



République Algérienne Démocratique et Populaire  
Ministre de l'Enseignement Supérieur et de Recherche  
Scientifique



UNIVERSIT KASDI MERBAH OUARGLA  
FACULTE DES MATHMATIQUES ET SCIENCES DE  
LA MATIERE

## DEPARTEMENT DE MATHMATIQUES

MEMOIRE PRESENTE EN VUE DE L'OTENTION DU DIPLOME:

**MASTER en Mathématiques**

option : Probabilités Et Statistique

par : **KOUIDRI AFAF**

Thème

La distribution de khi-deux dècentrè  
et l'estimation de la moyenne d'une loi normal multidimensionnelle

dater : 2020/2021

Membres du Comitè d'Examen:

- Koudri Mohamed    KASDI MERBAH UNIVERSTY OUARGLA Prèsident
- Mansoul Brahim    KASDI MERBAH UNIVERSTY OUARGLA Examineur
- Mezouar Nadia    KASDI MERBAH UNIVERSTY OUARGLA Supervisor

---

*Dédicaces*

*D'abord, je remercie mon DIEU qui ma donnée le courage et la force pour accomplir ce modeste travail.*

*Je présente mes dédicaces de ce travail, le fruit des années d'études à : ceux qui ont consacré toute leur vie pour la réussite de leurs fils et leurs filles et qui ont les bougies allumant la réussite, pour leur bonté, pour leur générosité et encouragement.*

*Mes chers parents*

*Mon cher grand-père*

*Ma chère grande-mère*

*Mes chers frères : Kemal, Walid*

*Mes chères soeurs : Samira, Soumaia, Yamina, Bachra, Hind*

*Mes chères poussins de la famille : Sohail, Hussam, Wael, Djori, Lina, Djoud*

*Mes oncles et mes tantes*

*Mes cousins et mes cousines*

*Toutes Mes amies sans exception*

*Tous ceux qui m'ont aidé de près ou de loin à réaliser ce travail.*

## REMERCIEMENTS

*La réalisation de ce mémoire a été grâce au possible au concours de plusieurs personnes que je voudrais témoigner toute ma reconnaissance.*

*Je voudrais tout d'abord adresser tout ma gratitude à l'encadreur de ce projet "MEZOUAR Nadia" pour sa disponibilité et surtout ses judicieux conseil.*

*Nous présentons nos plus sincères remerciements au "Kouidri Mohamed", qui nous a honoré, en acceptant de présider le jury de soutenance.*

*Nous exprimons une profonde gratitude et respect au "Mansoul Brahim" qui a eu l'amabilité d'examiner notre mémoire.*

*Je remercie tous mes enseignantes et enseignants.*

*Et enfin à tous nos collègues de la promotion 2020 / 2021 et à tous ceux qui ont contribué , de près ou de loin, à la réalisation de ce travail.*

# Table des matières

<i>Introduction</i>	<b>4</b>
<b>1</b> <i>Préliminaire</i>	<b>5</b>
1.1 <i>Vecteur aléatoire</i> . . . . .	5
1.1.1 <i>Variable aléatoire continue</i> . . . . .	5
1.1.2 <i>Vecteur aléatoire</i> . . . . .	11
1.2 <i>Vecteur gaussien</i> . . . . .	17
1.2.1 <i>Variable gaussien</i> . . . . .	17
1.2.2 <i>Vecteur gaussien</i> . . . . .	17
1.3 <i>l'estimation paramétrique</i> . . . . .	20
1.3.1 <i>Estimateur du maximum de vraisemblance</i> . . . . .	27
<b>2</b> <i>La distribution de Khi-deux centré et décentré</i>	<b>29</b>
2.1 <i>La distribution de Khi-deux centré</i> . . . . .	29
2.1.1 <i>Loi Gamma</i> . . . . .	29
2.2 <i>La distribution de Khi-deux décentré</i> . . . . .	34
<b>3</b> <i>Application sur le calcule de risque</i>	<b>37</b>
3.1 <i>Calcul de risque quadratique de l'estimateur usuel</i> . . . . .	37
3.2 <i>Calcul de risque quadratique de l'estimateur de James Stien</i>	38
<i>Conclusion</i>	<b>42</b>

## *Introduction*

*Dans ce travail, nous nous intéressons à l'estimation de la moyenne  $\mu$  d'une loi normal multidimensionnelle par des estimateurs à rétrécisseurs déduites de l'estimateur usuel  $X$ .*

*En 1956, Stein a annoncé l'inadmissibilité de l'estimateur usuel  $X$ . En 1961, James et Stein ont introduit la classe des estimateurs de la forme*

$$\delta_a(X) = \left(1 - a \frac{1}{\|X\|^2}\right)X, \quad \forall a \in \mathbb{R}_+,$$

*qui est une classe des estimateurs biaisés, mais a un risque quadratique inférieur à celui de l'estimateur usuel  $X$ .*

*Notre travail se décompose en trois chapitres :*

*Dans le premier chapitre, nous rappelons quelques préliminaires sur les vecteurs aléatoire, vecteurs gaussien et les notions de base de l'estimation paramétrique.*

*Dans le deuxième chapitre, nous présentons la distribution de khi-deux centré, décentré et tous ses propriétés.*

*Le troisième chapitre est consacré aux applications, notamment le calcul de risque quadratique de l'estimateur usuel et l'estimateur de type James-Stien. Finalement, nous déduisons le meilleur estimateur dans la classe des estimateurs de type James-Stien  $\delta_a$ , il s'appelle estimateur primitif de James-Stien  $\delta_{\hat{a}}$ .*

— chapitre 1 —  
**Préliminaire**

*L'objectif de ce chapitre est de rappeler l'essentiel des définitions et des notations ainsi d'introduire des résultats fondamentaux qui seront utilisés plus tard.*

## 1.1 Vecteur aléatoire

### 1.1.1 Variable aléatoire continue

**Définition 1.1** *Soit  $(\Omega, \mathcal{A}, P)$  probabilité.  $X$  est dit une variable aléatoire réelle si elle est une application mesurable sur  $(\mathbb{R}, B(\mathbb{R}))$  c-à-d*

$$X^{-1}(B) \subset \mathcal{A}, \quad \forall B \in B(\mathbb{R}).$$

**Définition 1.2** *Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité. Une variable aléatoire*

$$X : \Omega \longrightarrow D_X \subset \mathbb{R}$$

*est dite variable aléatoire continue si son support  $D_X$  est un intervalle ou bien réunion d'intervalles.*

**Exemple 1.1** *On désigne par  $Y$  : "le temps d'attente avant l'arrivée du prochain bus", donc  $Y$  est une variable continue de support  $[0, 1\text{min}]$ .*

**Remarque 1.1** *Dans ce qui suit on vas travailler uniquement avec les variables aléatoires continues donc on vas dire brièvement une variable aléatoire*

### Densité de probabilité d'une variable aléatoire

**Définition 1.3** Une fonction réelle positive est dite densité de probabilité ou bien densité si

$$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

**Définition 1.4** La variable aléatoire réelle  $X$  suit la loi de densité  $f$  ou bien la variable aléatoire réelle  $X$  a pour densité la fonction  $f$  si pour tout intervalle  $I$  de  $\mathbb{R}$

$$P(X \in I) = P(X^{-1}(I)) = \int_I f(x)dx.$$

Dans ce cas le support de  $X$  est l'ensemble

$$D_X = \{x \in \mathbb{R} : f(x) > 0\}.$$

**Exemple 1.2** La fonction  $f(x) = \frac{1}{\pi(1+x^2)}$  est une densité de probabilité d'une variable aléatoire  $X$  de support  $D_X = \mathbb{R}$ . En effet :

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x) dx &= \int_{-\infty}^{+\infty} \frac{1}{\pi(1+x^2)} dx \\ &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{1}{(1+x^2)} dx \\ &= \frac{1}{\pi} \arctan(x) \Big|_{-\infty}^{+\infty} = \frac{1}{\pi} \left( \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) \right) = 1. \end{aligned}$$

### Valeur moyenne ou espérance mathématique d'une variable aléatoire

**Définition 1.5** Soit  $X$  une variable aléatoire réelle de support  $D_X$  et de fonction de densité  $f$ . On dit que  $X$  est  $P$ -intégrable si

$$\int_{-\infty}^{+\infty} |x| dP_X(x) = \int_{-\infty}^{+\infty} |x| f(x) dx = \int_{D_X} |x| f(x) dx < +\infty.$$

Dans ce cas on dit que  $X$  possède une espérance mathématique (notée  $E(X)$ ) ou bien une moyenne (notée  $\mu$ ) égale à

$$\mu = E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_{D_x} x f(x) dx.$$

Si  $\mu = E(X) = 0$  alors  $X$  est dite une variable aléatoire réelle centré.

**Exemple 1.3** Soit  $X$  a pour densité de probabilité  $f(x) = \lambda \exp^{-\lambda x} I_{[0,+\infty[}$  alors

$$\int_{-\infty}^{+\infty} |x| f(x) dx = \int_0^{+\infty} |x| \lambda \exp^{-\lambda x} dx = \int_0^{+\infty} \lambda x \exp^{-\lambda x} dx.$$

En utilisant l'intégration par partie une fois on trouve,

$$\int_{-\infty}^{+\infty} |x| f(x) dx = \frac{1}{\lambda}$$

d'où  $X$  possède une espérance mathématique

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \frac{1}{\lambda}.$$

**Propriété 1.1** Soit  $X$  une variable aléatoire réelle de support  $D_X$  de fonction de densité  $f$  d'espérance  $E(X)$ . Pour tous réels  $a, b$  on a

- i)  $E(aX + b) = aE(X) + b$ .
- ii)  $E(a) = a$ .



