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

Mémoire présenté en vue de l'obtention du Diplôme **:**

MASTER en Mathématiques

Option : **Statistique**

Par

Hani Soumia

Titre :

Vecteurs Gaussiens

Membres du Comité d'Examen :

Dr.	**Berkane Hassiba**	UMKB	Président
Dr.	**Ouanoughi Yasmina**	UMKB	Encadreur
Dr.	**Kheireddine Souraya**	UMKB	Examinateur

Septembre 2020

Dédicace

Je dédie ce humble travail.

La plus proch de mon coeur celle que a fait l'impossible pour me donner le bonheur mon chère mère.

A mon père qui fait l'impossible pour me donner le bonheur et pour suivre mes études jusqu'à ce jour.

Je dédie aussi ce travail à mes frères et tout ma famille.

A tous mes amies de promotion 2019/2020

de mathématiques.

A toute personne que prendra de son temps pour lire ce document à parfaire.

.

REMERCIEMENTS

Avant tout choses, je remercie à **Dieu** *le tout puissant, pour m'avoir donnée la force et la patience, la santé et la volonté pour réaliser ce modeste travai*l.

*Mes premiers remerciement s'adressent à mon encadreur Mdm.***Ouanoughi Yasmina** *qui a bien voulu me proposer ce thème et m'aider au cours de sa réalisation.*

Je tiens à remercier :Dr. **Hacene Nacer** *pour ces vonseils, ses orientations, et la confiance qu'il m'accorde.*

Je tien à remercie les membres du Jury qui m'ont fait l'honneur de participer à ma soutenance.

Je remercie ma famille à qui je n'ai jamais su dire tout l'affection que j'ai pour aux, mon père, ma mère, mes frères et ma soeur qui ont été et seront toujours présents à mes côtés, merci pour votre soutient et vos encouragements.

Un grand merci particulier à mes collègues et mes amies pour les sympathiques moments qu'on a passés ensemble, on les remerci pour leur confiance, et leur soutien moral au cours de ces années.

Que tout ceux, que je n'ai pas nommés.

Merci

Table des matières

Liste des tableaux

Introduction

La popularité des modèles gaussiens tient à ce qu'ils se prêtent bien au calcul et les problèmes relatifs à de tels modèles reçoivent souvent une solution analytique complète sous une forme agréable. Cet avantage réside principalement dans les deux propriétés de stabilité suivantes : d'une part les transformations affines des vecteurs gaussiens produisent des vecteurs gaussiens, d'autre part lorsque deux vecteurs sont gaussiens dans leur ensemble, le conditionnement de l'un par l'autre conduit à un vecteur gaussien qui est une combinaison affine du vecteur conditionnant Cependant, le seul fait que le calcul gaussien soit agréable ne suffirait pas à justifier l'usage des modèles gaussiens.

Il y a une raison profonde qui fait que la distribution gaussienne apparait naturellement dans les phénomènes aléatoires dont la base physique est de nature microscopique et qu'on observe à l'échelle macroscopique. Les petite contributions supposées indépendantes (mais cette hypothèse n'est pas absolument nécessaire) de chaque électron, s'ajoute pour former un courant gaussien, et cela quelle que soit la distribution statistique des contributions individuelles.

Les variables gaussiennes sont des lois universelles qui sont d'une importance particulière. Il sera très utile en particulier en statistique de pouvoir passer du cas de la dimension 1 à des dimensions plus grandes. Ce passage nécessite de développer un formalisme qui sera l'objet du chapitre (1). Dans le chapitre 2, nous calculerons la fonction caractéristique des vecteurs gaussiens que nous aurons préalablement définis. Nous montrerons aussi comment nous pouvons alors aborder de manière très efficace et très simple les problèmes d'indépendance de variables gaussiennes. Nous énoncerons ensuite le théorème central limite vectoriel. Nous terminerons par la propriété suivante : le conditionnement de deux vecteurs gaussien conduit à un vecteur gaussien qui est une combinaison affine du vecteur conditionnant.

Chapitre 1

Variables aléatoires

1.1 Modèle de Probabilité

Définition 1.1 (Tribu) *On dit que T est une tribu de l'univers Ω si, et seulement si,*

1. $\Omega \in T$.
2. Pour tout $A \in T$, $\overline{A} \in T$.
3. Pour toute famille d'événements $(A_k)_{k \in \mathbb{N}}$ telle que, pour $k \in \mathbb{N}$ $A_k \in T$ $\cup_{k=0}^{\infty} A_k \in T$.[3]

Définition 1.2 (Probabilités) *On appelle probabilité définie sur (Ω, T) toute application $P : T \to \mathbb{R}$ telle que :*

i $P(\Omega) = 1$.

ii *Pour tout suit A_n d'événement incompatibles, soit $A_n \in T$ avec $A_n \cap A_m = \varnothing$ pour $m \neq n$:*

$$P(\cup_{n=0}^{+\infty} A_n) = \sum_{n=0}^{+\infty} P(A_n). \tag{1.1}$$

1.1.1 Espace de Probabilité

Le triplet(Ω, T, P) est appelé espace probabilisé si Ω est fini ou infini dénombrable (comme, par exemple, $\Omega = \mathbb{N}$), on prend $T = P(\Omega)$, l'ensemble des parties de Ω.

Propriètès élémentaires :

1. $P(\overline{A}) = 1 - P(A)$.
2. $P(A) = P(A \cap B) + P(A \cap \overline{B})$.
3. $A \subseteq B \Rightarrow P(A) \leq P(B)$.
4. Si Ω est finie ou infinie dénombrable :

$$P(A) = \sum_{\omega \in A} P(\{\omega\}), \quad \sum_{\omega \in \Omega} P(\{\omega\}) = 1. \tag{1.2}$$

1.2 Variables discrètes

Définition 1.3 (Variables discrètes) *Si X est une variable que ses valeurs dant un ensemble fini ou dénombrable I. On dit qu'elle est discrètes pour ce type de variable on définit la fonction poids en chaque point :*

$$P(i) = P(X = i), i \in I. \tag{1.3}$$

1.2.1 Fonction de répartition

Définition 1.4 (Fonction de répartition) *On appelle fonction de répartition de la v.a.X, la fonction F_X définite sur $\mathbb{R}$ par*

$$F_X\,(x) = P(X \leq x) = P_X(] - \infty, x]). \tag{1.4}$$

Propriétés :

- F_X fonction croissante au sens large. $\left(\text{i, e si } x \leq x^{'} \;\Rightarrow\; F_X\,(x) \leq F_{\acute{X}}\left(x^{'}\right)\right)$.
- $\lim_{x \to -\infty} F_X(x) = 0$.
- $\lim_{x \to +\infty} F_X(x) = 1$.
- F_X est continue à droite en tout point de $\mathbb{R}$.
- F_X n'est pas nécessairement continue à gauche.[11]

1.2.2 Moments d'une Variable aléatoire discète

Définition 1.5 (L'espérance mathématique) *L'espérence ou moyenne d'une v.a discrète X est, le réel*

$$E(X) = \sum_{k} kP[X = k], \tag{1.5}$$

ou on somme sur toutes les valeurs k que prendre X.

Propiétés Soient X et Y deux $v.a.$on a les propriétés :

1. $E(X)$ est finie ssi $E(|X|)$est finie.
2. $|X| \leq Y$ et $E(Y)$ est finie entrainent $E(X)$ finie.
3. -$\infty < a \leq X \leq b < +\infty \Rightarrow a \leq E(X) \leq b$.
4. $X = a$ p.s $\Rightarrow E(X) = a$.
5. $E(X)$ finie suivant consigne les propriétés de l'esperance mathématique.

Théorème 1.1 *Soit X et Y deux v.a.discrètes*

Si $E|X| < \infty$ et $E|Y| < \infty$, alors on a les propriétes :

A :lineairité :

(1)$E(X + Y) = E(X) + E(Y)$.

(2)$E(\lambda X) = \lambda E(X) \quad \forall \lambda \in \mathbb{R}$.

B :Monotonie :

(1)$X \geq 0 \Rightarrow E(X) \geq 0$.

(2)$X \geq Y \Rightarrow E(X) \geq E(Y)$.

(3)$X = Y$ p.s $\Rightarrow E(X) = E(Y)$.

C :Indépendance : Si X et Y sont indépendance, alors $E(XY)$ est finie et l'on a

$$E(XY) = E(X)E(Y).$$

[8]

Théorème 1.2 *Pour tout fonction g,*

$$E(g(X)) = \sum_k g(k)P(X = k). \tag{1.6}$$

Définition 1.6 (Variance) *Nous dirons qu'une v.a X de carré integrable si $E(X^2) < +\infty$. L'ensemble des v.a.de carré intégrable est noté L^2.*

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

$$Var(X) = E(X - E(X))^2. \tag{1.7}$$

On a

$$\begin{aligned} Var(X) &= E(X - E(X))^2 \\ &= \sum_k (k - E(X))^2 P[X = k] \\ &= E\left(X^2\right) - E(X)^2. \end{aligned}$$

La variance mesure le carré d'une distance a la moyenne.

