A Festschrift for Herman Rubin
Institute of Mathematical Statistics
Lecture Notes – Monograph Series
Vol. 45 (2004) 21–44

Estimation in restricted parameter spaces: A review

Eric Marchand*[1] and William E. Strawderman[2]

University of New Brunswick

Rutgers University

Abstract: In this review of estimation problems in restricted parameter spaces, we focus through a series of illustrations on a number of methods that have proven to be successful. These methods relate to the decision-theoretic aspects of admissibility and minimaxity, as well as to the determination of dominating estimators of inadmissible procedures obtained for instance from the criteria of unbiasedness, maximum likelihood, or minimum risk equivariance. Finally, we accompany the presentation of these methods with various related historical developments.

1. Introduction

Herman Rubin has contributed in deep and original ways to statistical theory and philosophy. He has selflessly shared his keen intuition into and extensive knowledge of mathematics and statistics with many of the researchers represented in this volume. The statistical community has been vastly enriched by his contributions through his own research and through his influence, direct and indirect, on the research and thinking of others. We are pleased to join in this celebration in honor of Professor Rubin.

This review paper is concerned with estimation of a parameter or vector parameter θ, when θ is restricted to lie in some (proper) subset of the "usual" parameter space. The approach is decision theoretic. Hence, we will not be concerned with hypothesis testing problems, or with algorithmic problems of calculating maximum likelihood estimators. Excellent and extensive sources of information on these aspects of restricted inference are given by Robertson, Wright and Dykstra (1988), Akkerboom (1990), and Barlow, Bartholomew, Bremner and Brunk (1972). We will not focus either on the important topic of interval estimation. Along with the recent review paper by Mandelkern (2002), here is a selection of interesting work concerning methods for confidence intervals, for either interval bounded, lower bounded, or order restricted parameters: Zeytinoglu and Mintz (1984, 1988), Stark (1992), Hwang and Peddada (1994), Drees (1999), Kamboreva and Mintz (1999), Iliopoulos and Kourouklis (2000), and Zhang and Woodroofe (2003).

We will focus mostly on point estimation and we will particularly emphasize finding estimators which dominate classical estimators such as the Maximum Likelihood or UMVU estimator in the unrestricted problem. Issues of minimaxity and admissibility will also naturally arise and be of interest. Suppose, for example, that the problem is a location parameter problem and that the restricted (and of course

*Research supported by NSERC of Canada.
[1]Department of Mathematics and Statistics, University of New Brunswick, P.O. Box 4400, Fredericton, N.B. Canada, E3B 5A3. e-mail: marchand@math.unb.ca
[2]Department of Statistics, Rutgers University, 501 Hill Center, Busch Campus, Piscataway, NJ 08855, USA. e-mail: straw@rutgers.edu
Keywords and phrases: restricted parameter spaces, minimaxity, admissibility, dominating procedures, maximum likelihood estimation.
AMS 2000 subject classifications: 62-02, 62C99, 62C15, 62C20, 62C10, 62F10, 62F30.

the original space) is non-compact. In this case it often happens that these classical estimators are minimax in both the original problem and the restricted problem. If the restriction is to a convex subset, projection of the classical procedure onto the space will typically produce an improved minimax procedure, but the resulting procedure will usually not be admissible because of violation of technical smoothness requirements. In these cases there is a natural interest in finding minimax generalized Bayes estimators. The original result in this setting is that of Katz (1961) who showed (among other things) for the normal location problem with the mean restricted to be non-negative, that the generalized Bayes estimator with respect to the uniform prior (under quadratic loss) is minimax and admissible and dominates the usual (unrestricted ML or UMVU) estimator. Much of what follows has Katz's result as an examplar. A great deal of the material in sections 2, 3 and 5 is focussed on extending aspects of Katz's result.

If, in the above normal example, the restricted space is a compact interval, then the projection of the usual estimator still dominates the unrestricted MLE but cannot be minimax for quadratic loss because it is not Bayes. In this case Casella and Strawderman (1981) and Zinzius (1981) showed that the unique minimax estimator of the mean θ for a restriction of the form $\theta \in [-m, m]$ is the Bayes estimator corresponding to a 2 point prior on $\{-m, m\}$ for m sufficiently small. The material in section 6 deals with this result, and the large related literature that has followed.

In many problems, as in the previous paragraph, Bayes or Generalized Bayes estimators are known to form a complete class. When loss is quadratic and the prior (and hence typically the posterior) distribution is not degenerate at a point, the Bayes estimator cannot take values on the boundary of the parameter space. There are many results in the literature that use this phenomenon to determine inadmissibility of certain estimators that take values on (or near) the boundary. Moors (1985) developed a useful technique which has been employed by a number of authors in proving inadmissibility and finding improved estimators. We investigate this technique and the related literature in section 4.

An interesting and important issue to which we will not devote much effort is the amount of (relative or absolute) improvement in risk obtained by using procedures which take the restrictions on the parameter space into account. In certain situations the improvement is substantial. For example, if we know in a normal problem that the variance of the sample mean is 1 and that the population mean θ is positive, then risk, at $\theta = 0$, of the (restricted) MLE is 0.5, so there is a 50% savings in risk (at $\theta = 0$). Interestingly, at $\theta = 0$, the risk of the Bayes estimator corresponding to the uniform prior on $[0, \infty)$ is equal to the risk of the MLE so there is no savings in risk at $\theta = 0$. There is, however, noticeable improvement some distance from $\theta = 0$. An interesting open problem is to find admissible minimax estimators in this setting which do not have the same risk at $\theta = 0$ as the unrestricted MLE, and, in particular, to find an admissible minimax estimator dominating the restricted MLE.

We will concern ourselves primarily with methods that have proven to be successful in such problems, and somewhat less so with cataloguing the vast collection of results that have appeared in the literature. In particular, we will concentrate on the following methods.

In Section 2, we describe a recent result of Hartigan (2003). He shows, if $X \sim N_p(\theta, I_p)$, loss is $L(\theta, d) = \|d - \theta\|^2$, and $\theta \in C$, where C is any convex set (with non empty interior), then the Bayes estimator with respect to the uniform prior distribution on C dominates the (unrestricted MRE, UMVU, unrestricted ML) estimator $\delta_0(X) = X$. Hartigan's result adds a great deal to what was already known and provides a clever new technique for demonstrating domination.

In Section 3, we study the Integral Expression of Risk Difference (IERD) method introduced by Kubokawa (1994a). The method is quite general as regards to loss function and underlying distribution. It has proven useful in unrestricted as well as restricted parameter spaces. In particular, one of its first uses was to produce an estimator dominating the James-Stein estimator of a multivariate normal mean under squared error loss.

In Section 4, following a discussion on questions of admissibility concerning estimators that take values on the boundary of a restricted parameter space, we investigate a technique of Moors (1985) which is useful in constructing improvements to such estimators under squared error loss.

Section 5 deals with estimating parameters in the presence of additional information. For example, suppose X_1 and X_2 are multivariate normal variates with unknown means θ_1 and θ_2, and known covariance matrices $\sigma_1^2 I$ and $\sigma_2^2 I$. We wish to estimate θ_1 with squared error loss $\|\delta - \theta_1\|^2$ when we know for example that $\theta_1 - \theta_2 \in A$ for some set A. We illustrate the application of a rotation technique, used perhaps first by Blumenthal and Cohen (1968a), as well as Cohen and Sackrowitz (1970), which, loosely described, permits to subdivide the estimation problem into parts that can be separately handled.

Section 6 deals with minimaxity, and particularly those results related to Casella and Strawderman (1981) and Zinzius (1981) establishing minimaxity of Bayes estimators relative to 2 point priors on the boundary of a sufficiently small one dimensional parameter space of the form $[a, b]$.

2. Hartigan's result

Let $X \sim N_p(\theta, I_p)$, $\theta \in C$ where C is an arbitrary convex set in $\Re^p$ with an open interior. For estimating θ under squared error loss, Hartigan (2003) recently proved the striking result that the (Generalized) Bayes estimator relative to the uniform prior distribution on C dominates the usual (unrestricted) MRE estimator X. It seems quite fitting to begin our review of methods useful in restricted parameter spaces by discussing this, the newest of available techniques. Below, ∇ and ∇^2 will denote respectively the gradient and Laplacian operators.

Theorem 1 (Hartigan, 2003). *For $X \sim N_p(\theta, I_p)$, $\theta \in C$ with C being a convex subset of $\Re^p$ with a non-empty interior, the Bayes estimator $\delta_U(X)$ with respect to a uniform prior on C dominates $\delta_0(X) = X$ under squared error loss $\|d - \theta\|^2$.*

Proof. Writing

$$\delta_U(X) = X + \frac{\nabla_X m(X)}{m(X)} \quad \text{with} \quad m(X) = (2\pi)^{-p/2} \int_C e^{-\frac{1}{2}|X-\nu|^2}\, d\nu,$$

we have following Stein (1981),

$$\begin{aligned}
&R\big(\theta, \delta_U(X)\big) - R\big(\theta, \delta_0(X)\big) \\
&\quad = E_\theta \left[\|X + \frac{\nabla_X m(X)}{m(X)} - \theta\|^2 - \|X - \theta\|^2 \right] \\
&\quad = E_\theta \left[\frac{\|\nabla_X m(X))\|^2}{m^2(X)} + 2(X-\theta)' \frac{\nabla_X m(X)}{m(X)} \right] \\
&\quad = E_\theta \left[\frac{\|\nabla_X m(X)\|^2}{m^2(X)} + \operatorname{div}\left(\frac{\nabla_X m(X)}{m(X)} \right) + \frac{(X-\theta)'\nabla_X m(X)}{m(X)} \right]
\end{aligned}$$

$$= E_\theta \left[\frac{\|\nabla_X m(X)\|^2}{m^2(X)} + \frac{m(X)\nabla_X^2 m(X) - \|\nabla_X m(X)\|^2}{m^2(X)} + \frac{(X-\theta)'\nabla_X m(X)}{m(X)} \right]$$
$$= E_\theta \left[\frac{1}{m(X)} H(X,\theta) \right],$$

where $H(x,\theta) = \nabla_x^2 \, m(x) + (x-\theta)'\nabla_x m(x)$.

