OLD AND NEW ASPECTS OF MINIMAX ESTIMATION OF A BOUNDED PARAMETER

GUIDO BADER AND WOLFGANG BISCHOFF
Universität Karlsruhe

In practice the unknown parameter of an experiment is often bounded. Therefore there is practical interest to include such additional knowledge into statistical procedures to improve them. In this paper we consider minimax estimation of a bounded parameter. The minimax principle is, at least from a theoretical point of view, very important. In the last two decades of the last century a sequence of papers has treated such problems for some densities that are smooth with respect to the unknown parameter and for specific convex losses. It occurred that in all examples in which the loss is strictly convex and the parameter interval is sufficiently small a minimax estimator exists which is Bayes with respect to a two-point prior with mass at the boundaries. In this paper we show that this result is true in general when the Lebesgue-densities are smooth with respect to the unknown parameter and the loss is strictly convex, but do not necessarily penalize equally underestimation and overestimation. Our result provides new classes of interesting losses and densities for which the above statement holds true.

1. Introduction

In practice there are typical situations where it is most appealing to apply the minimax principle. On the other hand, there are several reasons refraining the utilization of minimax rules, for example:

- minimax rules are difficult to calculate,
- there is no universal method to construct a minimax rule,
- examples are known for which a decision rule exists whose maximal risk is a little larger than the minimax risk but whose risk function is much better than the risk function of the minimax rule for large parts of the parameter space. This is especially the case when the minimax risk function is constant. In such a situation the minimax rule is not attractive to a practitioner.

Nevertheless, from a theoretical point of view it is important to obtain minimax rules, see Brown (1994). Note that only if the risk function of the minimax rule is known, rules that are more convenient with respect to the risk function can be developed. Moreover, the situation is a little different in case where the parameter space is restricted because then the risk function of the minimax estimator is typically not constant. Consequently the risk functions of minimax estimators for restricted problems are often quite more appealing than in the unbounded case where they are often constant.

Professor Constance van Eeden was one of the first scientists who made contributions to problems with restricted parameter spaces, see van Eeden (1956, 1957a–c, 1958). Since that time she (sometimes with coauthors) has provided main contributions to this topic. In this paper we consider only bounded parameter spaces. Some history for results on other restricted parameter spaces is given in van Eeden (1996).

In the statistical literature it is often assumed that the parameter space is unbounded which seems to be never fulfilled in practice if the parameter describes, for example, the mean or the variance of a real phenomenon (e.g., rainfall-runoff, weight or height of persons/animals, values for controlling an industrial process). Therefore there is practical interest to include such additional information into statistical procedures to improve them. In this paper we consider the minimax estimation of a bounded parameter. In the last two decades of the last century a sequence of papers has treated such problems for certain densities that are smooth with respect to the unknown parameter and for specific convex losses. It occurred that in all examples in which the loss is strictly convex and the parameter interval is sufficiently small a minimax estimator exists which is Bayes with respect to a two-point prior with mass at the boundaries. In this paper we show that this result is true in general when the Lebesgue-densities are smooth with respect to the unknown parameter and the loss is strictly convex, but do not necessarily penalize equally underestimation and overestimation. So, our result provides new classes of interesting losses and densities for which the above statement holds true.

In the next section we give some historical remarks on the problem and discuss the relevant literature. The main result and some examples are given in Section 3. The arguments are contained in Section 4 and Section 5. More precisely, in Section 4 the Bayes estimator of a two-point prior is determined and properties of this estimator are given and in Section 5 the properties of the risk function of that Bayes estimator are investigated and the proof of the main result is given.

2. Historical remarks on estimating a bounded parameter

The first results on exact minimax estimation of a bounded parameter were published by Casella and Strawderman (1981) and Zinzius (1979, 1981). They gave the minimax estimator of a bounded normal mean θ with squared error loss. More exactly, if $\theta \in [a, b] \subset \mathbb{R}$ and $b - a$ is sufficiently small, then the prior with mass $1/2$ in a and b, respectively, is least favorable and the corresponding Bayes estimator is minimax. They used the following well-known result (see, e.g., Lehmann and Casella, 1998, p. 310, Theorem 1.4):

Theorem 2.1. *If δ is a Bayes estimator with respect to a prior distribution π and if the Bayes risk $r(\pi, \delta)$ is equal to the supremum risk,* $\sup_\theta R(\theta, \delta)$, *then δ is minimax and π is least favorable.*

Casella and Strawderman investigated the risk function of the two-point Bayes estimator mentioned above using a result by Karlin (1957). So, they could establish that the condition for the risk function given in the above theorem is fulfilled if and only if $b - a < m_0$, where $m_0 \approx 2.1$ is calculated numerically. In contrast, Zinzius proved his result by using the following corollary to the above theorem.

