

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 pour obtenir le diplôme de

Master en “**Mathématiques Appliquées**”

Option : Statistique

Par Melle. SAAD Maroua

**Titre :**

Introduction aux variables aléatoires vectorielles

Devant le Jury :

Mme.	ABDELLI JIHANE	MCA	U. Biskra	Présidente
Mr.	CHERFAOUI MOULOUD	Prof.	U. Biskra	Rapporteur
Mme.	SOLTANE LUIZA	MCB	U. Biskra	Examinatrice

**Soutenu Publiquement le 28/06/2022**

## Dédicace

A chaque fois qu'on achève une étape importante dans notre vie, on fait une pose pour regarder en arrière et se rappeler toutes ces personnes qui ont partagés avec nous tous les bons moments de notre existence, mais surtout les mauvais. Ces personnes qui nous ont aidés sans le leur dire, soutenus sans réserve, aimé sans compter, ces personnes à qui notre bonheur devient directement le leur, à qui un malheur en nous, en eux se transforme en pleur.

**Je dédie ce modeste travail à tous ceux que j'aime mais surtout :**

La plus proche de mon coeur celle qui fait l'impossible pour me donner le bonheur ma chère mère.

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

A mes chers frères et soeurs : *Mahmoud, Mohamed, Khalifa, Zineb et Darine.*

A tous les membres de ma grande famille.

Pour les résidents dans mon coeur, mes amies et mes compagnons avec lesquels j'ai partagé des moments les plus agréables : *Riguet Lamia, Driham sofia, Sabegh Hind, Tourki Wafa, Dandogha Ahlam, Fatima ....*

A tous mes amies de promotion 2022 de mathématiques.

A toutes les personnes qui ont participé et soutenu à la réalisation de ce travail de près ou de loin.

# Remerciements

*Après avoir rendu grâce à Dieu le Tout puissant et le Miséricordieux.*

*Je tiens tout d'abord à exprimer mes profondes gratitude et remerciements à mon encadreur Prof. Cherfaoui Mouloud pour son aide et son soutien, surtout pour sa patience qu'il a faite preuve avec moi et ses précieux conseils.*

*Mes remerciements vont aussi à tous les membres de Jury Dr. Abdelli Jihane et Dr. Soltane Luiza, qui ont accepté de lire et d'évaluer ce travail.*

*Je remercie ma famille, à qui je n'ai jamais su dire tout l'affection que j'ai pour eux, mon père, ma mère, mes frères et ma soeur qui ont été et seront toujours présents à mes côtés. Merci du fond du coeur pour votre soutien infailible, vos encouragements sans cesse.*

*Un grand merci particulier à mes collègues et mes amies pour les sympathiques moments qu'on a partagés au fil des années, je les remercie pour leur confiance, et leur soutien moral.*

# Rèsumè du mèmòire

**Rèsumè :** Dans ce mèmòire, nous avons abordé la notion de variable aléatoire vectorielle au sens probabiliste et statistique. Nous nous sommes particulièrement intéressés à la présentation de ce type de variable, au calcul de sa moyenne et de sa variance-covariance ainsi qu'à la mesure de la dépendance entre ses composantes (coefficient de corrélation). Enfin, à titre d'illustration des différentes notions abordées, un exemple sur des données réelles (notes des étudiants de Master 2 Mathématiques, option Statistique, promotion 2021/2022) a été présenté.

**Mots-clés :** Variables aléatoires univariée, Variables aléatoires vectorielles, caractéristiques statistiques, corrélation.

**Abstract :** In this work, we approached the notion of vector random variable in the probabilistic and statistical sense. We were particularly interested in the presentation of this type of variable, the calculation of its mean and its variance-covariance as well as the measurement of the dependence between its components (correlation coefficient). Finally, as an illustration of the different concepts discussed, an example on real data (grades of Master 2 Mathematics students, Statistics option, Class of 2021/2022) was presented.

**Keywords :** Univariate random variables, vector random variables, statistical characteristics, correlation.

# Notations et symbols

$v.a$	variable aléatoire
$f(x)$	fonction de densité
$F(x)$	fonction répartition
$\mu = E(X)$	Espérance de $X$
$\sigma^2 = Var(X)$	variance de $X$
$\sigma$	l'écart-type
$i.i.d$	indépendantes et identiquement distribuées
$\Gamma$	fonction de Gamma
$\Phi_X(t)$	fonction caractéristique de $X$
$\xrightarrow{\mathcal{L}}$	converge en loi
$\emptyset$	l'ensemble vide
$T.C.L$	théorème central limite
$N(0, 1)$	loi normale centré réduite
$X'$	transposée de $X$
$\Sigma$	matrice de variance covariance
$\rho$	coefficient de corrélation
$N_d(\mu, \Sigma)$	loi normale à $d$ -dimonsion
$l(X; \mu, \Sigma)$	la vraisemblance
$\hat{\mu}$	estimateur de la moyenne
$\hat{\Sigma}$	estimateur de la variance

# Table des matières

Dédicace	i
Remerciements	ii
Résumé du mémoire	iii
Notations et symbols	iv
Table des matières	v
Table des figures	viii
Liste des tableaux	ix
Introduction	1
<b>1 Caractérisation d'une variable aléatoire univariée</b>	<b>3</b>
1.1 Le concept de variable aléatoire . . . . .	3
1.2 La distribution d'un variable aléatoire . . . . .	5
1.2.1 Cas d'une variable aléatoire discrète . . . . .	5
1.2.2 Cas d'une variable aléatoire continue . . . . .	5

1.3	La fonction de répartition (distribution cumulative) . . . . .	6
1.4	Espérance (moyenne) d'une variable aléatoire . . . . .	7
1.5	La variance et l'écart-type d'une <i>v.a</i> . . . . .	9
1.6	Quelques lois de probabilités usuelles . . . . .	10
1.6.1	La distribution de Gauss (Normale) . . . . .	11
1.6.2	La distribution $\chi^2$ ( <i>Khi – Deux</i> ) . . . . .	14
1.6.3	La distribution Student ( <i>t</i> ) . . . . .	15
1.6.4	La distribution Fisher (Fisher-Snedecor) <i>F</i> . . . . .	16
	Conclusion . . . . .	17
<b>2</b>	<b>Caractérisation d'une variable aléatoire vectorielle</b>	<b>18</b>
	Introduction . . . . .	18
2.1	Fonction de répartition et densité d'une <i>v.a</i> vectorielle . . . . .	19
2.1.1	Fonction de répartition : . . . . .	19
2.1.2	Fonction de densité : . . . . .	21
2.1.3	Changement de variables dans une densité . . . . .	21
2.2	Fonction caractéristique et Théorème de Cramer-Wold . . . . .	24
2.2.1	Fonction caractéristique . . . . .	24
2.2.2	Théorème de Cramer-Wold . . . . .	25
2.3	Espérance, variances-covariances et corrélations . . . . .	26
2.4	Transformations linéaires . . . . .	29
2.5	Vecteurs aléatoires Gaussiens . . . . .	30
2.5.1	La loi multinormale . . . . .	31
2.5.2	Estimation des paramètres $\mu$ et $\Sigma$ . . . . .	33

2.5.3	Lois dérivées de la loi normale multivariée . . . . .	35
	Conclusion . . . . .	41
<b>3</b>	<b>Exemple numérique illustratif d'analyse d'une <i>v.a.</i> vectorielle</b>	<b>42</b>
	Introduction . . . . .	42
3.1	Présentation des données . . . . .	42
3.2	Estimation de la moyenne . . . . .	44
3.3	Estimation de la matrice Variance-Covariance . . . . .	46
3.4	Estimation de la matrice de corrélation . . . . .	48
	Conclusion . . . . .	50
	<b>Conclusion générale</b>	<b>53</b>
	<b>Bibliographie</b>	<b>54</b>
	<b>Annexe A : Rappel sur quelques notions de l'algèbre linéaire</b>	<b>56</b>
	<b>Annexe B : Les données de l'application numérique</b>	<b>61</b>

# Table des figures

1.1	Classification des variables aléatoires. . . . .	4
1.2	Fonction répartition : (A) d'une variable discrète, (B) d'une variable continue et (C) d'une variable diffuse et d'une variable avec atome, mais non discrète. . . . .	6
1.3	Distribution de Gauss centrée et réduite . . . . .	12
1.4	Distribution de $\chi_2$ pour différentes degré de liberté. . . . .	14
1.5	Distribution de Student. . . . .	15
1.6	Distribution de Fisher. . . . .	16
2.1	Monotonie illustré par le biais d'ensembles. . . . .	21
3.1	Code source Matlab pour le calcul du $\bar{X}$ . . . . .	45
3.2	Code source Matlab pour le calcul de la Matrice de $\hat{\Sigma}_X$ . . . . .	47
3.3	Code source Matlab pour le calcul de la Matrice de $\hat{\rho}_X$ . . . . .	51

# Liste des tableaux

3.1	Présentation des données . . . . .	43
3.2	Estimateur sans biais de la Matrice variance-covariance. . . . .	47
3.3	Estimateur de la matrice de coefficients de corrélation. . . . .	50

# Introduction

Dans des domaines très différents comme les domaines scientifiques, sociologique ou médical, on s'intéresse à de nombreux phénomènes dans lesquels apparaît l'effet du hasard. Ces phénomènes sont caractérisés par le fait que les résultats des observations varient d'une expérience à l'autre. Une expérience est appelée "aléatoire" s'il est impossible de prévoir à l'avance son résultat et si, répétée dans des conditions identiques, elle peut donner des résultats différents.

L'analyse des résultats d'une expérience aléatoire ne peut se faire que via l'outil statistique. En effet, contrairement à la démarche mathématique qui est généralement déductive (on part d'une propriété générale pour démontrer une propriété particulière), celle des statistiques, souvent inductive où on extrapole les paramètres d'un échantillon à une population.

Dans la littérature statistiques on trouve un nombre important de techniques qui sont soit générale ou spécifique à des phénomènes plus précis. Enfin ces techniques peuvent être distinguées par divers facteurs. Parmi les distinctions possibles est la dimension de la variable considéré dans l'analyse. Ainsi, la première d'entre elles sépare les techniques univariées des multivariées où une technique univariée s'attache à une seule série d'un caractère donné ou à une seule mesure (même s'il y a plusieurs échantillons). Tandis qu'une technique multivariée analyse les

éventuelles relations existant entre plusieurs caractères.

Dans ce sens, l'objectif du présent mémoire est de se familiariser et d'expliquer la notion de variables aléatoires multivariées et sa caractérisation. Plus précisément on s'intéresse au cas de la variable aléatoire vectorielle. Pour répondre à notre objectif, en plus de la présente introduction, nous avons organisé notre mémoire comme suit :

- ✓ Dans le Chapitre 1, nous avons exposé quelques notions et définitions de base associées à une variable aléatoire univariée.
- ✓ Dans le Chapitre 2, la notion de variables aléatoires vectorielles et sa caractérisation a été abordée.
- ✓ Dans le Chapitre 3, une application numérique illustrative de la caractérisation d'une variable vectorielle a été exposée.

Le mémoire contient également une conclusion générale, une liste Bibliographique et deux Annexes. Les deux Annexes A et B contiennent respectivement, des notions de base sur les matrices et les données considérées dans l'application numérique.

# Chapitre 1

## Caractérisation d'une variable aléatoire univariée

### Introduction

Le calcul des probabilités s'occupe d'épreuves aléatoires et de phénomènes aléatoires c'est-à-dire d'expériences ou de phénomènes naturels qui, dans des conditions déterminées et stables, ne mènent pas toujours à la même issue. On observe, cependant, une certaine régularité statistique. L'étude de cette régularité fait l'objet d'une théorie mathématique. Dans ce chapitre nous nous limitons à la présentation de quelques notions sur le phénomène aléatoire unidimensionnel et sa caractérisation.

### 1.1 Le concept de variable aléatoire

Une variable aléatoire (*v.a*)  $X$  est une fonction qui associée à chaque résultat d'une expérience aléatoire un nombre réel. L'ensemble des réalisations (ou de résultats)

de  $X$  sera désigné par :

$$X = \{X_i : X_i = X(\omega); \omega \in \Omega\}$$

**Exemple 1.1.1** *La taille d'un individu extrait au hasard d'une population ou, encore le nombre de "face", dans une série de 10 jets d'une monnaie, et la valeur d'un Dé entre 1 et 6, la côte de la pièce dans un pile ou face représentent des variables aléatoires.*

Dans ce qui suit nous allons remarquer que l'outil mathématique utilisé pour décrire une caractéristique d'une variable aléatoire dépend de la nature de cette dernière. Par conséquent, avant toute analyse l'identification préalable du type de la variable aléatoire est nécessaire. La classification d'une variable aléatoire selon sa nature peut être schématisée comme suite :

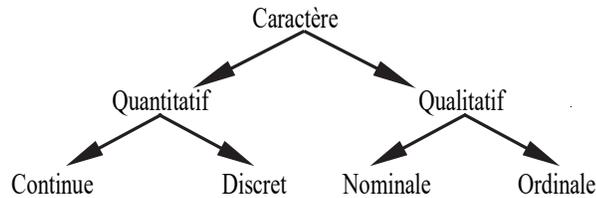


FIGURE 1.1 – Classification des variables aléatoires.

**Exemple 1.1.2** *Le résultat du jet d'un Dé, le nombre d'enfants dans une famille, sont des variables aléatoires quantitative discrètes. Tandis que la note moyenne des étudiants et la taille et le poids d'un individu, sont des variables aléatoires quantitatives continues.*

## 1.2 La distribution d'un variable aléatoire

Dans cette section nous allons intéresser à la définition et la présentation la distribution d'une variable aléatoire dans le cas discret et le cas continu.