Ecart-Type de X :

$$\sigma(X) = \sqrt{Var(X)}. \tag{1.8}$$

[7]

Propriétés de la variance :

– Soient a et b deux constantes et X $v.a$:

$$V(aX+b) = a^2.V(X).$$

En effet :

$$V(aX+b) = E((aX+b) - E(aX+b))^2,$$

or, d'aprées la première propriété de l'espérance mathématique et aprés simplification :

$$\begin{aligned} V(aX+b) &= E[a^2(X-E(X))^2] \\ &= a^2E(X-E(X))^2 \\ &= a^2V(X). \end{aligned}$$

– Soient X et Y deux $v.a$ independantes :

$$\begin{aligned} V(X+Y) &= E[X+Y-E(X+Y)]^2 \\ &= E[X-E(X)+Y-E(Y)]^2 \\ &= E[X-E(X)]^2 + E[Y-E(Y)]^2 + 2E[(X-E(X))(Y-E(Y))] \\ &= Var(X) + Var(Y). \end{aligned}$$

Ce résultat se généralisé pour n $v.a$ indépendante. Soient $X_1, X_2, ..., X_n$ n $v.a$ indépendante, nous obtenons :

$$V(X_1+X_2+...+X_n) = V(X_1)+V(X_2)+...+V(X_n).$$

Donc pour les $v.a$ indépendante, la variance d'une somme est égale à la somme des variances.

Définition 1.7 *Nous dirons qu'une v.a de carré intégrable X est centrée réduite si*

$$E(X) = 0 \quad ; \quad Var(X) = 1.$$

Proposition 1.1 *Soit Y v.a de carré intégrable. Alors, la v.a.X définie par :*

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

est centée réduite.

1.2.3 Lois discrètes classiques

Loi de Bernoulli

Définition 1.8 *La variable aléatoire X suit la loi de Bernoulli de paramètre p ($p \in [0,1]$) si elle ne prend que deux valeurs 0 et 1 avec :*

$$\begin{cases} P(X=1) & = & p. \\ P(X=0) & = & 1-p & = q. \end{cases} \tag{1.9}$$

Notation 1 $X \rightsquigarrow B(p)$.

Proposition 1.2 *Si X suit la loi de Bernoulli de paramétre p*

$$E(X) = p \qquad et \qquad V(X) = p(1-p)$$

Loi uniforme sur un ensemble fini de rèels

Définition 1.9 *La variable aléatoire X suit la loi uniforme sur l'ensemble de réels $\{x_1, ..., x_n\}$ si P_x est l'équiprobabilité sur cet ensemble.*

Autrement dit, l'ensemble des valeurs possibles de X est $X(\Omega) = \{x_1, ..., x_n\}$ et :

$$\forall k = 1, ..., n, \quad P(X = x_k) = \frac{1}{n}. \tag{1.10}$$

Par exemple le nombre de points indiqué par un dé suit la loi uniforme sur $\{1,2,3,4,5,6\}$.

Loi binomiale

Définition 1.10 *La variable aléatoire X suit la loi binomiale de paramétres n et p ($n \in \mathbb{N}^*$ et $p \in [0,1]$) si l'ensemble des valeurs possibles est $X(\Omega) = \{1,...,n\}$ et*

$$\forall k = 1,...,n, \;\; P(X = k) = C_n^k p^k (1-p)^{n-k}. \tag{1.11}$$

Notation 2 *$X \rightsquigarrow B(n,p)$.*

La formule ci-dessus définit bien une loi de probabilité puisque les $C_n^k p^k (1-p)^{n-k}$ sont positifs et :

$$\Sigma_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1.$$

Théorème 1.3 *La somme de n v.a discrètes de Bernoulli indépendant, de même paramétre p, suit la loi binomaile $B(n,p)$.*

Proposition 1.3 *Si X suit la loi binomaile $B(n,p)$*

$$E(X) = np \qquad et \qquad V(X) = np(1-p)$$

Loi géométrique

Définition 1.11 *On dit que X est le temps d'attente du premier événement A.*

Soit $p \in \,]0,1[$. On dit qu'une $v.a$ X suit la loi géométrique de paramétre p, (note $G(p)$) lorsque X prend les valeurs $n \in \mathbb{N}^*$avec les probabilités

$$P(X = n) = p(1-p)^{n-1}. \tag{1.12}$$

Proposition 1.4 *Si X suit la loi géométrique de paramétre p. On a :*

$$E(X) = \frac{1}{p} \qquad et \qquad V(X) = \frac{(1-p)}{p^2}$$

Loi de Poisson

Définition 1.12 *On dit que la variable alétoire discète X suit la loi de Poisson de paramétre $\lambda > 0$ si l'ensemble des valeurs possibles est $X(\Omega) = \mathbb{N}$ et*

$$\forall k \in \mathbb{N}, P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}. \tag{1.13}$$

Notation 3 $X \rightsquigarrow P(\lambda)$.

Proposition 1.5 *Si X suit la loi de Poisson $P(\lambda)$, on a :*

$$E(X) = \lambda \quad et \qquad V(X) = \lambda$$

1.2.4 Fonction caractéristique

Définition 1.13 (Fonction caractéristique) *Soit X une v.a.discrète de fonction de masse P. On appelle fonction caractéristique de X, la fonction de la variable t ($t \in \mathbb{R}$) définie par :*

$$\phi_X(t) = E(e^{itX}) = \sum_{x \in \mathbb{R}} e^{itx} p(x). \tag{1.14}$$

Théorème 1.4 *Si X et Y deux v.a indépendante alors :*

$$\phi_{(X+Y)}(t) = \phi_X(t)\phi_Y(t).$$

Preuve. On a

$$\begin{aligned}\phi_{(X+Y)}(t) &= E(e^{it(X+Y)}) = E(e^{itX}e^{itY}) \\ &= E(e^{itX})E(e^{itY}) \\ &= \phi_X(t)\phi_Y(t).\end{aligned}$$

■

Exemple de fonction caractéristique

Loi de Bernoulli B(1,p)	$\phi_X(t) = 1 - p(1 - e^{it})$
Loi binomiale B(n,p)	$\phi_X(t) = (1 - p + pe^{it})^n$
Loi de Poisson P(λ)	$\phi_X(t) = e^{\lambda(e^{it}-1)}$
Loi géométrique G(p)	$\phi_X(t) = \frac{pe^{it}}{(1-(1-p)e^{it})}$

TAB. 1.1 – fonction caractéristique de loi discrètes

[12]

1.2.5 Fonction génératrice

Définition 1.14 (Fonction génératrice) *La fonction génératrice de X est la fonction définie pour $s \in [0,1]$ par :*

$$G_X(s) = E(s^X) = \sum_{n\in\mathbb{N}} P(X = n)s^n. \tag{1.15}$$

C'est la somme d'une série entière dont le rayon de convergence est au moins 1*, car* $\sum_n P_n = 1$.
Cette fonction caractérise la loi de X.

1.2.6 Couple de v.a discrètes

Loi d'un couple

Un couple de *v.a.* discrètes est constitué de deux *v.a.* discrètes X et Y dont l'ensmble des valeurs possibles peut s'écrire respectivement sous la forme $\{x_i\}_{i\in I}$ et $\{y_j\}_{j\in J}$, où I et J sont des ensembles d'indices inclus dans $\mathbb{N}$, pouvant d'ailleurs être $\mathbb{N}$ tout entier. On convient de ne faire figurer que des valeurs de probabilité strictemeent positive. Comme dans le cas unidimensionnel, la loi d'un couple discrèt est définie par l'ansemble des valeurs possibles, soit ici $\{(x_i, y_j);\ (i,j) \in I \times J\}$, et par les probabilités associéés :

$$P_{ij} = P_{X,Y} = P(X = x_i \cap Y = y_j). \tag{1.16}$$

[6]

Lois marginales

La loi d'un couple est associé au deux lois marginales qui sont les lois de chacun des éléments du couple pris séparément, définies par l'ensemble des valeurs possibles et les probabilités associées obtenues par sommation, soit :

$$\begin{cases} P_X(X = x_i) = \sum_{j\in J} P_{(X,Y)}(X = x_i, Y = y_j) = \sum_{j\in J} P_{ij} = P_{i\cdot} \\ P_Y(Y = y_j) = \sum_{i\in I} P_{(X,Y)}(X = x_i, Y = y_j) = \sum_{i\in I} P_{ij} = P_{\cdot j} \end{cases} \tag{1.17}$$

[6]

Lois conditionneles

Les événements $\{X = x_i\}$ et $\{Y = y_j\}$ étant de probabilités non nulles on définit alors deux familles de lois conditionnelles selon que l'on connaît la « valeur » de X ou de Y. Rappelons qu'ici X et Y ne sont pas forcément des variables aléatoires réelles mais peuvent être des variables qualitatives

– Lois conditionneles de X si $Y = y_j$:

$$P(X = x_i/Y = y_j) = \frac{P_{ij}}{P_j} = \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}. \tag{1.18}$$

– Lois conditionneles de Y si $X = x_i$:

$$P(Y = y_j/X = x_i) = \frac{P_{ij}}{P_i} = \frac{P(X = x_i \cap Y = y_j)}{P(X = x_i)}. \tag{1.19}$$

le théorème des probabilités totales (deuxième forme) permet d'écrire :

$$\begin{aligned} P(Y = y_j \cap X = x_i) &= \textstyle\sum_{j \in J} P(X = x_i / Y = y_j) P(Y = y_j) \\ &= \textstyle\sum_{i \in I} P(Y = y_j / X = x_i) P(X = x_i). \end{aligned}$$

[6]

Loi d'une somme

Si X et Y sont deux $v.a.$discrètes de lois respectives $\{(x_i, p_i); i \in I\}$ et $\{(y_j, q_j); j \in J\}$. La $v.a$ $Z = X + Y$ est aussi une $v.a$ discrètes dont la loi de probabilité est définie par l'ensemble des valeurs possible, soit ici $\{(x_i + y_j); i \in I, j \in J\}$, et les probabilité associeés :

$$P(Z = z_k) = \sum \{P(X = x_i, Y = y_j) / x_i + y_j = z_k\}. \tag{1.20}$$

Dans la cas générale cette formule ne peus donc être définie que par la donnée de la loi du couple (X, Y).

