

**Bayes Estimators Associated with Uniform Distributions on Spheres (II):
the Hierarchical Bayes Approach**

by Christian Robert, Mary Ellen Bock and George Casella
Cornell University and Purdue University

BU-1002-M

January 1989

1. Introduction

As it has been seen in Bock and Robert (1988), Bayes estimators associated with uniform distributions on spheres can be used for the estimation of a normal mean. The use can be either in a direct or indirect way. The *empirical Bayes* version of these estimators produces an estimator which practically dominates the positive-part James-Stein estimator. In fact, it performs better than the positive-part James-Stein estimator for the values of the parameters which are of interest. However, from a theoretical point of view, the lack of 'tail domination' is still a drawback.

In this paper, we will present a second analysis of this rich family of Bayes estimators by using another approach, *hierarchical Bayes* analysis. In fact, if empirical Bayes estimators are often preferred in practice for their simplicity, it is well known that a hierarchical approach is generally preferable as it usually leads to robustness (see Berger (1985), Brown (1987), Berger and Robert (1988)).

For this particular problem, the hierarchical Bayes approach gives a broader scope of possibilities than the empirical Bayes approach. In fact, we will restrict ourselves to a special class of hyperpriors, i.e., the class of discrete hyperpriors. These priors give positive probability to a countable number of radii. Hierarchical Bayes estimators are then simply mixtures of Bayes estimators related to uniform distributions on spheres. However, this restricted family is rich enough to produce very interesting results. It is a *relatively essentially complete* family (Robert, 1988) in the sense that the closure of this family is essentially complete as any spherically symmetric Bayes estimator can be approximated by one of these mixtures and spherically symmetric Bayes estimators form an essentially complete class (Berger and Srivivasan, 1978). Furthermore, Kempthorne (1986, 1988a, 1988b) has stressed their importance in *compromise* and *restricted Bayes* decision rules. We will see in Section 3 how his results imply that the hierarchical Bayes estimators lack some undesirable properties of the original estimators; the resulting estimators are minimax and still admissible.

Another consequence of these completeness results deals with the domination of the positive-part James-Stein estimator. It is well known that this estimator is 'nearly admissible' (Efron and Morris (1973)) and that usual techniques such as integration by parts are useless in its case (Bock (1988), Brown (1988)). Through *tail minimarity* arguments (Berger (1976), Bock and Robert (1988)) we give, in Section 4, sufficient conditions for *tail equivalence* of the positive-part estimator with mixtures of Bayes estimators. In fact, *tail domination* of the positive-part estimator does not seem possible.

These estimators have still another interesting feature. It is well known that, for bounded parameter spaces, they are the "natural" estimators (Casella and Strawderman (1981), Bickel (1981), DasGupta (1985), Kempthorne (1988a)). In Section 2 we give a few results related to this property in order to make the connection with the tail results of Section 4.

2. Bounded parameter space

2.1. **Model.** We observe a normal random variable, $\mathbf{x} \sim N_p(\theta, I_p)$ with mean $\theta \in \Theta$. In this section (only), we will suppose that Θ is a compact set in \mathbb{R}^p ; in the next sections, $\Theta = \mathbb{R}^p$.

The usual least squares estimator, $\phi^\circ(\mathbf{x}) = \mathbf{x}$, is not minimax in this context as it is dominated by the maximum likelihood estimator $\delta_0(\mathbf{x}) = \mathbf{x}$ if $\mathbf{x} \in \Theta$, $P_\Theta(\mathbf{x})$ otherwise (where $P_\Theta(\mathbf{x})$ is the orthogonal projection of \mathbf{x} on Θ). However, δ_0 itself is neither minimax nor admissible (for lack of smoothness).

The usual way to find a minimax estimator for this problem is to determine a *least favorable prior*. This technique has been used by the authors who have studied the problem. Casella and Strawderman (1981) have considered the case $p = 1$ and $\Theta = [-m, m]$; and they have established that, when $m < 1.05$, the prior distribution which gives probability $\frac{1}{2}$ to $-m$ and m is least favorable. For larger values of m , the least favorable prior is still symmetric with discrete support but the number of points in this support increase. Kempthorne (1988a) gives an algorithm for the computation of the support of the minimax estimator for a given m (see also Eichenauer *et al.* (1988)). Bickel (1981) also studied this problem for large values of m , in order to evaluate the asymptotic behavior of the minimax risk. Note that, for this problem, the minimax estimator is unique (Ghosh (1964)). Das Gupta (1985) and Kempthorne (1988b) deal with higher dimensions. If Θ is a ball, it appears that the least favorable prior distribution will be a finite mixture of uniform distributions on spheres (shells) with the same center as Θ . It is not clear whether the Ghosh (1964) result can be generalized to higher dimensions, i.e., that the minimax estimator is unique for $p > 1$.

2.2. **Estimators.** If π_c is the uniform distribution on the sphere $\{\theta; \|\theta\|=c\}$, the Bayes estimator associated with π_c is

$$\delta_c(\mathbf{x}) = c r_{\frac{p}{2}-1}(\|\mathbf{x}\|c) \frac{\mathbf{x}}{\|\mathbf{x}\|} \tag{2.1}$$

where $r_{\frac{p}{2}-1}$ is the ratio of the modified Bessel functions $I_{\frac{p}{2}}$ and $I_{\frac{p}{2}-1}$ (see Robert (1988) and Bock and Robert (1988)).

If $\Theta = \{\theta; \|\theta\| \leq \rho\}$ is the ball of radius ρ , let us consider a finite sequence $0 \leq c_1 < \dots < c_n \leq \rho$ of radii and the corresponding sequence q_1, \dots, q_n of weights such that $\sum_{i=1}^n q_i = 1$. Thus the prior distribution $\pi = \sum_{i=1}^n q_i \pi_{c_i}$ is a hierarchical Bayes prior distribution in the sense that the first stage prior is π_c and the second stage prior on c (or *hyperprior*) is $\sum_{i=1}^n q_i \delta_{c_i}$. In other words, the sphere of radius c_i is chosen with probability q_i and then θ is chosen on the sphere of radius c_i according to a uniform distribution. The Bayes estimator associated with π is (George (1986))

$$\delta^\pi(x) = \sum_{i=1}^n \frac{q_i m_{c_i}(x)}{\sum_{j=1}^n q_j m_{c_j}(x)} \delta_{c_i}(x), \quad (2.2)$$

where m_c is the marginal density of π_c ,

$$m_c(x) \propto e^{-\|x\|^2/2} e^{-c^2/2} \frac{I_{\frac{p}{2}-1}(\|x\|c)}{(c\|x\|)^{\frac{p}{2}-1}} \quad (2.3)$$

(see Robert (1988)). Using obvious notation, δ^π can also be written

$$\delta^\pi(x) = \sum_{i=1}^n \tilde{q}_i(x) \delta_{c_i}(x). \quad (2.4)$$

This formula expresses the fact that $\tilde{q}_i(x)$ is the *posterior probability* of the sphere of radius c_i ; the prior information (given by q_1, \dots, q_n) is reconsidered in the light of the observation, x .

