1}^{n} \phi_{X_j}.(u_j)\\ &= \phi_X(u_1,...,u_n)\end{aligned}$$

se qui implique que

$$\phi_{X'} = \phi_X \Rightarrow loi\left(X'\right) = loi\left(X\right)$$

on a donc

$$loi\left(X'\right) = \prod_{j=1}^{n} loi\left(X'_j\right) = \prod_{j=1}^{n} loi\left(X_j\right) = loi\left(X\right)$$

alors les v.a X_i $i \in \{1,...,n\}$ sont indépendante. ■

2.2.2 Somme de variables aléatoires indépendantes

Proposition 2.2.4 *Si X et Y sont deux v.a indépendantes à valeurs dans $\mathbb{R}$*

$$\phi_{X+Y}(u) = E(e^{iu(X+Y)}) = E(e^{iuX}e^{iuY})$$

$$= \phi_X(u)\phi_Y(u)$$

Remarque 2.2.4 *Cette proposition est très utilisée dans la pratique pour trouver la loi d'une somme de v.a.*

Exemple 2.2.2 *X et Y suivent des lois normales $N(m, 6^2)$ et $N\left(m', (6')^2\right)$ indépendantes et si $Z = X + Y$, alors $Z \backsim N\left(m + m', \sigma^2 + (\sigma')^2\right)$.*
On a pour tout $u \in \mathbb{R}$

$$\phi_Z(u) = \phi_X(u)\phi_Y(u),$$

$$\phi_X(u) = e^{ium}e^{\frac{-\sigma^2}{2}},$$

$$\phi_Y(u) = e^{ium'}e^{\frac{-\sigma'^2}{2}},$$

alors

$$\phi_Z(u) = e^{iu(m+m')}e^{\frac{-\left(\sigma^2+\sigma'^2\right)}{2}},\text{ c'est une fonction caractéristique d'une loi normale}$$

et donc $Z \backsim N\left(m + m', \sigma^2 + (\sigma')^2\right)$.

2.2.3 Fonction Caractéristique et moments

Supposons que X soit à valeurs réelles et de carré intégrable et

$$\phi_X(u) = E\left(e^{iuX}\right)$$

La question qui se pose avons nous le droit d'écrire

$$\phi_X'(u) = E\left(iXe^{iUX}\right),\ \phi_X''(u) = E\left(-X^2e^{iuX}\right)$$

se qui implique que si $u = 0$ on aurait

$$\phi_X'(0) = E(iX) \text{ et } \phi_X''(0) = E\left(-X^2\right)$$

(calcule de moment d'une dérivation sous signe somme).

Sous quelle hypothèse pouvons-nous le faire ?

Proposition 2.2.5 *soit* $X = (X_1, ..., X_n)$. *Si la v.a* $|X|^m$ *est intégrable pour un entier* m, *alors la fonction* ϕ_X *est* m *fois continûment différentiable sur* $\mathbb{R}^n$ *et ou a pour tout choix des indices* $i_1, ..., i_m$

$$\frac{\partial^m}{\partial u_{i_1}, \partial u_{i_2}, ..., \partial u_{i_m}} \phi_X(u) = i^m E\left(e^{i\langle u,X\rangle} X_{i_1}...X_{i_m}\right).$$

Preuve. En fait la preuve dans le cas $m = 1$, on a pour tout $j \in \{1, ..., n\}$, $|X_j|$ est intégrable et nous étudions la dérivée première de la fonction ϕ_X.
Soit $v_j = (0, ...0, 1, 0, ..., 0)$, le j ème vecteur de base de $\mathbb{R}^n$, on a

$$\begin{aligned} \frac{\phi_X(u + tv_j) - \phi_X(u)}{t} &= \frac{E\left(e^{i<u+tv_j,X>}\right) - E\left(e^{i<u,X>}\right)}{t} \\ &= E\left(e^{i<u,X>} \frac{e^{itX_j} - 1}{t}\right) \end{aligned}$$

on veut calculer la limite quand t tend vers 0 ?.
Soit la suite $(t_p) \to 0$ quand $p \to +\infty$,
Les v.a $\frac{e^{it_pX_j}-1}{t_p}$converge simplement vers iX_j $\left(i.e\ \forall\omega \in \Omega : \frac{\varrho^{it_pX_j(\omega)}-1}{t_p} \to iX_j(\omega)\right)$ de plus, les v.a $\varrho^{i<u,X>}\left(\frac{\varrho^{itX_j}-1}{t}\right)$ son bornées en module par $\sqrt{2}\,|X_j| \in L^1$. Le théorème de convergence dominée entraine que la suite $E\left(\varrho^{i<u,X>}\left(\frac{\varrho^{itX_j}-1}{t}\right)\right)_p$ converge vers $iE\left(\varrho^{i<u,X>}X_j\right)$

quand $p \to +\infty$. ■

Application Calcul des moments $E\left(X_{i_1}X_{i_2}...X_{i_m}\right)$ en fonction des dérivées de ϕ_X en $(0, ..., 0)$.