### Variance d'une variable aléatoire

**Définition 1.6** Soit  $X$  une variable aléatoire de support  $D_X$  et de fonction de densité  $f$ . On dit que  $X$  possède un moment d'ordre  $k$  noté  $E(X^k)$  ( $k \in \mathbb{N}^*$ ) si

$$E(|X|^k) = \int_{-\infty}^{+\infty} |x|^k f(x) dx = \int_{D_X} |x|^k f(x) dx < +\infty.$$

Dans ce cas

$$E(X^k) = \int_{-\infty}^{+\infty} x^k f(x) dx = \int_{D_X} x^k f(x) dx.$$

**Définition 1.7** Soit  $X$  une variable aléatoire de support  $D_X$  et de fonction de densité  $f$ . Si  $X$  possède un moment d'ordre deux (i.e.  $E(|X|^2) < +\infty$ ) alors la variance de la variable aléatoire  $X$ , notée  $\text{Var}(X)$  ou bien  $\sigma^2$  est donnée par :

$$\text{Var}(X) = E[(X - E(X))^2] \in \mathbb{R}^+.$$

La quantité  $\sqrt{\text{Var}(X)}$  (ou bien  $\sigma$ ) est appelée l'écart quadratique moyen ou l'écart type de la variable aléatoire  $X$ .

Lorsque  $\text{Var}(X) = 1$ , on dit que  $X$  est une variable aléatoire réduite.

**Proposition 1.1** Soit  $X$  une variable aléatoire de support  $D_X$  et de fonction de densité  $f$ , alors

1.  $\text{Var}(X) = E(X^2) - [E(X)]^2$ .
2.  $\text{Var}(aX + b) = a^2 \text{Var}(X)$ ,  $\forall a, b \in \mathbb{R}$ .
3.  $\text{Var}(a) = 0$ ,  $\forall a \in \mathbb{R}$ .

**Preuve 1.1** 1. On a

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

D'après de la Propriété 1.1 ceci nous donne

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

2. D'après ce qui précède on a

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

En utilisant la Propriété 1.1 on trouve

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

3. D'après la définition de la variance et ii) de la Propriété 1.1 on obtient

$$\text{Var}(a) = E(a - E(a))^2 = E(a - a)^2 = E(0) = 0.$$

### La fonction caractéristique d'une variable aléatoire

**Définition 1.8** Soit  $X$  une variable aléatoire de support  $D_X$  et de fonction de densité  $f$ . On appelle fonction caractéristique de la variable aléatoire  $X$  la fonction  $\varphi_X$  définie par :

$$\begin{aligned} \varphi_x : \mathbb{R} &\longrightarrow \mathbb{C} \\ t &\longmapsto \varphi_x(t) = E(e^{iXt}), \end{aligned}$$

donc

$$\varphi_x(t) = \int_{-\infty}^{+\infty} e^{ixt} f(x) dx = \int_{D_X} e^{ixt} f(x) dx$$

où  $i$  est le nombre complexe qui vérifie  $i^2 = -1$ .

**Proposition 1.2** Soit  $X$  une variable aléatoire réelle de fonction caractéristique  $\varphi_X$ . Alors

i) pour tout  $t \in \mathbb{R} : |\varphi_X(t)| \leq 1$ .

ii)  $\varphi_X(0) = 1$ .

iii) pour tout  $t \in \mathbb{R} : \varphi_X(-t) = \overline{\varphi_X(t)}$  (où  $\overline{\varphi_X(-t)}$  est le conjugué du nombre complexe  $\varphi_X(t)$ ).

iv) pour tout réels  $a, b$  :  $\varphi_{aX+b}(t) = e^{ibt} \varphi_X(at), \forall t \in \mathbb{R}$ .

v)  $\varphi_X$  est continue sur  $\mathbb{R}$ .

**Théorème 1.1** Si  $\varphi_X$  est la fonction caractéristique d'une variable aléatoire  $X$  et si  $\varphi_X$  est intégrable sur  $\mathbb{R}$  i.e

$$\int_{-\infty}^{+\infty} |\varphi_X| dx < +\infty$$

alors la variable aléatoire  $X$  admet une densité de probabilité définie par

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ixt} \varphi_X(t) dt, \forall x \in \mathbb{R}.$$

**Exemple 1.4** Si  $X$  est une variable aléatoire réelle de fonction caractéristique  $\varphi_X(t) = e^{-|t|}$ , alors sa fonction de densité est

$$f_X(x) = \frac{1}{\pi(1+x^2)}, \forall x \in \mathbb{R}.$$

En effet

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ixt} \varphi_X(t) dt \\ &= \frac{1}{2\pi} \int e^{-ixt} e^{-|t|} dt \\ &= \frac{1}{2\pi} \left[ \int_{-\infty}^0 e^{-ixt+t} dt + \int_0^{+\infty} e^{-ixt-t} dt \right] \\ &= \frac{1}{2\pi} \left[ \frac{e^{(1-ix)t}}{(1-ix)} \right]_{-\infty}^0 + \frac{-e^{-(1+ix)t}}{(1+ix)} \right]_{-\infty}^0 \\ &= \frac{1}{2\pi} \left[ \frac{1}{(1-ix)} + \frac{1}{(1+ix)} \right] \\ &= \frac{1}{2\pi} \frac{(1+ix) + (1-ix)}{(1-ix)(1+ix)} = \frac{1}{\pi(1+x^2)}. \end{aligned}$$

### 1.1.2 Vecteur aléatoire

**Définition 1.9** Un vecteur aléatoire  $X = \begin{pmatrix} X_1 \\ \cdot \\ \cdot \\ \cdot \\ X_k \end{pmatrix}$  est une application

mesurable sur  $(\mathbb{R}^k, B(\mathbb{R}^k))$  où toutes les coordonnées du vecteur  $X$  sont des variables aléatoires réelles.

**Définition 1.10** Soit  $X$  un vecteur aléatoire tel que, pour tout  $j \in \{1, \dots, k\}$ , la variable aléatoire  $X_j$  est  $P$ -intégrable. L'espérance de  $X$  est le vecteur

$$E(X) = \begin{pmatrix} E(X_1) \\ \cdot \\ \cdot \\ \cdot \\ E(X_k) \end{pmatrix}.$$

On dit que le vecteur aléatoire  $X$  est centré si  $E(X)$  est le vecteur nul dans  $\mathbb{R}^k$ .

**Définition 1.11** Soient  $X$  et  $Y$  deux variables aléatoires, on dit que  $X$  et  $Y$  sont indépendants si et seulement si

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

**Définition 1.12** La covariance de deux variables aléatoires  $X$  et  $Y$ , notée  $Cov(X, Y)$ , égale à

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

#### Proposition 1.3

- 1)  $Cov(X, X) = Var(X)$
- 2)  $Cov(X, Y) = E(XY) - E(X)E(Y)$
- 3) Si  $X$  et  $Y$  sont indépendants  $\implies Cov(X, Y) = 0$

**Preuve 1.2**

1)

$$\begin{aligned} \text{Cov}(X, X) &= E[(X - E(X))(X - E(X))] \\ &= E[(X - E(X))^2] \\ &= \text{Var}(X). \end{aligned}$$

2)

$$\begin{aligned} \text{Cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E[(XY) - XE(Y) - YE(X) + E(X)E(Y)] \\ &= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

3) D'après 2) on a

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

et comme  $X$  et  $Y$  sont indépendants alors  $E(XY) = E(X)E(Y)$  ce qui nous donne

$$\text{Cov}(X, Y) = E(X)E(Y) - E(X)E(Y) = 0.$$

**Remarque 1.2** Le sens réciproque de la propriété 3) est faux. Voici un contre exemple.

**Exemple 1.5** • Soit  $X$  une variable aléatoire centrée (i.e.  $E(X)=0$ ) et soit  $\varepsilon$  une variable aléatoire de loi

$$P(\varepsilon = 1) = P(\varepsilon = -1) = \frac{1}{2}.$$

Supposons que  $X$  et  $\varepsilon$  soient des variables aléatoires indépendants. On pose  $Y = \varepsilon X$ . On a  $\text{Cov}(X, Y) = 0$ .

En effet la covariance est égale à

$$\begin{aligned} \text{Cov}(X, Y) &= E(XY) - E(X)E(Y) = E(XY) \quad (\text{car } E(X) = 0) \\ &= E(X^2\varepsilon) = E(X^2)E(\varepsilon) \\ &= 0 \quad (\text{car } E(\varepsilon) = 0). \end{aligned}$$

Or  $|Y| = |\varepsilon||X| = |X|$  donc les variables aléatoires  $X$  et  $Y$  ne sont pas indépendants.