Remark 1. m_c is defined for $c=0$ as $\lim_{t \rightarrow 0} t^{-\nu} I_\nu(t) = 2^{-\nu} / \Gamma(\nu+1)$.

It is not possible to determine where $\tilde{q}_i(x)$ is maximal, due to the complexity of \tilde{q}_i . However, it is likely that this maximum is obtained for $\|x\|$ close to c_i . We have, for $\tilde{q}_i(x) = \tilde{q}_i(\|x\|)$,

Lemma 2.1. $\frac{\partial}{\partial t} \tilde{q}_i(t) = \tilde{q}_i'(t) = \tilde{q}_i(t) \left\{ c_i r_{\frac{p}{2}-1}(tc_i) - \sum_{j=1}^n \tilde{q}_j(t) c_j r_{\frac{p}{2}-1}(tc_j) \right\}$

Proof: From (2.3), we deduce that

$$\tilde{q}_i(t) = q_i c_i^{-\nu} e^{-c_i^2/2} I_\nu(tc_i) / \sum_{j=1}^n q_j c_j^{-\nu} e^{-c_j^2} I_\nu(tc_j),$$

with $\nu = \frac{p}{2}-1$. And

$$\tilde{q}'_i(t) = q_i c_i^{-\nu} e^{-c_i^2/2} c_i \frac{I_{\nu+1}(tc_i) + \frac{\nu}{tc_i} I_\nu(tc_i)}{\sum_{j=1}^n q_j c_j^{-\nu} I_\nu(tc_j)}$$

$$- \tilde{q}_i(t) \sum_{j=1}^n \frac{q_j c_j^{-\nu} e^{-c_j^2/2} c_j \left\{ I_{\nu+1}(tc_j) + \frac{\nu}{tc_j} I_\nu(tc_j) \right\}}{\sum_{k=1}^n q_k c_k^{-\nu} e^{-c_k^2/2} I_\nu(tc_k)}$$

$$= \tilde{q}_i(t) \left\{ c_i r_\nu(rc_i) - \sum_{j=1}^n \tilde{q}_j(t) c_j r_\nu(rc_j) \right\} .$$

□

Remark 2. In order to determine if δ^π , defined in (2.4), is minimax, one has to check if the Bayes risk of δ^π , $r(\pi)$, is equal to the maximum of the frequentist risk, $R(\delta^\pi, \theta)$, on Θ . Once again, given the complexity of the estimator, this verification can only be done numerically.

For reasons given in the next section (Corollary 3.3), it appears that there exist sequences (c_1, \dots, c_n) and (q_1, \dots, q_n) such that the resulting estimator δ^π dominates the positive-part James-Stein estimator

$$\phi_{pp}^+(x) = \left(1 - \frac{p-2}{\|x\|^2} \right)^+ x \tag{2.5}$$

on Θ , as ϕ_{pp}^+ is not admissible. Therefore, for a fixed radius ρ , it is possible to find such sequences. Figure 1 gives the comparison of the risks of ϕ_{pp}^+ and δ^π for $p=5$ and $\|x\| \leq 5.5$; the radii are 0.5, 2.5, 4 and 5.5. The weights q_i have been determined by trial and error. Note that the resulting estimator, δ^π , is not necessarily minimax on Θ as the purpose was to dominate ϕ_{pp}^+ on Θ . A starting point for the choice of the radii c_i could still be the radii of the minimax estimator (which can be approximated using Kempthorne (1988a).) As the

minimax estimator δ_* has a risk equal to the minimax risk for all the points of its support, δ_* cannot dominate $\phi_{\mathbf{p}\mathbf{p}}^+$. But it is still possible that another choice of the weights q_i leads to the uniform domination of $\phi_{\mathbf{p}\mathbf{p}}^+$ on Θ ...

Remark 3. Strictly speaking, it is pointless to consider the domination of $\phi_{\mathbf{p}\mathbf{p}}^+$ on Θ as this estimator is dominated by its truncated counterpart. However, in practice, it is rarely the case that the radius ρ is known precisely. Thus, people use $\phi_{\mathbf{O}}$ and $\phi_{\mathbf{p}\mathbf{p}}^+$ rather than $\delta_{\mathbf{O}}$. Further, if ρ is large, restricted and unrestricted estimators have approximately the same behavior. Here, we are mainly interested in the domination of $\phi_{\mathbf{p}\mathbf{p}}^+$ on $\Theta = \mathbb{R}^P$; this numerical determination of the radii and the weights is only a first step (see Section 4).

3. General Case

When $\Theta = \mathbb{R}^P$, the finite mixtures of the previous section are replaced by infinite mixtures. Their interest is twofold. For some choices of the sequences, the resulting estimator is minimax and admissible. Therefore, the mixing is strong enough to eliminate some of the undesirable behavior of the single estimators. Secondly, these mixtures span the entire class of spherically symmetric Bayes estimators and, therefore, can be used to approximate any of these estimators.

3.1. **Compromise decision problem.** One reason why these infinite mixtures of uniform distributions on spheres are so important is because they are closely related to some *least favorable prior distributions* (see Berger (1985), DasGupta (1985)). This is also why infinite mixtures are related to Kempthorne's results.

Kempthorne (1986, 1988a, 1988b) considers the following *compromise decision problem*: given an estimator δ_0 of θ and a spherically symmetric prior distribution π_0 on Θ , find an estimator δ_* solving

$$\min_{\delta \in D} \int_{\Theta} R(\theta, \delta) \pi_0(d\theta) \quad (3.1)$$

where $D = \{ \delta; R(\theta, \delta) \leq R(\theta, \delta_0) \text{ for every } \theta \}$ and R is the risk associated with the usual quadratic loss. Of particular interest is the case where δ_0 is the least square estimator because the solution of (3.1) is then minimax. The relation with infinite mixtures is given by the following result.

Theorem 3.1. (Kempthorne (1988)). *The estimator δ_* solving (3.1) is admissible and generalized Bayes with respect to a measure π^* . The prior measure π^* is a mixture of π_0 and a possibly improper measure π_λ , a mixture of uniform distributions on shells centered at the origin whose radii have no accumulation point. Furthermore, if the mixture gives positive weight to π_λ , the risks of δ_0 and δ_* are equal at all θ in the support of π_λ .*

This very powerful theorem has immediate consequences for our problem. If we take

$\delta_{\mathbf{0}}$ to be the least squares estimator and $\pi_{\mathbf{0}}$ to be the point mass at zero, we have the following corollary.

