2001/2002

PROBABILITES

Chap 1 :

ELEMENTS DE BASE DE L'ETUDE DES PROBABILITES

I. Triplet de probabilité $(\Omega, \mathcal{C}, P)$

. Ω l'ensemble des résultats d'experience

exemples : jet d'une pièce : $\Omega = \{pile, face\}$

jet d'un dé : $\Omega = \{1, 2, 3, 4, 5, 6\}$

$\Omega = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{R}^n \dots$

. $\mathcal{C}$ l'ensemble des evenement, qui doit satisfaire les propriétés suivante.

① $\mathcal{C} \subset P(\Omega)$ ensemble des parties de Ω ie l'ens de tt les ss ensembles de Ω.

ex : $\Omega = \{pile, face\}$, $P(\Omega) = \{\{pile\}, \{face\}, \Omega, \phi\}$

② $\Omega \in \mathcal{C}$. Ω evenement certain

③ Si $A \in \mathcal{C}$, alors $\overline{A} = C_\Omega A \in \mathcal{C}$

$\overline{A}$ evenement contraire de A

④ Si $(A_i)_{i \in I}$, (I fini ou dénombrable) est un ensemble d'éléments de $\mathcal{C}$, ie $A_i \in \mathcal{C}$ $\forall i \in I$, alors $\bigcup_{i \in I} A_i \in \mathcal{C}$

exemples : jet d'1 pièce : $\mathcal{C} = \{\Omega, \phi\}$ présente peu d'intérêt

$\mathcal{C} = \{\Omega, \phi, \{pile\}, \{face\}\} = P(\Omega)$ retenu ds toutes les applicat°

jet d'1 dé : $\mathcal{C} = \{\Omega, \phi\}$ BOF !

$\mathcal{C} = \{\Omega, \phi, \underbrace{\{2,4,6\}}_{\text{evenement "pair"}}, \underbrace{\{1,3,5\}}_{\text{evenement "impair"}}\}$

$\mathcal{C} = P(\Omega) = \{\Omega, \phi, \{1\}, \{2\}, \dots$
$\{1,2\}, \{1,3\}, \dots$
$\{1,2,3\}, \dots$
$\dots\}$

modélisat° plus précise et sut + utile :

en g^al^ on choisit $\mathcal{C} = P(\Omega)$

Propriétés de $\mathcal{C}$: $\boxed{\phi \in \mathcal{C}}$ $\boxed{\bigcap_{i \in I} A_i \in \mathcal{C} \text{ fini ou dénombrable}}$

. P "probabilité" est une application de $\mathcal{C}$ dans $[0,1]$

$$P : \mathcal{C} \longrightarrow [0,1]$$
$$A \longrightarrow P(A)$$

qui verifie les propriétés suivantes :

$P(\Omega) = 1$

$P(\overline{A}) = 1 - P(A)$

$P(\bigcup_{i \in I} A_i) = \sum_{i \in I} P(A_i)$ si les evenements A_i sont disjoints

ie $A_i \cap A_j = \phi$
$i \neq j$

On a alors : $P(\phi) = 0$
si $A \subset B$ alors $P(A) \leq P(B)$
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Vocabulaire : - si $a \in \Omega$, alors $\{a\}$ est un evenement elementaire

- si $\Omega = \bigcup_{i \in I} A_i$ avec $A_i \cap A_j = \phi$, on dit que l'ensemble des A_i est un systeme complet d'evenements
- $(\Omega, \mathcal{C})$ espace probabilisable

$\mathcal{C}$: tribu ou σ-algèbre

$(\Omega, \mathcal{C}, P)$ espace probabilisé

I. Equiprobabilité - Dénombrement

Définition : lorsque | Ω est fini
| on ne veut privilegier aucun evenement (equiprobabilité)
élémentaire

on définit P de la façon suivante :

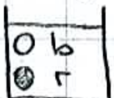
$$P : \mathcal{C} \longrightarrow [0,1]$$
$$A \longmapsto P(A) = \frac{\text{card } A}{\text{card } \Omega} \quad \left(\frac{\text{nb de cas favorables}}{\text{nb de cas possibles}}\right)$$

(d'où dénombrement)

exemples : ① jet d'1 dé : $P(\{2,4,6\}) = \frac{3}{6} = \frac{1}{2}$

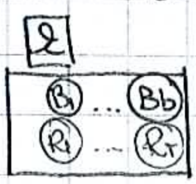
$P(\{2\}) = \frac{1}{6}$

② Tirages avec remise dans une urne à deux catégorie.



$r + b = N$
nb total de boules

Pb : on effectue n tirages (avec remise)
on recherche P_k la probabilité d'avoir k boules blanches sur ces n tirages avec $0 \leq k \leq n$



• Ω est l'ensemble des suites de n éléments pris dans $\{B_1, \dots, B_b, R_1 \dots R_r\}$ avec eventuelle répétition et ordre ex : $n=3$ $\{B_1, R_2, B_1\}$ possible $\neq \{R_2, B_1, B_1\}$

- $\mathcal{C} = \mathcal{P}(\mathcal{E})$ Rq : A_k l'evenement "k boules blanches" $\in \mathcal{C}$.
- $\boxed{P_k = P(A_k) = \dfrac{\text{card } A_k}{\text{card } \mathcal{E}}}$

$\boxed{\text{Card } \mathcal{E} = N^n}$

$\boxed{\text{Card } A_k = ?}$ ⊛ suites de la forme $\underbrace{B_1 B_2 \dots B_1}_{k\ BB} \underbrace{R \dots}_{n-k\ BR}$: $\boxed{b^k r^{n-k}}$

⊛ suites de la forme $\underbrace{R_1 \dots R_2}_{n-k} . \underbrace{B_2 \dots B_1}_{k}$: $\boxed{b^k r^{n-k}}$

etc...

Conclusion : $\boxed{\text{Card } A_k = b^k r^{n-k} \underbrace{C_n^k}_{\substack{\text{nb d'emp.} \\ \text{poss. pr boules blanches}}}}$.

Conclusion : $P_k = P(A_k) = \dfrac{C_n^k b^k r^{n-k}}{N^k N^{n-k}} = C_n^k \left(\dfrac{b}{N}\right)^k \left(\dfrac{r}{N}\right)^{n-k}$

on pose $p = \dfrac{b}{N}$ probabilité du "succès" sur 1 experience

$q = \dfrac{r}{N} = 1 - \dfrac{b}{N} = 1 - p$ probabilité de l'"echec" sur 1 experience.

LOI BINOMIALE $\boxed{P(k \text{ succès sur } n \text{ experiences identiques et indépendantes}) = C_n^k p^k q^{n-k}}$

exercice : n tirages sans remise. $P(A_k) = \dfrac{C_b^k C_r^{n-k}}{C_N^n}$ ~~$k \in \{0, \dots, n\}$~~

loi hypergéométriq peu utilisée.

Rq : $C_n^k = \dfrac{n!}{k!(n-k)!}$ $\boxed{\max(0, n-r) \leq k \leq \min(n, b)}$

III. Probabilites conditionnelles

1. Definition

La probabilité d'un evenement A sachant B (ou proba de A conditionnellement à B) est définie par

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$ en supposant $P(B) \neq 0$

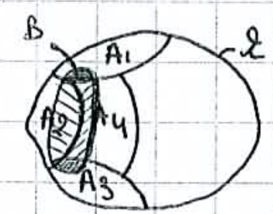
Rq : écriture equivalente : $\boxed{P(A \cap B) = P(A|B)P(B)}$

2. Théorème des probabilités totales

$(A_i)_{i\in I}$ système complet d'évènements (ie $\Omega = \bigcup_{i\in I} A_i$, $A_i \cap A_j = \emptyset$, $i \neq j$)

$P(B) = \sum_{i\in I} P(B|A_i)P(A_i)$

Démo :



$B = B\cap\Omega = B\cap\left(\bigcup_{i\in I} A_i\right)$

$= \bigcup_{i\in I}(B\cap A_i)$

or les A_i sont disjoints donc les $B\cap A_i$ aussi

donc $P\left(\bigcup_{i\in I}(B\cap A_i)\right) = \sum_{i\in I} \underbrace{P(B\cap A_i)}_{P(B|A_i)P(A_i)}$ CQFD.

3. Formule de BAYES.

$P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$

Démo : $P(A|B) = \dfrac{P(A\cap B)}{P(B)}$

$P(B|A) = \dfrac{P(B\cap A) = P(A\cap B)}{P(A)}$

$\Rightarrow P(A|B)P(B) = P(B|A)P(A)$ CQFD.

FIN COURS 1

4. Indépendance

Déf : Deux évènements A et B sont indépendants ssi :

$P(A\cap B) = P(A)P(B)$

Rq : . cette déf n'a de véritable sens que si $P(A)\neq 0$ et $P(B)\neq 0$.

. Dans ces conditions en utilisant la déf des probas conditionnelles on a : $P(A|B) = \dfrac{P(A\cap B)}{P(B)}$

A et B indépendants $\Leftrightarrow$ $P(A|B)P(B) = P(A)P(B)$

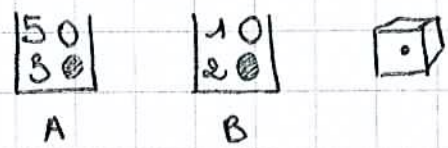
A et B indépendants $\Leftrightarrow$ $P(A|B) = P(A)$

Généralisation : on dit que $(A_i)_{i \in I}$ est une famille d'évènements mutuellement indépendants ssi :

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i) \quad \forall J \subset I .$$

exercice d'application.

ex5 (TD1)



→ 1 ou 3 : on tire 1 boule de A
on la regarde
on la met dans B
on tire ensuite 1 boule dans B

→ 2,4,5,6 : on tire 2 boules ds A (s remise)

1) Probabilité d'avoir 1 boule blanche au premier tirage.

$P_1 = \frac{5}{8}$

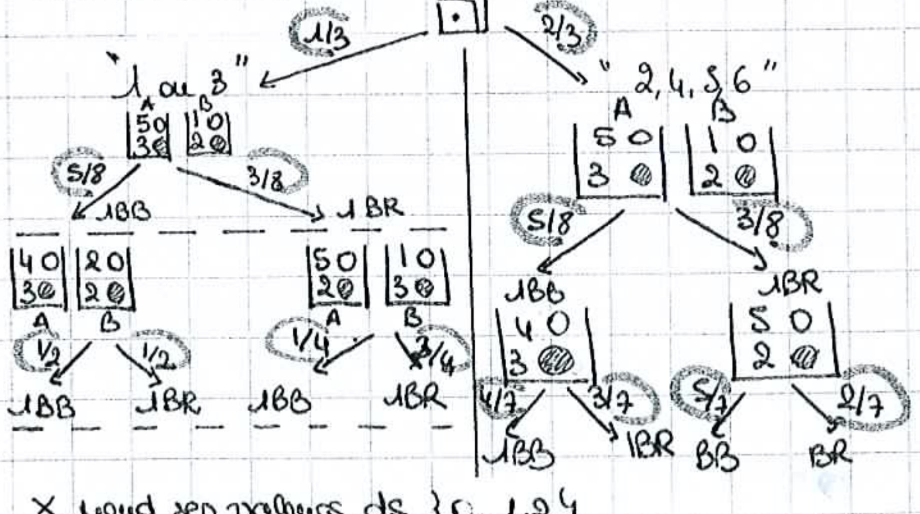
Rq $P_1 = P$("1 ou 3 et la 1ère boule est blanche" ou "2,4,5,6 et la 1ère boule est blanche")

$= P(\text{"1 ou 3 et 1ère BB"}) + P(\text{2,4,5,6 et 1ère BB})$

$= \underbrace{P(\text{1ère BB} \mid \text{1 ou 3})\, P(\text{1 ou 3}) + P(\text{1ère BB} \mid 2,4,5,6)\, P(2,4,5,6)}_{\text{Thm des probas totales}}$

$$\frac{5}{8} \times \frac{1}{3} \quad + \quad \frac{5}{8} \times \frac{2}{3} = \frac{5}{8}$$

2) X est le nb de BB tirées au cours de cette expérience. Déterminer les valeurs possibles de X et les probas. associées.



X prend ses valeurs ds {0, 1, 2}

$$P[X=0] = \left(\frac{1}{3} \times \frac{3}{8} \times \frac{3}{4}\right) + \left(\frac{2}{3} \times \frac{3}{8} \times \frac{2}{7}\right)$$

1 ou 3 — boule rouge | 1 ou 3 — boule rouge | boule rouge et 1 ou 3

$P[X=2] = \left(\frac{1}{3} \times \frac{5}{8} \times \frac{1}{2}\right) + \left(\frac{2}{3} \times \frac{5}{8} \times \frac{4}{7}\right)$

$P[X=1] = 1 - P[X=0] - P[X=2]$

Rq : on a défini la loi de la variable aléatoire X (cf chap 2)

3) à voir.

