République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

Université de Saida - Dr Moulay Tahar.

Faculté des Sciences.

Département de Mathématiques.

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Filière : MATHEMATIQUES

Spécialité : Analyse Stochastique, Statistique des Processus et Applications

par

Azzouz Amel Aicha [1]

Sous la direction de

Dr A. Benkhaled

Thème :

Estimateurs De Type Polynomial De La Moyenne D'une Loi Normale Multidimensionnelle Sous Une Fonction De Coût Quadratique

Soutenue le 06/07/2022 devant le jury composé de

T. Djebbouri	Université de Saïda Dr. Moulay Tahar	Président
A. Benkhaled	Université de Mustafa Stambouli Mascara	Encadreur
K. Djerfi	Université de Saïda Dr. Moulay Tahar	Examinateur

Année univ.: 2021/2022

1. e-mail : amelaicha14@gmail.com

Dédicace

Je dédie ce travail

À ma chère mère et à mon cher père qui n'ont jamais cessé de me supporter, me soutenir et m'encourager durant mes années d'études.
Qu'iles trouvent ici le témoignage de ma profonde gratitude et reconnaissance.

À mes frères, mes soeurs et ma famille qui me donnent de l'amour et de la vivacité.

À tous ceux qui m'ont aidé-de prés ou de loin-et ceux qui ont partagé avec moi les moments d'émotion lors de la réalisation de ce travail et qui m'ont chaleureusement supporté et encouragé tout au long de mon parcours.

À tous mes amis qui m'ont toujours encouragé, et à qui je souhaite plus de succés.

Merci !

Amel Aicha

Remerciements

Tout d'abord, nous tenons à remercier le "**BON DIEU**" le tout puissant de nous avoir accorder la patience, le courage et la volonté afin de réaliser ce modeste travail.

Je tiens à remercier tout particuliérement mon encadreur, **Dr. Abdelkader Benkhaled** pour ses conseils, sa grande disponibilité et sa générosité. La pertinence de ses questions et de ses remarques ont toujours su me motiver et me diriger.

Je voudrais également remercier tous les membres de jury d'avoir accepté d'évaluer et d'examiner ce travail, merci pour toutes leurs remarques et critiques.

Je remercie chaleureusement toute ma famille, qui m'a soutenu, encouragé et poussé durant toutes mes années d'étude. Je voudrais exprimer mes sincères remerciements à mes amis permanents, qui m'ont toujours entouré et m'ont motivé à continuer à meilleure.

J'exprime mes remerciements à tous mes enseignants du département de Mathématiques qui m'ont initié aux valeurs authentiques, en signe d'un profond respect et d'un profond amour, ainsi que le personnel de l'administration.

Je remercie tous les membres de laboratoire des Modèles Stochastique Statistique et Application pour leur accueil et leur sympathie.

Merci à tous

Résumé

Dans ce travail, on s'intéresse à l'étude de l'estimation de la moyenne d'une loi normale multidimensionnelle à variance connue. On prend comme critère adopté pour comparer deux estimateurs, le risque associé à une fonction de coût quadratique générale. On étudie plus particulièrement la minimaxité des estimateurs á rétrécisseurs de type James-Stein et de type Polynomial. A la fin du mémoire, on illustre les résultats théoriques par des représentations graphiques des fonctions des risques des estimateurs considérés.

Mots clés : Estimateur de type James-Stein, estimateur de type Polynomial, loi normale mutidimensionnelle.

Table des matières

Introdution

Dans ce mémoire, nous nous intéressons à l'estimation paramétrique de la moyenne d'une loi normale multidimentionnelle par deux formes d'estimateurs à rétrécisseur, de type James-Stein et de type Polynomial. Ce travail se présente en quatre chapitres, décrits successivement comme suit :

Le chapitre un est introductif, on présente un panorama général sur la théorie des estimateurs paramétrique, vecteurs gaussiens, Modèle Statistique, construction d'estimateurs, qualité d'un estimateur ect.

Dans le deuxième chapitre, nous introduisons les estimateurs de type James-Stein. Sous des hypothèses de régularité nous établissons la minimaxité.

Dans le troisième chapitre, on considère les estimateurs de type polynomiale avec l'indéterminée $1/\|X\|^2$ et on montre que si on augmente le degré du polynôme on peut construire un meilleur estimateur à partir de celui construit précédemment.

Le dernier chapitre sera consacré à l'étude de simulation. En premier temps, nous représentons graphiquement les rapport de risques des estimateurs δ_{JS}, $\delta_b^{(2)}$ et $\delta_c^{(3)}$ par rapport à X. En second temps, nous donnons un tableau contients les valeurs des rapport de risques des estimateurs δ_{JS}, $\delta_b^{(2)}$ et $\delta_c^{(3)}$ par rapport à X pour différentes valeurs de p et λ.

Finalement, le mémoire s'achève par une conclusion générale.

Chapitre 1

Introduction générale

Sommaire

1.1 Lois gaussiennes

Définition 1.1. Soit X une variable aléatoire réelle. On dit que X est une variable aléatoire gaussienne de paramètres (μ, σ^2) avec $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}^+$ (on note $X \sim N(\mu, \sigma^2)$) si et seulement si X vérifie une des deux conditions suivantes :

- $\sigma > 0$ et X admet pour densité

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

- $\sigma = 0$ et X est presque sûrement égale à μ.

Remarque 1.1. Dans le deuxième cas, on parle de lois gaussiennes dégénérées et donc la variable aléatoire X n'admet pas de densité par rapport à la mesure de Lebesgue.

Proposition 1.1. Une variable aléatoire X de loi $N(\mu, \sigma^2)$ a pour
- Espérance : $\mathbb{E}[X] = \mu$.
- Variance : $Var(X) = \sigma^2$.
- Fonction caractéristique

$$\phi_X(t) = \mathbb{E}(e^{itX}) = e^{it\mu} e^{-\frac{t^2\sigma^2}{2}}, \ t \in \mathbb{R}$$

Lorsque la moyenne μ vaut 0, et l'écart-type vaut 1, la loi sera notée $N(0,1)$ et sera appelée loi normale standard. Sa fonction caractéristique vaut $e^{-\frac{t^2}{2}}$. Seule la loi $N(0,1)$ est tabulée car les autres lois (c'està- dire avec d'autres paramètres) se déduise de celle-ci à l'aide du théorème suivant :

Théorème 1.1. *Si la variable aléatoire X suit une loi $N(\mu, \sigma^2)$, alors $Y := \frac{X-\mu}{\sigma}$ suit la loi $N(0,1)$*

1.2 Vecteurs gaussiens

Définition 1.2.
- Un ***vecteur aléatoire*** est un vecteur $(X_1, \ldots, X_n)$ composé de n variables aléatoires définies sur le même espace.
- Un vecteur aléatoire $(X_1, \ldots, X_n)$ est dit L^1, resp. L^2, si $\mathbb{E}[X_i] < +\infty$, resp. $\mathbb{E}[X_i^2] < +\infty$, pour tout $1 \leq i \leq n$.
- La ***matrice de covariance*** d'un vecteur aléatoire $X = (X_1, \ldots, X_n)$ L^2 est la matrice carrée symétrique, positive

$$\Sigma = (Cov(X_i, Y_j))_{1 \leq i,j \leq n}.$$

- ***L'espérance*** d'un vecteur aléatoire $X = (X_1, \ldots, X_n)$ L^1 est le vecteur des espérances de ses marginales

$$\mathbb{E}(X) = (\mathbb{E}(X_1), \ldots, \mathbb{E}(X_n)).$$

Définition 1.3. Un vecteur aléatoire $X = (X_1, \dots, X_n)^t$ est gaussien si et seulement si toutes les combinaisons linéaires de ses coordonnées $\langle a, X \rangle = a_1 X_1 + \dots + a_n X_n$ suit une loi gaussienne dans $\mathbb{R}$ (pour tout $a = (a_1, \dots, a_n)^t \in \mathbb{R}^n$).

Proposition 1.2. Si ψ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ et si X est un vecteur gaussien de dimension n alors $\psi(X)$ est aussi un vecteur gaussien de dimension m.

Remarque 1.2.

- Si X est un vecteur gaussien alors pour toute partie $\{i_1, \dots, i_p\}$ de $\{1, \dots, n\}$, le vecteur $(X_{i_1}, \dots, X_{i_p})$ est gaussien.
- Un vecteur gaussien est nécessairement L^2 puisque, par définition, chacune de ses marginales X_i est gaussienne donc L^2.

Théorème 1.2. *Un vecteur aléatoire X à valeurs dans $\mathbb{R}^n$ est un vecteur gaussien si et seulement si X est L^2 et il admet pour fonction caractéristique*

$$\phi_X(u) = \mathbb{E}(e^{itX}) = e^{iu^t\mu} e^{-\frac{1}{2}u^t\Sigma u}, \ u \in \mathbb{R}^n$$

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

Proposition 1.3. Soit $X \sim N_n(\mu, \Sigma)$ un vecteur gaussien de dimension n, de moyenne μ et de covariance Σ. Les variables aléatoires $X_1, \dots, X_n$ sont indépendantes si et seulement si la matrice Σ est diagonale.

Proposition 1.4. 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 a 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 $q \times q$.

Proposition 1.5. La densité d'un vecteur gaussien $X \sim N_n(\mu, \Sigma)$ non dégénéré (i.e $det\Sigma \neq 0$) est

$$f_X(x) = \frac{exp(-\langle (x-\mu), \Sigma^{-1}(x-\mu) \rangle / 2)}{((2\pi)^n det\Sigma)^{1/2}}, \ x \in \mathbb{R}^n$$

1.3 Loi du χ^2 (khi-deux)

Définition 1.4. Soit $Z_1, \ldots, Z_\nu$ une suite de variables aléatoires indépendantes de même loi $N(0,1)$. Alors la variable aléatoire $\sum_{i=1}^{\nu} Z_i^2$ suit une loi appelée loi du Khi-deux à ν degrés de liberté, notée χ^2_ν.

Proposition 1.6.

- La densité de la loi du χ^2_ν est

$$f_{\chi^2_\nu}(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2},\ x > 0$$

  où Γ est la fonction Gamma d'Euler définie par $\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx$
- L'espérance de la loi du χ^2_ν est égale au nombre ν de degrés de liberté et sa variance est 2ν.
- Sa fonction caractéristique est $\phi_{\chi^2_\nu}(t) = (1-2it)^{-\nu/2}$
- Pour $n \geq 30$, $\sqrt{2\chi^2_\nu} - \sqrt{2n-1}$ suit approximativement une loi $N(0,1)$.

1.4 Le moment d'ordre k

Définition 1.5. Soit X une variable aléatoire qui suit la loi χ^2_p. On appelle moment d'ordre k la quantité

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

où $f(u)$ est la densité de X.

Proposition 1.7. Soit X une variable aléatoire qui suit la loi χ^2_p. Alors

$$\mathbb{E}(X^k) = 2^k \frac{\Gamma(\frac{p}{2}+k)}{\Gamma(\frac{p}{2})}$$

D'après la proposition précédente $\mathbb{E}(\chi^2_p) = \frac{\frac{p}{2}}{\frac{1}{2}} = p$ et $Var(\chi^2_p) = \frac{\frac{p}{2}}{\frac{1}{4}} = 2p$.

