

STATISTIQUE ET ANALYSE DES DONNÉES

DOUKISSA CRITICOU

DIMITRIS TERZAKIS

**Une classe d'estimateurs à rétrécisseur bayésiens pour
la moyenne d'un vecteur normal**

Statistique et analyse des données, tome 16, n° 3 (1991), p. 1-23.

http://www.numdam.org/item?id=SAD_1991__16_3_1_0

© Association pour la statistique et ses utilisations, 1991, tous droits réservés.

L'accès aux archives de la revue « Statistique et analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

UNE CLASSE D'ESTIMATEURS A RETRECISSEUR BAYESIENS POUR LA MOYENNE D'UN VECTEUR NORMAL

Doukissa CRITICOU*, Dimitris TERZAKIS*

U.R.A. C.N.R.S. D 1378 - Université de Rouen, France

*Université de Crète, Département de Mathématiques
Heraklion - Crète - Grèce

Résumé.

On s'intéresse au traitement bayésien de l'estimation de la moyenne dans le modèle linéaire gaussien, quand la variance est connue à un facteur multiplicatif près, le coût utilisé est quadratique (inverse de la variance) et la probabilité a priori est un mélange de lois gaussiennes.

Les estimateurs de Bayes sont, dans ce cas, des estimateurs à rétrécisseurs, non nécessairement uniformément meilleurs que l'estimateur des moindres carrés. Dans cet article on présente des estimateurs qui dominent l'estimateur des moindres carrés tout en minimisant un risque bayésien au sein d'une certaine classe. Ces estimateurs apparaissent en fait comme des estimateurs bayésiens tronqués.

Mots clés.

estimateurs de James-Stein, estimateurs de Bayes, fonction de rétrécissement.

Abstract.

We are interested in the Bayesian estimation of the mean in a gaussian linear model, when the variance is unknown up to a multiplicative factor, the loss (the inverse of variance) is quadratic and the prior probability is a mixture of gaussian distributions.

In this case, the Bayesian estimators are shrinkage estimators, which are not necessarily better than the least squares estimator. In this paper we present the estimators which dominate the least squares estimator while minimizing the Bayesian risk within a certain class. These estimators can be expressed as truncated Bayesian estimators.

Key words.

James-Stein estimators, Bayesian estimators, shrinkage functions.

1. INTRODUCTION.

Après que Stein (1956) ait montré l'inadmissibilité de l'estimateur des moindres carrés de la moyenne d'un vecteur aléatoire $y \sim \mathcal{N}_n(\theta, \sigma^2 I_n)$ avec $n \geq 3$, beaucoup d'estimateurs à rétrécisseurs dominant l'estimateur des moindres carrés ont été proposés à la suite de l'estimateur de James-Stein primitif (1961). En particulier, Baranchick (1964) a montré que l'estimateur de James-Stein de partie positive, domine l'estimateur des moindres carrés. Bien que cet estimateur ne soit pas admissible, aucun estimateur uniformément meilleur n'est connu.

A la recherche d'estimateurs admissibles, Strawderman (1971, 1973) et Alam (1973, 1975) introduisent le point de vue bayésien en construisant certaines classes d'estimateurs propres. En particulier, Efron et Morris (1973) montrent que l'estimateur de Stein tronqué appartient à une classe de "bons" estimateurs qui ont de petit risque bayésien et de plus dominant l'estimateur des moindres carrés. Ensuite, Spruill (1986) considère le problème de l'estimation robuste de la moyenne normale, pour le cas $\sigma^2 = 1$. Il construit certains estimateurs sphériquement symétriques qui sont de risque borné et de petit risque bayésien par rapport aux lois a priori uniformes sur la sphère $\|\theta\|^2 \leq \lambda$. La performance de ces estimateurs robustes est comparée avec celle des estimateurs de Stein. Il se trouve que, quand λ est petit, l'amélioration sur le risque de Bayes apportée par l'estimateur de Stein de partie positive est plus grande que celle de l'estimateur de "Bayes restreints approchés" φ^c , pour c suffisamment petit. Par contre, si λ est grand, φ^c est préférable.

On s'intéresse ici au traitement bayésien de l'estimation de la moyenne θ dans le modèle linéaire gaussien, quand la variance est connue à un facteur multiplicatif près, le coût utilisé est quadratique (inverse de la variance), et la probabilité a priori est un mélange de lois gaussiennes. On sait bien que les estimateurs de Bayes sont, dans ce cas, des estimateurs à rétrécisseur, non nécessairement uniformément meilleurs que l'estimateur des moindres carrés. On adopte ici le point de vue (déjà introduit par [Efron B. et Morris C., 1973] et [Spruill C.M., 1986]) qui consiste à chercher l'estimateur de plus faible risque bayésien à l'intérieur d'une classe d'estimateurs dont on sait par ailleurs qu'ils sont uniformément meilleurs que l'estimateur des moindres carrés. C'est l'objet de la section 4. Dans cette section, on définit les estimateurs "bayésiens à rétrécisseur majoré" dont un cas particulier sont les estimateurs "bayésiens tronqués au niveau c ". On présente des résultats théoriques sur ces estimateurs et on les calcule explicitement pour certaines probabilités a priori. On trouve dans la section 4.d que, quand l'"hyperparamètre" ω , servant à indexer la distribution mélangeante, est près de 1, l'amélioration sur le risque bayésien apportée par les estimateurs "bayésiens tronqués au

niveau c " devient plus grande que celle de l'estimateur de James-Stein de partie positive pour certaines valeurs de c (voir la figure 1).

2. MODELES.

On effectue une observation y dans \mathbb{R}^n , suivant une loi normale $\mathcal{N}_n(\theta, \sigma^2 v)$ où θ appartient à un sous-espace vectoriel connu Θ de \mathbb{R}^n , de dimension k supérieure à 2, et où v est une forme bilinéaire symétrique définie positive (en abrégé f.b.s.d.p.) connue (il sera pratique de la considérer comme un automorphisme de \mathbb{R}^n et de noter \tilde{v} la forme quadratique associée). θ et σ (> 0) sont inconnus.

La forme v étant fixée dans toute l'étude, on pourra noter, sans risque de confusion

- II. la norme définie sur \mathbb{R}^n par v^{-1} , i.e. $\|y\|^2 = \overline{v^{-1}}(y)$
(et également, sur tout sous-espace de \mathbb{R}^n , par la restriction de v^{-1} à ce sous-espace).
- I. l'orthogonalité définie dans \mathbb{R}^n par v^{-1}
(en particulier on utilisera Θ^\perp , sous-espace v^{-1} -orthogonal de Θ).

Soit \mathcal{E} l'ensemble des estimateurs de θ , autrement dit des applications (évidemment mesurables) de \mathbb{R}^n dans Θ ; ils sont comparés au moyen du coût "usuel", qui, si (θ, σ^2) est la valeur exacte du couple de paramètres inconnus et $\varphi(y)$ l'estimation proposée pour θ , prend la valeur

$$\frac{1}{\sigma^2} \overline{v_\Theta^{-1}}(\varphi(y) - \theta)$$

(où v_Θ^{-1} est la restriction de v^{-1} à Θ) qu'on peut aussi noter $\frac{1}{\sigma^2} \|\varphi(y) - \theta\|^2$; le risque de φ est alors l'application

$$(\theta, \sigma) \rightarrow R(\varphi, \theta, \sigma) = \frac{1}{\sigma^2} \mathbb{E}_{\theta, \sigma}[\|\varphi(y) - \theta\|^2]$$

où $\mathbb{E}_{\theta, \sigma}$ désigne l'intégration par rapport à la loi de probabilité $\mathcal{N}_n(\theta, \sigma^2 v)$.