Par exemple, si X est à valeurs réelles et de carré intégrable, on a

$$E(X) = -i\phi'_X(0), \ E\left(X^2\right) = -\phi''_X(0)$$

2.2.4 Vecteurs Gaussiens

On a déjà rencontré des v.a gaussiennes sur $\mathbb{R}$ ce sont les v.a de loi $N\left(m, 6^2\right)$.

Dans la suite on considérera également les v.a.r presque surement constantes (donc de covariance nulle) comme gaussiennes.

Définition 2.2.3 *Un vecteur aléatoire à valeurs dans $\mathbb{R}^d$ est dit gaussien si toute combinaison linéaire de ses composantes est une v.a réelle gaussienne. La loi d'un vecteur gaussien est dite gaussienne.*

Proposition 2.2.6 *Toute image par une application linéaire ou affine d'une vecteur gaussien est encore vecteur gaussien.*

Exemple 2.2.3 *1)* $X = (X_1, ..., X_d)$ *où* $X_1, ..., X_d$ *des v.a.r indépendantes et* $X_i \backsim N\left(m_i, 6_i^2\right)$ *, pour* $1 \leqslant i \leqslant d$.

On a la v.a.r $\sum\limits_{i=1}^{d} a_i X_i$ *suit une loi gaussienne de moyenne* $\sum\limits_{i=1}^{d} a_i m_i$ *et d'écart type* $\sqrt{\sum\limits_{i=1}^{d} a_i^2 6_i^2}$

donc X *est un vecteur gaussien* $\forall t \in \mathbb{R}$,

$$\begin{aligned}
\phi_{\sum\limits_{i=1}^{n} a_k X_k}(t) &= \prod_{k=1}^{n} \phi_{a_k X_k}(t) \\
&= \prod_{k=1}^{n} \phi_{X_k}(a_k t) \\
&= \prod_{k=1}^{n} \exp(ita_k m_k - \frac{a_k^2 \sigma_k^2 t^2}{2}) \\
&= \exp\left(it \sum_{k=1}^{n} a_k m_k - \frac{t^2}{2} \sum_{k=1}^{n} a_k^2 \sigma_k^2 \right)
\end{aligned}$$

et donc

$$\sum_{k=1}^{n} a_k X_k \backsim N\left(\sum_{k=1}^{n} a_k m_k, \sum_{k=1}^{n} a_k^2 6_k^2 \right)$$

2)$X = (X_1, X_2)$ *où* $X_1 \backsim N(m, 6^2)$, $X_2 = b$ *p.s*

$\forall (a_1, a_2) \in \mathbb{R}^2$, *un calcule de loi rapide montre que la v.a*

$$a_1 X_1 + a_2 X_2 \backsim N\left(a_1, m + a_2 b, 6^2\right)$$

et donc X *est un vecteur gaussien.*

Remarque 2.2.5 Y *est un vecteur aléatoire gaussien, on montre qu'il s'écrit sous la forme* $Y = AX + b$ *où* A *est une matrice déterministe,* b *un vecteur déterministe et* X *un vecteur aléatoire dans les composantes sont des v.a gaussiennes indépendantes.*

Proposition 2.2.7 *Si* X *est un vecteur aléatoire gaussien à valeurs dans* $\mathbb{R}^d$ *de moyenne et de matrice de dispersion* D*, sa fonction caractéristique* ϕ *est égale à* $\phi(U) = \exp\left(iU^t m - \frac{1}{2} U^t D U\right)$ *où* U *est un vecteur de* $\mathbb{R}^d$

$$U^t m = \langle U, m \rangle = \sum_{j=1}^{d} U_j m_j$$

$$U^t DU = \langle U, DU \rangle = \sum_{j=1}^{d} \left(\sum_{k=1}^{d} U_j D_{jk} U_k \right)$$