Corollary 2.1 (Convexity technique). *If δ is a Bayes estimator with respect to a two-point prior π with mass in a and b and if $R(\cdot, \delta)$ is convex with $R(a, \delta) = R(b, \delta)$, then δ is minimax and π is least favorable.*

Obviously, Zinzius could only obtain a smaller bound than m_0 and the best known bound analytically calculated with the convexity technique was stated in Bischoff and Fieger (1992): $m_1 = \sqrt{2}$. Following these papers, minimax estimation of a bounded one dimensional parameter has been considered by many authors for different distributions and loss functions, see Eichenauer (1986), Chen and Eichenauer (1988), Eichenauer-Herrmann and Fieger (1989, 1992), Bischoff (1992), Bischoff and Fieger (1992), Bischoff, Fieger and Wulfert (1995), van Eeden and Zidek (1999), Marchand and MacGibbon (2000), Wan, Zou and Lee (2000) and the papers cited there. These papers have in common the following facts: if the length of the parameter interval is sufficiently small, then a two-point prior with mass at the boundaries of the parameter interval is least favorable, and the corresponding Bayes estimator is minimax. Furthermore, all papers assumed a strictly convex loss and densities being smooth with respect to the unknown parameter. It is worth mentioning that all these results have been proven using the convexity or a related technique. Furthermore, we like to mention that minimaxity results have also been obtained without the restriction that the parameter space has to be sufficiently small, see, e.g., Moors (1985) and Berry (1989).

It is worth noting that we assume a specific form of the loss function, see the next section. For instance the loss function considered by van Eeden and Zidek (1999) does not fulfill our assumptions. In a forthcoming paper, however, it will be shown that such loss functions and problems satisfy our conditions after a certain transformation.

In this paper we consider only the case where the parameter space is a subset of $\mathbb{R}$. However we will make the following remarks about multivariate spaces. The technique used by Casella and Strawderman could only be successfully applied to a multivariate normal with a mean vector bounded to a sufficient small sphere, see Berry (1990). For other problems, however, the technique used by Casella and Strawderman has not been applied up to now in contrast to the convexity technique. DasGupta (1985) rediscovered the convexity technique and generalized Corollary 2.1 to the case that the parameter region is multivariate, but he only applied the convexity technique to one dimensional problems. A two dimensional problem, which can not be handled with Berry's results mentioned above, was analyzed in Bischoff, Fieger and Ochtrop (1995) using the convexity technique.

In case where the loss is only convex but not strictly convex, Eichenauer-Herrmann and Ickstadt (1992), and Bischoff and Fieger (1993) under milder conditions, showed that there are cases where there does not exist a two-point prior which is least favorable and whose corresponding Bayes estimator is minimax. For example, this is the case for the bounded normal mean and the L_1-loss, $L(\theta, a) = |\theta - a|$. Hence the question arises under which conditions the following statement holds true:

Statement 2.2. *A two-point prior with mass in a and b is least favorable and the corresponding Bayes estimator is minimax.*

There are two papers that investigate this problem. Namely, Eichenauer-Herrmann and Fieger (1992) and Boratyńska (2001). The first paper investigates minimax estimation under convex loss when the parameter is bounded, this for a class of special truncated Lebesgue densities. Boratyńska (2001) proves Statement 2.2 in the case where: the parameter interval is sufficiently small, the loss function is strictly convex and three times continuously differentiable, the densities are twice differentiable with respect to the unknown parameter and certain integrability conditions are satisfied. There are many loss functions and some interesting classes of densities, see Section 3, that do not fulfill these assumptions. We show, under weaker/different assumptions than those given by Boratyńska (2001), that Statement 2.2 holds true. It is worth mentioning that to show convexity we use a different approach than the one used by Boratyńska (2001).

3. Main results and examples

We investigate a real random variable X distributed according to a probability measure P_θ, where θ is an unknown parameter that belongs to a bounded parameter space $\Theta = [\theta_0, \theta_0 + m] \subset \mathbb{R}$ with fixed θ_0 and $m \in (0, M)$, where $M > 0$. (Note that we introduce the constant M only to facilitate the statement of some assumptions.)

Remark 3.1. We like to emphasize, that in many cases our one-dimensional sample space $\mathbb{R}$ is not a restriction at all. Often one has n independent and identically distributed real random variables $X_1, \ldots, X_n$ and a sufficient statistic $X\colon \mathbb{R}^n \to \mathbb{R}$. So we can apply our result to X and it still holds true for the original problem by the Rao–Blackwell Theorem.

Equipping the sample space $\mathbb{R}$ with its Borel σ-algebra $\mathcal{B}$ and Θ with its Borel σ-algebra $\mathcal{B}(\Theta)$, the risk function for a given loss function $L_\theta\colon \mathbb{R} \to [0,\infty)$, $\theta \in \Theta$, and an arbitrary randomized estimator $\delta\colon \mathbb{R} \times \mathcal{B}(\Theta) \to [0,1]$, is defined by

$$R(\theta,\delta) = \int_{\mathbb{R}} \int_{\Theta} L_\theta(a)\, \delta(x,da)\, P_\theta(dx).$$

Hence, our estimators satisfy the natural condition

$$P_\theta(\{x \in \mathbb{R} \mid \delta(x,\Theta) = 1\}) = 1, \quad \theta \in \Theta,$$

that is for a nonrandomized estimator δ, $P_\theta(\{x \in \mathbb{R} \mid \delta(x) \in \Theta\}) = 1$, $\theta \in \Theta$.