Corollary 3.2. *There exist a sequence of radii $c_0=0 < c_1 < \dots$ with no accumulation point and a sequence of weights q_0, q_1, \dots such that the Bayes estimator corresponding to the (possibly improper) prior distribution $\sum_{i=0}^{\infty} q_i \pi_{c_i}$ is minimax and admissible.*

Note that the prior will be *improper* if and only if $\sum_{i=0}^{\infty} q_i = +\infty$. This result shows how important are the weights $\tilde{q}_i(x)$ (as it can be seen that formula (2.4) can be generalized to the case of infinite mixtures) and indicates, once more, the power of hierarchical Bayes techniques. However, due to the “wave property” of the risk indicated in Theorem 3.1, the estimator given in Corollary 3.2 cannot dominate the positive-part James-Stein estimator ϕ_{pp}^+ of (2.5). But, replacing $\delta_{\mathbf{0}}$ by ϕ_{pp}^+ (as ϕ_{pp}^+ is not admissible), we have the following.

Corollary 3.3. *There exist mixtures of uniform distributions on spheres which give generalized Bayes admissible estimators dominating the positive-part James-Stein estimator.*

Furthermore, this result allows us to take any arbitrary radius in the sequence, as there is no constraint on the prior distribution $\pi_{\mathbf{0}}$. Corollary 3.3 justifies the last developments of Section 2.2. We are sure that such radii and weights exist. Moreover, for a bounded parameter space, there are necessarily a finite number of radii. One can also note some indication of the “wave phenomenon” on Figure 1, as the risk of the Bayes estimator hits the risk of the positive-part James-Stein estimator. However, a major problem remains unsolved: the determination of the sequences (c_i) and (q_i) . We will see in Section 4 that we can determine these sequences for large values of i . It seems that the only way to find the first terms of the sequences is numerical. Kempthorne (1988a) also indicates a possible algorithm based on the fact that the risks must be equal on the support of π (see also Kempthorne (1987)).

3.2. **Complete class properties.** As the following result indicates, we can approximate any spherically symmetric estimator by a Bayes estimator associated with a mixture:

Theorem 3.4. [Robert (1988)]. For an arbitrary spherically symmetric distribution π , there exist sequences (q_i^n) and (c_i^n) such that, for every x ,

$$\delta^\pi(x) = \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} \tilde{q}_i^n(x) \delta_{c_i^n}(x) .$$

Furthermore, if we impose certain conditions on the sequence (c_i^n) , it is possible to determine the sequence (q_i^n) in terms of (c_i^n) (Robert (1988)):

$$q_i^n = e^{\left(\frac{(c_i^n)^2 - (c_{i+1}^n)^2}{2} \int_{c_i^n}^{c_{i+1}^n} f(r) r^{p-1} dr \right)} , \quad (3.2)$$

if the density of the prior distribution factors through f . This constructive aspect of the result is its main interest, as it allows the approximation of complicated Bayes estimators, like most hierarchical Bayes estimators (see Berger (1985) for examples). Note that the integrals (3.2) are usually more easily computable than the Bayes estimators themselves (which depend on two integrals).

4. Tail Comparison

Given an infinite mixture, $\sum_{i=1}^{\infty} q_i \pi_{c_i}$, we will give sufficient conditions on (q_i) and (c_i) so that the associated Bayes estimator has a risk asymptotically equivalent to the risk of the positive-part James-Stein estimator. In order to achieve this aim, we use “tail minimaxity” results. This technique has been introduced in Berger (1976) and has already been of use in the present context in Bock and Robert (1988).

4.1. Preliminary results. The major result about tail minimaxity is the following one:

Theorem 4.1 (Berger, (1976)). *If $\delta(x) = \left(1 - \frac{g(\|x\|)}{\|x\|^2}\right)x$ and if g satisfies some regularity conditions,*

$$R(\delta, \|\theta\|) = p - 2 \frac{g'(\|\theta\|)}{\|\theta\|} - \frac{g(\|\theta\|)}{\|\theta\|^2} \{2(p-2) - g(\|\theta\|)\} + o(\|\theta\|^{-2}).$$

This formula gives an approximation of the risk for large values of $\|\theta\|$ and can be used to establish tail minimaxity. Also, it reveals itself to be quite powerful for the comparison of two estimators when the usual techniques fail (see Bock and Robert, 1988). This is, in particular, the case for comparison with the positive-part James-Stein estimator (2.5).

For the mixture estimator $\sum_{i=1}^{+\infty} \tilde{q}_i(x) \delta_{c_i}(x)$, Theorem 4.1 also gives the following expression for the risk of the mixture:

Lemma 4.2: *If $\pi = \sum_{i=1}^{+\infty} q_i \pi_{c_i}$, then, for $\nu = \frac{p}{2} - 1$,*

$$R(\delta^\pi, \|\theta\|) = \sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) R(\delta_{c_i}, \|\theta\|) + \sum_{i=1}^{\infty} \tilde{q}_i(\|\theta\|) [\|\theta\| - c_i r_\nu(c_i \|\theta\|)] \left\{ \sum_{j=1}^{+\infty} \tilde{q}_j(\|\theta\|) c_j r_\nu(\|\theta\| c_j) - c_i r_\nu(\|\theta\| c_i) \right\} + o(\|\theta\|^{-2}).$$

Proof: We can always write $\delta_{c_i}(x) = \left(1 - \frac{g_i(\|x\|)}{\|x\|^2}\right)x$ and

$$\delta^\pi(x) = \sum_{i=1}^{+\infty} \tilde{q}_i(x) \delta_{c_i}(x) = \left[1 - \left(\sum_{i=1}^{+\infty} \tilde{q}_i(x) g_i(\|x\|)\right) \frac{1}{\|x\|^2}\right]x,$$

since $\sum_{i=1}^{\infty} \tilde{q}_i(\mathbf{x}) = 1$. Thus, from Theorem 4.1,

$$\begin{aligned} \mathbf{R}(\delta^\pi, \|\theta\|) &= p - 2\|\theta\|^{-1} \left\{ \sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \mathbf{g}_i(\|\theta\|) \right\}' \\ &\quad - \|\theta\|^{-2} \sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \mathbf{g}_i(\|\theta\|) \left\{ 2(p-2) - \sum_{j=1}^{+\infty} \tilde{q}_j(\|\theta\|) \mathbf{g}_j(\|\theta\|) \right\} + o(\|\theta\|^{-2}) \\ &= \sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \mathbf{R}(\delta_{\mathbf{c}_i}, \|\theta\|) - 2\|\theta\|^{-1} \sum_{i=1}^{+\infty} \tilde{q}_i'(\|\theta\|) \mathbf{g}_i(\|\theta\|) \\ &\quad + \|\theta\|^{-2} \left(\sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \mathbf{g}_i(\|\theta\|) \right)^2 - \|\theta\|^{-2} \sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \mathbf{g}_i^2(\|\theta\|) + o(\|\theta\|). \end{aligned}$$