### La variance d'un vecteur aléatoire

**Définition 1.13** Soit  $X$  un vecteur aléatoire dont toutes les coordonnées sont des variables réelles  $L^2$  intégrable. La variance de  $X$  est une matrice (noté  $\text{Var}(X)$  ou  $\Sigma_X$ ) constituée des éléments

$$\begin{aligned} \text{Cov}(X_i, X_j) &= E((X_i - E(X_i))(X_j - E(X_j))) \\ &= E(X_i X_j) - E(X_i)E(X_j), \forall 1 \leq i, j \leq k. \end{aligned}$$

La matrice  $\Sigma_X$  est appelée matrice de covariance de  $X$ . Elle peut aussi s'exprimer de la façon suivante :

$$\Sigma_X = E(XX^t) - E(X)E(X)^t.$$

**Exemple 1.6** Soit  $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$  un vecteur aléatoire.

$$\text{Var}(X) = \Sigma_X = \begin{pmatrix} \text{Cov}(X_1, X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_2, X_1) & \text{Cov}(X_2, X_2) \end{pmatrix}.$$

D'après 1) de la proposition 1.3 ceci nous donne

$$\text{Var}(X) = \Sigma_X = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) \end{pmatrix}$$

et si les variables sont indépendants alors d'après 2) de la proposition 1.3, on a

$$\text{Var}(X) = \Sigma_X = \begin{pmatrix} \text{Var}(X_1) & 0 \\ 0 & \text{Var}(X_2) \end{pmatrix}$$

**Proposition 1.4** Soit  $X = (X_1, \dots, X_k)^t$  un vecteur aléatoire où les  $X_i$  sont indépendants alors  $\text{Var}(X)$  est diagonale et égale à

$$\text{Var}(X) = \Sigma_X = \begin{pmatrix} \text{Var}(X_1) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \text{Var}(X_k) \end{pmatrix}$$

**Propriété 1.2** • Soit  $X$  un vecteur aléatoire et soient  $B \in \mathbb{R}^p$  et  $A$  une matrice  $(p \times k)$ . On pose  $Y = AX + B$  une transformation affine. Le vecteur aléatoire  $Y$  vérifie

- 1)  $E(Y) = AE(X) + B$ .  
 2)  $Var(Y) = AVar(X)A^t$ .

**Définition 1.14** Soient  $X$  et  $Y$  deux vecteurs aléatoires  $L^2$  intégrable. On suppose que  $X$  et  $Y$  sont respectivement à valeur dans  $\mathbb{R}^k$  et  $\mathbb{R}^p$ . La covariance entre les vecteurs  $X$  et  $Y$  est la matrice constituée des éléments  $Cov(X_i, Y_j)$  pour  $i = 1, \dots, k$  et  $j = 1, \dots, p$ .

**Proposition 1.5** Soient  $X$  et  $Y$  deux vecteurs aléatoires.

i) Si  $X$  et  $Y$  sont indépendants alors  $Cov(X, Y) = 0$ . la réciproque est fausse

ii) Posons  $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$ . Sa matrice de covariance est donnée par

$$\begin{pmatrix} Var(X) & Cov(X, Y) \\ Cov(Y, X) & Var(Y) \end{pmatrix}$$

iii) Si  $X$  et  $Y$  sont des vecteurs de même dimension alors

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

**Définition 1.15** Soit  $X$  un vecteur aléatoire. Sa fonction caractéristique est définie par

$$\Phi_X(u) = E(e^{iu^t X}) = E(e^{i \sum_{j=1}^k u_j X_j}) \quad \forall u \in \mathbb{R}^k.$$

le résultat suivant assure que la loi d'un vecteur aléatoire est déterminée par sa fonction caractéristique.

**Théorème 1.2** Soient  $X$  et  $Y$  deux vecteurs aléatoires. On note  $\Phi_X$  et  $\Phi_Y$  leurs fonctions caractéristiques. Si fonction  $\Phi_X = \Phi_Y$  alors  $X$  et  $Y$  admettent la même loi.

**Propriété 1.3** Soit  $X = (X_1, \dots, X_k)$  un vecteur aléatoire .

i) Si  $Y$  est une transformation affine du vecteur  $X$  c'est à dire  $Y = AX + B$  avec  $B \in \mathbb{R}^p$  et  $A$  une matrice ( $p \times k$ ), alors les fonctions caractéristiques de  $X$  et  $Y$  sont liées par la relation suivante

$$\Phi_Y(v) = e^{iv^t B} \Phi_X(A^t v)$$

pour tout  $v \in \mathbb{R}^p$ .

ii) Les variables aléatoires  $(X_1, \dots, X_k)$  sont indépendants si et seulement si

$$\Phi_X(u) = \prod_{j=1}^k \Phi_{X_j}(u_j) \quad \forall u = (u_1, \dots, u_k).$$

### Preuve 1.3

i)

$$\begin{aligned} \Phi_Y(v) &= E(e^{iv^t Y}) \\ &= E\left(e^{iv^t (AX+B)}\right) \\ &= e^{iv^t B} E(e^{iv^t (AX)}) \\ &= e^{iv^t B} \Phi_X(A^t v). \end{aligned}$$

ii) [ $\implies$ ] Supposons que les coordonnées de  $X$  sont indépendants. Pour tout  $u \in \mathbb{R}^k$ , on a

$$\Phi_X(u) := E\left(e^{i \sum_{j=1}^k u_j x_j}\right) = \prod_{j=1}^k E\left(e^{i u_j X_j}\right) = \prod_{j=1}^k \Phi_{X_j}(u_j)$$

[ $\impliedby$ ] Réciproquement, pour tout  $u \in \mathbb{R}^k$ , on a

$$\Phi_X(u) = \prod_{j=1}^k \Phi_{X_j}(u_j).$$

Alors, on a, pour tout  $u \in \mathbb{R}^k$

$$\int e^{i \sum_{j=1}^k u_j x_j} dP_{X_1, \dots, X_k}(x_1, \dots, x_k) = \prod_{j=1}^k \int e^{i u_j x_j} dP_{X_j}(x_j)$$



$$= \int e^{i \sum_{j=1}^k u_j X_j} \prod_{j=1}^k dP_{X_j}(x_j).$$

*L'unicité de la fonction caractéristique (voir Théorème 1.2) implique que la loi du vecteur  $X$  est la mesure produit, d'où l'indépendance des coordonnées du  $X$ .*

## 1.2 Vecteur gaussien

### 1.2.1 Variable gaussien

**Définition 1.16** Soit  $X$  une variable aléatoire réelle. On dit que  $X$  est une variable aléatoire gaussienne de paramètres  $(\mu, \sigma^2)$  avec  $\mu \in \mathbb{R}$  et  $\sigma \in \mathbb{R}_+^*$  (on note  $X \sim N(\mu, \sigma^2)$ ) si et seulement si  $X$  admet pour densité

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

Si  $X \sim N(0, 1)$  alors  $X$  est dite variable gaussienne centré réduit ou standard.

**Propriété 1.4** Soit  $X \sim N(\mu, \sigma^2)$ . On a

- l'espérance de  $X$  :  $E(X) = \mu$ .
- la variance de  $X$  :  $Var(X) = \sigma^2$ .
- la fonction caractéristique de  $X$  est égale à

$$\phi_X(t) = E(e^{itX}) = e^{it\mu} e^{-\frac{t^2\sigma^2}{2}}.$$

- Si  $X \sim N(\mu, \sigma^2)$  alors  $\frac{X - \mu}{\sigma} \sim N(0, 1)$ .
- Plus généralement, on a

$$aX + b \sim N(a\mu + b, a^2\sigma^2)$$

pour tout  $(a, b) \in \mathbb{R}^2$ .

### 1.2.2 Vecteur gaussien

**Définition 1.17** Un vecteur aléatoire  $X$  à valeurs dans  $\mathbb{R}^k$  est un vecteur gaussien si et seulement si toute combinaison linéaire de ses coordonnées est une variable aléatoire gaussienne i.e.,  $\forall \lambda_j \in \mathbb{R}$ ,  $\sum_{j=1}^k \lambda_j X_j$  suit une loi gaussienne.

### Propriétés 1.1

1) Si  $X$  est vecteur gaussien alors pour toute partie  $\{i_1, \dots, i_p\}$  de  $\{1, \dots, k\}$ , le vecteur  $(X_{i_1}, \dots, X_{i_p})$  est gaussien.

2) Soient  $A$  une matrice de dimension  $p \times k$  et  $X$  un vecteur gaussien,  $X \sim N_k(\mu, \Sigma)$ . Le vecteur  $AX$  est un vecteur gaussien de dimension  $p$  de plus, on a  $AX \sim N_p(A\mu, A\Sigma A^t)$ .

**Théorème 1.3** Un vecteur aléatoire  $X$  à valeur dans  $\mathbb{R}^k$  est un vecteur gaussien si et seulement si  $X \in L^2$  et il admet pour fonction caractéristique

$$\Phi_X(u) = e^{iu^t\mu - \frac{1}{2}u^t\Sigma u} \quad \forall u \in \mathbb{R}^k$$

avec  $\mu = E(X)$  et  $\Sigma = \text{Var}(X)$ .

**Preuve 1.4**  $[\Rightarrow]$  Par définition on a

$$\Phi_X(\mu) = E(e^{iu^tX}) = E(e^{iY}) = \Phi_Y(1)$$

où  $Y = \sum_{i=1}^k u_i X_i$ . La variable aléatoire  $Y$  est une variable aléatoire gaussienne puisque c'est une combinaison linéaire des coordonnées du vecteur gaussien  $X$ . On a donc

$$Q_Y(1) = e^{iE(Y)} e^{-\frac{\text{Var}(Y)}{2}}$$

avec

$$E(Y) = E\left(\sum_{i=1}^k u_i X_i\right) = \sum_{i=1}^k u_i E(X_i) = u^t E(X) = u^t \mu.$$

et

$$\begin{aligned} \text{Var}(Y) &= \text{Var}\left(\sum_{i=1}^k u_i X_i\right) = \sum_{i=1}^k \sum_{j=1}^k \text{Cov}(u_i X_i, u_j X_j) \\ &= \sum_{i=1}^k \sum_{j=1}^k u_i u_j \text{Cov}(X_i, X_j) = u^t \text{Var}(X) u = u^t \Sigma u. \end{aligned}$$

Ceci prouve que la fonction caractéristique de la forme annoncée.