1.5 Loi du khi-deux décentrée

Définition 1.6. Soit $X_1, \ldots, X_\nu$ une suite de variables aléatoires indépendantes suivent la loi $N(\theta_i, \sigma_i^2)$, $i = 1 : \nu$. Alors la variable aléatoire $X = \sum_{i=1}^{\nu} \left(\frac{X_i}{\sigma_i}\right)^2$ suit la loi du Khi-deux décentrée, elle dépend de deux paramètres :
ν : est le nombre de degrés de liberté.
λ : est le paramètre de décentrage, il est donné par $\lambda = \sum_{i=1}^{\nu} \left(\frac{\theta_i}{\sigma_i}\right)^2$ et on note $X \sim \chi_\nu^2(\lambda)$.

Proposition 1.8.

- La densité de la loi du $\chi_\nu^2(\lambda)$ est

$$f_{\chi_\nu^2(\lambda)}(x) = \sum_{k=0}^{+\infty} \chi_{p+2k}^2 \frac{e^{-\lambda/2}(\lambda/2)^k}{k!}, \; x > 0$$

- Sa fonction caractéristique est $\phi_{\chi_\nu^2(\lambda)}(t) = \dfrac{e^{\left(\frac{i\lambda t}{1-2it}\right)}}{(1-2it)^{\nu/2}}$

Définition 1.7. Soit h une fonction mesurable et $X \sim \chi_\nu^2(\lambda)$, on définit l'espérence de $h(X)$ par

$$\mathbb{E}(h(X)) = \sum_{k=0}^{+\infty} \left[\int_0^{+\infty} h(x)\chi_{p+2k}^2 dx\right] \frac{e^{-\lambda/2}(\lambda/2)^k}{k!} \tag{1.1}$$

où χ_{p+2k}^2 est la loi de Khi-deux centrée à $p + 2k$ degrés de liberté.

1.6 Estimation paramétrique

1.6.1 Modèle Statistique

Définition 1.8. ► Un échantillon d'une loi est une suite de v.a indépendantes identiquement distribuées (i.i.d).
► Un modèle statistique est la donnée de triplet $(\mathfrak{X}, \mathfrak{A}, (P_\theta)_{\theta \in \Theta})$ où : $\mathfrak{X}$ est l'espace de réalisations, $\mathfrak{A}$ tribu sur $\mathfrak{X}$, $P_\theta = P_X$ loi de X et Θ l'ensemble des paramétres θ.

Exemple 1.1. Soit un échantillonnage de $N(m,\sigma^2)$, c'est à dire une suites $X_1,\ldots,X_n$ de v.a i.i.d avec $\forall i,\ X_i \hookrightarrow N(m,\sigma^2)$, $\mathfrak{X} = \mathbb{R}^n$, $\mathfrak{A} = \mathfrak{B}_{\mathbb{R}^n}$, $P_\theta = N(m,\sigma^2)$ et $\theta = (m,\sigma^2) \in \Theta = \mathbb{R} \times \mathbb{R}^{+*}$.

Définition 1.9. Une statistique est une application T mesurable (v.a) de $(\mathfrak{X},\mathfrak{A})$ dans un espace mesurable $(\mathfrak{F},\mathfrak{H})$.

$$\begin{array}{rccc} T: & (\mathfrak{X},\mathfrak{A}) & \longrightarrow & (\mathfrak{F},\mathfrak{H}) \\ & (X_1,\ldots,X_n) & \longmapsto & T(X_1,\ldots,X_n). \end{array}$$

Définition 1.10. On appelle estimateur de θ, toute statistique T de $(\mathfrak{X},\mathfrak{A})$ à valeurs dans Θ.

1.6.2 Construction d'estimateurs

Méthode des moments

C'est une méthode naturelle dans la mesure où elle est intuitive. Supposons que l'on doive estimer le paramétre θ, la méthode des moments consiste à choisir comme estimateur $\widehat{\theta_n}$ la solution de l'équation obtenue en égalant le moment théorique d'ordre k et le moment empirique d'ordre k.

$$E(X^k) = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

Exemple 1.2. Soit $X \hookrightarrow G(1,\theta)$, donc $E(X) = \dfrac{1}{\theta}$.
$*$ pour $k = 1$ la méthode des moments nous donne $E(X) = \overline{X}_n$, alors un estimateur de θ est

$$\widehat{\theta_n} = \frac{1}{\overline{X}_n}$$

$*$ pour $k = 2$ la méthode des moments nous donne $E(X^2) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i^2$, or $E(X^2) = \theta \displaystyle\int_0^{+\infty} x^2 e^{-\theta x} dx = \dfrac{2}{\theta^2}$, alors un estimateur de θ est

$$\widehat{\theta_n} = \sqrt{\frac{2}{\frac{1}{n}\sum_{i=1}^{n} X_i^2}}$$

Méthode du maximum de vraisemblance

Définition 1.11. Soient $X = (X_1, \ldots, X_n)$ une suites de v.a i.i.d, on appelle fonction de vraisemblance pour X la fonction définie par :

$$L(X_1, \ldots, X_n, \theta) = \begin{cases} \prod_{i=1}^{n} P(X_i, \theta) & \text{si} \quad \text{les } X_i \text{ sont discrètes} \\ \\ \prod_{i=1}^{n} f(X_i, \theta) & \text{si} \quad \text{les } X_i \text{ sont continues} \end{cases}$$

Définition 1.12. l'estimateur de θ par la méthode du maximum de vraisemblance est la valeur $\widehat{\theta}_n$ qui rend maximale la fonction de vraisemblance L.

Les conditions requises pour assurer cette maximisation sont $\frac{dL}{d\theta} = 0$ et $\frac{d^2L}{d\theta^2} < 0$.
Il est par fois plus commode de maximiser le logarithme népérien de L par rapport à θ puisque cette fonction comporte souvent des puissances ou des formes exponentielles, les conditions deviennent alors $\frac{dlnL}{d\theta} = 0$ et $\frac{d^2lnL}{d\theta^2} < 0$.
lnL est une fonction croissante et elle aura sa valeur maximum pour la même valeur de θ qu'aurait la fonction L.

Remarque 1.3. L'estimateur du maximum de vraisemblance peut ne pas exister.

Exemple 1.3. Si les X_i sont de loi $N(m, \sigma^2)$, la fonction de vraisemblance est :

$$L(X_1, \ldots, X_n, m, \sigma^2) = \prod_{i=1}^{n} f(X_i, m, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X_i - m)^2}{2\sigma^2}} = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - m)^2}$$

D'où

$$lnL(X_1, \ldots, X_n, m, \sigma^2) = -\frac{n}{2} ln\sigma^2 - \frac{n}{2} ln2\pi - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - m)^2$$

On doit annuler les dérivées partielles de ce logarithme par rapport à m et σ^2. On a

$$\frac{\partial}{\partial m} lnL(X_1,\ldots,X_n,m,\sigma^2) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} -2(X_i-m) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} X_i - nm\right),$$

qui s'annule pour

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

$$\frac{\partial}{\partial \sigma^2} lnL(X_1,\ldots,X_n,m,\sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(X_i-m)^2,$$

qui s'annule pour

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

1.6.3 Qualité d'un estimateur

Biais d'un estimateur

Définition 1.13. Le biais d'un estimateur est la quantité

$$b_\theta(T) = E_\theta(T) - \theta$$

où E_θ espérance par rapport à P_θ.
* Si $b_\theta(T) = 0$, T est dit estimateur sans biais.
* Si $b_\theta(T) \neq 0$, T est dit estimateur biaisé.

Définition 1.14. Un estimateur $T(X) = (T_n(X))_{n\in\mathbb{N}}$ de θ, où $T_n(X)$ est intégrable pour tout n, est dit asymptotiquement sans biais si $E(T_n(X)) - \theta$ tend vers 0 lorsque n tend vers l'infini et ce pour tout θ dans Θ.

Propriétés 1.1. * La moyenne empirique $\overline{X}_n$ est un estimateur sans biais pour m, en éffet

$$E(\overline{X}_n) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}nm = m$$

$*$ La variance empirique S_e^2 est un estimateur biaisé pour σ^2 mais il est asymptotiquement sans biais, en éffet

$$\begin{aligned} E(S_e^2) &= E(X^2) - E(\overline{X}_n^2) \\ &= V(X) + E(X)^2 - V(\overline{X}_n) - E(\overline{X}_n)^2 \\ &= \frac{n-1}{n} V(X) \\ &= \frac{n-1}{n}\sigma^2 \rightarrow_{n\to\infty} \sigma^2. \end{aligned}$$

En revanche, on voit que $E(\frac{n}{n-1}S_e^2) = \frac{n}{n-1}E(S_e^2) = \sigma^2$. On pose donc

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

Par conséquent S^2 (appelée variance estimée) est un estimateur sans biais pour σ^2.

Estimateur convergent

Définition 1.15. Un estimateur T est dit convergent si $E(T)$ tend vers θ lorsque n tend vers l'infini. Il sera dit consistant si T converge en probabilité vers θ lorsque n tend vers l'infini.

Théorème 1.3. *Si T est convergent et de variance tendant vers* 0 *lorsque n tend vers l'infini alors T est consistant.*

Si T et θ) sont dans $\mathbb{R}$, la définition de la convergence de l'estimateur signifie que l'on a, pour tout $\epsilon > 0$:

$$P(|T - \theta| > \epsilon) \to 0,$$

quand $n \to +\infty$.

Exemple 1.4. Si les X_i sont de loi $B(\theta)$ alors l'estimateur $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ converge en probabilité vers θ lorsque n tend vers l'infini. En effet, soit $\epsilon > 0$

$$p(|\overline{X}_n - \theta| > \epsilon) \leq \frac{Var(\overline{X}_n)}{\epsilon^2} = \frac{1}{\epsilon^2}\frac{1}{n}\theta(1-\theta) \rightarrow_{n\to+\infty} 0$$

On peut considérer d'autres types de convergence, comme la convergence p.s. ou la convergence dans L^p, pour p fixé. Dans ces cas, on dira respectivement que l'estimateur est fortement consistant ou L^p-consistant.

Risque d'un estimateur

On se donne en premier lieu un critère mesurant et pénalisant l'écart entre l'estimateur δ et la vraie valeur θ. On parle de fonction de coût.

Définition 1.16. On appelle fonction de coût (ou de perte) toute fonction L mesurable de $\Theta \times \Theta$ dans $\mathbb{R}^+$.

$$\begin{array}{rccc} L : & \Theta \times \Theta & \longrightarrow & \mathbb{R}^+ \\ & (\delta, \theta) & \longmapsto & L(\delta, \theta). \end{array}$$

Quelques fonctions de coût classiques sont :
1– La fonction de coût valeur absolue : $L(\delta, \theta) = |\delta - \theta|$
2– La fonction de coût quadratique : $L(\delta, \theta) = (\delta - \theta)^2$
Le rôle de chaque fonction de coût est :
► de mesurer la qualité de l'estimation,
► d'aboutir à une solution en minimisant la fonction de coût.