From Lemma 2.1,

$$\tilde{q}_i'(\|\theta\|) = \tilde{q}_i(\|\theta\|) \left\{ \mathbf{c}_i \mathbf{r}_\nu(\|\theta\| \mathbf{c}_i) - \sum_{j=1}^{+\infty} \tilde{q}_j(\|\theta\|) \mathbf{c}_j \mathbf{r}_\nu(\|\theta\| \mathbf{c}_j) \right\}$$

and, since

$$\mathbf{g}_i(t) = t \left\{ t - \mathbf{c}_i \mathbf{r}_\nu(t \mathbf{c}_i) \right\},$$

$$\begin{aligned} \mathbf{R}(\delta^\pi, \|\theta\|) &= \sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \mathbf{R}(\delta_{\mathbf{c}_i}, \|\theta\|) + 2 \sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) (\mathbf{c}_i \mathbf{r}_\nu(\|\theta\| \mathbf{c}_i))^2 \\ &\quad - 2 \left(\sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \mathbf{c}_i \mathbf{r}_\nu(\|\theta\| \mathbf{c}_i) \right)^2 + \left(\sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \mathbf{c}_i \mathbf{r}_\nu(\|\theta\| \mathbf{c}_i) \right)^2 \\ &\quad - \sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) (\mathbf{c}_i \mathbf{r}_\nu(\|\theta\| \mathbf{c}_i))^2 + o(\|\theta\|^{-2}). \quad \square \end{aligned}$$

Note that, as $\sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) = 1$,

$$\sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) (\mathbf{c}_i \mathbf{r}_\nu(\|\theta\| \mathbf{c}_i))^2 \leq \left(\sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \mathbf{c}_i \mathbf{r}_\nu(\|\theta\| \mathbf{c}_i) \right)^2,$$

and, thus, that $\mathbf{R}(\delta^\pi, \|\theta\|) \geq \sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \mathbf{R}(\delta_{\mathbf{c}_i}, \|\theta\|)$. The introduction of $\|\theta\|$ in the second sum is not necessary (as $\sum_{i=1}^{+\infty} \tilde{q}_i(\|\theta\|) \left\{ \sum_{j=1}^{+\infty} \tilde{q}_j(\|\theta\|) \mathbf{c}_j \mathbf{r}_\nu(\|\theta\| \mathbf{c}_j) - \mathbf{c}_i \mathbf{r}_\nu(\|\theta\| \mathbf{c}_i) \right\} = 0$).

However, it provides an easier resolution of the following developments.

For a given radius c , we have the following representation for the Bayes estimator.

Lemma 4.3: $R(\delta_c, t) = -p + 2c^2 + t^2 - 2tcr_\nu(tc) - c^2r_\nu(tc) + o(t^{-2})$.

Proof: Since $g(t) = t(t - cr_\nu(tc))$, substitution using Theorem 4.1 yields the desired result. \square

Furthermore, using the following expansion (Robert, 1988),

$$r_\nu(t) = 1 - \frac{2\nu+1}{2t} + \frac{(2\nu+1)(2\nu-1)}{8t^2} + \frac{(2\nu+1)(2\nu-1)}{8t^3} + o(t^{-3}), \quad (4.1)$$

and substituting using Lemma 4.3 yields the following representation.

Lemma 4.4: For the Bayes estimator δ_c ,

$$R(\delta_c, t) = -1 + (c-t)^2 + \frac{1}{t} \left((p-1)c - \frac{(p-1)(p-3)}{4c} \right) - \frac{(p-1)(p-2)}{2t^2} - \frac{(p-1)(p-3)}{4t^2c^2} + r_m\left(\frac{1}{t^2}\right).$$

It is also a direct consequence of Theorem 4.1 that the risk of the positive-part estimator is equal to

$$R(\phi_{pp}^+, t) = p - \frac{(p-2)^2}{t^2} + o(t^{-2}). \quad (4.2)$$

4.2. The integral approximation. In order to get conditions under which $R(\delta^\pi, \|\theta\|) \leq R(\phi_{pp}^+, \|\theta\|)$ is satisfied, we will replace the sums of the previous sections with integrals. This approximation is justified for a particular choice of the sequences (q_i) and (c_i) .

First, it follows from Lemma 4.2, Lemma 4.4 and (4.2) that

$$\begin{aligned} R(\phi_{pp}^+, t) - R(\delta^\pi, t) &= \sum_{i=1}^{+\infty} \tilde{q}_i(t) \left\{ p+1 - \frac{(p-2)^2}{t^2} - (c_i-t)^2 - \frac{1}{t} \left\{ (p-1)c_i - \frac{(p-1)(p-3)}{tc_i} \right\} \right. \\ &\quad \left. + \frac{(p-1)(p-2)}{2t^2} + \frac{(p-1)(p-3)}{4t^2c_i^2} \right\} + \sum_{i=1}^{+\infty} \tilde{q}_i(t) (t - c_i r_\nu(tc_i)) c_i r_\nu(tc_i) \\ &\quad - \sum_{i=1}^{+\infty} \tilde{q}_i(t) (t - c_i r_\nu(tc_i)) \sum_{j=1}^{+\infty} \tilde{q}_j(t) c_j r_\nu(tc_j) + o(t^{-2}) \end{aligned}$$

and (4.1) implies

$$\begin{aligned}
 R(\phi_{pp}^+, t) - R(\delta^\pi, t) &= \sum_{i=1}^{+\infty} \tilde{q}_i(t) \left\{ p+1 - (c_i-t)^2 - \frac{1}{t} \left[(p-1)c_i - \frac{(p-1)(p-3)}{4c_i} \right] - \frac{(p-2)(p-3)}{2t^2} \right\} \\
 &+ \sum_{i=1}^{+\infty} \tilde{q}_i(t) \left\{ t-c_i + \frac{p-1}{2t} - \frac{(p-1)(p-3)}{8t^2 c_i} \right\} c_i \left\{ 1 - \frac{p-1}{2tc_i} \right\} \\
 &- \sum_{i=1}^{+\infty} \tilde{q}_i(t) \left\{ t-c_i + \frac{p-1}{2t} - \frac{(p-1)(p-3)}{8t^2 c_i} \right\} \sum_{j=1}^{+\infty} \tilde{q}_j(t) c_j \left\{ 1 - \frac{p-1}{2tc_j} \right\} + o(t^{-2}).
 \end{aligned} \tag{4.3}$$

We will see below why this order of approximation will suffice.

Note that

$$\tilde{q}_i(t) = \frac{q_i c_i^{-\frac{p-2}{2}} e^{c_i^2/2} I_{\frac{p}{2}-1}(tc_i)}{\sum_{j=1}^{+\infty} q_j c_j^{-\frac{p-2}{2}} e^{-c_j^2/2} I_{\frac{p}{2}-1}(tc_j)}$$

and (Robert, 1988)

$$I_{\frac{p}{2}-1}(t) = \frac{e^{-t}}{\sqrt{2\pi t}} \left\{ 1 - \frac{(p-1)(p-3)}{8t} + o(t^{-1}) \right\}.$$

Consider now the special case $q_i c_i^{-\frac{p-1}{2}} = (c_{i+1}-c_i)c_i^{-\alpha}$; we impose furthermore that the series $\sum_{i=1}^{+\infty} (c_{i+1}-c_i)^2$ converges.