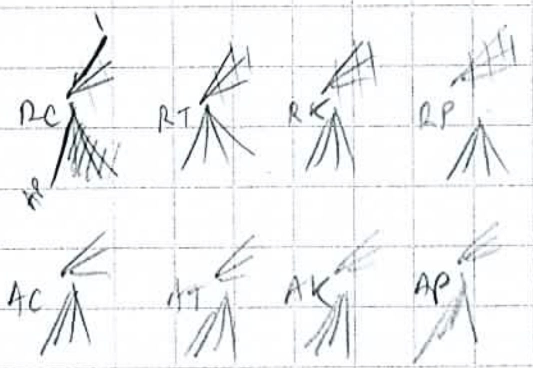
4) Sachant que la 1ère boule est blanche, proba pour que le 4 soit sorti ?

ie : $P(4 \mid 1^{ère}\ BB)$

$P(A|B) = P(B|A)\, P(A)$ BAYES

$P(4 \mid 1^{ère}\ BB) = \dfrac{P(1^{ère}\ BB \mid 4)\, P(4)}{P(1^{ère}\ BB)} = \dfrac{5/8 \cdot 1/6}{5/8} = 1/6$

Rq : $P(4 \mid 1^{ère}\ BB) = P(4)$: évenements indépendants.



$\dfrac{P(A \cap B)}{P(B)} =$

PROBABILITES.

Chap 2: **Variables aléatoires réelles**

I Définition

le triplet de proba $(\Omega, \mathcal{C}, P)$ est parfois "peu utile".

exemple : on jette 2 dés et on s'intéresse à la somme des 2 résultats de ces dés.

$\Omega = \{(1,1), (1,2), \dots, (6,6)\}$

$\mathcal{C} = P(\Omega)$

On a vite envie de travailler avec l'application X définie par :

$X : \Omega \longrightarrow \mathbb{N}$

$\omega = (m, n) \longmapsto m + n$

Il y a 1 lien évident entre Ω (les évènements de $\mathcal{C}$) et les valeurs que prend X. Par exemple, $X = 3$ (la somme vaut 3) correspond à $A = \{(1,2), (2,1)\} \in \mathcal{C}$.

Def : Soit $(\Omega, \mathcal{C}, P)$ un triplet de probabilité qui modélise l'expérience.
Soit $(\Omega', \mathcal{C}')$ un espace probabilisable qui "résume" les "quantités" qui nous intéressent avec $\Omega' \subset \mathbb{R}$ et $\mathcal{C}'$ construit comme l'ensemble des réunions et intersections des intervalles de Ω'.

On dit que $\mathcal{C}'$ est la tribu des Boréliens de $\mathbb{R}$.

X est une variable aléatoire (va) si c'est une application de Ω dans Ω' qui possède la propriété de mesurabilité (cf cours intégration) :

$\forall \underbrace{(a,b)}_{[a,b] \text{ ou }]a,b[\text{ ou} \dots} \in \mathcal{C}', \{\omega / X(\omega) \in (a,b)\} \in \mathcal{C}$

Rq : cette propriété sera admise ds toutes les applications.

II Loi d'une variable aléatoire.

On distingue 3 classes de variables aléatoires qui sont rencontrées dans la quasi totalité des applications

① les va discrètes

X est une va discrète si l'ensemble des valeurs possibles de X (ie $\{X(\omega), \omega \in$ ...) est fini ou dénombrable.

On notera alors $\{x_i, i \in I\}$ cet ensemble de valeurs possibles.

La loi de X est définie par : ① $\{x_i, i \in I\}$ ② $P[X = x_i]$ $i \in I$ ensemble des probas associées telles que $\sum_{i \in I} P[X = x_i] = 1$

Rq : $P[X \in \Delta] = \sum_{x_i \in \Delta} P[X = x_i]$

- Exemple : X est le résultat d'un jet de dé.

② Les va continues

X est une va continue si l'ensemble des valeurs possibles de X est infini non dénombrable et $P[X=x_i]=0\ \forall x_i$

ex : Taille, poids ....

La loi de X est définie par ① $\{X(\omega) \mid \omega \in \Omega\}$ qui est en g^al une réunion d'intervalles

② une densité de probabilité $p(x)$, $x\in\mathbb{R}$ tq

$$\begin{cases} \forall x\ p(x)\geq 0 \\ P[X\in\Delta] = \int_\Delta p(x)\,dx \quad \Delta\subset\mathbb{R} \end{cases}$$

Rq : ⊛ $P[X\in\mathbb{R}]=1=\int_\mathbb{R} p(x)\,dx$ ms on peut avoir $p(x)>1$ pr certains x

⊛ $P[X\in[x,x+dx[] = \int_x^{x+dx} p(u)\,du \underset{dx\text{ "petit"}}{\approx} p(x)\int_x^{x+dx} du = p(x)\,dx$

$p(x)dx$ s'interprète comme la proba d'appartenir à 1 petit intervalle.

d'où $p(x) = \dfrac{P[X\in[x,x+dx[]}{dx}$

FIN cours 2

Donner un exemple = loi uniforme

③ Les va mixtes

On dit que X est une va mixte si l'ens. des valeurs possibles de X est la réunion de 2 ensembles, le premier est un ensemble fini ou dénombrable $\{x_i, i\in I\}$ avec $P[X=x_i]>0\ \forall i\in I$, le second est un ens. infini non dénombrable E avec $\underset{\forall x_i}{P[X=x_i]=0}$

$\{X(\omega)\mid\omega\in\Omega\} = E\cup\{x_i, i\in I\}$

ex : tension aux bornes d'1 voltmètre

$\{x_i, i\in I\} = \{V_s, -V_s\}$ V_s tension de saturation.

$E =]-V_s, +V_s[$

la loi de X est définie par ① $\{x_i, i\in I\}$ et $P[X=x_i]\ \forall i\in I$

② E et la densité de proba associée.

avec $P[X\in\Delta] = \int_\Delta p(x)\,dx + \sum_{x_i\in\Delta} P[X=x_i]$

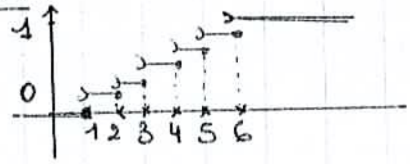
III. Fonction de répartition

La fonction de répartition d'une variable aléatoire réelle X est définie par $F(x) = P[X<x]$.

Propriétés : F est une fonction croissante tq $\lim_{x \to -\infty} F(x) = 0$
$\lim_{x \to +\infty} F(x) = 1$

- X va discrète à valeurs dans $\{x_i, i \in I\}$.

ex du jet de dés :



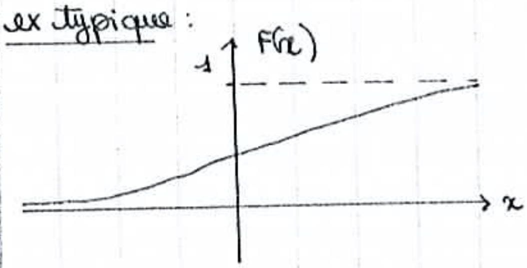
F(x) est une fonction en escalier dont les "sauts" se font aux observations x_i et sont d'amplitude p_i.

- X va continue de densité p(x)

$P[X<x] = P[X \in]-\infty, x[] = \int_{-\infty}^{x} p(u)du$

C'est une fonction continue (croissante telle que $\lim_{-\infty} F(x) = 0$ et $\lim_{+\infty} F(x) = 1$)

ex typique :



- X va mixte : F(x) est une fonction continue par morceaux, qui peut posséder des sauts aux points x_i tels que $P[X=x_i] > 0$.

PROPRIETE FONDAMENTALE

La fonction de répartition caractérise une va.

Application : souvent, pour déterminer la loi de X, on déterminera F(x).

Rq : cas continu $p(x) = F'(x)$

IV Exemples Fondamentaux.

voir poly.

ANNEXE 2 : Tables de lois

LOIS DE PROBABILITÉ DISCRÈTES

m : moyenne σ^2 : variance **F. C.** : fonction caractéristique

$P_k = P[X = k]$ $P_{1...m} = P[X_1 = k_1, ..., X_m = k_m]$

LOI	Probabilités	m	σ^2	F. C.
Uniforme	$P_k = \frac{1}{n}$ $k \in \{1, ..., n\}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\dfrac{e^{it}(1-e^{itn})}{n(1-e^{it})}$
Bernoulli	$P[X=1] = p$ $P[X=0] = q$ $p \in [0,1] \quad q = 1-p$	p	pq	pe^{it}
Binomiale $B(n,p)$	$P_k = C_n^k p^k q^{n-k}$ $p \in [0,1] \quad q = 1-p$ $k \in \{0, 1, ..., n\}$	np	npq	$(pe^{it}+q)^n$
Binomiale négative	$P_k = C_{n+k-1}^{n-1} p^n q^k$ $p \in [0,1] \quad q = 1-p$ $k \in \mathbb{N}$	$n\dfrac{q}{p}$	$n\dfrac{q}{p^2}$	$\left(\dfrac{p}{1-qe^{it}}\right)^n$
Multinomiale	$P_{1...m} = \frac{n!}{k_1!...k_m!} p_1^{k_1} ... p_m^{k_m}$ $p_j \in [0,1] \quad q_j = 1 - p_j$ $k_j \in \{0, 1, \ldots, n\}$ $\sum_{j=1}^{m} k_j = n \quad \sum_{j=1}^{m} p_j = 1$	np_j	Variance : $np_j q_j$ Covariance : $-np_j p_k$	$\left(\sum_{j=1}^{m} p_j e^{it}\right)^n$
Poisson $P(\lambda)$	$P_k = e^{-\lambda}\dfrac{\lambda^k}{k!}$ $\lambda > 0 \quad k \in \mathbb{N}$	λ	λ	$e^{\lambda(e^{it}-1)}$
Géométrique	$P_k = pq^{k-1}$ $p \in [0,1] \quad q = 1-p$ $k \in \mathbb{N}^*$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$	$\dfrac{pe^{it}}{1-qe^{it}}$

LOIS DE PROBABILITÉ CONTINUES

m : moyenne σ^2 : variance

F. C. : fonction caractéristique

LOI	**Densité de probabilité**	**m**	σ^2	**F. C.**
Uniforme	$f(x) = \frac{1}{b-a}$ $x \in]a, b[$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\dfrac{e^{itb} - e^{ita}}{it(b-a)}$
Gamma $\Gamma(\theta, \nu)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\theta x} x^{\nu-1}$ $\theta > 0,\ \nu > 0$ $x \geq 0$	$\dfrac{\nu}{\theta}$	$\dfrac{\nu}{\theta^2}$	$\dfrac{1}{\left(1 - i\frac{t}{\theta}\right)^\nu}$
Première loi de Laplace	$f(x) = \frac{1}{2} e^{-\lvert x \rvert}$	0	2	$\dfrac{1}{1+t^2}$
Normale $N(m, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$	m	σ^2	$e^{imt - \frac{\sigma^2 t^2}{2}}$
Khi$_2$ χ^2_ν $\Gamma\left(\frac{1}{2}, \frac{\nu}{2}\right)$	$f(x) = k e^{-\frac{x}{2}} x^{\frac{\nu}{2}-1}$ $k = \dfrac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}$ $\nu \in \mathbb{N}^*,\ x \geq 0$	ν	2ν	$\dfrac{1}{(1-2it)^{\frac{\nu}{2}}}$
Cauchy $c_{\lambda,\alpha}$	$f(x) = \dfrac{1}{\pi\lambda\left(1 + \left(\frac{x-\alpha}{\lambda}\right)^2\right)}$ $\lambda > 0,\ \alpha \in \mathbb{R}$	(-)	(-)	$e^{i\alpha t - \lambda\lvert t \rvert}$
Bêta	$f(x) = k x^{a-1} (1-x)^{b-1}$ $k = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ $a > 0,\ b > 0,\ x \in]0, 1[$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	(*)

(-) : n'existe pas

(*) : trop compliquée pour être utilisée !

V Espérance mathématique.