### 1.2.1 Cas d'une variable aléatoire discrète

D'une manière générale, si une variable aléatoire discrète  $X$  peut prendre les valeurs  $\Omega = \{x_1, x_2, \dots, x_n\}$  avec des probabilités respectives  $p_1, p_2, \dots, p_n$ , nous dirons que  $X$  a pour distribution de probabilité l'ensemble des couples :

$$(x_1, p_1), (x_2, p_2), \dots, (x_n, p_n).$$

Il est à noter que la distribution de probabilité d'une variable aléatoire discrète vérifiée les trois propriétés suivantes :

- $p_i \geq 0$  pour tout  $i$  ;
- $\sum_{i=1}^n p_i = 1$  ;
- $P(a \leq X \leq b) = \sum_{i=1}^n p_i \mathbf{1}_{\{a \leq x_i \leq b\}}$ , avec  $\mathbf{1}_{\{\cdot\}}$  est la fonction indicatrice.

On peut représenter graphiquement une distribution de probabilité d'une variable aléatoire discrète à l'aide d'un diagramme en bâtons ou en colonnes.

### 1.2.2 Cas d'une variable aléatoire continue

Pour calculer les probabilités afférentes à une variable continue on utilise sa fonction de densité, c'est-à-dire une fonction qui permet de calculer la probabilité que  $X$  soit dans un intervalle  $[a; b]$ . Plus précisément, la densité de probabilité (ou densité) de  $X$  est une fonction  $f_x$  définie de  $\Omega$  dans  $\mathbb{R}_+$  telle que :

1.  $f_X(x) \geq 0$  pour tout  $x$  ;
2.  $\int_{\Omega} f_X(x) dx = 1$ ;
3.  $p(a \leq x \leq b) = \int_a^b f_X(x) dx$ ;

S'il ne y'a pas de possibilité de confusion, on peut se contenté de l'utilisation uniquement du symbole  $f$  à la place de  $f_X$ .

### 1.3 La fonction de répartition (distribution cumulative)

La fonction répartition ou fonction distribution cumulative d'une variable aléatoire réelle  $X$  est la fonction  $F_x$  qui, a tout réel  $X$ , associée la probabilité d'obtenir une valeur inférieure ou égale à  $x$ . La fonction cumulative est une fonction continue dans le cas de *v.a* continues et elle est définie en escalier dans le cas de *v.a* discrètes (voir l'exemple de la figure 1.2).

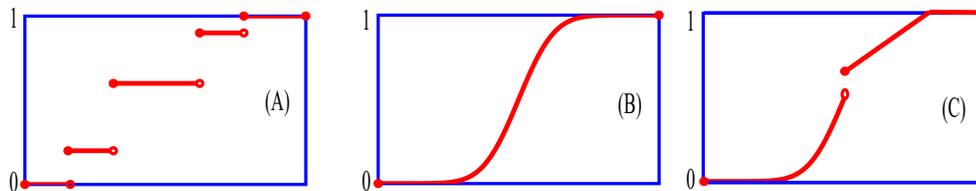


FIGURE 1.2 – Fonction répartition : (A) d'une variable discrète, (B) d'une variable continue et (C) d'une variable diffuse et d'une variable avec atome, mais non discrète.

Cette fonction est une caractéristique de la loi de probabilité de la *v.a*. Elle permet de calculer la probabilité de chaque intervalle semi-ouvert à gauche  $]a, b]$  où  $a \leq b$  par :

$$P(X \in ]a, b]) = P(a \leq X \leq b) = F_X(b) - F_X(a)$$

**Propriété 1.3.1 :**

- Elle est non décroissante.
- Elle prend des valeurs entre 0 et 1.
- Elle tend vers 0 lorsque  $x$  tend vers  $-\infty$  et vers 1 lorsque  $x$  tend vers  $+\infty$ .
- La fonction de répartition d'une variable continue est une fonction continue tandis que la fonction de distribution cumulative d'une variable discrète est une fonction discontinue en escalier.
- Entre la distribution de probabilité  $(x_1, p_1), (x_2, p_2), \dots, (x_n, p_n)$  et la fonction de répartition  $F$  d'une variable discrète il y a la relation suivante :

$$F(x) = \sum_{x_i \leq x} p_i.$$

- Entre la fonction de densité et la fonction répartition  $F$  d'une variable continue y a les deux relations suivantes :

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt, \\ f(x) &= \frac{dF(x)}{dx}, \quad \text{si } F \text{ est dérivable en } x. \end{aligned}$$

**Remarque 1.3.1** Pour une variable aléatoire continue  $P(X \leq x) = P(X < x)$ .

## 1.4 Espérance (moyenne) d'une variable aléatoire

L'espérance d'une variable aléatoire est, lorsqu'elle existe, la moyenne des valeurs de cette variable, pondérées par leurs probabilités de réalisations on voit bien comment traduire cette définition informelle dans le cas d'une *v.a* discrète  $X$  en posant :

$$\mu(X) = E(X) = \sum_{x \in \Omega} xP(X = x) \quad (1.1)$$

Essayons de traduire la définition informelle ci-dessus dans le cas d'une *v.a* à densité  $f$ . Partant de on remplace  $P(X = x)$  par  $P(X \in [x, x + dx])$  probabilité "valut  $f(x) dx$  " et on remplace la somme (ou série ) par une intégrale ce qui conduit à :

$$\mu(X) = E(X) = \int_{\Omega} x f(x) dx \quad (1.2)$$

**Propriété 1.4.1 :**

- Si on ajoute une constante  $a$  à une variable aléatoire  $X$ , il en est de même pour son espérance

$$E(X + a) = E(X) + a ; \quad \forall a \in \mathbb{R} \quad (1.3)$$

- Si on multiplie une *v.a*  $X$  par une constante  $a$ , il en est de même pour son espérance :

$$E(aX) = aE(X) ; \quad \forall a \in \mathbb{R} \quad (1.4)$$

- L'espérance d'une somme de deux *v.a*  $X$  et  $Y$  est la somme des espérances de ces deux *v.a* :

$$E(X + Y) = E(X) + E(Y) \quad (1.5)$$

On peut résumer ces trois propriétés en une seule expression. En effet, le fait que l'opérateur espérance est linéaire alors on aura :

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y) ; \quad \forall (\alpha, \beta) \in \mathbb{R}^2$$

Une autre propriété très utile, concerne l'espérance d'une transformation d'une variable aléatoire. Soit  $g$  une fonction réelle quelconque et  $Y = g(X)$  une transformation de  $X$ . Alors l'espérance de la variable  $Y$  est donnée par :

$$E(Y) = E(g(X)) = \begin{cases} \sum g(x_i)P(X = x_i), & \text{dans le cas de variable discrète;} \\ \int_{-\infty}^{+\infty} g(x)f(x)dx, & \text{dans le cas de variable continue.} \end{cases} \quad (1.6)$$

**Remarque 1.4.1** *On peut remarquer que l'espérance d'une variable aléatoire binaire est égale à la probabilité de valoir 1, autrement dit si  $X$  est une v.a binaire (prenant ses valeurs dans  $\{0, 1\}$ , on a alors  $E[X] = P(X = 1)$ .*

## 1.5 La variance et l'écart-type d'une v.a

En statistique et en théorie des probabilités, la variance est une mesure de la dispersion des valeurs d'un échantillon ou d'une distribution de probabilité au tour de leurs moyenne. Pour l'aspect calculatoire la variance est égale à la différence entre la moyenne des carrés des valeurs de la variable et le carré de la moyenne, selon le **théorème de König-Huygens**. Ainsi, à cause de la nature de la fonction caré alors plus l'écart à la moyenne est grand plus il est prépondérant dans le calcul total de la variance qui donnerait donc une bonne idée sur la dispersion des valeurs.

$$\sigma^2 = V = V(X) = Var(X),$$

est définie par :

$$\sigma^2(X) = E([X - E(X)]^2) \quad (1.7)$$

C'est-à-dire :

$$\sigma^2(X) = \begin{cases} \sum_{i=1}^n [X_i - E(X)]^2 P(X = x_i), & \text{si } X \text{ est discrète;} \\ \int_{\Omega} [X_i - E(X)]^2 f(x) dx, & \text{si } X \text{ est continue.} \end{cases} \quad (1.8)$$

La variance est toujours positive, et ne s'annule que si les valeurs sont tous égales.

Sa racine carrée définit l'écart-type  $\sigma$ , d'où la notation :

$$\sigma(X) = \sqrt{v(X)} \quad (1.9)$$

**Propriété 1.5.1 :**

- Pour une constante  $c$ , on a :  $Var(c) = 0$
- Pour un couple de v.a  $(X, Y)$  on a :

$$\begin{aligned} Var(X + Y) &= Var(X) + Var(Y) + 2 Cov(X, Y), \\ Var(\alpha X + \beta Y) &= \alpha^2 Var(X) + \beta^2 Var(Y) + 2 \alpha \beta Cov(X, Y), \end{aligned}$$

avec  $\alpha$  et  $\beta$  sont des constantes réelles.

- Si  $X$  et  $Y$  sont deux variables aléatoires indépendantes le terme  $Cov(X, Y)$  est nul.

## 1.6 Quelques lois de probabilités usuelles

Cette partie définit brièvement les modèles de distributions uni-variées les plus fréquemment utilisés comme descriptions approximatives de distributions en statistique. Comme ces modèles dépendent de paramètres qui doivent être déterminés à l'aide des données que l'on souhaite décrire on les appelle des modèles paramé-

triques.

### 1.6.1 La distribution de Gauss (Normale)

La distribution de Gauss constitue à plusieurs égards la pierre angulaire de la statistique. Elle est caractérisée par deux paramètres : sa moyenne  $\mu$  et sa variance  $\sigma^2$ .

**Définition 1.6.1** *On dit qu'une variable aléatoire  $X$  obéit à une loi normale de moyenne  $\mu$  (où  $-\infty < \mu < +\infty$ ) et de variance  $\sigma^2$  lorsqu'elle présente la densité*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[ \left( -\frac{1}{2} \right) \left( \frac{x - \mu}{\sigma} \right)^2 \right]; -\infty < x < +\infty. \quad (1.10)$$

En abrégé, on écrit  $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$ .

#### Les propriétés de la distribution de Gauss

La loi normale possède plusieurs propriétés importantes, énumérées ci-après,

1.  $\int_{-\infty}^{\infty} f(x) dx = 1$ ;
2.  $f(x) \geq 0$ , pour tout  $x$ .
3.  $\lim_{x \rightarrow \pm\infty} f(x) = 0$ ;
4.  $f(\mu + x) = f(\mu - x)$ , c'est-à-dire la courbe de la densité  $f$  est symétrique par rapport à  $\mu$ .
5. La courbe de  $f$  atteint son maximum au point  $x = \mu$ , c'est-à-dire Mode=la moyenne.
6. Les points d'inflexion de la courbe de  $F$  se situent en  $x = \mu \pm \sigma$ .

### La distribution normale centrée et réduite

On dit que la variable aléatoire  $X$  à (ou suit) une distribution normale centrée et réduite ou une distribution de Gauss centrée et réduite qu'on note  $X \rightsquigarrow N(0, 1)$  si elle a pour densité

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad (1.11)$$

Autrement dit, une distribution normale centrée et réduite et une distribution normale dont la moyenne vaut zéro ( $\mu = 0$ ) et la variance vale un ( $\sigma^2 = 1$ ). Le graphique de  $f$  dans ce cas est une courbe "en cloche" (voir figure 1.3).

La fonction de répartition de  $X$  est :

$$F(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt. \quad (1.12)$$

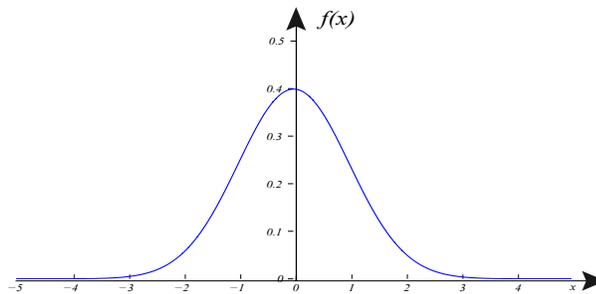


FIGURE 1.3 – Distribution de Gauss centrée et réduite

Pour déterminer les valeurs de  $F(x)$ , on se réfère à la table de la loi normal ou on utilise des programmes d'intégration numérique.

Si  $X \rightsquigarrow N(0, 1)$  alors, la variable aléatoire  $Y = \sigma X + \mu$  a une distribution de Gauss de moyenne  $\mu$  et de variance  $\sigma^2$ , c'est-à-dire  $Y \rightsquigarrow N(\mu, \sigma^2)$ . Tandis que, si la transformation précédente ( $Y = \sigma X + \mu$ ) est effectuée en sens inverse alors on passe d'une variable aléatoire  $Y$  de distribution  $N(\mu, \sigma^2)$  à la variable centrée et

réduite

$$X = \frac{Y - \mu}{\sigma}, \quad (1.13)$$

qui suit une distribution  $N(0, 1)$ .

Cette dernière transformation nous permet de calculer des probabilités relatives à la variable  $Y$  à l'aide de la fonction de répartition et des tables de  $N(0, 1)$ .

L'un des principaux résultats obtenus sur la distribution Normale est le théorème central limite résumé comme suit :

**Théorème 1.6.1** (*Théorème Limite Centrale (TCL)*)

Supposons que  $X_1, \dots, X_n$  soient *i.i.d.* (indépendantes et identiquement distribuées) selon une distribution  $F_X$  inconnue, telle que  $E(X_i) = \mu$  et  $Var(X_i) = \sigma^2$ .