An estimator δ^* is called minimax if

$$\sup_{\theta\in\Theta} R(\theta,\delta^*) = \inf_{\delta\in\Delta} \sup_{\theta\in\Theta} R(\theta,\delta),$$

where Δ denotes the set of all randomized estimators for the unknown parameter θ. Moreover, if π is a prior distribution on $\mathcal{B}(\Theta)$, then the Bayes risk of an arbitrary estimator δ is given by

$$r(\pi,\delta) = \int_{\Theta} R(\theta,\delta)\, \pi(d\theta).$$

We assume that the loss $L_\theta(a)$ can be written as a function $h\colon (-M,M) \to [0,\infty)$ of the difference between the parameter θ and the decision a:

$$\forall\, \theta, a \in \Theta: \quad L_\theta(a) = h(\theta - a).$$

In this case the function h is also called loss function. Further we assume that the family of probability measures $\{P_\theta \mid \theta \in \Theta\}$ has a density with respect to the Lebesgue-measure λ on $(\mathbb{R},\mathcal{B})$:

$$f_\theta(x) = f(x,\theta) := \frac{dP_\theta}{d\lambda}(x), \quad x \in \mathbb{R}.$$

To establish our main result we need the following assumptions on the loss function h and the densities f_θ.

Assumption on the loss function h. *The function $h\colon (-M,M) \to [0,\infty)$ has the following properties:*

A.1 $\forall\, s < t < 0$: $h(s) > h(t) > h(0) = 0$ *and* $\forall\, 0 < t < s$: $0 = h(0) < h(t) < h(s)$,

A.2 h *is continuously differentiable and* $h|_{(-M,0)\cup(0,M)}$ *is twice continuously differentiable with* $h''(s) > 0$, *and for* $s \in (-M, 0)$, *respectively* $s \in (0, M)$, h *satisfies one of the following conditions (not necessarily the same on both sides of the origin)*

A.3 $\exists K > 0\ \forall s \in (-M, 0)$ *(or* $\forall s \in (0, M)$ *resp.):* $h''(s) > K$,

A.4 $\exists K, k\colon (0, \infty) \to (0, \infty) \forall 0 < m < M \forall -m < s < 0$ *(or* $\forall 0 < s < m$ *resp.)*

$$\frac{h''(s)}{|h'(s)|} \geq K(m), \quad \frac{h(s)}{|h'(s)|} \leq k(m), \quad \lim_{m \to 0} K(m) = \infty, \quad \lim_{m \to 0} k(m) = 0.$$

As a consequence of these assumptions we get immediately that

$$h'(s) < 0,\ s \in (-M, 0); \quad h'(s) > 0,\ s \in (0, M); \quad h'(0) = 0, \tag{3.1}$$

and

$$G_m\colon (-m, 0) \to (0, \infty),\ s \mapsto \frac{-h'(s)}{h'(m+s)} \tag{3.2}$$

is bijective and strictly decreasing.

Assumption on the densities f_θ. *The density functions* $f_\theta\colon \mathbb{R} \times [\theta_0, \theta_0 + M) \to [0, \infty)$ *have the following properties:*

A.5 $\forall x \in \mathbb{R}$: $\partial f(x, \theta)/\partial \theta$ *exists up to countably many* $\theta \in (\theta_0, \theta_0 + M)$ *and is uniformly bounded,*

A.6 $\forall x \in \mathbb{R}$: $\partial^2 f(x, \theta)/\partial \theta^2$ *exists up to countably many* $\theta \in (\theta_0, \theta_0 + M)$ *and is uniformly bounded.*

A.7 (Only needed in combination with Assumption A.4.)
$\forall x \in \mathbb{R}$: $[\partial f(x, \theta)/\partial \theta]/f(x, \theta)$ *and* $[\partial^2 f(x, \theta)/\partial \theta^2]/f(x, \theta)$ *are uniformly bounded up to countably many* $\theta \in (\theta_0, \theta_0 + M)$ *where* $\frac{0}{0} = 0$.

We state now our main result which will be a consequence of Corollary 2.1 stated above and of Corollary 5.1 and Lemma 5.2 that will follow.

Theorem 3.1. *Under the Assumptions* A.1–A.7 *there exists a constant* $m_0 \in (0, M)$ *such that for each* $m \in (0, m_0]$ *Statement 2.2 holds true with* $a = \theta_0$ *and* $b = \theta_0 + m$, *if* $P_\theta(\{x \in \mathbb{R} \mid f(x, \theta_0) f(x, \theta_0 + m) > 0\}) > 0$ *for* $\theta \in \{\theta_0, \theta_0 + m\}$.

Note that $P_{\theta_0}(\{x \in \mathbb{R} \mid f(x, \theta_0) f(x, \theta_0 + m) > 0\}) = 0$ implies $R(\theta_0, \delta) = 0$ and $P_{\theta_0 + m}(\{x \in \mathbb{R} \mid f(x, \theta_0) f(x, \theta_0 + m) > 0\}) = 0$ implies $R(\theta_0 + m, \delta) = 0$ for each Bayes estimator with respect to a two-point prior with mass in $\{\theta_0, \theta_0 + m\}$, see Section 5. Thus the last assumption in the above theorem is a necessary condition for Statement 2.2 to be true for $a = \theta_0$ and $b = \theta_0 + m$.