It suffices to show tht $H(x,\theta) \leq 0$ for all $x \in \Re^p$ and $\theta \in C$. Now, observe that $\nabla_x(e^{-\frac{1}{2}\|x-\nu\|^2}) = -\nabla_\nu(e^{-\frac{1}{2}\|x-\nu\|^2})$ and $\nabla_x^2(e^{-\frac{1}{2}\|x-\nu\|^2}) = \nabla_\nu^2(e^{-\frac{1}{2}\|x-\nu\|^2})$ so that

$$\begin{aligned}
(2\pi)^{p/2} H(x,\theta) &= \nabla_x^2 \int_C e^{-\frac{1}{2}\|x-\nu\|^2} dv + (x-\theta)'\nabla_x \int_C e^{-\frac{1}{2}\|x-\nu\|^2} dv \\
&= \int_C \nabla_\nu^2 \big(e^{-\frac{1}{2}\|x-\nu\|^2}\big)\, d\nu - (x-\theta)' \int_C \nabla_\nu \big(e^{-\frac{1}{2}\|x-\nu\|^2}\big)\, d\nu \\
&= \int_C \nabla_\nu' \big\{ \big(\nabla_\nu \big(e^{-\frac{1}{2}\|x-\nu\|^2}\big) - (x-\theta) e^{-\frac{1}{2}\|x-\nu\|^2}\big) \big\}\, d\nu \\
&= \int_C \mathrm{div}\,_\nu \big[(\theta-\nu) e^{-\frac{1}{2}\|x-\nu\|^2} \big]\, d\nu. \qquad (1)
\end{aligned}$$

By the Divergence theorem, this last expression gives us

$$(2\pi)^{p/2} H(x,\theta) = \int_{\partial C} \eta(\nu)'(\theta-\nu) e^{-\frac{1}{2}\|x-\nu\|^2}\, d\sigma(\nu), \qquad (2)$$

where $\eta(\nu)$ is the unit outward normal to C at ν on ∂C, and $d\sigma(\nu)$ is the surface area Lebesgue measure on ∂C (for $p = 1$, see Example 1). Finally, since C is convex, the angle between the directions $\eta(\nu)$ and $\theta - \nu$ for a boundary point ν is obtuse, and we thus have $\eta(\nu)'(\theta-\nu) \leq 0$, for $\theta \in C,\ \nu \in \partial C$, yielding the result. □

Remark 1.

(a) If θ belongs to the interior C° of C; (as in part (a) of Example 1); notice that $\eta(v)'(\theta - v) < 0$ a.e. $\sigma(v)$, which implies $H(x,\theta) < 0$, for $\theta \in C^\circ$ and $x \in \Re^p$, and consequently $R(\theta, \delta_U(X)) < R(\theta, \delta_0(X))$ for $\theta \in C^\circ$.

(b) On the other hand, if C is a pointed cone at θ_0; (as in part (b) of Example 1); then $\eta(\nu)'(\theta_0 - \nu) = 0$ for all $\nu \in \partial C$ which implies $R(\theta_0, \delta_U(X)) = R(\theta_0, \delta_0(X))$.

As we describe below, Theorem 1 has previously been established for various specific parameter spaces C. However, Hartigan's result offers not only a unified and elegant proof, but also gives many non-trivial extensions with respect to the parameter space C. We pursue with the instructive illustration of a univariate restricted normal mean.

Example 1.

(a) For $X \sim N(\theta, 1)$ with $\theta \in C = [a,b]$, we have by (1),

$$\begin{aligned}
(2\pi)^{\frac{1}{2}} H(x,\theta) &= \int_a^b \frac{d}{d\nu}(\theta-\nu) e^{-\frac{1}{2}(x-\nu)^2}\, d\nu \\
&= \Big[(\theta-\nu)\, e^{-\frac{1}{2}(x-v)^2} \Big]_a^b \\
&= (\theta-b)\, e^{-\frac{1}{2}(x-b)^2} - (\theta-a)\, e^{-\frac{1}{2}(x-a)^2} \\
&< 0, \qquad \text{for all } \theta \in [a,b].
\end{aligned}$$

This tells us that $R(\theta, \delta_U(X)) < R(\theta, \delta_0(X))$ for all $\theta \in C = [a, b]$.

(b) For $X \sim N(\theta, 1)$ with $\theta \in C = [a, \infty)$ (or $C = (-\infty, a]$), it is easy to see that the development in part (a) remains valid with the exception that $H(x, a) = 0$ for all $x \in \Re$, which tells us that $R(\theta, \delta_U(X)) \leq R(\theta, \delta_0(X))$ for $\theta \in C$ with equality iff $\theta = a$.

The dominance result for the bounded normal mean in Example 1(a) was established by MacGibbon, Gatsonis and Strawderman (1987), in a different fashion, by means of Stein's unbiased estimate of the difference in risks, and sign change arguments following Karlin (1957). The dominance result for the lower bounded normal mean in Example 1(b) was established by Katz (1961), where he also showed that $\delta_U(X)$ is a minimax and admissible estimator of θ.[1] Notice that these results by themselves lead to extensions of the parameter spaces C where $\delta_U(X)$ dominates $\delta_0(X)$, for instance to hyperrectangles of the form $C = \{\theta \in \Re^p : \theta_i \in [a_i, b_i];\ i = 1, \ldots, p\}$, and to intersection of half-spaces since such problems can be expressed as "products" of one-dimensional problems.

Balls and cones in $\Re^p$ are two particularly interesting classes of convex sets for which Hartigan's result gives new and useful information. It is known that for balls of sufficiently small radius, (see e.g., Marchand and Perron, 2001, and Section 4.3 below), the uniform prior leads to dominating procedures (of the mle), but Hartigan's result implies that the uniform Bayes procedures always dominate $\delta_0(X) = X$. Also, for certain types of cones such as intersections of half spaces, Katz's result implies domination over X as previously mentioned. However, Hartigan's result applies to all cones, and, again, increases greatly the catalog of problems where the uniform Bayes procedure dominates X under squared error loss.

Now, Hartigan's result, as described above, although very general with respect to the choice of the convex parameter space C, is nevertheless limited to: (i) normal models, (ii) squared error loss, (iii) the uniform prior as leading to a dominating Bayes estimator; and extensions in these three directions are certainly of interest. Extensions to general univariate location families and general location invariant losses are discussed in Section 3.2. Finally, it is worth pointing out that in the context of Theorem 1, the maximum likelihood estimator $\delta_{\text{mle}}(X)$, which is the projection of X onto C, also dominates $\delta_0(X) = X$. Hence, dominating estimators of $\delta_0(X)$ can be generated by convex linear combinations of $\delta_U(X)$ and $\delta_{\text{mle}}(X)$. Thus the inadmissibility itself is obvious but the technique and the generality are very original and new.

3. Kubokawa's method

Kubokawa (1994a) introduced a powerful method, based on an integral expression of risk difference (IERD), to give a unified treatment of point and interval estimation of the powers of a scale parameter, including the particular case of the estimation of a normal variance. He also applied his method for deriving a class of estimators improving on the James-Stein estimator of a multivariate mean. As reviewed by Kubokawa (1998,1999), many other applications followed such as in: estimation of variance components, estimation of non-centrality parameters, linear calibration, estimation of the ratio of scale parameters, estimation of location and scale parameters under order restrictions, and estimation of restricted location

[1] Although the result is correct, the proof given by Katz has an error (see for instance van Eeden, 1995).

and scale parameters. As well, a particular strength resides in the flexibility of the method in handling various loss functions.

3.1. Example

Here is an illustration of Kubokawa's IERD method for an observable X generated from a location family density $f_\theta(x) = f_0(x-\theta)$, with known f_0, where $E_\theta[X] = \theta$, and $E_\theta[X^2] < \infty$. For estimating θ, with squared error loss $(d-\theta)^2$, under the constraint $\theta \geq a$ (hereafter, we will take $a = 0$ without loss of generality), we show that the Generalized Bayes estimator $\delta_U(X)$ with respect to the uniform prior $\pi(\theta) = 1_{(0,\infty)}(\theta)$ dominates the MRE estimator $\delta_0(X) = X$. As a preliminary to the following dominance result, observe that $\delta_U(X) = X + h_U(X)$, where

$$h_U(y) = \frac{\int_0^\infty (\theta - y) f_0(y-\theta)\, d\theta}{\int_0^\infty f_0(y-\theta)\, d\theta} = \frac{-\int_{-\infty}^{y} u f_0(u)\, du}{\int_{-\infty}^{y} f_0(u)\, du} = -E_0[X|X \leq y];$$

and that h_U is clearly continuous, nonincreasing, with $h_U(y) \geq -\lim_{y\to\infty} E_0[X|X \leq y] = -E_0[X] = 0$.

Theorem 2. *For the restricted parameter space $\theta \in \Theta = [0,\infty)$, and under squared error loss:*

(a) *Estimators $\delta_h(X) = \delta_0(X) + h(X)$ with absolutely continuous, non-negative, nonincreasing h, dominate $\delta_0(X) = X$ whenever $h(x) \leq h_U(x)$ (and $\delta_h \neq \delta_0$);*

(b) *The Generalized Bayes estimator $\delta_U(X)$ dominates the MRE estimator $\delta_0(X)$.*

Proof. First, part (b) follows from part (a) and the above mentioned properties of h_U. Observing that the properties of h and h_U imply $\lim\limits_{y\to\infty} h(y) = 0$, and following Kubokawa (1994a), we have

$$\begin{aligned}
(x-\theta)^2 - (x+h(x)-\theta)^2 &= (x+h(y)-\theta)^2 \,|_{y=x}^{\infty}. \\
&= \int_x^\infty \frac{\partial}{\partial y}(x+h(y)-\theta)^2\, dy \\
&= 2\int_x^\infty h'(y)(x+h(y)-\theta)\, dy,
\end{aligned}$$

so that

$$\begin{aligned}
\Delta_h(\theta) &= E_\theta[(X-\theta)^2 - (X+h(X)-\theta)^2] \\
&= 2\int_{-\infty}^{\infty} \left\{ \int_x^\infty h'(y)(x+h(y)-\theta)\, dy \right\} f_0(x-\theta)\, dx \\
&= 2\int_{-\infty}^{\infty} h'(y) \left\{ \int_{-\infty}^{y} (x+h(y)-\theta) f_0(x-\theta)\, dx \right\} dy.
\end{aligned}$$

