REVSTAT – Statistical Journal
Volume 7, Number 3, November 2009, 203–226

COMPARISONS OF THE PERFORMANCES OF ESTIMATORS OF A BOUNDED NORMAL MEAN UNDER SQUARED-ERROR LOSS

Authors: Yiping Dou
– Department of Statistics, University of British Columbia, Canada
ydou@stat.ubc.ca

Constance van Eeden
– Department of Statistics, University of British Columbia, Canada
cve@xs4all.nl

Received: January 2009 Revised: May 2009 Accepted: June 2009

Abstract:

- This paper is concerned with the estimation under squared-error loss of a normal mean θ based on $X \sim \mathcal{N}(\theta, 1)$ when $|\theta| \leq m$ for a known $m > 0$. Nine estimators are compared, namely the maximum likelihood estimator (mle), three dominators of the mle obtained from Moors, from Charras and from Charras and van Eeden, two minimax estimators from Casella and Strawderman, a Bayes estimator of Marchand and Perron, the Pitman estimator and Bickel's asymptotically-minimax estimator. The comparisons are based on analytical as well as on graphical results concerning their risk functions. In particular, we comment on their gain in accuracy from using the restriction, as well as on their robustness with respect to misspecification of m.

Key-Words:

- *admissibility; Bayes estimator; bounded Normal mean; restricted estimators; robustness; squared-error loss.*

AMS Subject Classification:

- 62F10, 62F30, 62F15.

1. INTRODUCTION

The problem considered in this paper is the estimation under squared-error loss of a normal mean θ based on $X \sim \mathcal{N}(\theta, 1)$ when $|\theta| \le m$ for a known $m > 0$.

This estimation problem is considered by Casella and Strawderman (1981), by Marchand and Perron (2001), by Bickel (1981) and by Gatsonis, MacGibbon and Strawderman (1987). Casella and Strawderman show that, when $0 < m \le m_o \approx 1.056742$, there exists a unique minimax estimator of θ with respect to a symmetric two-point least-favourable prior on $\{-m, m\}$. They give an explicit expression for it and show that it dominates the maximum likelihood estimator (mle) when $m \le 1$. They also give a class of minimax estimators for the case where $1.4 \le m \le 1.6$. These estimators are Bayes with respect to a symmetric three-point prior on $\{-m, 0, m\}$. Bickel gives an estimator which is asymptotically minimax as $m \to \infty$ and Gatsonis, MacGibbon and Strawderman graphically compare these estimators and the Pitman estimator for several values of m. Marchand and Perron consider the problem of estimating θ when $X \sim \mathcal{N}_k(\theta, I)$ with $\|\theta\| \le m$ and give conditions on m, k and the prior for Bayes estimators to dominate the mle. An example of their results is that the Bayes estimator with respect to the uniform prior on the boundary of the parameter space dominates the mle when $m \le \sqrt{k}$, generalizing the Casella–Strawderman result to $k > 1$. Dominators for the mle can also be obtained from results of Charras (1979), of Moors (1981, 1985) and of Charras and van Eeden (1991). These authors consider estimation in restricted parameter spaces in a very general setting, give conditions for inadmissibility for squared-error loss and either give methods of constructing dominators (Moors and Charras and van Eeden) or prove the existence of dominators within a given class of estimators (Charras). Their conditions are satisfied for the bounded-normal-mean problem and one of the purposes of this paper is to find explicit expressions for these dominators and compare their risk functions, analytically as well as graphically, with those of the mle, the Casella–Strawderman minimax estimators, Bickel's asymptotically minimax estimator, the Pitman estimator and one of the Marchand–Perron Bayes estimators. In these comparisons, questions of an estimator's gain in accuracy obtained from using the restriction are looked at, as well as how this gain depends on m and how robust the estimators are with respect to misspecification of m.

One of our analytical results shows that, if and only if $m \le 1$, Moors' dominator of the mle of a bounded normal mean is the Casella–Strawderman minimax estimator, implying (by the Casella–Strawderman result for $m \le 1$) that this Moors dominator of the mle is admissible when $m \le 1$. Another analytical result we have is that the dominators in the Charras–van Eeden class are all inadmissible. We also show, again analytically, that the estimator $\delta_o(x) \equiv 0$ (which we call the "trivial estimator") dominates the mle if and only if $0 < m \le m_1 \approx 0.5204372$. Marchand and Perron (2001) show, as a special case of their results for a $k \geq 1$-dimensional restricted normal mean, that (when $k=1$) every symmetric estimator dominates the mle when $m \leq m_o \approx .4837$. Finally we find, numerically, that the Marchand–Perron Bayes estimator considered by us has a risk function which is, on the whole interval $[-m, m]$, very close to that of one of the Casella–Strawderman minimax estimators.

Explicit expressions for the estimators are presented in Section 2. Our numerical comparisons are presented in the form of graphs and discussed in Section 3. The proofs of the lemmas and theorems are given in Appendix A.