It is worth mentioning that an L_p-loss

$$h\colon (-M, M) \to \mathbb{R}, \ s \mapsto |s|^p,$$

for $p > 1$ fulfills our Assumption on the loss function h whereas the assumptions of Boratyńska (2001) only cover the case $p = 2$. Moreover, uniform distributions satisfy our Assumption on the densities f_θ but not the assumption of Boratyńska (2001). Very often in practice we need a loss function where over- and underestimation is judged by different losses. As a consequence the loss function is not as smooth at the origin as often assumed. To end this section we present some important examples of loss functions.

Example 3.1. Here we have a look at L_p-losses with different p on the right- and the left-hand side of the origin. Let

$$h_1(s) = \begin{cases} |s|^{p_1}, & s \leq 0 \\ s^{p_2}, & s > 0, \end{cases}$$

with $p_1, p_2 > 1$.

Obviously Assumptions A.1 and A.2 are fulfilled; we see, that h_1 is not twice differentiable in the origin if $p_1 < 2$ or $p_2 < 2$. Moreover, we can easily verify Assumption A.3 for $1 < p_i \leq 2$, $i = 1$ or $i = 2$, and Assumption A.4 for $2 < p_i$, $i = 1$ or $i = 2$.

We can also combine L_p-losses with the well-known Linex Loss.

Example 3.2. Let

$$h_2(s) = \begin{cases} |s|^p, & s \leq 0 \\ \gamma \cdot (\exp\{\alpha s\} - \alpha s - 1), & s > 0, \end{cases}$$

with $p > 1$, $\alpha > 0$ and $\gamma > 0$.

Here Assumption A.1–A.3 (or A.4 for $p > 2$ on the left-hand side of the origin) holds true and h_2 is not twice differentiable in the origin except for the case where $p = \gamma\alpha^2 = 2$.

Example 3.3. Similarly to h_2 we consider

$$h_3(s) = \begin{cases} \gamma \cdot (\exp\{\alpha s\} - \alpha s - 1), & s \leq 0 \\ |s|^p, & s > 0, \end{cases}$$

with $p > 1$, $\alpha > 0$ and $\gamma > 0$.

Again Assumption A.1–A.3 (or A.4 for $p > 2$ on the right-hand side of the origin) holds true and h_3 is not twice differentiable in the origin except for the case $p = \gamma\alpha^2 = 2$.

4. The Bayes estimator with respect to the two-point prior at the boundary

In this section we determine the Bayes estimator $\delta^{\beta,m}$ with respect to the two-point prior $\pi_\beta, \beta \in [0,1]$, defined by

$$\pi_\beta(\{\theta_0\}) = \beta, \quad \pi_\beta(\{\theta_0+m\}) = 1-\beta.$$

The Bayes risk of an arbitrary randomized estimator δ with respect to the prior π_β is given by

$$\begin{aligned} r(\pi_\beta, \delta) & \\ &= \int_\Theta R(\theta,\delta)\,\pi_\beta(d\theta) = \int_\Theta\int_{\mathbb{R}}\int_\Theta h(\theta-a)\,\delta(x,da) f(x,\theta)\,\lambda(dx)\,\pi_\beta(d\theta) \\ &= \int_{\mathbb{R}}\int_\Theta h(\theta_0-a)f(x,\theta_0)\beta + h(\theta_0+m-a)f(x,\theta_0+m)(1-\beta)\,\delta(x,da)\,\lambda(dx). \end{aligned}$$

To determine a Bayes estimator $\delta^{\beta,m}$, $\beta\in[0,1]$, $m>0$ arbitrarily fixed, we discuss for each fixed $x\in\mathbb{R}$ under which conditions the function

$$\begin{aligned} H_{\beta,m,x}\colon [\theta_0,\theta_0+m] &\to [0,\infty) \\ a &\mapsto h(\theta_0-a)f(x,\theta_0)\beta + h(\theta_0+m-a)f(x,\theta_0+m)(1-\beta) \end{aligned}$$

has a minimum. Since the loss h is continuous

$$M(\beta,m,x) := \{s\in\Theta \mid H_{\beta,m,x}(s) = \min_{a\in\Theta} H_{\beta,m,x}(a)\} \neq \varnothing$$

for each $x\in\mathbb{R}$, and therefore each measurable function $\delta^{\beta,m}\colon \mathbb{R}\to\Theta$ with $\delta^{\beta,m}(x)\in M(\beta,m,x)$ is a Bayes estimator with respect to π_β. We have to discuss different situations:

- If $f(x,\theta_0) = f(x,\theta_0+m) = 0$, then the above function is minimal for each a.
- If $f(x,\theta_0)=0$, and $f(x,\theta_0+m)>0$, then the above function is minimal for $a=\theta_0+m$.
- If $f(x,\theta_0)>0$, and $f(x,\theta_0+m)=0$, then the above function is minimal for $a=\theta_0$.
- Finally we have to consider the case where $f(x,\theta_0) > 0$, and $f(x,\theta_0+m)>0$.

The following lemmas are useful to calculate $\delta^{\beta,m}(x)$ for $\beta\in(0,1)$.