On cherchera en particulier à comparer φ à φ_0 , estimateur des moindres carrés généralisés, qui est la projection v^{-1} -orthogonale, de \mathbb{R}^n sur Θ . On rappelle que la

variance de φ_0 est $\sigma^2 v_\Theta$ où v_Θ (f.b.s.d.p. sur Θ^*) est définie comme l'inverse de la restriction (v_Θ^{-1}) de v^{-1} au sous-espace vectoriel Θ . Son risque est l'application constante k ($= \dim \Theta$).

On notera, dans tout ce texte

$$\begin{aligned} x &= \varphi_0(y) \ (\in \Theta) & t &= \overline{v_\Theta^{-1}}(x) \ (= \|\varphi_0(y)\|^2) \ (\in \mathbb{R}_+) \\ z &= y - \varphi_0(y) \ (\in \Theta^\perp) & u &= \overline{v_\Theta^{-1}}(z) \ (= \|y - \varphi_0(y)\|^2) \ (\in \mathbb{R}_+). \end{aligned}$$

Une fois choisie une probabilité a priori π sur (θ, σ) , il est naturel d'associer à chaque estimateur φ de θ son risque bayésien, c'est-à-dire

$$r(\varphi, \Pi) = \int_{\mathbb{R}^n \times \Theta \times \mathbb{R}_+^*} \frac{1}{\sigma^2} \|\varphi(y) - \theta\|^2 P(dy, d\theta, d\sigma)$$

(où la probabilité P est définie par la probabilité a priori Π et la transition d'échantillonnage $\mathcal{N}_n(\theta, \sigma^2 v)$.)

Soit φ^* un estimateur de Bayes associé à Π . Dans ces conditions, nous présenterons des estimateurs qui, à la fois, sont meilleurs que φ_0 et ont de bonnes propriétés relativement à une probabilité a priori donnée Π . L'idée pour cela consiste à utiliser des résultats de nature "classique" (i.e. non bayésienne) permettant de fixer un sous-ensemble de \mathcal{E} , soit \mathcal{E}' , dont tous les éléments sont meilleurs que φ_0 , et à chercher l'estimateur de plus petit risque bayésien à l'intérieur de \mathcal{E}' .

Les probabilités a priori Π sur (θ, σ) , que nous utiliserons ici vérifieront la propriété suivante :

conditionnellement à σ , la loi de θ , est un mélange de lois normales, toutes de même moyenne (indépendante de σ) et à variance proportionnelle à v_Θ (la loi de mélange, qui porte sur le facteur de proportionnalité pouvant dépendre de σ).

Autrement dit, Π est caractérisée par le choix de μ ($\in \Theta$) et d'une probabilité π sur $(\sigma, \omega) \ (\in \mathbb{R}_+^* \times]0, 1[)$ vérifiant $\int_{\mathbb{R}_+^*} \frac{1}{\sigma^2} \pi_1(d\sigma) < +\infty$ (où π_j , $j = 1, 2$, sont les lois

marginales de π) ; alors la loi conditionnelle de θ sachant (σ, ω) est $\mathcal{N}_\Theta(\mu, \sigma^2 \frac{1-\omega}{\omega} v_\Theta)$

([Berger, J.O., 1985]).

L'origine de l'espace \mathbb{R}^n n'a, jusqu'ici, joué aucun rôle dans notre étude (on suppose seulement qu'elle appartient à Θ) ; on ne perd donc rien de la généralité de notre étude en posant $\mu = 0$.

On constatera (voir en section 3) que la recherche d'estimateurs bayésiens, dans le cadre que nous venons de présenter, conduit à des *estimateurs à rétrécisseurs* de la forme

$$\varphi(y) = [1 - f(y)]\varphi_0(y)$$

où f (application de \mathbb{R}^n dans \mathbb{R}) est appelé *le rétrécisseur*.

En fait, nous verrons que le choix de la densité a priori normale pour laquelle x et u sont des statistiques exhaustives conduit à se limiter à des rétrécisseurs ne dépendant de y que par l'intermédiaire des valeurs associées à $\varphi_0(y)$ et $y - \varphi_0(y)$ par la forme quadratique inverse de v_Θ ; autrement dit, on considère les estimateurs de la forme

$$\varphi(y) = [1 - g(\|\varphi_0(y)\|^2, \|y - \varphi_0(y)\|^2)]\varphi_0(y)$$

(par exemple, avec la notation introduite ci-dessus, $g(t,u) = \frac{k-2}{n-k+2} \frac{u}{t}$ est l'estimateur "primitif" de James-Stein).

Pour les rétrécisseurs de ce type, on dispose de théorèmes caractérisant des familles d'estimateurs uniformément meilleurs que φ_0 .

Nous utiliserons ici la proposition suivante due à Strawderman ([Strawderman E.W., 1973]).

Proposition 2.1.

Une condition suffisante pour que l'estimateur φ , de rétrécisseur $g(\mathbb{R}_+^2 \rightarrow \mathbb{R}_+)$ soit de risque uniformément inférieur ou égal à celui de φ_0 est

- (i) pour tout $u (\geq 0)$ l'application, définie sur \mathbb{R}_+ , $t \rightarrow tg(t,u)$ est croissante ;
- (ii) pour tout $\alpha (>0)$ l'application, définie sur \mathbb{R}_+ , $u \rightarrow g(\alpha u, u)$ est décroissante ;
- (iii) $\forall (t,u) \in \mathbb{R}_+ \times \mathbb{R}_+^*$, $\frac{t}{u} g(t,u) \leq 2 \frac{k-2}{n-k+2}$.

3. ESTIMATEURS BAYESIENS : FORME GENERALE, CALCUL DU RISQUE

Pour le calcul du risque bayésien, nous avons besoin de certains résultats classiques que nous rappelons ici sans démonstration (voir [Lindley D.V. et Smith A-F.M., 1972]).

La loi prédictive de y est

$$\mathcal{N}_n(0, \sigma^2(v + \frac{1-\omega}{\omega} \tilde{v}_\Theta))$$

(où \tilde{v}_Θ est définie par, pour tout couple (f_1, f_2) de forme linéaire sur \mathbb{R}^n , on a $\tilde{v}_\Theta(f_1, f_2) = v_\Theta(f_1|_\Theta, f_2|_\Theta)$) autrement dit, en écrivant v sous la forme $\tilde{v}_\Theta + \tilde{v}_{\Theta^\perp}$, et en écrivant y sous la forme (x, z) , la loi ci-dessus devient

$$\mathcal{N}_n(0, \frac{\sigma^2}{\omega} v_\Theta) \times \mathcal{N}_{\Theta^\perp}(0, \sigma^2 v_{\Theta^\perp}) \quad (3.1)$$

Il nous sera utile pour la suite d'utiliser la loi prédictive de (t, u) ; en considérant une base de \mathbb{R}^n , v^{-1} -orthogonale dont les k -premiers éléments constituent une base de Θ , il résulte de (3.1) que c'est

$$\frac{\sigma^2}{\omega} \chi_k^2 \times \sigma^2 \chi_{n-k}^2 \quad (3.1')$$

On peut adopter pour la loi a posteriori de θ quand on observe $y (= (x, z))$ la probabilité

$$\mathcal{N}_\Theta((1-\omega)x, \sigma^2(1-\omega)v_\Theta). \quad (3.2)$$

Alors, le risque bayésien d'un estimateur φ conditionnellement à (σ, ω) est

$$r(\varphi, \sigma, \omega) = k(1-\omega) + \frac{1}{\sigma^2} \int_{\mathbb{R}^n} \|\varphi(y) - (1-\omega)\varphi_0(y)\|^2 \mathcal{N}_n(0, \sigma^2(v + \frac{1-\omega}{\omega} \tilde{v}_\Theta)) ; dy \quad (3.3)$$

On retrouve (résultat déjà connu) que l'estimateur de Bayes ξ relativement à la probabilité a priori $\mathcal{L}_{\Theta}(0, \sigma^2 \frac{1-\omega}{\omega} \nu_{\Theta})$ quand σ est connu et ω fixé est $\xi(y) = (1-\omega)\varphi_0(y)$ et que son risque bayésien est $k(1-\omega)$.