We know of only one other family of distributions for which Charras's (1979) and Moors' (1981, 1985) dominators have been obtained and compared. These results can be found in Perron (2003). He compares the mle with its Charras and its Moors dominators, as well as with the Pitman estimator and the Bayes estimator with respect to a prior proportional to $\big(p(1-p)\big)^{-1}$ for the case where $X \sim \text{Bin}(n,p)$ when $p \in [a, 1-a]$ for a given $a \in (0, 1/2)$. He gives an algorithm for finding the Charras dominator.

2. THE ESTIMATORS

The problem of estimating a bounded normal mean based on $X \sim \mathcal{N}(\theta, 1)$ is a special case of the following problem: $(\mathcal{X}, \mathcal{A})$ is a measurable space and $\mathcal{P} = \{P_\theta, \theta \in D\}$ is a family of probability measures on $(\mathcal{X}, \mathcal{A})$ where $D \subset \mathbb{R}^k$ is a subset of the set of θ for which P_θ is such a probability measure. Further, D is convex and closed. The problem is to find, for a given loss function, "good" estimators of θ based on a random vector $X \in \mathbb{R}^n$ defined on $(\mathcal{X}, \mathcal{A})$, where $\delta(X)$ is an estimator of θ if it satisfies $P_\theta\big(\delta(X) \in D\big) = 1$ for all $\theta \in D$. Many analytical results concerning admissibility and minimaxity for such models have, for various loss functions, been obtained (see e.g. van Eeden (2006)).

The present section contains explicit expressions for each of the estimators of a bounded normal mean considered in this paper. It also contains their known, as well as our, analytical properties.

The maximum likelihood estimator

The mle of θ for our problem of estimating a bounded normal mean is given by

$$\delta^{\text{mle}}(X) = \begin{cases} -m & \text{if } X \leq -m \\ X & \text{if } -m < X < m \\ m & \text{if } X \geq m\,. \end{cases}$$

It is well-known that this estimator is inadmissible for our problem.

Casella and Strawderman's minimax estimators

Casella and Strawderman (1981) give conditions for a Bayes estimator to be minimax for estimating a bounded normal mean based on $X \sim \mathcal{N}(\theta, 1)$ with squared-error loss. They show that a two-point symmetric prior on $\{-m, m\}$ is least favourable if $m \leq m_0 \approx 1.056742$, implying that the corresponding Bayes estimator is minimax. This m_0 is the solution of $R(0, \delta^{\text{cs.2}}) - R(m, \delta^{\text{cs.2}}) = 0$, where $\delta^{\text{cs.2}}$ is the Bayes estimator which is given by

$$\delta^{\text{cs.2}}(X) \;=\; m \tanh(mX)\ .$$

The authors show that this minimax estimator dominates the mle when $m \leq 1$. They also give a class of minimax estimators for symmetric three-point priors as follows: for a three-point prior $\pi(0) = \alpha$ and $\pi(-m) = \pi(m) = (1-\alpha)/2$, the Bayes estimator under squared-error loss is given by

$$\delta^{\text{cs.3}}(X) \;=\; \frac{(1-\alpha)\; m \tanh(mX)}{1-\alpha+\alpha \exp(m^2/2)/\cosh(mX)}\ . \tag{2.1}$$

Casella and Strawderman show that, if α and m satisfy

$$(m^2-1)\left(m^2-1+\exp(m^2/2)\right)^{-1} \;\leq\; \alpha \;\leq\; 2\left(2+\exp(m^2/2)\right)^{-1}, \tag{2.2}$$

and α is such that $R(0, \delta^{\text{cs.3}}) - R(m, \delta^{\text{cs.3}}) = 0$, then $\delta^{\text{cs.3}}$ is minimax for estimating θ when $|\theta| \leq m$. They find numerically that these two conditions are satisfied when $1.4 \leq m \leq 1.6$.

A Bayes estimator of Marchand and Perron

Marchand and Perron (2001) consider the estimation, for squared-error loss, of θ based on $X \sim \mathcal{N}_k(\theta, I)$ when $\|\theta\| \leq m$ for a known $m > 0$. The purpose of their paper is finding dominators of the mle. One of their classes of dominators consists of Bayes estimators with respect priors of the form $\pi(\theta) = K e^{-h(\|\theta\|^2)}$, where K is a normalizing constant. Their Corollary 4 gives sufficient conditions on the triple (m, h, k) for the resulting Bayes estimator to dominate the mle. For our case where $k=1$, taking $m > \sqrt{k} = 1$, $h(\theta^2) = -a\,\theta^2/2$ and a the unique solution to (see their Example 3 on the lines 4–16 of their page 1088)

$$\int_0^m t^2\, e^{\frac{a-1}{2}t^2}\, dt \;=\; \int_0^m e^{\frac{a-1}{2}t^2}\, dt \tag{2.3}$$

assures that the first and second conditions of their Corollary 4 are satisfied.