Alors, si  $n \rightarrow \infty$ ,

$$\frac{\left( \sum_{i=1}^n X_i/n \right) - \mu}{\sigma/\sqrt{n}} \rightsquigarrow N(0, 1). \quad (1.14)$$

Ce théorème peut être interpréter comme suit : La distribution de la moyenne arithmétique centrée et réduite est donc approximativement Gaussienne  $N(0, 1)$ , indépendamment de la distribution  $F_X$ , pourvu que  $n$  soit suffisamment élevé. La distribution de la moyenne arithmétique est approximativement normale de moyenne  $\mu$  et variance  $\sigma^2/n$  et ces paramètres peuvent être estimés. Malheureusement, il n'y a pas en général une règle simple pour déterminer la valeur minimale de  $n$  pour que l'approximation soit bonne. Cette valeur dépend de la forme de  $F_X$ . Mais généralement, dans la pratique, on se contente de  $n \geq 30$ .

### 1.6.2 La distribution $\chi^2$ (*Khi – Deux*)

Soient  $X_1, \dots, X_n$ ,  $n$  variables aléatoires *i.i.d* selon une distribution normale standard. On dit que la variable aléatoire

$$Z = X_1^2 + X_2^2 + \dots + X_n^2,$$

à une *distribution*  $\chi^2$  à  $n$  *degrés de liberté* notée  $\chi_n^2$ . La densité de cette distribution est

$$f(z) = \frac{z^{(\frac{n}{2}-1)}}{2^{n/2}\Gamma(n/2)} e^{-z/2}, \quad z \geq 0, \quad (1.15)$$

avec  $\Gamma(\cdot)$  réfère à la fonction  $\Gamma$  (Gamma), définie par

$$\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx, \quad p > 0, \quad (1.16)$$

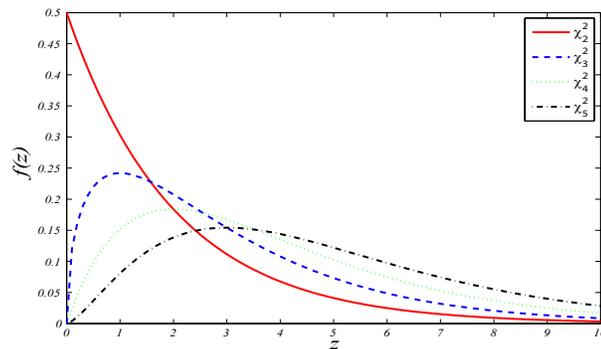


FIGURE 1.4 – Distribution de  $\chi_2^2$  pour différentes degré de liberté.

La fonction de répartition de la loi de  $\chi_n^2$  est généralement calculée à l'aide d'un programme informatique ou de "tables de la distribution  $\chi^2$ ". La moyenne et la variance de la distribution  $\chi^2$  sont  $E(Z) = n$  et  $\sigma^2(Z) = 2n$ .

**Remarque 1.6.1** Soit  $Z_1$  et  $Z_2$  deux variables aléatoires de distribution  $\chi^2$  de

degré liberté  $n$  et  $m$  respectivement, alors la variable aléatoire  $Z = Z_1 + Z_2$  est aussi une variable aléatoire d'une distribution de  $\chi^2$  de degré liberté  $n+m$  ( $Z \rightsquigarrow \chi^2_{(n+m)}$ ).

### 1.6.3 La distribution Student ( $t$ )

Supposons que  $X_0, X_1, \dots, X_n$ ,  $n$  variables aléatoires *i.i.d* selon une distribution normale standard. On dit que la variable aléatoire

$$T = \frac{X_0}{\sqrt{\frac{1}{n}(X_1^2 + \dots + X_n^2)}} = \frac{X_0}{\sqrt{Z/n}} \quad (\text{avec } Z \rightsquigarrow \chi_n^2) \quad (1.17)$$

à une distribution  $t$  (ou distribution de Student) à  $n$  degrés de liberté notée  $t_n$ .

La densité de cette distribution est

$$f(t) = \frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{n\pi}} (1 + t^2/n)^{-(n+1)/2}. \quad (1.18)$$

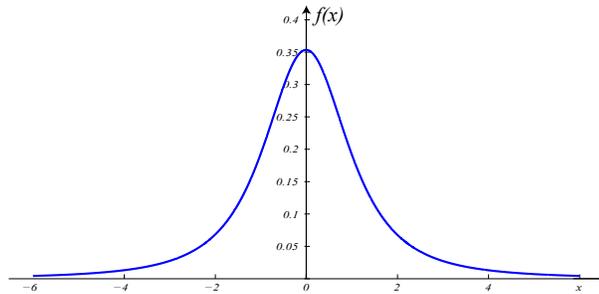


FIGURE 1.5 – Distribution de Student.

La fonction de distribution cumulative est généralement calculée à l'aide d'un programme informatique ou de "tables de la distribution  $t$ ". La moyenne et la variance de la distribution  $t$  sont  $E(T) = 0$  et  $\sigma^2(T) = n/(n-2)$ , pour  $n > 2$ .

### 1.6.4 La distribution Fisher (Fisher-Snedecor) $F$

Soit  $X_1, \dots, X_{n+m}$ ,  $n + m$  variables aléatoires indépendantes qui suivent une distribution normale centrée et réduite. On dit que la variable aléatoire

$$Y = \frac{\frac{1}{n}(X_1^2 + \dots + X_n^2)}{\frac{1}{m}(X_{n+1}^2 + \dots + X_{n+m}^2)} = \frac{Z_1/n}{Z_2/m} \quad (Z_1 \rightsquigarrow \chi_n^2 \text{ et } Z_2 \rightsquigarrow \chi_m^2) \quad (1.19)$$

a une distribution  $F$  avec  $n$  degrés de liberté au numérateur et  $m$  degrés de liberté au dénominateur notée  $F_{(n,m)}$  ou de degrés de libertés  $(n, m)$ . La densité de cette distribution est

$$f(y) = \frac{\Gamma((n+m)/2)}{\Gamma(n/2)\Gamma(m/2)} n^{n/2} m^{m/2} y^{(n/2)-1} (m+ny)^{-(n+m)/2}, \quad y \geq 0. \quad (1.20)$$

Les calculs concernant la distribution  $F$  sont généralement effectués à l'aide d'un programme informatique ou de "tables de la distribution  $F$ ". La moyenne de la distribution  $F$  est  $E(Y) = \frac{m}{m-2}$  pour  $m > 2$  et sa variance  $\sigma^2(Y) = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)}$ , pour  $m > 4$ .

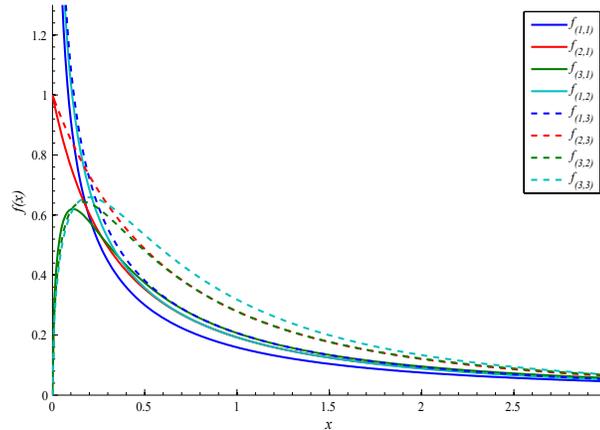


FIGURE 1.6 – Distribution de Fisher.

## Conclusion

Dans le présent chapitre nous avons focalisé uniquement sur la notion de variables aléatoire unidimensionnelle, où nous avons essayé de faire une brève description des fonctions (densité, répartition,  $\check{E}$ ) et des quantités (moyenne, variance,...) qui peuvent caractérisée cette variable.

Dans le chapitre suivant, l'objectif est de revoir l'adaptation de ces notions dans le cadre d'une variable aléatoire vectorielle.

## Chapitre 2

# Caractérisation d'une variable aléatoire vectorielle

### Introduction

Dans le présent chapitre nous allons exposer dans un premier lieu la notion d'une variable aléatoire multivariée plus précisément lorsque cette dernière se présente sous forme vectorielle. Par la suite, nous allons intéresser aux principales notions de la caractérisation de telles variables (fonction de densité, fonction de répartition, moyenne, ...). Enfin, la notion du vecteur gaussien sera exposée ainsi que les lois statistiques qu'on peut dériver de ce dernier.

## 2.1 Fonction de répartition et densité d'une v.a vectorielle

Un vecteur aléatoire  $X$  une application de  $(\Omega, P)$  dans un espace vectoriel réel, en général  $\mathbb{R}^d$  muni de sa tribu borélienne. En pratique  $\mathbb{R}^d$  est muni de sa canonique et on identifiera  $X$  au  $d$ -uple de variables aléatoires formées par des composantes sur cette base  $X = (X_1, \dots, X_n)$ .

### 2.1.1 Fonction de répartition :

De même que dans le cas univarié le comportement d'une variable aléatoire vectorielle peut être caractérisé par une fonction de répartition notée  $F$ . La fonction de répartition  $F$  dans ce cas est une application de  $\mathbb{R}^d$  dans  $[0; 1]$  définie par :

$$F(x_1, \dots, x_d) = P(X_1 < x_1, \dots, X_d < x_d). \quad (2.1)$$

**Propriété 2.1.1** *Par ailleurs, il est à noter que toute fonction de répartition  $F$  associée à une variable aléatoire vectorielle satisfait les propriétés suivantes :*

1.  $\lim_{x_1 \rightarrow \infty} \cdots \lim_{x_d \rightarrow \infty} F(x_1, \dots, x_d) = 1$  ;
2.  $\lim_{x_i \rightarrow -\infty} F(x_1, \dots, x_i, \dots, x_d) = 0, \quad \forall i \in \{1, 2, \dots, d\}$ . Ceci découle du fait que l'intersection de l'ensemble vide  $\emptyset$  avec n'importe lequel autre ensemble est l'ensemble  $\emptyset$ , dont la probabilité est nulle.
3.  $F$  est une fonction non-décroissante et continue à droite dans chacun de ces  $d$  arguments.

*Par exemple, si on considère que  $x_2, \dots, x_d \in \mathbb{R}$  sont fixés, alors  $t \rightarrow$*

$F(t, x_2, \dots, x_d)$  est une fonction non-décroissante et continue à droite.

4.  $F$  est  $d$ -monotone.

Cette dernière propriété peut être interprétée par exemple :

- Dans le cas  $d = 2$ , elle signifie que pour

$$x_1 < x_1^* \text{ et } x_2 < x_2^* \Rightarrow F(x_1^*, x_2^*) - F(x_1, x_2^*) - F(x_1^*, x_2) + F(x_1, x_2) \geq 0.$$

- Enfin, et en dimension  $d$  arbitraire on aura :

$$\sum_{c_i \in \{x_i, x_i^*\}_{i \in \{1, \dots, d\}}} (-1)^{v(c)} \times F(c_1, \dots, c_d) \geq 0,$$

où  $v(c)$  est le nombre de constantes  $c_i$  sans étoiles. Inversement, toute fonction  $F : \mathbb{R}^d \rightarrow [0, 1]$  qui satisfait (1) – (4) est une fonction de répartition valide, c'est-à-dire qu'il existe une mesure de probabilité qui engendre  $F$ . La preuve est la même que dans le cas univarié ; l'élément clé est que  $]x_1, x_1^*] \times ]x_2, x_2^*] \times \dots \times ]x_d, x_d^*]$  sont précisément des générateurs de la tribu borélienne sur  $\mathbb{R}^d$ .

- La figure 2.1 est un exemple d'illustration de la notion de monotonie de  $F(x)$  via des ensembles. Supposons qu'on s'intéresse au calcul de la probabilité d'un événement décrit comme suit :

$$P(A \cup B \cup C \cup D) - P(A \cup C) - P(A \cup B) + P(A).$$

dans ce cas, puisqu'il s'agit d'ensembles disjoints alors en termes d'aire, c'est  $(A + B + C + D) - (A + C) - (A + B) + A = D$ , ainsi la probabilité  $P(a_1 < X_1 < b_1, a_2 < X_2 < b_2) = D$  est positive.

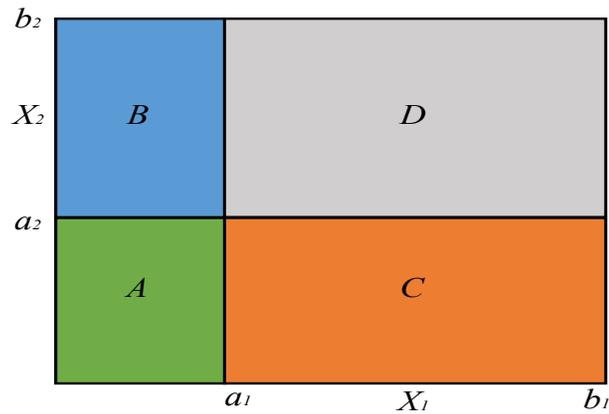


FIGURE 2.1 – Monotonie illustré par le biais d'ensembles.

Les propriétés (1) – (3) sont suffisantes en une dimension, mais la propriété essentielle pour la construction d'une fonction de répartition multivariée est bien que la (4). On dénomme loi marginale la loi correspondant à une sous-composante de dimension  $d - q$  du  $d$ -vecteur  $X$ , pour  $1 \leq q \leq d$ .

### 2.1.2 Fonction de densité :

La fonction de densité  $f$ , associée à une variable aléatoire vectorielle, si elle existe elle est définie par :

$$f(x_1, \dots, x_d) = \frac{\partial^d F}{\partial x_1 \partial x_2 \dots \partial x_d}. \quad (2.2)$$

### 2.1.3 Changement de variables dans une densité

Effectuons le changement de variables défini par :

$$Y_i = \varphi_i(X_1, \dots, X_d)$$

Les fonctions  $\varphi_i$  étant telles que le passage de  $(X_1, \dots, X_d)$  à  $(Y_1, \dots, Y_d)$  est biuni-

voque. Nous désignerons en abrégé par  $\varphi$  la transformation  $X \xrightarrow{\varphi} Y$  :

$$Y = \varphi(X).$$

Ainsi, la densité  $g$  du nouveau vecteur  $Y$ , s'obtient alors par la formule :

$$g(y) = \frac{f[\varphi^{-1}(y)]}{|\det J|},$$

où  $\det J$ , appelé Jacobien de la transformation, est tel que :

$$\det J = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \cdots & \frac{\partial y_d}{\partial x_1} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_d} & \cdots & \cdots & \frac{\partial y_d}{\partial x_d} \end{vmatrix} \quad \text{et} \quad (\det J)^{-1} = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \cdots & \frac{\partial x_1}{\partial y_d} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial x_d}{\partial y_1} & \cdots & \cdots & \frac{\partial x_d}{\partial y_d} \end{vmatrix} = \det J^{-1}$$

La démonstration de cette propriété figure dans tous les ouvrages consacrés à l'intégration, et ceci au niveau changement de variables dans les intégrales multiples (voir l'exemple de [6, 7]).