Le risque bayésien d'un estimateur φ est

$$\begin{aligned} r(\varphi, \Pi) &= \int_{\mathbb{R}_+^* \times]0,1]} r(\varphi, \sigma, \omega) \pi(d\sigma, d\omega) \\ &= k \int_0^1 (1-\omega) \pi_2(d\omega) + \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}_+^* \times]0,1]} \frac{1}{\sigma^2} \|\varphi(y) - (1-\omega)\varphi_0(y)\|^2 \tau(d\sigma, d\omega|y) \right) p(dy) \end{aligned} \quad (3.4)$$

où $\tau(\cdot|y)$ est la loi a posteriori de (σ, ω) quand on observe y et p la loi prédictive de y .

Il résulte de l'hypothèse $\int_{\mathbb{R}_+^*} \frac{1}{\sigma^2} \pi_1(d\sigma) < +\infty$ que, presque tout y

(l'expression "presque tout" s'entend bien sûr, relativement à la probabilité prédictive p),

on a $\int_{\mathbb{R}_+^* \times]0,1]} \frac{1}{\sigma^2} \tau(d\sigma, d\omega|y) < +\infty$. Alors, par un calcul élémentaire d'intégration, on

montre que l'estimateur de Bayes de θ relativement à Π est $\varphi^*(y) = \varphi_0(y) - f^*(y)\varphi_0(y)$,

où

$$f^*(y) = \frac{\int_{\mathbb{R}_+^* \times]0,1]} \frac{\omega}{\sigma^2} \tau(d\sigma, d\omega|y)}{\int_{\mathbb{R}_+^* \times]0,1]} \frac{1}{\sigma^2} \tau(d\sigma, d\omega|y)} \quad (3.5)$$

c'est un estimateur à rétrécisseur scalaire.

Précisons le calcul du rétrécisseur f^* . Dans le modèle statistique bayésien, avec paramètre (σ, ω) , l'application $y = (x, z) \rightarrow (t, u)$ est exhaustive et donc $f^*(y)$ définie par la formule (3.5), ne dépend de y que par l'intermédiaire du couple (t, u) ; on notera $f^*(y) = g^*(t, u)$. Posant $y = (x, z)$, on déduit de (3.1) que la probabilité de (σ, ω) sachant (x, z) , $\tau(\cdot|x, z)$, admet pour densité par rapport à π , l'application

$$\ell(\sigma, \omega|x, z) = \frac{\omega^{k/2} \sigma^{-n} \exp\left\{-\frac{1}{2\sigma^2}(\omega t + u)\right\}}{\int_{\mathbb{R}_+^* \times]0,1]} \omega^{k/2} \sigma^{-n} \exp\left\{-\frac{1}{2\sigma^2}(\omega t + u)\right\} \pi(d\sigma, d\omega)}$$

et donc

$$f^*(y) = g^*(t,u) = \frac{\int_{\mathbb{R}_+^* \times]0,1]}{\int_{\mathbb{R}_+^* \times]0,1]} \frac{\omega^{k/2+1} \sigma^{-(n+2)} \exp\{-\frac{1}{2\sigma^2}(\omega t+u)\} \pi(d\sigma, d\omega)}{\omega^{k/2} \sigma^{-(n+2)} \exp\{-\frac{1}{2\sigma^2}(\omega t+u)\} \pi(d\sigma, d\omega)} \quad (3.5')$$

qu'on peut noter aussi, compte tenu de la probabilité conjointe de t, u, σ, ω sur $\mathbb{R}_+^2 \times \mathbb{R}_+^* \times]0,1]$ définie par π et par $\frac{\sigma^2}{\omega} \chi_k^2 \times \sigma^2 \chi_{n-k}^2$ (voir (3.1')) par

$$f^*(y) = g^*(t,u) = \frac{\int_{\mathbb{R}_+^* \times]0,1]}{\int_{\mathbb{R}_+^* \times]0,1]} \frac{\frac{\omega}{\sigma^2} \tau(d\sigma, d\omega | t, u)}{\frac{1}{\sigma^2} \tau(d\sigma, d\omega | t, u)} \quad (3.5'')$$

Ce résultat fait l'objet du lemme suivant.

Lemme 3.1.

On adopte la probabilité a priori Π sur (θ, σ) définie par la probabilité π sur (σ, ω) ($\in \mathbb{R}_+^* \times]0,1]$) et la transition, de $\mathbb{R}_+^* \times]0,1]$ vers Θ , $\eta_{\Theta}(0, \sigma^2 \frac{1-\omega}{\omega} v_{\Theta})$; alors si $\int_{\mathbb{R}_+^*} \frac{1}{\sigma^2} \pi_1(d\sigma) < +\infty$, l'estimateur bayésien de θ (relativement au coût $\frac{1}{\sigma^2} \overline{v_{\Theta}^{-1}}(\theta - \hat{\theta})$) est l'estimateur à rétrécisseur $\varphi^*(y) = [1 - f^*(y)] \varphi_0(y)$ de rétrécisseur $f^*(y) = g^*(t, u)$, à valeurs dans l'intervalle $]0,1]$, donné par (3.5') (où $t = \overline{v_{\Theta}^{-1}}(\varphi_0(y))$ et $u = \overline{v_{\Theta}^{-1}}(y - \varphi_0(y))$).

Ce lemme justifie que nous nous limitons désormais à la considération d'estimateurs à rétrécisseurs de la forme

$$\varphi(y) = [1 - g(t, u)] \varphi_0(y)$$

où g est à valeurs dans l'intervalle $]0,1]$.

La formule (3.4), qui donne le risque bayésien, devient

$$r(\varphi, \Pi) = k \int_0^1 (1-\omega) \pi_2(d\omega) + \int_{\mathbb{R}_+^2} t \left(\int_{\mathbb{R}_+^* \times]0,1]} \frac{1}{\sigma^2} [\omega - g(t,u)]^2 \tau(d\sigma, d\omega t, u) \right) p(dt, du) \quad (3.6)$$

autrement dit, en utilisant le rétrécisseur g^* de l'estimateur de Bayes φ^* ,

$$r(\varphi, \Pi) = k \int_0^1 (1-\omega) \pi_2(d\omega) + \int_{\mathbb{R}_+^2} t \left(\int_{\mathbb{R}_+^* \times]0,1]} \frac{1}{\sigma^2} [\omega - g^*(t,u)]^2 \tau(d\sigma, d\omega t, u) \right) p(dt, du) \quad (3.6')$$

$$+ \int_{\mathbb{R}_+^2} t [g(t,u) - g^*(t,u)]^2 \left(\int_{\mathbb{R}_+^* \times]0,1]} \frac{1}{\sigma^2} \tau(d\sigma, d\omega t, u) \right) p(dt, du)$$