Définition 1.17. On appelle risque d'un estimateur δ de θ associé à la fonction de coût L, la fonction R de Θ vers $\mathbb{R}^+$ définie par

$$R(\delta, \theta) = E(L(\delta, \theta)),$$

pour tout θ de Θ, sous réserve que cette espérance existe.

Remarque 1.4. Quand la fonction de coût est quadratique on parle de risque quadratique.

Proposition 1.9. Soit T un estimateur de θ, si la fonction de coût $L(\delta, \theta)$ est quadratique on a :

$$R(T, \theta) = V_\theta(T) + b_\theta^2(T)$$

Remarque 1.5. Entre deux estimateurs sans biais, le "meilleur" sera celui dont la variance est minimale (on parle d'efficacité).

Exemple 1.5. Soient X_1 et X_2 deux variables aléatoires i.i.d de moyenne θ et de variance σ^2. Soient $\delta_1$et δ_2 deux estimateurs non biasés de θ telle que :

$$\delta_1 = \frac{X_1 + X_2}{2} \text{ et } \delta_2 = \frac{aX_1 + bX_2}{a+b} \text{ où } a, b \in \mathbb{R}$$

$$\begin{aligned}
Var(\delta_1) - Var(\delta_2) &= Var(\frac{X_1 + X_2}{2}) - Var(\frac{aX_1 + bX_2}{a+b}) \\
&= \frac{1}{4} Var(X_1 + X_2) - \frac{a^2 + b^2}{(a+b)^2} Var(X_1) \\
&= \left(\frac{1}{2} - \frac{a^2 + b^2}{(a+b)^2} \right) \sigma^2 \\
&= \left(\frac{(a+b)^2 - 2a^2 - 2b^2}{2(a+b)^2} \right) \sigma^2.
\end{aligned}$$

Comme $2(a+b)^2 > 0$ et $(a+b)^2 - 2a^2 - 2b^2 = -(a-b)^2 < 0$, alors

$$Var(\delta_1) < Var(\delta_2).$$

Donc δ_1 est meilleur que δ_2.

Définition 1.18. Soient δ_1 et δ_2 deux estimateurs de θ. On dit que δ_1 est préférable (domine) à δ_2 si l'on a :

$$R(\delta_1, \theta) \leq R(\delta_2, \theta)$$

pour tout θ de Θ et avec une inégalité stricte pour au moins un θ de Θ.

Définition 1.19. Un estimateur T de θ est dit admissible s'il n'existe pas d'estimateur de θ qui lui soit préférable.

Définition 1.20. Un estimateur T_m de θ est appelé minimax s'il atteint le plus petit risque maximum pour tout autre estimateurs T, ce qui signifie qu'il satisfait

$$sup_{\theta \in \Theta} R(T_m, \theta) = inf_{T \in D} sup_{\theta \in \Theta} R(T, \theta)$$

avec $D = \{T/ \ T \text{ estimateur de } \theta\}$

Information de Fisher

Au vu d'un échentillon $X = (X_1, \ldots, X_n)$ on peut obtenir une certain information sur le paramètre θ, il s'agit de contifier cette information et de montrer qu'il a un intérêt pour les statistiques.

Définition 1.21. L'information de Fisher ($I_X(\theta)$) apporté par X sur le paramètre θ est définie par :

$$I_X(\theta) = E\left(\left(\frac{\partial}{\partial\theta} lnL(X,\theta)\right)^2\right).$$

On peut établir une autre écriture de l'information de Fisher.

Proposition 1.10. L'information de Fisher est aussi égale à

$$I_X(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2} lnL(X,\theta)\right).$$

Proposition 1.11. Soit T une statistique de θ. Alors

$$I_{T(X)}(\theta) \leq I_X(\theta)$$

► **Cas vectoriel** $\theta = (\theta_1, \dots \theta_p)$
On définit L'information de Fisher par la matrice suivantes

$$I_X(\theta) = \left(I_{i,j}(\theta)\right)_{i,j=1,\dots,p}$$

où

$$I_{i,j}(\theta) = -E\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j} lnL(X,\theta)\right).$$

Borne de Cramer-Rao

Le resultat suivant affirme l'existence d'une borne inférieure pour la variance de n'importe qu'el estimateure. Dans la suite on supposera les hypothèses suivantes.

H_1 : Le domaine des réalisation de $X = (X_1, \dots, X_n)$ ne dépent pas de θ.
H_2 : La densité de X est 2 fois dérivable par rapport à θ.
H_3 : On peut dériver par rapport à θ sou le signe d'integrale.

Théorème 1.4. *Soit T un estimateur sans biais de θ. Alors sou les hypothèses H_1, H_2 et H_3, on a :*

$$Var(T) \geq \frac{1}{I_X(\theta)}.$$

La borne $\frac{1}{I_X(\theta)}$ est la borne de Cramer-Rao

Définition 1.22. (**Estimateur efficace**) Un estimateur sans biais T est dit efficace s'il atteint la borne de Cramer-Rao, c'est-à-dire si

$$Var(T) = \frac{1}{I_X(\theta)}.$$

Il est dit **asymptotiquement efficace** si

$$lim_{n\to+\infty}\frac{1}{I_X(\theta)Var(T)} = 1$$

1.6.4 Amélioration d'estimateurs

Statistique exhaustive

Il s'agie de construire une statistique $T(X)$ à partire d'un échantillon $X = (X_1, \ldots, X_n)$ qui vont nous renseigner sur le paramètre θ, sans entraîner de perte d'information.

Définition 1.23. Soit $(\mathfrak{X}, \mathfrak{A}, (P_\theta)_{\theta\in\Theta})$ un modèle paramétrique et $X = (X_1, \ldots, X_n)$ un échantillon dans ce modèle. Une statistique $T(X)$ est dite exhaustive pour le paramètre θ si la loi de X conditionnelle à $T(X)$ est indépendante du paramètre θ.

Le calcul de la loi conditionnelle n'étant pas toujours facile, on utilisera souvent le théorème suivant qui donne un moyen plus aisé pour prouver l'exhaustivité d'une statistique.

Théorème 1.5. *Soit $(\mathfrak{X}, \mathfrak{A}, (P_\theta)_{\theta\in\Theta})$ un modèle paramétrique et $X = (X_1, \ldots, X_n) \sim f(X, \theta)$ un échantillon dans ce modèle. Une statistique $T(X)$ est exhaustive si, et seulement si, la densité $f(X, \theta)$ s'ecrit :*

$$f(X, \theta) = g(X)h(T(X), \theta)$$

où g et h sont des fonction mesurable et positive.

Exemple 1.6. soit $X = (X_1, \ldots, X_n) \sim f(X,\theta) = \prod_{i=1}^{n} f(X_i,\theta)$ où $\forall i = 1 : n$, $X_i \sim \mathcal{U}_{[0,\theta]}$ and $f(X_i,\theta) = \frac{1}{\theta}\mathbb{I}_{[0,\theta]}(x)$, on a

$$\begin{aligned}
f(X,\theta) = \prod_{i=1}^{n} f(X_i,\theta) &= \frac{1}{\theta^n}\prod_{i=1}^{n}\mathbb{I}_{[0,\theta]}(x_i) \\
&= \frac{1}{\theta^n}\prod_{i=1}^{n}\mathbb{I}_{\{0\le x_i\le\theta\}} \\
&= \frac{1}{\theta^n}\prod_{i=1}^{n}\mathbb{I}_{\{x_i\le\theta\}}\prod_{i=1}^{n}\mathbb{I}_{\{x_i\ge 0\}} \\
&= \prod_{i=1}^{n}\mathbb{I}_{\{x_i\ge 0\}}\frac{1}{\theta^n}\prod_{i=1}^{n}\mathbb{I}_{\{x_i\le\theta\}} \\
&= \mathbb{I}_{\{inf_i x_i\ge 0\}}\frac{1}{\theta^n}\mathbb{I}_{\{sup_i x_i\le\theta\}} \\
&= g(x)h(T(x),\theta).
\end{aligned}$$

Danc la statistique $T(X_1,\ldots,X_n) = sup_{1\le i\le n}X_i$ est une statistique exhaustive pour θ.

La statistique $T(X) = X$ est toujours une statistique exhaustive. Mais elle n'est pas d'un grand intérât et ne réduit absolument pas l'information. Il ne s'agit donc pas seulement de trouver une statistique exhaustive mais plutôt de trouver parmi les statistiques exhaustives celle(s) qui réduit(sent) au maximum l'information. En d'autres termes, le problème est de trouver une statistique exhaustive qui soit minimale.

Définition 1.24. On dit qu'une statistique exhaustive est minimale, si elle est une fonction mesurable de toutes les autres statistiques exhaustives.

Autrement dit, la statistique T est minimale si pour toute statistique exhaustive S il existe une fonction h telle que $T = h(S)$.

Théorème 1.6. ***(Théorème de Rao-Blackwell)*** *Soit $(\mathfrak{X}, \mathfrak{A}, (P_\theta)_{\theta\in\Theta})$ un modèle paramétrique et $X = (X_1, \ldots, X_n)$ un échantillon dans ce modèle. Soit*

$T(X)$ un estimateur de θ de carré intégrable. Si le modèle possède une statistique exhaustive $S(X)$ pour le paramètre θ, alors l'estimateur $E_\theta(T(X)|S(X))$ de θ a un risque quadratique inférieur à $T(X)$, c'est à dire que l'on a :

$$R(E_\theta(T(X)|S(X)),\theta) \leq R(T(X),\theta),$$

pour tout θ dans Θ. De plus cette inégalité est stricte pour au moins un θ de Θ, i.e. $E_\theta(T(X)|S(X))$ est préférable à $T(X)$, sauf si $T(X)$ est sans biais et une fonction de la statistique exhaustive $S(X)$. Si $T(X)$ est un estimateur sans biais de θ alors $E_\theta(T(X)|S(X))$ est également sans biais pour θ et l'inégalité sur les risques quadratiques se traduit également sur les variances.

Le théorème précédent nous permet déjà d'améliorer la qualité d'un estimateur. Mais il ne nous assure pas de tomber sur un estimateur optimal. L'obtention directe d'un estimateur optimal sera possible grâce au Théorème de Lehmann-Scheffé donné ci-dessous. Mais il nous faut auparavant introduire la notion de statistique complète qu'il utilise.