### Exemple 2.1.1

On fait l'hypothèse d'une répartition de densité (gaussienne) dans le quart de plan :

$$f(x, y) = \exp(-x - y) ; x > 0 \text{ et } y > 0$$

Supposons qu'on désire réaliser le changement de variable :  $u = x + y$  et  $v = \frac{x}{y}$ .

On cherche la loi de probabilité du couple  $(u, v)$  dans son domaine d'application.

Les formules de passage d'un espace à l'autre, pour ce cas sont données comme suit :

$$\begin{cases} u = x + y \\ v = \frac{x}{y} \end{cases} \Rightarrow \begin{cases} x = \frac{uv}{1+v} \\ y = \frac{u}{1+v} \end{cases}$$

Ainsi, après calcul des différentes dérivées nécessaires, on aura la matrice Jacobien suivante :

$$J = \begin{pmatrix} \frac{v}{1+v} & \frac{u}{(1+v)^2} \\ \frac{1}{1+v} & \frac{-u}{(1+v)^2} \end{pmatrix},$$

par conséquent,

$$\det J = \frac{-u}{(1+v)^2}.$$

Cependant, le fait que  $u > 0$ , alors il faut prendre :  $|J| = \frac{u}{(1+v)^2}$ .

D'où l'expression de la densité :

$$f(x, y) = \exp(-x - y) \longrightarrow \frac{u \exp(-u)}{(1+v)^2},$$

et les distributions marginales sont :

$$\begin{aligned}
 f_u(u) &= u \exp(-u) \int_0^{+\infty} \frac{dv}{(1+v)^2} \\
 &= u \exp(-u), \quad u > 0 \\
 f_v(v) &= \frac{1}{(1+v)^2} \int_0^{+\infty} u \exp(-u) du \\
 &= \frac{1}{(1+v)^2}, \quad v > 0.
 \end{aligned}$$

**Remarque 2.1.1** *Si la transformation  $\varphi$  est linéaire de matrice constante  $A$ , c'est-à-dire  $Y = AX$ , alors sous la condition de la régularité de la matrice  $A$ , on a  $\det J = |A|$ . En particulier si  $A$  est une transformation orthogonale le Jacobien vaut 1.*

## 2.2 Fonction caractéristique et Théorème de Cramer-Wold

### 2.2.1 Fonction caractéristique

Soit  $t$  un vecteur non aléatoire. On définit la fonction caractéristique  $\Phi_X : \mathbb{R}^d \rightarrow \mathbb{C}$  d'un vecteur aléatoire  $X$  la fonction de l'argument vectoriel a définie par :

$$\Phi_X(t) = E[\exp(it'X)], \quad \forall t \in \mathbb{R}^d, \quad (2.3)$$

avec  $t'$  est la transposée du vecteur  $t$  et  $i$  est le nombre imaginaire tel que  $i^2 = -1$ .

Autrement dit, si  $X = (X_1, \dots, X_d)$  et  $t = (t_1, \dots, t_d)$  alors :

$$\Phi_X(t) = E[\exp(i(t_1X_1 + \dots + t_dX_d))] \quad (2.4)$$

**Théorème 2.2.1** *Les composantes  $X_1, \dots, X_d$  de  $X$  sont indépendantes si et seulement si la fonction caractéristique de  $X$  est égale au produit des fonctions caractéristiques de ses composantes*

$$\Phi_X(t) = \prod_{i=1}^d \Phi_{X_i}(t_i) \quad (2.5)$$

*Si les  $X_i$  sont indépendantes alors l'espérance d'un produit de fonctions de caractéristique des  $X_i$  est égale au produit de leurs espérances, c'est-à-dire :*

$$E[\exp(it'X)] = E[\exp(it_1X_1)] \times E[\exp(it_2X_2)] \times \dots \times E[\exp(it_dX_d)]. \quad (2.6)$$

La réciproque plus délicate utilise l'inversion de la fonction caractéristique et est omise. Dans ce qui suit nous allons présenter un résultat fondamental qui nous permet de définir des lois de probabilités à  $d$ -dimensions à partir des lois unidimensionnelles.

## 2.2.2 Théorème de Cramer-Wold

Cramer-Wold on énoncé un résultats de grande importance où ils indiquent que la loi du vecteur aléatoire  $X$  est entièrement déterminée par celles de toutes les combinaisons linéaires de ses composantes.

En effet, Il suffit de poser  $Y = it' = \sum_{i=1}^d i_i t_i$  et de chercher la fonction caractéristique de la variable aléatoire univarié  $Y$  :

$$\Phi_Y(u) = E[\exp(iuY)] = E[\exp(iut'X)]. \quad (2.7)$$

d'où  $\Phi_Y(1) = \Phi_X(t)$ .

Ainsi, si la loi de la variable univarié  $Y$  est connue pour tout  $t$  on connaît donc la fonction caractéristique de la loi de la variable vectorielle  $X$ .

## 2.3 Espérance, variances-covariances et corrélations

Sois  $\mu = E(X)$ , qui désigne l'espérance de la variable aléatoire vectorielle  $X$ . On appelle par définition espérance du vecteur  $X = (X_1, \dots, X_d)$ , le vecteur certain défini par :

$$E(X) = (E(X_1), E(X_2), \dots, E(X_d))' \quad (2.8)$$

$$= (\mu_1, \mu_2, \dots, \mu_d)' \quad (2.9)$$

La matrice de variances-covariances, appelée aussi matrice de covariances,  $\Sigma$  de  $X$  est définie par :

$$\begin{aligned} \Sigma_X &= Cov(X) = E[(X - E(X))(X - E(X))'] \\ &= E(XX') - E(X)[E(X)]' \\ &= E(XX') - \mu\mu'. \end{aligned}$$

**Propriété 2.3.1** *En développant les termes de l'expression de  $\Sigma_X$  on obtient la forme suivante :*

$$\Sigma_x = \begin{pmatrix} \sigma_{X_1}^2 & Cov(X_1, X_2) & \cdots & Cov(X_1, X_d) \\ Cov(X_2, X_1) & \sigma_{X_2}^2 & & \\ \vdots & & \ddots & \vdots \\ Cov(X_d, X_1) & \cdots & & \sigma_{X_d}^2 \end{pmatrix}$$

**Propriété 2.3.2** *La matrice de Variances-Covariances  $\Sigma_X$  possède les propriétés suivantes :*

- La matrice est symétrique car on a  $Cov(X, Y) = Cov(Y, X)$
- La matrice est semi-définie positive (ses valeurs propres sont positives ou nulles).
- Les éléments de sa diagonale ( $\Sigma_{i,j}$ ) représentent la variance de chaque éléments, étant donné la propriété  $Cov(X_i, X_i) = Var(X_i)$
- Les éléments en dehors de la diagonale ( $\Sigma_{i,j}, i \neq j$ ) représentent la covariance entre les variables  $i$  et  $j$ .
- La matrice de covariance d'un vecteur non corrélé ou indépendant est diagonale ( $\Sigma_{i,j} = 0; \forall i \neq j$ ).

item C'est une matrice carrée symétrique d'ordre  $d$ , c'est-à-dire  $\forall i, j$  on a  $\Sigma_{i,j} = \Sigma_{j,i}$ .

**Définition 2.3.1** *Soit  $X$  et  $Y$  deux variables aléatoire univarié. Le coefficient de corrélation entre  $X$  et  $Y$  est défini par :*

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{V(X)V(Y)}}. \quad (2.10)$$

Ci-dessous quelque propriétés du coefficient de corrélation :

1.  $-1 \leq \rho(X, Y) \leq 1$
2. Si  $X$  et  $Y$  sont indépendantes alors  $Cov(X, Y) = 0$  et  $\rho(X, Y) = 0$  (l'inverse n'est pas forcément vrai).

3. Si  $Y = a + bX$  alors

- $\rho(X, Y) = -1$  si  $b < 0$ .
- $\rho(X, Y) = +1$  si  $b > 0$ .

Soit  $S$  une matrice carrée de dimension  $d$  dont ses éléments sont donnés par :

$$S_{ij} = \begin{cases} \sqrt{\Sigma_{ii}}, & \text{si } i = j; \\ 0, & \text{sinon.} \end{cases} \quad (2.11)$$

**Définition 2.3.2** Soit  $X$  variable aléatoire vectorielle. La matrice des coefficients de corrélation entre les éléments de  $X$  est définie par :

$$\rho_X = S^{-1}\Sigma_X S^{-1}, \quad (2.12)$$

avec  $S^{-1}$  est l'inverse de la matrice définie dans (2.11).

**Remarque 2.3.1**

1. Les propriétés du coefficient de corrélation introduites précédemment reste vraie pour les éléments de la matrice de corrélation  $\rho_X$ .
2. Si les variable  $X_i$  du  $X$  sont réduites, alors  $\Sigma$  s'identifie avec la matrice de corrélation :

$$\Sigma_X = \rho_X = \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1d} \\ \rho_{21} & 1 & & \vdots \\ \vdots & & \ddots & \\ \rho_{d1} & \dots & & 1 \end{bmatrix}$$

## 2.4 Transformations linéaires

Supposons que nous effectuons un changement de variable linéaire  $Y = AX$  où  $A$  est une matrice quelconque de constantes (pas nécessairement carrée), alors :

$$\mu_Y = A\mu_X \quad (2.13)$$

$$\Sigma_Y = A\Sigma_X A', \quad (2.14)$$

ce qui se démontre en appliquant les définitions.

En particulier si  $A$  est une matrice uniligne,  $Y$  est alors une variable aléatoire unidimensionnelle. Si  $t'$  désigne cette ligne  $Y = \sum_{i=1}^d t_i X_i$  et  $Var(Y) = t'\Sigma t$ . On a donc pour tout  $t$ ,  $t'\Sigma t \geq 0$  car une variance est non négative. On en déduit le résultat suivant :

**Théorème 2.4.1** *Une condition nécessaire et suffisante pour qu'une matrice  $\Sigma$  symétrique soit la matrice de variance d'un vecteur aléatoire est que  $\Sigma$  soit une matrice positive.*

La réciproque s'établit à partir de la propriété classique suivante élaborée sur des matrices symétriques définies positives :

**Propriété 2.4.1** *Tout matrice symétrique positive  $\Sigma$  peut s'écrire sous la forme  $\Sigma = TT'$  où  $T$  est définie à une transformation orthogonale près (si  $T$  convient,  $S = TU$ , où  $U$  est orthogonale, convient aussi ; une solution particulière est fournie par  $T = \Sigma^{1/2} = P\Lambda^{1/2}P'$  où  $P$  est la matrice des vecteurs propres normés de  $T$  et  $\Lambda$  la matrice diagonale des valeurs propres). Il suffit donc de partir d'un vecteur aléatoire  $X$  de matrice de variance  $I$ , (par exemple un  $p$ -uplet de variables indépendantes centrées réduites) et de faire la transformation  $Y = TX$  pour obtenir*

un vecteur aléatoire de matrice de variance  $\Sigma$ .

Si  $\Sigma$  est régulière, c'est-à-dire si les composantes de  $X$  ne sont pas linéairement dépendantes on peut trouver une transformation inverse qui "normalise" le vecteur  $X$ .

**Théorème 2.4.2** *Si  $\Sigma$  est régulière il existe une infinité de transformations linéaires  $A$ , telles que  $Y = AX$  soit un vecteur de matrice de variance  $I$ .*

Il suffit de prendre  $A = T^{-1}$ . Un choix particulièrement intéressant est celui de  $T = \Sigma^{1/2}$ .

**Définition 2.4.1** *On appelle transformation de Mahalanobis la transformation définie par  $\Sigma^{-1/2}$ .*

$Y = \Sigma^{-1/2}(X - \mu)$  est alors un vecteur aléatoire centré-réduit à composantes non corrélées. On en déduit aisément le résultat suivant :

**Théorème 2.4.3** *La variable aléatoire  $(X - \mu)' \Sigma^{-1} (X - \mu) = D^2$  a pour espérance  $d$ .*

En effet  $D^2 = \sum_{i=1}^d Y_i^2$  sont d'espérance nulle et de variance 1. La fonctionnelle  $D$  est appelée distance de Mahalanobis de  $X$  et  $\mu$ .

## 2.5 Vecteurs aléatoires Gaussiens

Dans cette section nous allons s'intéresser à une variable aléatoire vectorielle, lorsque ses composantes sont de variables aléatoire gaussiennes ainsi on aura la notion de la loi multinormale. La définition détaillée et ses caractéristiques de cette dernière feront l'objet du reste de la présente section.

### 2.5.1 La loi multinormale

**Définition 2.5.1** *On dit que  $X$  est un vecteur gaussien à  $d$  dimensions, si toute combinaison linéaire de ses composantes  $t'X$  suit une loi normale à une (01) dimension.*

**Le théorème de Crame-Wold** permet d'établir que la loi de  $X$  est ainsi parfaitement déterminée. On remarquera que la normalité de chaque composante ne suffit nullement à définir un vecteur gaussien. La fonction caractéristique de  $X$  s'en déduit aisément (ou supposera ici que  $X$  est centrée ce qui ne nuit pas à la généralité).