Now, since $h'(y) \leq 0$ (h' exists a.e.), it suffices in order to prove that $\Delta_h(\theta) \geq 0$; $\theta \geq 0$; to show that

$$G_h(y,\theta) = \int_{-\infty}^{y} (x+h(y)-\theta) f_0(x-\theta)\, dx \leq 0$$

for all y, and $\theta \geq 0$. But, this is equivalent to

$$\begin{aligned}
&\frac{\int_{-\infty}^{y}(x+h(y)-\theta)f_0(x-\theta)\,dx}{\int_{-\infty}^{y}f_0(x-\theta)\,dx} \leq 0 \\
&\Leftrightarrow \frac{\int_{-\infty}^{y-\theta}(u+h(y))f_0(u)\,du}{\int_{-\infty}^{y-\theta}f_0(u)\,du} \leq 0 \\
&\Leftrightarrow h(y) \leq -E_0[X|X \leq y-\theta]; \quad \text{for all } y, \quad \text{and} \quad \theta \geq 0; \\
&\Leftrightarrow h(y) \leq \inf_{\theta \geq 0}\{-E_0[X|X \leq y-\theta]\}; \quad \text{for all } y; \\
&\Leftrightarrow h(y) \leq -E_0[X|X \leq y] = h_U(y); \quad \text{for all } y;
\end{aligned}$$

given that $E_0[X|X \leq z]$ is nondecreasing in z. This establishes part (a), and completes the proof of the theorem. □

Remark 2. In Theorem 2, it is worth pointing out, and it follows immediately that $G_h(y,\theta) \leq 0$, for all y, with equality iff $h = h_U$ and $\theta = 0$, which indicates that, for the dominating estimators of Theorem 2, $R(\theta,\delta_h(X)) \leq R(\theta,\delta_0(X))$ with equality iff $h = h_U$ and $\theta = 0$. As a consequence, $\delta_U(X)$ fails to dominate any of these other dominating estimators $\delta_h(X)$, and this includes the case of the truncation of $\delta_0(X)$ onto $[0,\infty)$, $\delta^+(X) = \max(0,\delta_0(X))$ (also see Section 4.4 for a discussion on a normal model $\delta^+(X)$).

3.2. Some related results to Theorem 2

For general location family densities $f_0(x-\theta)$, and invariant loss $L(\theta,d) = \rho(d-\theta)$ with strictly convex ρ, Farrell (1964) established: (i) part (b) of Theorem 2, and (ii) the minimaxity of $\delta_U(X)$, and (iii) the admissibility of $\delta_U(X)$ for squared error loss ρ. Using Kubokawa's method, Marchand and Strawderman (2003,a) establish extensions of Theorem 2 (and of Farrell's result (i)) to strictly bowl-shaped losses ρ. They also show, for quite general (f_0,ρ), that the constant risk of the MRE estimator $\delta_0(X)$ matches the minimax risk. This implies that dominating estimators of $\delta_0(X)$, such as those in extensions of Theorem 2, which include $\delta_U(X)$ and $\delta^+(X)$, are necessarily minimax for the restricted parameter space $\Theta = [0,\infty)$. Marchand and Strawderman (2003,a,b) give similar developments for scale families, and for cases where the restriction on θ is to an interval $[a,b]$. Related work for various models and losses includes Jozani, Nematollahi and Shafie (2002), van Eeden (2000, 1995), Parsian and Sanjari (1997), Gajek and Kaluszka (1995), Berry (1993), and Gupta and Rohatgi (1980), and many of the references contained in these papers.

Finally, as previously mentioned, Kubokawa's method has been applied to a wide range of problems, but, in particular for problems with ordered scale or location parameters (also see Remark 4), results and proofs similar to Theorem 2 have been established by Kubokawa and Saleh (1994), Kubokawa (1994b), and Ilioupoulos (2000).

4. Estimators that take values on the boundary of the parameter space

Theoretical difficulties that arise in situations when estimating procedures take values on, or close to the boundary of constrained parameter spaces are well documented. For instance, Sacks (1963), and Brown (1971), show for estimating under squared error loss a lower bounded normal mean θ with known variance, that the maximum likelihood estimator is an inadmissible estimator of θ. More recently, difficulties such as those encountered with interval estimates have been addressed in Mandelkern (2002). In this section, we briefly expand on questions of admissibility and on searches for improved procedures, but we mostly focus on a method put forth by Moors (1985) which is useful in providing explicit improvements of estimators that take values on, or close to the boundary of a restricted parameter space.

4.1. Questions of admissibility

Here is a simple example which illustrates why, in many cases, estimators that take values on the boundary of the parameter space are inadmissible under squared error loss. Take $X \sim N_p(\theta, I_p)$ where θ is restricted to a ball $\Theta(m) = \{\theta \in \Re^p : \|\theta\| \leq m\}$. Complete class results indicate that admissible estimators are necessarily Bayes for some prior π (supported on $\Theta(m)$, or a subset of $\Theta(m)$). Now observe that prior and posterior pairs $(\pi, \pi|x)$ must be supported on the same set, and that a Bayes estimator takes values $\delta_\pi(x) = E(\theta|x)$ on the interior of the convex $\Theta(m)$, as long as $\pi|x$, and hence π, is not degenerate at some boundary point θ_0 of $\Theta(m)$. The conclusion is that non-degenerate estimators $\delta(X)$ which take values on the boundary of $\Theta(m)$ (i.e., $\mu\{x : \delta(x) \in \partial(\Theta(m)\} > 0$, with μ as Lebesgue measure); which includes for instance the MLE; are inadmissible under squared error loss. In a series of papers, Charras and van Eeden (1991a, 1991b, 1992, 1994) formalize the above argument for more general models, and also provide various results concerning the admissibility and Bayesianity under squared error loss of boundary estimators in convex parameter spaces. Useful sources of general complete class results, that apply for bounded parameter spaces, are the books of Berger (1985), and Brown (1986).

Remark 3. As an example where the prior and posterior do not always have the same support, and where the above argument does not apply, take $X \sim Bi(n, \theta)$ with $\theta \in [0, m]$. Moreover, consider the MLE which takes values on the boundary of $[0, m]$. It is well known that the MLE is admissible (under squared error loss) for $m = 1$. Interestingly, again for squared error loss, Charras and van Eeden (1991a) establish the admissibility of the MLE for cases where $m \leq 2/n$, while Funo (1991) establishes its inadmissibility for cases where $m < 1$ and $m > 2/n$. Interestingly and in contrast to squared error loss, Bayes estimators under absolute-value loss may well take values on the boundary of the parameter space. For instance, Iwasa and Moritani (1997) show, for a normal mean bounded an interval $[a, b]$ (known standard deviation), that the MLE is a proper Bayes (and admissible) estimator under absolute-value loss.

The method of Moors, described in detail in Moors (1985), and further illustrated by Moors (1981) and Moors and van Houwelingen (1987), permits the construction of improved estimators under squared error loss of invariant estimators that take values on, or too close to the boundary of closed and convex parameter spaces. We next pursue with an illustration of this method.

4.2. The method of Moors

Illustrating Moors' method, suppose an observable X is generated from a location family density $f(x - \theta)$ with known positive and symmetric f. For estimating $\theta \in \Theta = [-m, m]$ with squared error loss, consider invariant estimators (with respect to sign changes) which are of the form

$$\delta_g(X) = g(|X|)\frac{X}{|X|}.$$

The objective is to specify dominating estimators of $\delta_g(X)$, for cases where $\delta_g(X)$ takes values on or near the boundary $\{-m, m\}$ (i.e., $|m - g(x)|$ is "small" for some x).

Decompose the risk of $\delta_g(X)$ by conditioning on $|X|$ (i.e., the maximal invariant) to obtain (below, the notation $E_\theta^{|X|}$ represents the expectation with respect to $|X|$)

$$\begin{aligned} R(\theta, \delta_g(X)) &= E_\theta^{|X|} \left\{ E_\theta \left[\left(g(|X|)\, \frac{X}{|X|} - \theta \right)^2 \middle| \, |X| \right] \right\} \\ &= E_\theta^{|X|} \left\{ \theta^2 + g^2(|X|) - 2E_\theta \left[\frac{\theta X}{|X|} g(|X|) \middle| \, |X| \right] \right\} \\ &= E_\theta^{|X|} \left\{ \theta^2 + g^2(|X|) - 2g(|X|) A_{|X|}(\theta) \right\}, \end{aligned}$$

where

$$A_{|X|}(\theta) = \theta E_\theta \left[\frac{X}{|X|} \middle| \, |X| \right] = \theta \left\{ \frac{f(|X| - \theta) - f(|X| + \theta)}{f(|X| - \theta) + f(|X| + \theta)} \right\},$$

(as in (6) below) by symmetry of f. Now, rewrite the risk as

$$R(\theta, \delta_g(X)) = E_\theta^{|X|} \left[\theta^2 - A^2_{|X|}(\theta) \right] + E_\theta^{|X|} \left[\left(g(|X|) - A_{|X|}(\theta) \right)^2 \right], \qquad (3)$$

to isolate with the second term the role of g, and to reflect the fact that the performance of the estimator $\delta_g(X)$, for $\theta \in [-m, m]$, is measured by the average distance $\left(g(|X|) - A_{|X|}(\theta) \right)^2$ under $f(x - \theta)$. Continue by defining the $A_{|x|}$ as the convex hull of the set $\{A_{|x|}(\theta) : -m \le \theta \le m\}$. Coupled with the prior representation (3) of $R(\theta, \delta_g(X))$, we can now state the following result.

Theorem 3. *Suppose $\delta_g(X)$ is an estimator such that $\mu\{x : g(|x|) \notin A_{|x|}\} > 0$, then the estimator $\delta_{g_0}(X)$ with $g_0(|x|)$ being the projection of $g(|x|)$ onto $A_{|x|}$ dominates $\delta_g(X)$, with squared error loss under f, for $\theta \in \Theta = [-m, m]$.*

Example 2. (Normal Case) Consider a normal model f with variance 1. We obtain $A_{|x|}(\theta) = \theta \tanh(\theta|x|)$, and $A_{|x|} = [0, m\tanh(m|x|)]$, since $A_{|x|}(\theta)$ is increasing in $|\theta|$. Consider an estimator $\delta_g(X)$ such that $\mu\{x : g(|x|) > m\tanh(m|x|)\} > 0$. Theorem 3 tells us that $\delta_{g_0}(X)$, with $g_0(|X|) = \min(m\tanh(m|X|), g(|X|))$, dominates $\delta_g(X)$.