(1) Définition

X étant une va réelle, on définit l'espérance mathématique de $\alpha(X)$ (α étant une fonction $\mathbb{R} \to \mathbb{R}$) par :

$$E[\alpha(X)] = \begin{cases} \text{X va discrète : } \sum_{i \in I} \alpha(x_i) P[X = x_i] \\ \text{X va continue : } \int_{\mathbb{R}} \alpha(u) p(u) du \\ \text{X va mixte : } \sum_i + \int_{\mathbb{R}} \dots \end{cases}$$

(2) Propriétés.

- $E[\text{cste}] = \text{cette cste}$

Plus généralement E [fonction déterministe] = cette fonction déterministe
(non aléatoire)

ex : $E[\cos(t)] = \cos(t)$

$E[\cos(t+\phi)] = \int \cos(t+\varphi) \, p(\varphi) \, d\varphi$.

ϕ uniforme sur $[0, 2\pi[$. densité de ϕ

- linéarité :

$E[aX+b] = a E[X] + b$. (vient de la linéarité de $\sum$ et $\int$)

(3) Exemples fondamentaux

- les moments non centrés. $m_n = E[X^n]$ $n \in \mathbb{N}$

$n = 0$ $m_0 = 1$

$n = 1$ $m_1 = E[X]$ s'appelle la moyenne de X

on verra à la fin du cours que si $x_1 \dots x_n$ sont des réalisations / observations de X

que $\underbrace{\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i}_{\text{moyenne arithmétiq}} \xrightarrow[n \to \infty]{} E[X]$.

- les moments centrés : $\mu_n = E[(X - E(X))^n]$

$n = 0$ $\mu_0 = 1$

$n = 1$ $\mu_1 = E(X - E(X)] = E[X] - E[\underbrace{E(X)}_{\text{cste}}] = E[X] - E(X) = 0$.

$n = 2$ $\mu_2 = E[X - E(X))^2]$ variance de X on verra plus tard que $\underbrace{\frac{1}{n}\sum_{i=1}^{n}[x_i - \bar{x}]^2}_{\substack{\text{moyenne arithmétiq}\\ \text{des écarts}\\ \text{quadratiques}}} \xrightarrow[n\infty]{} \mu_2$

μ_2 mesure la dispersion autour de la valeur moyenne.

$\sqrt{\mu_2}$ s'appelle l'écart type de X.

propriétés de la variance :

- on note souvent $\mathrm{Var}\,X = E[(X - E(X))^2]$
- $\mathrm{Var}\,X = E[X^2 - 2X E(X) + (E(X))^2] = E[X^2] - 2E[X]E(X) + E(X)^2$

$\mathrm{Var}\,X = E[X^2] - E(X)^2$ utile pour le calcul.

- $\mathrm{Var}(aX+b) = \mathrm{Var}(aX) = a^2\,\mathrm{Var}(X)$

- la fonction caractéristique : $\underset{\in\mathbb{C}}{\varphi_X(t)} = E[e^{itX}]$
 caractérise la loi.

Rq :. dans le cas continu on a $\varphi_X(t) = \int_{\mathbb{R}} e^{itx} p(x)\,dx$.

s'appelle la transformée de Fourier de $p(x)$. On a de bonnes tables qui donnent $p(x) = \frac{1}{2i\pi}\int_{\mathbb{R}} \varphi_X(t)\, e^{-itx}\,dt$

- $\varphi_X(0) = 1$, $|\varphi_X(t)| \leqslant 1$

FIN COURS 3

④ Exemples de calculs.

ex 1 : $X \sim P(\lambda)$ (suit la loi de Poisson de paramètre λ)

$P[X=k] = \frac{\lambda^k}{k!} e^{-\lambda}$ $k \in \mathbb{N}$

$E(X) = \sum_{k\in\mathbb{N}} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda}\underbrace{\sum_{k=1}^{+\infty}\frac{\lambda^{k-1}}{(k-1)!}}_{e^{\lambda}} = \lambda$.

la moyenne de X est λ.

$\mathrm{Var}\,X = E(X^2) - \underbrace{E(X)^2}_{\lambda^2}$

$E(X^2) = \sum_{k=0}^{\infty} \underbrace{k^2}_{k(k-1)+k} \frac{\lambda^k}{k!} e^{-\lambda} = \underbrace{\sum_{k=2}^{+\infty}\frac{\lambda^k}{(k-2)!}e^{-\lambda}}_{\lambda^2 e^{-\lambda}\sum_{k=2}^{+\infty}\frac{\lambda^{k-2}}{(k-2)!} = \lambda^2} + \underbrace{\sum_{k=1}^{+\infty}\frac{\lambda^k}{(k-1)!}e^{-\lambda}}_{\lambda}$

$= \lambda^2 - \lambda^2 + \lambda = \lambda$.

$\phi_X(t) = E[e^{itX}] = \sum_{k\in\mathbb{N}} e^{itk}\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k\in\mathbb{N}}\frac{(\lambda e^{it})^k}{k!}$

$= \exp[-\lambda + \lambda e^{it}] = \exp[\lambda(e^{it}-1)]$. voir utilité + tard

ex 2 : $X \sim N(m, \sigma^2)$ ie $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x-m)^2}{2\sigma^2}$ $x \in \mathbb{R}$

$$E[X] = \int_{\mathbb{R}} x f(x)\,dx = \int_{\mathbb{R}} \frac{x}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx$$

$u = \frac{x-m}{\sigma}$
$x = \sigma u + m$
$dx = \sigma du$

$$\stackrel{=}{} \int_{\mathbb{R}} \frac{\sigma u + m}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \frac{\sigma}{\sqrt{2\pi}} \underbrace{\int_{\mathbb{R}} u e^{-\frac{u^2}{2}} du}_{\substack{=0 \\ \text{(impaire)}}} + m \underbrace{\int \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du}_{\substack{f(x) \text{ avec} \\ \sigma = 1 \text{ et } m = 0}}$$

$\boxed{\text{Loi normale centrée réduite}}$

$E[X] = m$

$\mathrm{Var}\, X = E[X^2] - E[X]^2$

$$E[X^2] = \int_{\mathbb{R}} x^2 f(x)\,dx = \int_{\mathbb{R}} x^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x-m)^2}{2\sigma^2} dx$$

$$\stackrel{u = \frac{x-m}{\sigma}}{=} \int_{\mathbb{R}} (\sigma u + m)^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

$$= \sigma^2 \underbrace{\int_{\mathbb{R}} \frac{u^2}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du}_{\text{par parties} \Rightarrow 1} + 2m\sigma \int_{\mathbb{R}} \frac{u}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du + m^2 \underbrace{\int \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du}_{1}$$

$= \sigma^2 + m^2$

$\mathrm{Var}\, X = \sigma^2$

Rq : $\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$ <u>loi normale centrée réduite</u>

$$\phi_X(t) = E[e^{itX}] = \int_{\mathbb{R}} e^{itx} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] dx$$

$$\phi_X(t) \underset{\substack{\text{après qq} \\ \text{calculs} \\ \text{fastidieux}}}{=} \exp\left[imt - \frac{\sigma^2}{2} t^2\right]$$

VI Changements de variable

$X \in \{0, 1\} \longrightarrow \oplus \longrightarrow Y = X + N$, avec N entrant en $\oplus$

N : Bruit typiquement supposé $X \sim N(m, \sigma^2)$

Problème : étant donnée une variable aléatoire réelle X, on cherche à déterminer la loi de $Y = g(X)$ où g est une fonction de $\mathbb{R}$ ds $\mathbb{R}$.

1er cas X est une va à valeurs dans $\{x_i, i \in I\}$ I fini ou dénombrable alors $Y = g(X)$ est aussi une va discrète à valeurs ds

$\{y_j, j \in J\} = \{g(x_i), i \in I\}$. On a alors :

$$P[Y = y_j] = \sum_{i / g(x_i) = y_j} P[X = x_i]$$

Exemple. $X \sim P(\lambda)$ ie $P[X = i] = \frac{\lambda^i}{i!} e^{-\lambda}$ $i \in \mathbb{N}$

Quelle est la loi de $Y = (X-2)^2$?

Y est à valeurs dans $\{j^2, j \in \mathbb{N}\}$

$P[Y = j^2] = P[(X-2)^2 = j^2] = P[X-2 = j \text{ ou } X-2 = -j] = P[X = 2+j \text{ ou } X = 2-j]$

- si $2+j = 2-j$ ie $j = 0$. $\quad P[Y=0] = P[X=2] = \frac{\lambda^2}{2} e^{-\lambda}$
- si $j \geq 1$ on a $2+j \neq 2-j$ donc $P[Y = j^2] = P[X = 2+j] + P[X = 2-j]$

$$P[X = 2+j] = \frac{\lambda^{2+j}}{(2+j)!} e^{-\lambda} \quad \text{car } 2+j \in \mathbb{N} \ \forall j \in \mathbb{N}$$

$$P[X = 2-j] = \begin{cases} j=1 & P[X=1] = \lambda e^{-\lambda} \\ j=2 & P[X=0] = e^{-\lambda} \\ j \geq 3 & P[X=2-j] = 0 \end{cases}$$

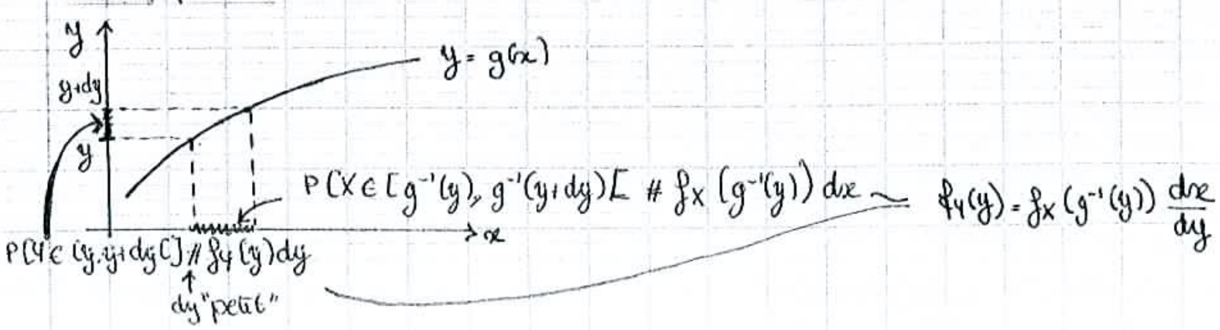
$$\boxed{\begin{aligned} P[Y=0] &= \frac{\lambda^2}{2} e^{-\lambda} \\ P[Y=1] &= \frac{\lambda^3}{6} e^{-\lambda} + \lambda e^{-\lambda} \\ P[Y=4] &= \frac{\lambda^4}{24} e^{-\lambda} + e^{-\lambda} \\ P[Y=j^2] &= \frac{\lambda^{2+j}}{(2+j)!} e^{-\lambda} \quad j \geq 3. \end{aligned}}$$

2ème cas : X est une va continue à valeurs dans un ouvert $O_X \subset \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ application bijective de O_X dans $O_Y \subset \mathbb{R}$ différentiable ainsi que son inverse g^{-1}

alors $Y = g(X)$ est une va continue de densité :

$$f_Y(y) = f_X(g^{-1}(y)) \underbrace{\left| \frac{dx}{dy} \right|}_{\text{Jacobien de la transformation}}$$

Idée de preuve:



Exemple : $X \sim N(m, \sigma^2)$

$Y = aX + b \quad a \neq 0$ loi de Y ?

$y = g(x)$ est une transformation bijective $\mathbb{R} \to \mathbb{R}$

densité de X : $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x-m)^2}{2\sigma^2} \quad x \in \mathbb{R}$

densité de Y : $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left(\frac{y-b}{a} - m\right)^2\right) \left|\frac{dx}{dy}\right|$

$$f_Y(y) = \frac{1}{|a|} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}\left(\frac{y-(b+am)}{a}\right)^2\right]$$

on pose $M = b + am$

$\Sigma^2 = a^2\sigma^2$

Y suit une loi normale $Y \sim N(b+am, a^2\sigma^2)$

Rq : $E[Y] = E[aX+b] = aE[X] + b = am + b$

$Var[Y] = Var[aX+b] = Var[aX] = a^2 Var[X] = a^2\sigma^2$

3ème cas : $Y = g(X)$ est bijectif par morceaux et les hypothèses précédentes (différentiabilité ...) sont vérifiées sur chaque morceau. Alors la densité de Y est la somme des contributions de chaque morceau.