Lemma 4.1. *Let* $\beta\in(0,1)$, $m>0$, *and* $x\in\{x\in\mathbb{R} \mid f(x,\theta_0)f(x,\theta_0+m)>0\}$. *Then we have:*

(a) *$s \in M(\beta, m, x)$ if and only if $\forall\, a \in [\theta_0, s)$:*

$$\frac{h(\theta_0 - a) - h(\theta_0 - s)}{s - a} f(x, \theta_0)\beta \geq \frac{h(\theta_0 + m - a) - h(\theta_0 + m - s)}{a - s} f(x, \theta_0 + m)(1 - \beta),$$

and $\forall\, a \in (s, \theta_0 + m]$:

$$\frac{h(\theta_0 - s) - h(\theta_0 - a)}{s - a} f(x, \theta_0)\beta \geq \frac{h(\theta_0 + m - s) - h(\theta_0 + m - a)}{a - s} f(x, \theta_0 + m)(1 - \beta).$$

(b) *If the right hand derivative h'^{+} of h exists in $-m$, then a necessary condition for $\theta_0 + m \in M(\beta, m, x)$ is given by:*

$$-h'^{+}(-m) f(x, \theta_0)\beta \leq h'(0) f(x, \theta_0 + m)(1 - \beta).$$

(c) *If the left hand derivative h'^{-} of h exists in m, then a necessary condition for $\theta_0 \in M(\beta, m, x)$ is given by:*

$$-h'(0) f(x, \theta_0)\beta \geq h'^{-}(m) f(x, \theta_0 + m)(1 - \beta).$$

(d) *If the derivative h' of h exists in $\theta_0 - s \in (-m, 0)$ and in $\theta_0 + m - s \in (0, m)$, then a necessary condition for $s \in M(\beta, m, x)$ is given by:*

$$-h'(\theta_0 - s) f(x, \theta_0)\beta = h'(\theta_0 + m - s) f(x, \theta_0 + m)(1 - \beta). \tag{4.1}$$

(e) *Let $s \in (\theta_0, \theta_0 + m)$ be arbitrary fixed and h differentiable. Then a necessary condition for $s \in M(\beta, m, x)$ is given by:*

$$\exists\, \epsilon > 0 : (\theta_0 - s - \epsilon, \theta_0 - s + \epsilon) \to \mathbb{R}, \quad t \mapsto G_m(t) \text{ is decreasing,} \tag{4.2}$$

where G_m is defined in (3.2).

Proof. Note that $s \in M(\beta, m, x)$ if and only if for each $a \in [\theta_0, \theta_0 + m]$ we have

$$h(\theta_0 - s) f(x, \theta_0)\beta + h(\theta_0 + m - s) f(x, \theta_0 + m)(1 - \beta) \leq h(\theta_0 - a) f(x, \theta_0)\beta + h(\theta_0 + m - a) f(x, \theta_0 + m)(1 - \beta).$$

Then (a) follows easily. (a) implies (b) and (c) by taking first limits from the right and the left, and then putting $s = \theta_0 + m$ and $s = \theta_0$, respectively. Result (d) is a consequence of the proof of b) and c). Let the conditions of e) be fulfilled. Then by (4.1) a necessary and sufficient condition for $H_{\beta,m,x}(\cdot)$ to have a local minimum in s is given by (4.2). □

Lemma 4.1(b), (c) and (3.1) imply immediately the following result.

Lemma 4.2. *Let* $\beta \in (0,1)$, $m \in (0,M)$, $x \in \{x \in \mathbb{R} \mid f(x,\theta_0)f(x,\theta_0+m) > 0\}$. *Let the function* h *fulfill Assumption* A.1–A.4. *Then*

$$M(\beta,m,x) \subset (\theta_0,\theta_0+m). \tag{4.3}$$

Lemma 4.3. *Let* $\beta \in (0,1)$, $m \in (0,M)$, $x \in D_{\theta_0,\theta_0+m} := \{x \in \mathbb{R} \mid f(x,\theta_0)f(x,\theta_0+m) > 0\}$. *Let the function* h *fulfill Assumption* A.1–A.4. *Then* $\delta^{\beta,m} : D_{\theta_0,\theta_0+m} \to [\theta_0,\theta_0+m]$ *is measurable and* $\delta^{\beta,m}(x), x \in D_{\theta_0,\theta_0+m}$, *is uniquely determined by the equation*

$$-h'\big(\theta_0 - \delta^{\beta,m}(x)\big)f(x,\theta_0)\beta = h'\big(\theta_0 + m - \delta^{\beta,m}(x)\big)f(x,\theta_0+m)(1-\beta).$$

Furthermore, $\delta^{\beta,m}(x), x \in D_{\theta_0,\theta_0+m}$, *is strictly decreasing with respect to* β. *Moreover,* $\delta^{\beta,m}(x), x \in D_{\theta_0,\theta_0+m}$, *is continuous with respect to* $\beta \in (0,1)$.

Proof. The continuity of h and (4.3) imply $\varnothing \neq M(\beta,m,x) \subset (\theta_0,\theta_0+m)$. Hence, we can apply Lemma 4.1(d) and see that each element of $M(\beta,m,x)$ fulfills equation (4.1). This equation can be written as

$$G_m(\theta_0 - s) = \frac{f(x,\theta_0+m)}{f(x,\theta_0)} \cdot \frac{1-\beta}{\beta}, \tag{4.4}$$

which by (3.2), has a unique solution $s \in (-m,0)$. Thus $\delta^{\beta,m}(x)$ is uniquely determined by

$$G_m(\theta_0 - \delta^{\beta,m}(x)) = \frac{f(x,\theta_0+m)}{f(x,\theta_0)} \cdot \frac{1-\beta}{\beta}. \tag{4.5}$$