– **Cas particulier :** X et Y sont indépendantes.

  On parle alors de convolution des lois X et Y, qui est définie par :

$$\begin{aligned} P(Z = z_k) &= \textstyle\sum_{i \in I} P(X = x_i) P(Y = z_k - x_i) \\ &= \textstyle\sum_{j \in J} P(Y = y_j) P(X = z_k - y_j). \end{aligned}$$

Exemple 1.1 *Convolution de lois de poisson.*

Si X et Y sont deux $v.a$ indépendantes de lois de poisson respectives $P(\lambda)$ et $P(\mu)$,alors

$Z = X + Y$ est une v.a à valeurs dans $\mathbb{N}$, avec pour tout entier k :

$$\begin{aligned}
P(Z = k) &= \sum_{x+y=k} P(X = x)P(Y = y) = \sum_{x \in \mathbb{N}} P(X = x)P(Y = k - x) \\
&= \sum_{x=0}^{k} P(X = x)P(Y = k - x) \\
&= \sum_{x=0}^{k} e^{-\lambda} \frac{\lambda^x}{x!} e^{-\mu} \frac{\mu^{k-x}}{(k-x)!} \\
&= e^{-(\lambda+\mu)} \sum_{x=0}^{k} \frac{k!}{x!(k-x)!} \lambda^x \mu^{k-x} \\
&= \frac{e^{-(\lambda+\mu)}}{k!} (\lambda + \mu)^k.
\end{aligned}$$

on retrouve bien comme résultat : $X + Y \rightsquigarrow P(\lambda + \mu)$.

1.2.7 Relation entre deux v.a

Théorème 1.5 *Soit f une application de $\mathbb{R}^2 dans \mathbb{R}$, telle que $f(X, Y)$ admette une espérance. Alors*

$$\begin{aligned}
E(f(X, Y)) &= \sum_{i,j} f(x_i, y_j) P([X = x_i] \cap [Y = y_j]) \\
&= \sum_{i,j} f(x_i, y_j) P_{ij}.
\end{aligned}$$

Propriété : Si X et Y sont indépendantes. Alors $E(XY) = E(X)E(Y)$.

Remarque 1.1 *La réciproque est engénérale fausse.*

Définition 1.15 (Covariance) *Soient X et Y deux v.a r d'ordre 2 .On appelle :covariance de X et Y, le réel $Cov(X, Y)$ définie par :*

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y). \tag{1.21}$$

***cofficient de corrélation** (linèaire) de X et de Y le réel $\varphi(X,Y)$ est définie par*

$$\varphi_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}. \tag{1.22}$$

Matrice de covariance : De x et de y, la matrice $\Gamma(X,Y)$ définie par :

$$\Gamma(X,Y) = \begin{pmatrix} Var(X) & Cov(X,Y) \\ Cov(Y,X) & Var(Y) \end{pmatrix} = \begin{pmatrix} Cov(X,X) & Cov(X,Y) \\ Cov(Y,X) & Cov(Y,Y) \end{pmatrix}. \tag{1.23}$$

Propriétés

1. $Cov(X,Y) = Cov(Y,X) = E(XY) - E(X)E(Y)$.
2. $Var(X+Y) = Var(X) + Var(Y) + Cov(X,Y)$.
3. Si X et Y sont indépendance, alors $Cov(X,Y) = 0$.
4. $Cov(aX + bY, Z) = aCov(X,Z) + bCov(Y,Z)$.
5. $\rho(X,Y) \in [-1,1]$ et $\rho(aX+b, cY+d) = \frac{ac}{|ac|}\rho(X,Y)$.
6. $\Gamma(X,Y)$ est matrice réelle symétrique et, pour tout $(u,v) \in \mathbb{R}^2$

$$(u,v)\Gamma(X,Y)(u,v)^t \geq 0.$$

1.3 Variable continue

Il existe des variable alétoire dont les valeurs possibles ne sont pas dans une ensemble fini ou dénombrable mais dans un ensemble dénombrable.

Soit X ce genre de v.on dit qu'elle est continue s'il existe une fonction positive f definie pour tout réel $x \in \mathbb{R}$tel que pour tout partie B de nombres réels $P(X \in B) = \int_B f(x)dx$

$$P(X \in]+\infty, -\infty[) = \int_{-\infty}^{+\infty} f(x)dx. \tag{1.24}$$

$$\begin{cases} si \quad B = [a,b] \qquad P(a < X < b) = \int_a^b f(x)dx. \\ si \qquad a = b \qquad P(X = a) = \int_a^a f(x) = dx = 0. \end{cases}$$

La probabilité qu'une variable continue soit égale à une valeur fixée est nulle

$$P(X < a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx. \tag{1.25}$$

1.3.1 Fonction de répartition

D'une façon générale, désignons par X une $v.a$ continue prenant ses valeurs sur l'ensemble des nombres réel $\mathbb{R}$. Soit x un nombre réel particulier, la probabilité que X prenne une valeur inférieure ou égale à x est exprimée par :

$$F(x) = P(X \leq x).$$

La fonction $F(x)$ est appelée la fonction de répartition de X.

Propriétés de la fonction de répartition

1. $F(x)$ est une fonction continue dérivable.
2. $F(x)$ est une fonction croissante pour tout x.

1.3.2 Fonction de densité

Définition 1.16 (Fonction de densité) *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 tout $x \in \mathbb{R}$ on a :*

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt. \tag{1.26}$$

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

$$\forall A \in T,\ P_X(A) = P(X \in A) = \int_A f_X(t)dt.$$

Théorème 1.6 *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

$$\int_{\mathbb{R}} f(t)dt = 1.$$

Proposition 1.6 *Si f_X est continue sur une intevalle $[a, b]$, alors F_X est dérivable sur $[a, b]$ est on a $f_X = F_X^{'}$.*

La probabilité que la $v.a$ X prenne une valeur dans l'intervalle $[a, b]$ est :

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x)dx.$$

On en déduit que $P(X = a) = 0$.

1.3.3 Caractéristiques des variables aléatoires

Espérance mathématique

Soit X un variable aléatoire. On appelle espérance mathématique moment d'ordre 1 par rapport à l'origine la quantité $E[X]$ définie par :

$$E[X] = \int_{-\infty}^{+\infty} xf(x)dx. \tag{1.27}$$

L'espérance d'une fonction aléatoire

Si X est une variable continue de densité de probabilité$f(x)$ alor pour tout fonction réelle g

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx. \tag{1.28}$$

Variance

Si X est une v.a de moyenne $E[X]$, alors la variance de X est définie par :

$$\begin{aligned} V(X) &= E[(X - E(X))^2] \\ &= E[X^2] - E[X]^2. \end{aligned}$$

1.3.4 Loi de probabilité continue

Loi uniforme

Définition 1.17 (Loi uniforme) *La v.a.Continue X suit la loi uniforme sur l'intervalle $[a, b]$, $(-\infty < a < b < +\infty)$ si elle a une densité f constante sur cet intervalle et nulle en dehors .On a alors :*

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

La fonction de répartition F est affine par morceaux :

$$F(x) = \begin{cases} 0 & si \quad -\infty < x \leq a. \\ \frac{x-a}{b-a} & si \quad a < x \leq b. \\ 1 & si \quad b < x < +\infty. \end{cases} \tag{1.29}$$

Proposition 1.7 *Si X suit la loi uniforme sur $[a, b]$, alors :*

$$E(X) = \frac{b+a}{2} \qquad et \qquad V(X) = \frac{(b-a)^2}{12}$$

Loi exponentielle

Définition 1.18 (Loi exponentielle) *Soit α une réel strictement positive. La v.a continue X suit la loi exponentielle de paramétre α si elle admet pour densité*

$$f(x) = \alpha e^{-\alpha t} 1_{[0,+\infty[}(x). \tag{1.30}$$

Proposition 1.8 *Si X suit la loi $exp(\alpha)$ alors :*

$$E(x) = \frac{1}{\alpha} \qquad et \qquad V(X) = \frac{1}{\alpha^2}$$

Loi normale

Définition 1.19 (Loi normale) *On dit que la v.a.X suit la loi normale $N(m,\sigma)$ si elle a pour densité la fonction*

$$\begin{array}{ccc} f_{m,\sigma} & : & \mathbb{R} \to \mathbb{R}^+ \\ x & \to & \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-m}{\sigma})^2} \end{array}. \tag{1.31}$$

Définition 1.20 *Une v.a continue X suit la loi normale centrée réduite notée $N(0,1)$ lors que X admet pour densité la fonction définite par :*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{1.32}$$

Proprétés :

1. Si X suit la loi normale $N(m, \sigma)$, on a :

$$E(X) = m \quad \text{et} \quad V(X) = \sigma^2$$

2. Si X suit la loi normale $N(0,1)$ on a :

$$E(X) = 0 \quad \text{et} \qquad V(X) = 1$$

3. Si X une $v.a$ continue

$$X \sim N(m, \sigma) \Longleftrightarrow Y = \frac{X - m}{\sigma} \sim N(0,1)$$

Loi Gamma

Définition 1.21 (Loi Gamma) *Une $v.a$ continue X suit une loi gamma $\gamma(r, \lambda)$, de paramétres positifs r et λ, si sa densité de probabilité est donné par*

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{r-1}}{\Upsilon(x)} & si \quad x \geq 0 \\ 0 & sinon \end{cases} \tag{1.33}$$

Υ *est la fonction enlérienne définie par l'intégrale pour $r > 0$*

$$\Upsilon(r) = \int_0^{+\infty} e^{-y} y^{r-1} dy.$$

– En particulier, si $r = 1$, on retrouve la loi exp.