Exemple : $X \sim N(0,1) \quad f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$Y = X^2$ loi de Y ?

bijection 1 : $\mathbb{R}^+ \to \mathbb{R}^+$
$x \to y = x^2 \Leftrightarrow x = \sqrt{y} \quad g_1(y) = \frac{1}{\sqrt{2\pi}} e^{-y/2} \left|\frac{1}{2\sqrt{y}}\right| \quad y \in \mathbb{R}^+$

bijection 2 : $\mathbb{R}^- \to \mathbb{R}^+$
$x \to y = x^2 \Leftrightarrow x = -\sqrt{y} \quad g_2(y) = \frac{1}{\sqrt{2\pi}} e^{-y/2} \left|-\frac{1}{2\sqrt{y}}\right| \quad y \in \mathbb{R}^+$

la densité de Y est $f_Y(y) \begin{cases} = \frac{1}{\sqrt{2\pi y}} e^{-y/2} & y > 0 \\ 0 & y < 0 \end{cases}$

noté : $f_Y(y) = \frac{1}{\sqrt{2\pi y}} e^{-y/2} \mathbb{1}_{\mathbb{R}^+}(y)$

Remarque : on dit que Y suit une loi du χ^2 à 1 degré de liberté et on note $Y \sim \chi^2_1$

FIN COURS 4

Couples de Variables Aléatoires Réelles

1. Définition

Soit (Ω, C, P) un espace probabilisé et (Ω', C') un espace probabilisable avec $\Omega' \subset \mathbb{R}^2$ et C' construit à partir des réunions et intersections finies ou dénombrables des pavés $(a, b) \times (c, d)$ de $\mathbb{R}^2$. Un couple (X, Y) de variables aléatoires réelles est une application mesurable de Ω dans Ω'.

On notera $P[(X, Y) \in \Delta], \Delta \subset \mathbb{R}^2$, la probabilité que le couple (X, Y) prenne ses valeurs dans Δ.

Dans la plupart des applications, on rencontre les couples de variables aléatoires suivants :

1-1 Les couples de va discrètes

La loi du couple (X, Y) est définie par l'ensemble des valeurs possibles du couple (qui est un ensemble fini ou dénombrable) noté $\{(x_i, y_j), i \in I, j \in J\}$ et par les probabilités associées $p_{ij} = P[X = x_i, Y = y_j]$, $i \in I, j \in J$ telles que $p_{ij} \geq 0$ et $\sum_{i,j} p_{ij} = 1$.

1-2 Les couples de va continues

La loi du couple (X, Y) est définie par l'ensemble des valeurs possibles du couple (qui est un ensemble infini non dénombrable), en général une réunion d'intervalles de $\mathbb{R}^2$, et par une densité de probabilité $f(x, y)$ telle que

$$\begin{aligned} f(x, y) &\geq 0 \\ \int\int_{\mathbb{R}^2} f(x, y) dx dy &= 1 \end{aligned}$$

et de manière plus générale

$$P\left[(X,Y)\in\Delta\right]=\int\int_{\Delta}f(x,y)dxdy$$

Remarque : signification de $f(x,y)$

2. Fonction de Répartition

Définition

$$F(x,y)=P\left[X<x,Y<y\right]$$

Remarques :

- C'est une fonction étagée lorsque (X,Y) est un couple de va discrètes
- C'est une fonction continue lorsque (X,Y) est un couple de va continues avec

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dudv$$

d'où

$$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y}$$

3. Lois Marginales

Les lois marginales d'un couple (X, Y) sont les lois de X et de Y telles que

- **Cas discret**

$$P[X = x_i] = p_{i.} = \sum_{j \in J} p_{ij}$$

$$P[Y = y_j] = p_{.j} = \sum_{i \in I} p_{ij}$$

- **Cas continu**

densité de X : $\quad f(x, .) = \int_{\mathbb{R}} f(x, y) dy$

densité de Y : $\quad f(., y) = \int_{\mathbb{R}} f(x, y) dx$

ex p128

4. Lois Conditionnelles

Les lois conditionnelles d'un couple (X, Y) sont les lois de $X|\, Y = y$ et de $Y|\, X = x$ telles que

- **Cas discret**

$$P[X = x_i|\, Y = y_j] = \frac{p_{ij}}{p_{.j}}$$

$$P[Y = y_j|\, X = x_i] = \frac{p_{ij}}{p_{i.}}$$

- **Cas continu**

densité de $X|\, Y = y$: $\quad f(x|\, y) = \dfrac{f(x, y)}{f(., y)}$

densité de $Y|\, X = x$: $\quad f(y|\, x) = \dfrac{f(x, y)}{f(x, .)}$

5. Indépendance

On dit que X et Y sont des variables aléatoires indépendantes si

$$P[X \in \Delta, Y \in \Delta'] = P[X \in \Delta] P[Y \in \Delta'], \forall\Delta, \forall\Delta'$$

De manière équivalente, X et Y sont des variables aléatoires indépendantes si

- **Cas discret**

$$p_{ij} = p_{i.}p_{.j} \qquad \forall i \in I, \forall j \in J$$

ou

$$P[X = x_i | Y = y_j] = p_{i.} \qquad \forall i \in I, \forall j \in J$$

- **Cas continu**

$$f(x, y) = f(x, .)f(., y) \qquad \forall x, \forall y$$

ou

$$f(x| y) = f(x, .), \qquad \forall x, \forall y$$

Propriété : si X et Y sont des variables aléatoires et α et β sont des applications continues de $\mathbb{R}$ dans $\mathbb{R}$, alors $\alpha(X)$ et $\beta(Y)$ sont des variables aléatoires indépendantes. La réciproque est vraie si α et β sont des applications bijectives. Par contre, dans le cas où α et β ne sont pas bijectives, la réciproque est fausse. On vérifiera par exemple que le couple $(X\,Y)$ de densité

$$f(x, y) = \begin{cases} \frac{1}{4}(1 + xy) \text{ si } |x| < 1 \text{ et } |y| < 1 \\ 0 \text{ sinon} \end{cases}$$

est tel que X^2 et Y^2 sont indépendantes alors que X et Y ne le sont pas.

Début cours 5

PROBABILITES

Chap 3

Couples de variables aléatoires réelles

I. Définition

$(\mathcal{E}, \mathcal{C}, P)$ un triplet de probabilités et $(\mathcal{E}', \mathcal{C}')$ un espace probabilisable avec $\mathcal{E}' \subset \mathbb{R}^2$ et $\mathcal{C}'$ est construit par réunion et intersection finies ou dénombrables des pavés $(a,b) \times (c,d)$ de $\mathbb{R}^2$. Un couple (x,y) de va réelles est une application de $\mathcal{E}$ dans $\mathcal{E}'$ qui possède la propriété de mesurabilité.

On notera $P[(X,Y) \in \Delta]$, $\Delta \in \mathbb{R}^2$, la probabilité que le couple (X,Y) prenne ses valeurs dans Δ.

Les principaux couples de va sont les suivants :

1. Les couples de va discrètes

définis par $\{(x_i, y_j)\ i \in I, j \in J\}$ des valeurs possibles de (X,Y) et par $p_{ij} = P[X = x_i, Y = y_j]$ avec $p_{ij} > 0$ et $\sum_{\substack{i\in I \\ j \in J}} p_{ij} = 1$

2. Les couples de va continues

définis par l'ensemble des valeurs prises par le couple (en général une $\cup$ ou $\cap$ de pavés) et par une densité de probabilité $p(x,y)$ telle que $p(x,y) \geq 0$, $P[(X,Y) \in \Delta] = \iint_\Delta f(x,y)\,dx\,dy$

d'où : $P[(X,Y) \in \mathbb{R}^2] = \iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = 1$

$P[X \in [x, x+dx[, Y \in [y, y+dy[] \underset{\substack{dx \text{ petit} \\ dy \text{ petit}}}{\approx} f(x,y)\,dx\,dy$

II Fonction de répartition

La fonction de répartition d'un couple (X,Y) est définie par

$F_{X,Y}(u,v) = P[X < u \text{ "et" } Y < v]$

C'est une fonction étagée (constante sur des pavés) dans le cas discret

continue dans le cas continu avec $F_{XY}(u,v) = \int_{-\infty}^{u}\int_{-\infty}^{v} p_{XY}(x,y)\,dx\,dy$

d'où $p_{X,Y}(u,v) = \frac{\partial}{\partial u}\frac{\partial}{\partial v}[F_{XY}(u,v)] = \frac{\partial^2 F}{\partial u \partial v}(u,v)$

III Lois marginales et lois conditionelles.

Les lois marginales du couple (X,Y) sont les lois de X et de Y définies par:

CAS DISCRET $P[X=x_i] = P[\bigcup_j \{X=x_i, Y=y_j\}] = \sum_j P(X=x_i, Y=y_j]$

Notation: $p_{i\cdot} = \sum_{j\in J} p_{ij}$

de même $p_{\cdot j} = \sum_{i\in I} p_{ij}$

CAS CONTINU $\underbrace{f(x,\cdot)}_{\text{densité de } X} = \int_{\mathbb{R}} \underbrace{f(x;y)}_{\text{densité du couple}(X,Y)} dy$ et $f(\cdot,y) = \int_{\mathbb{R}} f(x,y)dx$.

(Preuve à voir)

Les lois conditionelles sont définies par:

CAS DISCRET $P[X=x_i | Y=y_j] = \frac{p_{ij}}{p_{\cdot j}}$ $\left(= \frac{P(A\cap B)}{P(B)}\right)$

$P[Y=y_j | X=x_i] = \frac{p_{ij}}{p_{i\cdot}}$

CAS CONTINU Les lois de $Y|X=x$ et $X|Y=y$ sont des lois continues de densités:

$f(y|x) = \frac{f(x,y)}{f(x,\cdot)}$ et $f(x|y) = \frac{f(x,y)}{f(\cdot,y)}$

⚠ $f(y|x)\,dy \underset{dy \text{ petit}}{\#} P[Y\in[y,y+dy[\,|X=x] = \frac{\overset{=0}{P[Y\in[y,y+dy[, X=x]}}{\underset{=0}{P[X=x]}}$

indétermination.

donc pour résoudre ce pb on utilise $f(y|x)dy = \lim_{dx\to 0} \frac{P[Y\in[y,y+dy[, X\in[x,x+dx[]}{P[X\in[x,x+dx[]}$

etc..

IV. Indépendance

On dit que X et Y sont deux va indépendantes si $P[X\in\Delta, Y\in\Delta'] = P(X\in\Delta]P(Y\in\Delta']$
$\forall \Delta \in \mathbb{R}$ $\forall \Delta' \in \mathbb{R}$

(à rapprocher de $P(A\cap B) = P(A)P(B)$)

Cette définition peut se simplifier de la façon suivante :

INP ENSEEIHT

Probabilités et Statistiques
Complément de Cours : Densités Conditionnelles

Cette note présente la technique permettant d'obtenir la densité de $Y|X = x$ dans le cas de variables aléatoires continues (moyennant certaines hypothèses comme intervertion des limites et des intégrales ...).