**Théorème 2.5.1** *Sois  $\Sigma$  est la matrice de variance du vecteur aléatoire  $X \rightarrow N(\mu, \Sigma)$ , alors  $\Phi_X(t) = \exp\left(-\frac{1}{2}t'\Sigma t\right)$ .*

La démonstration de ce dernier théorème est immédiate. En effet d'après la théorème de Crame-Wold :  $\Phi_X(t) = \Phi_Y(1)$  où  $Y = t'X$ .

Ainsi, la loi de  $Y$  est par définition une gaussienne centrée de variance  $v(Y) = t'\Sigma t$  est la fonction caractéristique de  $Y$  est  $\Phi_Y(t) = \exp\left(-\frac{u^2}{2}v(Y)\right)$  ce qui établit le résultat. On en déduit le résultat fondamental suivant :

**Théorème 2.5.2** *Les composantes d'un vecteur gaussien  $X$  sont indépendantes si et seulement si  $\Sigma$  est diagonale, c'est-à-dire si elles sont non corrélées.*

On a en effet, si  $\Sigma$  est diagonale de termes  $\sigma_i^2$  :

$$\Phi_X(t) = \exp\left(-\frac{1}{2}\sum_{i=1}^d t_i^2 \sigma_i^2\right) = \prod_{i=1}^d \Phi_{X_i}(t_i) \quad (2.15)$$

On notera  $\mathcal{N}_d(\mu; \Sigma)$  la loi normale à  $d$  dimensions d'espérance  $\mu$  et de matrice Variance  $\Sigma$ .

### Densité de la loi normale à $d$ -dimensions

La densité d'une loi multi-normale n'existe que lorsque la matrice  $\Sigma$  est régulière, le théorème suivant en est une preuve.

**Théorème 2.5.3** *Si  $\Sigma$  est régulière  $X$  admet pour densité :*

$$f(x_1, \dots, x_d) = \frac{1}{(2\pi)^{d/2} (\det \Sigma)^{1/2}} \exp \left[ -\frac{1}{2} (x - \mu)' \Sigma^{-1} (x - \mu) \right]. \quad (2.16)$$

D'après ce dernier résultat, on déduit que la variable  $Y = \Sigma^{-1/2}(x - \mu)$  est alors un vecteur gaussien dont les composantes sont centrées-réduites et indépendantes,  $Y$  a pour densité :

$$g(y) = \frac{1}{(2\pi)^{d/2}} \exp \left( -\frac{1}{2} \sum_{i=1}^d y_i^2 \right) \quad (2.17)$$

### Cas particulier de la loi normale à deux dimensions

Si l'on introduit  $\rho$  coefficient de corrélation linéaire entre  $X_1$  et  $X_2$ , alors on aura :

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

d'où,

$$\det \Sigma = (\sigma_1\sigma_2)^2 (1 - \rho^2),$$

et

$$\Sigma^{-1} = \frac{1}{\det \Sigma} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix}$$

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{\left(\frac{x_1-m_1}{\sigma_1}\right)^2 - 2\rho\frac{(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-m_2}{\sigma_2}\right)^2}{2(1-\rho^2)} \right\} \quad (2.18)$$

### Théorème central-limite multidimensionnel

Soit  $X_1, \dots, X_n$  une suite de vecteurs aléatoires indépendants de même loi, d'espérance  $\mu = (\mu_1, \dots, \mu_d)$  et de matrice de variance  $\Sigma$  alors de même que pour des lois à une dimension on peut établir le résultat suivant :

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}_d(0, \Sigma) \quad (2.19)$$

### 2.5.2 Estimation des paramètres $\mu$ et $\Sigma$

Dans cette section, on dérive les estimations des paramètres de la loi multinormale à l'aide de la méthode du maximum de vraisemblance. Soit un échantillon  $X_1, \dots, X_n$ . La log-vraisemblance  $l = \log(L)$  est donnée par :

$$\begin{aligned} l(X; \mu, \Sigma) &= -\frac{np}{2} \log(2\pi) + \frac{n}{2} \log \det(\Sigma^{-1}) - \frac{1}{2} \sum_{i=1}^n (X_i - \mu)' \Sigma^{-1} (X_i - \mu) \\ &= -\frac{np}{2} \log(2\pi) + \frac{n}{2} \log \det(\Sigma^{-1}) - \frac{1}{2} \text{tr} \left( \Sigma^{-1} \sum_{i=1}^n (X_i - \mu)(X_i - \mu)' \right) \end{aligned}$$

On s'intéresse à la trace et tout particulièrement à la somme des termes centres.

L'on a pour cette dernière

$$\sum_{i=1}^n (X_i \pm \bar{X} - \mu)(X_i \pm \bar{X} - \mu)' = \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' + n(\bar{X} - \mu)(\bar{X} - \mu)',$$

puisque les termes croisés  $\sum_{i=1}^n (X_i - \bar{X})(\bar{X} - \mu)'$  sont nuls. La log-vraisemblance par rapport à  $\mu$  est proportionnelle à

$$l(X; \mu, \Sigma) \propto \frac{-1}{2} \text{tr} \left( \Sigma^{-1} \left( \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' + n(\bar{X} - \mu)(\bar{X} - \mu)' \right) \right),$$

et en simplifiant l'écriture, on aura que

$$l(X; \mu, \Sigma) = -\frac{np}{2} \log(2\pi) + \frac{n}{2} \log \det(\Sigma^{-1}) - \frac{1}{2} \text{tr}(\Sigma^{-1} S_{XX}) - \frac{n}{2} (\bar{X} - \mu)' \Sigma^{-1} (\bar{X} - \mu). \quad (2.20)$$

Le maximum de vraisemblance de la moyenne découle immédiatement de cette dernière expression où on trouve  $\hat{\mu} = \bar{X}$ . En dérivant par rapport à la matrice de précision, on obtient :

$$\frac{\partial l}{\partial \Sigma^{-1}} = \frac{n}{2} \Sigma^{-1} - \frac{1}{2} S'_{XX}, \quad (2.21)$$

et en égalant à zéro, on obtient  $\hat{\Sigma} = n^{-1} S_{XX}$ . On procède maintenant au calcul du biais des estimateurs du maximum de vraisemblance (*EMV*) :

$$E(\hat{\mu}) = E \left( \frac{1}{n} \sum_{i=1}^n X_i \right) = \frac{n}{n} \mu = \mu \quad (2.22)$$

$$\begin{aligned}
 E(\hat{\Sigma}) &= \frac{1}{n} E \left( \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' \right) \\
 &= \frac{1}{n} E \left( \sum_{i=1}^n (X_i - \mu)(X_i - \mu)' - n(\mu - \bar{X})(\mu - \bar{X})' \right) \\
 &= \frac{1}{n} n \Sigma - \text{Var}(\bar{X}) = \frac{n-1}{n} \Sigma
 \end{aligned} \tag{2.23}$$

D'après ces deux résultats on déduit que l'estimateur du vecteur moyen est un estimateur sans biais, tandis que l'estimateur de  $\Sigma$  a un biais de  $n^{-1}\Sigma$ . Pour ce dernier, on peut prendre alors l'estimateur corrigé (sans biais) suivant :

$$\hat{\Sigma}_c = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})'. \tag{2.24}$$

### 2.5.3 Lois dérivées de la loi normale multivariée

Du même principe que dans le cas de la loi normale univariée, plusieurs lois peuvent être dérivées de la loi normale multivariée et qui sont d'une grande utilité dans les statistiques mathématique multidimensionnelle. Dans ce qui suit nous allons présenter les lois statistiques dérivées les plus usitées dans la pratique.

#### Loi de Wishart

**Définition 2.5.2** Une matrice  $M(d, d)$  a une distribution de Wishart  $W_d(n, d)$  si  $M$  peut s'écrire  $M = X'X$  où  $X$  est une matrice  $(n, d)$  aléatoire définie de la façon suivante : les  $n$  lignes de  $X$  sont des vecteurs aléatoires gaussiens de même loi  $N_d(0; \Sigma)$  indépendants.

$X$  représente donc un échantillon de  $n$  observations indépendantes d'une loi normale multidimensionnelle.

Nous allons voir que cette loi généralise d'une certaine façon la loi du  $\mathcal{X}^2$ . Si  $d = 1$  on a en effet :

$$W_l(n, \sigma^2) = \sigma^2 X_n^2 = \sum_{i=1}^n x_i^2 \quad (2.25)$$

On montre que la densité de la loi de Wishart est :

$$f(M) = \frac{|M|^{(n-d-1)/2} \exp\left(-\frac{1}{2}Tr\Sigma^{-1}M\right)}{2^{nd/2} \pi^{d(d-1)/4} |\Sigma|^{n/2} \prod_{i=1}^d \Gamma\left(\frac{1}{2}(n+i-1)\right)}, \quad (2.26)$$

avec  $M > 0$  pour la mesure de Lebesgue dans  $\mathbb{R}^{d(d+1)/2}$  (avec  $M$  doit être symétrique et semi-définie positive). On rapprochera cette formule par celle de la densité d'un  $\mathcal{X}^2$ . On note également que la fonction caractéristique de la loi Wishart  $W_d(n, \Sigma)$  est :

$$E[\exp(iT\Sigma)] = |I - iT\Sigma|^{-n/2},$$

où  $T$  est une matrice de dimension  $d \times d$ .

Rappelons que la fonction caractéristique d'un  $\mathcal{X}_n^2$  est  $\varphi_{\mathcal{X}_n^2}(t) = (1 - 2it)^{-n/2}$ .

On a :  $E(M) = n\Sigma$  et  $E(M^{-1}) = \frac{1}{n-d-1}\Sigma^{-1}$  si  $n-d-1 > 0$

Pour tout vecteur constant  $a$  :

$$\frac{a'Ma}{a'\Sigma a} \text{ suit une loi } \mathcal{X}_n^2$$

En effet on vérifie sans peine que  $a'Ma$  est une matrice de Wishart  $W_1(n; a'\Sigma a)$  car  $a'Ma = a'X'X$  où  $Xa$  suit  $N_1(0; a'\Sigma a)$ .

On peut montrer également que  $\frac{a'\Sigma^{-1}a}{a'M^{-1}a}$  suit une loi  $\mathcal{X}_{n-d-1}^2$ . Les deux résultats suivants se généralisent avec des vecteurs aléatoires de lois quelconques.

**Propriété 2.5.1** *Soit  $x$  un vecteur aléatoire (de loi quelconque) indépendant de*

$M$  alors :

**Proposition 2.5.1**

$$\frac{x^T M x}{x^T \Sigma x} \text{ et } \frac{x^T \Sigma^{-1} x}{x^T M^{-1} x} \quad (2.27)$$

suivent les lois  $\mathcal{X}_n^2$  et  $\mathcal{X}_{n-p+1}^2$  respectivement et sont des variables indépendantes de  $a' M a$  et  $b' M^{-1} b$  sont indépendantes si  $a' \Sigma b = 0$ .

**La loi du  $\mathbf{T}^2$  de Hotelling**

Cette distribution généralise celle de Student (ou plutôt son carré). C'est celle d'un variable unidimensionnelle.

**Définition 2.5.3** Soit  $x$  un vecteur aléatoire normale  $N_d(0; I_d)$ , où  $I_d$  désigne la matrice identité  $d \times d$ , et une matrice de Wishart  $W_d(n; I_d)$ , indépendante de  $X$ ; alors la quantité  $n x' M^{-1} x$  suit par définition une loi du  $\mathbf{T}^2$  de Hotelling de paramètres  $d$  et  $n$ .

Par abus de notation, on posera :  $\mathbf{T}_d^2(n) = n x^T M^{-1} x$ .

**Propriété 2.5.2** Si  $x$  suit une loi  $N_d(\mu, \Sigma)$  et  $M$  une loi de Wishart  $W_d(n; \Sigma)$  indépendante de  $X$ , alors  $n(x - \mu)' M^{-1} (x - \mu)$  suit une loi  $\mathbf{T}_d^2(n)$ .

La démonstration évidente elle découle de l'utilisation de la transformation de Mahalanobis  $y = \Sigma^{-1/2}(x - \mu)$  et le fait que  $\Sigma^{-1/2} M \Sigma^{-1/2}$  est une  $W_d(n; I_d)$ . la variable  $n x' M^{-1} x$  suit ce qu'on appelle une loi de Hotelling décentrée  $\mathbf{T}_d^2(n, \lambda^2)$  où  $\lambda^2 = \mu' \Sigma \mu$  est le paramètre de décentrement.

La loi du  $\mathbf{T}^2$  de Hotelling s'identifie à celle de Fisher-Snedecor selon la formule :

$$\mathbf{T}_d^2(n) = \frac{nd}{n-d+1} F(d; n-d+1) \quad (2.28)$$

En effet, on peut écrire avec  $x$ ,  $N_d(0; I)$  :

$$\mathbf{T}_d^2(n) = \frac{nx' M^{-1} x}{x' x} \quad (2.29)$$

où  $\frac{x' x}{x' M^{-1} x}$  est un  $\mathcal{X}_{n-d+1}^2$  indépendant de  $x$  donc de  $x' x$  qui est un  $\mathcal{X}_d^2$  d'où :

$$\mathbf{T}_d^2(n) = n \frac{\mathcal{X}_d^2}{\mathcal{X}_{n-d+1}^2} \quad (2.30)$$

On voit que pour  $d = 1$ ,  $\mathbf{T}_d^2(n) = F(1; n)$  c'est-à-dire le carré de la variable de Student à  $n$  degrés de liberté.

Notons que :

$$E(\mathbf{T}_d^2(n)) = \frac{nd}{n-d+1} \quad (2.31)$$

**Remarque 2.5.1** *On peut généraliser les résultats obtenus avec la loi normale à une classe de vecteurs aléatoires dont la densité est de la forme générale  $f(x)\alpha g((x-\mu)' \Sigma^{-1}(x-\mu))$ . Les principaux exemples sont la loi multinormale et la loi de Student- $t$ .*

### Loi du Lambda $\Lambda$ de Wilks (généralisation de la loi de Fisher)

La loi du lambda de Wilks s'associe à la validation d'hypothèses de l'Analyse de Variance multivariée en traitant directement sur les matrices résiduelles.