But, it shows that the third condition of Corollary 4 is not satisfied for this triple (m, h, k) by using Corollary 4 on the authors' page 1090 together with the fact (see their Table 1 and the Remark 2 on their page 1088) that $\Delta_E(p)$ is empty when $k = 1$. So, this Bayes estimator might not dominate the mle. In order to get some insight into this question of domination, we compare (in Section 3) this Bayes estimator with a satisfying (2.3), with our other estimators for $m = 1.5$, as well as for $m = 1.8$.

Marchand and Perron (2001) show that (2.3) has a unique solution and that the corresponding Bayes estimator is given by

$$\delta^{\mathrm{E}}(X) = \frac{X}{|X|} \frac{\displaystyle\int_0^m t^{\frac{3}{2}}\, I_{1/2}(t|X|)\, e^{\frac{(a-1)t^2}{2}}\, dt}{\displaystyle\int_0^m t^{\frac{1}{2}}\, I_{-1/2}(t|X|)\, e^{\frac{(a-1)t^2}{2}}\, dt}\,, \tag{2.4}$$

where $I_\nu(t)$ is the modified Bessel function of order ν (see e.g. Robert (1990)).

The following theorem gives an alternate expression for the estimator valid for the case when $a \in (0, 1)$. The theorem also gives an equality which is equivalent to, and easier to solve than, (2.3) when $a \in (0, 1)$.

Theorem 2.1. *When $a \in (0, 1)$, an alternate expression for the estimator is*

$$\delta^{\mathrm{E}}(X) = \frac{X}{|X|} \frac{1}{\sqrt{1-a}} \frac{\displaystyle\int_0^{m\sqrt{1-a}} u \sinh\Big(u \frac{|X|}{\sqrt{1-a}}\Big)\, e^{-u^2/2}\, du}{\displaystyle\int_0^{m\sqrt{1-a}} \cosh\Big(u \frac{|X|}{\sqrt{1-a}}\Big)\, e^{-u^2/2}\, du}\,. \tag{2.5}$$

Moreover, when $a \in (0, 1)$, (2.3) is equivalent to

$$a\Big(\Phi\big(m\sqrt{1-a}\big) - \frac{1}{2}\Big) = m\sqrt{1-a}\; \phi\big(m\sqrt{1-a}\big)\,, \tag{2.6}$$

where $\Phi(t)$ and $\phi(t)$ are the standard normal distribution function and density function.

Moors' dominating estimator of the mle

Moors (1981, 1985) considers the problem described in the beginning of this section and gives sufficient conditions for "boundary estimators" to be inadmissible for squared-error loss. Here, a boundary estimator is an estimator which takes values on or near the boundary of D with positive probability for some $\theta \in D$. He assumes that the problem is invariant with respect to a finite group

$G = (g_1, ..., g_p)$ of measure preserving transformations from $\mathcal{X}$ to $\mathcal{X}$ and that the induced group $\tilde{G}$ is commutative and satisfies

$$\tilde{g}(ad_1 + bd_2) = a\tilde{g}(d_1) + b\tilde{g}(d_2) \qquad \text{for all} \ \ d_1, d_2 \in D, \quad \text{all} \ \ \tilde{g} \in \tilde{G} \ .$$

He then constructs random, closed, convex subsets D_X of D with the property that an estimator δ for which $P_\theta\big(\delta(X) \notin D_X\big) > 0$ for some $\theta \in D$ is inadmissible. These sets D_X are defined as follows. Let p_θ be the density of P_θ with respect to a σ-finite measure ν defined on $(\mathcal{X}, \mathcal{A})$ and let

$$\alpha\big(X, \bar{g}_j(\theta)\big) = \frac{p_{\bar{g}_j(\theta)}(X)}{S(X;\theta)} \ , \qquad j = 1, ..., p \ ,$$

where $S(X;\theta) = \sum_{j=1}^{p} p_{\bar{g}_j(\theta)}(X) > 0$. Further, he defines

$$h_X(\theta) \ = \ \begin{cases} \displaystyle\sum_{j=1}^{p} \alpha\big(X, \bar{g}_j(\theta)\big)\, \bar{g}_j(\theta) & \text{when} \ \ S(X;\theta) > 0 \\ \theta & \text{when} \ \ S(X;\theta) = 0 \ . \end{cases}$$

Then D_X is the convex closure of the range of $h_X(\theta)$ for $\theta \in D$ and boundary estimators $\delta(X)$, i.e. estimators $\delta(X)$ for which $P_\theta\big(\delta(X) \notin D_X\big) > 0$ for some $\theta \in D$, are inadmissible and are dominated by their projection unto D.