Statistique complète

Définition 1.25. Soit $(\mathfrak{X},\mathfrak{A},(P_\theta)_{\theta\in\Theta})$ un modèle paramétrique et $X=(X_1,\ldots,X_n)$ un échantillon dans ce modèle. Une statistique $T(X)$ est dite complète (ou totale) si toute fonction borélienne φ vérifiant

$$E_\theta|\varphi(T(X))| < +\infty \;\; et \;\; E_\theta(\varphi(T(X))) = 0$$

pour tout θ de Θ est nécessairement telle que

$$\varphi(T(X)) = 0,\; P_\theta - p.s.$$

pour tout θ de Θ

Théorème 1.7. *Toute statistique exhaustive et complète est minimale.*

Théorème 1.8. ***(Théorème de Lehmann-Scheffé)*** *Soit $(\mathfrak{X},\mathfrak{A},(P_\theta)_{\theta\in\Theta})$ un modèle paramétrique et $X=(X_1,\ldots,X_n)$ un échantillon dans ce modèle. Soit $T(X)$ un estimateur de θ de carré intégrable et $S(X)$ une statistique exhaustive et complète de θ. Alors l'estimateur amélioré de Rao-Blackwell $E_\theta(T(X)|S(X))$ est optimal dans la classe des estimateurs sans biais de θ.*

Cas des familles exponentielles

Définition 1.26. Soit $(\mathfrak{X}, \mathfrak{A}, (P_\theta)_{\theta\in\Theta})$ un modèle paramétrique et $X = (X_1, \ldots, X_n)$ un échantillon dans ce modèle. La famille des loi $(P_\theta)_{\theta\in\Theta}$ est dit famille exponentielle si P_θ admet une densité $f(x,\theta)$ et $f(x,\theta)$ admet la représentation suivante :

$$f(x,\theta) = exp\left[a(x)\alpha(\theta) + b(x) + \beta(\theta)\right] \quad x \in \mathbb{R}$$

$$\prod_{i=1}^{n} f(x_i,\theta) = exp\left(\sum_{i=1}^{n} a(x_i)\alpha(\theta) + \sum_{i=1}^{n} b(x_i) + n\beta(\theta)\right) \quad x \in \mathbb{R}^n$$

où a, b, α et β sont des fonction mesurables.

Exemple 1.7. Soit $X \sim p_\theta = b(m,\theta)$, i.e :

$$p_\theta(k) = p(X = k) = C_m^k \theta^k (1-\theta)^{m-k}.$$

on a

$$\begin{aligned} lnp(X = k) &= ln\left(C_m^k \theta^k (1-\theta)^{m-k}\right) \\ &= lnC_m^k + kln\left(\frac{\theta}{1-\theta}\right) + mln(1-\theta). \end{aligned}$$

Donc $(b(m,\theta))_{\theta\in[0,1]}$ est une famille exponentielle avec $a(k) = k$, $\alpha(\theta) = ln\left(\frac{\theta}{1-\theta}\right)$, $b(k) = lnC_m^k$ et $\beta(\theta) = mln(1-\theta)$.

Théorème 1.9. *(**Théorème de Darmois-Koopmans**) Soit $(\mathfrak{X}, \mathfrak{A}, (P_\theta)_{\theta\in\Theta})$ un modèle paramétrique dont le domaine des valeurs ne dépend pas de θ et $X = (X_1, \ldots, X_n)$ un échantillon dans ce modèle. Alors : il existe une statistique exhaustive de θ si et seulement si la famille $(P_\theta)_{\theta\in\Theta}$ est exponentielle. De plus $T(X) = \sum_{i=1}^{n} a(X_i)$ est la statistique exhaustive.*

Exemple 1.8. Soit $X \sim p_\theta = b(m,\theta)$, i.e :

$$p_\theta(k) = p(X = k) = C_m^k \theta^k (1-\theta)^{m-k}.$$

on a $T(X) = \sum_{i=1}^{n} X_i$ est une statistique exhaustive de θ.

Chapitre 2

Minimaxité

Sommaire

2.1 Inadmissibilité de l'stimateur usuel

Soit X une variable aléatoire qui suit une loi normale multidimentionelle $N_p(\theta, I_p)$, avec $\theta \in \mathbb{R}^p$. Pour tout estimateur $\delta(X)$ de θ, on définit la fonction de coût quadratique par :

$$L(\delta(X), \theta) = \|\delta(X) - \theta\|_p^2$$

où $\|.\|_p$ est la norme usuelle dans $\mathbb{R}^p$: Ainsi son risque quadratique est :

$$\begin{aligned} R(\delta(X), \theta) &= \mathbb{E}_\theta(L(\delta(X), \theta)) \\ &= \mathbb{E}_\theta(\|\delta(X) - \theta\|_p^2) \\ &= \int_{\mathbb{R}^p} \frac{\|\delta(X) - \theta\|_p^2}{(2\pi)^{p/2}} \exp\left(-\frac{1}{2}\|x - \theta\|_p^2\right) dx. \end{aligned}$$

On sait que l'estimateur usuel de θ est $\delta_0(X) = X$, il est minimax et son risque quadratique est :

$$R(\delta_0(X), \theta) = \mathbb{E}(\|\delta_0(X) - \theta\|^2) = p$$

en effet :

$$X - \theta \sim N_p(0, I_p)$$

donc

$$\|X - \theta\|^2 = \langle X - \theta, X - \theta \rangle = \sum_{i=1}^{p} (X_i - \theta_i)^2 \sim \mathcal{X}_p^2$$

car pour tout $i \in \{1, ..., p\}$ la variable aléatoire réelle $X_i - \theta_i \sim N(0,1)$ et les variables X_i sont indépendantes. Donc

$$\mathbb{E}(\delta_0(X)) = \mathbb{E}(\mathcal{X}_p^2) = p$$

Il est clair que l'estimateur usuel X est admisible pour $p \leq 2$. Stein a anoncé que quand $p \geq 3$ l'estimateur de la forme :

$$\delta_{a,b}^{JS}(X) = \left(1 - \frac{a}{b + \|X\|^2}\right) X$$

a un risque uniformément inférieur au risque de $\delta_0(X)$, pour a suffisamment petit et b suffisamment grand.

2.2 Estimateur de James -Stein

Lemme 2.1. *(**Stein**[9]) Si $Y \sim N(0,1)$, alors pour toute fonction dérivable h, telle que $|\mathbb{E}(h'(Y))| < \infty$ alors :*

$$\mathbb{E}[Y h(Y)] = \mathbb{E}[h'(Y)]$$

Démonstration : On pose

$$f_Y(y) = \frac{1}{2\pi} \exp(-\frac{1}{2} y^2)$$

la densité de la loi normale centrée réduite, et par dérivation on trouve $f'_Y(y) = -yf_Y(y)$ Alors

$$\begin{aligned}
\mathbb{E}[h'(Y)] &= \int_{-\infty}^{+\infty} h'(y)f(y)dy \\
&= -\int_{0}^{+\infty} h'(y)\left(\int_{y}^{+\infty} -zf(z)dz\right)dy + \int_{-\infty}^{0} h'(y)\left(\int_{-\infty}^{y} -zf(z)dz\right)dy \\
&= \int_{0}^{+\infty} h'(y)\left(\int_{y}^{+\infty} zf(z)dz\right)dy - \int_{-\infty}^{0} h'(y)\left(\int_{-\infty}^{y} zf(z)dz\right)dy \\
&= \int_{0}^{+\infty} zf(z)\left(\int_{0}^{z} h'(y)dy\right)dz - \int_{-\infty}^{0} zf(z)\left(\int_{z}^{0} h'(y)dy\right)dz \text{ (d'aprés Fubini)} \\
&= \int_{0}^{+\infty} zf(z)[h(z)-h(0)]dz + \int_{-\infty}^{0} zf(z)[h(z)-h(0)]dz \\
&= \left(\int_{0}^{+\infty} + \int_{-\infty}^{0}\right)\{zf(z)[h(z)-h(0)]\}dz \\
&= \int_{-\infty}^{+\infty} zh(z)f(z)dz - h(0)\int_{-\infty}^{+\infty} zf(z)dz \\
&= \int_{-\infty}^{+\infty} zh(z)f(z)dz - h(0)\mathbb{E}(Y) \\
&= \int_{-\infty}^{+\infty} zh(z)f(z)dz \text{ (car}\mathbb{E}(Y)=0).
\end{aligned}$$

D'où

$$\mathbb{E}(h'(Y)) = \mathbb{E}(Yh(Y)).$$

■

Corollaire 2.1. *Si $X \sim \mathcal{N}(\mu, 1)$, alors pour toute fonction dérivable h, telle que $|\mathbb{E}(h'(X))| < \infty$ on a :*

$$\mathbb{E}((X-\mu)h(X)) = \mathbb{E}(h'(X))$$

Démonstration : On pose $Y = X - \mu$, alors $Y \sim N(0,1)$, et donc d'aprés le lemme 2.1 on trouve

$$\begin{aligned}
\mathbb{E}((X-\mu)h(X)) &= \mathbb{E}(Yh(Y+\mu)) \\
&= \mathbb{E}\left(\frac{\partial}{\partial Y}h(Y+\mu)\right) \\
&= \mathbb{E}(h'(Y+\mu)) = \mathbb{E}(h'(X)).
\end{aligned}$$

■

Lemme 2.2. *Soit $X \sim N_p(\theta, I_p)$, alors*

$$\mathbb{E}\left(\frac{1}{\|X^2\|}\right) = \mathbb{E}\left(\frac{1}{p-2+2K}\right)$$

où $K \sim \mathcal{P}\left(\frac{\|\theta\|^2}{2}\right)$ est la loi de Poisson du paramètre $\frac{\|\theta\|^2}{2}$.

Démonstration : On pose $U = \|X\|^2$. Il est clair que $U \sim \chi_p^2(\lambda = \|\theta\|^2)$, alors :

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{U}\right) &= \sum_{k=0}^{+\infty}\left[\int_0^{+\infty}\frac{1}{u}\chi^2_{p+2k}du\right]\exp(-\lambda)\frac{\lambda^k}{k!}\\
&= \sum_{k=0}^{+\infty}\left[\frac{(\frac{1}{2})^{\frac{p+2k}{2}}}{\Gamma(\frac{p+2k}{2})}\int_0^{+\infty}\frac{1}{u}\exp(\frac{-1}{2}u)u^{\frac{p+2k}{2}-1}du\right]\exp(-\lambda)\frac{\lambda^k}{k!}\\
&= \sum_{k=0}^{+\infty}\left[\frac{(\frac{1}{2})^{\frac{p+2k}{2}}}{\Gamma(\frac{p+2k}{2})}\int_0^{+\infty}\exp(\frac{-1}{2}u)u^{\frac{p+2k}{2}-2}du\right]\exp(-\lambda)\frac{\lambda^k}{k!}.
\end{aligned}$$

Posons $t = \frac{1}{2}u \Leftrightarrow 2t = u$ et $du = 2dt$ alors on a :