**Définition 2.5.4** *Si  $A \rightsquigarrow W_d(m, \Sigma)$  et  $B \rightsquigarrow W_d(n, \Sigma)$  indépendantes et  $m, n \geq d$ , alors la loi du lambda de Wilks peut s'obtenir comme suit :*

$$U = \frac{|A|}{|A+B|} = \frac{1}{|A^{-1}B + I|} \sim \Lambda(d, m, n) \quad (2.32)$$

Ceci est en lien avec la loi de Fisher dans quatre circonstances :

1. Quand le paramètre  $d = 1$ . C'est-à-dire, que les matrices  $A$  et  $B$  suivent une loi khi-deux avec  $m$  et  $n$  degrés de liberté respectivement. Alors,

$$\frac{1 - \Lambda(1, m, n)}{\Lambda(p, m, n)} = \frac{1 - \frac{|A|}{|A+B|}}{\frac{|A|}{|A+B|}} = \frac{|A+B| - |A|}{|A|} \quad (2.33)$$

Rappelons-nous que si  $d = 1$ , nous avons  $A = X'X = \sum_{i=1}^m x_i^2$  et  $B = Y'Y = \sum_{i=1}^n y_i^2$  alors

$$\frac{1 - \Lambda(1, m, n)}{\Lambda(p, m, n)} = \frac{|A+B| - |A|}{|A|} = \frac{\sum_{i=1}^n y_i^2}{\sum_{i=1}^m x_i^2} = \frac{\mathcal{X}_n^2}{\mathcal{X}_m^2} = \frac{n}{m} F(n, m), \quad (2.34)$$

c'est-à-dire,

$$U = \frac{\mathcal{X}_m^2}{\mathcal{X}_m^2 + \mathcal{X}_n^2} = \frac{1}{1 + \frac{\mathcal{X}_n^2}{\mathcal{X}_m^2}} = \frac{1}{1 + \frac{n}{m} F(n, m)}, \quad (2.35)$$

donc,

$$\frac{n}{m} F(n, m) = \frac{1 - U}{U} \quad (2.36)$$

De plus, il est à noter que la relation entre la distribution de la loi Fisher et la distribution de la loi Bêta comme suit : si  $X \sim F(d_1, d_2)$  alors  $\frac{\frac{d_1}{d_2} X}{1 + \frac{d_1}{d_2} X} \sim$

$Beta\left(\frac{d_1}{2}, \frac{d_1}{2}\right)$ . Celle-ci nous permet d'observer que

$$\frac{\frac{d_1}{d_2} X}{1 + \frac{d_1}{d_2} X} = \frac{\frac{d_1}{d_2} \frac{d_2}{d_1} \frac{\mathcal{X}_{d_1}^2}{\mathcal{X}_{d_1}^2}}{1 + \frac{d_1}{d_2} \frac{d_2}{d_1} \frac{\mathcal{X}_{d_1}^2}{\mathcal{X}_{d_1}^2}} = \frac{\frac{\mathcal{X}_{d_1}^2}{\mathcal{X}_{d_1}^2}}{1 + \frac{\mathcal{X}_{d_1}^2}{\mathcal{X}_{d_1}^2}} = \frac{\mathcal{X}_{d_1}^2}{\mathcal{X}_{d_1}^2 + \mathcal{X}_{d_2}^2} \quad (2.37)$$

qui est similaire à la formule (3.14) et nous permet de conclure que lorsque  $d = 1$  alors  $U \sim \text{Beta}\left(\frac{m}{2}, \frac{n}{2}\right)$ .

2. Quand le paramètre  $n = 1$  ( $Y$  est un vecteur  $1 \times d$ ). Nous savons que,

$$U = \frac{|A|}{|A+B|} = |A^{-1}B + I|^{-1} = |A^{-1}Y'Y + I|^{-1} \quad (2.38)$$

Considérons  $C = A^{-1}Y^T$  et  $D = Y$ . Il est clair que  $C$  est de dimension  $d \times 1$  puisque  $A$  est  $d \times d$  et  $Y^T$  est  $d \times 1$ .

**Lemme 2.5.1** *Soit  $C_{d \times n}$  et  $D_{d \times n}$  des matrices complexes. Alors,*

$$|CD + I_d| = |DC + I_n| \quad (2.39)$$

Nous obtenons

$$U = \frac{|A|}{|A+B|} = |A^{-1}B + I|^{-1} = |A^{-1}Y'Y + I|^{-1} = (YA^{-1}Y' + 1)^{-1} \quad (2.40)$$

Maintenant, trouvons la distribution de  $YA^{-1}Y'$ . On peut observer que

$$YA^{-1}Y' = \frac{YA^{-1}Y'}{YY'}YY'. \quad (2.41)$$

Donc,  $\frac{YY'}{YA^{-1}Y'}$  suit une loi  $\chi_{m-d+1}^2$  et  $YY' \rightsquigarrow \chi_d^2$ . Alors, il s'en suit que :

$$U = \left(1 + YA^{-1}Y'\right)^{-1} = \frac{1}{1 + \frac{\chi_d^2}{\chi_{m-d+1}^2}} \quad (2.42)$$

Par conséquent,  $U \sim \text{Beta}\left(\frac{m-d+1}{2}, \frac{d}{2}\right)$

3. Quand le paramètre  $d = 2$ . Nous avons

$$\frac{1 - \sqrt{\Lambda(2, m, n)}}{\sqrt{\Lambda(2, m, n)}} = \frac{n}{m - 1} F(2n, 2(m - 1)) \quad (2.43)$$

4. Quand le paramètre  $n = 2$ . Nous avons

$$\frac{1 - \sqrt{\Lambda(p, m, 2)}}{\sqrt{\Lambda(p, m, 2)}} = \frac{p}{m - p + 1} F(2p, 2(m - p + 1)) \quad (2.44)$$

## Conclusion

Dans le présent chapitre nous avons abordé la notion de variables aléatoire vectorielle où quelques caractéristiques associées à de telles variables ont été exposées. Cependant, une relecture minutieuse des expressions de certaines nous permet de rendre compte qu'elles ne sont que des notations et qui ne vérifient pas le conditionnement d'un calcul mathématique juste en particulier le calcul matricielle. A cet effet, il nous paraît qu'il est très important d'expliquer les changements qu'un utilisateur doit apporter aux expressions pour un calcul mathématique juste. Ceci fera l'objet du chapitre 3.

## Chapitre 3

# Exemple numérique illustratif d'analyse d'une *v.a.* vectorielle

### Introduction

Dans ce chapitre, nous allons présenter un exemple numérique sur des données réelles pour illustrer comment procéder pour réaliser une analyse statistiques préliminaire (moyenne, variance-covariance et coefficient de corrélation) d'une variable aléatoire vectorielle. Plus précisément, nous essayons de mettre au claire deux approches, à savoir : celle qui fais recours aux calculs basics destinés aux variables univariées et celle qui fais recours aux calculs matricielle.

### 3.1 Présentation des données

Les données dont on dispose sont procurées au niveau de la scolarité du département Mathématiques de l'université de Biskra. Elles présentent les notes détaillées (TD : Travaux Dirigés, C : Contrôle, Moy : Moyenne) obtenues par les 33 étu-

dians de Master 2 Mathématiques option Statistiques durant le premier semestre de l'année 2021/20222 dans les cinq modules de leurs formation (AC : Analyse convexe, PE&SO : Processus empirique & Statistiques d'ordre, SNP : Statistiques non Paramétrique, MR : Méthodologie de Recherche, S& MN : Simulation & Méthode Numérique). Leur présentation est illustrée dans la table 1 et leurs détails sont exposés dans l'Annexe B.

Il est claire que si on considère la variable  $X$  qui représente les notes d'un étudiants alors  $X$  est une variable aléatoire de dimension  $d = 13$  (voir la Table 3.1).

Analyse convexe			Processus Empirique et Statistiques d'ordre			Statistiques Non Paramétrique			Méthodologie de Recherche	Simulation et Méthodes Numériques		
Moy	Contrôle	TD	Moy	Contrôle	TD	Moy	Contrôle	TD	Moy	Moy	Contrôle	TD
$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$	$X_{10}$	$X_{11}$	$X_{12}$	$X_{13}$
$x_{1,1}$	$x_{2,1}$	$x_{3,1}$	$x_{4,1}$	$x_{5,1}$	$x_{6,1}$	$x_{7,1}$	$x_{8,1}$	$x_{9,1}$	$x_{10,1}$	$x_{11,1}$	$x_{12,1}$	$x_{13,1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_{1,j}$	$x_{2,j}$	$x_{3,j}$	$x_{4,j}$	$x_{5,j}$	$x_{6,j}$	$x_{7,j}$	$x_{8,j}$	$x_{9,j}$	$x_{10,j}$	$x_{11,j}$	$x_{12,j}$	$x_{13,j}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_{1,33}$	$x_{2,33}$	$x_{3,33}$	$x_{4,33}$	$x_{5,33}$	$x_{6,33}$	$x_{7,33}$	$x_{8,33}$	$x_{9,33}$	$x_{10,33}$	$x_{11,33}$	$x_{12,33}$	$x_{13,33}$

TABLE 3.1: Présentation des données

**Remarque 3.1.1** *Il est à souligner, bien qu'une brève discussion des résultats obtenus a été exposée l'objectif de l'application réalisée, n'est pas l'analyse statistique au sens propre du mot des données, mais plutôt c'est l'illustration des procédés des calculs dans le cas de variables aléatoires vectorielles.*

## 3.2 Estimation de la moyenne

L'estimateur empirique du vecteur moyen peut être calculer comme suite :

$$\bar{X} = \left( \frac{1}{33} \sum_{j=1}^{33} X_{1,j}; \frac{1}{33} \sum_{j=1}^{33} X_{2,j}; \dots ; \frac{1}{33} \sum_{j=1}^{33} X_{13,j} \right)'.$$

Cette dernière peut être implémenter sous Matlab tel qu'il est illustré dans la figure 3.1, son exécution sur nos données nous fournis le vecteur suivant :

$$\bar{X} = ( 9.8106 \quad 8.4394 \quad 11.1818 \quad 12.7239 \quad 11.5152 \quad 13.9318 \quad 11.5609 \quad 9.1970 \quad 13.9242 \quad 15.7273 \quad 8.9000 \quad 4.9394 \quad 12.8561 )'.$$

Si on procède au calcul de ce dernier estimateur par l'approche matricielle alors on doit calculer le produit matriciel suivant :

$$\bar{X} = X' \delta_n,$$

où  $X$  est l'échantillon (matrice des données) et  $\delta_n$  est un vecteur colonne de dimension  $n$  dont tous ses éléments vaux  $1/n$ .

D'après le vecteur  $X$  on constate que :

- La note moyenne la plus faible correspond à la variable  $X_{12}$  tandis que la plus élevée correspond à la variable  $X_{10}$ . Ceci signifie, que les étudiants ont obtenus, en moyenne, d'une part des notes faibles dans le contrôle du module "Simulation et Méthodes Numériques", d'autre part, les notes les plus élevée ont été obtenues dans le module Méthodologie de Recherche.
- Les notes en moyenne sont faible ( $\bar{X}_i < 10$ ) dans :
  1. Le contrôle et la moyenne du module Analyse convexe.
  2. Contrôle du module Statistiques Non Paramétrique.

```
function Moy = Moyenne(X,d,n)
%Programme de calcul de la Moyenne
for i=1:d
    Moy(i)=0;
end
for i=1:d
    for j=1:n
        Moy(i)=Moy(i)+X(j,i);
    end
    Moy(i)=Moy(i)/n;
end
```

FIGURE 3.1 – Code source Matlab pour le calcul du  $\bar{X}$ .

3. Le contrôle et la moyenne du module Simulation et Méthodes Numériques.

- On remarque également dans les quatre (04) modules Analyse convexe, Processus empirique et statistiques d'ordre, et Simulation et Méthode Numérique, que la propriété de la linéarité de l'espérance est vérifiée. En effet, on peut vérifier facilement dans ces 04 modules que :

$$E(Moy) = E(0.5C + 0.5TD) = 0.5E(C) + 0.5E(TD).$$

### 3.3 Estimation de la matrice Variance-Covariance

Pour obtenir l'estimateur sans biais,  $\hat{\Sigma}$ , de la matrice des variances-covariances il suffit de quantifier pour toutes les combinaisons  $(i, j)_{i=\overline{1:d}, j=\overline{1:d}}$ , ce qui suit :

$$\hat{\Sigma}_{ij} = \frac{1}{n-1} \sum_{k=1}^n (X_{ik} - \bar{X}_i) (X_{jk} - \bar{X}_j), \quad \text{pour } i = \overline{1:d}; j = \overline{1:d}. \quad (3.1)$$

avec  $X$  et  $\bar{X}_i (i = \overline{1:d})$  sont respectivement l'échantillon et la  $i$ ème composante du vecteur moyen (voir section 3.2).

La traduction de cette dernière expression en langage Matlab est présentée dans la 3.2. Par ailleurs, les calculs effectués sur nos données via l'expression (3.1), nous a fournis les résultats rangés dans la Table 3.2 :

La version matricielle de l'expression (3.1) se résume à :

$$\hat{\Sigma} = (X - \bar{X}\delta_{n-1})' (X - \bar{X}\delta_{n-1}),$$

où  $X$  est l'échantillon,  $\bar{X}$  est l'estimateur empirique du vecteur moyen et  $\delta_{n-1}$  est un vecteur colonne de dimension  $n$  dont toutes ses composantes égales à  $1/(n-1)$ .