For the problem of estimating a bounded normal mean under squared-error loss, Moors' conditions are satisfied with $p = 2$, $g_1(x) = x$ and $g_2(x) = -x$ which gives $h_X(\theta) = \theta \tanh(\theta X)$, because $p_\theta(x) = \exp\big(-(x-\theta)^2/2\big)/\sqrt{2\pi}$. So the subset D_X is given by

$$D_X = \ \big(-m \tanh(m|X|), \, m \tanh(m|X|)\big) \ ,$$

which implies by Moors that any estimator δ for which

$$P_\theta\Big(\delta(X) \notin \big(-m \tanh(m|X|), \, m \tanh(m|X|)\big)\Big) \ > \ 0 \qquad \text{for some} \ \ \theta \in D$$

is inadmissible and is dominated by its projection unto D. Hence, Moors' dominator of the mle is given by

$$\text{(2.7)} \qquad \delta^{\text{mr}}(X) \ = \begin{cases} -m \tanh(m|X|) & \text{if} \ \ X \leq -m \tanh(m|X|) \\ X & \text{if} \ \ -m \tanh(m|X|) < X < m \tanh(m|X|) \\ m \tanh(m|X|) & \text{if} \ \ X \geq m \tanh(m|X|) \ . \end{cases}$$

The following theorem shows that, for $m \leq 1$, Moors' dominating estimator of the mle is Casella and Strawderman's minimax estimator. We also obtain there a more explicit expression for this dominator for the case where $m > 1$. The proof of the theorem is given in Appendix A.

Theorem 2.2. *Moors' dominator of the mle can also be written as*

(i) *if* $0 < m \leq 1$ *then* $\delta^{\mathrm{mr}}(X) = m \tanh(mX)$;

(ii) *if* $m > 1$, *then*

$$\delta^{\mathrm{mr}}(X) = \begin{cases} m \tanh(mX) & \text{if } X \geq \xi(m) \text{ or } X \leq -\xi(m) \\ X & \text{if } -\xi(m) < X < \xi(m) \ , \end{cases}$$

where $\xi(m)$, $r(m) < \xi(m) < m$, *is the unique root of* $u(x) = x - m \tanh(mx) = 0$ *for* $x > 0$ *and* $r(m) = \frac{1}{m} \ln\big[m + \sqrt{m^2-1}\big]$.

Charras's and Charras and van Eeden's dominators of the mle

Charras (1979) considers the problem as described in the beginning of this section. He gives, for squared-error loss, conditions for boundary estimators to be non-Bayes as well as conditions for them to be inadmissible, where a boundary estimator is, for him, an estimator δ for which $P_\theta\big(\delta(X) \in B\big) > 0$ for some $\theta \in D$ and B is the boundary of D. For the case where $k = 1$ and $\theta \in [a, b]$ for known $-\infty < a < b < \infty$, he gives conditions for the existence of classes of dominators of his boundary estimators and ways to construct them.

The inadmissibility results of Charras (1979) are published in Charras and van Eeden (1991), but his dominators are only mentioned there. Instead, Charras and van Eeden study a different class of dominators (proposed by a referee of this Charras and van Eeden paper) of Charras' boundary estimators. The authors construct, for squared-error loss, a class of dominators δ^{cve} for boundary estimators $\delta(X)$ of θ when $\theta \in [a, b]$ with $-\infty < a < b < \infty$, where they suppose that these boundary estimators δ satisfy

$$\left.\begin{aligned} P_\theta\big(\delta(X) = a\big) &> 0 \\ P_\theta\big(\delta(X) = b\big) &> 0 \end{aligned}\right\} \quad \text{for all} \quad \theta \in [a, b] \ .$$

They further suppose that, for each $\theta_o \in D$,

$$\lim_{\theta \to \theta_o} \int_{\mathcal{X}} |p_\theta - p_{\theta_o}| \, d\nu \;=\; 0 \ , \tag{2.8}$$

where p_θ is the density of P_θ with respect to the σ-finite measure ν.

The authors then show that there exists estimators of the form

$$\delta^{\mathrm{cve}}(X) = \begin{cases} a + \varepsilon_1 & \text{if } \delta(X) \leq a \\ \delta(X) & \text{if } a < \delta(X) < b \\ b - \varepsilon_2 & \text{if } \delta(X) \geq b \end{cases} \tag{2.9}$$

where $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ and $\varepsilon_1 + \varepsilon_2 \leq b - a$, which dominate δ.