– Si la paramètre λ est différent de 1, la v.a $Y = \lambda X$ suit une loi $\gamma(r, 0)$ ou $\gamma(1)$ de densité :

$$f(x) = \begin{cases} \frac{e^{-y} y^{r-1}}{\Upsilon(r)} & si \quad y \geq 0 \\ 0 & \text{si non} \end{cases}$$

Proprétés de la fonction $\Upsilon(r)$

– $\Upsilon(r) = (r-1)\Upsilon(r-1);$

– $\Upsilon(1) = 1;$

– $\Upsilon(n) = (n-1);$

Moment

Par intégration par parties et en utilisant les propriétés de la fonction Υ, on obtient :

$$E(X) = \frac{r}{\lambda} \qquad \text{et} \qquad V(X) = \frac{r}{\lambda^2}$$

Loi Bêta

Définition 1.22 (Loi Bêta) *Une v.a.absolument continue suit une loi bêta si sa densité de probabilité est :*

$$f(x) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} & si \quad 0 < x < 1 \\ 0 & sinon \end{cases} \tag{1.34}$$

ou α et β sont constantes et $B(\alpha, \beta)$ est la fonction béta

$$\begin{aligned} B(\alpha, \beta) &= \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \\ &= \frac{\Upsilon(\alpha).\Upsilon(\beta)}{\Upsilon(\alpha+\beta)}. \end{aligned}$$

On a

$$E(X) = \frac{\alpha}{\alpha+\beta} \qquad et \qquad V(X) = \frac{\alpha.\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Loi de Cauchy

Définition 1.23 (Loi de Cauchy) *Une v.a.X absolument continue suit une loi de cauchy $C(0, a)$ si sa densité de probabilité est :*

$$f(x) = \frac{a}{\pi(x^2+a^2)}, a > 0, x \in \mathbb{R}. \tag{1.35}$$

$E(X^k)$ *n'existe pas ; $k \geq 1$.*

Loi de khi-deux

Définition 1.24 (Loi de khi-deux) *Soient k v.a.$X_1, X_2, ..., X_k$ indépendante suivant une loi normal $N(m_i, \sigma_i)$, $i = 1, ..., k$.*

La variable aléatoire

$$Z = \sum_{i=1}^{k} (\frac{X_i - m_i}{\sigma_i})^2. \tag{1.36}$$

suit un loi de khi-deux (χ_k^2) *à* k *dégrés de liberté*

– La densité de probabilité de la $v.a$ Z est :

$$f(x) = \begin{cases} \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} e^{\frac{-x}{2}} x^{\frac{k}{2}-1} & si \quad x \geq 0. \\ 0 & si \quad x < 0. \end{cases}$$

On a $E(Z) = k$ et $V(Z) = 2k$.

– Cette loi est sur tout utile dans les tests statistique.

Loi de Student

Soient deux $v.a$ indépendantes X et Y distribuées suivant une loi normale $N(0,1)$ et celle du χ_k^2 respectivement. La variable

$$T = \frac{X}{\sqrt{\frac{Y}{k}}}. \tag{1.37}$$

Suit une loi de student à k dégres de liberté.

On a

$$E(T) = 0, \ k > 0 \quad \text{et} \quad V(T) = \frac{k}{k-2}, \ k > 0$$

Si on a $k = 1$, alors on obtient la loi de Cauchy dont la densité est :

$$g(x) = \frac{1}{\pi(1 + x^2)}.$$

D'une façon générale la densité T s'écrit :

$$g(x) = \frac{1}{\sqrt{k}B(\frac{1}{2}, \frac{k}{2})} . \frac{1}{(1 + \frac{x^2}{k})^{\frac{k+1}{2}}}.$$

La loi T de student est utilisée sur tout dans les tests statistique.

Exemples de fonction caractéristique

loi uniform $U[a,b]$	$\phi_x(t) = \frac{e^{itb}-e^{ita}}{it(b-a)}$
loi normale$N(0,1)$	$\phi_x(t) = e^{\frac{-t^2}{2}}$
loi normale $N\,(m,\sigma)$	$\phi_x(t) = e^{itm}e^{-\frac{t^2\sigma^2}{2}}$
loi exp $\xi(\lambda)$	$\phi_x(t) = \frac{\lambda}{\lambda-it}$
loi gamma $\Gamma(r,\lambda)$	$\phi_x(t) = (\frac{\lambda}{\lambda-it})^r$

TAB. 1.2 – fonction caractéristique de loi continues

[12]

1.3.5 Couple de v.a continue

Loi du couple

Si X et Y sont deux $v.a.r$ continue, la loi de probabilité du couple (X,Y) est déterminée par sa fonction de répartition F définie sur $\mathbb{R}^2$ par :

$$F(x,y) = P(X < x, Y < y). \tag{1.38}$$

Propriété : les fonction de répartition des $v.a.r.X$ et Y vérfient

$$F_X(x) = \lim_{y\to+\infty} F_{X,Y}(x,y) \qquad \text{et} \qquad F_Y(y) = \lim_{x\to+\infty} F_{X,Y}(x,y).$$

Définition 1.25 *La loi de couple (X,Y) est dite absolument continue s'il existe une application $f_{X,Y}$ de $\mathbb{R}^2$ sur $\mathbb{R}$, appelée densité du couple (X,Y), continue sur l'intérieur d'un sous ensemble D de $\mathbb{R}^2$ et nulle sur son complémentaire, telle que, pour tout $(x,y) \in \mathbb{R}^2$:*

$$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)dudv. \tag{1.39}$$

Propriétés :

1. $\int_{-\infty}^{+\infty}(\int_{-\infty}^{+\infty} f_{X,Y}(u,v)du)dv = \int_{-\infty}^{+\infty}(\int_{-\infty}^{+\infty} f_{X,Y}(u,v)dv)du = 1.$
2. En tout (x_0,y_0) ou $f_{X,Y}$ est continue, on a $f_{X,Y}(x_0y_0) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x_0,y_0).$

Loi marginales

Les fonction de répartition marginales de X et Y sont définies à partir de la fonction de répartition.du couple par :

$$\begin{cases} F_X(x) = P(X < x) = P_X\left(-\infty, x\right). \\ F_Y(y) = P(Y < y) = P_Y\left(-\infty, y\right). \end{cases} \tag{1.40}$$

c'est-à-dire en faisant tendre y, respactivenemt x, vers plus l'infini. Dans le cas d'une loi absolument continue, les densités marginales sont alors obtenues par dérivation de ces fonction de répartition marginales.

Cependant, si la loi du couple est définie par sa densité, les densités marginales sont obtenues plutôt par intégration :

$$\begin{cases} f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy. \\ f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx. \end{cases} \tag{1.41}$$

Lois conditionnelles

Définition 1.26 *Si l'une des v.a.X ou Y a une valeur fixée, les lois conditionnelles sont définies par les densités conditionnelles :*

$$\begin{cases} f_X(x/Y=y) = \frac{f(x,y)}{f_Y(y)}. \\ f_Y(y/X=x) = \frac{f(x,y)}{f_X(x)}. \end{cases} \tag{1.42}$$

à condition bien sûr que $f_Y(y) > 0$ et $f_X(x) > 0$.

Moment associés à un couple

Si $h : \mathbb{R}^2 \to \mathbb{R}$ est une application continue, l'espérance de $h(X,Y)$se calcule pour une loi de densité f par l'intégrale :

$$E[h(X,Y)] = \int_{\mathbb{R}} \int_{\mathbb{R}} h(x,y) f(x,y) dxdy. \tag{1.43}$$

Dans le cas particulier où $h(X,Y) = [X - E(X)][Y - E(Y)]$, ceci définit la covariance :

$$Cov = E\left(\left(X - E(X)\right)\left(Y - E(Y)\right)\right). \tag{1.44}$$

Dans le cas particulier où $v.a.X$ et Y sont indépendantes :

$$\begin{aligned} E(XY) &= \int_{\mathbb{R}}\int_{\mathbb{R}} xyf(x,y)dxdy \\ &= \int_{\mathbb{R}} xf_X(x)dx \int_{\mathbb{R}} yf_Y(y)dy \\ &= E(X)E(Y). \end{aligned}$$

et par conséquent

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

Indépendace

Définition 1.27 *Deux $v.a.r.X$ et Y sont dites indépendante si, pour tout $(x,y) \in \mathbb{R}^2$, on a :*

$$F_{X,Y}(x,y) = F_X(x)F_Y(y). \tag{1.45}$$

On admet que ceci équivaut à $f_{X,Y}(x,y) = f\ _X(x)f_Y(y)$ pour tout $(x,y) \in \mathbb{R}^2$ ou bien à $E(g(X)h(Y)) = E(g(X))E(h(Y))$ pour tout fonction g et h pour vu que ces espérance existent.