Remark 4. This condition is totally compatible with the condition $\lim_{i \rightarrow +\infty} c_i = +\infty$. For instance, this is the case when $c_{i+1}-c_i = \frac{1}{i}$, as we get the Riemann series $\sum_{i=1}^{+\infty} i^{-2}$.

Remark 5. For such a choice of (q_i) , α small enough leads to a generalized Bayes estimator.

However, following Theorem 3.1 of Kempthorne (1988a), there is no reason to restrict the study to the class of proper Bayes estimators.

Then, with the above representations we have

$$\tilde{q}_i(t) = \frac{\left(c_{i+1}-c_i\right)c_i^{-\alpha} e^{-(c_i-t)^2/2} \left\{1 - \frac{(p-1)(p-3)}{8tc_i} + o(t^{-1})\right\}}{\sum_{j=1}^{+\infty} \left(c_{j+1}-c_j\right)c_j^{-\alpha} e^{-(c_j-t)^2/2} \left\{1 - \frac{(p-1)(p-3)}{8tc_j} + o(t^{-1})\right\}}$$

and it follows from Appendix 1 that the RHS of (4.3) can be replaced by

$$\begin{aligned} & \frac{1}{D} \int_{c_1}^{+\infty} \left\{ p+1 - \frac{(p-2)(p-3)}{2t^2} - (u-t)^2 - \frac{(p-1)u}{t} + \frac{(p-1)(p-3)}{4ut} + u \left\{ 1 - \frac{p-1}{2tu} \right\} \right. \\ & \quad \cdot \left. \left\{ t-u + \frac{p-1}{2t} - \frac{(p-1)(p-3)}{8t^2u} \right\} \right\} \left\{ 1 - \frac{(p-1)(p-3)}{8tu} \right\} u^{-\alpha} e^{-(u-t)^2/2} du \\ & - \frac{1}{D^2} \int_{c_1}^{+\infty} \left\{ t-u + \frac{p-1}{2t} - \frac{(p-1)(p-3)}{8t^2u} \right\} u^{-\alpha} e^{-(u-t)^2/2} \left\{ 1 - \frac{(p-1)(p-3)}{8tu} \right\} du \\ & \quad \cdot \int_{c_1}^{+\infty} \left\{ 1 - \frac{p-1}{2tu} \right\} \left\{ 1 - \frac{(p-1)(p-3)}{8tu} \right\} u^{-\alpha+1} e^{-(u-t)^2/2} du + o(t^{-2}) \end{aligned} \tag{4.4}$$

where

$$D = \int_{c_1}^{+\infty} \left\{ 1 - \frac{(p-1)(p-3)}{8tu} \right\} u^{-\alpha} e^{-(u-t)^2/2} du .$$

The change of variable $\omega = \frac{u}{t}$ gives

$$\begin{aligned} & \frac{1}{\bar{M}} \int_{\frac{c_1}{t}}^{+\infty} \left\{ p+1 - \frac{(p-2)(p-3)}{2t^2} - t^2(\omega-1)^2 - (p-1)\omega + \frac{(p-1)(p-3)}{4t^2\omega} + \omega \left\{ 1 - \frac{p-1}{2t^2\omega} \right\} \right. \\ & \quad \cdot \left. \left\{ t^2(\omega-1) + \frac{p-1}{2} - \frac{(p-1)(p-3)}{8t^2\omega} \right\} \right\} \left\{ 1 - \frac{(p-1)(p-3)}{8t^2\omega} \right\} \omega^{-\alpha} e^{-(\omega-1)^2t^2/2} d\omega \end{aligned} \tag{4.5}$$

$$\begin{aligned}
 & -\frac{1}{M^2} \int_{\frac{c_1}{t}}^{+\infty} \left\{ -t^2(\omega-1) + \frac{p-1}{2} - \frac{(p-1)(p-3)}{8t^2\omega} \right\} \omega^{-\alpha} e^{-(\omega-1)^2 t^2 / 2} \left\{ 1 - \frac{(p-1)(p-3)}{8t^2\omega} \right\} d\omega \\
 & \cdot \int_{\frac{c_1}{t}}^{+\infty} \left\{ 1 - \frac{p-1}{2t^2\omega} \right\} \left\{ 1 - \frac{(p-1)(p-3)}{8t^2\omega} \right\} \omega^{-\alpha+1} e^{-(\omega-1)^2 t^2 / 2} d\omega + o(t^{-2})
 \end{aligned}$$

where

$$M = \int_{\frac{c_1}{t}}^{+\infty} \left\{ 1 - \frac{(p-1)(p-3)}{8t^2\omega} \right\} \omega^{-\alpha} e^{-(\omega-1)^2 t^2 / 2} d\omega .$$

The justification of the previous handling of $o(\cdot)$'s in (4.3), (4.4) and (4.5) is given by the following lemma:

Lemma 4.6. *For every $\gamma > 0$, $\omega_1 < 1$, $\omega_2 > 1$,*

$$\lim_{t \rightarrow +\infty} t \int_{\frac{c_1}{t}}^{\omega_1} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega = \lim_{t \rightarrow +\infty} t \int_{\omega_2}^{+\infty} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega = 0$$

and

$$\lim_{t \rightarrow +\infty} t \int_{\omega_1}^{\omega_2} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega = \lim_{t \rightarrow +\infty} t \int_{\frac{c_1}{t}}^{+\infty} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega = \sqrt{2\pi} .$$

Proof. It is straightforward to establish

$$0 \leq \int_{\frac{c_1}{t}}^{\omega_1} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega \leq \omega_1 \frac{t^\gamma}{c_1} e^{-(\omega_1-1)^2 t^2 / 2}$$

for any t , and the upper bound is going to 0 as t goes to $+\infty$ for any value of γ . On the other hand,

$$0 \leq t \int_{\omega_2}^{+\infty} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega \leq t \omega_2^\gamma e^{-(\omega_2-1)^2 t^2 / 4} \int_{\omega_2}^{+\infty} e^{-(\omega-1)^2 t^2 / 4} d\omega$$

and the upper bound is also going to 0 as t goes to $+\infty$. If we consider now

$$t \int_{\omega_1}^{\omega_2} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega ,$$

it is bounded from below by $t \int_{\omega_1}^{\omega_2} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega$ and from above by $t \omega_1^{-\gamma} \int_{\omega_1}^{\omega_2} e^{-(\omega-1)^2 t^2 / 2} d\omega$. Further, for every $\omega_1 < 1 < \omega_2$.

$$\lim_{t \rightarrow +\infty} t \int_{\omega_1}^{\omega_2} e^{-(\omega-1)^2 t^2 / 2} d\omega = \lim_{t \rightarrow +\infty} t \int_{-\infty}^{+\infty} e^{-(\omega-1)^2 t^2 / 2} d\omega = \sqrt{2\pi} . \quad \square$$

Note that this result, which implies

$$\lim_{t \rightarrow +\infty} \frac{1}{M} \int_{\frac{c_1}{t}}^{+\infty} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega = 1 ,$$

allows for the manipulation of $o(\cdot)$ outside the integrals and the sums (due to the equivalence shown in Appendix 1). The next result shows precisely the order of approximation of these integrals.