Here are some additional observations related to the previous example (but also applicable to the general case of this section).

(i) The set $A_{|x|} = [0, m\tanh(m|x|)]$ can be interpreted as yielding a complete class of invariant estimators with the upper envelope corresponding to the Bayes estimator $\delta_{BU}(X)$ associated with the uniform prior on $\{-m, m\}$.

(ii) In Example 2, the dominating estimator $\delta_{g_0}(X)$ of Theorem 3 will be given by the Bayes estimator $\delta_{BU}(X)$ if and only if $m\tanh(m|x|) \le g(|x|)$, for all x. In particular, if $\delta_g(X) = \delta_{\text{mle}}(X)$, with $g(|X|) = \min(|X|, m)$, it is easy to verify that $\delta_{g_0}(X) = \delta_{BU}(X)$ iff $m \le 1$.

(iii) It is easy to see that improved estimators $\delta_{g'}(X)$ of $\delta_g(X)$ can be constructed by projecting $g(|x|)$ a little bit further onto the interior of $A_{|x|}$, namely by selecting g' such that $\frac{1}{2}\left[g'(|x|) + g(|x|) \right] \le g_0(|x|)$ whenever $g(|x|) > g_0(|x|)$.

4.3. Some related work

Interestingly, the dominance result in (iii) of the normal model MLE was previously established, in a diferent manner, by Casella and Strawderman (1981) (see also Section 6). As well, other dominating estimators here were provided numerically by Kempthorne (1988).

For the multivariate version of Example 2: $X \sim N_p(\theta, I_p)$; $(p \geq 1)$; with $\|\theta\| \leq m$, Marchand and Perron (2001) give dominating estimators of $\delta_{\text{mle}}(X)$ under squared error loss $\|d-\theta\|^2$. Namely, using a similar risk decomposition as above, including argument (ii), they show that $\delta_{BU}(X)$ (Bayes with respect to a boundary uniform prior) dominates $\delta_{\text{mle}}(X)$ whenever $m \leq \sqrt{p}$. By pursuing with additional risk decompositions, they obtain various other dominance results. In particular, it is shown that, for sufficiently small radius m, $\delta_{\text{mle}}(X)$ is dominated by all Bayesian estimators associated with orthogonally invariant priors (which includes the uniform Bayes estimator δ_U). Finally, Marchand and Perron (2003) give extensions and robustness results involving δ_{BU} to spherically symmetric models, and Perron (2003) gives a similar treatment for the model $X \sim Bi(n, \theta)$ with $|\theta - \frac{1}{2}| \leq m$.

4.4. Additional topics and the case of a positive normal mean

Other methods have proven useful in assessing the performance of boundary estimators in constrained parameter spaces, as well as providing improvements. As an example, for the model $X_i \sim \text{Bin}(n_i, \theta_i)$; $i = 1, \ldots, k$; with $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_k$, Sackrowitz and Strawderman (1974) investigated the admissibility (for various loss functions) of the MLE of $(\theta_1, \ldots, \theta_k)$, while Sackrowitz (1982) provided improvements (under sum of squared error losses) to the MLE in the cases above where it is inadmissible. Further examples consist of a series of papers by Shao and Strawderman (1994,1996a,1996b) where, in various models, improvements under squared error loss to truncated estimators are obtained. Further related historical developments are given in the review paper of van Eeden (1996).

We conclude this section by expanding upon the case of a positive (or lower-bounded) normal mean θ, for $X \sim N(\theta, 1), \theta \geq 0$. While a plausible and natural estimator is given by the MLE $max(0, X)$, its efficiency requires examination perhaps because it discards part of the sufficient statistic X (i.e., the MLE gives a constant estimate on the region $X \leq 0$). Moreover, as previously mentioned, the MLE has long been known to be inadmissible (e.g., Sacks, 1963) under squared error loss. Despite the age of this finding, it was not until the paper of Shao and Strawderman (1996a) that explicit improvements were obtained (under squared error loss), and there still remains the open question of finding admissible improvements. As well, Katz's (1961) uniform Generalized Bayes estimator remains (to our knowledge) the only known minimax and admissible estimator of θ (under squared error loss).

5. Estimating parameters with additional information

In this section, we present a class of interesting estimation problems which can be transformed to capitalize on standard solutions for estimation problems in constrained parameter spaces. The key technical aspect of subdividing the estimation problem into distinct pieces that can be handled separately is perhaps due to the early work of Blumenthal, Cohen and Sackrowitz. As well, these types of problems have been addressed in some recent work of Constance van Eeden and Jim Zidek.

Suppose X_j; $j = 1, 2$; are independently distributed as $N_p(\theta_j, \sigma_j^2 I_p)$, with $p \geq 1$ and known σ_1^2, σ_2^2. Consider estimating θ_1 under squared error loss $L(\theta_1, d) =$

$\|d-\theta_1\|^2$ with the prior information $\theta_1-\theta_2 \in A$; A being a proper subset of $\Re^p$. For instance, with order restrictions of the form $\theta_{1,i} \geq \theta_{2,i}, i=1,\dots p$, we would have $A=(\Re^+)^p$. Heuristics suggest that the independently distributed X_2 can be used in conjunction with the information $\theta_1-\theta_2 \in A$ to construct estimators that improve upon the unrestricted MLE (and UMVU estimator) $\delta_0(X_1,X_2)=X_1$. For instance, suppose $\frac{\sigma_2^2}{\sigma_1^2} \approx 0$, and that A is convex. Then, arguably, estimators of θ_1 should shrink towards $A+x_2=\{\theta_1 : \theta_1-x_2 \in A\}$. The recognition of this possibility (for $p=1$ and $A=(0,\infty)$) goes back at least as far as Blumenthal and Cohen (1968a), or Cohen and Sackrowitz (1970); and is further discussed in some detail by van Eeden and Zidek (2003).

Following the rotation technique used by Blumenthal and Cohen (1968a), Cohen and Sackrowitz (1970), van Eeden and Zidek (2001,2003) among others, we illustrate in this section how one can exploit the information $\theta_1-\theta_2 \in A$; for instance to improve on the unrestricted MLE $\delta_0(X_1,X_2)=X_1$. It will be convenient to define $\mathcal{C}_1$ as the following subclass of estimators of θ_1:

Definition 1.

$$\mathcal{C}_1=\Big\{\delta_\phi : \delta_\phi(X_1,X_2)=Y_2+\phi(Y_1),\quad \text{with} \quad Y_1=\frac{X_1-X_2}{1+\tau},\ Y_2=\frac{\tau X_1+X_2}{1+\tau}, \quad \text{and } \tau=\sigma_2^2/\sigma_1^2\Big\}.$$

Note that the above defined Y_1 and Y_2 are independently normally distributed (with $E[Y_1]=\mu_1=\frac{\theta_1-\theta_2}{1+\tau}$, $E[Y_2]=\mu_2=\frac{\tau\theta_1+\theta_2}{1+\tau}$, $Cov(Y_1)=\frac{\sigma_1^2}{1+\tau}I_p$, and $Cov(Y_2)=\frac{\tau\sigma_1^2}{1+\tau}I_p$). Given this independence, the risk function of δ_ϕ (for $\theta=(\theta_1,\theta_2)$) becomes

$$\begin{aligned} R\big(\theta,\delta_\phi(X_1,X_2)\big) &= E_\theta\big[\|Y_2+\phi(Y_1)-\theta_1\|^2\big] \\ &= E_\theta\left[\left\|\left(Y_2-\frac{\tau\theta_1+\theta_2}{1+\tau}\right)+\left(\phi(Y_1)-\frac{\theta_1-\theta_2}{1+\tau}\right)\right\|^2\right] \\ &= E_\theta\big[\|Y_2-\mu_2\|^2\big]+E_\theta\big[\|\phi(Y_1)-\mu_1\|^2\big]. \end{aligned}$$

Therefore, the performance of $\delta_\phi(X_1,X_2)$ as an estimator of θ_1 is measured solely by the performance of $\phi(Y_1)$ as an estimator of μ_1 under the model $Y_1 \sim N_p(\mu_1,\frac{\sigma_1^2}{1+\tau}I_p)$, with the restriction $\mu_1 \in C=\{y:(1+\tau)y \in A\}$. In particular, one gets the following dominance result.

Proposition 1. *For estimating θ_1 under squared error loss, with $\theta_1-\theta_2 \in A$, the estimator $\delta_{\phi_1}(X_1,X_2)$ will dominate $\delta_{\phi_0}(X_1,X_2)$ if and only if*

$$E_{\mu_1}\big[\|\phi_1(Y_1)-\mu_1\|^2\big] \leq E_{\mu_1}\big[\|\phi_0(Y_1)-\mu_1\|^2\big],$$

for $\mu_1 \in C$ (with strict inequality for some μ_1).

We pursue with some applications of Proposition 1, which we accompany with various comments and historical references.

(A) *Case where $\delta_{\phi_0}(X_1,X_2)=X_1$ (i.e., the unrestricted mle of θ_1), and where A is convex with a non empty interior.*

This estimator arises as a member of $\mathcal{C}_1$ for $\phi_0(Y_1)=Y_1$. Hartigan's result (Theorem 1) applies to $\phi_0(Y_1)$ (since A convex implies C convex), and tells us that the Bayes estimator $\phi_{U_C}(Y_1)$ of μ_1 with respect to a uniform prior on C dominates $\phi_0(Y_1)$ (under squared-error loss). Hence, Proposition 1 applies with $\phi_1 = \phi_{U_C}$, producing the following dominating estimator of $\delta_0(X_1, X_2)$:

$$\delta_{\phi_{U_C}}(X_1, X_2) = Y_2 + \phi_{U_C}(Y_1). \tag{4}$$

For $p = 1$ and $A = [-m, m]$, the dominance of δ_{ϕ_0} by the estimator given in (4) was established by van Eeden and Zidek (2001), while for $p = 1$ and $A = [0, \infty)$ (or $A = (-\infty, 0]$), this dominance result was established by Kubokawa and Saleh (1994). In both cases, Kubokawa's IERD method, as presented in Section 3, was utilized to produce a class of dominating estimators which includes $\delta_{\phi_{U_C}}(X_1, X_2)$. As was the case in Section 2, these previously known dominance results yield extensions to sets A which are hyperrectangles or intersection of half-spaces, but Hartigan's result yields a much more general result.