Preuve. Si $X = (X_1, ..., X_d)$ est un vecteur gaussien, sa fonction caractéristique est

$$\phi_X(U) = \phi_X(U_1, ..., U_d) = E\left(e^{i \sum\limits_{k=1}^{d} U_k X_k} \right)$$

$$= \phi_{\sum\limits_{k=1}^{d} U_k X_k}(1)$$

■

or $\sum\limits_{k=1}^{d} U_k X_k = U^t X$ est gaussienne car X est vecteur gaussien de moyenne $U^t m$ et de variance $U^t DU$ sa fonction caractéristique au point **1** est

$$\exp\left(iU^t m - \frac{1}{2} U^t DU \right).$$

Remarque 2.2.6 *On constate que la loi d'un vecteur gaussien est entièrement déterminer par D, et on note cette loi*

$$X\tilde{\ } N_d(m, D).$$

Une propriété importante des vecteurs gaussiens est que l'indépendance des composantes est équivalente à la nullité des covariances.

Proposition 2.2.8 *Soit $X = (X_1, ..., X_d)$ un vecteur gaussien de loi $N_d(m, D)$. Les v.a.r $X_1, ..., X_d$ sont indépendantes si et seulement si la matrice D est diagonale.*

Preuve. Si les v.a.r $X_1, ..., X_d$ sont indépendantes, on sait que les covariances de X_i et X_j sont nulles si $i \neq j$

Réciproquement si D est diagonale et si on désigne σ_j^2 les termes diagonaux de D,

$$\begin{aligned}\phi_{(X_1,\dots,X_d)}(U) &= \exp\left(i\sum_{j=1}^{d} U_j m_j - \frac{1}{2}\sum_{j=1}^{d} U_j^2\sigma_j^2\right) \\ &= \prod_{j=1}^{d} \exp\left(iU_j m_j - \frac{1}{2}U_j^2\sigma_j^2\right) \\ &= \prod_{j=1}^{d} \phi_{X_j}(U_j)\end{aligned}$$

$X_j\tilde{}N\left(m_j, 6_j^2\right)$, et donc les X_j sont indépendantes. ■

Remarque 2.2.7 *Il faut prendre bien garde d'appliquer cette proposition avec toutes ses hypothèses, deux v.a.r gaussiennes U et V de covariance nulle ne sont pas forcément indépendant il faut que le couple (U, V) soit gaussien.*

Lorsque la matrice de dispersion d'un vecteur gaussien est inversible, la loi de se vecteur admet une densité par rapport a la mesure de Lebesgue et on a la formule explicite de cette densité

Proposition 2.2.9 *Soit X un vecteur gaussien de loi $N_d(m, D)$. La loi de X admet une densité si et seulement si la matrice D est inversible, et dans ce cas la densité est*

$$\frac{1}{(2\pi)^{\frac{-d}{2}}}(\det(D))^{\frac{-1}{2}}\exp\left(\frac{-1}{2}(x-m)^t D^{-1}(x-m)\right)$$

Preuve. Nous démontrons uniquement la condition suffisante

La matrice $d \times d$ symétrique positive D étant de rang d, il existe une matrice $d \times d$ B telle que :

$$D = BB^t,$$

posons

$$Y = B^{-1}(X-m),$$

le vecteur Y est un vecteur gaussien et on a

$$E(Y) = B^{-1}E(X-m) = 0$$

■

$$D_Y = B^{-1}D\left(B^{-1}\right)^t = Id$$

les composantes de Y sont indépendante et sont toute de loi gaussienne centrée réduite

$$f_Y(x) = \frac{1}{(2\pi)^{\frac{d}{2}}}\exp\left(-\frac{1}{2}\|x\|^2\right).$$

Pour trouver la loi de X on applique le théorème de transfert

$$\begin{aligned} E(\varphi(X)) &= E(\varphi(m+BY)) \\ &= \int_{\mathbb{R}^d}\varphi(m+By)\,f(y)\,dy \end{aligned}$$

par le changement de variable

$$x = m + By$$

$$E(\varphi(X)) = \int_{\mathbb{R}^d}\varphi(x)\exp\left(\frac{-1}{2}(x-m)^t\left(B^t\right)^{-1}B^{-1}(x-m)\right)dx$$

d'où le résultat

$$E(\varphi(X)) = (2\pi)^{\frac{d}{2}}\det\left(B^{-1}\right)\int_{\mathbb{R}^d}\varphi(x)\exp\left(\frac{-1}{2}(x-m)^t\left(B^t\right)^{-1}B^{-1}(x-m)\right).$$

2.3 Sommes de variables aléatoires indépendantes

Proposition 2.3.1 *Soit $X_1, ..., X_n$ des variables aléatoires de carré intégrable. Alors*

1.