This Charras–van Eeden result with $a = -m$ and $b = m$ clearly applies to our problem of dominating the mle of a bounded normal mean, where because of the symmetry of the problem, one can take $0 < \varepsilon_1 = \varepsilon_2 = \varepsilon \leq m$. This gives a class of dominators of the mle of a bounded normal mean for squared-error loss and using the results of Charras and van Eeden (1991) one finds that $\varepsilon \in (0, \varepsilon_o]$, where $\varepsilon_o = \min\bigl(m\bigl(8\,\Phi(-2m)/(1 + 2\phi(-2m))\bigr),\, m\bigr)$, gives a dominator of the mle. However, each of these dominators is inadmissible. This follows from Brown (1986)'s necessary condition for admissibility for squared-error loss in the estimation of the mean of an exponential-family distribution. He shows that an admissible estimator has to be non-decreasing and our estimator $\delta^{\text{cve}}(X)$ is clearly not non-decreasing, while the $\mathcal{N}(\theta, 1)$ is an exponential-family distribution. This inadmissibility result is summarized in the following theorem:

Theorem 2.3. *Let $X \sim \mathcal{N}(\theta, 1)$ with $|\theta| \leq m$ for a known positive m. Then the Charras–van Eeden dominators (2.9) of the mle are inadmissible for squared-error loss.*

We have not been able to find dominators for these inadmissible dominators and so will not consider them any further in this paper.

We now present Charras's (1979) method of obtaining dominators for his boundary estimators and use it to find dominators of the mle in the bounded-normal-mean problem.

Let δ be a Charras boundary estimator, then Charras considers the following class of estimators

$$\delta^{\text{t}}(X) = \begin{cases} a(t) & \text{if } \delta(X) \leq a(t) \\ \delta(X) & \text{if } a(t) < \delta(X) < b(t) \\ b(t) & \text{if } \delta(X) \geq b(t)\ , \end{cases} \tag{2.10}$$

where $a(t)$ and $b(t)$, $t \in [0,1]$, take values in $[a, b]$ with $a(0) = a$, $b(0) = b$, $a(1) = b(1)$, $a(t)$ is non-decreasing and $b(t)$ is non-increasing. He then gives sufficient conditions on the functions $a(t)$ and $b(t)$, on the distribution of X and of $\delta(X)$ and on the loss function, for δ^{t} to dominate δ. These conditions are given in Appendix A. Here we give this domination result for the special case of the bounded normal mean when $a(t) = -m(1-t)$ and $b(t) = m(1-t)$, $t \in [0,1]$. Obviously, Charras's conditions are satisfied in the bounded-normal-mean case and his dominator of the mle can then be written as follows:

$$\delta^{\text{ch}}(X) = \begin{cases} -m(1-t) & \text{if } X \leq -m(1-t) \\ X & \text{if } -m(1-t) < X < m(1-t) \\ m(1-t) & \text{if } X \geq m(1-t)\ . \end{cases}$$

For simplicity of the proof we let $\varepsilon = mt \in [0, m]$ and rewrite this dominator as follows:

$$\delta^{\mathrm{ch}}(X) \;=\; \begin{cases} -(m-\varepsilon) & \text{if } X \le -(m-\varepsilon) \\ X & \text{if } -(m-\varepsilon) < X < m-\varepsilon \\ m-\varepsilon & \text{if } X \ge m-\varepsilon \;. \end{cases}$$

Then the following theorem holds:

Theorem 2.4. *Let $X \sim \mathcal{N}(\theta, 1)$ with $|\theta| \le m$ for a known $m > 0$. Then, for squared-error loss, $\{\delta^{\mathrm{ch}} \colon 0 < \varepsilon \le \varepsilon_m\}$ is a class of dominating estimators of the mle, where ε_m is the unique root to $\psi_m(x) = 0$, with $\psi_m(x) = g(2m - x) + g(x) - 2x$ and $g(x) = 2x\,\Phi(-x)$.*

The proof of this theorem is given in Appendix A. It is Charras's proof applied to our special case.

The trivial estimator

For the estimator $\delta_o(X) \equiv 0$ the following theorem holds. Its proof is in Appendix A.

Theorem 2.5. *Let m_1 be the unique positive solution to $u(2m) + 1/2 - m^2 = 0$, where $u(x) = x^2\,\Phi(-x) - \Phi(-x) - x\,\phi(x)$. Then, for squared-error loss, the estimator δ_o dominates the mle if and only if $0 < m \le m_1 \approx 0.5204372$.*

A related result can be found in Marchand and Perron (2001). They present dominators of the mle of θ when $X \sim \mathcal{N}_k(\theta, I)$ with $k \ge 1$, $\|\theta\| \le m$ and squared-error loss. One of their results says that, when $k = 1$, any symmetric estimator dominates the mle when $m \le m_o \approx .4837$.