4.3361	5.8944	2.7777	2.6916	3.6553	1.7269	0.6620	1.0268	0.2938	0.5053	2.1191	2.2186	2.0168
5.8944	9.6993	2.0895	3.7324	4.2041	3.2614	1.0294	1.3561	0.6984	0.6001	2.6323	3.3321	1.9285
2.7777	2.0895	3.4659	1.6508	3.1065	0.1925	0.2945	0.6974	-0.1108	0.4105	1.6059	1.1051	2.1051
2.6916	3.7324	1.6508	8.0422	9.1646	6.9174	-0.1143	-0.6549	0.4234	-0.3303	3.2435	4.0543	2.4321
3.6553	4.2041	3.1065	9.1646	12.2420	6.0831	-0.1749	-0.7375	0.3840	-0.6442	4.5020	5.7900	3.2132
1.7269	3.2614	0.1925	6.9174	6.0831	7.7511	-0.0534	-0.5721	0.4634	-0.0153	1.9843	2.3180	1.6500
0.6620	1.0294	0.2945	-0.1143	-0.1749	-0.0534	2.8781	4.1416	1.6126	0.2087	0.9094	0.6952	1.1224
1.0268	1.3561	0.6974	-0.6549	-0.7375	-0.5721	4.1416	6.7608	1.5193	0.3913	0.6966	0.0846	1.3085
0.2938	0.6984	-0.1108	0.4234	0.3840	0.4634	1.6126	1.5193	1.7050	0.0256	1.1200	1.3039	0.9341
0.5053	0.6001	0.4105	-0.3303	-0.6442	-0.0153	0.2087	0.3913	0.0256	2.3764	0.2588	0.0884	0.4283
2.1191	2.6323	1.6059	3.2435	4.5020	1.9843	0.9094	0.6966	1.1200	0.2588	4.6789	6.0178	3.3373
2.2186	3.3321	1.1051	4.0543	5.7900	2.3180	0.6952	0.0846	1.3039	0.0884	6.0178	10.1446	1.8875
2.0168	1.9285	2.1051	2.4321	3.2132	1.6500	1.1224	1.3085	0.9341	0.4283	3.3373	1.8875	4.7853

TABLE 3.2: Estimateur sans biais de la Matrice variance-covariance.

```

function Var_Cov = Covariance(X,d,n,Moy)
%Programme de calcul de la Matrice Variance-Covariance
for i=1:d
    for j=1:d
        Var_Cov(j,i)=0;
    end
end
for i=1:d
    for k=1:d
        for j=1:n
            Var_Cov(i,k)=Var_Cov(i,k)+(X(j,i)-Moy(i))*(X(j,k)-Moy(k));
        end
        Var_Cov(i,k)=Var_Cov(i,k)/(n-1);
    end
end
end

```

FIGURE 3.2 – Code source Matlab pour le calcul de la Matrice de  $\hat{\Sigma}_X$ .

D'après la matrice  $\hat{\Sigma}$ , on constate que :

- La plus grande variabilité est associée à la variable  $X_5$ , ceci signifie que les notes obtenues par les étudiants dans le contrôle du module Processus Empirique et Statistiques d'Ordre sont très diversifier et dispersé autour de la moyenne du module.
- La plus faible variabilité (variance) est associée à la variable  $X_9$ . Ceci signifie, que les étudiants ont pratiquement la même note dans le TD du module Statistiques Non Paramétrique qui est au voisinage de la moyenne du TD de ce module.
- La propriété

$$Var(\alpha X + \beta Y) = \alpha^2 Var(X) + \beta^2 Var(Y) + 2 \times \alpha \times \beta \times Cov(X, Y),$$

est vérifiée dans le cas des 04 modules qui ont la note de TD et du contrôle.

En effet, il est facile de vérifier numériquement sur la matrice  $\hat{\Sigma}$  que :

$$\begin{aligned} Var(Moy) &= Var(0.5C + 0.5TD) \\ &= (0.5)^2 Var(C) + (0.5)^2 Var(TD) + 2 \times 0.5 \times 0.5 \times Cov(C, TD). \end{aligned}$$

### 3.4 Estimation de la matrice de corrélation

Pour le calcul de l'estimateur de la matrice des coefficients de corrélations,  $\hat{\rho}$ , on doit disposer de la matrice  $\hat{\Sigma}$  préalablement. Par ailleurs, l'élément  $\hat{\rho}_{ij}$  de la matrice  $\hat{\rho}$  sera quantifier comme suit :

$$\hat{\rho}_{ij} = \frac{\hat{\Sigma}_{ij}}{\sqrt{\hat{\Sigma}_{ii} \hat{\Sigma}_{jj}}}, \quad \text{pour } i = \overline{1 : d}; j = \overline{1 : d}. \quad (3.2)$$

La traduction de cette dernière expression en langage Matlab est présentée dans la 3.3. Ainsi, la quantification de cette dernière expression appliquée à nos données, pour toute les couples  $(i, j)_{i=\overline{1:d}; j=\overline{1:d}}$ , nous a fourni la matrice présentée dans la table 3.3.

Pour la version matricielle de  $\hat{\rho}$  est définie comme suit :

$$\hat{\rho} = S^{-1}\hat{\Sigma}S^{-1},$$

avec  $S$  est une matrice diagonale de dimension  $d \times d$  dont l'élément  $S_{ii}$  correspond à  $\sqrt{\hat{\Sigma}_{ii}}$ . Il est claire que le fait que la matrice  $S$  est diagonale alors  $S^{-1}$  l'ai aussi de plus ses éléments ne sont rien que l'inverse des éléments de la matrice  $S$  c'est-à-dire  $(S^{-1})_{ii} = 1/S_{ii}$ .

D'après les résultats obtenus (voir table 3.3) on constate que :

- La note moyenne d'un module est fortement corrélée avec la note du contrôle et du TD du même module. Ceci est toute à fait logique et évident le fais que d'une part la moyenne d'un module est une combinaison linéaire de sa note de TD et sa note du contrôle

$$Moy = 0.5C + 0.5TD,$$

et d'autre par le coefficient de corrélation mesure l'association linéaire entre deux variables aléatoire.

- Le module Méthodologie de Recherche est linéairement indépendant avec le reste des modules. Tandis que le module Analyse convexe est linéairement dépendant (moyennement dépendant) avec le module Processus Empirique et Statistiques d'ordre et le module Simulation et Méthodes Numériques.

- La plus grande colinéarité est entre le module Simulation et Méthodes Numériques et le module Processus Empirique et Statistiques d'ordre
- Malgré la non existence d'un lien linéaire entre le module Processus Empirique et Statistiques d'Ordre avec les deux module Statistiques Non Paramétrique et Méthodologie de Recherche; la négativité du coefficient de corrélation entre ces modules ( $\hat{\rho}_{4,7} < 0$  et  $\hat{\rho}_{4,10} < 0$ ); signifie que globalement les étudiants qui ont une bonne moyenne dans le module Processus Empirique et Statistiques d'Ordre ont de faible moyenne dans le module Statistiques Non Paramétrique et le module Méthodologie de Recherche et le contraire est vrai.

1.0000	0.9089	0.7165	0.4558	0.5017	0.2979	0.1874	0.1896	0.1081	0.1574	0.4705	0.3345	0.4428
0.9089	1.0000	0.3604	0.4226	0.3858	0.3761	0.1948	0.1675	0.1717	0.1250	0.3907	0.3359	0.2831
0.7165	0.3604	1.0000	0.3127	0.4769	0.0371	0.0933	0.1441	-0.0456	0.1430	0.3988	0.1864	0.5169
0.4558	0.4226	0.3127	1.0000	0.9236	0.8761	-0.0238	-0.0888	0.1143	-0.0756	0.5288	0.4489	0.3920
0.5017	0.3858	0.4769	0.9236	1.0000	0.6245	-0.0295	-0.0811	0.0840	-0.1194	0.5949	0.5196	0.4198
0.2979	0.3761	0.0371	0.8761	0.6245	1.0000	-0.0113	-0.0790	0.1275	-0.0036	0.3295	0.2614	0.2709
0.1874	0.1948	0.0933	-0.0238	-0.0295	-0.0113	1.0000	0.9389	0.7280	0.0798	0.2478	0.1287	0.3024
0.1896	0.1675	0.1441	-0.0888	-0.0811	-0.0790	0.9389	1.0000	0.4475	0.0976	0.1239	0.0102	0.2301
0.1081	0.1717	-0.0456	0.1143	0.0840	0.1275	0.7280	0.4475	1.0000	0.0127	0.3965	0.3135	0.3270
0.1574	0.1250	0.1430	-0.0756	-0.1194	-0.0036	0.0798	0.0976	0.0127	1.0000	0.0776	0.0180	0.1270
0.4705	0.3907	0.3988	0.5288	0.5949	0.3295	0.2478	0.1239	0.3965	0.0776	1.0000	0.8735	0.7053
0.3345	0.3359	0.1864	0.4489	0.5196	0.2614	0.1287	0.0102	0.3135	0.0180	0.8735	1.0000	0.2709
0.4428	0.2831	0.5169	0.3920	0.4198	0.2709	0.3024	0.2301	0.3270	0.1270	0.7053	0.2709	1.0000

TABLE 3.3: Estimateur de la matrice de coefficients de corrélation.

## Conclusion

Dans ce chapitre nous avons mis l'accent sur deux approches de calcul pour l'évaluation de caractéristiques statistiques (statistiques descriptives) d'un vecteur aléatoire, à savoir : l'approche de calcul basique (dédiée principalement pour le

```
function R = Correlation(d,Var_Cov)
    % Programme de calcul de la matrice des
    % Coefficients de Corrélation
for i=1:d
    for j=1:d
        R(j,i)=Var_Cov(j,i)/sqrt(Var_Cov(i,i)*Var_Cov(j,j));
    end
end
```

FIGURE 3.3 – Code source Matlab pour le calcul de la Matrice de  $\hat{\rho}_X$ .

cas univarié) et l'approche matricielle (dédiée principalement aux cas multivarié).

Dans le cadre pratique, le choix entre ces deux approches dépend de l'outil de calcul dont on dispose (manuel, calculatrice, micro-ordinateur,...) et la maîtrise de cet outil. Cependant, dans le cadre théorique généralement il est plus intéressant d'opter pour l'approche matricielle qui peut nous éviter des calculs intermédiaire fastidieux. La meilleure justification réside dans l'exemple de la construction de la loi d'un vecteur aléatoire réel où le calcul d'intégrales multiple est inévitable si on opte pour l'approche basic hors l'approche matricielle peut nous fournir directement la loi de ce vecteur (voir Section 2.5 Chapitre 2).

# Conclusion générale

Le choix d'une méthode d'analyse de données est une première étape primordiale dans l'exploration des données. Le nombre de méthodes statistiques disponibles augmente de jour en jour avec l'avènement d'outil informatique. L'un des critères de choix de la méthode dépend de l'espace de la variable candidate pour l'analyse : unidimensionnel ou multidimensionnel. Ainsi, dans le présent mémoire l'objectif est de distinguer le cas de variables uni-variés (unidimensionnel) et le cas de variables vectorielles (multidimensionnel).

Pour répondre à notre objectif, dans un premier lieu nous avons exposé la notion de la variable aléatoire et les mesures auxquelles on s'intéresse pour caractériser cette variable (la densité, la distribution, la moyenne, la variance,...) où nous avons différencié le cas de variables quantitatives discrètes et quantitatives continue. Par la suite, avec la même démarche, que le cas d'univarié, nous avons essayé de mettre en évidence l'adaptation de ces caractéristiques au cas de variables vectorielles.

Enfin, un exemple numérique illustratif de la caractérisation d'une variable aléatoire vectorielle a été présenté où nous avons essayé de mettre l'accent sur deux approches de calcul pour l'évaluation de caractéristiques statistiques (statistiques descriptives) d'un vecteur aléatoire, à savoir : l'approche de calcul basique (dédiée principalement pour le cas univarié) et l'approche matricielle (dédiée principale-

ment aux cas multivarié). L'exemple concerne l'application de certaines notions théoriques introduites au début du mémoire, sur des données procurées au niveau de la scolarité du département Mathématiques de l'université de Biskra. Plus précisément, les données présentent les notes obtenues par les 33 étudiants de Master 2 Mathématiques option Statistiques durant le premier semestre de l'année 2021/20222 dans les cinq modules de leurs formation (Analyse convexe, Processus empirique & Statistiques d'ordre, Statistiques non Paramétrique, Méthodologie de Recherche, Simulation & Méthode Numérique).

**Il sera intéressant d'enrichir le présent travail par la considération dans le cadre de variables vectorielles :**

- L'estimation ponctuelle et par région de confiance.
- Les tests de conformité et tests d'homogénéité.
- ACP, AFG,...

# Bibliographie

- [1] B. Jourdain, Probabilités et statistique pour l'ingénieur, CERMICS, 10 Janvier 2018.
- [2] W.W. Hines, D.C. Montgomery, D.M. Goldsman and C.M. Borror. Probabilité et statistique pour ingénieurs 3 édition, 2017.
- [3] P.D. Lax and M.S. Terrell, Multivariable Calculus with Applications, 1st edition, Springer, 2017.
- [4] L. Lebart, A. Morineau et M. Piron, Statistique exploratoire multidimensionnelle, Dunod, 1995.
- [5] L. Lebart, A. Morineau et N. Tabard, Technique de la description statistique, Dunod, 1977.
- [6] N. Piskounov. Calcul Differentiel Et Integral, Tome 1, EDITION MIR T MOSCOU. 1998.
- [7] N. Piskounov. Calcul Differentiel Et Integral, Tome 2, EDITION MIR T MOSCOU. 1998.
- [8] G. Saporta. Probabilités, analyse des données et statistique; Editions Technique Paris, 2011.