On pose

$$P[Y \in \Delta | X = x] = \lim_{h \to 0} \frac{P[Y \in \Delta, X \in [x, x+h[]}{P[X \in [x, x+h[]}$$

c'est-à-dire

$$\begin{aligned} P[Y \in \Delta | X = x] &= \lim_{h \to 0} \frac{\iint_{\Delta \times [x,x+h[} f(x,y)dxdy}{\iint_{\mathbb{R} \times [x,x+h[} f(x,y)dxdy} \\ &= \lim_{h \to 0} \frac{\int_{\Delta} \left[\int_{[x,x+h[} f(x,y)dx\right] dy}{\int_{\mathbb{R}} \left[\int_{[x,x+h[} f(x,y)dx\right] dy} \end{aligned}$$

et on applique le théorème de la moyenne pour y fixé :

$$\int_{[x,x+h[} f(x,y)dx = hf(x + \theta_y h, y) \qquad \theta_y \in [0,1[$$

On en déduit (en supposant $f(x,.) \neq 0$) :

$$P[Y \in \Delta | X = x] = \lim_{h \to 0} \frac{\int_{\Delta} [hf(x + \theta_y h, y)] dy}{\int_{\mathbb{R}} [hf(x + \theta_y h, y)] dy} = \frac{\int_{\Delta} f(x,y)dy}{f(x,.)}$$

c'est-à-dire :

$$P[Y \in \Delta | X = x] = \int_{\Delta} \frac{f(x,y)}{f(x,.)} dy$$

Cette petite note explique comment dans un cas simple, on peut montrer que la densité de $Y|X = x$ pour $f(x,.) \neq 0$ est

$$f(y|x) = \frac{f(x,y)}{f(x,.)}$$

CAS DISCRET X et Y indépendants $\Leftrightarrow p_{ij} = p_{i\cdot}\, p_{\cdot j} \quad \forall i \forall j$

Rq : $\boxed{X \text{ et } Y \text{ ind} \Leftrightarrow f_{i|j} = P[X = x_i | Y = y_j] = f_{i\cdot} \quad \forall i \forall j}$

en d'autres termes $f_{i|j}$ indépendant de j

CAS CONTINU X et Y indépendants $\Leftrightarrow f(x,y) = f(x,\cdot) f(\cdot,y) \quad \forall x \forall y$ (voir TD3)

Rq : $\boxed{X \text{ et } Y \text{ ind} \Leftrightarrow f(x|y) = f(x,\cdot) \quad \forall x \forall y}$

ie $f(x|y)$ indépendant de y

Propriété : Si X et Y sont indépendantes et si $\alpha : \mathbb{R} \to \mathbb{R}$ et $\beta : \mathbb{R} \to \mathbb{R}$

alors en général $\alpha(X)$ et $\beta(Y)$ sont des va indépendantes

V. Espérance mathématique

1. Définition.

De manière analogue au cas 1D, on définit l'espérance mathématique de $\alpha(X,Y)$, où α est une fonction de $\mathbb{R}^2 \to \mathbb{R}$ par :

$$E[\alpha(X,Y)] = \begin{cases} \text{cas discret :} \\ \sum_{i,j} \alpha(x_i, y_j)\, p_{ij} \\ \text{cas continu} \\ \iint_{\mathbb{R}^2} \alpha(x,y) f(x,y)\, dx\, dy \\ \text{cas mixte : } X \text{ discrète, } Y \text{ continue} \\ \sum_i \left(\int_{\mathbb{R}} \alpha(x_i,y) f(y)\, dy\right) P[X = x_i] \end{cases} = \sum_i \int_{\mathbb{R}} \alpha(x_i, y)\, f_i(y)\, dy$$

2. Propriétés.

- linéarité $E[a\alpha(X,Y) + b\beta(X,Y)] = a\,E[\alpha(X,Y)] + b\,E[\beta(X,Y)]$
- constante ou fonction déterministe $E[\text{cste}] = \text{cste}$, $E[\cos t] = \cos t$ (t : temps)
- $\underset{\text{cas continu}}{E[\alpha(X)]} = \iint_{\mathbb{R}^2} \alpha(x) f(x,y)\, dx\, dy = \int_{\mathbb{R}} \alpha(x) \underbrace{\left[\int_{\mathbb{R}} f(x,y)\, dy\right]}_{f(x,\cdot)} dx$

  la définition est cohérente avec le cas 1 variable.
- indépendance $E[\alpha(X).\beta(Y)] = E[\alpha(X)]\,E[\beta(Y)]$ si X et Y indépendantes alors

démo (cas continu)

$$E[\alpha(X)\beta(Y)] = \iint \alpha(x)\beta(y) f(x,y)\, dx\, dy \underset{X \text{ et } Y \text{ ind}}{=} \iint \alpha(x)\beta(y) f(x,\cdot) f(\cdot,y)\, dx\, dy$$

$$\underset{\text{var séparables}}{=} \int \alpha(x) f(x,\cdot)\, dx \int \beta(y) f(\cdot,y)\, dy = E[\alpha(X)]\,E[\alpha(Y)].$$

3. Exemples.

les moments centrés $\mu_{ij} = E[(X - E(X))^i (Y - E(Y))^j]$

les moments non centrés $m_{ij} = E[X^i Y^j]$ $\forall i \in \mathbb{N}$ $\forall j \in \mathbb{N}$

Rq : pour $i=0$ ou $j=0$, on se ramène au cas 1D.
pour $i=1$ et $j=1$, on a $\mu_{11} = E[(X-E(X))(Y-E(Y))]$
μ_{11} s'appelle la covariance entre X et Y notée $Cov(X,Y)$

Calcul de $Cov(X,Y)$: $\mu_{11} = E[(X-E(X))(Y-E(Y))]$
$= E[XY - XE(Y) - YE(X) + E(X)E(Y)]$
$= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y]$
$Cov(X,Y) = E[XY] - E[X]E[Y]$.

Matrice de covariance (ou matrice des variances-covariances)

$$E\left[\underbrace{\begin{pmatrix} X-E(X) \\ Y-E(Y)\end{pmatrix}}_{V} \underbrace{\begin{pmatrix} X-E(X) & Y-E(Y)\end{pmatrix}}_{{}^tV}\right] = \begin{pmatrix} Var X & Cov(X,Y) \\ Cov(X,Y) & Var Y\end{pmatrix}$$

Rq : si X et Y sont indépendantes, alors $Cov(X,Y) = 0$ et donc la matrice de covariance est diagonale.
la réciproque est fausse (voir TD)

le coefficient de corrélation $r_{X,Y} = \dfrac{cov(X,Y)}{\sigma_X \sigma_Y}$

avec σ_X = écart type de $X = \sqrt{Var X}$
σ_Y = ______ $Y = \sqrt{Var Y}$

On va voir que $r_{X,Y}$ est une mesure imparfaite mais très pratique du lien existant entre X et Y.

propriétés : $-1 \leqslant r_{X,Y} \leqslant +1$

démo : on pose $g(\alpha) = E[((X-E(X)) + \alpha(Y-E(Y)))^2] \geqslant 0$ $\forall \alpha$
mais $g(\alpha) = Var X + 2\alpha\, cov(X,Y) + \alpha^2 Var Y \geqslant 0$ $\forall \alpha$
$\Rightarrow \Delta' \leqslant 0$.
$\Rightarrow \Delta' = cov(X,Y)^2 - Var X\, Var Y \leqslant 0$
$\Rightarrow \dfrac{cov(X,Y)^2}{Var X\, Var Y} \leqslant 1 \Rightarrow \boxed{r_{X,Y}^2 \leqslant 1}$

Rq : on montre que $\langle X,Y \rangle = E[XY]$ est un produit scalaire.
on applique Cauchy-Schwartz à $\tilde{X} = X - E(X)$, $\tilde{Y} = Y - E(Y)$
Inégalité de Cauchy-Schwartz : $|\langle X,Y\rangle|^2 \leqslant \|X\|^2 \|Y\|^2$

$$Y = \sum_{i=1}^{n} X_i \qquad X_i \begin{cases} \to 1 & p \\ \to 0 & q \end{cases} \qquad E[X_i] = p$$

$$E[Y] = np$$

si N v.a. aléatoire

$$E\left[\sum_{i=1}^{N} X_i\right] = E\left[E\left[\sum X_i \mid N\right]\right] = E[Np] = p\,E(N)$$

$$|\mathrm{cov}(X,Y)|^2 \leqslant \mathrm{var}X \ \mathrm{var}Y$$

que se passe-t-il lorsque $r_{XY} = \pm 1$?

alors $\Delta' = 0 \Rightarrow g(\alpha) = \mathrm{Var}Y(\alpha - \alpha_0)^2$

$g(\alpha_0) = 0$.

d'où $E[((X-E(X)) + \alpha_0(Y-E(Y)))^2] = 0$

$X - E(X) + \alpha_0(Y - E(Y)) = 0$

$r_{XY} = \pm 1 \Rightarrow$ X et Y sont liés par une relation affine.

la réciproque est vraie (évident)

conclusion.
$$\boxed{\begin{array}{l} X \text{ et } Y \text{ ind} \underset{\not\Leftarrow}{\Rightarrow} r_{XY} = 0 \\ X \text{ et } Y \text{ liés par une relation affine} \Leftrightarrow r_{XY} = \pm 1 \\ -1 \leqslant r_{XY} \leqslant +1 \end{array}}$$

r_{XY} est une mesure imparfaite du lien entre X et Y

. Fonction caractéristique

Rappel (1D) $\phi_X(u) = E[e^{iuX}]$ $u \in \mathbb{R}$

la fonction caractéristique d'un couple est définie par :

$$\phi_{X,Y}(u_1, u_2) = E[e^{i(u_1X + u_2Y)}] \quad (u_1, u_2) \in \mathbb{R}^2$$

on pose souvent $u = \binom{u_1}{u_2}$ d'où $\phi_{X,Y}(u) = E[e^{i\,{}^tu\binom{X}{Y}}]$ $u \in \mathbb{R}^2$

Fin cours 5

4. Espérances conditionnelles

Thm: $\underset{(X,Y)}{E}[\alpha(X,Y)] = E_X[E_Y[\alpha(X,Y)|X]]$

Preuve

$\iint \alpha(x,y) f(x,y)\, dx\, dy =$

$\iint \alpha(x,y)\, f(y|x) f(x,\cdot)\, dx\, dy =$

$\int [\int \alpha(x,y) f(y|x) dy]\, f(x,\cdot)\, dx$

$E_X[E[\alpha(x,y)|X]]$

Exemple d'application : le "secteur"

$X(t) = A\cos(2\pi f_0 t + \phi)$ Calculer plutôt $E[X^2|t]$

Modèle 1 $\left| \begin{array}{l} A = 220\sqrt{2} \\ f_0 = 50\text{Hz} \end{array} \right.$

. autre exemple plus simple $Y = \sum_{i=1}^{N} X_i$ $X_i \to 1$ p, $\to 0$ q, $N \sim P(\lambda)$

ϕ uniforme sur $[0, 2\pi[$ $f(\varphi) = \frac{1}{2\pi} 1_{[0,2\pi[}(\varphi)$

$E[X(t)] = \int_{\mathbb{R}} A\cos(2\pi f_0 t + \varphi) f(\varphi)\, d\varphi = \frac{1}{2\pi}\int_0^{2\pi} A\cos(2\pi f_0 t + \varphi)\, d\varphi = 0$.

Modèle 2 $\left| \begin{array}{l} A = 220\sqrt{2} \\ \phi \text{ uniforme sur } [0, 2\pi[\\ f_0 \text{ uniforme sur } [50 - \Delta f, 50 + \Delta f] \end{array} \right.$

(ϕ et f_0 en général indépendants)

$$E[X(t)] = E_{\varphi_0}\left[\underbrace{E_{\varphi}[X(t) \mid \varphi_0 \text{ fixe}]}_{0} \right] = E[\varphi_0] = 0.$$

VI Changements de variable de $\mathbb{R}^2 \to \mathbb{R}^2$

Problème : étant donné un couple (X,Y) de loi connue, comment peut-on déterminer la loi de $\begin{pmatrix} U \\ V \end{pmatrix} = \alpha(X,Y)$ où α est une fonction de $\mathbb{R}^2 \to \mathbb{R}^2$ et U, V deux fonctions de $\mathbb{R}^2 \to \mathbb{R}$.

1. Cas discret

Si (X,Y) est un couple de va discrètes à valeur dans $\{(x_i, y_j)\ i \in I,\ j \in J\}$ alors (U,V) est aussi un couple de va discrètes à valeurs dans $\{\alpha(x_i, y_j),\ i \in I, j \in J\} = \{(u_k, v_\ell),\ k \in K, \ell \in L\}$ avec :

$$P\left[\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} u_k \\ v_\ell \end{pmatrix}\right] = \sum_{\{(i,j) / \alpha(x_i, y_j) = \binom{u_k}{v_\ell}\}} \overbrace{P[X = x_i, Y = y_j]}^{p_{ij}}$$ c'est un cas "simple".

2. Cas continu

Théorème : si (X,Y) est un couple de va continues à valeurs dans un ouvert $\Delta \subset \mathbb{R}$ et si $\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ est bijective continûment différentiable ainsi que son inverse α^{-1}, alors $\begin{pmatrix} U \\ V \end{pmatrix} = \alpha\begin{pmatrix} X \\ Y \end{pmatrix}$ est un couple de va continues à valeurs dans $\alpha(\Delta)$ de densité :

$$f_{(U,V)}(u,v) = f_{(X,Y)}\left(\alpha^{-1}\begin{pmatrix} U \\ V \end{pmatrix}\right) |\det(J)|$$

avec J matrice Jacobienne $J = \begin{pmatrix} \frac{\partial X}{\partial U} & \frac{\partial Y}{\partial U} \\ \frac{\partial X}{\partial V} & \frac{\partial Y}{\partial V} \end{pmatrix}$

Exemple : $\left| \begin{array}{l} X \sim N(0,1) \\ Y \sim N(0,1) \\ X \text{ et } Y \text{ indépendantes} \end{array} \right.$ $\Leftrightarrow$ $\boxed{\begin{array}{l} f(X,Y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \\ (x,y) \in \mathbb{R}^2 \end{array}}$

Quelle est la loi de $\begin{pmatrix} R \\ \Phi \end{pmatrix}$ avec $\begin{cases} X = R\cos\Phi \\ Y = R\sin\Phi \end{cases}$

On sait que le chgt de va est bijectif de $\mathbb{R}^2 \setminus \{(0,0)\}$ dans $]0, +\infty[\times [0, 2\pi[$

$J = \begin{pmatrix} \cos\Phi & \sin\Phi \\ -R\sin\Phi & R\cos\Phi \end{pmatrix} \Rightarrow \det J = R$

d'où : $\boxed{f(r,\varphi) = \frac{1}{2\pi} e^{-\frac{1}{2}r^2}\, r \qquad (r,\varphi) \in \mathbb{R}^+ \times [0, 2\pi[}$

Rq: si α est bijective par morceaux, on ajoute la contribution de chaque bijection.