On peut caractériser l'estimateur φ par la mesure de performance "Relative Saving Loss" ($RSL(\varphi, \cdot)$) (déjà utilisée en [Efron B. et Morris C., 1972], [Efron B. et Morris C., 1973] et [Spruill C.M., 1986] ; en [Spruill C.M., 1986] l'estimateur des moindres carrés est remplacé par l'estimateur de Stein), définie par

$$RSL(\varphi, \sigma, \omega) = \frac{r(\varphi, \sigma, \omega) - r(\xi, \omega)}{r(\varphi_0, \omega) - r(\xi, \omega)} = \frac{r(\varphi, \sigma, \omega) - k(1-\omega)}{k\omega}$$

Il résulte de la formule (3.3), d'après (3.1'), que

$$RSL(\varphi, \sigma, \omega) = \int_{\mathbb{R}_+^2} \left[1 - \frac{1}{\omega} g(t,u) \right]^2 \left(\frac{\sigma^2}{\omega} \chi_{t+2}^2 \right) (dt) \left(\sigma^2 \chi_{n-k}^2 \right) (du) \quad (3.7)$$

et dans le cas particulier où g factorise à travers $\frac{t}{u}$, c'est-à-dire $g(t,u) = \bar{g}\left(\frac{t}{u}\right)$, on a

$$RSL(\varphi, \omega) = \int_{\mathbb{R}_+} \left[1 - \frac{1}{\omega} \bar{g}\left(\frac{k+2}{n-k} v\right) \right]^2 \left(\frac{1}{\omega} \mathcal{F}_{k+2, n-k} \right) (dv) \quad (3.7')$$

$RSL(\varphi, \sigma, \omega)$ est évidemment nul si g est la constante ω , et égal à 1 si g est la constante 0. C'est ce qui en fait l'intérêt pratique : en effet, le rétrécisseur constant ω est le "meilleur" si ω est connu (puisqu'il est alors bayésien) mais est en général inemployable (car ω est inconnu) ; inversement le rétrécisseur nul est toujours utilisable (il définit l'estimateur des moindres carrés φ_0) mais notre but est justement de l'améliorer. Il y aura donc intérêt à n'utiliser que des rétrécisseurs de RSL compris entre

0 et 1, et la proximité de $RSL(\varphi, \sigma, \omega)$ à 0 caractérisera dans quelle mesure, pour la valeur (σ, ω) considérée, le rétrécisseur g apporte l'amélioration recherchée.

Lien avec l'estimateur primitif de James-Stein.

Il est bien connu [Judge, G.G. et Bock, E.M. (1978)] que les estimateurs à rétrécisseurs de la forme

$$\varphi_c(y) = [1 - c \frac{u}{t}] \varphi_0(y)$$

sont meilleurs que φ_0 si $c \leq 2(k-2)/(n-k+2)$ avec $k > 2$ et $n-k > 0$ (c'est un cas particulier évident de la Proposition 2.1), et que l'un d'entre eux, celui pour lequel $c = (k-2)/(n-k+2)$, est de risque uniformément inférieur à celui de tous les autres.

$$\text{Posons} \quad h_c(t, u) = \tilde{h}_c\left(\frac{t}{u}\right) = c \frac{u}{t}.$$

Le calcul de RSL de l'estimateur associé est élémentaire [Efron, B. et Morris, C. (1973)] ; il est constant et vaut

$$RSL(\varphi_c) = \frac{2n}{k(n-k+2)} + \left(1 - \frac{2n}{k(n-k+2)}\right) \left(\frac{n-k+2}{k-2} c - 1\right)^2 \quad (3.8)$$

Il faut noter que h_c ne rentre pas dans le type de rétrécisseurs auxquels nous avons décidé de limiter notre attention, car il n'est pas majoré par 1 ; de ce fait, il ne peut pas, en particulier, être obtenu comme définissant un estimateur de Bayes relativement à une probabilité a priori du type de celles que nous étudions ici. Mais on peut trouver dans la littérature (voir [Efron B. et Morris C., 1973] et [Lin P.E. et Tsai H.L., 1973]) l'estimateur des James-Stein comme estimateur de Bayes empirique.

Bien sûr, on retrouve des rétrécisseurs du type de ceux auxquels nous nous intéressons en prenant

$$h_c^+(t, u) = \inf(h_c(t, u), 1) \quad (= \inf(c \frac{u}{t}, 1))$$

et on peut vérifier (voir [Judge G.G. et Bock, E.M., 1978]) que h_c^+ est uniformément meilleur que h_c ; mais les calculs de risque sont plus délicats (C. Robert en [12] donne une formule exacte du risque usuel quand $c = (k-2)/(n-k+2)$ après un calcul assez

difficile) ainsi que la comparaison de ces estimateurs entre eux ; nous les retrouverons en section 4.

4. ESTIMATEURS BAYESIENS A RETRECISSEURS DOMINES.

a) Estimateurs à rétrécisseur majoré.

A la recherche des estimateurs de petit risque bayésien, Spruill en [Spruill C.M., 1986] introduit les estimateurs de "Bayes restreints approchés" φ^c qui minimisent le risque de Bayes, parmi tous les rétrécisseurs $h(t) = \frac{k-2}{t} - 2\Phi(t)$ pour lesquels $D\Phi(t) \leq c$ quel que soit $t > 0$, quand $c (> 0)$ est une constante suffisamment petite et D est un opérateur différentiel, tandis que Efron et Morris en [Efron B. et Morris C., 1973], introduisent les estimateurs de "Bayes tronqués" qui minimisent le risque de Bayes parmi tous les rétrécisseurs $h(t)$ qui sont majorés. Pour le même but, d'après [Efron B. et Morris C., 1973], on définit les estimateurs suivants.

Nous serons amenés à nous limiter à des classes d'estimateurs dont le rétrécisseur vérifie

$$g(t,u) \leq s(t,u)$$

où s est une application donnée de \mathbb{R}_+^2 dans \mathbb{R}_+ . On note \mathcal{E}_s l'ensemble de ces estimateurs.

On appelle, s'il existe, *estimateur bayésien à rétrécisseur majoré par s* , relativement à une probabilité π sur $\mathbb{R}_+^* \times]0,1]$ (vérifiant $\int_{\mathbb{R}_+^*} \frac{1}{\sigma^2} \pi_1(d\sigma) < +\infty$) tout estimateur $\varphi_s^* \in \mathcal{E}_s$ vérifiant

$$r(\varphi_s^*, \pi) = \inf_{\varphi \in \mathcal{E}_s} r(\varphi, \pi).$$

Reprenant la formule (3.6') du risque bayésien, on voit qu'on obtient un tel estimateur en construisant son rétrécisseur, g_s^* , de telle sorte que, pour tout y ,

$$[g_s^*(t,u) - g^*(t,u)]^2 = \inf_{a \leq s(t,u)} [a - g^*(t,u)]^2$$

ce qui équivaut évidemment à

$$g_s^*(t,u) = \inf(s(t,u), g^*(t,u)) . \quad (4.1)$$

La différence entre les risques bayésiens de φ_s^* et φ^* est

$$r(\varphi_s^*, \pi) - r(\varphi^*, \pi) = \int_{[g^* > s]} t [g^*(t,u) - s(t,u)]^2 \left(\int_{\mathbb{R}_+^* \times]0,1]} \frac{1}{\sigma^2} \tau(d\sigma, d\omega | t, u) \right) p(dt, du) \quad (4.2)$$

Il résulte immédiatement de (4.2) que, si $s_1 \leq s_2$

$$r(\varphi_{s_1}^*, \pi) \geq r(\varphi_{s_2}^*, \pi) .$$

On a donc l'énoncé suivant.

Lemme 4.1.