# Annexe A : Rappel sur quelques notions de l'algèbre linéaire

Dans cette Annexe on va présenter quelques notions sur l'Algèbre linéaire (matrice, l'inverse d'une matrice, ...) qui sont très utiles pour la compréhension des différentes notations et les calculs matricielles introduite dans le cadre de variables aléatoires vectorielles.

**Définition 1 (Matrice)** *Une matrice  $X$  est un tableau rectangulaire de nombre. On dit que  $X$  est de dimension  $n \times d$ , si  $X$  a  $n$  lignes et  $d$  colonnes. Une telle matrice est représentée de la manière suivante :*

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & \vdots & \vdots \\ x_{n1} & \cdots & x_{nd} \end{pmatrix},$$

où  $x_{ij}$  est l'élément de la  $i^{\text{ème}}$  ligne et la  $j^{\text{ème}}$  colonne, on le note aussi  $(X)_{ij}$ .

**Définition 2 (Vecteur)** *Une vecteur colonne (respectivement, ligne)  $X$  de dimension  $d$  est une matrice de dimension  $d \times 1$  (respectivement,  $1 \times d$ ).*

**Définition 3 (La matrice identité)** *La matrice identité est une matrice de dimension  $n$ , notée  $I_n$ , telle que :*

$$x_{ij} = \begin{cases} 1, & \text{si } i = j; \\ 0, & \text{si } i \neq j. \end{cases},$$

*c'est-à-dire :*

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

**Définition 4 (La matrice unitaire)** *La matrice unitaire est une matrice de dimension  $n$ , notée  $\mathbf{1}_{n \times d}$ , telle que :*

$$\forall i, j \quad x_{ij} = 1,$$

*c'est-à-dire :*

$$\mathbf{1}_{n \times d} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \dots & 1 & 1 \end{pmatrix}$$

**Définition 5 (La matrice nulle)** *Une matrice est dite nulle si ses éléments sont tous nuls, c'est-à-dire :*

$$x_{ij} = 0 \quad \forall i = \overline{1, n} \quad \forall j = \overline{1, d}.$$

**Définition 6 (Transposée)** *La transposée de la matrice  $X$ , notée  $X'$ , est obtenue*

nue à partir de  $X$  en inter-changeant les lignes et les colonnes, c'est-à-dire :

$$(X)'_{ij} = (X)_{ji}.$$

Remarquer que si  $X$  est une matrice de taille  $n \times d$ , alors  $X'$  est de taille  $d \times n$ .

**Définition 7 (Symétrie)** On dit qu'une matrice est symétrique si :

$$X' = X.$$

**Définition 8 (La trace)** La trace d'une matrice carrée  $X$  d'ordre  $d$ , notée  $tr(X)$  est la somme de ses éléments diagonaux, c'est-à-dire :

$$tr(X) = \sum_{i=1}^d x_{ii}.$$

**Définition 9 (Le déterminant)** A toute matrice carrée d'ordre  $d$ , correspond un nombre réel, noté  $|X|$  ou  $\det(X)$ ; appelé le déterminant de  $X$ , qui peut être obtenu comme suit :

- Pour une matrice d'ordre 1 :

$$|x_{11}| = x_{11};$$

- Pour une matrice d'ordre 2 :

$$\begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} = x_{11}x_{22} - x_{21}x_{12};$$

- D'une façon générale, le déterminant d'une matrice carrée  $X$  peut être calculé

grâce aux notions de mineurs ou de cofacteurs.

Le mineur de l'élément  $x_{ij}$  d'une matrice  $X$  est le déterminant que l'on obtient en éliminant la  $i^{\text{ème}}$  ligne et la  $j^{\text{ème}}$  colonne de  $X$ .

Le cofacteur  $X_{ij}$  de  $x_{ij}$  est égal au mineur de  $x_{ij}$  multiplié par  $(-1)^{i+j}$ .

Le déterminant de  $X$  peut être calculé en effectuant la somme des produits des différents éléments d'une même colonne par leurs cofacteurs respectifs :

$$\begin{aligned} \det(X) &= \sum_{i=1}^d x_{ij} X_{ij} \text{ (pour tout } j); \\ &= \sum_{i=1}^d x_{ij} X_{ij} \text{ (pour tout } i). \end{aligned}$$

**Définition 10 (l'inverse d'une matrice)** Notons que la transposée de la matrice des cofacteurs  $X_{ij}$  est appelée "matrice adjointe" et désignée par  $\text{adj}(X)$ .

La matrice inverse de  $X$ ; notée  $X^{-1}$ ; peut être obtenue notamment en divisant la matrice adjointe de  $X$  par le  $\det(X)$  :

$$X^{-1} = \frac{\text{adj}(X)}{\det(X)}; \quad \det(X) \neq 0.$$

**Définition 11** Une matrice  $X$  de dimension  $n \times d$  est dite inversible (ou régulière ou encore non singulière) si :

1.  $X$  est une matrice carrée ( $n = d$ ),
2.  $\det(X) \neq 0$ ,
3. il existe une matrice  $Y$  de même taille que  $X$  avec laquelle les produits  $XY = YX = I_n$ .

**Lemme 1** Soit  $X$  une matrice de taille  $n \times n$ , les trois conditions suivantes sont équivalentes :

1.  $\det X = 0$ .
2. Il existe un vecteur colonne  $V \in \mathbb{R}^n, V \neq 0_{\mathbb{R}^n}$ , tel que :  $XV = 0$ , (un vecteur colonne est une matrice de taille  $n \times 1$ ).
3.  $X$  ne possède pas d'inverse, c'est-à-dire il n'existe pas de matrice  $X^{-1}$  telle que  $X^{-1}X = I_n$ .

**Lemme 2** Soit  $X$  une matrice de taille  $n \times n$ , alors les trois conditions suivantes sont équivalentes :

1.  $\det(X) \neq 0$ .
2. Si  $V \in \mathbb{R}^n$  est un vecteur colonne de taille  $n$ , alors  $XV = 0 \implies V = 0_{\mathbb{R}^n}$ .
3.  $X$  possède une matrice inverse, c'est-à-dire il existe une matrice  $X^{-1}$  telle que :  $X^{-1}X = XX^{-1} = I_n$ .

**Définition 12 (Les valeurs propres)** Les valeurs propres ou les valeurs caractéristiques d'une matrice carrée  $X$  d'ordre  $d$ , sont les  $d$  solutions de l'équation caractéristique :

$$\det(X - \lambda I_d) = 0.$$

**Remarque 3.4.1** Les valeurs propres peuvent être réelles ou complexes, positives, nulles ou négatives, distinctes ou confondues, mais on s'intéresse ici aux valeurs propres réelles, alors que leurs sommes est égale à la trace de la matrice  $X$  et leurs produits à son déterminant :

$$\sum_{i=1}^p \lambda_i = \text{tr}(X);$$

$$\prod_{i=1}^d \lambda_i = \det(X).$$

# Annexe B : Les données de l'application numérique

Analyse convexe			Processus Empirique et Statistiques d'ordre			Statistiques Non Paramétrique			Méthodologie de Recherche	Simulation et Méthodes Numériques		
Moy	Contrôle	TD	Moy	Contrôle	TD	Moy	Contrôle	TD	Moy	Moy	Contrôle	TD
10.25	10.5	10	15.5	13	18	13.75	11	16.5	18	9.5	6	13
10.25	10.5	10	13.75	10	17.5	10.5	7.5	13.5	15	6.38	0.5	12.25
11.5	9	14	16.5	16.5	16.5	9	6	12	18	9.25	3.5	15
9.25	8.5	10	12.5	10	15	10.25	7	13.5	15	12.38	9.25	15.5
10.75	5.5	16	11.88	12.5	11.25	11.25	10	12.5	15	7.5	1.5	13.5
7.25	4.5	10	5.75	3.5	8	15.5	15	16	15	7	0.5	13.5
16.25	15.5	17	16.75	17.5	16	15.13	14.75	15.5	16.5	13.13	7.5	18.75
7.25	4.5	10	11.5	10	13	10.75	9	12.5	18	8	4.25	11.75
9.25	8.5	10	10	8	12	9.5	6	13	16.5	7.75	5	10.5
10	8	12	16.88	17.5	16.25	11.25	10	12.5	12	8.75	6.5	11
7.5	5	10	13.5	11	16	10.25	8.5	12	16.5	6.75	2.5	11
11.25	11.5	11	11.5	11	12	13	11.5	14.5	16	10	7	13
11.25	9.5	13	9.25	8.5	10	10	7	13	16.5	9	7.5	10.5
9	7	11	12.75	8	17.5	12.75	11	14.5	18	8.38	2.75	14
8.75	7.5	10	13	11.5	14.5	12	11.5	12.5	15	9	3.5	14.5
11.25	10.5	12	8	6	10	11.75	11	12.5	15	7.13	4.25	10
11	10	12	12.75	12	13.5	10.5	6	15	15	10.38	6.25	14.5
8.75	7.5	10	12.25	10	14.5	11	8	14	16.5	6.5	2	11
10.25	8.5	12	12	10	14	10	6	14	16.5	8.38	3.25	13.5
8.75	7.5	10	13.5	12	15	11.25	9	13.5	16.5	11.25	8	14.5
7.25	4.5	10	8.5	7	10	14.5	13	16	16.5	8.13	4.75	11.5
7.25	4.5	10	11.25	10	12.5	11.75	10	13.5	16.5	7	3.25	10.75
10.25	10.5	10	11	11.5	10.5	12	9.5	14.5	16.5	6.88	2.25	11.5
14.25	15.5	13	17.75	17.5	18	14.38	13.75	15	16.5	11.38	8.75	14
7.5	5	10	15	14	16	11.25	7	15.5	11	9.25	4.5	14
7.5	5	10	13.5	13	14	10.5	7.5	13.5	13.5	8	4.5	11.5
12	14	10	13	11.5	14.5	9	6	12	15	6.75	3	10.5
8.75	5.5	12	10.75	12.5	9	10.5	7.5	13.5	16.5	12.13	9.25	15
10.75	11.5	10	13	11	15	13.5	12	15	15	10.25	4.75	15.75
12	10	14	17.13	18	16.25	11	7	15	16.5	12.88	8.75	17
9.25	8.5	10	10.25	7.5	13	11	8.5	13.5	15	5.88	1.25	10.5
7.25	4.5	10	12	11.5	12.5	10.25	7	13.5	15	5.88	1.25	10.5
10	10	10	17.25	16.5	18	12.5	9	16	15	12.88	15.25	10.5

Les données présentées dans la table ci-dessus, sont procurées au niveau de la scolarité du département Mathématiques de l'université de Biskra. Elles présentent les notes détaillées (TD : Travaux Dirigés, C : Contrôle, Moy : Moyenne) obtenues par les 33 étudiants de Master 2 Mathématiques option Statistiques durant le premier semestre de l'année 2021/20222 dans les cinq modules de leurs formation (Analyse convexe, Processus empirique & Statistiques d'ordre, Statistiques non Paramétrique, Méthodologie de Recherche, Simulation & Méthode Numérique).

## Résumé

Dans ce mémoire, nous avons abordé la notion de variable aléatoire vectorielle au sens probabiliste et statistique. Nous nous sommes particulièrement intéressés à la présentation de ce type de variable, au calcul de sa moyenne et de sa variance-covariance ainsi qu'à la mesure de la dépendance entre ses composantes (coefficient de corrélation). Enfin, à titre d'illustration des différentes notions abordées, un exemple sur des données réelles (les notes des étudiants de Master 2, option statistique, promotion 2021/2022) a été présenté.

**Mots-clés :** Variables aléatoires univariée, Variables aléatoires vectorielle, caractéristiques statistiques, corrélation.

## Abstract

In this work, we approached the notion of vector random variable in the probabilistic and statistical sense. We were particularly interested in the presentation of this type of variable, the calculation of its mean and its variance-covariance as well as the measurement of the dependence between its components (correlation coefficient). Finally, as an illustration of the different concepts discussed, an example on real data (grades of Master 2 Mathematics students, Statistics option, Class of 2021/2022) was presented.

**Keywords :** Univariate random variables, vector random variables, statistical characteristics, correlation.

## Résumé

Dans ce mémoire, nous avons abordé la notion de variable aléatoire vectorielle au sens probabiliste et statistique. Nous nous sommes particulièrement intéressés à la présentation de ce type de variable, au calcul de sa moyenne et de sa variance-covariance ainsi qu'à la mesure de la dépendance entre ses composantes (coefficient de corrélation). Enfin, à titre d'illustration des différentes notions abordées, un exemple sur des données réelles (les notes des étudiants de Master 2, option statistique, promotion 2021/2022) a été présenté.

**Mots clés :** Variables aléatoires univariée, Variables aléatoires vectorielle, Caractéristiques statistiques, Corrélation.

## Abstract

In this work, we approached the notion of vector random variable in the probabilistic and statistical sense. We were particularly interested in the presentation of this type of variable, the calculation of its mean and its variance-covariance as well as the measurement of the dependence between its components (correlation coefficient). Finally, as an illustration of the different concepts discussed, an example of real data (grades of Master 2 Mathematics students, Statistics option, Class of 2021/2022) was presented.

**Keywords:** Univariate random variables, Vector random variables, Statistical characteristics, Correlation.

## المخلص

في هذا العمل، اقتربنا من مفهوم المتغير العشوائي المتجه بالمعنى الاحتمالي والإحصائي. كنا مهتمين بشكل خاص بعرض هذا النوع من المتغيرات، وحساب متوسطه وتباين، وكذلك قياس التبعية بين مكوناته (معامل الارتباط). أخيراً، كتوضيح للمفاهيم المختلفة التي تمت مناقشتها، تم تقديم مثال لبيانات حقيقية (درجات طلاب ماستر 2 الرياضيات، تخصص الإحصاء، لسنة 2022/2021).

**الكلمات الرئيسية:** المتغيرات العشوائية أحادية المتغير، الشعاع العشوائي، الخصائص الإحصائية، الارتباط.