The Pitman estimator

The Pitman estimator of θ for our problem is defined as the Bayes estimator with respect to a uniform prior on $[-m, m]$ and squared-error loss. This Bayes estimator is the posterior mean of θ given X. Since the marginal density of X is given by

$$p(X) \;=\; \int_{-m}^{m} p_\theta(X)\,\pi(\theta)\,d\theta \;=\; \frac{1}{2m}\left[\Phi(m-X) - \Phi(-m-X)\right],$$

the posterior density of θ given X is given by

$$\begin{aligned} p(\theta|X) &= \frac{p_\theta(X)\,\pi(\theta)}{p(X)} \\ &= \frac{1}{\Phi(m-X)-\Phi(-m-X)}\,\frac{1}{\sqrt{2\pi}}\ \exp\left\{-\frac{(\theta-X)^2}{2}\right\}\,1_{\{|\theta|\le m\}}\ . \end{aligned}$$

Hence the Pitman estimator of θ is given by

$$\begin{aligned} \delta^{\mathrm{p}}(X) \;=\; E(\theta|X) &= X + \frac{\displaystyle\int_{-m-X}^{m-X} z\,\phi(z)\,dz}{\Phi(m-X)-\Phi(-m-X)} \\ &= X - \frac{\phi(m-X)-\phi(m+X)}{\Phi(m-X)-\Phi(-m-X)}\ . \end{aligned}$$

Bickel's asymptotically minimax estimator

Bickel (1981) constructs, for squared-error loss, a class of asymptotically minimax estimators for estimating a bounded normal mean. He constructs this class in the following way:

Let, for $|x|<1$, $\bar\psi = \pi\tan(\frac{\pi}{2}x)$ and let

$$\psi_m(x) \;=\; \begin{cases} \bar\psi(x) & \text{if } |x|\le 1-a_m^2 \\ \Big(\bar\psi(1-a_m^2)+\bar\psi'(1-a_m^2)\big(x^2-(1-a_m^2)\big)\Big)\,\mathrm{sgn}\,x & \text{if } |x|>1-a_m^2\ . \end{cases}$$

He then shows that an asymptotically minimax estimator δ^{b} is given by

$$\delta^{\mathrm{b}}(X) \;=\; X-\frac{1}{n}\,\psi_m\!\left(\frac{X}{n}\right),$$

where $n=m(1-a_m)^{-1}$, $a_m<1$ and $ma_m\to\infty$ as $m\to\infty$. Bickel (1981) suggests taking $a_m=m^{\frac{1}{8}}$ which gives the following expression for $\psi_m(x)$:

$$\begin{cases} \pi\ \tan\!\left(\dfrac{\pi}{2}\,x\right) & \text{if } |x|\le g(m) \\ \left[\pi\ \tan\!\left(\dfrac{\pi}{2}\,g(m)\right)+\dfrac{\pi^2}{2}\ \sec^2\!\left(\dfrac{\pi}{2}\,g(m)\,\big(x-g(m)\big)\right)\right]\mathrm{sgn}\,x & \text{if } |x|>g(m)\ , \end{cases}$$

where $g(m)=1-m^{-1/4}$.

Then Bickel's asymptotically minimax estimator of θ is given by

$$\delta^{\mathrm{b}}(X) \;=\; X-\frac{1-m^{\frac{1}{8}}}{m}\ \psi_m\!\left(\frac{(1-m^{-\frac{1}{8}})\,X}{m}\right)$$

and the minimax value is given by $1-\frac{\pi^2}{m^2}+o(m^2)$ as $m\to\infty$.

3. NUMERICAL COMPARISONS

Appendix B contains graphs of the risk functions for squared-error loss of the estimators $\delta^{\rm mle}$, $\delta^{\rm cs.2}$, $\delta^{\rm cs.3}$, $\delta^{\rm E}$, $\delta^{\rm mr}$, $\delta^{\rm ch}$, $\delta^{\rm p}$ and $\delta^{\rm b}$ for several values of m. For the estimator $\delta^{\rm cs.3}$ the value $\alpha = .341$ (see (2.1)) is used, while for $\delta^{\rm ch}$, $\varepsilon = \varepsilon_m$ (see Theorem 2.4) is used. For the estimator $\delta^{\rm E}$ a value of $m > 1$ is needed.

Because of their symmetries, the risk functions are plotted only on the positive part of the parameter space. Moreover to check the robustness of the estimators with respect to misspecification of m, the risk functions are plotted on a somewhat wider interval, namely the interval $[0, 5m/4]$.

Figure 1 gives the risk functions for $m = .5$, .8, 1.0 and 1.5, while Figure 2 gives them for $m = 1.8$, 3, 5 and 10. The values of ε_m for these m (needed for the estimator $\delta^{\rm ch}$) are given in Table 1, while the values of a needed for $\delta^{\rm E}$ are given in Table 2.