Le risque bayésien de l'estimateur bayésien à rétrécisseur majoré est une fonction décroissante du majorateur (les majorateurs étant ordonnées selon la relation d'ordre usuelle des applications de \mathbb{R}^n dans \mathbb{R}_+^*).

Un cas particulièrement fréquent est celui où on considère une famille de majorateurs de la forme $c/S(t,u)$, où S est une application fixée de \mathbb{R}^n dans \mathbb{R}_+^* , et où seule la constante c varie (autrement dit on impose à g_S d'être majoré par c).

Alors

$$g_{c/S}^*(t,u) = \frac{1}{S(t,u)} \inf(c, g^*(t,u)S(t,u))$$

et

$$r(\varphi_{c/S}^*, \pi) = r(\varphi^*, \pi) + \int_{[g^*S > c]} \frac{t}{S^2(t,u)} [g^*(t,u)S(t,u) - c]^2 \left(\int_{\mathbb{R}_+^* \times]0,1]} \frac{1}{\sigma^2} \tau(d\sigma, d\omega | t, u) \right) p(dt, du) \quad (4.3)$$

On dira que $\varphi_{c/S}^*$ est un *estimateur bayésien à rétrécisseur S-tronqué* et c est appelé *son niveau de troncature*. Ces estimateurs vérifient la proposition suivante.

Proposition 4.1.

Le risque bayésien de l'estimateur bayésien à rétrécisseur S-tronqué et une fonction décroissante et convexe du niveau de troncature.

Preuve. La décroissance a déjà été obtenue par le lemme 4.1. La convexité résulte de la formule (4.3) et de la convexité, pour tout $t \in \mathbb{R}_+$ de l'application

$$c \rightarrow 1_{]c, +\infty[}(t) (t-c)^2.$$

b) Estimateurs bayésiens tronqués

Tous les estimateurs considérés dans cette section satisfont aux hypothèses de la Proposition 2.1. Ils sont donc de risque uniformément inférieur ou égal à celui de φ_0 .

Notre démarche pour les obtenir est la suivante :

1°- Considérer une probabilité π sur (σ, ω) , et s'assurer que l'estimateur bayésien associé de rétrécisseur g^* , satisfait les hypothèses (i) et (ii) de la Proposition 2.1.

2°- Poser $S(t, u) = \frac{t}{u}$, et considérer les rétrécisseurs $g_{c/S}^*$, déduits de g^* par S-troncature au niveau c , autrement dit (voir en a)

$$g_{c/S}^*(t, u) = \inf \left(c \frac{u}{t}, g^*(t, u) \right) = \frac{u}{t} \inf \left(c, \frac{t}{u} g^*(t, u) \right) ;$$

il est clair que $g_{c/S}^*$ satisfait, comme g^* , aux hypothèses (i) et (ii) et il suffit de prendre $c \leq 2(k-2)/(n-k+2)$ pour qu'il satisfasse aussi à l'hypothèse (iii).

Dans toute cette partie, le renvoi aux conditions (i), (ii) et (iii) les désignera sous la forme qui vient d'être donnée.

On notera μ la probabilité a priori sur (σ, ω) à partir de laquelle on détermine un estimateur bayésien dont on envisage les troncatures. Notons que, une fois déterminés de tels estimateurs, on pourra être amené à considérer leur risque bayésien relativement à d'autres probabilités a priori sur (σ, ω) .

On adopte alors la définition et la notation suivantes :

Définition.

Etant donnée une probabilité μ sur $\mathbb{R}_+^* \times]0, 1]$ telle que g_μ^* , rétrécisseur de l'estimateur bayésien associé, vérifie les conditions (i) et (ii) et un nombre $c \in]0, 2 \frac{k-2}{n-k+2}]$, on appelle *estimateur μ -bayésien tronqué au niveau c* l'estimateur $\varphi_{\mu, c}^*$ de rétrécisseur

$$g_{\mu, c}^*(t, u) = \inf \left(c \frac{u}{t}, g_\mu^*(t, u) \right)$$

(voir le lemme 3.1. pour l'expression de g_{μ}^*).

Dans la suite nous comparons ces estimateurs et mettons en évidence des probabilités μ fournissant de tels estimateurs.

c) Comparaison des estimateurs bayésiens tronqués.

On rappelle que le "Relative Saving Loss" de l'estimateur φ_c en (3.8), est constant et on le note $RSL(\varphi_c)$.

Alors un lemme utile dans la comparaison des estimateurs est le suivant.

Lemme 4.2.

Pour tout estimateur φ à rétrécisseur g satisfaisant aux hypothèses (i) et (ii) et

$$\forall (t,u) \in (\mathbb{R}_+ \times \mathbb{R}_+^*), \frac{t}{u} g(t,u) \leq c$$

on a

$$\lim_{\substack{\omega \rightarrow 0 \\ \sigma \rightarrow 0}} RSL(\varphi, \sigma, \omega) = RSL(\varphi_c).$$

Preuve. Il résulte de la formule (3.7), dans laquelle on effectue le changement de variable $t = \frac{\sigma^2 v}{\omega}$ et $u = \sigma^2 w$ que

$$RSL(\varphi, \sigma, \omega) = \int_{\mathbb{R}_+^2} \left[1 - \frac{w}{v} \frac{\sigma^2 v}{\omega} \frac{1}{\sigma^2 w} g\left(\frac{\sigma^2 v}{\omega}, \sigma^2 w\right) \right]^2 \chi_{k+2}^2(dv) \chi_{n-k}^2(dw).$$

Il résulte des hypothèses (i) et (ii) que, si, v et w étant fixés, ω et σ tendent vers 0, $\frac{\sigma^2 v}{\omega} \frac{1}{\sigma^2 w} g\left(\frac{\sigma^2 v}{\omega}, \sigma^2 w\right)$ tend en croissant vers c .

Il résulte donc du théorème de la convergence dominée de Lebesgue que

$$\begin{aligned} \lim_{\substack{\omega \rightarrow 0 \\ \sigma \rightarrow 0}} RSL(\varphi, \sigma, \omega) &= \int_{\mathbb{R}_+^2} \left(1 - \frac{w}{v} c \right)^2 \chi_{k+2}^2(dv) \chi_{n-k}^2(dw) \\ &= \int_{\mathbb{R}_+^2} \left(1 - \frac{1}{\omega} c \frac{u}{t} \right)^2 \left(\frac{\sigma^2}{\omega} \chi_{k+2}^2 \right)(dt) (\sigma^2 \chi_{n-k}^2)(du) \\ &= RSL(\varphi_c) \end{aligned}$$

Proposition 4.2.

Les estimateurs μ -bayésiens tronqués, de niveau c appartenant à l'intervalle $[\frac{k-2}{n-k+2}, 2\frac{k-2}{n-k+2}]$, sont tous incomparables (au sens du risque usuel).

Preuve. On démontre que ces estimateurs $\varphi_{\mu,c}^*$, sont incomparables quand on les caractérise par leurs risques conditionnels à σ et ω (ou, ce qui est équivalent par leurs RSL). Il s'en suit a fortiori qu'ils sont incomparables au sens du risque usuel.

Soient φ_{μ_j,c_j}^* où $j \in \{1,2\}$, deux estimateurs μ_j -bayésiens tronqués, de niveaux c_j ($\in [\frac{k-2}{n-k+2}, 2\frac{k-2}{n-k+2}]$).