$$Var\left(\sum_{i=1}^{n}\right)=\sum_{i=1}^{n}Var\left(X_i\right)+\sum_{1\leq i,j\leq n}^{n}Cov\left(X_i,X_j\right).$$

2. Si de plus les X_i sont indépendantes , on a

$$Var\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}Var\left(X_i\right).$$

Preuve. On a

$$\begin{aligned}
Var\left(\sum_{i=1}^{n}X_i\right) &= Cov\left(\sum_{i=1}^{n}X_i,\sum_{i=1}^{n}X_i\right)\\
&=\sum_{i=1}^{n}\sum_{j=1}^{n}Cov\left(X_i,X_j\right)\\
&=\sum_{i=1}^{n}Cov\left(X_i,Xi\right)+\sum_{\substack{i,j=1\\ i\neq j}}^{n}Cov\left(X_i,X_j\right)\\
&=\sum_{i=1}^{n}Var\left(X_i\right)+\sum_{\substack{i,j=1\\ i\neq j}}^{n}Cov\left(X_i,X_j\right)
\end{aligned}$$

■

Corollaire 2.3.1 *Si les X_i sont indépendantes et de même loi, de variance σ^2, alors*

$$Var\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{\sigma^2}{n}$$

Preuve. On a d'après la proposition (?), puisque les X_i sont indépendantes

$$\begin{aligned}
Var\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) &= \frac{1}{n^2}Var\left(\sum_{i=1}^{n}X_i\right)\\
&=\frac{1}{n^2}\sum_{i=1}^{n}Var\left(X_i\right)\\
&=\frac{\sigma^2}{n}.
\end{aligned}$$

■

Remarque 2.3.1 *Une propriété importante est que si les v.a X_i sont indépendantes alors quand n est grand $Var\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) \to 0$ et puisque la Var mesure les fluctuation de la v.a auteur de sa moyenne, et donc ici on a disparition des fluctuation aléatoires de $\frac{1}{n}\sum_{i=1}^{n}X_i$ lorsque n tend vers $+\infty$ la v.a est de plus en plus proche de sa moyenne et donc la v.a va se rapprocher d'une variable déterministe.*

Somme de variables aléatoires indépendantes et convolution

Proposition 2.3.2 *Soient X et Y deux variables aléatoires réelles indépendantes de densités respectives f_X et f_Y a. Alors $Z = X + Y$ admet la densité f_Z donnée par*

$$f_Z(z) = \int_{\mathbb{R}} f_X(z-y)\, f_Y(y)\, dy = \int_{\mathbb{R}} f_X(x)\, f_Y(z-x)\, dx.$$

La fonction f_Z est appelée produit de convolution de f_X et de f_Y est souvent notée $f_X * f_Y$.

Preuve. Pour calculer la loi de Z en utilise la fonction test $g \in C_b(\mathbb{R})$ (continue bornée) on a d'après la proposition (?)
si il existe une fonction f positive et d'intégrale égale à 1 telle que

$$E(g(Z)) = \int_{\mathbb{R}} g(z)\ f(z)\, dz$$

alors f est la densité de Z. On a

$$\begin{aligned} E(g(Z)) &= E(g(X+Y)) \\ &= \int_{\mathbb{R}^2} g(X+Y)\, f_{(X,Y)}(x,y)\, dxdy \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} g(X+Y)\, \underbrace{f_X(x)\, f_Y(y)}_{X\ ind\ Y} dxdy, \end{aligned}$$

par un changement de variable $x+y=z;y$

$$\begin{aligned}E\left(g\left(X+Y\right)\right) &= \int_{\mathbb{R}}\int_{\mathbb{R}} g\left(z\right) f_X\left(z-y\right) f_Y\left(y\right) dzdy \\ &= \int_{\mathbb{R}} g\left(z\right)\left(\int_{\mathbb{R}} f_X\left(z-y\right) f_Y\left(y\right) dy\right) dz \\ &= \int_{\mathbb{R}} g\left(z\right) f_Z\left(z\right) dz\end{aligned}$$

et donc la fonction

$$f_Z\left(z\right) = \int_{\mathbb{R}} f_X\left(z-y\right) f_Y\left(y\right) dy$$

est la densité de Z. ■

Nous allons maintenant donner un exemple qui sert à expliquer l'application de cette proposition.