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{U}\right) &= \sum_{k=0}^{+\infty}\left[\frac{\left(\frac{1}{2}\right)^{\frac{p+2k}{2}}}{\Gamma\left(\frac{p+2k}{2}\right)}\int_0^{+\infty}\exp(-t)(2t)^{\frac{p+2k}{2}-2}2dt\right]\exp(-\lambda)\frac{\lambda^k}{k!}\\
&= \sum_{k=0}^{+\infty}\left[\frac{\left(\frac{1}{2}\right)^{\frac{p+2k}{2}}}{\Gamma\left(\frac{p+2k}{2}\right)}2^{\frac{p+2k}{2}-2}2\int_0^{+\infty}\exp(-t)t^{\frac{p+2k}{2}-2}dt\right]\exp(-\lambda)\frac{\lambda^k}{k!}\\
&= \sum_{k=0}^{+\infty}\left[\frac{\frac{1}{2}}{\Gamma(\frac{p+2k}{2})}\Gamma\left(\frac{p+2k}{2}-1\right)\right]\exp(-\lambda)\frac{\lambda^k}{k!}\\
&= \sum_{k=0}^{+\infty}\left[\frac{\frac{1}{2}\Gamma\left(\frac{p+2k}{2}-1\right)}{\Gamma\left(\frac{p+2k}{2}-1+1\right)}\right]\exp(-\lambda)\frac{\lambda^k}{k!}\\
&= \sum_{k=0}^{+\infty}\left[\frac{\frac{1}{2}\Gamma\left(\frac{p+2k}{2}-1\right)}{\frac{p+2k-2}{2}\Gamma\left(\frac{p+2k}{2}-1\right)}\right]\exp(-\lambda)\frac{\lambda^k}{k!}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{+\infty}\left[\frac{1}{p-2+2k}\right]\exp(-\lambda)\frac{\lambda^k}{k!} \\
&= \mathbb{E}\left(\frac{1}{p-2+2K}\right).
\end{aligned}
$$

où $K \sim \mathcal{P}\left(\frac{\|\theta\|^2}{2}\right)$ est la loi de Poisson de paramètre $\frac{\|\theta\|^2}{2}$. ■

Lemme 2.3. *Soit $U \sim \chi_p^2(\lambda)$ la loi du Khi-deux décentrée avec p degrés de liberté et le paramètre de décentrage λ alors,*

i) pour tout nombre réel s et r où $-\frac{p}{2} < s \le r < 0$, la fonction réel

$$H_{p,r,s}(\lambda) = \frac{E(U^r)}{E(U^s)} = \frac{\int_{R_+} x^r \chi_p^2(\lambda; dx)}{\int_{R_+} x^s \chi_p^2(\lambda; dx)}$$

est non décroissante sur λ.

ii) De plus, si $X \sim N_p\left(\theta, I_p\right)$, on a

$$\sup_{\|\theta\|}\left(\frac{E(\|X\|^{-2r+2})}{E(\|X\|^{-r})}\right) = 2^{\frac{-r+2}{2}}\frac{\Gamma(\frac{p}{2}-r+1)}{\Gamma(\frac{p-r}{2})}.$$

Démonstration : : On montre d'abord que, pour tout réel v

$$\frac{\partial}{\partial\lambda}E(U^v) = \frac{\partial}{\partial\lambda}\int_{R_+} x^v \chi_p^2(\lambda; dx) = v2^{v-1}\sum_{k=0}^{+\infty}\frac{\Gamma(\frac{p}{2}+v+k)}{\Gamma(\frac{p}{2}+1+k)}P\left(\frac{\lambda}{2}; dk\right),$$

où $P(\frac{\lambda}{2})$ étant la distribution de Poisson du paramètre $\frac{\lambda}{2}$.
Utilisant la formule (1.1) nous avons, pour tout réel v

$$E(U^v) = E[(\chi_p^2(\lambda))^v] = E[(\chi_{p+2K}^2)^v] = 2^v E\left[\frac{\Gamma(\frac{p}{2}+K+v)}{\Gamma(\frac{p}{2}+K)}\right], \tag{2.1}$$

où $K \sim P(\frac{\lambda}{2})$. Alors

$$\begin{aligned}
\frac{\partial}{\partial\lambda}E(U^v) &= \frac{\partial}{\partial\lambda}\int_{R_+} x^v \chi_p^2(\lambda;dx) \\
&= 2^v \sum_{k=0}^{+\infty}\left[\frac{\Gamma(\frac{p}{2}+k+v)}{\Gamma(\frac{p}{2}+k)}\right]\frac{1}{k!}\frac{\partial}{\partial\lambda}\left[\left(\frac{\lambda}{2}\right)^k exp\left(-\frac{\lambda}{2}\right)\right] \\
&= 2^{v-1} \sum_{k=0}^{+\infty}\left[\frac{\Gamma(\frac{p}{2}+k+v)}{\Gamma(\frac{p}{2}+k)}\right]\frac{1}{k!}exp\left(-\frac{\lambda}{2}\right)\left[-\left(\frac{\lambda}{2}\right)^k + k\left(\frac{\lambda}{2}\right)^{k-1}\right] \\
&= 2^{v-1}exp\left(-\frac{\lambda}{2}\right)\left\{-\sum_{k=0}^{+\infty}\left[\frac{\Gamma(\frac{p}{2}+k+v)}{\Gamma(\frac{p}{2}+k)}\right]\frac{1}{k!}\left(\frac{\lambda}{2}\right)^k\right\} \\
&+ 2^{v-1}exp\left(-\frac{\lambda}{2}\right)\left\{\sum_{k=0}^{+\infty}\left[\frac{\Gamma(\frac{p}{2}+k+v+1)}{\Gamma(\frac{p}{2}+k+1)}\right]\frac{1}{k!}\left(\frac{\lambda}{2}\right)^k\right\} \\
&= 2^{v-1}exp\left(-\frac{\lambda}{2}\right)\left\{\sum_{k=0}^{+\infty}\frac{1}{k!}\left(\frac{\lambda}{2}\right)^k\left[\frac{\Gamma(\frac{p}{2}+k+v)}{\Gamma(\frac{p}{2}+k+1)}\right]\left[-\left(\frac{p}{2}+k\right)+\left(\frac{p}{2}+v+k\right)\right]\right\} \\
&= v2^{v-1}\sum_{k=0}^{+\infty}\frac{\Gamma(\frac{p}{2}+v+k)}{\Gamma(\frac{p}{2}+1+k)}P\left(\frac{\lambda}{2};dk\right).
\end{aligned}$$

Soit la fonction

$$\begin{aligned}
K_{p,r,s}(\lambda) &= \left(\frac{\partial}{\partial\lambda}\int_{R_+} x^r\chi_p^2(\lambda;dx)\right)\left(\int_{R_+} x^s\chi_p^2(\lambda;dx)\right) \\
&- \left(\frac{\partial}{\partial\lambda}\int_{R_+} x^s\chi_p^2(\lambda;dx)\right)\left(\int_{R_+} x^r\chi_p^2(\lambda;dx)\right).
\end{aligned}$$

Pour que la fonction $H_{p,r,s}$ soit strictement croissante, il suffit que la fonction $K_{p,r,s}$ prend des valeurs positives. De l'égalité (2.1), on obtient

$$\begin{aligned}
K_{p,r,s}(\lambda) &= 2^{r+s-1}r\sum_{i=0}^{+\infty}\sum_{j=0}^{+\infty}\frac{\Gamma(\frac{p}{2}+r+i)}{\Gamma(\frac{p}{2}+i+1)}\frac{\Gamma(\frac{p}{2}+s+j)}{\Gamma(\frac{p}{2}+j)}P\left(\frac{\lambda}{2};di\right)P\left(\frac{\lambda}{2};dj\right) \\
&- 2^{r+s-1}s\sum_{i=0}^{+\infty}\sum_{j=0}^{+\infty}\frac{\Gamma(\frac{p}{2}+r+j)}{\Gamma(\frac{p}{2}+j)}\frac{\Gamma(\frac{p}{2}+s+i)}{\Gamma(\frac{p}{2}+i+1)}P\left(\frac{\lambda}{2};dj\right)P\left(\frac{\lambda}{2};di\right).
\end{aligned}$$

Comme, $r > s$ alors

$$K_{p,r,s}(\lambda) \geq r2^{r+s-1}\sum_{i=0}^{+\infty}\sum_{j=0}^{+\infty}l_{p,r,s}(i,j)P\left(\frac{\lambda}{2};di\right)P\left(\frac{\lambda}{2};dj\right),$$

où

$$l_{p,r,s}(i,j) = \frac{\Gamma(\frac{p}{2}+r+i)\Gamma(\frac{p}{2}+s+j) - \Gamma(\frac{p}{2}+r+j)\Gamma(\frac{p}{2}+s+i)}{\Gamma(\frac{p}{2}+i+1)\Gamma(\frac{p}{2}+j)}.$$

Nous notons que, pour tout i, $l_{p,r,s}(i,j) = 0$; alors nous avons

$$K_{p,r,s}(i,j) \geq r2^{r+s-1}\sum_{i=0}^{+\infty}\sum_{j>i}^{+\infty}(l_{p,r,s}(i,j)+l_{p,r,s}(j,i))P\left(\frac{\lambda}{2};di\right)P\left(\frac{\lambda}{2};dj\right).$$

Mais si $i<j$, on a

$$\begin{aligned}
l_{p,r,s}(i,j)+l_{p,r,s}(j,i) &= \left(\Gamma\left(\frac{p}{2}+r+i\right)\Gamma\left(\frac{p}{2}+s+j\right)-\Gamma\left(\frac{p}{2}+r+j\right)\Gamma\left(\frac{p}{2}+s+i\right)\right)\\
&\times \left[\frac{1}{\Gamma(\frac{p}{2}+i+1)\Gamma(\frac{p}{2}+j)} - \frac{1}{\Gamma(\frac{p}{2}+j+1)\Gamma(\frac{p}{2}+i)}\right]\\
&= \frac{\Gamma(\frac{p}{2}+r+i)\Gamma(\frac{p}{2}+s+i)}{\Gamma(\frac{p}{2}+i)\Gamma(\frac{p}{2}+j)}\left[\frac{1}{\frac{p}{2}+i}-\frac{1}{\frac{p}{2}+j}\right]\\
&\times \left[\prod_{t=0}^{j-i-1}\left(\frac{p}{2}+s+i+t\right)-\prod_{t=0}^{j+i-1}\left(\frac{p}{2}+r+i+t\right)\right]\\
&\leq 0,
\end{aligned}$$

car pour tout t, $\frac{p}{2}+s+i+t<\frac{p}{2}+r+i+t$. Comme en hypothèse $r<0$, on a $K_{p,r,s}(\lambda)>0$. Ainsi, nous obtenons le résultat souhaité.

ii) En utilisant i) il est clair que la fonction $H^1_{p,r}(\lambda) = \frac{E(\|X\|^{-r})}{E(\|X\|^{-2r+2})}$ est non décroissante sur λ, alors la fonction $\frac{1}{H^1_{p,r}(\lambda)}$ n'est pas croissant sur λ, Donc

$$\begin{aligned}
\sup_{\|\theta\|}\left(\frac{E(\|X\|^{-2r+2})}{E(\|X\|^{-r})}\right) &= \sup_{\|\theta\|}\left(\frac{1}{H^1_{p,r}(\lambda)}\right)\\
&= \frac{1}{H^1_{p,r}(0)}\\
&= 2^{\frac{-r+2}{2}}\frac{\Gamma(\frac{p}{2}-r+1)}{\Gamma(\frac{p-r}{2})}.
\end{aligned}$$

■

On considére l'estimateur

$$\delta_{a,r} = \left(1-a\frac{1}{\|X\|^r}\right)X = X - a\frac{1}{\|X\|^r}X, \tag{2.2}$$

où $2 \leq r < \frac{p+2}{2}$ et la constante réelle positive a peut dépendre de p.