Démontrons qu'il existe (σ_1, ω_1) et (σ_2, ω_2) tels que

$$r(\varphi_{\mu_1,c_1}^*, \sigma_1, \omega_1) < r(\varphi_{\mu_2,c_2}^*, \sigma_1, \omega_1) \text{ et } r(\varphi_{\mu_2,c_2}^*, \sigma_2, \omega_2) < r(\varphi_{\mu_1,c_1}^*, \sigma_2, \omega_2).$$

En effet, si $c_1 = c_2$ on a

$$r(\varphi_{\mu_1,c_1}^*, \sigma_1, \omega_1) < r(\varphi_{\mu_2,c_1}^*, \sigma_1, \omega_1)$$

De même,

$$\int_{\mathbb{R}_+^* \times]0,1]} r(\varphi_{\mu_2,c_1}^*, \sigma, \omega) \mu_2(d\sigma, d\omega) < \int_{\mathbb{R}_+^* \times]0,1]} r(\varphi_{\mu_1,c_1}^*, \sigma, \omega) \mu_2(d\sigma, d\omega) ;$$

donc, il existe (σ_2, ω_2) tel que

$$r(\varphi_{\mu_2,c_1}^*, \sigma_2, \omega_2) < r(\varphi_{\mu_1,c_1}^*, \sigma_2, \omega_2) .$$

Soit $c_1 \neq c_2$. Sans restriction de généralité, on peut supposer que $c_1 < c_2$. Alors, d'après le lemme ci-dessus,

$$\lim_{\substack{\omega \rightarrow 0 \\ \sigma \rightarrow 0}} \text{RSL}(\varphi_{\mu_1,c_1}^*, \sigma, \omega) = \text{RSL}(\varphi_{c_1}) < \text{RSL}(\varphi_{c_2}) = \lim_{\substack{\omega \rightarrow 0 \\ \sigma \rightarrow 0}} \text{RSL}(\varphi_{\mu_2,c_2}^*, \sigma, \omega).$$

Donc, il existe (σ_1, ω_1) tel que

$$\text{RSL}(\varphi_{\mu_1, c_1}^*, \sigma_1, \omega_1) < \text{RSL}(\varphi_{\mu_2, c_2}^*, \sigma_1, \omega_1) ;$$

alors d'après la définition du RSL, il existe (σ_1, ω_1) tel que

$$r(\varphi_{\mu_1, c_1}^*, \sigma_1, \omega_1) < r(\varphi_{\mu_2, c_2}^*, \sigma_1, \omega_1) .$$

Inversement, d'après la Proposition 4.1., puisque l'application

$$c \rightarrow \int_{\mathbb{R}_+^* \times]0,1]} r(\varphi_{\mu, c}^*, \sigma, \omega) \mu(d\sigma, d\omega)$$

est strictement décroissante, on a

$$\int_{\mathbb{R}_+^* \times]0,1]} r(\varphi_{\mu_2, c_1}^*, \sigma, \omega) \mu_2(d\sigma, d\omega) > \int_{\mathbb{R}_+^* \times]0,1]} r(\varphi_{\mu_2, c_2}^*, \sigma, \omega) \mu_2(d\sigma, d\omega).$$

Mais d'après la définition de φ_{μ_2, c_1}^*

$$\int_{\mathbb{R}_+^* \times]0,1]} r(\varphi_{\mu_2, c_1}^*, \sigma, \omega) \mu_2(d\sigma, d\omega) < \int_{\mathbb{R}_+^* \times]0,1]} r(\varphi_{\mu_1, c_1}^*, \sigma, \omega) \mu_2(d\sigma, d\omega).$$

Donc

$$\int_{\mathbb{R}_+^* \times]0,1]} r(\varphi_{\mu_1, c_1}^*, \sigma, \omega) \mu_2(d\sigma, d\omega) > \int_{\mathbb{R}_+^* \times]0,1]} r(\varphi_{\mu_2, c_2}^*, \sigma, \omega) \mu_2(d\sigma, d\omega).$$

Il en résulte qu'il existe (σ_2, ω_2) tel que

$$r(\varphi_{\mu_1, c_1}^*, \sigma_2, \omega_2) > r(\varphi_{\mu_2, c_2}^*, \sigma_2, \omega_2).$$

d) Exemples.

1) Efron et Morris en [7] étudient (pour le cas où σ^2 est connu) certaines propriétés des estimateurs à rétrécisseur $\inf(\frac{c}{t}, b)$ où $b \in]0,1]$ et $c \in]0, 2(k-2)]$, relativement à la probabilité de Dirac δ_b sur ω .

Généralisons ces propriétés, pour le cas où σ^2 est inconnu, en considérant le cas où $\pi_2 = \delta_b$ (probabilité de Dirac en $b \in]0,1[$) et π_1 quelconque.

g^* est la constante b (vérifie évidemment (i) et (ii)).

On notera $\varphi_{c,b}^*$ l'estimateur bayésien tronqué de niveau c ; dans ce cas, son rétrécisseur est

$$g_{c,b}^*(t,u) = g_{c,b}^*\left(\frac{t}{u}\right) = \frac{u}{t} \inf\left(c, \frac{t}{u} b\right).$$

Si $b = 1$, on notera φ_c^* par $\varphi_{c,1}^*$.

Proposition 4.3.

Quel que soit $\omega \leq b$, on a

$$\begin{aligned} \text{RSL}(\varphi_c) - \text{RSL}(\varphi_{c,b}^*, \omega) &= \int_{\left\{b < \frac{n-k}{k+2} \frac{c}{v}\right\}} \left[\left(\frac{n-k}{k+2} \frac{c}{\omega v} - 1\right)^2 - \left(\frac{b}{\omega} - 1\right)^2 \right] \left(\frac{1}{\omega} \mathfrak{F}_{k+2, n-k}\right)(dv) \\ &> 0 \end{aligned} \quad (4.4)$$

Preuve. Il en résulte de la formule (3.7') que

$$\begin{aligned} \text{RSL}(\varphi_c) - \text{RSL}(\varphi_{c,b}^*, \omega) &= \\ \int_{\mathbb{R}_+} \left[\left(\frac{n-k}{k+2} \frac{c}{\omega v} - 1\right)^2 - \left(\frac{1}{\omega} g_{c,b}^*\left(\frac{k+2}{n-k} v\right) - 1\right)^2 \right] \left(\frac{1}{\omega} \mathfrak{F}_{k+2, n-k}\right)(dv). \end{aligned}$$

Or $g_{c,b}^*\left(\frac{k+2}{n-k} v\right)$ est égal à $\frac{n-k}{k+2} \frac{c}{v}$ si $\frac{n-k}{k+2} \frac{c}{v} \leq b$ et à b si $\frac{n-k}{k+2} \frac{c}{v} > b$; l'égalité (4.4) en résulte immédiatement. La stricte positivité de $\text{RSL}(\varphi_c) - \text{RSL}(\varphi_{c,b}^*, \omega)$ résulte alors de ce que, si $b < \frac{n-k}{k+2} \frac{c}{v}$, et $\omega \leq b$, on a $\frac{n-k}{k+2} \frac{c}{\omega v} > \frac{b}{\omega} \geq 1$ et d'autre part

$$\left(\frac{1}{\omega} \mathfrak{F}_{k+2, n-k}\right)\left(\left\{v : b < \frac{n-k}{k+2} \frac{c}{v}\right\}\right) > 0.$$

Une conséquence importante de cette proposition est que, pour toute probabilité a priori π sur (σ, ω) dont la loi marginale π_2 est concentrée sur $]0, b]$, le risque bayésien de $\varphi_{c,b}^*$ est inférieur à celui de φ_c (et même strictement inférieur si π_2 n'est pas δ_b). Considérant le cas où $b = 1$, on peut dire en particulier que, pour la famille des

probabilités a priori Π sur (θ, σ) considérées dans cette étude, φ_c^* est uniformément meilleur que φ_c .