Exemple 2.3.1 (Somme de deux $v.a$ indépendantes de lois normales) *La somme de deux $v.a$ indépendantes de lois normales $N\left(m,\sigma^2\right)$ et $N\left(\mu,\tau^2\right)$ est une $v.a$ normale de loi $N\left(m+\mu,\sigma^2+\tau^2\right)$*

Preuve. Supposons que $m=\mu=0$ et $\sigma^2=\tau^2=1$, on a d'après la proposition (?)

$$\begin{aligned}f_{X}+_{Y}\left(z\right) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{\frac{-(z-x)^2}{2}} e^{\frac{-x^2}{2}} dx \\ &= \frac{1}{2\pi} e^{\frac{-z^2}{2}} \int_{\mathbb{R}} e^{\frac{-x^2}{2}} e^{\frac{-x^2}{2}} e^{zx} dx \\ &= \frac{1}{2\pi} e^{\frac{-z^2}{2}} \int_{\mathbb{R}} e^{-\frac{1}{2}\left(2x^2-2zx\right)} dx \\ &= \frac{1}{2\pi} e^{-\frac{Z^2}{2}} \int_{\mathbb{R}} e^{-\frac{1}{2}\left(\sqrt{2}x-\frac{Z}{\sqrt{2}}\right)^2} e^{\frac{Z^2}{4}} dx \\ &= \frac{1}{2\pi} e^{\frac{-Z^2}{4}} \int_{\mathbb{R}} e^{-\frac{1}{2}\left(\sqrt{2}x-\frac{Z}{\sqrt{2}}\right)^2} dx\end{aligned}$$

on fait le changement de variable $\sqrt{2}x = u$

$$
\begin{aligned}
f_{X} +_{Y}(z) &= \frac{1}{2\pi} e^{-\frac{Z^2}{4}} \underbrace{\frac{1}{\sqrt{2}} \int\limits_{\mathbb{R}} e^{-\frac{1}{2}\left(u-\frac{Z}{\sqrt{2}}\right)^2} du}_{\sqrt{2\pi}\times\frac{1}{\sqrt{2}}=\sqrt{\pi}} \\
&= \frac{1}{2\sqrt{\pi}} e^{-\frac{Z^2}{4}}.
\end{aligned}
$$

■

Exemple 2.3.2 (Somme de deux $v.a$ **indépendantes de loi uniformes)** *La somme de deux variables aléatoires indépendantes de loi uniforme sur* $[-1, 1]$ *est une loi a densité triangulaire.*

Chapitre 3

Convergences et loi des grands nombres

Le but de cette section est :

– Justifier «les lois empiriques du hasard»

– Modèle mathématique : les fréquences empiriques «convergent» vers la probabilité

– Nous allons justifier a posteriori cet axiome : loi des grands nombres

– Mathématiquement : Soient $X_1, ..., X_n$ des $v.a.i.d$ (i.e indépendantes et de même loi) modélisant n résultat d'une même expérience (résultats d'un sondage) la question est quel est le comportement de la moyenne empirique $\frac{X_1+...+X_n}{n}$ quand n tend vers $+\infty$ résultat compliquer que a était résolut par Jacque Bernoulli (1713 Ars conjectandi) : jeu de pile ou face et la Généralisation a était faite par Chebychev , Kolmogrov (20ème siècle).

La question est donc : comment définir la limite de la fonction

$$\omega \to \frac{X_1(\omega) + ... + X_n(\omega)}{n}$$

quand n tend vers l'infini ?

3.1 Convergences de variables aléatoires

Soit $(X_n)_n$ une suite de $v.a$ comment définir « la convergence » de la suite $(X_n)_n$ quand n tend vers l'infini ?

On a déférents notions de convergences non équivalentes :

– On peut décrire la proximité des lois des $v.a$ ($c.v$ en loi).

– On peut décrire la proximité des $v.a$ (suite de $v.a$ définie sur le même espace).

Considérons sur $(\Omega, \mathcal{A}, P)$ des $v.a$ $(X_n)_n$ et X

Définition 3.1.1 *Supposons que $X_n \in L^1, X \in L^1$*

1. *On dit que la suite $(X_n)_n$ converge en moyenne vers X si et seulement si*

$$E\left(|X_n - X|\right) \to 0 \text{ \textbf{\textit{quand}} } \quad n \to +\infty.$$

2. *La suite $(X_n)_n$ converge en probabilité vers X si et seulement si*

$$\forall \varepsilon > 0 : P(|X_n - X| \geq \varepsilon) \underset{n \to +\infty}{\to} 0.$$

3. *La suite (X_n) converge presque sûrement vers X si et seulement si*

$$P\left(\left\{\omega \ / \quad |X_n(\omega) - X(\omega)| \underset{n \to +\infty}{\to} 0\right\}\right) = 1.$$

Ces définitions ne sont pas équivalentes comme nous allons le voir dans les exemples suivants :