Proposition 2.1. La fonction de risque de l'estimateur $\delta_{a,r}$ donné en (2.2) sous la fonction de perte L, est

$$\begin{aligned} R(\delta_{a,r},\theta) &= p - 2a(p-r)E\left(\frac{1}{\|X\|^r}\right) \\ &+ a^2 E\left(\frac{1}{\|X\|^{2r-2}}\right). \end{aligned}$$

Démonstration : On a :

$$\begin{aligned} R(\delta_{a,r},\theta) &= \mathbb{E}(\|\delta_{a,r}-\theta\|^2) \\ &= \mathbb{E}\left\|\left(1-\frac{a}{\|X\|^r}\right)X-\theta\right\|^2 \\ &= \mathbb{E}\left(\left\|X-\theta-\frac{a}{\|X\|^r}X\right\|^2\right) \\ &= \mathbb{E}\left\langle X-\theta-\frac{a}{\|X\|^r}X, X-\theta-\frac{a}{\|X\|^r}X\right\rangle \\ &= \mathbb{E}\left[\|X-\theta\|^2+\left\|\frac{a}{\|X\|^r}X\right\|^2-2\left\langle X-\theta,\frac{a}{\|X\|^r}X\right\rangle\right] \\ &= \mathbb{E}(\|X-\theta\|^2)+\mathbb{E}\left(\left\|\frac{a}{\|X\|^r}X\right\|^2\right)-2\mathbb{E}\left\langle X-\theta,\frac{a}{\|X\|^r}X\right\rangle \\ &= p+\mathbb{E}\left\langle\frac{a}{\|X\|^r}X,\frac{a}{\|X\|^r}X\right\rangle-2\mathbb{E}\left\langle X-\theta,\frac{a}{\|X\|^r}X\right\rangle \\ &= p+\mathbb{E}\left[\frac{a^2}{\|X\|^{2r}}\|X\|^2\right]-2\mathbb{E}\left[\frac{a}{\|X\|^r}\langle X-\theta,X\rangle\right] \\ &= p+a^2\mathbb{E}\left(\frac{1}{\|X\|^{2r-2}}\right)-2a\mathbb{E}\left[\sum_{i=1}^{p}(X_i-\theta_i)\frac{X_i}{\|X\|^r}\right]. \end{aligned}$$

D'aprés le Corollaire 2.1, on a :

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{p}(X_i-\theta_i)\frac{X_i}{\|X\|^r}\right] &= \sum_{i=1}^{p}\mathbb{E}\left[(X_i-\theta_i)\frac{X_i}{\|X\|^r}\right] \\
&= \sum_{i=1}^{p}\mathbb{E}\left(\frac{\partial}{\partial X_i}\left[\frac{X_i}{\|X\|^r}\right]\right) \\
&= \sum_{i=1}^{p}\mathbb{E}\left(\frac{\partial}{\partial X_i}\left[\frac{X_i}{X_1^r+X_2^r+....+X_p^r}\right]\right) \\
&= \sum_{i=1}^{p}\mathbb{E}\left[\frac{\|X\|^r-rX_i^r}{\|X\|^{2r}}\right] \\
&= \sum_{i=1}^{p}\mathbb{E}\left[\frac{1}{\|X\|^r}-\frac{rX_i^r}{\|X\|^{2r}}\right] \\
&= \mathbb{E}\sum_{i=1}^{p}\frac{1}{\|X\|^r}-r\mathbb{E}\left[\sum_{i=1}^{p}\frac{X_i^r}{\|X\|^{2r}}\right] \\
&= p\mathbb{E}\left[\frac{1}{\|X\|^r}\right]-r\mathbb{E}\left[\frac{1}{\|X\|^{2r}}\sum_{i=1}^{p}X_i^r\right] \\
&= p\mathbb{E}\left[\frac{1}{\|X\|^r}\right]-r\mathbb{E}\left[\frac{1}{\|X\|^r}\right]=(p-r)\mathbb{E}\left[\frac{1}{\|X\|^r}\right].
\end{aligned}$$

D'où

$$R(\delta_{a,r},\theta)=p+a^2\mathbb{E}\left[\frac{1}{\|X\|^{2r-2}}\right]-2(p-r)a\mathbb{E}\left[\frac{1}{\|X\|^r}\right]$$

■

Théorème 2.1. *Soit l'estimateur $\delta_{a,r}$ défini en (2.2).*

i) Une condition suffisante pour que l'estimateur $\delta_{a,r}$ domine E.M.V (donc c'est minimax), est

$$0\leq a\leq 2^{\frac{r}{2}}(p-r)\frac{\Gamma(\frac{p-r}{2})}{\Gamma(\frac{p-2r+2}{2})},$$

ii) la valeur optimale de a qui minimise la fonction de risque $R(\delta_{a,r},\theta)$, est

$$\widehat{a}=2^{\frac{r-2}{2}}(p-r)\frac{\Gamma(\frac{p-r}{2})}{\Gamma(\frac{p-2r+2}{2})}.$$

Démonstration : i) En utilisant la proposition 2.1, nous avons

$$\begin{aligned} R(\delta_{a,r},\theta) &= p-2a(p-r)E\left(\frac{1}{\|X\|^r}\right) \\ &+ a^2\left(\frac{E\left(\frac{1}{\|X\|^{2r-2}}\right)}{E\left(\frac{1}{\|X\|^r}\right)}\right)E\left(\frac{1}{\|X\|^r}\right). \end{aligned}$$

Du lemme 2.3, on a

$$\begin{aligned} R(\delta_{a,r},\theta) &\leq p-2a(p-r)E\left(\frac{1}{\|X\|^r}\right) \\ &+ a^2 2^{-\frac{r-2}{2}}\frac{\Gamma(\frac{p-2r+2}{2})}{\Gamma(\frac{p-r}{2})}E\left(\frac{1}{\|X\|^r}\right) \\ &= p-2a(p-r)E\left(\frac{1}{\|X\|^r}\right) \\ &+ 2^{-\frac{r-2}{2}}a^2\frac{\Gamma(\frac{p-2r+2}{2})}{\Gamma(\frac{p-r}{2})}E\left(\frac{1}{\|X\|^r}\right). \end{aligned} \tag{2.3}$$

A partir du côté droit de la dernière égalité, il est facile de montrer qu'une condition suffisante pour que $R(\delta_{a,r},\theta) \leq R(X,\theta) = p$ et par conséquent $\delta_{a,r}$ domine l'E.M.V (donc c'est minimax), est

$$-2a(p-r)E\left(\frac{1}{\|X\|^r}\right)+2^{-\frac{r-2}{2}}a^2\frac{\Gamma(\frac{p-2r+2}{2})}{\Gamma(\frac{p-r}{2})}E\left(\frac{1}{\|X\|^r}\right) \leq 0,$$

qui équivaut à

$$2aE\left(\frac{1}{\|X\|^r}\right)\left[-(p-r)-2^{-\frac{r-2}{2}}a\frac{\Gamma(\frac{p-2r+2}{2})}{\Gamma(\frac{p-r}{2})}\right] \leq 0,$$

qui conduit à

$$0 \leq a \leq 2^{\frac{r}{2}}(p-r)\frac{\Gamma(\frac{p-r}{2})}{\Gamma(\frac{p-2r+2}{2})}.$$

ii) En utilisant la convexité sur a de la fonction donnée en côté droit de l'égalité (2.3), on peut facilement obtenir le résultat. La preuve du théorème est terminée. ∎

Pour $r = 2$, on note $\widehat{a}$ par $d := (p-2)$, on obtient l'estimateur de James-Stein

$$\delta_{JS} = \delta_{d,2} = \left(1 - d\frac{1}{\|X\|^2}\right)X. \tag{2.4}$$

Il découle de la proposition 2.1 que la fonction de risque de δ_{JS} est donné par

$$R(\delta_{JS},\theta) = p - (p-2)^2 E\left(\frac{1}{\|X\|^2}\right). \tag{2.5}$$

De la formule (2.5), on remarque que

$$R(\delta_{JS},\theta) \leq p = R(X,\theta),$$

alors δ_{JS} domine E.M.V X, donc c'est aussi minimax.

Chapitre 3

Estimateur de type Polynomial

Sommaire

Puisque l'estimateur $\delta_a = X - a\dfrac{1}{\|X\|^2}X$ domine l'E.M.V X pour certaines valeurs de a, nous pensons ajouter le terme $b\left(\dfrac{1}{\|X\|^2}\right)^2 X$ à l'estimateur de James-Stein δ_{JS} pour obtenir un estimateur plus performant que δ_{JS}, puis nous construisons les classes d'estimateurs à rétrécisseur qui dominent l'estimateur de James-Stein δ_{JS}. Notre idée principale est d'ajouter à chaque fois un terme de la forme $\gamma(1/\|X\|^2)^m X$ où γ est une constante réelle peut dépendre de p et le paramètre m est un entier, et on construit les estimateurs qui dominent les estimateurs de la classe définie précédemment. Ainsi, dans cette section, nous traitons des estimateurs à rétrécisseur de forme polynomiale avec l'indéterminé $1/\|X\|^2$.

TABLE DES MATIÈRES

3.1 Estimateurs améliorant l'estimateur de James-Stein

on considère l'estimateur

$$\begin{aligned}\delta_b^{(2)} &= \delta_{JS} + b\left(\frac{1}{\|X\|^2}\right)^2 X \\ &= X - (p-2)\frac{1}{\|X\|^2}X + b\left(\frac{1}{\|X\|^2}\right)^2 X, \qquad (3.1)\end{aligned}$$

où la constante réelle positive b peut dépendre de p.