[ $\Leftarrow$ ] Soit  $\Lambda^t = (\lambda_1, \dots, \lambda_k)$ . On pose  $Y = \sum_{i=1}^k \lambda_i X_i = \Lambda^t X$ . Montrons que  $Y$  est une variable aléatoire gaussienne. Pour tout  $u \in \mathbb{R}$ , la fonction caractéristique de  $Y$  est donnée par

$$\begin{aligned} \Phi_Y(u) &= E \left( e^{iu\Lambda^t X} \right) = \Phi_X(u\Lambda) \\ &= e^{i(u\Lambda)^t \mu - \frac{1}{2} (u\Lambda)^t \Sigma u\Lambda} \\ &= e^{iu\Lambda^t \mu - \frac{1}{2} u^2 \Lambda^t \Sigma \Lambda}. \end{aligned}$$

On reconnaît la fonction caractéristique de la loi gaussienne  $N(\Lambda^t \mu, \Lambda^t \Sigma \Lambda)$ . Donc  $Y$  est bien une variable aléatoire gaussienne.

### vecteur gaussien et indépendance

**Proposition 1.6** Soit  $X$  un vecteur gaussien de dimension  $k$ , de moyenne  $\mu$  et de covariance  $\Sigma$ . Les variables aléatoires  $(X_1, \dots, X_k)$  sont indépendants si et seulement si la matrice  $\Sigma$  est diagonale.

**Preuve 1.5** Notons  $\sigma_{i,j}$  les éléments de la matrice  $\sigma$ .

D'après ii) de la propriété 1.3, les variables aléatoires  $(X_1, \dots, X_k)$  sont indépendants si et seulement si.

$$\Phi_X(u) = \prod_{j=1}^k \Phi_{X_j}(u_j) \quad \forall u \in \mathbb{R}^k. \quad (1.1)$$

Or la fonction caractéristique de  $X$  s'écrit

$$\Phi_X(u) = e^{iu^t \mu - \frac{1}{2} u^t \Sigma u}$$

et le terme de droite dans (1.1) s'exprime comme

$$\prod_{j=1}^k \Phi_{X_j}(u_j) = \prod_{j=1}^k e^{i\mu_j u_j - \frac{1}{2} u_j^2 \sigma_{j,j}} = e^{iu^t \mu} e^{-\frac{1}{2} \sum_{j=1}^k u_j^2 \sigma_{j,j}} \quad (1.2)$$

De plus,  $\sum_{j=1}^k u_j^2 \sigma_{j,j} = u^t D u$  en prenant  $D$  la matrice diagonale construite à partir de la diagonale de  $\Sigma$  i.e

$$D = \begin{pmatrix} \sigma_{1,1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \sigma_{k,k} \end{pmatrix}$$

L'égalité (1.1) est donc équivalente à  $\Sigma = D$ . D'où l'indépendance des coordonnées de  $X$  si et seulement si  $\Sigma = D$ .

**Proposition 1.7** Soit  $X$  un vecteur gaussien écrit de la forme  $(Y, Z)$  avec  $Y \in \mathbb{R}^p$  et  $Z \in \mathbb{R}^q$ . Les vecteurs  $Y$  et  $Z$  sont indépendants si et seulement si la matrice de covariance de  $X$  est diagonale par blocs c'est à dire

$$\begin{pmatrix} A & 0_{p,q} \\ 0_{q,p} & B \end{pmatrix}$$

avec  $A$  une matrice de dimension  $p \times p$  et  $B$  une matrice de dimension  $q \times q$ .

**Densité d'un vecteur gaussien**

**Définition 1.18** Soit  $X \sim N_k(\mu, \sigma)$  un vecteur gaussien alors sa densité égale à

$$f(x) = \frac{1}{(2\pi)^{\frac{k}{2}} \det(\Sigma)^{\frac{1}{2}}} e^{(-\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu))}.$$

### 1.3 l'estimation paramétrique

On commence cette partie par un exemple de l'estimation paramétrique réel pour expliquer les notions posées.

**Exemple 1.7** L'estimation du taux moyenne de cholestérol mesuré sur 200 femmes de 50 ans en Algérie (48 Wilayas).

Le statisticien, pour répondre à cette question, doit avoir :

- Un modèle statistique et  $n$ -échantillon
- Une estimation paramétrique
- Puis il cherche une meilleur estimation.

### Modèle statistique et $n$ -échantillon

**Définition 1.19** On appelle modèle statistique, la donnée d'un espace des observations  $\Omega$ , d'une tribu  $\mathcal{A}$  d'événements sur  $\Omega$  et d'une famille de probabilités  $P$  sur l'espace probabilisable  $(\Omega, \mathcal{A})$ , on le note  $(\Omega, \mathcal{A}, P)$ .

**Définition 1.20** On dit qu'un modèle statistique est paramétrique s'il existe un entier  $d$  et un sous ensemble  $\Theta$  de  $\mathbb{R}^d$  tels que la famille de probabilité  $P$  puisse être paramétrée par  $\theta$  tels que l'application :

$$\begin{aligned} \Theta &\longrightarrow P \\ \theta &\longrightarrow P_\theta \end{aligned}$$

est surjective . On note :  $P = \{P_\theta / \theta \in \Theta \subset \mathbb{R}^d\}$ . Dans le cas contraire, on dit que le modèle est non paramétrique.

**Remarque 1.3** le paramétrage n'est par forcément unique.

**Exemple 1.8** Dans l'exemple 1.7, l'espace  $\Omega$  est l'Algerie,  $\mathcal{A}$  : Wilayas et  $d = 200$  femmes.

**Définition 1.21** • On appelle  $n$ -échantillon de la loi  $P_\theta$ , la donnée d'un vecteur  $(X_1, X_2, \dots, X_n)$  constituée de  $n$  variables aléatoires indépendantes et identiquement distribuées (i.i.d) de loi  $P_\theta$ . Le nombre  $n$  est appelé taille de l'échantillon.

• On appelle modèle d'échantillonnage, le modèle  $(\Omega^n, \mathcal{A}^{\otimes n}, P_\theta^{\otimes n})$ , où  $\mathcal{A}^{\otimes n}$  est la tribu produit ( engendrée par les pavés ) sur  $\Omega^n$  et  $P_\theta^{\otimes n} = P_\theta \otimes \dots \otimes P_\theta$  est la probabilité produit sur  $(\Omega^n, \mathcal{A}^{\otimes n})$  qui est la loi de vecteur  $(X_1, X_2, \dots, X_n)$ .

En effet :

Un modèle d'échantillonnage est donc un modèle statistique particulier , où l'espace des observations est de la forme  $\Omega^n$  muni de la tribu produit classique et de la probabilités de la forme  $P_\theta^{\otimes n}$ .

**Exemple 1.9** Dans l'exemple 1.7 :

- $n = 48$  Wilayas.
- $X_i$  : le taux de colestérol de la  $i^{\text{eme}}$  femme.
- $\omega_j$  : la  $j^{\text{eme}}$  Wilaya.
- $X_i(\omega_j)$  : le taux de colestérol de la  $i^{\text{eme}}$  femme dans la  $j^{\text{eme}}$  Wilaya.

**Définition 1.22** Soit  $(\Omega^n, \mathcal{A}^{\otimes n}, P_\theta^{\otimes n} = P_\theta \otimes \dots \otimes P_\theta)$  un modèle d'échantillonnage. On appelle statistique, la variable aléatoire  $T(X) = T(X_1, \dots, X_n)$  où  $T$  est une fonction mesurable de  $(\Omega^n, \mathcal{A}^{\otimes n}, P_\theta^{\otimes n})$  vers un espace probabilisable  $(F, \mathcal{F})$

$$\begin{aligned} T : \Omega^n &\longrightarrow F \\ X &\longrightarrow T(X) \end{aligned}$$

### Estimation paramétrique

L'estimation paramétrique consiste à estimer ou évaluer un ou plusieurs paramètres inconnues émanant de  $X$  à partir des données  $(x_1, x_2, \dots, x_n)$  qui sont les réalisations de la variable aléatoire  $X$  (i. e.  $x_i = X_i(\omega)$ ).

**Définition 1.23** Soit  $g$  une application mesurable de  $\Theta$  dans  $\mathbb{R}^k$ . Un estimateur de  $g(\theta)$  est une fonction qui fait correspondre à chaque réalisation possible de  $n$ -échantillon  $X_1, X_2, \dots, X_n$  de loi  $P_\theta$ , la valeur  $\delta_n$  que l'on nomme estimation

$$\delta_n = h(X_1, X_2, \dots, X_n)$$

où

$$\begin{aligned} h : \Omega^n &\longrightarrow g(\Theta) \\ (X_1, X_2, \dots, X_n) &\longmapsto h(X_1, X_2, \dots, X_n) \end{aligned}$$

**Remarque 1.4** Si on arrive à estimer le  $\theta$  alors on peut avoir l'estimation de n'importe quelle valeur  $g(\theta)$  donc on peut considérer  $g(\theta) = \theta$  et on peut généraliser la valeur ailleurs.