Propriété : Si X et Y sont indépendantes, alors$f_Y(y/X = x) = f_Y(y)$ pour tout x tel que $f_X(x) \neq 0$ et $f_X(x/Y = y) = f_X(x)$ pour tout y tel que $f_Y(y) \neq 0$.

Régression

Les densités conditinnelles permettent de calculer les moments conditionnelles, comme par exemple les espérance ; on peut définir notamment la fonction

$$x \longmapsto E(Y/X = x)$$

que s'appelle fonction de régession, son graphe étant la courbe de régression de Y en X.

Pour une loi absolument continue, on obtient :

$$E(Y/X=x)=\int_{\mathbb{R}} y f_Y(y/X=x)dy=\int_{\mathbb{R}} y\frac{f(x,y)}{f_X(x)}dy. \tag{1.46}$$

Changement de variables

Théorème 1.7 *Soit (X,Y)un couple aléatoire de densité f_{XY} et δ une fonction de $\mathbb{R}^2$ sur $\mathbb{R}$. Si $f_{X,Y}$ est continue sur l'interieur d'un ensemble D et nulle sur son comlémentaire, si δ est un bijection de D*

$$J(\delta^{-1})=\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \neq 0.$$

sur E, alors le couple aléatoire $(u,v)=\delta(X,Y)$ admet pour densité la fonction $f_{U,V}$ définie par

$$f_{U,V}(u,v)=f_{X,Y}(\delta^{-1}(u,v))\left|J(\delta^{-1})(u,v)\right| \text{ si } (u,v)\in E. \tag{1.47}$$

Loi d'une somm

La loi de $v.a.Z=X+Y$ se détermine par sa $f.r$, définie par :

$$H(z)=P(Z\le z)=P(X+Y\le z),$$

qui ne peut se calculer que si l'on connaît la loi du couple (X,Y).

Dans le cas particulier où ce couple admet une densité f, on obient.

$$H(z)=\int_D\int f(x,y)dxdy, \tag{1.48}$$

où le domaine d'intégration est $D=\{(x,y)/x+y<z\}$. On effectue le changement de variable suivant :

$$\begin{cases} u=x \\ v=x+y \end{cases} \quad \begin{cases} x=u \\ y=v-u \end{cases}$$

le jacobien de cette trensformation est :

$$\frac{D(x,y)}{D(u;v)} = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1,$$

d'où l'intégrale :

$$H(z) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} f(x,y)dy = \int_{-\infty}^{+\infty} du \int_{-\infty}^{z} f(u,v-u)dv,$$

ce qui permet de mettre en évidence la densité de Z, soit :

$$h(v) = \int_{-\infty}^{+\infty} f\left(u, v-u\right) du.$$

[5]

Chapitre 2

Vecteurs Gaussiens

2.1 Vecteurs aléatoires

Définition 2.1 *On appelle vecteur aléatoire où variable aléatoire à valeurs dans* $\mathbb{R}^n$*, une application mesurable de* (Ω, T)*dans* $(\mathbb{R}^n, B(\mathbb{R}^n))$.

2.1.1 Loi d'un vecteur aléatoire

Soit (Ω, T, P) un espace de probabilité, si X est un vecteur aléatoire à valeurs dans $\mathbb{R}^n$, on appelle loi de X, la probabilité P_X définit par :

$$\forall B \in B(\mathbb{R}^n),\ P_X(B) = P(X^{-1}(B)) = P(X \in B). \tag{2.1}$$

Définition 2.2 *Si* $X = (X_1, ..., X_n)$ *est un vecteur à valeurs dans* $\mathbb{R}^n$*. La loi de la v.a* X_i $(1 \leq i \leq n)$ *est appelée la ième loi marginal.*

- La loi P_X de X est la loi joint du d-uplet $(X_1, ..., X_n)$ probabilité sur $\mathbb{R}^n$.
- La loi P_{X_i} de X_i est la ième loi marginale probabilité sur $\mathbb{R}$.

Définition 2.3 *La loi de* X *admet la densité* f *positive, intégrable sur* $\mathbb{R}^n$ *et d'intégrale* 1 *si* *et seulement si*

$$F(x_1,...,x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f(y_1,...,y_d)d_{y_1}...d_{y_n}. \tag{2.2}$$

Remarque 2.1 *On a pour tout $B \in B(\mathbb{R})$*

$$\begin{aligned} P_{X_i}(B) &= P(X_i \in B) = P\left(X_1 \in \mathbb{R}, ..., X_i \in B, ..., X_n \in \mathbb{R}\right) \\ &= P(X \in \mathbb{R} \times .... \times B \times ....\mathbb{R}) \\ &= P_X(\mathbb{R} \times ... \times B \times ...\mathbb{R}), \end{aligned}$$

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

Proposition 2.1 *Si le vecteur $X = (X_1, ..., X_n)$ admet une densité f, alors X_i $(1 \leq i \leq n)$ a pour densité la fonction f_{X_i} donnée par*

$$f_{X_i}(u) = \int_{\mathbb{R}^{n-1}} f(x_1, ..., x_{i-1}, u, x_{i+1}, ..., x_n)d_{x_1}...d_{x_{i-1}}d_{x_{i+1}}...d_{x_n}, \tag{2.3}$$

on appelle cette densité la ième densité marginale.

2.1.2 Moments d'un vecteur aléatoire

Comme pour les $v.a.r$, on définе pour les vecteurs aléatoires des moments d'ordre 1et 2.

Définition 2.4 *Si $X = (X_1, ..., X_n)$ un vecteur aléatoire lorsque chaque composante est une v.a positive ou nulle, on appelle espérance (ou moyenne d'ordre 1) de X, le vecteur de $\mathbb{R}^d$ égal a :*

$$E(X) = (E(X_1), ..., E(X_n)). \tag{2.4}$$

Lorsque $\parallel X \parallel$ est intégrables, chaque composante X_i est intégrable et la définition a bien un sens.

On peut calculer $E(X_i)$ à l'aide de la loi de X_i où à l'aid de la loi de X

$$E(X_i) = \int_\Omega X_i(\omega)dP(\omega) = \int_{\mathbb{R}} x_i dP_{X_i}(x_i) = \int_{\mathbb{R}^n} x_i dP_X(x), \tag{2.5}$$

le vecteur $E(X_i)$ décrit le comportement moyen du vecteur aléatoire X.

Proposition 2.2 *Si $X_1, ..., X_n$ sont des v.a.r de carré intégrable, on a la formule*

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) + 2 \sum_{1 \leq i < j \leq n} Cov(X_i, X_j). \tag{2.6}$$

Définition 2.5 *Si $X = (X_1, X_2, ..., X_n)$un vecteur aléatoire à valeurs dans $\mathbb{R}^n$ tel que $E(\| X \|^2) < \infty$. On appelle matrice de dispersion de X ou (matrice de variance covariance) la matrice $n \times n$ de terme générale*

$$(\Gamma_X)_{i,j} = Cov(X_i, X_j). \tag{2.7}$$

Cette matrice est bien définie car, si la v.a $\| X \|^2$ est intégrable, chaque variable X_i est intégrable.
La matrice de dispersion est une matrice symetrique. Sa daigonale principale est formée des variance de X_i.

Exemple 2.1 *Soit $X = (X_1, X_2)$un vecteur aléatoire de densité :*

$$f(x, y) = \frac{1}{2\pi\sqrt{1-b^2}} \exp(-\frac{x^2 - 2bxy + y^2}{2(1-b^2)}).$$

avec $|b| < 1$. Nous avons vu que X_1 et X_2 suivent une loi gaussinne centrée réduite. Donc $E(X_1) = E(X_2) = 0$ et $Var(X_1) = Var(X_2) = 1$. Nous connaissons donc le vecteur espérance de X et la diagonale principale de la matrice de dispersion de X. Il reste à calculer $Cov(X_1, X_2)$. Comme les moyennes de X_1 et X_2 sont nulles, on a :

$$Cov(X_1, X_2) = Cov(X_2, X_1) = E(X_1 X_2)$$

et :

$$\begin{aligned}E(X_1X_2) &= \int xy\frac{1}{2\pi\sqrt{1-b^2}}\exp(-\frac{x^2-2bxy+y^2}{2(1-b^2)})dxdy\\ &= \int \frac{1}{\sqrt{2\pi}}xe^{-\frac{x^2}{2}}\left(\int \frac{1}{\sqrt{2\pi(1-b^2)}}y\exp\left(-\frac{(y-bx)^2}{2(1-b^2)}\right)dy\right)dx\\ &= b.\end{aligned}$$

Car la parenthèse représente la moyenne d'une loi gaussienne de moyenne bx et de variance $1-b^2$ et vaut donc bx ; en factorisante par b, on reconnaît ensuite la variance d'une loi gaussienne centrée réduite qui vaut 1*. On a donc*

$$E(X) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \; et\ \Gamma_X = \begin{pmatrix} 1 & b \\ b & 1 \end{pmatrix}$$

Si on représente les vecteurs X et $E(X)$ par des matrices colonnes, on peut écrire matriciellement la matrice Γ_X sous la forme

$$\Gamma_X = E((X-E(X))(X-E(X))^t).$$

où, étant donné un vecteur (ou un matrice) $u.u^t$ désigne sont tranposé.

L'espérance d'un vecteure aléatoire est lineaire alors que la disperion est une quntité quadratique. Cela se voit, entre autre, dans les formules de la proposition suivante.