Exemple 3.1.1 *Soit $(X_n)_n$ une suite de variables aléatoires de Bernoulli telles que*

$$P(X_n = 1) = \frac{1}{n} \ ; \ P(X_n = 0) = 1 - \frac{1}{n}$$

On a $(X_n)_n$ *tend en probabilité vers* $X = 0$ *car on a pour tout* $\varepsilon \in \left]0, 1\right]$

$$P\left(|X_n| \geq \varepsilon\right) = \frac{1}{n} \underset{n \to +\infty}{\to} 0.$$

Comme on a aussi $(X_n)_n$ *tend en moyenne* $X = 0$ *puisque on a*

$$E\left(X_n\right) = \frac{1}{n} \underset{n \to +\infty}{\to} 0.$$

Exemple 3.1.2 *Soit* (X_n) *une suite de v.a de Bernoulli indépendantes telles que :*

$$P\left(X_n = 1\right) = \frac{1}{n}; P\left(X_n = 0\right) = 1 - \frac{1}{n}.$$

On a pour tout $\varepsilon > 0$,

$$P\left(X_n \geq \varepsilon\right) = \frac{1}{n},$$

et donc la série de terme générale $\frac{1}{n}$ *est divergente. Si l'on pose* $A_n = \{X_n \geq \varepsilon\}$ *, alors* $\sum P(A_n) = +\infty$ *et puisque les* A_n *sont indépendantes, par le théorème de Borelle-Cantelli implique que pour presque tout* ω*, une infinité de* $X_n(\omega)$ *serons supérieurs à* ε *(une infinité de* n *tq* $P(A_n) = 1$*) et donc La suite* $(X_n)_{n\in\mathbb{N}}$ *ne peut pas converger presque surement vers* 0.

Exemple 3.1.3 *Considérons un cas particulier de l'exemple 1. Soit* U *une variable aléatoire uniforme sur* $[0, 1]$*. Posons* $Z_n = 1_{\left\{U \leq \frac{1}{n}\right\}}$*. Alors*

$$P\left(Z_n = 1\right) = P\left(U \leq \frac{1}{n}\right) = \frac{1}{n},$$

$$P\left(Z_n = 0\right) = 1 - \frac{1}{n}.$$

Nous avons déjà vu que la suite (Z_n) *converge en probabilité et en moyenne vers* 0*. Nous* *pouvons montrer de plus que* (Z_n) *converge presque-sûrement vers* 0

$$P(\{\omega : Z_n(\omega) \to 0\}) = 1?.$$

En effet , si ω *est fixé , alors*

$$U(\omega)\rangle 0 \Rightarrow \exists n_0 \in \mathbb{N} \ : \ U(\omega)\rangle \frac{1}{n_0}.$$

Cela entraine

$$Z_n(\omega) = 0 \ , \ \forall n \geq n_0,$$

car

$$Z_n(\omega) = 1_{\left\{U(\omega) \ \langle \frac{1}{n}\right\}} = 0,$$

et donc

$$\forall \omega \leftarrow \Omega : Z_n(\omega) \to 0 \quad i.e\ Z_n \overset{P.p.s}{\to} \ 0 \ .$$

Remarque 3.1.1 *il n'y a pas de contradiction avec l'exemple2 car dans cette exemple les v.a* $(Z_n)_{n\in\mathbb{N}}$ *ne sont pas indépendantes (car elles dépend tous de la v.a* U*). En effet par exemple*

$$P(Z_n = 1; Z_{n+1} = 1) = P(Z_{n+1} = 1) = \frac{1}{n+1}$$

$$P(Z_n = 1)\, P(Z_{n+1} = 1) = \frac{1}{n(n+1)}.$$

Il existe des liens entre les différentes notions de convergences.

Théorème 3.1.1 *Supposons que les v.a* $(X_n)_n$ *et* X *soient intégrables. Alors la convergence en moyenne implique la convergence en probabilité.*

Preuve. On obtient le résultat en appliquant l'inégalité de Markov dans L^1 :

$$P\left(|X_n - X| \geq \varepsilon\right) \leq \frac{E\left(|X_n - X|\right)}{\varepsilon}.$$

■

La réciproque est fausse comme nous allons le voir dans l'exemple suivant :

Exemple 3.1.4 *Soit* $(Y_n)_n$ *une suite de variables aléatoires telles que*

$$P\left(Y_n = n^2\right) = \frac{1}{n};\ P\left(Y_n = 0\right) = 1 - \frac{1}{n}$$

La suite $(Y_n)_n$ *tend en probabilité vers* 0 *car*

$$\forall \varepsilon > 0 \quad P\left(Y_n > \varepsilon\right) = \frac{1}{n}.$$

Par contre La suite $(Y_n)_n$ *ne tend pas en moyenne vers* 0 *puisque* $E\left(Y_n\right) = n$.