**Définition 1.24** • Soit  $\delta_n$  est un estimateur. Le biais de  $\delta_n$  est une mesure de l'écart moyen entre  $\delta_n$  et  $\theta$  et on le note  $B(\delta_n)$  i. e.,

$$B(\delta_n) = E(\delta_n) - \theta$$

•  $\delta_n$  est un estimateur sans biais (ou non biaisé) du paramètre  $\theta$  si

$$E(\delta_n) = \theta.$$

**Définition 1.25** Un estimateur  $\delta_n$  est asymptotiquement sans biais pour le paramètre  $\theta$  si

$$\lim_{n \rightarrow +\infty} E(\delta_n) = \theta.$$

**Définition 1.26** On appelle variance d'un estimateur  $\delta_n$  la valeur

$$\text{Var}(\delta_n) = E(\delta_n - E(\delta_n))^2.$$

**Définition 1.27** On appelle risque quadratique d'un estimateur  $\delta_n$  par rapport à  $\theta$  la quantité

$$R(\delta_n, \theta) = E(\delta_n - \theta)^2.$$

**Remarque 1.5** Si  $\delta$  est un estimateur sans biais alors

$$R(\delta_n, \theta) = E(\delta_n - E(\delta_n))^2 = \text{Var}(\delta_n).$$

**Moyenne, variance, moments empiriques**

**Définition 1.28** Soit  $X_1, X_2, \dots, X_n$  un  $n$ -échantillon d'une loi connue  $P_\theta$ , on appelle moyenne empirique la statistique  $\overline{X}_n$  définie par.

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

**Proposition 1.8** Soient  $m_\theta$  et  $\sigma_\theta^2$  respectivement la moyenne et la variance de l'échantillon  $X_1, X_2, \dots, X_n$  alors

$$E(\overline{X}_n) = m_\theta \text{ et } \text{Var}(\overline{X}_n) = \frac{\sigma_\theta^2}{n}.$$



**Preuve 1.6**

$$\begin{aligned}
 \bullet E(\overline{X}_n) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\
 &= \frac{1}{n} E\left(\sum_{i=1}^n X_i\right) \\
 &= \frac{1}{n} \sum_{i=1}^n E(X_i) \\
 &= \frac{nm_\theta}{n} = m_\theta.
 \end{aligned}$$

D'où  $\overline{X}_n$  est un estimateur sans biais de  $m_\theta$ .

$$\begin{aligned}
 \bullet \text{Var}(\overline{X}_n) &= \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\
 &= \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n X_i\right) \\
 &= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) \\
 &= \frac{n\sigma_\theta^2}{n^2} = \frac{\sigma_\theta^2}{n}
 \end{aligned}$$

d'où  $\text{Var}(\overline{X}_n) = \frac{\sigma_\theta^2}{n}$ .

**Définition 1.29** Soit  $X_1, \dots, X_n$  un  $n$ -échantillon d'une loi connue  $P_\theta$ , on appelle variance empirique la statistique  $S_n^2$  définie par

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2.$$

**Proposition 1.9** La moyenne de la loi de la variance empirique de l'échantillon  $X_1, X_2, \dots, X_n$  est

$$E(S_n^2) = \frac{n-1}{n} \sigma_\theta^2.$$

**Preuve 1.7**

$$\begin{aligned}
\bullet E(S_n^2) &= E \left[ \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \right] \\
&= \frac{1}{n} E \left( \sum_{i=1}^n (X_i - \bar{X}_n)^2 \right) \\
&= \frac{1}{n} E \left[ \sum_{i=1}^n (X_i^2 - 2X_i\bar{X}_n + \bar{X}_n^2) \right] \\
&= \frac{1}{n} E \left[ \sum_{i=1}^n X_i^2 - 2\bar{X}_n \sum_{i=1}^n X_i + n\bar{X}_n^2 \right] \\
&= \frac{1}{n} E \left[ \left( \sum_{i=1}^n X_i^2 \right) - n\bar{X}_n^2 \right] \\
&= \frac{1}{n} \left[ \sum_{i=1}^n E(X_i^2) - nE(\bar{X}_n^2) \right] \\
&= \frac{1}{n} \sum_{i=1}^n E(X_i^2) - E(\bar{X}_n^2) \\
&= \frac{1}{n} \sum_{i=1}^n (Var(X_i) + E(X_i)^2) - (Var(\bar{X}_n) + E(\bar{X}_n)^2) \\
&= \sigma_\theta^2 + m_\theta^2 - \frac{\sigma_\theta^2}{n} - m_\theta^2 = \sigma_\theta^2 - \frac{\sigma_\theta^2}{n} = \frac{n-1}{n} \sigma_\theta^2
\end{aligned}$$

d'où  $E(S_n^2) = \frac{n-1}{n} \sigma_\theta^2$ .

**Remarque 1.6** on remarque que  $S_n^2$  est un estimateur biaisé car

$$E(S_n^2) - \sigma_\theta^2 = \frac{n-1}{n} \sigma_\theta^2 - \sigma_\theta^2 = \frac{\sigma_\theta^2}{n} \neq 0.$$

D'autre part on remarque que  $S_n^2$  est un estimateur asymptotiquement sans biais de la variance  $\sigma_\theta^2$  car

$$\lim_{n \rightarrow +\infty} \frac{n-1}{n} \sigma_\theta^2 = \sigma_\theta^2.$$

**Définition 1.30** On appelle variance empirique modifiée de l'échantillon  $X_1, X_2, \dots, X_n$  la statistique

$$\widetilde{S}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2.$$

**Proposition 1.10** La moyenne de la loi de la variance empirique modifiée de l'échantillon  $X_1, X_2, \dots, X_n$  est

$$E(\widetilde{S}_n^2) = \sigma_\theta^2.$$

**Preuve 1.8** D'après sa définition

$$\begin{aligned} \widetilde{S}_n^2 &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2 = \frac{1}{n-1} \left[ \frac{n}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2 \right] \\ &= \frac{n}{n-1} \left[ \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2 \right] \\ &= \frac{n}{n-1} S_n^2 \end{aligned}$$

alors

$$\begin{aligned} E(\widetilde{S}_n^2) &= E\left(\frac{n}{n-1} S_n^2\right) = \frac{n}{n-1} E(S_n^2) \\ &= \frac{n}{n-1} \frac{n-1}{n} \sigma_\theta^2 = \sigma_\theta^2 \end{aligned}$$

d'où  $\widetilde{S}_n^2$  est un estimateur sans biais du  $\sigma_\theta^2$ .

**Définition 1.31** • Soient  $X_1, X_2, \dots, X_n$  un  $n$ -échantillon d'un loi connu  $P_\theta$ , on appelle moment empirique d'ordre  $k$ , la statistique  $M_n$  définie par

$$M_n = \frac{1}{n} \sum_{i=1}^n X_i^k.$$

• On appelle moment centré empirique d'ordre  $k$ , la statistique  $M_n^*$  définie par

$$M_n^* = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X})^k.$$

### 1.3.1 Estimateur du maximum de vraisemblance

**Définition 1.32** Soit un modèle statistique  $(\Omega, \mathcal{A}, P)$  où  $P = P_\theta/\theta \in \Theta \subset \mathbb{R}^2$ . On appelle vraisemblance de l'observation  $x$ , la fonction

$$\begin{aligned} L(x; \cdot) : \Theta &\longrightarrow \mathbb{R}_+ \\ \theta &\longrightarrow L(x; \theta) \end{aligned}$$

où  $L(x; \theta)$  représente la densité de probabilité du vecteur  $(X_1, X_2, \dots, X_n)$  au point  $(x \in \mathbb{R}^n)$ . Si les  $X_i$  ont la densité  $f(x_i, \theta)$  alors

$$L(x; \theta) = L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta).$$

**Définition 1.33** Soit  $X_1, X_2, \dots, X_n$  un  $n$ -échantillon d'une loi connue  $P_\theta$ . Estimer le paramètre  $\theta$  par la méthode de vraisemblance c'est trouver quantité  $\hat{\theta}_n$  (en fonction de  $X_1, X_2, \dots, X_n$ ) qui maximise la vraisemblance  $L(x; \theta)$ ,  $\hat{\theta}_n$  est appelé estimateur de maximum de vraisemblance (EMV) et on écrit

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} L(x_1, \dots, x_n; \theta).$$

**Remarque 1.7** • L'estimation par la méthode de maximum de vraisemblance vérifie (pour  $\Theta \subseteq \mathbb{R}$ ) :

$$\begin{cases} \frac{\partial \ln L(x_1, \dots, x_n; \theta)}{\partial \theta} = 0 \\ \frac{\partial^2 \ln L(x_1, \dots, x_n; \theta)}{\partial \theta^2} < 0 \end{cases}$$

ce dernier système d'équations s'appelle l'équation de vraisemblance quand doit la résoudre pour trouver l'EMV.