Lemma 4.7. $\frac{t}{\sqrt{2\pi}} \int_{\frac{c_1}{t}}^{+\infty} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega = 1 + \frac{\gamma(\gamma+1)}{2t^2} + o(t^{-2})$ for $\gamma \geq 0$.

Proof. By L'Hospital's rule, we have

$$\begin{aligned} & \lim_{t \rightarrow +\infty} t^2 \left[\frac{t}{\sqrt{2\pi}} \int_{\frac{c_1}{t}}^{+\infty} \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega - 1 \right] \\ &= \lim_{t \rightarrow +\infty} \frac{t^4}{3} \left[\frac{-1}{\sqrt{2\pi}} \frac{c_1}{t^2} \left(\frac{t}{c_1} \right)^\gamma e^{-(c_1-t)^2 / 2} + \frac{t}{\sqrt{2\pi}} \int_{\frac{c_1}{t}}^{+\infty} (\omega-1)^2 \omega^{-\gamma} e^{-(\omega-1)^2 t^2 / 2} d\omega - \frac{1}{t^2} \right] \end{aligned}$$

$$\begin{aligned}
 &= \lim_{t \rightarrow +\infty} \frac{t^2}{3} \left[\frac{t}{\sqrt{2\pi}} \int_{\frac{c_1}{t}}^{+\infty} t^2(\omega-1)^2 \omega^{-\gamma} e^{-(\omega-1)^2 t^2/2} d\omega - 1 \right] \\
 &= \lim_{t \rightarrow +\infty} \frac{t^2}{3} \left[\frac{t}{\sqrt{2\pi}} \int_{\frac{c_1}{t}}^{+\infty} \left\{ \omega^{-\gamma} - \gamma \omega^{-(\gamma+1)} (\omega-1) \right\} e^{-(\omega-1)^2 t^2/2} d\omega - 1 \right]
 \end{aligned}$$

by an integration of parts. Thus

$$\begin{aligned}
 \lim_{t \rightarrow +\infty} t^2 \left[\frac{t}{\sqrt{2\pi}} \int_{\frac{c_1}{t}}^{+\infty} \omega^{-\gamma} e^{-(\omega-1)^2 t^2/2} d\omega - 1 \right] &= \frac{1}{2} \lim_{t \rightarrow +\infty} \frac{t}{\sqrt{2\pi}} \int_{\frac{c_1}{t}}^{+\infty} (-\gamma) t^2 (\omega-1) \omega^{-\gamma-1} e^{-(\omega-1)^2 t^2/2} d\omega \\
 &= \frac{\gamma(\gamma+1)}{2} \lim_{t \rightarrow +\infty} \frac{t}{\sqrt{2\pi}} \int_{\frac{c_1}{t}}^{+\infty} \omega^{-\gamma-2} e^{-(\omega-1)^2 t^2/2} d\omega
 \end{aligned}$$

□

Remark 6. The integration by parts used in the previous proof explains why we do not need to further expand $I_{\frac{p}{2}-1}$ in the terms where $t^2(\omega-1)$ appears. In fact,

$$\lim_{t \rightarrow +\infty} \frac{1}{M} \int_{\frac{c_1}{t}}^{+\infty} t^2 (\omega-1) \omega^{-\gamma} e^{-(\omega-1)^2 t^2/2} d\omega = -\gamma.$$

4.3 Equivalence condition. We will now consider the comparison with ϕ_{pp}^+ for large values of $\|\theta\|$. As we have seen in the previous section, a sufficient condition for the equivalence is that (4.5) must be positive up to the order $o(t^{-2})$ or $MA_1 - A_2A_3 \geq o(t^{-2})$, where

$$\begin{aligned}
 A_1 &= t \int_{\frac{c_1}{t}}^{+\infty} \left\{ p+1 - \frac{(p-2)(p-3)}{2t^2} - t^2(\omega-1)^2 - (p-1)\omega + \frac{(p-1)(p-3)}{4t^2\omega} + \omega \left\{ 1 - \frac{p-1}{2t^2\omega} \right\} \right. \\
 &\quad \cdot \left. \left\{ \frac{p-1}{2} - t^2(\omega-1) - \frac{(p-1)(p-3)}{8t^2\omega} \right\} \right\} \left[1 - \frac{(p-1)(p-3)}{8t^2\omega} \right] \omega^{-\alpha} e^{-(\omega-1)^2 t^2/2} d\omega,
 \end{aligned}$$

$$A_2 = t \int_{\frac{c_1}{t}}^{+\infty} \left\{ -t^2(\omega-1) + \frac{p-1}{2} - \frac{(p-1)(p-3)}{8t^2\omega} \right\} \left\{ 1 - \frac{(p-1)(p-3)}{8t^2\omega} \right\} \omega^{-\alpha} e^{-(\omega-1)^2 t^2/2} d\omega ,$$

$$A_3 = t \int_{\frac{c_1}{t}}^{+\infty} \left[1 - \frac{p-1}{2t^2\omega} \right] \left[1 - \frac{(p-1)(p-3)}{8t^2\omega} \right] \omega^{-\alpha+1} e^{-(\omega-1)^2 t^2/2} d\omega .$$

We have

$$\begin{aligned} & t \int_{\frac{c_1}{t}}^{+\infty} t^2(\omega-1)^2 \left[1 - \frac{(p-1)(p-3)}{8t^2\omega} \right] \omega^{-\alpha} e^{-(\omega-1)^2 t^2/2} d\omega \\ &= t \int_{\frac{c_1}{t}}^{+\infty} \left[1 - \frac{(p-1)(p-3)}{8t^2\omega} \right] \omega^{-\alpha} e^{-(\omega-1)^2 t^2/2} d\omega \\ &+ t \int_{\frac{c_1}{t}}^{+\infty} \left[-\alpha + (\alpha+1) \frac{(p-1)(p-3)}{8t^2\omega} \right] \omega^{-\alpha-1} (\omega-1) e^{-(\omega-1)^2 t^2/2} d\omega + o(t^{-2}) \\ &= t \int_{\frac{c_1}{t}}^{+\infty} \left\{ \left[1 - \frac{(p-1)(p-3)}{8t^2\omega} \right] \omega^{-\alpha} + \alpha(\alpha+1) \frac{\omega^{-\alpha-2}}{t^2} \right\} e^{-(\omega-1)^2 t^2/2} d\omega + o(t^{-2}) . \end{aligned} \quad (4.6)$$