Table 1: Values of ε_m for $\delta^{\rm ch}$.

m	.5	.8	1.0	1.5	1.8	≥ 3
ε_m	.276	.195	.101	.008	.001	.000

Table 2: Values of a for $\delta^{\rm E}$.

m	1.5	1.8	3	5	10
a	2.02	0.82	0.03	1.48×10^{-5}	6.91×10^{-17}

From the graphs one sees that

1. For $m = .5$, .8 and 1, $\delta^{\rm mr}$ has the same risk function as Casella and Strawderman's minimax estimator $\delta^{\rm cs.2}$. This is in accordance with our Theorem 2.2 which says that, for $m \leq 1$, these estimators are the same estimator.
2. Our Theorem 2.4 says that $\delta^{\rm ch}$ dominates $\delta^{\rm mle}$. This is clearly visible in the graphs for $m = .5$, .8 and 1. For larger m there is little difference between the risk functions of these two estimators and, in fact, little difference between the risk functions of $\delta^{\rm mle}$, $\delta^{\rm mr}$ and $\delta^{\rm ch}$, verifying the intuitively obvious result that, as $m \to \infty$, the differences

between these estimators converge almost surely to zero. This asymptotic result also holds for the estimators δ^{mle}, δ^{p} and δ^{b}, but for these estimators it takes a larger m for the risk functions to be close.

3. From the graph for $m = 1.5$ it is seen that, risk-functionwise, there is very little difference between δ^{E} and $\delta^{\text{cs.3}}$. But δ^{E} is computationally more complicated — two numerical integrations are needed to find a and two more to compute the estimator, while $\delta^{\text{cs.3}}$ is easily computable from (2.1). For $m = 1.8$, no minimax estimator is available, but for this value of m, δ^{E} behaves relative to δ^{P}, as it does for $m = 1.5$ — better for the smaller values of $|\theta|$, worse for values of $|\theta|$ closer to m with a fairly constant risk function. But the computational problems with this estimator relative to the others might, for a user, well be the determining factor concerning the question of which estimator to use.

4. In each of the graphs for $m \geq 3$, each of δ^{mle}, δ^{mr}, δ^{ch} and δ^{p} are close to being minimax with a minimax value ≈ 1. This is another example of the above-mentioned asymptotic result because, for the unrestricted case, the minimax value equals 1. Further, from the graphs for $m \leq 3$ one sees that, for these estimators, the maximum value of the risk function increases with m.

5. If one does not use the information that $|\theta| \leq m$, the best estimator of θ is X. Its risk function (for squared-error loss) is constant and equal to 1. From the graphs one can observe the gain in accuracy with which one can estimate θ when the restriction on θ is used in the construction of the estimator. One also sees that this gain (of course) decreases as m increases. For $m = .5$, e.g., one can get a minimum gain (over Θ) of about 80.1%, for $m = .8$ this is about 62.6%, for $m = 1$ about 55.0%, for $m = 1.5$ about 42.4%, for $m = 1.8$ about 28.7%, and for $m = 3$ about 3.9%. For the other values of m, this minimum gain is about 0 for all the restricted estimators except the Bickle's asymptotic estimator δ^{b}. The risk function of Bickle's estimator is, for large m, parallel to the one for the unrestricted estimator, X, under the squared-error loss. For $m = 5$, the minimum gain is about 39.5% and for $m = 10$ about 9.9% for δ^{b}. So, it is "worth the trouble" to use the information that $|\theta| \leq m$ at least for values of m that are not too large. Of course this increase in accuracy also occurs in other restricted-parameter-space models, but there are not many cases where numerical results about the gain in accuracy have been obtained (see e.g. van Eeden (2006), Chapter 7, which also contains robustness results for models other than the present one).

6. The graph for $m = .5$ gives an example of our Theorem 2.5, where it is shown that, for $m \leq m_1 \approx 0.5204372$, the trivial estimator dominates δ^{mle}: in the graph the risk function of δ^{mle} is, on the whole interval $[-m, m]$, $>$ than θ^2.

7. In each of the graphs we see that δ^{p} dominates each of the other estimators, except δ^{b}, on the middle part of the parameter space, but not near the endpoints.

8. For the estimator δ^{b}, the graph of its risk function is given for $m = 5$ and for $m = 10$. For those m, it dominates all the other estimators on the middle part of the parameter space, but not near the endpoints.

9. Graphs of the risk functions of δ^{mle}, $\delta^{\mathrm{cs.2}}$, δ^{p} and $\delta^{\mathrm{cs.3}}$ for $m = .5$, .75, 1.5 and 2 can be found in Gatsonis, MacGibbon and Strawderman (1987).

10. The robustness of the domination results with respect to miss-specification of the parameter space can be observed by studying the behaviour of the risk functions for values of θ in the neighbourhood of the value of m used to construct the estimators. For δ^{mle}, δ^{mr} and δ^{ch}, for instance, one sees that, for those m for which the risk functions are not too close (i.e. for $m = .5$, .8 and 1), the domination results hold on a small interval outside the parameter space.

11. The graphs for $m = .5$, .8 and 1 seem to indicate that δ^{mr} dominates δ^{ch}. We do not know whether this holds in general.

APPENDIX

A. PROOFS OF THE RESULTS IN SECTION 2

In this section proofs are given for the results in Section 2.