D'après la Proposition 4.2., les estimateurs $\varphi_{c,b}^*$ sont incomparables. Mais on peut démontrer que, si $c_1 \in]0, \frac{k-2}{n-k+2}[$ [et $c_2 \in [\frac{k-2}{n-k+2}, 2 \frac{k-2}{n-k+2} - c_1]$, $RSL(\varphi_{c_1,b}^*, \omega) > RSL(\varphi_{c_2,b}^*, \omega)$ pour tout $\omega \leq b$. En effet, pour $j \in \{1, 2\}$,

$$RSL(\varphi_{c_j}) = \frac{2n}{k(n-k+2)} + (1 - \frac{2n}{k(n-k+2)}) (\frac{n-k+2}{k-2} c_j - 1)^2$$

déjà introduit en (3.8). Alors $RSL(\varphi_{c_1}) \geq RSL(\varphi_{c_2})$.

Puisque, pour tout $\frac{n-k}{k+2} \frac{c_1}{v\omega} < \frac{n-k}{k+2} \frac{c_2}{v\omega}$, le résultat suit de la Proposition 4.3.

Une expression de $RSL(\varphi_{c,b}^*, \omega)$ est donnée par le théorème suivant.

Théorème 4.1.

Si $b = 1$,

$$RSL(\varphi_c^*, \omega) = RSL(\varphi_c) - \int_{\mathbb{R}_+} \left\{ \left[\frac{c^2 y^2}{k(k-2)} - \frac{2cy}{k} - \frac{1}{\omega} \left(\frac{1}{\omega} - 2 \right) \right] I_{k/2} \left(\frac{c\omega}{2} y \right) + \left[\frac{c^2 y^2}{k(k-2)} + \left(\frac{1}{\omega} - 2 \right) \frac{cy}{k} \right] i_{k/2} \left(\frac{c\omega}{2} y \right) \right\} \chi_{n-k}^2(dy) \quad (4.5)$$

où $I_{k/2}(s)$ est la fonction de répartition de la distribution Gamma de paramètres $k/2$ et 1, et $i_{k/2}(s)$ est sa densité.

Si $b < 1$ et $\omega \leq b$,

$$RSL(\varphi_{c,b}^*, \omega) = RSL(\varphi_c^*, \frac{\omega}{b}) \quad (4.6)$$

où le $RSL(\varphi_c^*, \frac{\omega}{b})$ est calculé par (4.5).

La démonstration du Théorème 4.1. est très longue (voir [Criticou, D., 1986]) mais simple et nous l'omettons.

Corollaire 4.1.

L'application définie sur $]0,1[$, $\omega \rightarrow \text{RSL}(\varphi_{c,b}^*, \omega)$ (définie par (4.5) et (4.6)) prend sa valeur minimum en $\omega = b$; c'est

$$\begin{aligned} \text{RSL}(\varphi_{c,b}^*, b) = \text{RSL}(\varphi_c) - \int_{\mathbb{R}_+} \left\{ \left[\frac{c^2 y^2}{k(k-2)} - \frac{2cy}{k} + 1 \right] I_{k/2} \left(\frac{c}{2} y \right) \right. \\ \left. + \left[\frac{c^2 y^2}{k(k-2)} - \frac{cy}{k} \right] i_{k/2} \left(\frac{c}{2} y \right) \right\} \chi_{n-k}^2(dy) \end{aligned} \quad (4.7)$$

(qui ne dépend pas de b).

Preuve. Si $\omega \in]0,1[$ et $b \in]0,1[$, $\text{RSL}(\varphi_{c,\omega}^*, \omega) \leq \text{RSL}(\varphi_{c,b}^*, \omega)$. Ensuite, $\text{RSL}(\varphi_{c,\omega}^*, \omega) = \text{RSL}(\varphi_c^*, 1) = \text{RSL}(\varphi_{c,b}^*, b)$ et la formule (4.7) résulte de la formule (4.5) du Théorème 4.1.

Une autre interprétation pour la constante $c \in \left[\frac{k-2}{n-k+2}, 2 \frac{k-2}{n-k+2} \right]$ est donnée par le corollaire suivant.

Corollaire 4.2.

Soit $\frac{\sigma^2}{\omega} \chi_k^2 \times \sigma^2 \chi_{n-k}^2$ la loi prédictive du couple (t, u) relativement à la probabilité a priori sur θ , $\mathcal{N}_{\theta}(0, \sigma^2 \frac{1-\omega}{\omega} v_{\theta})$. Alors,

$$\left(\frac{\sigma^2}{\omega} \chi_k^2 \times \sigma^2 \chi_{n-k}^2 \right) (\{(t, u) : g_c^*(t, u) \geq \omega\}) = \mathfrak{F}_{k, n-k}(\{z : 0 < z \leq c(n-k)/k\}).$$

Un choix raisonnable pour la constante $\frac{c(n-k)}{k}$ apparaît être la médiane de la distribution $\mathfrak{F}_{k, n-k}$ pour les valeurs suivantes de k et $n-k$:

$$\begin{array}{l|l|l|l} k = 4 & n-k \geq 5 & 6 \leq k \leq 7 & n-k \geq 6 \\ k = 5 & n-k \geq 8 & 8 \leq k \leq 13 & n-k \geq 5 \end{array} \quad \left| \begin{array}{l} k \geq 14 \\ n-k \geq 4 \end{array} \right.$$

(les autres valeurs de k et $n-k$ ne sont pas acceptées).

La figure 1 ci-dessous, montre la fonction de "Relative Saving Loss" donnée par la formule

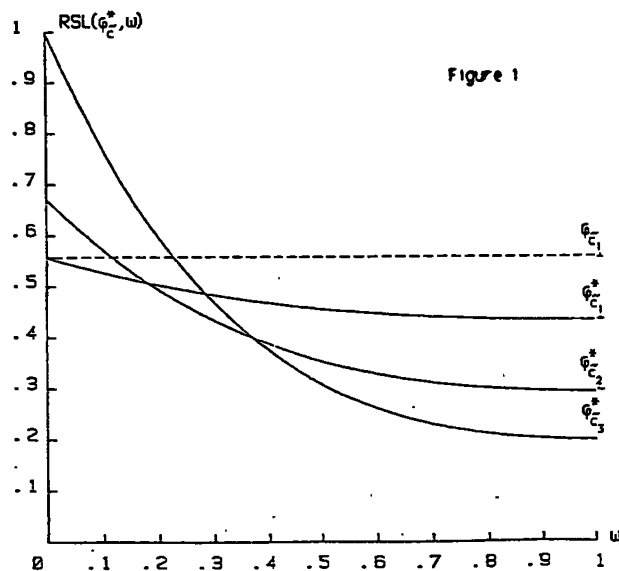
$$RSL(\varphi_{\bar{c}}^*, \omega) = RSL(\varphi_{\bar{c}}) - \int_0^{+\infty} \left\{ \left[\frac{\bar{c}^2 y^2}{k(k-2)(n-k)^2} - \frac{2\bar{c}y}{k(n-k)} - \frac{1}{\omega} \left(\frac{1}{\omega} - 2 \right) \right] I_{k/2} \left(\frac{\bar{c}\omega y}{2(n-k)} \right) + \left[\frac{\bar{c}^2 y^2}{k(k-2)(n-k)^2} + \left(\frac{1}{\omega} - 2 \right) \frac{\bar{c}y}{k(n-k)} \right] i_{k/2} \left(\frac{\bar{c}\omega y}{2(n-k)} \right) \right\} \chi_{n-k}^2(dy)$$

pour les trois valeurs de \bar{c} (notant ici la constante $c(n-k)$)

$$\bar{c}_1 = \frac{(n-k)(k-2)}{n-k+2}, \quad \bar{c}_2 = \frac{3(n-k)(k-2)}{2(n-k+2)}, \quad \bar{c}_3 = 2 \frac{(n-k)(k-2)}{n-k+2}$$

dans le cas où $k = 4$ et $n-k = 16$.