3. Cas continu non bijectif

VII Changements de variables de $\mathbb{R}^2 \to \mathbb{R}$

$U = \alpha(X,Y) \quad \alpha : \mathbb{R}^2 \to \mathbb{R}$

Methode 1 : on ajoute une variable fictive $V = \beta(X,Y)$, de façon à rendre le changement de va bijectif (si possible), on détermine la loi de $\binom{U}{V}$ $f(u,v)$, puis celle de U grace à $f(u,.) = \int f(u,v)\,dv$ en général $V = X$ ou $V = Y$.

Ex: $X \sim U(0,1) \quad f(x,.) = \mathbb{1}_{[0,1]}(x)$

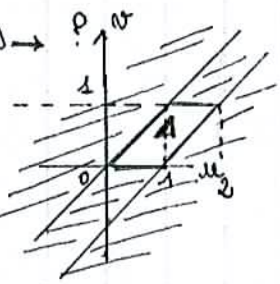
$Y \sim U(0,1) \quad f(y,.) = \mathbb{1}_{[0,1]}(y)$

X et Y indep.

Quelle est la loi de $U = X+Y$?

on pose par exemple $V = X \Rightarrow \begin{cases} U = X+Y \\ V = X \end{cases} \Leftrightarrow \begin{cases} X = V \\ Y = U-V \end{cases}$

donc le chgt de variable est bijectif de $[0,1]\times[0,1] \to$?

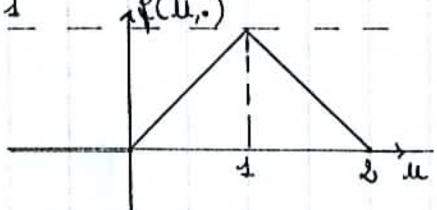


det de Δ : $\begin{cases} 0 \leq X \leq 1 \\ 0 \leq Y \leq 1 \end{cases} \Leftrightarrow \begin{cases} 0 \leq V \leq 1 \\ 0 \leq U-V \leq 1 \end{cases} \Leftrightarrow \boxed{\begin{cases} 0 \leq V \leq 1 \\ V \leq U \leq V+1 \end{cases}}$

$\det J = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$

$\boxed{f(u,v) = \mathbb{1}_\Delta(u,v)}$

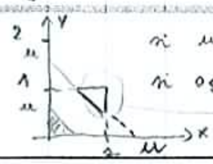
loi de U : $f(u,.) = \int_{\mathbb{R}} f(u,v)\,dv = \begin{cases} 0 \text{ si } u \leq 0 \text{ ou } u \geq 2 \\ \int_0^u 1\,dv \text{ si } u \in [0,1] \\ \int_{u-1}^1 dv = 2-u \text{ si } u \in [1,2] \end{cases}$



Rq: $\int_{\mathbb{R}} f(u,.)\,du = 1$

Methode 2 : on calcule directement la fonction de répartition de U
$F(u) = P[U<u] = P[\alpha(X,Y)<u] = P[(X,Y)\in D_u]$
$= \iint_{D_u} f(x,y)\,dx\,dy.$

ex précédent $F(u) = P[X+Y<u]$



si $u<1$ $F(u) = \frac{1}{2}u^2$

si $0\leq 1\leq 2$ $F(u) = 1-(1-(u-1))^2$

Méthode 3 pour $U = X+Y$, X et Y indépendants
<u>utilisation des fonctions caractéristiques</u>

Ex : $\begin{cases} X \sim N(m_1, \sigma_1^2) \\ Y \sim N(m_2, \sigma_2^2) \\ X \text{ et } Y \text{ indep.} \end{cases}$ loi de $U = X+Y$?

$$\phi_U(t) = E[e^{itU}] = E[e^{itX} e^{itY}] \underset{\substack{\uparrow \\ X \text{ et } Y \\ \text{indépendants}}}{=} E[e^{itX}]\, E[e^{itY}] = \phi_X(t)\, \phi_Y(t)$$

Dans les tables, on lit $X \sim N(m_1, \sigma_1^2)$ $\phi_X(t) = \exp(i m_1 t - \frac{\sigma_1^2}{2} t^2)$

$Y \sim N(m_2, \sigma_2^2)$ $\phi_Y(t) = \exp(i m_2 t - \frac{\sigma_2^2}{2} t^2)$

d'où $\phi_U(t) = \exp\left(i \underbrace{(m_1 + m_2)}_{m} t - \frac{1}{2} \underbrace{(\sigma_1^2 + \sigma_2^2)}_{\sigma^2} t^2 \right)$

$\Rightarrow \boxed{U \sim N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)}$

Chap 4

Vecteurs gaussiens

I. Définition

on dit que $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ suit une loi normale à n dimensions (ou que X est un vecteur gaussien) et on note $X \sim N_n(m, \Sigma)$ si la densité de X s'écrit $f(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} \exp\left(-\frac{1}{2} {}^t(x-m)\Sigma^{-1}(x-m)\right) \quad \forall x \in \mathbb{R}^n$

où : $m = \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} \in \mathbb{R}^n$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, et $\Sigma \in M_n(\mathbb{R})$ une matrice symétrique définie positive.

. Σ symétrique ${}^t\Sigma = \Sigma$
. Σ positive ${}^tx\Sigma x \geq 0 \; \forall x$
. Σ définie ${}^tx\Sigma x = 0 \Rightarrow x = 0$

Σ est diagonalisable avec des v.p $d_i > 0$
donc Σ "se comporte" comme $\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$
et $\det \Sigma > 0$, Σ inversible.

STOP 6 octobre 2003

$n = 1$: on a $m = (m)$ et $\Sigma = (\sigma^2)$, $\sigma^2 > 0$.
d'où $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right\}$

on retrouve la densité d'une loi normale à 1D. Cohérent.

Σ diagonale : $\Sigma = \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix}$ donc $\Sigma^{-1} = \begin{pmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_n^2 \end{pmatrix}$

d'où $f(x) = \frac{1}{(\sqrt{2\pi})^n \sqrt{\sigma_1^2 \ldots \sigma_n^2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \frac{(x_i - m_i)^2}{\sigma_i^2}\right)$

ie $\boxed{f(x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp -\frac{1}{2\sigma_i^2}(x_i - m_i)^2}$ produit de n densités normales à 1D.

Cas général. $f(x) = \text{cste} . \exp\{\underbrace{\text{forme quadratique}}_{\frac{1}{2}\sum a_{ij}x_ix_j} + \underbrace{\text{terme linéaire}}_{\sum b_i x_i} + \text{cste}\}$

Rq : on note souvent $f(x) = \text{cste} \exp\left\{\frac{{}^tx A x}{2} + {}^tBx + K\right\}$

Théorème : Si $X \sim N_n(m, \Sigma)$, alors on a

fonction génératrice des moments $\}$ $\Theta(u) = E\left[e^{{}^tuX}\right] = \int \cdots \int_{\mathbb{R}^n} e^{{}^tux} f(x)dx = \exp\left({}^tum + \frac{1}{2}{}^tu\Sigma u\right)$

fonction caractéristique $\}$ $\phi(u) = E\left[e^{i{}^tuX}\right] = \int \cdots \int_{\mathbb{R}^n} e^{i{}^tux} f(x)dx = \exp\left(i{}^tum - \frac{1}{2}{}^tu\Sigma u\right)$

Signification de m et de Σ

En utilisant ce théorème, on peut déterminer tous les moments de X (d'où le nom de $\Theta(u)$). En effet :

Moments d'ordre 1 : $E[X_i]$

$\Theta(u) = E[e^{{}^t u X}] = E\left[\sum_{k=0}^{\infty} \frac{({}^t u X)^k}{k!}\right] \underset{\text{permis}}{=} \sum_{k=0}^{\infty} \frac{1}{k!} E[({}^t u X)^k]$

d'où $\Theta(u) = 1 + E[u_1 X_1 + .. + u_n X_n] + \frac{1}{2} E[(u_1 X_1 + .. + u_n X_n)^2] + \dots$

$= 1 + u_1 E(X_1) + u_2 E(X_2) + \dots + u_n E(X_n) + \frac{1}{2} E[(\ \ )^2] + \dots$

on "voit" donc que $\boxed{E[X_i] = \left.\frac{\partial \Theta(u)}{\partial u_i}\right|_{u=0}}$

Puisqu'on connait $\Theta(u) = \exp[{}^t u m + \frac{1}{2} {}^t u \Sigma u]$ on en déduit :

$E(X_i) = \left[\exp[{}^t u m + \frac{1}{2} {}^t u \Sigma u] . m_i\right] \Big|_{u=0} = m_i$

Conclusion : $m = \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix}$ est le vecteur des moyennes.

On fait de même à l'ordre 2 et on obtient $\boxed{E[X_i X_j] = m_i m_j + \Sigma_{ij}}$

d'où $\Sigma_{ij} = E[X_i X_j] - m_i m_j = \begin{cases} i=j & \mathrm{Var} X_i \\ i \neq j & \mathrm{Cov}(X_i, X_j) \end{cases}$

d'où $\Sigma = \begin{array}{c c} & \\ i \longrightarrow & \begin{pmatrix} \mathrm{Var} X_1 & & \\ & \mathrm{Var} X_2 & \\ \mathrm{Cov}(X_i,X_j) & \ddots & \mathrm{Var} X_n \end{pmatrix} \\ & \uparrow j \end{array}$ s'appelle la matrice de covariance.

Rq : Σ s'écrit de la façon suivante $\Sigma = E[(X - E(X))\, {}^t(X - E(X))]$

2. Transformation affine

Problème : soit X un vecteur gaussien $X \sim \mathcal{N}_n(m, \Sigma)$

quelle est la loi de $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_p \end{pmatrix} = AX + B$ avec $p \leq n$

$A = \overbrace{\begin{pmatrix} & \end{pmatrix}}^{n} \}p$, $B \in \mathbb{R}^p$

Idée : on connait la fonction caractéristique X

$\phi(u) = \exp(i\, {}^t u m - \frac{1}{2} {}^t u \Sigma u)$

on calcule celle de $Y = AX + B$ et on espère reconnaitre le résultat

Preuve : fonction caractéristique de Y : $\tilde{\phi}(\tilde{u}) = E[e^{i {}^t \tilde{u} Y}]$

on a $\tilde{\phi}(\tilde{u}) = E[e^{i {}^t \tilde{u}(AX+B)}] = E[e^{i \underbrace{{}^t(\ )}_{{}^t A \tilde{u}} X} e^{i {}^t \tilde{u} B}]$

Or ${}^t\tilde{u}A = {}^t({}^tA\tilde{u})$

donc $\tilde{\phi}(\tilde{u}) = e^{i\,{}^t\tilde{u}B}\phi({}^tA\tilde{u}) = e^{i\,{}^t\tilde{u}B}\exp[i\,{}^t\tilde{u}Am - \frac{1}{2}\,{}^t\tilde{u}A\Sigma{}^tA\tilde{u}]$

$\Rightarrow Y \sim N_p(B+Am, A\Sigma{}^tA)$ dans la mesure où $A\Sigma{}^tA$ est symétrique définie positive

Rq : $Y = AX+B \Rightarrow E[Y] = A\underbrace{E[X]}_{m} + B = Am+B$.

$\Sigma_Y = E[(Y-E[Y])^t(Y-E[Y])]$

$= E[(AX+B-Am-B)^t(A(X-m))] = E[A(X-m)^t(X-m)^tA]$

$= A\underbrace{E[(X-m)^t(X-m)]}_{\Sigma}{}^tA = A\Sigma{}^tA$.

Conclusion : $X \sim N_n(m, \Sigma) \Rightarrow Y = AX+B \sim N_p(Am+B, A\Sigma{}^tA)$

pourvu que $A\Sigma{}^tA$ soit symétrique défini positive

i.e A de rang maximum.