In the same way,

$$\begin{aligned} & t \int_{\frac{c_1}{t}}^{+\infty} t^2(\omega-1) \left(1 - \frac{(p-1)}{2t^2\omega} \right) \left(1 - \frac{(p-1)(p-3)}{8t^2\omega} \right) \omega^{-\alpha+1} e^{-(\omega-1)^2 t^2/2} d\omega \\ &= t \int_{\frac{c_1}{t}}^{+\infty} \left\{ -\alpha+1 + \alpha \frac{(p-1)(p+1)}{8t^2\omega} \right\} \omega^{-\alpha} e^{-(\omega-1)^2 t^2/2} d\omega + o(t^{-2}) \end{aligned} \quad (4.7)$$

and

$$t \int_{\frac{c_1}{t}}^{+\infty} t^2(\omega-1) \left(1 - \frac{(p-1)(p-3)}{8t^2\omega} \right) \omega^{-\alpha} e^{-(\omega-1)^2 t^2/2} d\omega$$

$$= t \int_{\frac{c_1}{t}}^{+\infty} \left(-\alpha + (\alpha+1) \frac{(p-1)(p+1)}{8t^2\omega} \right) \omega^{-\alpha-1} e^{-(\omega-1)^2 t^2/2} d\omega + o(t^{-2}). \quad (4.8)$$

Using these integrations by parts, we get

Theorem 4.8. $\lim_{\|\theta\| \rightarrow +\infty} R(\phi_{pp}^+, \|\theta\|) - R(\delta^\pi, \|\theta\|) = 0$ for every α .

Proof. From (4.6), (4.7)

$$\begin{aligned} \lim_{t \rightarrow +\infty} A_1 &= p+1 - 1 - p+1 + \frac{p-1}{2} + \alpha-1 = \frac{p-1}{2} + \alpha, \\ \lim_{t \rightarrow +\infty} tM &= 1 = \lim_{t \rightarrow +\infty} A_3 \end{aligned}$$

and, from (4.8),

$$\lim_{t \rightarrow +\infty} A_2 = \alpha + \frac{p-1}{2}. \quad \square$$

This shows that, as long as the integral approximations is justified for the choice

$$q_i c_i^{-\frac{p-1}{2}} = (c_{i+1} - c_i) c_i^{-\alpha}, \quad (4.9)$$

the estimators ϕ_{pp}^+ and δ^π have risks equivalent in the tails.

For particular choices of α the equivalence is even stronger. Using (4.6), (4.7) and (4.8), we get the following expressions.

$$\begin{aligned} A_1 &= t \int_{\frac{c_1}{t}}^{+\infty} \left\{ p+1 - \frac{(p-2)(p-3)}{2t^2} - (p-1)\omega + \frac{(p-1)(p-3)}{4t^2\omega} - \frac{(p+1)(p-1)(p-3)}{8t^2\omega} \right. \\ &\quad \left. + \frac{(p-1)^2(p-3)}{8t^2} - 1 + \frac{(p-1)(p-3)}{8t^2\omega} - \alpha(\alpha+1) \frac{1}{t^2\omega^2} + \alpha-1 - \alpha \frac{(p-1)(p+1)}{8t^2\omega} \right\} \omega^{-\alpha-1} e^{-(\omega-1)^2 t^2/2} d\omega \end{aligned}$$

$$\begin{aligned}
& + \omega \frac{p-1}{2} - \frac{(p-1)^2}{4t^2} - \frac{(p-1)(p-3)}{8t^2} - \frac{(p-1)^2(p-3)}{16t^2} \left\} \omega^{-\alpha} e^{-(\omega-1)^2 t^2 / 2} d\omega \\
& = t \int_{\frac{c_1}{t}}^{+\infty} \left\{ p-1 + \alpha + \frac{(p-1)^2(p-3)}{16t^2} - \frac{7p^2-28p+29}{8t^2} - \frac{p-1}{2} \omega \right. \\
& \quad \left. - \frac{1}{\omega t^2} \left[\frac{(p-1)(p-2)(p-3)}{8} + \alpha \frac{(p-1)(p+1)}{8} \right] - \frac{\alpha(\alpha+1)}{t^2 \omega^2} \right\} \omega^{-\alpha} e^{-(\omega-1)^2 t^2 / 2} d\omega, \\
A_2 & = t \int_{\frac{c_1}{t}}^{+\infty} \left\{ \frac{p+1}{2} - \frac{(p+1)(p-1)(p-3)}{16t^2 \omega} + \frac{\alpha}{\omega} - (\alpha+1) \frac{(p-1)(p-3)}{8t^2 \omega^2} \right\} \omega^{-\alpha} e^{-(\omega-1)^2 t^2 / 2} d\omega, \\
A_3 & = t \int_{\frac{c_1}{t}}^{+\infty} \left\{ 1 - \frac{(p-1)(p+1)}{8t^2 \omega} \right\} \omega^{-\alpha+1} e^{-(\omega-1)^2 t^2 / 2} d\omega.
\end{aligned}$$

From Lemma 4.7, it then follows that

$$\begin{aligned}
\frac{tM}{\sqrt{2\pi}} & = 1 + \frac{\alpha(\alpha+1)}{2t^2} - \frac{(p-1)(p-3)}{8t^2} + o(t^{-2}), \\
\frac{A_1}{\sqrt{2\pi}} & = p-1 + \alpha + \frac{(p-1)^2(p-3)}{16t^2} - \frac{7p^2-28p+29}{8t^2} - \frac{p-1}{2} - \frac{1}{t^2} \left\{ \frac{(p-1)(p-2)(p-3)}{8} + \alpha \frac{p^2-1}{8} \right\} \\
& \quad - \frac{\alpha(\alpha+1)}{t^2} + \frac{\alpha(\alpha+1)}{2t^2} (p-1 + \alpha) - \frac{p-1}{4t^2} \alpha(\alpha-1) + o(t^{-2}), \\
\frac{A_2}{\sqrt{2\pi}} & = \frac{p-1}{2} - \frac{(p+1)(p-1)(p-3)}{16t^2} + \alpha - (\alpha+1) \frac{(p-1)(p-3)}{8t^2} \\
& \quad + \frac{(p-1)}{4t^2} \alpha(\alpha+1) + \frac{\alpha(\alpha+1)(\alpha+2)}{2t^2} + o(t^{-2}), \\
\frac{A_3}{\sqrt{2\pi}} & = 1 - \frac{(p-1)(p+1)}{8t^2} + \frac{\alpha(\alpha-1)}{2t^2} + o(t^{-2}).
\end{aligned}$$