On peut voir qu'il existe un bon cas si on utilise $\bar{c} > \bar{c}_2$. Plus précisément, si ω est près de 1, l'amélioration sur le risque bayésien apportée par les estimateurs "bayésiens tronqués", $\varphi_{\bar{c}}^*$, devient plus grand que celle de l'estimateur de James-Stein de partie positive $\varphi_{\bar{c}_1}^*$ comme \bar{c} approche à 3.5556 (i.e. quand c approche à 0.222.).



$k=4, n-k=16, \bar{c}_1=1.7778, \bar{c}_2=2.6667, \bar{c}_3=3.5556$

2) Soit $\pi = \pi_1 \times \pi_2$ où π_1 est a.c. par rapport à la mesure de Lebesgue sur \mathbb{R}_+ de densité proportionnelle à

$$\sigma \rightarrow 1_{[0, \sigma_0]}(\sigma) \sigma^{-b} \quad (\text{avec } b < -1)$$

et π_2 est a.c. par rapport à la mesure de Lebesgue sur $]0, 1[$, de densité proportionnelle à $\omega \rightarrow \omega^{-\alpha}$ (avec $\alpha < 1$) ; (en posant $\sigma_0 = \frac{1}{\gamma}$ et $b = -2(\nu-1)$ on retrouve la densité de probabilité $g(t) = ct^{-(\nu+1/2)}$ si $t \geq \gamma^2$ (avec $\nu > \frac{3}{2}$) qui est celle de Strawderman en [Strawderman, E.W., 1973]. La densité sur ω a été utilisée par beaucoup d'auteurs (voir par exemple en [Alam K., 1973], [Efron B. et Morris C., 1973], [Strawderman E.W., 1971] et [Strawderman E.W., 1973]). Alors d'après le Lemme 3.1., puisque $\int_{\mathbb{R}_+^*} \frac{1}{\sigma^2} \pi_1(d\sigma) < +\infty$, l'estimateur de Bayes associé est de rétrécisseur

$$f^*(y) = g^*(t, u) = \frac{1}{t} \frac{\int_{[0, \sigma_0] \times]0, 1[} \omega^{k/2 - \alpha + 1} \sigma^{-(n+b+2)} \exp\left\{-\frac{1}{2\sigma^2}(\omega+u)\right\} d\sigma d\omega}{\int_{[0, \sigma_0] \times]0, 1[} \omega^{k/2 - \alpha} \sigma^{-(n+b+2)} \exp\left\{-\frac{1}{2\sigma^2}(\omega+u)\right\} d\sigma d\omega} \quad (4.8)$$

Strawderman en [16], a démontré que ces rétrécisseurs satisfont aux hypothèses (i) et (ii) de la Proposition 2.1., et a donné une majoration de $\frac{1}{u} g^*(t, u)$ qui permet de considérer le cas où l'hypothèse (iii) est aussi satisfaite. Nous reprenons ses calculs et nous résumons les résultats dans le théorème suivant.

Théorème 4.2.

L'estimateur $\varphi^*(y) = [1 - f^*(y)]\varphi_0(y)$ avec f^* donné par (4.8) (où $t = \frac{-1}{v_{\Theta}}$ ($\varphi_0(y)$) et $u = v_{\Theta}^{-1}(y - \varphi_0(y))$) est l'estimateur de Bayes de θ relativement à la probabilité a priori π (qui dépend de α et b comme il a été indiqué au début de l'exemple) ; il est de risque inférieur ou égal à celui de φ_0 si on peut choisir k satisfaisant

$$k \geq 5, \quad 1 > \alpha \geq \frac{k^2 - nk + 6n + 4k - 12 - 2(1+b)(k-2)}{2(n+k-2)}$$

En particulier des estimateurs bayésiens de risque uniformément inférieur à celui de φ_0 existent pour $k \geq 5$ quand $n-k > \frac{6k-8}{k-4}$, c'est-à-dire

$n-k \geq 23$ pour $k = 5$	$n-k \geq 10$ pour $k = 9$
$n-k \geq 15$ pour $k = 6$	$n-k \geq 9$ pour $10 \leq k \leq 12$
$n-k \geq 12$ pour $k = 7$	$n-k \geq 8$ pour $13 \leq k \leq 20$
$n-k \geq 11$ pour $k = 8$	$n-k \geq 7$ pour $k \geq 21$

Remarque : Le tableau de validité donné dans le théorème ci-dessus généralise celui de Strawderman en [Strawderman E.W., 1973], qui se limitait au cas où n est un multiple de k . Un calcul détaillé pour cet exemple est donné en [Criticou, D., 1986].

REFERENCES

- [1] **Alam, K.** A family of admissible minimax estimators of the mean of a multivariate normal distribution. *Ann. Stat.* 1 (3), 517-525, 1973.
- [2] **Alam, K.** Minimax and admissible minimax estimators of the mean of a multivariate normal distribution for unknown covariance matrix. *Journal of Multivariate Analysis*, 5, 83-95, 1975.
- [3] **Baranchick, A.** Multiple regression and estimation of the mean of a multivariate normal distribution. *Techn. Rep. N° 51*, Stanford Univ., 1964.
- [4] **Berger, J.O.** *Statistical Decision Theory and Bayesian Analysis*. 2nd ed. Springer Verlag, New York, 1985.
- [5] **Criticou, D.** Estimateurs à rétrécisseurs (cas de distributions normales). Une classe d'estimateurs bayésiens. Doctorat, Univ. de Rouen, 1986.
- [6] **Efron, B. et Morris, C.** Limiting the risk of Bayes and empirical Bayes estimators - Part. II : The empirical Bayes case. *JASA*, vol. 67, n° 337, 1972.
- [7] **Efron, B. et Morris, C.** Stein's estimation rule and its competitors. An empirical Bayes approach. *JASA*, vol. 68, n° 343, 1973.

- [8] **Judge, G.G. et Bock, E.M.** The Statistical implications of Pre-Test and Stein-Rule Estimators in Econometrics. North-Holland, 1978.
- [9] **Lin, P.E. et Tsai, H.L.** Generalized Bayes minimax estimators of the multivariate normal mean with unknown covariance matrix. *Ann. Stat.* 1(1), 142-145, 1973.
- [10] **Lindley, D.V. et Smith, A.-F.M.** Bayes Estimates for the Linear Model. *J. Royal Statist. Soc. B.* 34, 1-42, 1972.
- [11] **James, W. et Stein, C.** Estimation with quadratic loss. *Proc. Fourth Berk. Symp. Math. Prob.* 1, 361-379, 1961.
- [12] **Roberts, C.** An explicit formula for the risk of the positive-part James-Stein estimator. *Canadian J. Stat.* Paru. 1988.
- [13] **Spruill, C.M.** Some approximate restricted Bayes estimators of a normal mean. *Statistics & Decisions* 4, 337-351, 1986.
- [14] **Stein, C.** Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc., Third Berk. Symp.* 1, 197-206, 1956.
- [15] **Strawderman, E.W.** Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Stat.* 42 (1), 385-388, 1971.
- [16] **Strawderman, E.W.** Proper Bayes minimax estimators of the multivariate normal mean vector for the case of common unknown variances. *Ann. Stat.* 1(6), 1189-1194, 1973.