Exemple 3.1.5 *Prenons* $(Z_n)_n$ *avec*

$$P\left(Z_n = 1\right) = \frac{1}{n^2};\ P\left(Z_n = 0\right) = 1 - \frac{1}{n^2}.$$

On a convergence en moyenne et en probabilité vers 0. *Soit* $B_n = \{Z_n \geq \varepsilon\}$. *Alors* $\sum_n P\left(B_n\right) < +\infty$, *par le théorème de Borel-Cantelli on a pour presque tout* ω, *un nombre fini au plus de* $Z_n\left(\omega\right)$ *seront supérieurs a* ε.

$$\exists n_0, \forall n \geq n_0 \quad Z_n\left(\omega\right) < \varepsilon,$$

alors la suite (Z_n) *converge presque sûrement vers* 0

Conclusion : Attention, ces notions sont très délicates !

3.1.1 Thérème de convergence dominée

Ce théorème donne le lien entre la convergence presque sûre et la convergence en moyenne (appelé aussi théorème de Lebesgue) :

Théorème 3.1.2 *Si la suite* $(X_n)_n$ *converge presque-sûrement vers* X *et si*

$$\forall n, |X_n| \leq Z \text{ avec } Z \in \mathbb{L}^1$$

alors X_n *et* X *sont dans* $\mathbb{L}^1$, *et*

$$E(|X_n - X|) \underset{n \to +\infty}{\to} 0$$

En particulier , $E(X_n) \to E(X)$.

Attention, la réciproque est fausse dans l'exemple2 $\left(X_n \overset{E}{\to} 0 \text{ et } X_n \overset{P.s}{\nrightarrow} 0\right)$
bien que X_n soit bornée par 1
L'hypothèse de domination est nécessaire comme nous allons le voir dans l'exemple suivant :

Exemple 3.1.6 *Soit* $(T_n)_n$ *une suite de variables aléatoires telles que*

$$P\left(T_n = n^2\right) = \frac{1}{n\sqrt{n}} = 1 - P(T_n = 0).$$

On a $P(T_n \rangle \varepsilon) = \frac{1}{n\sqrt{n}}$ *est le terme générale d'une série C.V, et* (T_n) *converge presque-sûrement vers* 0 *(Théorème de Borel-Cantelli) où* $E(T_n) = \sqrt{n}$ *et donc la suite* $(T_n)_n$ *ne peut pas converger en moyenne.*

On a également le résultat suivant

Théorème 3.1.3 *La convergence presque sûre implique la convergence en probabilité*

Remarque 3.1.2 *La réciproque est fausse voir l'exemple 60*

Relation entre modes de convergences

$$\begin{array}{ccc} \text{convergence presque sûre} & \overrightarrow{\text{si dominée}} & \text{convergence en moyenne} \\ \searrow & & \swarrow \\ & \text{convergence en probabilité} & \end{array}$$

3.2 Loi des grands nombres

Convergence de la suite $M_n = \frac{X_1+...+X_n}{n}$?

On a deux types de résultats :

1. Loi faible des grands nombres : facile a avoir, résultat peu informatif (contrôle d'erreur, estimation de la limité).
2. Loi forte des grands nombres : convergence presque-sûr résultat plus fors preuve plus délicate (presque une convergence ponctuelle).

Ces théorème sont vrais pour une suite de variables aléatoires indépendantes : La suite $(X_n)_n$ est dite indépendante si pour tout n la famille finie $X_1, ..., X_n$ est indépendante.

3.2.1 Loi faible des grands nombres

Théorème 3.2.1 *Soit une suite $(X_n)_n$ de v.a indépendantes, de même loi, intégrables on pose $E(X_1) = m$. Alors*

$$M_n = \frac{X_1 + ... + X_n}{n} \underset{n \to +\infty}{\to} m,$$

en probabilité et en moyenne.