Moreover, (3.2) implies the existence and measurability of G_m^{-1}. Hence,

$$\delta^{\beta,m}(x) = \theta_0 - G_m^{-1}\left(\frac{f(x,\theta_0+m)}{f(x,\theta_0)} \cdot \frac{1-\beta}{\beta}\right) \tag{4.6}$$

is measurable. By equation (4.6) we see that $\delta^{\beta,m}(x)$ is strictly decreasing with respect to β. Moreover, it is easy to see that G_m^{-1} is continuous. Thus $\delta^{\beta,m}(x)$ is continuous with respect to $\beta \in (0,1)$. □

The discussion at the very beginning of this section and Lemma 4.3 imply the following

Corollary 4.1. *Let* $\beta \in [0,1]$, $m \in (0,M)$, *and let the function* h *fulfill Assumption* A.1–A.4. *Then*

$$\delta^{0,m}\colon \mathbb{R} \to [\theta_0,\theta_0+m], \quad x \mapsto \begin{cases} \theta_0, & \textit{if } f(x,\theta_0) > 0,\ f(x,\theta_0+m) = 0 \\ \theta_0 + m, & \textit{else} \end{cases}$$

and

$$\delta^{1,m}\colon \mathbb{R} \to [\theta_0, \theta_0+m], \quad x \mapsto \begin{cases} \theta_0+m, & \text{if } f(x,\theta_0)=0,\ f(x,\theta_0+m)>0 \\ \theta_0, & \text{else} \end{cases}$$

are Bayes estimators with respect to π_β, $\beta \in \{0,1\}$. *For* $\beta \in (0,1)$

$$\delta^{\beta,m}\colon \mathbb{R} \to [\theta_0, \theta_0+m],$$
$$x \mapsto \begin{cases} \theta_0, & \text{if } f(x,\theta_0)>0,\ f(x,\theta_0+m)=0 \\ \theta_0 - G_m^{-1}((1-\beta)f(x,\theta_0+m)/(\beta f(x,\theta_0))), & \\ & \text{if } f(x,\theta_0)f(x,\theta_0+m)>0 \\ \beta\theta_0+(1-\beta)(\theta_0+m), & \text{if } f(x,\theta_0)=0,\ f(x,\theta_0+m)=0 \\ \theta_0+m, & \text{if } f(x,\theta_0)=0,\ f(x,\theta_0+m)>0 \end{cases}$$

is a Bayes estimator with respect to π_β, $\beta \in (0,1)$. *The function* $[0,1] \to [\theta_0, \theta_0+m]$, $\beta \mapsto \delta^{\beta,m}$ *is continuous. Moreover,* $\delta^{\beta,m}$ *is strictly decreasing with respect to* $\beta \in [0,1]$ *for* $x \in \{x \in \mathbb{R} \mid f(x,\theta_0)f(x,\theta_0+m) > 0$ *or* $f(x,\theta_0) = 0,\ f(x,\theta_0+m) = 0\}$, *elsewhere* $\delta^{\beta,m}$ *is constant with respect to* $\beta \in [0,1]$.

5. The risk function of the Bayes estimator $\delta^{\beta,m}$

In this section we investigate the properties of the risk function

$$R(\theta, \delta^{\beta,m}) = \int_{\mathbb{R}} h(\theta - \delta^{\beta,m}(x))\, P_\theta(dx).$$

of the Bayes estimator $\delta^{\beta,m}$ determined in Corollary 4.1.

Lemma 5.1. *Let Assumption* A.1–A.4 *be fulfilled. Then, for* $m \in (0,M)$, $\theta \in \Theta$ *the function*

$$[0,1] \to [0,\infty), \quad \beta \mapsto R(\theta, \delta^{\beta,m})$$

is continuous. Furthermore,

$$\lim_{\beta\to 0} R(\theta,\delta^{\beta,m}) = h(\theta-\theta_0-m)P_\theta(\{x \in \mathbb{R} \mid f(x,\theta_0)=0 \text{ or } f(x,\theta_0+m)>0\}) + h(\theta-\theta_0)P_\theta(\{x \in \mathbb{R} \mid f(x,\theta_0)>0, f(x,\theta_0+m)=0\}),$$

and

$$\lim_{\beta\to 1} R(\theta,\delta^{\beta,m}) = h(\theta-\theta_0)P_\theta(\{x \in \mathbb{R} \mid f(x,\theta_0)>0 \text{ or } f(x,\theta_0+m)=0\}) + h(\theta-\theta_0-m)P_\theta(\{x \in \mathbb{R} \mid f(x,\theta_0)=0, f(x,\theta_0+m)>0\}).$$