Vérifions que $A\Sigma{}^tA$ est symétrique définie positive

en effet : . ${}^t(A\Sigma{}^tA) = \underbrace{{}^t({}^tA)}_{A}\underbrace{{}^t\Sigma}_{\Sigma}{}^tA = A\Sigma{}^tA$: $A\Sigma{}^tA$ toujours symétrique.

. ${}^ty(A\Sigma{}^tA)\underbrace{y}_{v} = {}^tv\Sigma v \geq 0$ car Σ positive.

. ${}^ty(A\Sigma{}^tA)y = 0 \underset{?}{\Rightarrow} y = 0$.

${}^ty(A\Sigma{}^tA)y = {}^tv\Sigma v = 0 \underset{\Sigma \text{ définie}}{\Rightarrow} v = 0$

$\Rightarrow {}^tAy = 0 \Rightarrow y = 0$

si ${}^tA$ est de rang maximum.

Rq : A l'aide du résultat précédent on montre facilement que si $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim N_n(m, \Sigma)$ alors $X_i \sim N(m_i, \Sigma_{ii})$

Réciproque fausse (voir TD3).

3. Lois marginales.

$X \sim N_n(m, \Sigma)$

$X = \begin{pmatrix} X' \\ X'' \end{pmatrix} \begin{matrix} \updownarrow p \\ \updownarrow n-p \end{matrix} \quad m = \begin{pmatrix} m' \\ m'' \end{pmatrix} \quad \Sigma = \left(\begin{array}{c|c} \Sigma' & A \\ \hline {}^tA & \Sigma'' \end{array}\right)$

conclusion : $X' \sim N_p(m', \Sigma')$

demo :

il suffit de voir que $X' = \underbrace{(I_p, 0)}_{C}\begin{pmatrix} X' \\ X'' \end{pmatrix} \begin{matrix} \updownarrow p \\ \updownarrow n-p \end{matrix}$ $\quad C = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \end{pmatrix} \updownarrow p$ ($\overset{p}{\leftrightarrow}\ \overset{n-p}{\leftrightarrow}$) de rang max p.

donc on peut appliquer les résultats du I, d'où :

$$X' \sim N_p\left((I_p, 0)\begin{pmatrix} m' \\ m''\end{pmatrix} + 0,\; C_{(I_p,0)} \begin{pmatrix}\Sigma' & A \\ {}^tA & \Sigma''\end{pmatrix} {}^tC_{\begin{pmatrix} I_p \\ 0\end{pmatrix}}\right)$$

d'où $X' \sim N_p(m', \Sigma')$ cqfd.

I Lois conditionnelles.

Si $X = \begin{pmatrix} X' \\ X''\end{pmatrix}\begin{matrix} \}p \\ \}n-p\end{matrix} \sim N_n(m,\Sigma)$ alors $X' | X'' = x''$ est aussi un vecteur gaussien.

admis.

: . Indépendance.

Si $X = \begin{pmatrix} X' \\ X''\end{pmatrix} \sim N_n(m, \Sigma)$ avec $m = \begin{pmatrix} m' \\ m''\end{pmatrix}$ et $\Sigma = \begin{pmatrix}\Sigma' & A \\ {}^tA & \Sigma''\end{pmatrix}$

alors : X' et X'' indépendantes $\Leftrightarrow A = 0$. (ie $cov(X_i, X_j) = 0$ $\forall j \in \{p+1, \dots, n\}$ $\forall i \in \{1, \dots p\}$)

démo :

X' et X'' indépendantes $\overset{def}{\Leftrightarrow} f(x) = f(x', \cdot) f(\cdot, x'')$ $\forall x \in \mathbb{R}^n$ $\forall x' \in \mathbb{R}^p$ $\forall x'' \in \mathbb{R}^{n-p}$

$$\underset{(ADMIS)}{\Leftrightarrow} \Phi_X(u) = \Phi_{X'}(u') \Phi_{X''}(u'') \quad \forall (u, u', u'')$$

(Rq: $\Phi_X(u) = E[e^{i\,{}^tuX}] = E[e^{i\,{}^t\begin{pmatrix}u'\\u''\end{pmatrix}\begin{pmatrix}X'\\X''\end{pmatrix}}] = E[e^{i\,{}^tu'X' + {}^tu''X''}]$

$= E[e^{i\,{}^tu'X'} e^{i\,{}^tu''X''}] = \phi_{X'}(u')\phi_{X''}(u'')$)

donc $\Rightarrow$ évident, on admet seulement $\Leftarrow$

$$\Leftrightarrow \exp(\cancel{i\,{}^tum} - \tfrac{1}{2}{}^tu\Sigma u) = \exp(\cancel{i\,{}^tu'm'} - \tfrac{1}{2}\cancel{{}^tu'\Sigma' u'})$$
$$\exp(\cancel{i\,{}^tu''m''} - \tfrac{1}{2}\cancel{{}^tu''\Sigma'' u''})$$

or ${}^tum = {}^t\begin{pmatrix}u'\\u''\end{pmatrix} \cdot \begin{pmatrix}m'\\m''\end{pmatrix} = {}^tu'm' + {}^tu''m''$ donc

$${}^tu\Sigma u = ({}^tu'\ {}^tu'')\begin{pmatrix}\Sigma' & A \\ {}^tA & \Sigma''\end{pmatrix}\begin{pmatrix}u'\\u''\end{pmatrix}$$
$$= ({}^tu'\ {}^tu'')\begin{pmatrix}\Sigma'u' + Au'' \\ {}^tAu' + \Sigma''u''\end{pmatrix}$$
$$= \cancel{{}^tu'\Sigma'u'} + {}^tu'Au'' + {}^tu''\,{}^tAu' + \cancel{{}^tu''\Sigma''u''} \text{ donc}$$

$$\Leftrightarrow {}^tu'Au'' + \underbrace{{}^tu''\,{}^tAu'}_{\in \mathbb{R}} = 0 \quad \forall (u', u'').$$

donc ${}^tu''\,{}^tAu' = {}^t({}^tu''\,{}^tAu') = {}^tu'Au''$

$$\Leftrightarrow 2\,{}^tu'Au'' = 0 \quad \forall (u', u'')$$

$$\Leftrightarrow A = 0.$$

INP ENSEEIHT

COMPLEMENTS DE COURS

- **Loi du Chi2 à n degrés de liberté**

  Définition : Si $X_1, ..., X_n$ sont n variables aléatoires indépendantes (variables aléatoires i.i.d.) et de même loi $\mathcal{N}(0,1)$,

  alors $Y = \sum_{k=1}^{n} X_k^2$ suit une loi du Chi2 à n degrés de liberté

  Notation :

$$Y \sim \chi_n^2$$

  Densité de Probabilité :

$$f_n(y) = \frac{y^{\frac{n}{2}-1}e^{-\frac{y}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} 1_{\mathbb{R}^+}(y)$$

  Ce résultat peut se démontrer par récurrence. Pour $n = 1$, nous l'avons déjà fait dans la partie changement de variables continues. Ensuite, on cherche la loi de $Y + X_{n+1}^2$ (remarque : pour $n = 2$ ou $n = 3$, on peut faire un changement de variables avec les coordonnées polaires et les coordonnées sphériques respectivement).

  Fonction Caractéristique

$$\varphi_Y(t) = \frac{1}{(1-2it)^{\frac{n}{2}}}$$

  Moyenne et Variance

$$\begin{aligned} E[Y] &= n \\ VarY &= 2n \end{aligned}$$

  En effet

  $E[Y] = E\left[\sum_{k=1}^{n} X_k^2\right] = \sum_{k=1}^{n} E[X_k^2] = n$

  $VarY = \sum_{k=1}^{n} VarX_k^2 = \sum_{k=1}^{n} \left(E[X_k^4] - E[X_k^2]^2\right) = \sum_{k=1}^{n} (3-1) = 2n$

  Propriété d'additivité

$$\left\{ \begin{array}{l} X \sim \chi_n^2 \\ Y \sim \chi_m^2 \\ X \text{ et } Y \text{ indépendantes} \end{array} \right. \Longrightarrow X + Y \sim \chi_{n+m}^2$$

- **Loi de Student à n degrés de liberté ($Z \sim t_n$)**

  Définition

$$\left\{\begin{array}{l} X \sim \mathcal{N}(0,1) \\ Y \sim \chi^2_n \\ X \text{ et } Y \text{ indépendantes} \end{array}\right. \Longrightarrow Z = \frac{X}{\sqrt{\frac{Y}{n}}} \sim t_n$$

  Densité

$$f_n(z) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1+\frac{z^2}{n}\right)^{\frac{n+1}{2}}} \qquad z \in \mathbb{R}$$

  Moyenne et Variance

  Dès que $n > 2$, on a

$$\begin{aligned} E[Z] &= 0 \\ VarZ &= \frac{n}{n-2} \end{aligned}$$

  Loi de Cauchy

  Pour $n = 1$, la loi suivie par $Z = \frac{X}{\sqrt{Y}}$ porte le nom de la loi de CAUCHY. Sa densité est

$$f(z) = \frac{1}{\pi} \frac{1}{1+z^2} \qquad z \in \mathbb{R}$$

  Une telle loi ne possède aucun moment et en particulier pas de moyenne et pas de variance.

- **Loi de Fisher-Snédécor à n et m degrés de libertés ($Z \sim f_{n,m}$)**

  Définition

$$\left\{\begin{array}{l} X \sim \chi^2_n \\ Y \sim \chi^2_m \\ X \text{ et } Y \text{ indépendantes} \end{array}\right. \Longrightarrow Z = \frac{X/n}{Y/m} \sim f_{n,m}$$

  Densité

$$f_n(z) = \frac{n}{m} \frac{1}{\beta\left(\frac{n}{2}, \frac{m}{2}\right)} \frac{\left(\frac{n}{n}z\right)^{\frac{n}{2}-1}}{\left(1+\frac{n}{n}z\right)^{\frac{n+m}{2}}} \qquad z \in \mathbb{R}^+$$

  avec $\beta(n,m) = \frac{\Gamma(n)\Gamma(m)}{\Gamma(n+m)}$.

  Moyenne et Variance

  Dès que $m > 4$, on a

$$\begin{aligned} E[Z] &= \frac{m}{m-2} \\ VarZ &= \frac{2m^2(n+m-2)}{n(m-4)(m-2)^2} \end{aligned}$$

Chap 5 Convergence et théorèmes limites

I. Convergence en loi

Définition : on dit que la suite de va $X_1, \dots X_n$ cv en loi vers la va X si la suite des fonctions de répartition $F_n(x)$ cv simplement vers la fonction de répartition de X notée $F(x) = P[X<x]$ en tout point x où F est continue

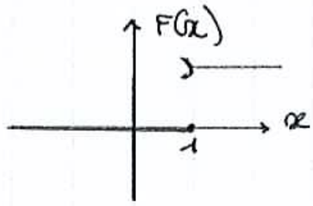
Notation : $X_n \xrightarrow[n\infty]{L} X$

exemple :

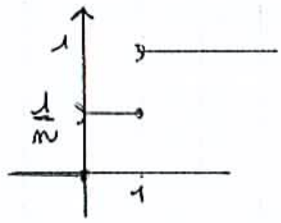
$$X_n \begin{cases} \nearrow 0 & P[X_n=0] = \frac{1}{n} \\ \searrow 1 & P[X_n=1] = \frac{n-1}{n} = 1-\frac{1}{n} \end{cases}$$

MQ $X_n \xrightarrow[n\infty]{L} 1 = X$

$F(x) = P[X<x]$



$$F_n(x) = P[X_n<x] = \begin{cases} 0 \text{ si } x \leq 0 \\ \frac{1}{n} \text{ si } x \in]0,1] \\ 1 \text{ si } x > 1 \end{cases}$$



- si $x \leq 0$ $F_n(x) = 0 \xrightarrow[n\infty]{} 0 = F(x)$
- si $x \in]0,1[$ $F_n(x) = \frac{1}{n} \xrightarrow[n\infty]{} 0 = F(x)$
- si $x > 1$ $F_n(x) = 1 \xrightarrow[n\infty]{} 1 = F(x)$

} il y a cv simple de $F_n(x)$ vers $F(x)$ en tout point $x \neq 1$.

Propriétés : ① THÉORÈME DE LEVY

$$X_n \xrightarrow[n\infty]{L} X \Leftrightarrow \begin{cases} \varphi_n(t) = E[e^{itX_n}] \xrightarrow[n\infty]{} \varphi(t) = E[e^{itX}] \quad \forall t \\ \varphi \text{ continue en } t=0 \end{cases}$$

Dans l'exemple précédent $\varphi_n(t) = E[e^{itX_n}] = e^{it0}P[X_n=0] + e^{it1}P[X_n=1]$

$$\boxed{\varphi_n(t) = \frac{1}{n} + e^{it}\left(1-\frac{1}{n}\right)}$$

$$\boxed{\varphi_n(t) \xrightarrow[n\infty]{} e^{it}}$$

or $\varphi(t) = E[e^{it1}] = e^{it}$ continue en $t=0$.