Proposition 3.1. Sous la fonction de perte L, la fonction de risque de l'estimateur $\delta_b^{(2)}$ donné en (3.1), est

$$R(\delta_b^{(2)}, \theta) = R(\delta_{JS}, \theta) - 4bE\left(\frac{1}{\|X\|^4}\right) + b^2 E\left(\frac{1}{\|X\|^6}\right).$$

Démonstration : Utilisation de la fonction de risque, on a

$$\begin{aligned}R(\delta_b^{(2)}, \theta) &= E\left(\left\|\delta_{JS} + b\frac{1}{\|X\|^2}X - \theta\right\|^2\right) \\ &= E\left(\left\|\delta_{JS} - \theta\right\|^2 + b^2\frac{1}{(\|X\|^2)^3} + 2\left\langle \delta_{JS} - \theta, b\frac{1}{(\|X\|^2)^2}X\right\rangle\right) \\ &= R(\delta_{JS}, \theta) + b^2 E\left(\frac{1}{(\|X\|^2)^3}\right) \\ &+ 2E\left(\left\langle X - \theta - (p-2)\frac{1}{\|X\|^2}X, b\frac{1}{(\|X\|^2)^2}X\right\rangle\right) \\ &= R(\delta_{JS}, \theta) + b^2 E\left(\frac{1}{\|X\|^6}\right) \\ &+ 2b\sum_{i=1}^{p} E\left((X_i - \theta_i)\frac{X_i}{\|X\|^4}\right) - 2b(p-2)E\left(\frac{1}{\|X\|^4}\right).\end{aligned}$$

Utilisant Corollaire 2.1, on a

$$\begin{aligned}
\sum_{i=1}^{p} E\left((X_i-\theta_i)\frac{X_i}{\|X\|^4}\right) &= \sum_{i=1}^{p} E\left(\frac{\partial}{\partial X_i}\frac{X_i}{\|X\|^4}\right)\\
&= \sum_{i=1}^{p} E\left(\frac{1}{\|X\|^4}-4\frac{X_i^2}{\|X\|^6}\right)\\
&= (p-4)E\left(\frac{1}{\|X\|^4}\right).
\end{aligned}$$

Alors,

$$\begin{aligned}
R(\delta_b^{(2)},\theta) &= R(\delta_{JS},\theta)+b^2E\left(\frac{1}{\|X\|^6}\right)\\
&+ 2b(p-4)E\left(\frac{1}{\|X\|^4}\right)-2b(p-2)E\left(\frac{1}{\|X\|^4}\right)\\
&= R(\delta_{JS},\theta)-4bE\left(\frac{1}{\|X\|^4}\right)+b^2E\left(\frac{1}{\|X\|^6}\right).
\end{aligned}$$

■

Théorème 3.1. *Sous la fonction de perte L, l'estimateur $\delta_b^{(2)}$ avec*

$$b = 2(p-6),$$

domine l'estimateur de James-Stein δ_{JS}.

Démonstration : Utilisation de la proposition 3.1, on a

$$R(\delta_b^{(2)},\theta) = R(\delta_{JS},\theta)-4bE\left(\frac{1}{\|X\|^4}\right)+b^2\frac{E\left(\frac{1}{\|X\|^6}\right)}{E\left(\frac{1}{\|X\|^4}\right)}E\left(\frac{1}{\|X\|^4}\right).$$

De ii) du lemme 2.3, on obtient

$$\frac{E\left(\frac{1}{\|X\|^6}\right)}{E\left(\frac{1}{\|X\|^4}\right)} = \frac{E\left(\|X\|^{-6}\right)}{E\left(\|X\|^{-4}\right)} \leq 2^{\frac{-4+2}{2}}\frac{\Gamma(\frac{p}{2}-4+1)}{\Gamma(\frac{p-4}{2})} = \frac{1}{p-6}.$$

Alors,

$$R(\delta_b^{(2)},\theta) \leq R(\delta_{JS},\theta)-4bE\left(\frac{1}{\|X\|^4}\right)+b^2\frac{1}{(p-6)}E\left(\frac{1}{\|X\|^4}\right). \tag{3.2}$$

La valeur optimale pour b qui minimise le côté droit de la dernière inégalité, est

$$\widehat{b} = 2(p-6). \tag{3.3}$$

Si nous remplaçons b par $\widehat{b}$ dans l'inégalité (3.2), on a

$$\begin{aligned} R(\delta_{\widehat{b}}^{(2)}, \theta) &\leq R(\delta_{JS}, \theta) - 4(p-6)E\left(\frac{1}{\|X\|^4}\right) \\ &\leq R(\delta_{JS}, \theta). \end{aligned}$$

■

3.2 Estimateurs dominant l'estimateur $\delta_b^{(2)}$

On considère maintenant l'estimateur

$$\begin{aligned} \delta_c^{(3)} &= \delta_b^{(2)} + c\left(\frac{1}{\|X\|^2}\right)^3 X \\ &= X - (p-2)\frac{1}{\|X\|^2}X + b\left(\frac{1}{\|X\|^2}\right)^2 X + c\left(\frac{1}{\|X\|^2}\right)^3 X, \end{aligned} \tag{3.4}$$

où la constante réelle positive c peut dépendre de p.

Proposition 3.2. Sous la fonction de perte L, la fonction de risque de l'estimateur $\delta_c^{(3)}$ donné en (3.4), est

$$\begin{aligned} R(\delta_c^{(3)}, \theta) &= R(\delta_b^{(2)}, \theta) + c^2 E\left(\frac{1}{\|X\|^{10}}\right) + 4c(p-6)E\left(\frac{1}{\|X\|^8}\right) \\ &- 8cE\left(\frac{1}{\|X\|^6}\right). \end{aligned}$$

Démonstration : Utilisation de la fonction de risque, on obtient

$$\begin{aligned}
R(\delta_c^{(3)},\theta) &= E\left(\left\|\delta_{\widehat{b}}^{(2)} + c\left(\frac{1}{\|X\|^2}\right)^3 X - \theta\right\|^2\right) \\
&= E\left(\left\|\delta_{\widehat{b}}^{(2)} - \theta\right\|^2 + c^2\frac{1}{(\|X\|^2)^5} + 2\left\langle \delta_{\widehat{b}}^{(2)} - \theta, c\frac{1}{(\|X\|^2)^3}X\right\rangle\right) \\
&= R(\delta_{\widehat{b}}^{(2)},\theta) + c^2E\left(\frac{1}{(\|X\|^2)^5}\right) \\
&+ 2E\left\langle X - \theta - \widehat{a}\frac{1}{\|X\|^2}X + \widehat{b}\left(\frac{1}{\|X\|^2}\right)^2 X, c\left(\frac{1}{\|X\|^2}\right)^3 X\right\rangle \\
&= R(\delta_{\widehat{b}}^{(2)},\theta) + c^2E\left(\frac{1}{\|X\|^{10}}\right) - 2c\widehat{a}E\left(\frac{1}{\|X\|^6}\right) \\
&+ 2c\widehat{b}E\left(\frac{1}{\|X\|^8}\right) + 2c\sum_{i=1}^{p} E\left((X_i - \theta_i)\frac{X_i}{\|X\|^6}\right).
\end{aligned}$$

Utilisant Corollaire 2.1, on a

$$\begin{aligned}
\sum_{i=1}^{p} E\left((X_i - \theta_i)\frac{X_i}{\|X\|^6}\right) &= \sum_{i=1}^{p} E\left(\frac{\partial}{\partial X_i}\frac{X_i}{\|X\|^6}\right) \\
&= (p-6)E\left(\frac{1}{\|X\|^6}\right).
\end{aligned}$$

Alors

$$\begin{aligned}
R(\delta_c^{(3)},\theta) &= R(\delta_{\widehat{b}}^{(2)},\theta) + c^2E\left(\frac{1}{\|X\|^{10}}\right) - 2c(p-2)E\left(\frac{1}{\|X\|^6}\right) \\
&+ 4c(p-6)E\left(\frac{1}{\|X\|^8}\right) + 2c(p-6)E\left(\frac{1}{\|X\|^6}\right) \\
&= R(\delta_{\widehat{b}}^{(2)},\theta) + c^2E\left(\frac{1}{\|X\|^{10}}\right) + 4c(p-6)E\left(\frac{1}{\|X\|^8}\right) \\
&- 8cE\left(\frac{1}{\|X\|^6}\right).
\end{aligned}$$

■

Théorème 3.2. *Sous la fonction de perte L, l'estimateur $\delta_c^{(3)}$ avec*

$$c = 2(p-10)^2,$$

domine l'estimateur $\delta_b^{(2)}$.

Démonstration : En utilisant la dernière proposition, nous avons

$$\begin{aligned}
R(\delta_c^{(3)},\theta) &= R(\delta_{\widehat{b}}^{(2)},\theta)+c^2\frac{E\left(\frac{1}{\|X\|^{10}}\right)}{E\left(\frac{1}{\|X\|^{6}}\right)}E\left(\frac{1}{\|X\|^6}\right)\\
&+ 4c(p-6)\frac{E\left(\frac{1}{\|X\|^{8}}\right)}{E\left(\frac{1}{\|X\|^{6}}\right)}E\left(\frac{1}{\|X\|^6}\right)\\
&- 8cE\left(\frac{1}{\|X\|^6}\right).
\end{aligned}$$

De ii) du lemme 2.3, on obtient

$$\frac{E\left(\frac{1}{\|X\|^{10}}\right)}{E\left(\frac{1}{\|X\|^{6}}\right)}=\frac{E\left(\|X\|^{-10}\right)}{E\left(\|X\|^{-6}\right)}\leq 2^{\frac{-6+2}{2}}\frac{\Gamma(\frac{p}{2}-6+1)}{\Gamma(\frac{p-6}{2})}=\frac{1}{(p-8)(p-10)},$$

et de i) du lemme 2.3, on a

$$\begin{aligned}
\frac{E\left(\frac{1}{\|X\|^{8}}\right)}{E\left(\frac{1}{\|X\|^{6}}\right)} &= \frac{E\left((\chi_p^2(\lambda))^{-4}\right)}{E\left((\chi_p^2(\lambda))^{-3}\right)}\\
&\leq \frac{E\left((\chi_p^2)^{-4}\right)}{E\left((\chi_p^2)^{-3}\right)}=\frac{2^{-4}\frac{\Gamma(\frac{p}{2}-4)}{\Gamma(\frac{p}{2})}}{2^{-3}\frac{\Gamma(\frac{p}{2}-3)}{\Gamma(\frac{p}{2})}}=\frac{1}{p-8}.
\end{aligned}$$

Alors,

$$\begin{aligned}
R(\delta_c^{(3)},\theta) &\leq R(\delta_{\widehat{b}}^{(2)},\theta)+c^2\frac{1}{(p-8)(p-10)}E\left(\frac{1}{\|X\|^6}\right)\\
&+ 4c(p-6)\frac{1}{p-8}E\left(\frac{1}{\|X\|^6}\right)-8cE\left(\frac{1}{\|X\|^6}\right)\\
&= R(\delta_{\widehat{b}}^{(2)},\theta)+c^2\frac{1}{(p-8)(p-10)}E\left(\frac{1}{\|X\|^6}\right)\\
&- 4c\frac{p-10}{p-8}E\left(\frac{1}{\|X\|^6}\right). \qquad (3.5)
\end{aligned}$$

La valeur optimale de c qui minimise le côté droit de l'inégalité (3.5), est

$$\widehat{c}=2(p-10)^2. \qquad (3.6)$$

Si nous remplaçons c par $\widehat{c}$ dans l'inégalité (3.5), on a

$$\begin{aligned} R(\delta_c^{(3)}, \theta) &\leq R(\delta_{\widehat{b}}^{(2)}, \theta) - 4\frac{(p-10)^3}{p-8} E\left(\frac{1}{\|X\|^6}\right) \\ &\leq R(\delta_{\widehat{b}}^{(2)}, \theta). \end{aligned}$$

■

Chapitre 4

Résultats de simulation

Dans ce chapitre, nous prenons le modèle $X \sim N_p\left(\theta, I_p\right)$ et on rappelle les estimateurs δ_{JS}, $\delta_b^{(2)}$ et $\delta_c^{(3)}$, i.e.,

$$\delta_{JS} = \left(1 - \frac{p-2}{\|X\|^2}\right)X,$$

$$\delta_b^{(2)} = X - (p-2)\frac{1}{\|X\|^2}X + 2(p-6)\left(\frac{1}{\|X\|^2}\right)^2 X$$

et

$$\delta_c^{(3)} = X - (p-2)\frac{1}{\|X\|^2}X + 2(p-6)\left(\frac{1}{\|X\|^2}\right)^2 X + 2(p-10)^2\left(\frac{1}{\|X\|^2}\right)^3 X.$$

On représente graphiquement les rapport de risques des estimateurs cités ci-dessus, par rapport au E.M.V associé aux fonctions de pertes L noté respectivement : $\dfrac{R(\delta_{JS},\theta)}{R(X,\theta)}$, $\dfrac{R(\delta_b^{(2)},\theta)}{R(X,\theta)}$ et $\dfrac{R(\delta_c^{(3)},\theta)}{R(X,\theta)}$ en fonction de $\lambda = \|\theta\|^2$ pour différentes valeurs de p.