Proposition 2.3 *Si X est un vecteur aléatoire à valeurs dans $\mathbb{R}^n$, tel que $E(\| X \|^2) < \infty$. Si A est la matrice représentant une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et si B est un vecteur de $\mathbb{R}^n$, alors le vecteur aléatoire $Y = AX + B$ a valeur dans $\mathbb{R}^n$ est tel que $E(\| Y \|^2) < \infty$ et on a :*

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

$$D_Y = AD_XA^t.$$

Preuve. On a

$$\forall_{1\leq i\leq n}, X_i \in L^2(\Omega, T, P) \Rightarrow \forall_{1\leq i\leq n}, Y_i \in L^2(\Omega, T, P),$$
$$E(\parallel Y \parallel^2) = \sum_{i=1}^{n} E({Y_i}^2),$$

est finie puisque l'intégrale est linéaire alors on a

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

Calculons D_Y :

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

■

2.1.3 Fonction caractéristique

Définition 2.6 *Si X est un vecteur aléatoire à valeurs dans* $\mathbb{R}^n$ $(n \geq 1)$, *on appelle fonction caractéristique de* X *la fonction* ϕ_X *de* $\mathbb{R}^n$ *dans* $\mathbb{C}$ *définie par :*

$$\phi_X(t) = E(\exp(i \langle t, X \rangle)). \tag{2.8}$$

($\langle t, X \rangle$ *désigne le produit scalaire de* t *et de* X *dans* $\mathbb{R}^n$: $\langle t, X \rangle = \sum_{i=1}^{n} t_i X_i$).
La fonction $\exp(i \langle t, X \rangle)$ *est de norme égale à* 1 *donc est P-intégrable. Si on note* P_X *la loi de* X, *on peut écrire la fonction caractéristique* ϕ_X *de la maniére suivente :*

$$\phi_X(t) = \int_{\mathbb{R}^n} \exp(i \langle t, X \rangle) dP_X(x).$$

La fonction caractéristique de X est donc la transformée de Fourier de la loi de X. C'est une fonction continue bornée par 1.
Si X admet une densité f alors $\phi_X(t) = \int_{\mathbb{R}^n} \exp(i \langle t, X \rangle) f(x)dx$ et la fonction caractéristique de X est la transformée de Fourier de la fonction f.[5]

Propriété : Deux vecteurs aléatoires X et Y ont même loi si et seulement si

$$\phi_X = \phi_Y.$$

Propriété : Soit $X = (X_1, ..., X_n)^t$ un vecteur aléatoire de dimension n. Les coordonnées de X sont indépendantes si et seulement si la fonction caractéristique de X est le produit des fonctions caractéristiques de ses coordonnées i.e

$$\forall t = (t_1, ..., t_n)^t \in \mathbb{R}^n,\ \phi_X(t) = \prod_{k=1}^{n} \phi_{X_k}(t_k).$$

[2]

2.2 Vecteurs Gaussiens

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

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

2.2.1 Exemple fondamental

Considérons n variables aléatoires $X_1, ..., X_n$ indépendantes et de loi respectivement $N(m_1, {\sigma_1}^2), ..., N(m_n, {\sigma_n}^2)$.
Pour $i = 1, ..., n$, la variable aléatoire X_i est donc de densité

$$f_{X_i}(x) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left(-\frac{1}{2}\left(\frac{x - m_i}{\sigma_i}\right)^2\right). \tag{2.9}$$

En raison de l'indépendance des $v.a.X_i$, la densité conjointe du vecteur $X_1, ..., X_n$ est :

$$f_{X_i,...,X_n}(x_1, ..., x_n) = \frac{1}{\left(\sqrt{2\pi}\right)^n} \frac{1}{\prod\limits_{i=1}^{n} \sigma_i} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x - m_i}{\sigma_i}\right)^2\right). \tag{2.10}$$

Le vecteur espérance du vecteur $X = (X_1, ..., X_d)$ et sa matrice de covariance sont :

$$E(X) = m = \begin{pmatrix} m_1 \\ . \\ . \\ . \\ m_n \end{pmatrix} \qquad \text{et} \qquad \Gamma_X = \begin{pmatrix} \sigma_1^2 & & & 0 \\ & . & & \\ & & . & \\ & & & . \\ 0 & & & \sigma_n^2 \end{pmatrix}.$$

Notons que la matrice Γ_X est daigonnale en raison de l'indépendance des $v.a.r.(X_i)_{i=1,...,n}$.Comme toutes les variances σ_i sont strictement positives, on obtient aisément la matrice inverse

$$\Gamma_X^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2} & & & 0 \\ & . & & \\ & & . & \\ & & & . \\ 0 & & & \frac{1}{\sigma_n^2} \end{pmatrix}.$$

On peut alors récrire la densité conjointe du vecteur $X = (X_1, ..., X_n)$ sous la forme

$$f_{X_i,...,X_n}(x_1, ..., x_n) = \frac{1}{\left(\sqrt{2\pi}\right)^n} \times \frac{1}{\sqrt{\det(\Gamma_X)}} \exp\left(-\frac{1}{2}(x - m)^t \Gamma_X^{-1}(x - m)\right),$$

puisque

$$(x-m)^t\Gamma_X^{-1}(x-m) = (x_1-m_1,...,x_n-m_n)\begin{pmatrix} \frac{1}{\sigma_1^2} & & & 0 \\ & . & & \\ & & . & \\ & & & . \\ 0 & & & \frac{1}{\sigma_n^2} \end{pmatrix}\begin{pmatrix} x_1-m_1 \\ . \\ . \\ . \\ x_n-m_n \end{pmatrix}$$

$$= \sum_{i=1}^{n} \frac{(x-m_i)^2}{\sigma_i^2}.$$

Intéressons-nous maintenant à la fonction caractéristique du vecteur X.

Toujours en raison de l'indépendance, on a, pour tout $t=(t_1,...,t_n)$ de $\mathbb{R}^n$:

$$\phi_X(t) = \phi_{X_1,...,X_n}(t_1,...,t_n) = \prod_{j=1}^{n}\phi_{X_j}(t_j).$$

Or, que la fonction caractéristique d'une $v.a.r.$de loi $N(m_j, {\sigma_j}^2)$ est :

$$\phi_{X_j}(t_j) = \exp\left(it_jm_j - \frac{1}{2}t_j^2\sigma_j^2\right)$$

d'où on tire :

$$\begin{aligned}\phi_{X_1,...,X_n}(t) &= \exp\left(i\sum_{j=1}^{n}t_jm_j - \frac{1}{2}\sum_{j=1}^{n}t_j^2\sigma_j^2\right) \\ &= \exp\left(it^tm - \frac{1}{2}t^t\Gamma_X t\right).\end{aligned}$$

Remarquons enfin que toute combinaison linéaire des X_j, pour $j = 1, ..., n$, est de loi normale dans $\mathbb{R}$. Une combinaison linéaire des X_j s'écrit en effet de manière générale sous la forme :

$$\begin{aligned}\phi_{\langle t,X\rangle}(u) &= E\left(e^{iu\langle t,X\rangle}\right) = E\left(e^{i\langle ut,X\rangle}\right)\\ &= \phi_X(ut) = \phi_X(ut_1, ..., ut_n)\\ &= \exp\left(iut^t m - \frac{1}{2}u^2 t^t \Gamma_X t\right).\end{aligned}$$

La fonction caractéristique de la $v.a.r.$ $\langle t, X\rangle$est donc de la forme :

$$\phi_{\langle t,X\rangle}(u) = \exp\left(iua - \frac{1}{2}u^2 b\right),$$

avec $a = t^t m$ et $b = t^t\Gamma_X t$. Par caractérisation, la $v.a.r.$ $\langle t, X\rangle$est donc de loi $N(t^t m, t^t\Gamma_X t)$.[4]

2.2.2 Caractéristique des vecteus gaussiens

Définition 2.7 *Un vecteur aléatoire $X = (X_1, ..., X_n)$ de $\mathbb{R}^n$ est dit vecteur gaussien si, pour tout $t = (t_1, ..., t_n)$ de $\mathbb{R}^n$, la v.a.r.*

$$t^t X = \sum_{i=1}^{n} t_i X_i, \tag{2.11}$$

est une v.a.r. de loi normale. Autrement dit, si toute combinaison linéaire des composantes de $X = (X_1, ..., X_n)$ est de loi normale.

Si son vecteur des espérances est m et sa matrice de covariance est Γ_X, on note

$X \rightsquigarrow N_n(m, \Gamma_X)$.

Remarquons que l'on peut en particulier en déduire que toutes les composantes du vecteur X sont des v.a.r. de loi normale. En revanche, la réciproque est fausse. Un vecteur dont toutes les composantes sont de loi normale, n'est pas nécessairement un vecteur gaussien.

La définition précédente implique également que tout sous vecteur d'un vecteur gaussien est encore un vecteur gaussien.

Propriété : Soit $X_1, ..., X_n$ des variables aléatoires réelles indépendantes de loi normale centrée réduite. Alors, le vecteur $X = (X_1, ..., X_n)^t$ est un vecteur aléatoire gaussien.[2]

Proposition 2.4 *Soit X est vecteur gaussien à valeurs dans $\mathbb{R}^n$, de moyenne $m = E(X) = (E(X_1), ..., E(X_n))^t$ et de matrice de variance-covariance*

$$\Gamma = (Cov(X_i, X_j))_{i,j\in\{1,...,n\}} = E((X - E(X))(X - E(X))^t). \tag{2.12}$$

Soient $A \in M_{d,n}$ et $b \in \mathbb{R}^d$. Alors, $Y = AX + b$ est un vecteur gaussien à valeurs dans $\mathbb{R}^d$ de moyenne $Am + b$ et de matrice de variance-covariance $A\Gamma A^t$.