• Dans le cas d'un estimateur vectoriel (i.e.  $\Theta \subseteq \mathbb{R}^d$ ), l'estimation de maximum de vraisemblance  $\hat{\theta}_n$  est la solution du système d'équations

$$\begin{cases} \frac{\partial \ln L(x_1, \dots, x_n; \theta)}{\partial \theta_i} = 0, \forall i = \overline{1, n} \\ \frac{\partial^2 \ln L(x_1, \dots, x_n; \theta)}{\partial \theta_i \partial \theta_j} < 0, \forall i, j = \overline{1, n} \end{cases}$$

**Exemple 1.10** Soit  $X_1, X_2, \dots, X_n$  un  $n$ -échantillon d'une loi connue  $P_\theta = B(\theta)$  Bernoulli de paramètre  $\theta$ . La vraisemblance est définie par

$$\begin{aligned} L(x; \theta) &= L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n P(X_i = x_i) \\ &= \prod_{i=1}^n \theta^{x_i} (1 - \theta)^{1-x_i} \\ &= \theta^{\sum_{i=1}^n x_i} (1 - \theta)^{n - \sum_{i=1}^n x_i} \end{aligned}$$

Donc

$$\ln L(x; \theta) = \sum_{i=1}^n x_i \ln \theta + (n - \sum_{i=1}^n x_i) \ln(1 - \theta),$$

d'où

$$\begin{aligned} \frac{\partial \ln L(x_1, \dots, x_n; \theta)}{\partial \theta} = 0 &\Leftrightarrow \frac{\sum_{i=1}^n x_i}{\theta} - \frac{n - \sum_{i=1}^n x_i}{1 - \theta} = 0 \\ &\Rightarrow \hat{\theta} = \frac{\sum_{i=1}^n x_i}{n}. \end{aligned}$$

De plus

$$\frac{\partial^2 \ln L(x_1, \dots, x_n; \theta)}{\partial \theta^2} = -\frac{\sum_{i=1}^n x_i}{\theta^2} - \frac{n - \sum_{i=1}^n x_i}{(1 - \theta)^2}$$

qui est négative puisque les  $X_i$  se trouvent dans l'ensemble  $\{0, 1\}$  c'est-à-dire la fonction  $\ln L(x; \theta)$  est concave. Ainsi l'EMV est  $\hat{\theta} = \frac{\sum_{i=1}^n x_i}{n} = \bar{X}_n$ .

# La distribution de Khi-deux centré et décentré

## 2.1 La distribution de Khi-deux centré

### 2.1.1 Loi Gamma

#### Fonction Gamma

**Définition 2.1** Soit  $r$  un réel strictement positif. La fonction Gamma noté  $\Gamma(r)$  est l'intégral généralisée

$$\Gamma(r) = \int_0^{+\infty} e^{-x} x^{r-1} dx.$$

**Proposition 2.1** Soit  $r$  un réel strictement positif, alors

- i)  $\Gamma(r) = (r - 1)\Gamma(r - 1)$ .
- ii)  $\Gamma(n) = (n - 1)!$ , pour tout  $n$  entier naturel strictement positif.
- iii)  $\Gamma(1) = 1$  et  $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ .

**Définition 2.2** Une variable aléatoire réelle  $X$  suit la loi Gamma de paramètres  $\lambda$  et  $r$  ( $\lambda > 0, r > 0$ ) si sa fonction de densité s'écrit

$$f_X(x) = \begin{cases} \frac{\lambda^r}{\Gamma(r)} \exp^{-\lambda x} x^{r-1} & : x \geq 0 \\ 0 & : \text{si non} \end{cases} = \frac{\lambda^r}{\Gamma(r)} \exp^{-\lambda x} x^{r-1} I_{[0,+\infty]}$$

On notera  $X \sim \Gamma_{\lambda,r}$ .

**Proposition 2.2** Si  $X$  suit la loi Gamma de paramètres  $\lambda, r$  son espérance et sa variance sont données par :

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

**Preuve 2.1**

$$\begin{aligned} \bullet E(X) &= \int_{-\infty}^{+\infty} x f_X(x) dx = \int_0^{+\infty} x \frac{\lambda^r}{\Gamma(r)} e^{-\lambda x} x^{r-1} dx \\ &= \frac{\lambda^r}{\Gamma(r)} \int_0^{+\infty} e^{-\lambda x} x^r dx. \end{aligned}$$

En utilisant le changement de variable  $\langle y = \lambda x, dy = \lambda dx \rangle$ , on trouve

$$\begin{aligned} E(X) &= \frac{\lambda^r}{\Gamma(r)} \int_0^{+\infty} e^{-y} \left(\frac{y}{\lambda}\right)^r \frac{dy}{\lambda} \\ &= \frac{1}{\lambda \Gamma(r)} \int_0^{+\infty} e^{-y} y^r dy \\ &= \frac{\Gamma(r-1)}{\lambda \Gamma(r)} = \frac{r \Gamma(r)}{\lambda \Gamma(r)} = \frac{r}{\lambda}. \end{aligned}$$

$$\begin{aligned} \bullet E(X^2) &= \int_{-\infty}^{+\infty} x^2 f_X(x) dx = \int_0^{+\infty} x^2 \frac{\lambda^r}{\Gamma(r)} e^{-\lambda x} x^{r-1} dx \\ &= \frac{\lambda^r}{\Gamma(r)} \int_0^{+\infty} \exp^{-\lambda x} x^{r-1} dx. \end{aligned}$$

En utilisant le changement de variable  $\langle y = \lambda x, dy = \lambda dx \rangle$ , on trouve

$$\begin{aligned} E(X^2) &= \frac{\lambda^r}{\Gamma(r)} \int_0^{+\infty} e^{-y} \left(\frac{y}{\lambda}\right)^{r+1} \frac{dy}{\lambda} = \frac{1}{\lambda^2 \Gamma(r)} \int_0^{+\infty} e^{-y} y^{r+1} dy \\ &= \frac{\Gamma(r+2)}{\lambda \Gamma(r)} = \frac{r(r+1) \Gamma(r)}{\lambda^2 \Gamma(r)} = \frac{r(r+1)}{\lambda^2} \end{aligned}$$

d'où

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{r(r+1)}{\lambda^2} - \left(\frac{r}{\lambda}\right)^2 = \frac{r}{\lambda^2}.$$

**La loi de Khi-deux centré**

**Définition 2.3** Soit  $X_1, \dots, X_n$  une suite de variable aléatoire indépendants et identiquement distribuées de la loi  $N(0, 1)$ , alors la variable aléatoire

$\sum_{i=1}^n X_i^2$  suit une loi Khi-deux à  $n$  degrés de liberté notée  $\chi_n^2$  et on écrit

$$\sum_{i=1}^n X_i^2 \sim \chi_n^2.$$

**Remarque 2.1** La loi Khi-deux à  $n$  degrés de liberté est la loi Gamma de paramètres  $\lambda = \frac{1}{2}$ ,  $r = \frac{n}{2}$  ( $n \in \mathbb{N}^*$ ), alors sa densité est donnée par

$$f(x) = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} e^{\left(\frac{-x}{2}\right)} x^{\frac{n}{2}-1} I_{[0,+\infty]}(x).$$

Leur fonction caractéristique est

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

**Moment d'ordre  $k$**

**Définition 2.4** Soit  $U$  la variable aléatoire qui suit la loi  $\chi_n^2$ . On appelle moment d'ordre  $k$  la quantité

$$E(U^k) = \int_0^{+\infty} u^k f(u) du$$

où  $f(u)$  est la densité de  $U$  définie par

$$f(u) = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} e^{\left(\frac{-u}{2}\right)} u^{\frac{n}{2}-1} I_{[0,+\infty]}(u).$$

**Proposition 2.3** Soit  $U$  une variable aléatoire qui suit la loi  $\chi_n^2$ . Alors

$$E(U^k) = 2^k \frac{\Gamma\left(\frac{n}{2} + k\right)}{\Gamma\left(\frac{n}{2}\right)}.$$



**Preuve 2.2**

$$\begin{aligned}
E(U^k) &= \int_0^{+\infty} u^k \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} e^{\left(\frac{-u}{2}\right)} u^{\frac{n}{2}-1} du \\
&= \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} \int_0^{+\infty} u^k e^{\left(\frac{-u}{2}\right)} u^{\frac{n}{2}-1} du \\
&= \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} \int_0^{+\infty} u^{\frac{n}{2}+k-1} e^{\left(\frac{-u}{2}\right)} du.
\end{aligned}$$

Le changement de variable  $t = \frac{1}{2}u$  nous donne :

$$\begin{aligned}
E(U^k) &= \frac{2\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} \int_0^{+\infty} (2t)^{\frac{n}{2}+k-1} e^{-t} dt \\
&= \frac{2^{\frac{n}{2}+k}\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} \int_0^{+\infty} t^{\frac{n}{2}+k-1} e^{-t} dt \\
&= 2^k \frac{\Gamma\left(\frac{n}{2} + k\right)}{\Gamma\left(\frac{n}{2}\right)}.
\end{aligned}$$

**Corollaire 2.1**

i)  $E(\chi_n^2) = n.$

ii)  $Var(\chi_n^2) = 2n.$

**Preuve 2.3**

i) D'après la proposition précédente et la proposition 2.1, on a

$$\begin{aligned}
E(\chi_n^2) &= E(U) = 2 \frac{\Gamma\left(\frac{n}{2} + 1\right)}{\Gamma\left(\frac{n}{2}\right)} \\
&= 2 \frac{\frac{n}{2} \Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = n
\end{aligned}$$

ii) D'après la proposition précédente et la proposition 2.1, on a

$$\begin{aligned} E(U^2) &= 2^2 \frac{\Gamma(\frac{n}{2} + 2)}{\Gamma(\frac{n}{2})} \\ &= 4 \frac{(\frac{n}{2} + 1)\Gamma(\frac{n}{2} + 1)}{\Gamma(\frac{n}{2})} \\ &= 4 \frac{(\frac{n}{2} + 1)\frac{n}{2}\Gamma(\frac{n}{2})}{\Gamma(\frac{n}{2})} \\ &= (n + 2)n \end{aligned}$$

d'où

$$\text{Var}(\chi_n^2) = \text{Var}(U) = E(U^2) - (E(U))^2 = (n + 2)n - n^2 = 2n.$$

## 2.2 La distribution de Khi-deux décentré

**Définition 2.5** Soit  $X_1, \dots, X_n$   $n$  variables aléatoires indépendents de la loi normal  $N(\mu_i, \sigma_i)$  alors la variable aléatoire réelle  $\sum_{i=1}^n \left(\frac{X_i}{\sigma_i}\right)^2$  suit une loi

de Khi-deux décentré, notée  $\chi_n^2(\lambda)$ , dépend de deux paramètres

- $n$  : le nombre de degrés de liberté,
- $\lambda = \sum_{i=1}^n \left(\frac{\mu_i}{\sigma_i}\right)^2$  : appelé le paramètre de décentrage.