Thus for the special parameter $\theta = \theta_0$ *and* $\theta = \theta_0 + m$ *we have, under the additional assumption* $P_{\theta_0}(\{x \in \mathbb{R} \mid f(x, \theta_0 + m) > 0\}) > 0$, *that*

$$[0,1] \to [0,\infty), \quad \beta \mapsto R(\theta_0, \delta^{\beta,m})$$

is strictly decreasing with

$$\lim_{\beta\to 0} R(\theta_0, \delta^{\beta,m}) = h(-m)P_{\theta_0}(\{x \in \mathbb{R} \mid f(x, \theta_0 + m) > 0\}) = R(\theta_0, \delta^{0,m}),$$
$$\lim_{\beta\to 1} R(\theta_0, \delta^{\beta,m}) = h(0) = 0 = R(\theta_0, \delta^{1,m}),$$

and under the additional assumption $P_{\theta_0+m}(\{x \in \mathbb{R} \mid f(x, \theta_0) > 0\}) > 0$, *that*

$$[0,1] \to [0,\infty), \quad \beta \mapsto R(\theta_0 + m, \delta^{\beta,m})$$

is strictly increasing with

$$\lim_{\beta\to 0} R(\theta_0 + m, \delta^{\beta,m}) = h(0) = 0 = R(\theta_0 + m, \delta^{0,m}),$$
$$\begin{aligned}\lim_{\beta\to 1} R(\theta_0 + m, \delta^{\beta,m}) &= h(m)P_{\theta_0+m}(\{x \in \mathbb{R} \mid f(x, \theta_0) > 0\}) \\ &= R(\theta_0 + m, \delta^{1,m}).\end{aligned}$$

Proof. We have

$$\begin{aligned}R(\theta, \delta^{\beta,m}) = &\int_{\{f(x,\theta_0)f(x,\theta_0+m)>0\}} h(\theta - \delta^{\beta,m}(x))\, P_\theta(dx) + \int_{\{f(x,\theta_0)>0, f(x,\theta_0+m)=0\}} h(\theta - \theta_0)\, P_\theta(dx) \\ &+ \int_{\{f(x,\theta_0)=0, f(x,\theta_0+m)>0\}} h(\theta - \theta_0 - m)\, P_\theta(dx) \\ &+ \int_{\{f(x,\theta_0)=0, f(x,\theta_0+m)=0\}} h(\theta - \beta\theta_0 - (1-\beta)(\theta_0 + m))\, P_\theta(dx).\end{aligned}$$

Therefore, $[0,1] \to [0,\infty)$, $\beta \mapsto R(\theta, \delta^{\beta,m})$ is continuous, because $[0,1] \to [\theta_0, \theta_0 + m]$, $\beta \mapsto \delta^{\beta,m}$ is continuous by Corollary 4.1 and $\theta_0 \le \delta^{\beta,m}(x) \le \theta_0 + m$ by Lebegue's theorem on dominated convergence. The other assertions follow easily. □

The above lemma immediately implies the following

Corollary 5.1. *Let Assumptions* A.1–A.4 *be fulfilled, and let* $P_\theta(\{x \in \mathbb{R} \mid f(x, \theta_0)f(x, \theta_0 + m) > 0\}) > 0$ *for* $\theta \in \{\theta_0, \theta_0 + m\}$, *then for each* $m \in (0, M)$ *there exists a unique* $\beta^*(m) \in (0,1)$ *such that*

$$R(\theta_0, \delta_m) = R(\theta_0 + m, \delta_m),$$

where $\delta_m = \delta^{\beta^*(m),m}$.

Next we show that our assumptions given in Theorem 3.1 imply that the risk function is convex.

Lemma 5.2. *Let Assumptions* A.1–A.7 *be fulfilled. Then a constant* $m_0 > 0$ *exists such that for all* $m \in (0, m_0]$ *the risk function* $R(\theta, \delta_m)$, $\theta \in [\theta_0, \theta_0+m]$, *is convex.*

Proof. We define

$$G(x,\theta) := h(\theta - \delta_m(x))f(x,\theta).$$

Obviously, $G(\cdot,\theta)$ is integrable for each $\theta \in \Theta$. Next we distinguish the following three cases.

(a) Let Assumptions A.1–A.3 (with A.3 on both sides of the origin) and A.4, A.5 be fulfilled. Then for each $x \in \mathbb{R}$ we have for $\theta \notin N_x := \{\theta \in \Theta \mid \theta = \delta_m(x)$ or $\partial^2 f(x,\theta)/\partial\theta^2$ does not exist$\}$:

$$\frac{\partial^2}{\partial\theta^2}G(x,\theta) = h''(\theta - \delta_m(x))f(x,\theta) + 2h'(\theta - \delta_m(x))\frac{\partial}{\partial\theta}f(x,\theta) + h(\theta - \delta_m(x))\frac{\partial^2}{\partial\theta^2}f(x,\theta).$$

(b) Let Assumption A.1, A.2, A.4 (with A.4 on both sides of the origin) and A.5–A.7 be fulfilled. Then for each $x \in \mathbb{R}$ we have for $\theta \notin N_x := \{\theta \in \Theta \mid \theta = \delta_m(x)$ or $\partial^2 f(x,\theta)/\partial\theta^2$ does not exist$\}$:

$$\frac{\partial^2}{\partial\theta^2}G(x,\theta) \geq |h'(\theta - \delta_m(x))|f(x,\theta) \times \left(\frac{h''(\theta-\delta_m(x))}{|h'(\theta-\delta_m(x))|} - 2\frac{|\frac{\partial}{\partial\theta}f(x,\theta)|}{f(x,\theta)} - \frac{h(\theta-\delta_m(x))}{|h'(\theta-\delta_m(x))|}\frac{|\frac{\partial^2}{\partial\theta^2}f(x,\theta)|}{f(x,\theta)}\right).$$

(c) Let Assumption A.1–A.7 (with A.3 on one and A.4 on the other side of the origin) be fulfilled. Then

$$\frac{\partial^2}{\partial\theta^2}G(x,\theta) \geq 0$$

can be proved by showing $\partial^2 G(x,\theta)/\partial\theta^2 \geq 0$ for θ with $\theta - \delta_m(x) > 0$ and θ with $\theta - \delta_m(x) < 0$ separately. These two inequalities can be shown according to the proofs of (a) and (b), respectively.