Remark 4. Here are some additional notes on previous results related to the case $p = 1$ and $A = [0, \infty)$. Kubokawa and Saleh (1994) also provide various extensions to other distributions with monotone likelihood ratio and to strict bowl-shaped losses, while van Eeden and Zidek (2003) introduce an estimator obtained from a weighted likelihood perspective and discuss its performance in comparison to several others including $\delta_{\phi_{U_C}}(X_1, X_2)$. The admissibility and minimaxity of $\delta_{\phi_{U_C}}(X_1, X_2)$ (under squared error loss) were established by Cohen and Sackrowitz (1970). Further research concerning this problem, and the related problem of estimating jointly θ_1 and θ_2, has appeared in Blumenthal and Cohen (1968b), Brewster and Zidek (1974), and Kumar and Sharma (1988) among many others. There is equally a substantial body of work concerning estimating a parameter θ_1 (e.g., location, scale, discrete family) under various kinds of order restrictions involving k parameters $\theta_1, \ldots, \theta_k$ (other than work referred to elsewhere in this paper, see for instance van Eeden and Zidek, 2001, 2003 for additional annotated references).

Another dominating estimator of $\delta_{\phi_0}(X_1, X_2) = X_1$, which may be seen as a consequence of Proposition 1, is given by $\delta_{\phi_1}(X_1, X_2) = Y_2 + \phi_{\text{mle}}(Y_1)$, where $\phi_{\text{mle}}(Y_1)$ is the mle of μ_1, $\mu_1 \in C$. This is so because, as remarked upon in Section 2, $\phi_1(Y_1) = \phi_{\text{mle}}(Y_1)$ dominates under squared error loss, as an estimator of $\mu_1; \mu_1 \in C$; $\phi_0(Y_1) = Y_1$. Observe further that the maximum likelihood estimator $\delta_{\text{mle}}(X_1, X_2)$ of θ_1 for the parameter space $\Theta = \{(\theta_1, \theta_2) : \theta_1 - \theta_2 \in A, \tau\theta_1 + \theta_2 \in \Re^p\}$ is indeed given by: $\delta_{\text{mle}}(X_1, X_2) = (\hat{\mu_2})_{\text{mle}} + (\hat{\mu_1})_{\text{mle}} = Y_2 + \phi_{\text{mle}}(Y_1)$, given the independence and normality of Y_1 and Y_2, and the fact that Y_2 is the MLE of μ_2 ($\mu_2 \in \Re^p$).

Our next two applications of Proposition 1 deal with the estimator $\delta_{\text{mle}}(X_1, X_2)$.

(B) *Case where A is a ball and $\delta_{\phi_0}(X_1, X_2) = \delta_{mle}(X_1, X_2)$.*

For the case where $\theta_1 - \theta_2 \in A$, with A being a p-dimensional ball of radius m centered at 0, the estimator $\delta_{\text{mle}}(X_1, X_2)$ arises as a member of $\mathcal{C}_1$ for $\phi_0(Y_1) = \phi_{\text{mle}}(Y_1) = (\|Y_1\| \wedge \frac{m}{1+\tau})\frac{Y_1}{\|Y_1\|}$. By virtue of Proposition 1, it follows that dominating estimators $\phi_*(Y_1)$ of $\phi_{\text{mle}}(Y_1)$ (for the ball with $\|\mu_1\| \leq m/(1+\tau)$), such as those given by Marchand and Perron (2001) (see Section 4.3 above), yield dominating estimators $\delta_{\phi_*}(X_1, X_2) = \frac{\tau X_1 + X_2}{1+\tau} + \phi_*(\frac{X_1 - X_2}{1+\tau})$ of $\delta_{\text{mle}}(X_1, X_2)$.

(C) *Case where* $A = [0, \infty)$, *and* $\delta_{\phi_0}(X_1, X_2) = \delta_{mle}(X_1, X_2)$. This is similar to (B), and dominating estimators can be constructed by using Shao and Strawderman's (1996) dominating estimators of the MLE of a positive normal mean (see van Eeden and Zidek, 2001, Theorem 4.3)

Observe that results in (B) (for $p = 1$) and (C) lead to further applications of Proposition 1 for sets A which are hyperrectangles or intersection of half-spaces. We conclude by pointing out that the approach used in this section may well lead to new directions in future research. For instance, the methods used above could be used to specify dominating estimators for the case $p \geq 3$, (of $\delta_{\phi_0}(X_1, X_2) = Y_2 + \phi_0(Y_1)$) of the form $\phi_2(Y_2) + \phi_1(Y_1)$ where, not only is $\phi_1(Y_1)$ a dominating estimator of $\phi_0(Y_1)$ for $\mu_1 \in C$, but for $p \geq 3$, $\phi_2(Y_2)$ is a Stein-type estimator of μ_2 which dominates Y_2.

6. Minimax estimation

This section presents an overview of minimax estimation in compact parameter spaces, with a focus on the case of an interval constraint of the type $\theta \in [a, b]$ and analytical results giving conditions for which the minimax estimator is a Bayesian estimator with respect to a boundary prior on $\{a, b\}$. Historical elements are first described in Section 6.1, a somewhat novel expository example is presented in Section 6.2., and we further describe complementary results in Section 6.3.

6.1. Two point least favourable priors

With the criteria of minimaxity playing a vital role in the development of statistical theory and practice; as reviewed in Brown (1994) or Strawderman (2000) for instance; the results of Casella and Strawderman (1981), as well as those of Zinzius (1981) most certainly inspired a lot of further work. These results presented analytically obtained minimax estimators, under squared error loss, of a normal model mean θ, with known variance, when θ is known to be restricted to a small enough interval. More precisely, Casella and Strawderman showed, for $X \sim N(\theta, 1)$ with $\theta \in \Theta = [-m, m]$; (there is no loss of generality in assuming the variance to be 1, and the interval to be symmetric about 0); that the uniform boundary Bayes estimator $\delta_{BU}(x) = m \tanh(mx)$ is unique minimax iff $m \leq m_0 \approx 1.0567$. Furthermore, they also investigated three-point priors supported on $\{-m, 0, m\}$, and obtained sufficient conditions for such a prior to be least favourable. It is worth mentioning that these results give immediately minimax multivariate extensions to rectangular constraints where $X_i \sim N(\theta_i, 1); i = 1, \ldots, p$; with $|\theta_i| \leq m_i \leq m_0$, under losses $\sum_{i=1}^{p} \omega_i (d_i - \theta_i)^2$, (with arbitrary positive weights ω_i), since the least favourable prior for estimating $(\theta_1, \ldots, \theta_p)$ is obtained, in such a case, as the product of the least favourable priors for estimating $\theta_1, \ldots, \theta_p$ individually. Now, the above minimaxity results were obtained by using the following well-known criteria for minimaxity (e.g., Berger, 1985, Section 5.3, or Lehmann and Casella, 1998, Section 5.1).

Lemma 1. *If* δ_π *is a Bayes estimator with respect to a prior distribution* π, *and* $S_\pi = \{\theta \in \Theta : \sup_\theta \{R(\theta, \delta_\pi)\} = R(\theta, \delta_\pi)\}$, *then* δ_π *is minimax whenever* $P_\pi(\theta \in S_\pi) = 1$.

Casella and Strawderman's work capitalized on Karlin's (1957) sign change arguments for implementing Lemma 1 while, in contrast, the sufficient conditions obtained by Zinzius concerning the minimaxity of $\delta_{BU}(X)$ were established using the "convexity technique" as stated as part (b) of the following Corollary to Lemma 1. Part (a), introduced here as an alternative condition, will be used later in this section.

Corollary 1. *If δ_π is a Bayes estimator with respect to a two-point prior on $\{a, b\}$ such that $R(a, \delta_\pi) = R(b, \delta_\pi)$, then δ_π is minimax for the parameter space $\Theta = [a, b]$ whenever, as a function of $\theta \in [a, b]$,*

(a) $\frac{\partial}{\partial\theta} R(\theta, \delta_\pi)$ *has at most one sign change from $-$ to $+$; or*

(b) $R(\theta, \delta_\pi)$ *is convex.*

Although the convexity technique applied to the bounded normal mean problem gives only a lower bound for m_0; (Bader and Bischoff (2003) report that the best known bound using convexity is $\frac{\sqrt{2}}{2}$, as given by Bischoff and Fieger (1992)); it has proven very useful for investigating least favourable boundary supported priors for other models and loss functions. In particular, DasGupta (1985) used subharmonicity to establish, for small enough compact and convex parameter spaces under squared error loss, the inevitability of a boundary supported least favourable prior for a general class of univariate and multivariate models. As well, the work of Bader and Bischoff (2003), Boratyńska (2001), van Eeden and Zidek (1999), and Eichenauer-Hermann and Fieger (1992), among others, establish this same inevitability with some generality with respect to the loss and/or the model. Curiously, as shown by Eichenauer-Hermann and Ickstadt (1992), and Bischoff and Fieger (1993), there need not exist a boundary least favourable prior for convex, but not strictly convex, losses. Indeed, their results both include the important case of a normal mean restricted to an interval and estimated with absolute-value loss, where no two-point least favourable prior exists.

6.2. Two-point least favourable priors in symmetric location families

We present here a new development for location family densities (with respect to Lebesgue measure on $\Re^1$) of the form

$$f_\theta(x) = e^{-h(x-\theta)}, \quad \text{with convex and symmetric } h. \tag{5}$$

These assumptions on h imply that such densities f_θ are unimodal, symmetric about θ, and possess monotone increasing likelihood ratio in X. For estimating θ with squared error loss under the restriction $\theta \in [-m, m]$, our objective here is to present a simple illustration of the inevitability of a boundary supported least favourable prior for small enough m, i.e., $m \leq m_0(h)$. Namely, we give for densities in (5) with concave $h'(x)$ for $x \geq 0$ (this implies convex $h'(x)$ for $x \leq 0$) a simple lower bound for $m_0(h)$. We pursue with two preliminary lemmas; the latter one giving simple and general conditions for which a wide subclass of symmetric estimators (i.e., equivariant under sign changes) $\delta(X)$ of θ have increasing risk $R(\theta, \delta(X))$ in $|\theta|$ under squared error loss.