Sa fonction caractéristique est

$$\phi(u) = \frac{e^{\left(\frac{i\lambda u}{1-2iu}\right)}}{(1-2iu)^{\frac{n}{2}}}.$$

**Définition 2.6** Soit  $h$  une fonction mesurable et  $U \sim \chi_n^2(\lambda)$ . On définit l'expression de l'espérance de  $h$  par

$$E(h(U)) = \sum_{k=0}^{+\infty} \left[ \int_0^{+\infty} h(u) \chi_{n+2k}^2 du \right] \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!}$$

où  $\chi_{n+2k}^2 \approx \Gamma_{\frac{n+2k}{2}, \frac{1}{2}}$ .

**Proposition 2.4** Soit  $U \sim \chi_n^2(\lambda)$ , alors

- i)  $E(U) = n + \lambda$ .
- ii)  $Var(U) = 2n + 4\lambda$ .

### Preuve 2.4

i) De la définition précédente on pose  $h(U) = U$ , on obtient

$$\begin{aligned} E(U) &= \sum_{k=0}^{+\infty} \left[ \int_0^{+\infty} u \chi_{n+2k}^2 du \right] \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!} \\ &= \sum_{k=0}^{+\infty} E(\chi_{n+2k}^2) \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!} \end{aligned}$$

D'après le corollaire 2.1, on arrive à

$$\begin{aligned}
E(U) &= \sum_{k=0}^{+\infty} (n + 2k) \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!} \\
&= \sum_{k=0}^{+\infty} n \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!} + \sum_{k=0}^{\infty} 2k \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!} \\
&= n e^{-\frac{\lambda}{2}} \sum_{k=0}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{k!} + 2 e^{-\frac{\lambda}{2}} \sum_{k=0}^{+\infty} k \frac{\left(\frac{\lambda}{2}\right)^k}{k!} \\
&= n e^{-\frac{\lambda}{2}} \sum_{k=0}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{k!} + 2 e^{-\frac{\lambda}{2}} \sum_{k=1}^{+\infty} k \frac{\left(\frac{\lambda}{2}\right)^k}{k!} \\
&= e^{-\frac{\lambda}{2}} \left( n \sum_{k=0}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{k!} + 2 \sum_{k=1}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^{(k-1+1)}}{(k-1)!} \right) \\
&= e^{-\frac{\lambda}{2}} \left( n \sum_{k=0}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{k!} + 2 \sum_{k=1}^{+\infty} \frac{\lambda \left(\frac{\lambda}{2}\right)^{(k-1)}}{2 (k-1)!} \right) \\
&= e^{-\frac{\lambda}{2}} \left( n \sum_{k=0}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{k!} + \lambda \sum_{k=0}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{k!} \right) \\
&= e^{-\frac{\lambda}{2}} (n + \lambda) \left( \sum_{k=0}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{k!} \right).
\end{aligned}$$

On sait que :  $e^{\frac{\lambda}{2}} = \sum_{k=0}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{k!}$ , alors

$$E(U) = n + \lambda.$$

ii) De la définition précédente on pose  $h(U) = U^2$ , on obtient

$$\begin{aligned}
E(U^2) &= \sum_{k=0}^{+\infty} \left[ \int_0^{+\infty} u^2 \chi_{n+2k}^2 du \right] \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!} \\
&= \sum_{k=0}^{+\infty} \left( \text{Var}(\chi_{n+2k}^2) + (E(\chi_{n+2k}^2))^2 \right) \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!}.
\end{aligned}$$

D'après le corollaire 2.1, on arrive à

$$\begin{aligned} E(U^2) &= \sum_{k=0}^{+\infty} \left( 2(n+2k) + (n+2k)^2 \right) \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!} \\ &= \sum_{k=0}^{+\infty} \left( (2n+n^2) + 4(n+1)k + 4k^2 \right) \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!}. \end{aligned}$$

On suit le même calcul de  $i$ ), on obtient

$$\begin{aligned} E(U^2) &= \left( (2n+n^2) + 4(n+1) \frac{\lambda}{2} \right) + 4e^{-\frac{\lambda}{2}} \sum_{k=0}^{+\infty} k^2 \frac{\left(\frac{\lambda}{2}\right)^k}{k!} \\ &= \left( (2n+n^2) + 4(n+1) \frac{\lambda}{2} \right) + 4e^{-\frac{\lambda}{2}} \sum_{k=1}^{+\infty} k^2 \frac{\left(\frac{\lambda}{2}\right)^k}{k!} \\ &= \left( (2n+n^2) + 4(n+1) \frac{\lambda}{2} \right) + 4e^{-\frac{\lambda}{2}} \sum_{k=1}^{+\infty} k \frac{\left(\frac{\lambda}{2}\right)^k}{(k-1)!} \\ &= \left( (2n+n^2) + 4(n+1) \frac{\lambda}{2} \right) \\ &\quad + 4e^{-\frac{\lambda}{2}} \sum_{k=1}^{+\infty} (k-1+1) \frac{\left(\frac{\lambda}{2}\right)^k}{(k-1)!} \\ &= \left( (2n+n^2) + 4(n+1) \frac{\lambda}{2} \right) \\ &\quad + 4e^{-\frac{\lambda}{2}} \left( \sum_{k=2}^{+\infty} (k-1) \frac{\left(\frac{\lambda}{2}\right)^k}{(k-1)!} + \sum_{k=1}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{(k-1)!} \right) \\ &= \left( (2n+n^2) + 4(n+1) \frac{\lambda}{2} \right) \\ &\quad + 4e^{-\frac{\lambda}{2}} \sum_{k=2}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{(k-2)!} + 4 \frac{\lambda}{2} \\ &= \left( (2n+n^2) + 4(n+1) \frac{\lambda}{2} \right) + 4 \left(\frac{\lambda}{2}\right)^2 + 4 \frac{\lambda}{2} \\ &= n(n+2) + 2\lambda(n+2) + \lambda^2. \end{aligned}$$

Alors

$$\text{Var}(U) = E(U^2) - (E(U))^2 = n(n+2) + 2\lambda(n+2) + \lambda^2 - (n+\lambda)^2 = 2n + 4\lambda.$$

## *Application sur le calcul de risque*

### 3.1 Calcul de risque quadratique de l'estimateur usuel

Soit  $X$  un vecteur gaussien,  $X \sim N_p(\mu, \Sigma)$  où  $(\Sigma = I_p)$  et  $\mu$  est un vecteur inconnu. Il est clair que l'estimateur usuel de  $\mu$  est  $X$ , sa fonction de risque d'après la définition 1.27 est

$$R(X, \mu) = E(\|X - \mu\|^2).$$

On va chercher sa valeur.

Comme  $X \sim N_p(\mu, I_p)$  alors  $X - \mu \sim N_p(0, I_p)$  par conséquence

$$\|X - \mu\|^2 \sim \chi_p^2.$$

D'après le corollaire 2.1

$$E(\|X - \mu\|^2) = E(\chi_p^2) = p$$

ce qui veut dire

$$R(X, \mu) = E(\|X - \mu\|^2) = p. \tag{3.1}$$

## 3.2 Calcul de risque quadratique de l'estimateur de James Stien

**Définition 3.1** On dit qu'un estimateur  $\delta_1$  domine un estimateur  $\delta_2$  si et seulement si

$$R(\delta_1, \mu) \leq R(\delta_2, \mu).$$

**Définition 3.2** Un estimateur  $\delta_1$  est dite inadmissible s'il existe un autre estimateur  $\delta_2$  qui domine  $\delta_1$  c-à-d  $R(\delta_2, \mu) \leq R(\delta_1, \mu)$ .

**Définition 3.3** (estimateur de James- Stein)

Soit  $X \sim N_p(\mu, \Sigma)$ . James et Stein 1961, on introduit la classe des estimateurs noté

$$\delta_a(X) = \left(1 - a \frac{1}{\|X\|^2}\right)X, \quad \forall a \in \mathbb{R}_+$$

qui sont appelés estimateur à rétrécisseur.

### Importance de l'estimateur de type James-Stein

Stein 1956, a montré l'inadmissibilité de l'estimateur usuel  $X$  quand la dimension de l'espace des paramètres  $p \geq 3$ , où James et Stien ont montré plus tard que si  $0 \leq a \leq 2(p-2)$  alors l'estimateur  $\delta_a(X)$  domine l'estimateur  $X$  (c-à-d  $R(\delta_a(X), \mu) \leq R(X, \mu)$ ).