Then by our assumptions in each of the three cases (a), (b), (c) a constant m_0 exists which is independent of x such that $\partial^2 G(x,\theta)/\partial\theta^2 > 0$ up to countably many $\theta \in [\theta_0, \theta_0 + m_0]$. Hence, $G(x,\cdot)$ is convex for each $x \in \mathbb{R}$ implying that

$$R(\theta,\delta_m) = \int h(\theta - \delta_m(x))f(x,\theta)\,\lambda(dx)$$

is convex. □

Proof of Theorem 3.1. For the Bayes estimator $\delta^{\beta,m}$ with respect to π_β the assumptions of Corollary 2.1 are fulfilled by Corollary 5.1 and Lemma 5.2. Hence the assertion of Theorem 3.1 holds. □

Acknowledgements. We like to thank Professor Marc Moore for improving the language of our paper.

REFERENCES

Berry, J.C. (1989). Bayes estimation of a Bernoulli p in a restricted parameter space. *Comm. Statist. Theory Models* 18, 4606–4616.

Berry, J.C. (1990). Minimax estimation of a bounded normal mean vector. *J. Multivariate Anal.* 35, 130–139.

Bischoff, W. (1992). Minimax and Γ-minimax estimation for functions of the bounded parameter of a scale parameter family under L_p-loss. *Statist. Decisions* 10, 45–61.

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.

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

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

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

Boratyńska, A. (2001). Two point priors and minimax estimation of a bounded parameter. (Preprint)

Brown, L. (1994). Minimaxity, more or less. In *Statistical Decision Theory and Related Topics*, V. (S.S. Gupta and J.O. Berger, eds.) pp. 1–18. Springer, New York.

Casella, G. and Strawderman, W.E. (1981). Estimating a bounded normal mean. *Ann. Statist.* 9, 868–876.

Chen, L.X. and Eichenauer, J. (1988). Two point priors and Γ-minimax estimating in families of uniform distributions. *Statist. Hefte* 29, 45–57.

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

Eichenauer, J. (1986). Least favourable two point priors in estimating the bounded location parameter of a noncentral exponential distribution. *Statist. Decisions* 4, 389–392.

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

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

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

Karlin, S. (1957). Pólya type distributions. II. *Ann. Math. Statist.* 28, 281–309.

Lehmann, E. and Casella, G. (1998). *Theory of Point Estimation* (Second edition). Springer, New York.

Marchand, É. and MacGibbon, B. (2000). Minimax estimation of a constrained binomial proportion. *Statist. Decisions* 18, 129–167.

Moors, J.J.A. (1985). *Estimation in Truncated Parameter Spaces.* PhD thesis, Tilburg University, Tilburg, The Netherlands.

van Eeden, C. (1956). Maximum likelihood estimation of ordered probabilities. *Nederl. Akad. Wetensch. Proc. Ser. A.* 59, 444–455.

van Eeden, C. (1957a). A least squares inequality for maximum likelihood estimates of ordered parameters. *Nederl. Akad. Wetensch. Proc. Ser. A* 60, 513–521.

van Eeden, C. (1957b). Maximum likelihood estimation of partially or completely ordered parameters. I. *Nederl. Akad. Wetensch. Proc. Ser. A*, 60, 128–136; II. 201–211.

van Eeden, C. (1957c). Note on two methods for estimating ordered parameters of probability distributions. *Nederl. Akad. Wetensch. Proc. Ser. A*, *60*, 506–512.

van Eeden, C. (1958). *Testing and estimating ordered parameters of probability distributions.* Ph.D., University of Amsterdam.

van Eeden, C. (1996). Estimation in restricted parameter spaces—some history and some recent developments. *CWI Quaterly* 9, 69–76.

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

Wan, A.T.K., Zou, G. and Lee, A.H. (2000). Minimax and Γ-minimax estimation for the Poisson distribution under LINEX loss when the parameter space is restricted. *Statist. Probab. Lett.* 50, 23–32.

Zinzius, E. (1979). *Beiträge zur Theorie der nichtsequentiellen Parameterschätzprobleme.* Ph.D., Universität Karlsruhe.

Zinzius, E. (1981). Minimaxschätzer für den Mittelwert θ einer normalverteilten Zufallsgröße mit bekannter Varianz bei vorgegebener oberer und unterer Schranke für θ. *Math. Operationsforsch. Statist. Ser. Statist.* 12, 551–557.

Guido Bader
Institut für Mathematische Stochastik
Universität Karlsruhe
Englerstrasse 2
D-76128 Karlsruhe
Germany
Guido.Bader@math.uni-karlsruhe.de

Wolfgang Bischoff
Institut für Mathematische Stochastik
Universität Karlsruhe
Englerstrasse 2
D-76128 Karlsruhe
Germany
Wolfgang.Bischoff@math.uni-karlsruhe.de