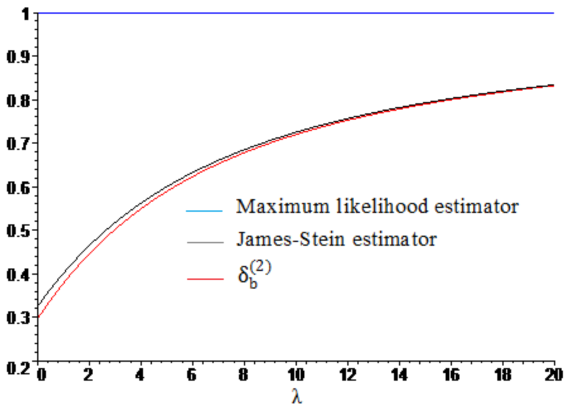

FIGURE 4.1 – Graphique des rapport de risques de $\dfrac{R(\delta_{JS},\theta)}{R(X,\theta)}$ et $\dfrac{R(\delta_b^{(2)},\theta)}{R(X,\theta)}$ comme fonctions de λ pour $p = 4$

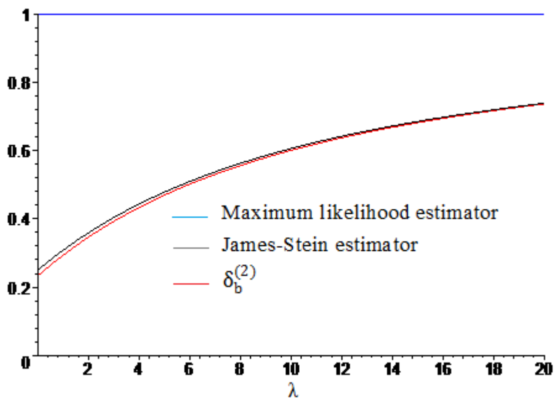

FIGURE 4.2 – Graphique des rapport de risques de $\dfrac{R(\delta_{JS},\theta)}{R(X,\theta)}$ et $\dfrac{R(\delta_b^{(2)},\theta)}{R(X,\theta)}$ comme fonctions de λ pour $p = 12$

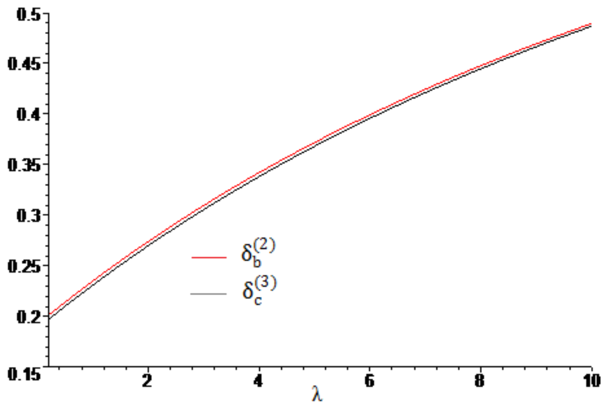

FIGURE 4.3 – Graphique des rapport de risques de $\dfrac{R(\delta_b^{(2)},\theta)}{R(X,\theta)}$ et $\dfrac{R(\delta_c^{(3)},\theta)}{R(X,\theta)}$ comme fonctions de λ pour $p = 18$

En Figure 4.1, Figure 4.2 et Figure 4.3, on remarque que les rapport de risques $\dfrac{R(\delta_{JS},\theta)}{R(X,\theta)}$, $\dfrac{R(\delta_b^{(2)},\theta)}{R(X,\theta)}$ et $\dfrac{R(\delta_c^{(3)},\theta)}{R(X,\theta)}$ sont inférieurs á 1, ainsi les estimateurs δ_{JS}, $\delta_b^{(2)}$ et $\delta_c^{(3)}$ sont minimax pour $p = 4$, $p = 12$ et $p = 18$. Nous remarquons aussi que, d'une part, plus p augment plus le gain augmente et d'autre part, plus la valeur de λ augmente, plus le gain diminue.

Dans le tableau suivants, nous donnons les valeurs des rapport de risques $\dfrac{R(\delta_{JS},\theta)}{R(X,\theta)}$, $\dfrac{R(\delta_b^{(2)},\theta)}{R(X,\theta)}$ et $\dfrac{R(\delta_c^{(3)},\theta)}{R(X,\theta)}$ pour des valeurs différentes de p et λ. La première entrée est $\dfrac{R(\delta_{JS},\theta)}{R(X,\theta)}$, la deusième entrée $\dfrac{R(\delta_b^{(2)},\theta)}{R(X,\theta)}$ et la troisième entrée $\dfrac{R(\delta_c^{(3)},\theta)}{R(X,\theta)}$.

TABLE 4.1 – Les valeurs des rapport de risques $R(\delta_{JS},\theta)/R(X,\theta)$, $R(\delta_b^{(2)},\theta)/R(X,\theta)$ et $R(\delta_c^{(3)},\theta)/R(X,\theta)$ comme fonctions de λ.

λ	rapport de risques	$p=14$	$p=18$	$p=24$
1.2418	δ_{JS}	0.2134	0.1688	0.1326
	$\delta_b^{(2)}$	0.2010	0.1608	0.1246
	$\delta_c^{(3)}$	0.1973	0.1563	0.1201
10.4311	δ_{JS}	0.5218	0.4535	0.3749
	$\delta_b^{(2)}$	0.5168	0.4418	0.3632
	$\delta_c^{(3)}$	0.5150	0.4390	0.3604
20.0000	δ_{JS}	0.6653	0.5923	0.5090
	$\delta_b^{(2)}$	0.6628	0.5901	0.5068
	$\delta_c^{(3)}$	0.6621	0.5888	0.5055

Dans le tableau précédent, on note que : si λ et p sont petits, le gain des rapports de risques $\frac{R(\delta_{JS},\theta)}{R(X,\theta)}$, $\frac{R(\delta_b^{(2)},\theta)}{R(X,\theta)}$ et $\frac{R(\delta_c^{(3)},\theta)}{R(X,\theta)}$ est très important. On observe également que, si les valeurs de p augmentent, le gain diminue et ce pour chaque valeur fixe de λ. On voit aussi que, si les valeurs de λ augmentent et p et fixée, les rapports de risques augmentent et le gain diminue.

Annexe

Proposition 4.1.

1. Soit $X_1, \dots, X_p$ une suites de variables aléatoires indépendantes et identiquement distrbuées, de loi $N(0,1)$ alors
$$\sum_{i=1}^{p} X_i^2 \sim \chi_p^2$$
et la loi du Chi-Deux (centré) à p degré de liberté.
2. Si $Y_1, \dots, Y_p$ des variables aléatoires indépendantes telle que $\forall i = 1, \dots, p$,$Y_i \sim N(\theta_i, 1)$ alors :
$$\sum_{i=1}^{p} Y_i^2 \sim \chi_p^2(\|\theta\|^2)$$
et la loi du Chi-Deux décentré à p degré de liberté et de paramère de décentrage $\|\theta\|^2$.

Théorème 4.1. ***(Inégalité de Jensen)*** *Soit f une fonction convexe sur un intervalle réel I, et X une variable aléatoire réel dont l'espérance $\mathbb{E}(X)$ existe. Alors :*
$$f(\mathbb{E}(X)) \leq \mathbb{E}[f(X)]$$

loi forte des grands nombres

Théorème 4.2. *Soit $(X_n)_{n>0}$ une suite de variables aléatoires indépendantes et identiquement distrbuées. On note $S_n = X_1 + \ldots + X_n$, on a :*
$$\left\{ \text{la suite} \frac{S_n}{n} \text{est convergente presque sûrement} \right\} \Leftrightarrow \{\mathbb{E}[X_1] < +\infty\}.$$

De plus, si l'une de ces deux conditions éqivalentes est remplie, on a :

$$\mathbb{P}\left(\omega \in \Omega / \lim_n \frac{S_n(\omega)}{n} = \mathbb{E}[X_1]\right) = 1.$$

conclusion

Dans ce travail, nous avons étudié les estimateurs de type James-Stein et de type polynomial de la moyenne θ, d'une loi normale multidimensionelle $N_p(\theta, I_p)$. Premièrement nous avons discuté l'inadmissibilité de l'estimateur usuel X quand la dimension de l'espace des paramètres $p \geq 3$. Ensuite nous avons présenté la classe des estimateurs de type James-Stein qui est une classe trés importantes des estimateurs biaisés bien sûr mais a un risque quadratique, uniformément meilleurs que celui de l'estimateur usuel X. Enfin nous avons développer des classes d'estimateurs qui domine l'estimateur de James-Stein.

Bibliographie

[1] A.J. Baranchik, (1964). *Multiple regression and estimation of the mean of a multivariate normal distribution*. Stanford Univ. Technical Report 51.

[2] A. Benkhaled, M. Terbeche and A. Hamdaoui (2021). Polynomials Shrinkage Estimators of a Multivariate Normal Mean. *Statistics, Optimization Information Computing*.

[3] D. Bernard, (2017). *Estimation paramétrique*, cours de Master 2. Université Rennes1, France.

[4] G. Casella and J.T. Hwang, (1982). *Limit expressions for the risk of James-Stein estimator*, Canad.J.Statist. 10(4), 305–309.

[5] W. James and C. Stein, (1961), *Estimation of quadratique loss Proc 4th Berkeley Symp*,Math. Statist.Prob. 1, 361–379.

[6] D. Jean-Yves. (2011–2012), CTU, *Licence de Mathématiques Statistique Inférentielle*, université de Franche-Comté, France.

[7] A. Hamdaoui, (2004), *Estimateur de James-Stein généralisé*, Mémoire de magister en mathématique, université Abou Bekr Belkaid, Tlemcen.

[8] C. Stein, (1956), *Inadmissibilty of the usual estimator for the mean of a multivariate normal distribution*. Proc 3th Berkeley Symp, Math.Statist.Prob., Vol.1, 197–206, Univ of california Press,Berkeley

[9] C. Stein, (1981), *Estimation of the mean of multivariate normal distribution*, Annals of Statistics, Vol.9, 1135–1151.