Lemma 2. *If g is a bounded and almost everywhere differentiable function, then under* (5)*:*

$$\frac{d}{d\theta} E_\theta\,[\,g(X)] = E_\theta\,[\,g'(X)].$$

Proof. First, interchange derivative and integral to obtain $\frac{d}{d\theta} E_\theta[\,g(X)\,] = E_\theta\,[g(X)\,h'(X-\theta)]$. Then, integrating by parts yields the result. □

Lemma 3. *For models in* (5)*, and estimators $\delta(X)$ with the properties:* (a) $\delta(x) = -\delta(-x)$*;* (b) $\delta'(x) \geq 0$*; and* (c) $\delta'(x)$ *decreasing in* $|x|$*; for all* $x \in \Re$*; either one of the following conditions is sufficient for* $R(\theta, \delta(X))$ *to be increasing in* $|\theta|$*;* $|\theta| \leq m$*:*

(i) $E_\theta[\delta(X)] \leq \theta(1 - E_\theta[\delta'(X)])$*, for* $0 \leq \theta \leq m$*;*

(ii) $E_\theta[\delta(X)] \leq \theta(1 - \delta'(0))$*, for* $0 \leq \theta \leq m$*;*

(iii) $\delta'(0) \leq \frac{1}{2}$.

Proof. It will suffice to work with the condition $\frac{\partial}{\partial\theta} R(\theta, \delta(X)) \geq 0$, $0 \leq \theta \leq m$, since $R(\theta, \delta(X))$ is an even function of θ, given property (a) and the symmetry of h. Differentiating directly the risk and using Lemma 2, we obtain

$$\frac{1}{2}\frac{\partial}{\partial\theta} R\big(\theta, \delta(X)\big) = \theta - E_\theta[\delta(X)] - \theta E_\theta\big[\delta'(X)\big] + E_\theta\big[\delta(X)\delta'(X)\big].$$

With properties (a) and (b), the function $\delta(x)\delta'(x)$ changes signs once (at $x = 0$) from $-$ to $+$, and, thus, sign change properties under h imply that $E_\theta[\delta(X)\delta'(X)\,]$ changes signs at most once from $-$ to $+$ as a function of θ. Since $E_0[\delta(X)\delta'(X)] = 0$ by symmetry of $\delta(x)\delta'(x)$ and h, we must have $E_\theta[\delta(X)\delta'(X)] \geq 0$ for $\theta \geq 0$. It then follows that

$$\frac{1}{2}\frac{\partial}{\partial\theta} R\big(\theta, \delta(X)\big) \geq \theta - E_\theta\big[\delta(X)\big] - \theta E_\theta\big[\delta'(X)\big];$$

and this yields directly sufficient condition (i). Now, property (c) tells us that $\delta'(x) \leq \delta'(0)$, and this indicates that condition (ii) implies (i), hence its sufficiency. Finally, condition (iii) along with Lemma 2 and the properties of $\delta(X)$ implies (ii) since $\frac{\partial}{\partial\theta} E_\theta[\delta(X) + \theta(\delta'(0) - 1)] = E_\theta[\delta'(X) + (\delta'(0) - 1)] \leq E_\theta[2\delta'(0) - 1] \leq 0$, and $E_\theta[\delta(X) + \theta(\delta'(0) - 1)]|_{\theta=0} = 0$.

We pursue by applying Lemma 3 to the case of the boundary uniform Bayes estimator $\delta_{BU}(X)$ to obtain, by virtue of Corollary 1, part (a), a minimaxity result for $\delta_{BU}(X)$. □

Corollary 2. *For models in* (5)*, $\delta_{BU}(X)$ is a unique minimax estimator of θ* (*under squared error loss*) *for the parameter space* $[-m, m]$ *when either one of the following situations arises:*

(A) *Condition* (i) *of Lemma* 3 *holds;*

(B) $h'(x)$ *is concave for* $x \geq 0$*, and condition* (ii) *of Lemma* 3 *holds;*

(C) $h'(x)$ *is concave for* $x \geq 0$*, and* $m \leq m^*(h)$ *where* $m^*(h)$ *is the solution in* m *of the equation* $mh'(m) = \frac{1}{2}$.

Proof. We apply Corollary 1, part (a), and Lemma 3. To do so, we need to investigate properties (b) and (c) of Lemma 3 for the estimator $\delta_{BU}(X)$ (property (a) is necessarily satisfied since the uniform boundary prior is symmetric). Under model (5), the Bayes estimator $\delta_{BU}(X)$ and its derivative (with respect to X) may be expressed as:

$$\delta_{BU}(x) = \frac{me^{-h(x-m)} - me^{-h(x+m)}}{e^{-h(x-m)} + e^{-h(x+m)}} = m\tanh\left(\frac{h(x+m) - h(x-m)}{2}\right); \tag{6}$$

and

$$\delta'_{BU}(x) = \{m^2 - \delta_{BU}(x)^2\}\left\{\frac{h'(x+m) - h'(x-m)}{2m}\right\}. \tag{7}$$

Observe that $|\delta_{BU}(x)| \le m$, and $h'(x+m) \ge h'(x-m)$ by the convexity of h, so that $\delta'_{BU}(x) \ge 0$ given (7). This establishes property (b) of Lemma 3, and part (A). Now, $m^2 - \delta^2_{BU}(x)$ is decreasing in $|x|$, and so is $h'(x+m) - h'(x-m)$ given, namely, the concavity of $h'(x)$ for $x \ge 0$. This tells us that $\delta_{BU}(X)$ verifies property (c) of Lemma 3, and (B) follows. Hence, condition (iii) of Lemma 3 applies becoming equivalent to $mh'(m) \le \frac{1}{2}$, as $\delta'(0) = mh'(m)$ by (7). Finally, the result follows by the fact that $mh'(m)$ is a continuous and increasing of function of m, $m > 0$. □

Remark 5. As the outcome of the above argument, combining both sign change arguments and convexity considerations, containing other elements which may be independent interest, part (C) of Corollary 2 gives a simple sufficient condition for the minimaxity of δ_{BU}, and is applicable to a wide class of models in (5). Namely, for Exponential Power families where, in (5), $h(y) = \alpha|y|^\beta$ with $\alpha > 0$ and $1 \le \beta \le 2$, part (C) of Corollary 2 applies, and tells us that $\delta_{BU}(X)$ (which may be derived from (6)) is unique minimax whenever $m \le m^*(h) = (\frac{1}{2\alpha\beta})^{1/\beta}$. In particular for double-exponential cases, (i.e., $\beta = 1$), we obtain $m^*(h) = \frac{1}{2\alpha}$; and for the standard normal case, (i.e., $(\alpha, \beta) = (\frac{1}{2}, 2)$, we obtain $m^*(h) = \frac{\sqrt{2}}{2}$. Observe that the normal case $m^*(h)$ matches the one given by Bischoff and Fieger (1992); and that, as expected with the various lower bounds used for the derivative of the risk, it falls somewhat below Casella and Strawderman's necessary and sufficient cutoff point of $m_0 \approx 1.05674$.

6.3. Some additional results and comments

The problem considered in Section 6.2, was studied also by Eichenauer-Hermann and Ickstadt (1992), who obtained similar results using a convexity argument for the models in (5) with $L^p, p > 1$ loss. Additional work concerning least favourable boundary priors for various models can be found in: Moors (1985), Berry (1989), Eichenauer (1986), Chen and Eichenauer (1988), Eichenauer-Hermann and Fieger (1989), Bischoff (1992), Bischoff and Fieger (1992), Berry (1993), Johnstone and MacGibbon (1992), Bischoff, Fieger and Wurlfert (1995), Bischoff, Fieger, and Ochtrop (1995), Marchand and MacGibbon (2000), and Wan, Zou and Lee (2000).

Facilitated by results guaranteeing the existence of a least favourable prior supported on a finite number of points (e.g., Ghosh, 1964), the dual problem of searching numerically for a least favourable prior π, as presented in Lemma 1, is very much the standard approach for minimax estimation problems in compact parameter spaces. Algorithms to capitalize on this have been presented by Nelson (1965), and Kempthorne (1987), and have been implemented by Marchand and MacGibbon (2000), for a restricted binomial probability parameter, MacGibbon, Gourdin, Jaumard, and Kempthorne (2000) for a restricted Poisson parameter, among others. Other algorithms have been investigated by Gourdin, Jaumard, and MacGibbon (1994).

Analytical and numerical results concerning the related criteria of Gamma-Minimaxity in constrained parameter spaces have been addressed by Vidakovic and DasGupta (1996), Vidavovic (1993), Lehn and Rummel (1987), Eichenauer and Lehn (1989), Bischoff (1992), Bischoff and Fieger (1992), Bischoff, Fieger and Wurlfert (1995), and Wan et al. (2000).

For spherical bounds of the form $\|\theta\| \leq m$, Berry (1990) generalized Casella and Strawderman's minimaxity of δ_{BU} result for multivariate normal models $X \sim N_p(\theta, I_p)$. He showed with sign change arguments that δ_{BU} is unique minimax for $m \leq m_0(p)$, giving defining equations for $m_0(p)$. Recently, Marchand and Perron (2002) showed that $m_0(p) \geq \sqrt{p}$, and that $m_0(p)/\sqrt{p} \approx 1.15096$ for large p. For larger m, least favourable distributions are mixtures of a finite number of uniform distributions on spheres (see Robert, 2001, page 73, and the given references), but the number, position and mixture weights of these spheres require numerical evaluation.

Early and significant contributions to the study of minimax estimation of a normal mean restricted to an interval or a ball of radius m, were given by Bickel (1981) and Levit (1980). These contributions consisted of approximations to the minimax risk and least favourable prior for large m under squared error loss. In particular, Bickel showed that, as $m \to \infty$, the least favourable distributions rescaled to $[-1,1]$ converge weakly to a distribution with density $\cos^2(\pi x/2)$, and that the minimax risks behave like $1 - \frac{\pi^2}{8m^2} + o(m^{-2})$. Extensions and further interpretations of these results were given by Melkman and Ritov (1987), Gajek and Kaluszka (1994), and Delampady and others (2001). There is also a substantial literature on the comparative efficiency of mimimax procedures and affine linear minimax estimators for various models, restricted parameter spaces, and loss functions. A small sample of such work includes Pinsker (1980), Ibragimov and Hasminskii (1984), Donoho, Liu and MacGibbon (1990), and Johnstone and MacGibbon (1992,1993).