② si $(X_n)_{n\in\mathbb{N}}$ est une suite de va continues de densité $f_n(x)$ et que $f_n(x) \xrightarrow[n\infty]{} f(x)$ presque partout.

alors $X_n \xrightarrow[n\infty]{L} X$ où X est une v.a. de densité $f(x)$.

Rq : c'est une CS et pas une CNS.

(3) si $X_n \xrightarrow[n\infty]{L} X$ et $g : \mathbb{R} \to \mathbb{R}$ continue, alors $g(X_n) \xrightarrow[n\infty]{L} g(X)$

Commentaire : la cv en loi est souvent utilisée ds les thm limite (cf fin du chapitre).

II Convergence en probabilité

Définition : on dit que X_n cv en proba vers X et on note

$$X_n \xrightarrow[n\infty]{P} X \text{ ssi } \forall \varepsilon > 0 \quad P[|X_n - X| > \varepsilon] \xrightarrow[n\infty]{} 0$$

exemple : X_n de densité $f_n(x) = \dfrac{n e^{-nx}}{(1+e^{-nx})^2}$ $\quad x \in \mathbb{R}$

on vérifie bien $\lim\limits_{x \to +\infty} f_n(x) = 0$

$\lim\limits_{x \to -\infty} f_n(x) = 0$

MQ $X_n \xrightarrow[n\infty]{P} 0$

$$P[|X_n - X| > \varepsilon] = P[|X_n| > \varepsilon] = 1 - \int_{-\varepsilon}^{+\varepsilon} f_n(x)\,dx$$

$$= 1 - \int_{-\varepsilon}^{+\infty} \frac{n e^{-nx}}{(1+e^{-nx})^2} = 1 - \left[\frac{1}{1+e^{-nx}}\right]_{\varepsilon}^{+\varepsilon}$$

$\varepsilon > 0$.

$$= 1 - \underbrace{\frac{1}{1+e^{-n\varepsilon}}}_{\downarrow\, 1} + \underbrace{\frac{1}{1+e^{n\varepsilon}}}_{\downarrow n\infty\, 0} \xrightarrow[n\infty]{} 0.$$

Propriété : si $X_n \xrightarrow[n\infty]{P} X$, $g : \mathbb{R} \to \mathbb{R}$ continue, alors $g(X_n) \xrightarrow[n\infty]{P} g(X)$.

III Convergence en moyenne quadratique.

c'est celle-là qui sert !

Définition : on dit que X_n cv en moyenne quadratique vers X et on note $X_n \xrightarrow[n\infty]{MQ} X$ ssi $E[|X_n - X|^2] \xrightarrow[n\infty]{} 0$

exemple

(1) $\begin{cases} P[X_n = n] = \dfrac{1}{n^3} \\ P[X_n = 0] = 1 - \dfrac{1}{n^3} \\ X_n \xrightarrow{MQ} 0\,? \end{cases}$

(2) $\begin{cases} P[X_n = n] = \dfrac{1}{n^2} \\ P[X_n = 0] = 1 - \dfrac{1}{n^2} \\ X_n \xrightarrow{MQ} 0\,? \end{cases}$

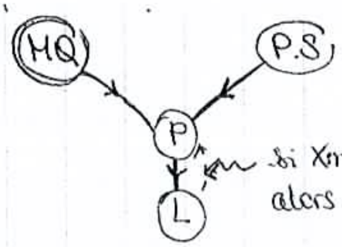
① $E[X_n^2] = 0^2\left(1-\frac{1}{n^3}\right) + n^2\left(\frac{1}{n^3}\right) = \frac{1}{n} \xrightarrow[n\infty]{} 0$ donc $X_n \xrightarrow[n\infty]{MQ} 0$

② $E[X_n^2] = 0^2\left(1-\frac{1}{n^3}\right) + n^2\left(\frac{1}{n^2}\right) = 1 \underset{n\infty}{\not\rightarrow} 0$ donc $X_n \underset{n\infty}{\overset{MQ}{\not\rightarrow}} 0$

II Convergence presque sûre

Définition : on dit que X_n cv presque sûrement vers X et on note $X_n \xrightarrow[n\infty]{PS} X$ si $X_n(w) \xrightarrow[n\infty]{} X(w)$ $\forall w \in A \,/\, P(A) = 1$

I Comparaison



si X_n cv en loi vers une constante alors X_n cv en proba vers une constante.

Compléments de Cours
Théorèmes Limites

Loi faible des grands nombres

Theorem 1 *Si $X_1, ..., X_n$ sont des variables aléatoires indépendantes et de même loi de moyenne $E[X_k] = m$, alors la variable aléatoire $\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$ converge en probabilité vers m.*

Preuve : Il est équivalent de montrer que $\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$ converge en loi vers $m = E[X_k]$. Pour ce, étudions la fonction caractéristique de $\overline{X}$:

$$\begin{aligned}
\varphi_{\overline{X}}(t) &= E\left[e^{it\frac{1}{n}\sum_{k=1}^{n} X_k}\right] \\
&= E\left[\prod_{k=1}^{n} e^{i\frac{t}{n}X_k}\right] \\
&= \prod_{k=1}^{n} E\left[e^{i\frac{t}{n}X_k}\right] \text{ (indépendance des variables } X_k\text{)} \\
&= \left[\varphi\left(\frac{t}{n}\right)\right]^n \text{ (les varriables } X_k \text{ ont la même fonction caractéristique)} \qquad (1)
\end{aligned}$$

Lorsque le premier moment d'une variable aléatoire X existe, la fonction caractéristique $\varphi_X(t)$ est une fois dérivable et on a $\varphi^{(k)}(0) = i^k E\left[X^k\right]$ pour $k \in \{0, 1\}$. En appliquant la formule de Tayor à la fonction φ, on obtient :

$$\begin{aligned}
\varphi(t) &= \varphi(0) + t\varphi'(0) + t\lambda(t) \\
&= 1 + itm + t\lambda(t)
\end{aligned}$$

avec $\lim\limits_{t\to 0} \lambda(t) = 0$. En utilisant le développement de Taylor de φ et l'expression (1), on obtient :

$$\begin{aligned}
\ln\left[\varphi_{\overline{X}}(t)\right] &= n \ln\left[1 + i\frac{t}{n}m + \frac{t}{n}\lambda\left(\frac{t}{n}\right)\right] \\
&= n\left[i\frac{t}{n}m + \frac{t}{n}\lambda\left(\frac{t}{n}\right)\right] \\
&= itm + t\lambda\left(\frac{t}{n}\right)
\end{aligned}$$

On a donc,

$$\lim_{n\to\infty} \varphi_{\overline{X}}(t) = e^{itm} \qquad \forall t$$

Mais, e^{itm} est la fonction caractéristique de la variable constant $X = m$. On en conclut d'après le théorème de Levy que

$$\overline{X} \xrightarrow{L} m \Leftrightarrow \overline{X} \xrightarrow{P} m$$

Loi des grands nombres dans L_2

Theorem 2 *Si $X_1, ..., X_n$ sont des variables aléatoires indépendantes et de même loi de moyenne $m < \infty$ et de variance $\sigma^2 < \infty$, alors la variable aléatoire $\overline{X}$ converge en moyenne quadratique vers m.*

Preuve : On a

$$\begin{aligned} E\left[\left(\overline{X}-m\right)^2\right] &= E\left[\left(\frac{1}{n}\sum_{k=1}^{n}X_k - \frac{1}{n}\sum_{k=1}^{n}m\right)^2\right] \\ &= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n}E\left[\left(X_k-m\right)\left(X_l-m\right)\right] \end{aligned}$$

Mais

$$E\left[\left(X_k-m\right)\left(X_l-m\right)\right] = \begin{cases} \sigma^2 \text{ si } k=l \\ 0 \text{ si } k \neq l \end{cases}$$

Donc

$$E\left[\left(\overline{X}-m\right)^2\right] = \frac{\sigma^2}{n} \underset{n\to\infty}{\to} 0$$

Remark 1 *Convergence de la moyenne empirique vers la moyenne de la loi.*

Théorème de la limite centrale

Theorem 3 *: Si $X_1, ..., X_n$ sont des variables aléatoires indépendantes et de même loi de moyenne $m < \infty$ et de variance $\sigma^2 < \infty$, alors la variable aléatoire centrée réduite $Y_n = \frac{\sum_{k=1}^n X_k - nm}{\sqrt{n\sigma^2}}$ converge en loi vers la loi normale $\mathcal{N}(0,1)$.*

Preuve : La fonction caractéristique de Y_n s'écrit :

$$\begin{aligned}
\varphi_{Y_n}(t) &= E\left[e^{itY_n}\right] = e^{-\frac{itm\sqrt{n}}{\sigma}} E\left[\prod_{k=1}^{n} e^{i\frac{t}{\sqrt{n\sigma^2}}X_k}\right] \\
&= e^{-\frac{itm\sqrt{n}}{\sigma}} \prod_{k=1}^{n} E\left[e^{i\frac{t}{\sigma\sqrt{n}}X_k}\right] \text{ (indépendance des variables } X_k) \\
&= e^{-\frac{itm\sqrt{n}}{\sigma}} \left[\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n \text{ (les variables } X_k \text{ ont la même fonction caractéristique)} \quad (2)
\end{aligned}$$

Lorsque les deux premiers moments d'une variable aléatoire X existent, la fonction caractéristique $\varphi_X(t)$ est deux fois dérivable et on a $\varphi^{(k)}(0) = i^k E\left[X^k\right]$, pour $k \in \{0,1,2\}$. En appliquant la formule de Tayor à la fonction φ, on obtient :

$$\begin{aligned}
\varphi(t) &= \varphi(0) + t\varphi'(0) + \frac{t^2}{2}\varphi''(0) + t^2\lambda(t) \\
&= 1 + itm - \left(m^2+\sigma^2\right)\frac{t^2}{2} + t^2\lambda(t)
\end{aligned}$$

avec $\lim_{t\to 0} \lambda(t) = 0$. En utilisant le développement de Taylor de φ et l'expression (2), on obtient :

$$\begin{aligned}
\ln\left[\varphi_{Y_n}(t)\right] &= -\frac{itm\sqrt{n}}{\sigma} + n\ln\left[1 + i\frac{t}{\sigma\sqrt{n}}m - \frac{(m^2+\sigma^2)}{2}\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + \left(\frac{t}{\sigma\sqrt{n}}\right)^2\lambda\left(\frac{t}{\sigma\sqrt{n}}\right)\right] \\
&= -\frac{itm\sqrt{n}}{\sigma} + n\left[i\frac{t}{\sigma\sqrt{n}}m - \frac{(m^2+\sigma^2)}{2}\left(\frac{t}{\sigma\sqrt{n}}\right)^2 - \frac{1}{2}\left(i\frac{t}{\sigma\sqrt{n}}m\right)^2 + \frac{t^2}{n}\lambda\left(\frac{t}{\sigma\sqrt{n}}\right)\right] \\
&= -\frac{t^2}{2} + \frac{t^2}{n}\lambda\left(\frac{t}{\sigma\sqrt{n}}\right)
\end{aligned}$$

Donc, on a

$$\lim_{n\to\infty} \varphi_{Y_n}(t) = e^{-\frac{t^2}{2}} \qquad \forall t$$

Puisque $e^{-\frac{t^2}{2}}$ est la fonction caractéristique de la loi normale $\mathcal{N}(0,1)$, en utilisant le théorème de Levy, on en déduit

$$\boxed{Y_n \xrightarrow{L} \mathcal{N}(0,1)}$$

Remark 2 *La convergence vers la loi normale est assurée sous des hypothèses moins fortes comme les conditions de Liapounov ou Lindeberg (voir livre de Renyi, Calcul de Probabilités, Dunod, Paris, 1966, page 414).*

Remark 3 *Théorème de la limite centrale multivarié : Soit* X_k *une suite de vecteurs aléatoires définis sur* $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$ *indépendants et de même loi de vecteur moyenne* $m \in \mathbb{R}^p$ *et de matrice de covariance* $\Sigma \in \mathcal{M}_p(\mathbb{R})$*, alors*

$$\sqrt{n}\left[\sum_{k=1}^{n} X_k - nm\right] \xrightarrow{L} \mathcal{N}(0, \Sigma)$$