Preuve. Supposons que les X_i sont de carrée intégrable et que

$$\sigma^2 = var(X_1)$$

on a

$$E\left(M_n\right) = E\left(\frac{X_1 + ... + X_n}{n}\right) = m,$$

et puisque les (X_i) indépendantes

$$\begin{aligned} Var\left(M_n\right) &= \frac{1}{n^2}\ Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var\left(X_i\right) \\ &= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}, \end{aligned}$$

et donc par l'inégalité de Bienaymé-Tchebychev on a pour tout $\varepsilon > 0$

$$P\left(\left|M_n - E\left(M_n\right)\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2} Var\left(M_n\right),$$

$$P\left(\left|M_n - m\right| > \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2} \underset{n \to +\infty}{\to} 0,$$

alors M_n converge vers m en probabilité ($i.e$ $M_n \xrightarrow{P} m$), de plus on a

$$E\left(\left|M_n - m\right|\right) \leq \sqrt{E\left(M_n - m\right)^2} = \sqrt{\frac{\sigma^2}{n}} \underset{n \to +\infty}{\to} 0,$$

on obtient M_n converge en moyenne vers m ■

3.2.2 Loi fort des grands nombres

Théorème 3.2.2 *Soit une suite (X_n) de v.a indépendantes, de même loi, intégrables et $E(X_1) = m$ est finie . Alors*

$$M_n = \frac{X_1 + ... + X_n}{n} \underset{n \to +\infty}{\to} m,$$

presque-sûrement.

Preuve. Dans le cas où $E\left(X_1^2\right) < +\infty$. On remplace X_n par $X_n - m$ et on pose $m = 0$. Montrons d'abord que la sous-suite $\left(M_{n^2}\right)_n$ tend $p.s$ vers 0. Par l'inégalité de Bienaymé

-Chebyshev

$$\forall\ q \in \mathbb{N}^*, P\left(|M_{n^2}| \geq \frac{1}{q}\right) \leq \frac{\sigma^2 q^2}{n^2}$$

le terme de la droite est le terme général d'une série convergente. Soit

$$A_{n,q} = \left\{|M_{n^2}| \geq \frac{1}{q}\right\}.$$

Alors $\sum\limits_{n\geq 1} P\left(A_{n,q}\right) < +\infty$. Par le Théorème de Borel-Cantelli

$$P\left(\limsup_n \left(A_{n,q}\right)\right) = 0.$$

Soit

$$N_q = \limsup_n \left(A_{n,q}\right) = \cap_n \cup_{m\geq n} A_{m,q},$$

posons $N = \cup_q N_q$

$$P\left(N\right) \leq \sum_q P\left(N_q\right) = 0 \text{ et } P\left(N^c\right) = 1$$

et comme $N^c = \cap_q \left(N_q\right)^c = \cap_q \cup_n \cap_{m\geq n} A^c_{m,q}$, alors

$$\begin{aligned} \omega \in N^c &\Rightarrow \forall q\ , \exists n_0 \in \mathbb{N}, \forall_m \geq n_0 : \omega \in A^c_{m,q} \\ &\Rightarrow \forall q\ , \exists n_0 \in \mathbb{N}, \forall_m \geq n_0 : |M_{n^2}\left(\omega\right)| \leq \frac{1}{q}. \end{aligned}$$

On en déduit que presque sûrement, (M_{n^2}) converge presque-sûrement vers 0.

Montrons maintenant que la suite (M_n) tend vers 0 $p.s.$

Pour un entier n soit $P\left(n\right)$ le nombre d'entier telque :

$$P\left(n\right)^2 \leq n < \left(P\left(n\right) + 1\right)^2 = P\left(n\right)^2 + 2P\left(n\right) + 1.$$

on a

$$M_n - \frac{P(n)^2}{n} M_{P(n)^2} = \frac{1}{n} \sum_{P=P(n)^2+1}^{n} X_P.$$

On en déduit

$$\begin{aligned} E\left(\left(M_n - \frac{P(n)^2}{n} M_{P(n)^2}\right)^2\right) &= \frac{n - P(n)^2}{n^2}\sigma^2 \\ &\leq \frac{2P(n)+1}{n^2} \leq \frac{2\sqrt{n}+1}{n^2}\sigma^2 \end{aligned}$$

parce que on a $P(n) \leq \sqrt{n}$.

D'après Bienaymé-Chebyshev (encore), on a

$$P\left(\left| M_n - \frac{P(n)^2}{n} M_{P(n)^2} \right| \geq a\right) \leq \frac{2\sqrt{n}+1}{n^2} \frac{\sigma^2}{a^2}$$

Comme la série $\sum_n \frac{2\sqrt{n}+1}{n^2}$ converge, le même raisonnement que ci-dessus pour montrer que

$$M_n - \frac{P(n)^2}{n} M_{P(n)^2} \to 0 \text{ p.s}$$

Par ailleurs $M_{P(n)^2} \to 0$ p.s, et $\frac{P(n)^2}{n} \to 1$. On en déduit que $M_n \to 0$ p.s.

■