Acknowledgements

The authors are grateful to John Hartigan for showing us his result (Theorem 1), and permitting us to make use of it in this review paper. The authors also benefited from discussions with François Perron, helpful in particular for the development presented in Section 6.2. Finally, the authors are thankful to Anirban DasGupta for a diligent reading of a first draft of this review paper, as well as constructive comments and suggestions.

References

[1] Akkerboom, J. C. (1990). *Testing problems with linear or angular inequality constraints*, Lecture Notes in Statistics, **62**, Springer-Verlag, New York. MR1044097

[2] Bader, G. and Bischoff, W. (2003). Old and new aspects of minimax estimation of a bounded parameter, *Mathematical Statistics and Applications, Festschrift for Constance van Eeden.* IMS Lecture Notes-Monograph Series, **42**, pp. 15–30.

[3] Barlow, R. E., Bartholomew, J. M., Bremner, J. M., and Brunk, H. D. (1972). *Statistical Inference under Order Restrictions, The Theory and Application of Isotonic Regression.* Wiley, New York. MR326887

[4] Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, New York, 2nd edition. MR804611

[5] Berry, C. (1989). Bayes minimax estimation of a Bernouilli p in a restricted parameter space, *Communications in Statistics: Theory and Methods*, **18**, 4607–4616. MR1046729

[6] Berry, C. (1990). Minimax estimation of a bounded normal mean vector, *Journal of Multivariate Analysis*, **35**, 130–139. MR1084946

[7] Berry, C. (1993). Minimax estimation of a restricted exponential location parameter, *Statistics and Decisions*, **11**, 307–316. MR1261841

[8] Bickel, P. (1981). Minimax estimation of the mean of a normal distribution when the parameter space is restricted, *Annals of Statistics*, **9**, 1301–1309. MR630112

[9] Bischoff, W. (1992). Minimax estimation and Γ-minimax estimation for functions of a scale parameter family under L_p loss, *Statistics and Decisions*, **29**, 45–61. MR1165702

[10] Bischoff, W. and Fieger, W. (1992). Minimax estimators and Γ-minimax estimators for a bounded normal mean under the loss $\ell_p(\theta, d) = |\theta - d|^p$, *Metrika*, **39**, 185–197. MR1173577

[11] Bischoff, W. and Fieger, W. (1993). On least favourable two-point priors and minimax estimators under absolute error loss, *Metrika*, **40**, 283–298. MR1242740

[12] Bischoff, W., Fieger, W., and Ochtorp, S. (1995). Minimax estimation for the bounded mean of a bivariate normal distribution, *Metrika*, **42**, 379–394. MR1366381

[13] Bischoff, W., Fieger, W., and Wulfert, S. (1995). Minimax and Γ-minimax estimation of a bounded normal mean under linex loss, *Statistics and Decisions*, **13**, 287–298. MR1345373

[14] Blumenthal, S. and Cohen, A. (1968a). Estimation of the larger translation parameter, *Annals of Mathematical Statistics*, **39**, 502–516. MR223006

[15] Blumenthal, S. and Cohen, A. (1968b). Estimation of two ordered translation parameters, *Annals of Mathematical Statistics*, **39**, 517–530. MR223007

[16] Boratyńska, A. (2001). Two point priors and minimax estimation of a bounded parameter, *Unpublished manuscript*.

[17] Brewster, J. F. and Zidek, J. K. (1974). Improving on equivariant estimators, *Annals of Statistics*, **2**, 21–38. MR381098

[18] Brown, L. D. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems, *Annals of Mathematical Statistics*, **42**, 855-903. MR286209

[19] Brown, L. D. (1994). Minimaxity, more or less, in *Statistical Decision Theory and Related Topics, V*, editors S. S. Gupta and J. O. Berger. Springer, New York, pp. 1–18. MR1286291

[20] Brown, L. D. (1996). *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory*. IMS Lecture Notes-Monograph Series, **9**.

[21] Casella, G. and Strawderman, W. E. (1981). Estimating a bounded normal mean, *Annals of Statistics*, **9**, 870–878. MR619290

[22] Charras, A. and van Eeden, C. (1991,a). Bayes and admissibility properties of estimators in truncated parameter spaces, *Canadian Journal of Statistics*, **19**, 121–134. MR1128402

[23] Charras, A. and van Eeden, C. (1991,b). Limit of Bayes estimators in convex truncated parameter spaces, Statistics & Probability Letters, **11**, 479–483. MR1116740

[24] Charras, A. and van Eeden, C. (1992). Bayes properties of estimators of location parameters in truncated parameter spaces, *Statistics and Decisions*, **10**, 81–86. MR1165705

[25] Charras, A. and van Eeden, C. (1994). Inadmissibility for squared loss when the parameter to be estimated is restricted to the interval $[a, \infty)$, *Statistics and Decisions*, **12**, 257–266. MR1309666

[26] Chen, L. and Eichenauer, J. (1988). Two point priors and Γ-minimax estimation in families of uniform distributions, *Statistical Papers*, *29*, 45–57. MR947507

[27] Cohen, A. and Sackrowitz, H. B. (1970). Estimation of the last mean of a monotone sequence, *Annals of Mathematical Statistics*, **41**, 2021–2034. MR270483

[28] DasGupta, A. (1985). Bayes minimax estimation in multiparameter families when the parameter space is restricted to a bounded convex set, *Sankhyā: Series A*, **47**, 326–332. MR863726

[29] Delampady, M., DasGupta, A., Casella, G., Rubin, H., and Strawderman, W. E. (2001). A new approach to default priors and robust Bayes methodology. *Canadian Journal of Statistics*, **29**, 437–450. MR1872645

[30] Drees, H. (1999). On fixed-length confidence intervals for a bounded normal mean, *Statistics & Probability Letters*, **44**, 399–404. MR1721440

[31] Donoho, D. L., Liu, R. C., and MacGibbon, B. (1990). Minimax estimation over hyperrectangles with implications, *Annals of Statistics*, **18**, 1416–1437. MR1062717

[32] Eichenauer, J. (1986). Least favourable two-point priors in estimating the bounded location parameter of a noncentral exponential distribution, *Statistics and Decisions*, **4**, 389–392. MR876878

[33] Eichenauer-Herrmann, J. and Fieger, W. (1989). Minimax estimation in scale parameter families when the parameter interval is bounded, *Statistics and Decisions*, **7**, 363–376. MR1038495

[34] Eichenauer-Hermann, J. and Lehn, J. (1989). Gamma-minimax estimators for a bounded normal mean under squared error loss, *Statistics and Decisions*, **7**, 37–62. MR1000926

[35] Eichenauer-Herrmann, J. and Fieger, W. (1992). Minimax estimation under convex loss when the parameter interval is bounded, *Metrika*, **39**, 27–43. MR1162684

[36] Eichenauer-Hermann, J. and Ickstadt, K. (1992). Minimax estimators for a bounded location parameter, *Metrika*, **39**, 227–237. MR1173580

[37] Farrell, R. H. (1964). Estimators of a location parameter in the absolutely continuous case, *Annals of Mathematical Statistics*, **35**, 949–998. MR171359

[38] Funo, E. (1991). Inadmissibility results of the MLE for the multinomial problem when the parameter space is restricted or truncated, *Communications in Statistics: Theory and Methods*, **20**, 2863–2880. MR1141983

[39] Gatsonis, C., MacGibbon, B., and Strawderman, W. E. (1987). On the estimation of a restricted normal mean, *Statistics & Probability Letters*, **6**, 21–30. MR907255

[40] Gajek, L., Kaluszka, M. (1994). Lower bounds for the asymptotic Bayes risk in the scale model (with applications to the second order minimax estimation), *Annals of Statistics*, **22**, 1831–1839. MR1329170

[41] Gajek, L., Kaluszka, M. (1995). Nonexponential applications of a global Cramèr-Rao inequality, *Statistics*, **26**, 111–122. MR1318207

[42] Ghosh, M. N. (1964). Uniform approximation of minimax point estimates, *Annals of Mathematical Statistics*, **35**, 1031–1047. MR164418

[43] Gourdin, E., Jaumard, B., and MacGibbon, B. (1994). Global optimization decomposition methods for bounded parameter minimax risk evaluation, *Siam Journal of Scientific Computing*, **15**, 16–35. MR1257152

[44] Gupta, A. K. and Rohatgi, V. K. (1980). On the estimation of a restricted mean, *Journal of Statistical Planning and Inference*, **4**, 369–379. MR596771

[45] Hartigan, J. (2003). Uniform priors on convex sets improve risk, *Statistics & Probability Letters*. To appear.