Proposition 2.5 *Pour qu'un vecteur X de $\mathbb{R}^n$ soit un vecteur gaussien il faut et il suffit qu'il existe un vecteur m de $\mathbb{R}^n$et une matrice Γ symétrique et positif de dimension $n \times n$ tels que, pour tout vecteur t de $\mathbb{R}^n$, on ait :*

$$\phi_X(t_1, ..., t_n) = \exp\left(it^t m - \frac{1}{2}t^t\Gamma t\right). \tag{2.13}$$

Dans ce cas, on a : $E(X) = m$ et $\Gamma_X = \Gamma$.

Preuve. Supposons que X soit un vecteur gaussien. Toute $v.a.r.$de la forme t^tX, pour t dans $\mathbb{R}^n$, est donc de loi $N(t^tm, t^t\Gamma t)$.Ainsi sa fonction caractéristique est :

$$\phi_{t^tX}(u) = E(e^{iut^tX}) = \exp\left(iut^tm - \frac{1}{2}u^2t^t\Gamma t\right).$$

En posant$u = 1$ dans cette équation, on obtient :

$$E(e^{it^tX}) = \exp\left(it^tm - \frac{1}{2}t^t\Gamma t\right).$$

Ce qui est bien l'expression annoncée pour la fonction caractéristique.

Réciproquement, soit X un vecteur aléatoire dans $\mathbb{R}^n$ ayant pour fonction caractéristique.

$$\phi_X(t) = \exp\left(it^t m - \frac{1}{2}t^t\Gamma t\right) = E(e^{i\langle t,X\rangle}),$$

pour tout t dans $\mathbb{R}^n$. Notons maintenant $Y = \langle t, X\rangle$ la variable aléatoire réelle dont la fonction caractéristique est, pour tout u dans $\mathbb{R}$: ■

$$\begin{aligned}\phi_Y(u) &= E(e^{iuY}) = E(e^{iut^tX}) = E(e^{i\langle ut,X\rangle})\\ &= \exp\left(iut^t m - \frac{1}{2}u^2 t^t\Gamma_X t\right)\\ &= \exp\left(iua - \frac{1}{2}u^2 b\right),\end{aligned}$$

où $a = t^t m$ et $b = t^t\Gamma_X t$. Par caractérisation, la $v.a.r.Y$ est donc de loi $N(t^t m, t^t\Gamma t)$. On a donc démontré que toute combinaison linéaire des composantes du vecteur X est de loi normale, et par définition il s'agit bien d'un vecteur gaussien.[4]

Remarque 2.2 *On constate que la loi d'un vecteur gaussien est entièrement déterminer par* Γ, *et on note cette loi*

$$X \rightsquigarrow N_n(m, \Gamma).$$

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

Remarque 2.3 *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.

2.2.3 Quelques propriétés des vecteurs gaussiens

En général, deux variables aléatoires réelles non corrélées ne sont pas nécessairement indépendantes. Pour des vecteurs gaussiens, indépendance des coordonnées et absence de corrélation sont deux notions équivalentes.

Propriétés : Soit X un vecteur gaussien dans $\mathbb{R}^n$: Pour que ses composantes $X_1, ..., X_n$ soient indépendantes, il faut et il suffit que la matrice de covariance soit diagonale.

Preuve. Il suffit, bien sûr, de montrer la réciproque. Supposons donc que Γ_X soit diagonale, i.e

$$\Gamma_X = \begin{pmatrix} \sigma_1^2 & 0 & . & . & 0 \\ 0 & . & . & & . \\ . & . & . & . & . \\ . & & . & . & 0 \\ 0 & . & . & 0 & \sigma_n^2 \end{pmatrix}.$$

Comme X est un vecteur gaussien de loi $N(m, \Gamma_X)$, chacune de ses composantes X_j, pour $j = 1, ..., n$, est de loi normale $N(m_j, \sigma_j^2)$ et de fonction caractéristique :

$$\phi_X(t) = \exp\left(it_j m_j - \frac{1}{2}\sigma_j t_j^2\right),$$

pour tout t_j dans $\mathbb{R}$.

Par ailleurs, la fonction caractéristique du vecteur X est, pour tout t dans $\mathbb{R}^n$:

$$\begin{aligned}
\phi_X(t) &= \exp\left(it^t m - \frac{1}{2}t^t\Gamma_X t\right) \\
&= \exp\left(i\sum_{j=1}^{n} t_j m_j - \frac{1}{2}\sum_{j=1}^{n}\sigma_j t_j^2\right) \\
&= \exp\left(\sum_{j=1}^{n}\left(it_j m_j - \frac{1}{2}\sigma_j t_j^2\right)\right) \\
&= \prod_{j=1}^{n}\exp\left(it_j m_j - \frac{1}{2}\sigma_j t_j^2\right) \\
&= \prod_{j=1}^{n}\phi_{X_j}(t_j).
\end{aligned}$$

■

Propriété : Soit $X = (X_1, ..., X_n)^t$ un vecteur aléatoire gaussien d'espérance $m \in \mathbb{R}^n$ et de matrice de covariance Γ_X. La loi de X admet une densité si et seulemet sila matrice Γ_X est inversible et définie positive (i.e. de déterminant $det(\Gamma_X) > 0$ non nul) et dans ce cas la densité est :

$$\forall x \in \mathbb{R}^n, \quad f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}}}\frac{1}{\sqrt{\det(\Gamma_X)}}\exp\left(-\frac{1}{2}(x-m)^t\Gamma^{-1}(x-m)\right). \tag{2.14}$$

Preuve. Nous démontrons uniquement la condition suffsante.

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

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

et sa matrice de dispersion est

$$\Gamma_Y = B^{-1}E(B^{-1})^t = Id,$$

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

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

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

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

d'où le résultat

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

■

Remarque 2.4 *analogie avec le cas uni-dimensionel* $(n = 1)$*. Si* $X \in \mathbb{R}$ *suit loi normal* $N(m, \sigma^2)$*, alors sa densité est :*

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}, x \in \mathbb{R}. \tag{2.15}$$

Propriétés :

La transformée d'un vecteur gaussien de $\mathbb{R}^n$ par une application linéaire de $\mathbb{R}^n$ vers $\mathbb{R}^p$ est encore un vecteur gaussien.

Preuve :

Soit X un vecteur gaussien de $\mathbb{R}^n$, de vecteur des espérances m et de matrice de covariance Γ_X. Soit A la matrice associée à une transformation linéaire quelconque de $\mathbb{R}^n$ vers $\mathbb{R}^p$. La matrice A est donc de dimension $p \times n$: Calculons la fonction caractéristique du vecteur aléatoire $Y = AX$, pour tout t de $\mathbb{R}^p$, on a :

$$\begin{aligned}\phi_Y(t) = \phi_{AX}(t) = E\left(e^{i\langle t, AX\rangle}\right) = E\left(e^{i\left\langle A^t t, X\right\rangle}\right) \\ = \phi_X(A^t t) = \exp\left(it^t Am - \frac{1}{2}t^t A\Gamma_X A^t t\right).\end{aligned}$$

Par caractérisation, le vecteur Y est donc un vecteur gaussien dans $\mathbb{R}^p$ de vecteur des espérances Am et de matrice de covariance $A\Gamma_X A^t$, i.e

$$Y \rightsquigarrow N_p(Am, A\Gamma_X A^t).$$

2.2.4 Le théorème central limite vectoriel

Théorème 2.1 (Le théorème central limite vectoriel) *Soit $(X_n)_{n\geq 1}$ une suite de vecteurs aléatoires i.i.d. de $\mathbb{R}^d$ appartenant à L^2, d'espérance $m \in \mathbb{R}^d$ et de matrice de covariance $\Gamma_X \in M_d(\mathbb{R})$. Notons pour tout $n \geq 1$*

$$S_n = X_1 + ... + X_n.$$

On a la convergence en loi

$$loi\left(\frac{S_n - nm}{\sqrt{n}}\right) \longrightarrow_{n\longrightarrow\infty} N(0, \Gamma_X).$$

Remarque 2.5 *On pourra écrire ainsi une sorte de développement limité des moyennes empiriques de vecteurs i.i.d. sous la forme*

$$\frac{S_n}{n} \backsimeq m + \frac{1}{\sqrt{n}} G$$

où G est vecteur gaussien de loi $N(0, \Gamma)$.[1]

2.2.5 Espérance conditionnelle et projection orthogonale

On va maintenant revenir sur la notion d'espérance conditionnelle et montrer qu'elle se caractérise trés facilement dans le cas gaussien. Soit (Ω, T, P) un espace probabilisé, et $X : (\Omega, T) \to (\mathbb{R}, B_{\mathbb{R}})$ une $v.a.$ réelle. On rappelle que l'on note par $L^2(\Omega, T, P)$ l'espace des $v.a.r.X$ de carré intégrable $i.e$ telles que $E(X^2) < +\infty$. De plus, l'espace $L^2(\Omega, T, P)$ est un espace de Hilbert lorsqu'il est muni du produit scalaire $\langle X, Y \rangle = E(XY)$, et de la norme quadratique associée $\|X\|_2 = \sqrt{E(X^2)}$.