La proposition suivante nous donne la valeur de risque de l'estimateur de James  $\delta_a$ .

### Proposition 3.1

$$R(\delta_a(X), \mu) = p + a^2 E\left(\frac{1}{\|X\|^2}\right) - 2a(p-2)E\left(\frac{1}{\|X\|^2}\right).$$

Pour la preuve de cette proposition on aura besoin du lemme suivant :

**Lemme 3.1** (de Stein)

Soit  $f$  une fonction mesurable si  $X$  un variable gaussian  $X \sim N(\mu, 1)$  alors

$$E((X - \mu)f(X)) = E(f'(X)).$$

**Preuve 3.1** (Preuve de la proposition 3.1)

on a :

$$\begin{aligned}
R(\delta_a(X), \mu) &= E(\|\delta_a(X) - \mu\|^2) \\
&= E\left(\left\|\left(1 - \frac{a}{\|X\|^2}\right)X - \mu\right\|^2\right) \\
&= E\left(\|X - \mu - \frac{a}{\|X\|^2}X\|^2\right) \\
&= E\left(\|X - \mu\|^2\right) + a^2 E\left(\frac{1}{\|X\|^2}\right) - 2E\left(\langle X - \mu, \frac{a}{\|X\|^2}X \rangle\right) \\
&= p + a^2 E\left(\frac{1}{\|X\|^2}\right) - 2a \sum_{i=1}^p E\left((x_i - \mu_i) \frac{1}{\|X\|^2} x_i\right)
\end{aligned}$$

D'après le lemme de stein

$$E\left((x_i - \mu_i) \frac{1}{\|X\|^2} x_i\right) = E\left(\frac{\partial}{\partial x_i} \left(\frac{1}{\|X\|^2} x_i\right)\right)$$

où  $x_i$  sont les composantes du  $X$  et  $\|X\|^2 = \sum_{i=1}^p x_i^2$ .

Donc

$$\begin{aligned}
E\left((x_i - \mu_i) \frac{1}{\|X\|^2} x_i\right) &= E\left(\left(\frac{\partial}{\partial x_i} x_i\right) \frac{1}{\|X\|^2} + x_i \frac{\partial}{\partial x_i} \left(\frac{1}{\|X\|^2}\right)\right) \\
&= E\left(\frac{1}{\|X\|^2} + x_i \left(\frac{-2x_i}{\|X\|^4}\right)\right) \\
&= E\left(\frac{1}{\|X\|^2} - 2\left(\frac{x_i^2}{\|X\|^4}\right)\right).
\end{aligned}$$



Alors

$$\begin{aligned}
R(\delta_a(X), \mu) &= p + a^2 E\left(\frac{1}{\|X\|^2}\right) - 2a \sum_{i=1}^p E\left(\frac{1}{\|X\|^2} - 2\frac{x_i^2}{\|X\|^4}\right) \\
&= p + a^2 E\left(\frac{1}{\|X\|^2}\right) - 2a \sum_{i=1}^p \left(E\left(\frac{1}{\|X\|^2}\right) - 2E\left(\frac{x_i^2}{\|X\|^4}\right)\right) \\
&= p + a^2 E\left(\frac{1}{\|X\|^2}\right) - 2a \left[\sum_{i=1}^p E\left(\frac{1}{\|X\|^2}\right) - 2 \sum_{i=1}^p E\left(\frac{x_i^2}{\|X\|^4}\right)\right] \\
&= p + a^2 E\left(\frac{1}{\|X\|^2}\right) - 2a \left[pE\left(\frac{1}{\|X\|^2}\right) - 2E\left(\frac{\sum_{i=1}^p x_i^2}{\|X\|^4}\right)\right] \\
&= p + a^2 E\left(\frac{1}{\|X\|^2}\right) - 2a \left[pE\left(\frac{1}{\|X\|^2}\right) - 2E\left(\frac{\|X\|^2}{\|X\|^4}\right)\right] \\
&= p + a^2 E\left(\frac{1}{\|X\|^2}\right) - 2a \left[pE\left(\frac{1}{\|X\|^2}\right) - 2E\left(\frac{1}{\|X\|^2}\right)\right] \\
&= p + a^2 E\left(\frac{1}{\|X\|^2}\right) - 2a(p-2)E\left(\frac{1}{\|X\|^2}\right)
\end{aligned}$$

et la preuve est terminée.

### **Théorème 3.1**

i)  $\delta_a(X)$  domine  $X$  et seulement si  $0 \leq a \leq 2(p-2)$ .

ii) La valeur optimale de  $a$  qui minimise la fonction de  $R(\delta_a(X), \mu)$  est

$$\hat{a} = p - 2$$

(c-à-d  $\arg \min R(\delta_a(X), \mu) = \hat{a}$ ).

**Preuve 3.2**

i)  $\delta_a(X)$  domine  $X \Leftrightarrow R(\delta_a(X), \mu) \leq R(X, \mu)$ .

D'après l'équation (3.1) et la proposition précédente ceci nous donne

$$\begin{aligned}
 \delta_a(X) \text{ domine } X &\Leftrightarrow p + (a^2 - 2a(p - 2))E\left(\frac{1}{\|X\|^2}\right) \leq p \\
 &\Leftrightarrow (a - 2a(p - 2))E\left(\frac{1}{\|X\|^2}\right) \leq 0 \\
 &\Leftrightarrow a^2 - 2a(p - 2) \leq 0 \\
 &\Leftrightarrow a(a - 2(p - 2)) \leq 0 \\
 &\Leftrightarrow a - 2(p - 2) \leq 0 \\
 &\Leftrightarrow 0 \leq a \leq 2(p - 2).
 \end{aligned}$$

ii) La fonction  $R(\delta_a(X), \mu)$  est une fonction convexe par rapport à  $a$  alors la valeur optimale de  $a$  qui minimise  $R(\delta_a(X), \mu)$  est  $\hat{a}$  tel que

$$\begin{aligned}
 \frac{\delta R(\delta_a(X), \mu)}{\delta a}(\hat{a}) = 0 &\Rightarrow (2\hat{a} - 2(p - 2))E\left(\frac{1}{\|X\|^2}\right) = 0 \\
 &\Rightarrow 2\hat{a} - 2(p - 2) = 0 \\
 &\Rightarrow \hat{a} = p - 2.
 \end{aligned}$$

## *Conclusion*

*Dans ce mémoire, nous avons intéressé à l'estimation de la moyenne d'une loi normal multidimensionnelle par des estimateurs à rétrécisseurs déduite de l'estimateur usuel. Pour comparer entre deux estimateur, nous avons utilisé le risque quadratique. Nous avons étudié la fameuse classe des estimateurs à rétrécisseur, c'est la classe des estimateurs introduite par James et Stein 1961 puis nous déduison la forme de l'estimateur primitif de james- stein qui est le meilleur dans cette classe puis nous montrons que cet estimateur domine l'estimateur usuel et par conséquent il est minimax.*

## *Bibliographie*

- [1] *J. Y. Dauxois*, Statistique inférentielle, *CTU, Licence de Mathématiques, université de France-Comté, 2012.*
- [2] *O. Gaudoin*, Statistique inférentielle Avancée, *Notes de cours.*
- [3] *W. James and C. Stien*, Estimation of quadratic loss, *Proc 4th Berkeley Symp. Math. Statist. Prob, Vol 1, (1961), 361-379.*
- [4] *A. Philipe and M. C. Viano*, Cours de Probabilités et Statistique de base , *université de Lille 1, 2004.*
- [5] *C. Stien*, Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, *Proc 3th Berkeley Symp. Math. Statist. Prob, Vol 1, (1956), 197-206.*

## ملخص

في هذه المذكرة نحن مهتمون بمقدار متوسط التوزيع الطبيعي متعدد الأبعاد بواسطة مقدرات الانكماش المستخلصة من المقدار المعتاد، وذلك باستخدام المخاطر التربيعية للمقدار المعتاد، والفئة العامة لجيمس شتاين "james-stein" بتركيز على فئة من الدرجة الثانية له والتي تم اكتشافها 1961.

في الأخير استخلصنا ان المقدار البدائي لجيمس في هذه الفئة هو المقدار الأفضل و المهيمن على المقدار المعتاد، وبالتالي فهو الحد الأدنى.

## Résumé

Dans ce mémoire, nous avons intéressé à l'estimation de la moyenne d'une loi normale multidimensionnelle par des estimateurs à rétrécisseurs déduite de l'estimateur usuel. Pour comparer entre deux estimateur, nous avons utilisé le risque quadratique. Nous avons étudié la fameuse classe des estimateurs à rétrécisseur, c'est la classe des estimateurs introduit par James et Stein 1961 puis nous déduison la forme de l'estimateur primitif de james-stein qui est le meilleur dans cette classe puis nous montrons que cet estimateur domine l'estimateur usuel et par conséquent il est minimax.

## Abstract

In this note, we are interested in studying the amount of the mean of a multi-dimensional normal distribution by means of contraction estimators extracted from the usual amount, using the quadratic risks of the normal amount and the general category of Gemish Stein with a focus on a second-order category for him, which was discovered in 1961. In this remembrance, we concluded that the primitive amount of Gemesh in this category is the best and dominant over the usual amount, and therefore it is the minimum.