[46] Hwang, G. and Peddada, S. D. (1994). Confidence intervals subject to order restrictions, *Annals of Statistics*, **22**, 67–93. MR1272076

[47] Ibragimov, I. A. and Hasminskii, R. Z. (1984). Nonparametric estimation of the value of a linear functional in a Gaussian white noise, *Theory of Probability and Its Applications*, **29**, 1–32. MR739497

[48] Iliopoulos, G. (2000). A note on decision theoretic estimation of ordered parameters, *Statistics & Probability Letters*, **50**, 33–38. MR1804623

[49] Iliopoulos, G. and Kourouklis, S. (2000). Interval estimation for the ratio of scale parameters and for ordered scale parameters, *Statistics and Decisions*, **18**, 169–184. MR1782343

[50] Iwasa, M. and Moritani, Y. (1997). A note on the admissibility of the maximum likelihood estimator for a bounded normal mean, *Statistics & Probability Letters*, **32**, 99–105. MR1439501

[51] Johnstone, I. and MacGibbon, B. (1992). Minimax estimation of a constrained Poisson vector, *Annals of Statistics*, **20**, 807–831. MR1165594

[52] Johnstone, I. and MacGibbon, B. (1993). Asymptotically minimax estimation of a constrained Poisson vector via polydisc transforms, *Annales de l'Institut Henri Poincaré- Probabilité et Statistiques*, **2**, 289–319. MR1227421

[53] Jafari Jozani, M., Nematollahi, N., Shafie, K. (2002). An admissible minimax estimator of a bounded scale-parameter in a subclass of the exponential family under scale-invariant squared error loss, *Statistics & Probability Letters*, **60**, 437–444. MR1947183

[54] Kamberova, G. and Mintz, M. (1999). Minimax rules under zero-one loss for a restricted location parameter, *Journal of Statistical Planning and Inference*, **79**, 205–221. MR1700760

[55] Karlin, S. (1957). Pólya type distributions, II, *Annals of Mathematical Statistics*, **28**, 281–309. MR88148

[56] Katz. M. (1961). Admissible and Minimax Estimates of Parameters in Truncated Spaces, *Annals of Mathematical Statistics*, **32**, 136–142. MR119287

[57] Kempthorne, P. J. (1987). Numerical specification of discrete least favorable prior distributions, *Siam Journal of Scientific Computing*, **8**, 171–184. MR879409

[58] Kempthorne, P. J. (1988). Dominating inadmissible procedures using compromise decision theory, *Statistical Decision Theory and Related Topics IV*, editors S. S. Gupta and J. O. Berger. Springer-Verlag, New York, pp. 381–396. MR927113

[59] Kubokawa, T., (1994a). A Unified Approach to Improving Equivariant Estimators, *Annals of Statistics*, **22**, 290–299. MR1272084

[60] Kubokawa, T., (1994b). Double shrinkage estimation of ratio of scale parameters, *Annals of the Institute of Statistical Mathematics*, **46**, 95–119. MR1272751

[61] Kubokawa, T. and Saleh, A. K. MD. E. (1994). Estimation of location and scale parameters under order restrictions. *Journal of Statistical Research*, **28**, 41–51. MR1370413

[62] Kubokawa, T., (1998). The Stein phenomenon in simultaneous estimation: A review. *Applied Statistical Science III* (editors S. E. Ahmed, M. Ahsanullah, and B. K. Sinha), NOVA Science Publishers, New York, pp. 143–173. MR1673649

[63] Kubokawa, T., (1999). Shrinkage and modification techniques in estimation of variance and the related problems: A review. *Communications in Statistics: Theory and Methods*, **28**, 613–650. MR1703774

[64] Kumar S. and Sharma, D. (1988). Simultaneous estimation of ordered parameters. *Communications in Statistics: Theory and Methods*, **17**, 4315–4336. MR981031

[65] Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation.* Springer-Verlag, New York, 2nd edition. MR1639875

[66] Lehn, J. and Rummel, F. (1987). Gamma-minimax estimation of a binomial probability under squared error loss, *Statistics and Decisions*, **5**, 229–249. MR905239

[67] Levit, B.Y.A. (1980). On asymptotic minimax estimation of the second order, *Theory Probab. Appl.*, **25**, 552–568. MR582585

[68] MacGibbon, B., Gourdin, E., Jaumard, B., and Kempthorne, P. J. (2000). Minimax estimation of exponential family means over l_p bodies under quadratic loss, in *CMS Conference Proceedings*, **26**, 319–330. MR1765018

[69] Marchand, É. and MacGibbon, B. (2000). Minimax estimation of a constrained binomial proportion, *Statistics and Decisions*, **18**, 129–167. MR1782342

[70] Marchand, É. and Perron, F. (2001). Improving on the MLE of a bounded normal mean. *Annals of Statistics*, **29**, 1078–1093. MR1869241

[71] Marchand, É. and Perron, F. (2002). On the minimax estimator of a bounded normal mean. *Statistics & Probability Letters*, **58**, 327–333. MR1923455

[72] Marchand, É. and Perron, F. (2003). Improving on the mle of a bounded location parameter for spherical distributions, *Journal of Multivariate Analysis*. To appear. MR1869241

[73] Marchand, É. and Strawderman, W. E. (2003a). On improving on the minimum risk equivariant estimator of a location parameter which is constrained to an interval or a half-interval, *Annals of the Institute of Statistical Mathematics*. To appear.

[74] Marchand, É. and Strawderman, W. E. (2003b). On improving on the minimum risk equivariant estimator of a scale parameter under a lower-bound constraint, *Journal of Statistical Planning and Inference*. To appear.

[75] Mandelkern, M. (2002). Setting Confidence Intervals for Bounded Parameters with discussion, *Statistical Science*, **17**, 149–172. MR1939335

[76] Melkman, A. A. and Ritov, Y. (1987). Minimax estimation of the mean of a general distribution when the parameter space is restricted, **15**, 432–442. MR885749

[77] Moors, J. J. A. (1981). Inadmissibility of linearly invariant estimators in the truncated parameter spaces, *Journal of the American Statistical Association*, **76**, 910–915. MR650903

[78] Moors, J. J. A. (1985). *Estimation in truncated parameter spaces*, Ph.D. thesis, Tilburg University.

[79] Moors, J. J. A. and van Houwelingen, J. C. (1993). Estimation of linear models with inequality restrictions, Statistica Neerlandica, **47**, 185–198. MR1243854

[80] Nelson, W. (1965). Minimax solution of statistical decision problems by iteration, *Annals of Mathematical Statistics*, **37**, 1643–1657. MR198635

[81] Parsian, A., Sanjari Farsipour, N. (1997). Estimation of parameters of exponential distribution in the truncated space using asymmetric loss function. *Statistical Papers*, 38, 423–443. MR1488641

[82] Perron, F. (2003). Improving on the mle of p for a Binomial(n, p) when p is around $1/2$, *Mathematical Statistics and Applications, Festschrift for Constance van Eeden*, IMS Lecture Notes-Monograph Series, **42**, pp. 45–64.

[83] Pinsker, M. S. (1980). Optimal filtering of square integrable signals in Gaussian white noise, *Problems Inform. Transmission*, 120–133.

[84] Robert, C. (2001). *The Bayesian Choice: From decision-theoretic foundations to computational implementation*. Springer-Verlag, New York, 2nd edition. MR1835885

[85] Robertson, T., Wright, F. T., and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. Wiley, New York. MR961262

[86] Rukhin, A. L. (1990). Comments on 'Developments in decision-theoretic variance estimation', by Maatta J. M. and Casella, G., *Statistical Science*, **5**, 113–116.

[87] Sacks, J. (1963). Generalized Bayes solutions in estimation problems, *Annals of Mathematical Statistics*, **34**, 751–768. MR150908

[88] Sackrowitz, H. B. and Strawderman, W. (1974). On the admissibility of the MLE for ordered binomial parameters, *Annals of Statistics*, **2**, 822–828. MR368242

[89] Sackrowitz, H. B. (1982). Procedures for improving the MLE for ordered binomial parameters, *Journal of Statistical Planning and Inference*, **6**, 287–296. MR664351

[90] Shao, P. Y.-S. and Strawderman, W. (1994). Improving on truncated estimators, in *Statistical Decision Theory and Related Topics, V*, editors S. S. Gupta and J. O. Berger. Springer, New York, pp. 369–376. MR1286313

[91] Shao, P. Y.-S. and Strawderman, W. (1996a). Improving on the mle of a positive normal mean, *Statistica Sinica*, **6**, 275–287. MR1379061

[92] Shao, P. Y.-S. and Strawderman, W. (1996b). Improving on truncated linear estimates of exponential and gamma scale parameters, *Canadian Journal of Statistics*, **24**, 105–114. MR1394744

[93] Stark, P. (1992). Affine minimax confidence intervals for a bounded normal mean, *Statistics & Probability Letters*, **13**, 39–44. MR1147637

[94] Stein, C. (1981). Estimation of the mean of a multivariate normal distribution, *Annals of Statistics*, **9**, 1135–1151. MR630098

[95] Strawderman, W. (2000). Minimaxity, *Journal of the American Statistical Assocation*, **95**, 1364–1368. MR1825295

[96] van Eeden, C. (1995). Minimax estimation of a lower-bounded scale parameter of a gamma distribution for scale-invariant squared error loss, *Canadian Journal of Statistics*, **23**, 245–256. MR1363590

[97] van Eeden, C. (1996). Estimation in restricted parameter spaces- some history and recent developments, *CWI Quarterly*, **9**, 69–76. MR1420018

[98] van Eeden, C. and Zidek, J. V. (1999). Minimax estimation of a bounded scale parameter for scale-invariant squared-error loss, *Statistics and Decisions*, **17**, 1–30. MR1701333

[99] van Eeden, C. (2000). Minimax estimation of a lower-bounded scale parameter of an F distribution, *Statistics & Probability Letters*, **46**, 283–286.

[100] van Eeden, C. and Zidek, J. V. (2001). Estimating one of two normal means when their difference is bounded, *Statistics & Probability Letters*, **51**, 277–284. MR1822735

[101] van Eeden, C. and Zidek, J. V. (2003). Combining sample information in estimating ordered normal means, *Sankhyā, Series A*, **64**, 588–610. MR1985401

[102] Vidakovic, B. (1993). On the efficiency of affine minimax rules in estimating a bounded multivariate normal mean, *Communications in Statistics: Theory and Methods*, **22**, 655–669. MR1235867

[103] Vidakovic, B. and DasGupta, A. (1996). Efficiency of linear rules for estimating a bounded normal mean, *Sankhyā* , **58**, 81–100. MR1659051

[104] Wald, A. (1950). *Statistical Decision Functions*. Wiley, New York. MR36976

[105] Wan, T., Zou, G., and Lee, A. (2000). Minimax and Γ-minimax estimation for a Poisson distribution under linex loss when the parameter space is restricted, *Statistics & Probability Letters*, **50**, 23–32. MR1804622

[106] Zeytinoglu, M. and Mintz, M. (1984). Optimal fixed size confidence procedures for a restricted parameter space, *Annals of Statistics*, **12**, 945–957. MR751284

[107] Zeytinoglu, M. and Mintz, M. (1988) Robust fixed size confidence procedures for a restricted parameter space, *Annals of Statistics*, **16**, 1241–1253. MR959199

[108] Zhang, T. and Woodroofe, M. (2003). Credible confidence sets for restricted parameter spaces, *Journal of Statistical Planning and Inference*, **115**, 479–490. MR1985880

[109] Zinzius, E. (1981). *Minimaxschätzer für den Mittelwert θ einer normalverteilen Zufallsgröß e mit bekannter Varianz bei vorgegebener oberer und unterer Schranke für θ*, *Matheamatische Operationsforschung Statistik, Series Statistics*, **12**, 551–557. MR639251