Une remarque trés importante est alors la suivante. Si $\binom{X}{Y} \in \mathbb{R}^2$ est un vecteur gaussien d'espérance nulle, alors l'orthogonalité de X et Y dans l'espace $L^2(\Omega, T, P)$ i.e.

$$\langle X, Y \rangle = E(XY) = 0,$$

et **l'indépendance de** X **et** Y sont des propriétés équivalentes car

$$Cov(X, Y) = E(XY) - E(X)E(Y) = E(XY) = \langle X, Y \rangle . \tag{2.16}$$

Utilisons donc de nouveau le cadre de la géométrie hilbertienne pour caractériser l'espérance conditionnelle. Soit $Z = (Z_1, ..., Z_n)^t \in \mathbb{R}^n$ un vecteur aléatoire gaussien d'espérance nulle, i.e.$E(Z_i) = 0$ pour tout $i = 1, ..., n$. Introduisons la définition suivante :

Définition 2.8 *Soient $\{i_1, ..., i_p\} \subset \{1, ..., n\}$ avec $p \leq n$. Le sous-espace vectoriel (fermé) engendré par les $v.a.Z_{i_1}, ..., Z_{i_p}$ est*

$$Vect(Z_{i_1}, ..., Z_{i_p}) = \overline{\left\{ \sum_{k=1}^{p} \alpha_k Z_{i_k}, \qquad \alpha_k \in \mathbb{R} \right\}}. \tag{2.17}$$

On a alors les propriétés suivantes :

Propriété : Soit $Y \in L^2(\Omega, T, P)$une $v.a.r.$telle que $(Y, Z_{i_1}, ..., Z_{i_p}) \in \mathbb{R}^{p+1}$ est un vecteur aléatoire gaussien d'espérance nulle, l'espérance conditionnelle de Y sachant le vecteur $\widetilde{Z} = (Z_{i_1}, ..., Z_{i_p})$, notée $E(Y \backslash Z_{i_1}, ..., Z_{i_p})$ ou $E(Y \backslash \widetilde{Z})$, est la projection orthogonale de Y sur l'espace

$$Vect(Z_{i_1}, ..., Z_{i_p}) = \overline{\left\{ \sum_{k=1}^{p} \alpha_k Z_{i_k}, \qquad \alpha_k \in \mathbb{R} \right\}}, \tag{2.18}$$

et dans le cas $p = 1$

$$E(Y \backslash Z_{i_1}) = \frac{Cov(Z_{i_1}, Y)}{Var(Z_{i_1})} Z_{i_1}.$$

On peut alors remarquer que $E(Y \backslash Z_{i_1}, ..., Z_{i_p})$ est donc une $v.a.$ gaussienne ce qui montre la **stabilité de la loi gaussienne par conditionnement**. De plus on a également la décomposition suivante lors du conditionnement par rapport à deux vecteurs gaussions indépendants.

Propriété : Soit $Z \in \mathbb{R}^n$ et $W \in \mathbb{R}^m$ deux vecteurs aléatoires gaussien indépendants et d'spérance nulle. Soit $Y \in L^2(\Omega, T, P)$une $v.a.r.$ telle que $(Y, Z'W') \in \mathbb{R}^{n+m+1}$ est un vecteur aléatoire gaussien d'espérance nulle. Alors, l'espérance conditionnelle de Y sachant le vecteur $\binom{Z}{W}$, notée $E(Y \backslash Z, W)$ est la projection orthogonale de Y sur l'espace

$$Vect(Z, W) = \overline{\left\{ \sum_{k=1}^{n} \alpha_k Z_k + \sum_{j=1}^{m} \beta_j W_j, \quad \alpha_k, \beta_j \in \mathbb{R} \right\}} = Vect(Z) \oplus Vect(W), \tag{2.19}$$

et donc

$$E(Y \backslash Z, W) = E(Y \backslash Z) + E(Y \backslash W).$$

2.2.6 Lois conditionnelles et prédiction

Etudions maintenant les lois conditionnelles de certaines coordonnées d'un vecteur gaussien sachant la valeur des autres coordonnées. Soit Z un vecteur aléatoire gaussien de dimension n divisé en deux blocs de coordonnées X et Y de dimensions $n_X \geq 1$ et $n_Y \geq 1$ telles que $n = n_X + n_Y$, i.e.$Z = \binom{X}{Y}$. L'espérance de Z s'écrit alors sous la forme

$$E(Z) = \begin{pmatrix} E(X) \\ E(Y) \end{pmatrix},$$

et la matrice de covariance Γ_Z de Z peut se décomposer de la façon suivante

$$\Gamma_Z = \begin{pmatrix} \Gamma_X & \Gamma_0 \\ \Gamma_0 & \Gamma_Y \end{pmatrix},$$

où

$$\Gamma_0 = E\left((X - E(X))(Y - E(Y))^t\right). \tag{2.20}$$

La notion de densité conditionnelle vue précédemment pour les couples de variables aléatoires réelles peut être étendue au cas des vecteurs aléatoires. La densité conditionnelle de Y sachant $X = x$ est alors la fonction

$$f_Y^{X=x}(y) = \frac{f_Z(x,y)}{f_X(x)}, \tag{2.21}$$

et on a le résultat suivant.

Propriété : Supposons que $Z = \binom{X}{Y}$soit un vecteur gaussien, et que la matrice de covariance Γ_X de X est définie positive. Soit $x \in \mathbb{R}^{n_X}$ La loi contionnelle de Y sachant $X = x$ est alors une loi gausienne d'espérance

$$E(Y \backslash X = x) = E(Y) + \Gamma_0^t \Gamma_X^{-1}(x - E(X)),$$

et de matrice de covariance

$$\Gamma_Y^{X=x} = \Gamma_Y - \Gamma_0^t \Gamma_X^{-1} \Gamma_0,$$

et donc

$$\forall y \in \mathbb{R}^{n_Y}, f_Y^{X=x}(y) = \frac{1}{(2\pi)^{\frac{n_Y}{2}}} \frac{1}{\sqrt{(\det(\Gamma_Y^{X=x}))}} \exp\left(-\frac{1}{2}(y - m_Y^{X=x})^t \left[\Gamma_Y^{X=x}\right]^{-1}(y - m_Y^{X=x})\right). \tag{2.22}$$

avec $m_Y^{X=x} = E(Y) + \Gamma_0^t \Gamma_X^{-1}(x - E(X))$.[2]

Bibliographie

[1] Benoit, Mselati et Benaych-Georges, Introduction aux Probabilités et aux Statistiques, 3 octobre 2004.

[2] Bigot, Jérémie. Notes de cours de Probabilités, 18 mars 2014.

[3] Chesneau,Christophe. Probabiliés et variables aléatoies discrètes.

[4] Dauxois.Jean-Yves Cours de Probabilités. Septembre 2011.

[5] D. Foata-A. Fuchs. Calcul des probabilités. Dunod 2003.

[6] Lecoutre, Jean-Pierre. Statistique et probabilités Cours et exercices corrigés. Dunod, 2016.

[7] Perryt, A. Cours de probabilités et statistiques, 2010.

[8] Popier Alexandre,Wintenberger Olivier. Introduction au calcul des probabilités, 2006, 2007.

[9] Saporta, Gilbert. Probabilités, analyse des données et statistique. Editions Technip, 2006

[10] Saussereau, Bruno. Cours de théorie des probabilités avec exercices corrigés et devoirs, 2013, 2014.

[11] Smolarz, André. Modélisation probabiliste pour l'ingénieur. 2009.

[12] Suquet, Charles. Introduction au Calcul des Probabilités, 2002, 2003.

[13] Ycart,Bernard, Modélées probabilistes simulation de variables aléatoies.

.

Annexe A : Abréviations et Notations

Les différentes abréviations et notations utilisées tout au long de ce mémoire sont expliquées ci-dessous.

Ω	:	événement certain.
ϕ	:	ensemble vide.
$\overline{A}$	:	événement contraire à A.
$P(A)$	:	probabilité de A.
T	:	tribu des événement.
(Ω, T)	:	espace probabilisable.
(Ω, T, P)	:	espace probabilisé.
P_X	:	loi de probabilité de X.
$v.a$	:	variable aléatoire.
$F(X)$	:	fonction de répartition.de X
$E(X)$	:	espérance mathématique du X.
$V(X)$	:	la variance de X.
σ	:	l'écart-type de X.
$Cov(X,Y)$	:	covariance de X et Y.
ρ_X	:	cofficient de corrélation.
$\Gamma(X,Y)$	:	matrice de covariance.
ϕ_X	:	la fonction caractéristique de X.
$G_X(s)$	:	fonction génératrice.
P_{ij}	:	loi joint du couple (X,Y).
P_i	:	loi marginale de X.

P_{ij}	:	loi de probabilité conjointe.
P_i	:	loi marginale.
$Cov(X, Y)$	:	la covariance entre deux variable aléatoires X et Y.
ssi	:	si et seulement si.
$f_{X,Y}$	:	densité du couple (X, Y).
$N(m, \sigma)$	:	la loi normale de paramètre m et σ.
$N(0, 1)$	:	la loi normale centrée réduite.
$f.r$	:	fonction de répartition.
P_{X_i}	:	la iéme loi marginale probabilité sur $\mathbb{R}$.
f_{X_i}	:	la iéme densité marginale.
Id	:	matrice identité.
iid	:	indépandante indentiquement distribué.