Thus

$$\frac{1}{\sqrt{2\pi}} (tMA_1 - A_2A_3)$$

$$\begin{aligned}
&= \left[\frac{p-1}{2} + \alpha - \frac{p-1}{2} - \alpha \right] + \frac{1}{t^2} \left[\frac{(p-1)^2(p-3)}{16} - \frac{7p^2-28p+2q}{8} - \frac{(p-1)(p-2)(p-3)}{8} \right] \\
&\quad - \alpha \frac{p^2-1}{8} - \alpha(\alpha+1) + \frac{\alpha^2(\alpha+1)}{2} + (p-1) \frac{\alpha(\alpha+1)}{2} - \frac{p-1}{4} \alpha(\alpha-1) + \frac{p-1}{4} \alpha(\alpha+1) \\
&\quad - \frac{(p-1)^2(p-3)}{16} + \frac{\alpha^2(\alpha+1)}{2} - \alpha \frac{(p-1)(p-3)}{8} + \frac{(p+1)(p-1)(p-3)}{16} \\
&\quad + (\alpha+1) \frac{(p-1)(p-3)}{8} - \alpha(\alpha+1) \frac{p-1}{4} - \frac{\alpha(\alpha+1)(\alpha+2)}{2} + \frac{(p-1)^2(p+1)}{16} \\
&\quad \left[- \frac{p-1}{4} \alpha(\alpha-1) + \alpha \frac{p^2-1}{8} - \frac{\alpha^2(\alpha-1)}{2} \right] + o(t^{-2}) \\
&= \frac{-\alpha^2 + (p-3)\alpha - \left(\frac{p-3}{2}\right)^2}{t^2} + o(t^{-2}).
\end{aligned}$$

Using these expressions, we obtain the following theorem.

Theorem 4.9. For $\alpha = \frac{p-3}{2}$, (i.e., $q_i \propto (c_{i+1}-c_i)c_i$),

$$\lim_{t \rightarrow +\infty} t^2 \left[R(\phi_{pp}^+, t) - R(\delta^\pi, t) \right] = 0.$$

This result shows that, for this particular choice of α , the two estimates are then equivalent in the tails up to the second order. Using the techniques introduced in Section 2, we can then build an estimator $\delta^\pi(x) = \sum_{i=1}^{+\infty} \tilde{q}_i(x) \delta_{c_i}(x)$ which behaves quite properly for the small values of $\|\theta\|$ (i.e., the values which are actually interesting) and still remains equivalent to the positive-part James-Stein estimator in the tail.

A second consequence of this result is to reinforce the conviction that the positive-part James-Stein estimator is “optimal in the tail”, in the sense that it gives the maximal improvement over the least squares estimator a *minimax* estimator can give.

Appendix 1: Approximation of the Series by the Integral

Let us consider the series $\sum_{i=1}^{+\infty} (c_{i+1}-c_i)c_i^{-\gamma} e^{-(c_i-t)^2/2}$, $\gamma > 0$, arising in (4.2) when

we substitute $(c_{i+1}-c_i)c_i^{\frac{p-1}{2}-\alpha}$ for q_i .

It is a well-known calculus result that this series is of the same kind as the integral

$$\int_{c_1}^{+\infty} u^{-\gamma} e^{-(u-t)^2/2} du, \quad (\text{A.1})$$

that is, that they are both convergent or divergent. In this case, (A.1) is convergent (see Lemma 4.6) and so is the series.

Using a first order Taylor expansion, it follows that

$$\left| \int_{c_i}^{c_{i+1}} u^{-\gamma} e^{-(u-t)^2/2} du - (c_{i+1}-c_i)c_i^{-\gamma} e^{-(c_i-t)^2/2} \right| < \bar{m}_i(t)(c_{i+1}-c_i)^2$$

where

$$\bar{m}_i(t) = \frac{1}{2} \sup_{[c_i, c_{i+1}]} |-\gamma + (u-t)u| u^{-\gamma-1} e^{-(u-t)^2/2}. \quad (\text{A.2})$$

We have then $\bar{m}_i(t) = o(t^{-d})$, for any d and

$$\left| \int_{c_1}^{+\infty} u^{-\gamma} e^{-(u-t)^2/2} du - \sum_{i=1}^{\infty} (c_{i+1}-c_i)c_i^{-\gamma} e^{-(c_i-t)^2/2} \right| < o(t^{-d}) \sum_{i=1}^{\infty} (c_{i+1}-c_i)^2$$

which is still $o(t^{-d})$ as the series is convergent.

5. References

- Berger, J.O. (1976). Tail minimaxity in location vector problems and its applications, *Ann. Statist.* 4(1), 33-50.
- Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer-Verlag.
- Berger, J.O. and Robert, C. (1988). Subjective hierarchical Bayes estimation of a multivariate normal mean: on the frequentist interface. Technical Report #88-24C, Purdue University.
- Berger, J.O. and Srinivasan, C. (1978). Generalized Bayes estimators in multivariate problems. *Ann. Statist.* 6(4), 783-801.
- Bickel, P. (1981). Minimax estimation of the mean of a normal distribution when the parameter space is restricted. *Ann. Statist.* 9(6), 1301-1309.
- Bock, M.E. (1988). Shrinkage estimators: pseudo-Bayes rules for normal mean vectors. In *Statistical Decision Theory and Related Topics IV* (S.S. Gupta and J.O. Berger, eds.). Springer-Verlag.
- Bock, M.E. and Robert, C. (1989). Bayes estimators associated with uniform distributions on spheres (I): the empirical Bayes approach (in preparation).
- Brown, L.D. (1987). An ancillarity paradox which appears in multiple linear regression. Technical Report, Dept. of Mathematics, Cornell University.
- Brown, L.D. (1988). The differential inequality of a statistical estimation problem. In *Statistical Decision Theory and Related Topics IV* (S.S. Gupta and J.O. Berger, eds.). Springer-Verlag.
- Casella, G. and Strawderman, W. (1981). Estimating a bounded normal mean. *Ann. Statist.* 3, 869-876.
- Das Gupta, A. (1985). Bayes minimax estimation in multiparameter families when the parameter space is restricted to a bounded convex set. *Sankhyā* 47, 326-332.

- Eichenauer, J., Kirschgarth, P., and Lehn, J. (1987). Gamma-minimax estimators for a bounded normal mean. *Statistics and Decisions* 6(4), .
- Georges, E. (1986). Minimax multiple shrinkage estimation. *Ann. Statist.* 14(1), 188-205.
- Gosh, M.N. (1964). Uniform approximation of minimax point estimates. *Ann. Math. Statist.* 35, 1031-1047.
- Kempthorne, P.J. (1987). Numerical specification of discrete least favorable prior distributions. *Siam J. S. Stat. Comput.*
- Kempthorne, P.J. (1988a). Dominating inadmissible procedures using compromise decision theory. In *Statistical Decision Theory and Related Topics IV* (S.S. Gupta and J.O. Berger, eds.). Springer-Verlag.
- Kempthorne, P.J. (1988b). Controlling risks under different loss functions: the compromise decision problem. *Ann. Statist.* 16(4), 1594-1608.
- Robert, C. (1988). Modified Bessel functions and their applications in Probability and Statistics. *Statistics and Probability Letters* (to appear).