A.1. Proofs for the Marchand–Perron estimator

Proof of Theorem 2.1: From Berry (1990) we have, for $\nu \geq 0$,

$$I_\nu(t) \;=\; \left(\frac{t}{2}\right)^{\nu} \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\,t\right)^{2k}}{k!\;\Gamma(\nu+k+1)}$$

and

$$I_{1/2}(t) \;=\; \sqrt{\frac{2t}{\pi}}\;\frac{\sinh(t)}{t}\;. \tag{A.1}$$

Further,

$$I_{3/2}(t) = I_{-1/2}(t) - \frac{1}{t}\, I_{1/2}(t) \qquad \text{and} \qquad I_{3/2}(t) = -\sqrt{\frac{2t}{\pi}}\,\frac{\sinh(t)}{t^2} + \sqrt{\frac{2t}{\pi}}\,\frac{\cosh(t)}{t}$$

give

$$I_{-1/2}(t) \;=\; I_{3/2}(t) + \frac{1}{t}\, I_{1/2}(t) \;=\; \sqrt{\frac{2t}{\pi}}\,\frac{\cosh(t)}{t}\;. \tag{A.2}$$

Then, using (A.1) and (A.2) and putting, when $a \in (0,1)$, $t = u(\sqrt{1-a})^{-1}$ in (2.4), gives (2.5).

For the proof of (2.6), note that, when $a \in (0,1)$, the right-hand side of (2.3) can be written as

$$\sqrt{\frac{2\pi}{1-a}}\,\left\{\Phi\big(m\sqrt{1-a}\big) - \frac{1}{2}\right\} \tag{A.3}$$

and the left-hand side of (2.3) as

$$\frac{\sqrt{2\pi}}{(1-a)^{3/2}}\,\left\{\Phi\big(m\sqrt{1-a}\big) - \frac{1}{2} - m\sqrt{1-a}\;\phi\big(m\sqrt{1-a}\big)\right\}\;. \tag{A.4}$$

The result then follows from (A.3) and (A.4). □

A.2. Proofs for Moors' dominator δ^{mr}

The following lemma is needed for the proof of Theorem 2.2.

Lemma A.1. *Let $u(x) = x - m\tanh(mx)$ and $v(x) = x + m\tanh(mx)$. Then*

- **(a)** *For $0 < m \le 1$, $u(x)$ and $v(x)$ are increasing in x and have the same sign as x.*
- **(b)** *For $m > 1$, let $r(m) = \frac{1}{m}\ln\big[m + \sqrt{m^2-1}\,\big]$.*
  Then:
  - **(i)** *$u(x)$ increases in x for $x > r(m)$ and for $x < -r(m)$. It decreases for $-r(m) < x < r(m)$.*
  - **(ii)** *$0 < r(m) < m$.*
  - **(iii)** *$u\big(r(m)\big) < 0$.*
  - **(iv)** *There exists a unique $\xi(m)$, $r(m) < \xi(m) < m$ such that $u\big(-\xi(m)\big) = u\big(\xi(m)\big) = 0$.*

Proof:

(a) For $0 < m \le 1$, since $u'(x) = 1 - m^2\,\text{sech}^2(mx)$, we have

$$u'(x) > 0 \;\Leftrightarrow\; \exp(2mx) - 2m\exp(mx) + 1 = (e^{mx} - m)^2 + 1 - m^2 > 0\,.$$

Consequently, when $0 < m < 1$, $u'(x) > 0$ for $x \in (-\infty, \infty)$ and, when $m = 1$, $u'(x) > 0$ for $x \neq 0$ and $u'(0) = 0$. So, $u(x)$ increases in x and $u(x)$ has the same sign as x because $u(0) = 0$. Since $v'(x) = 1 + m^2\,\text{sech}^2(mx) > 0$ for $x \in (-\infty, \infty)$, we have $v(x)$ increases in x and $v(x)$ has the same sign as x because $v(0) = 0$.

(b) **(i)** Since $u'(x) = 1 - m^2\,\text{sech}^2(mx)$, we have

$$\begin{aligned} u'(x) > 0 \;&\Leftrightarrow\; \big|\exp(mx) - m\big| \;>\; \sqrt{m^2 - 1} \\ &\Leftrightarrow\; \begin{cases} x > r(m) > 0 & \text{if } \exp(mx) > m \\ x < -r(m) < 0 & \text{if } \exp(mx) < m\,. \end{cases} \end{aligned}$$

So, $u(x)$ increases in x when $x > r(m)$ and when $x < -r(m)$. It decreases in x when $-r(m) < x < r(m)$.

(ii) Let $p(x) = x - \ln\big[x + \sqrt{x^2 - 1}\,\big]/x$ for $x > 1$. Then $p(m) = m - r(m)$ for $m > 1$. Further, note that

$$p(x) \;=\; \frac{1}{x}\,\ln\!\left(\frac{\exp(x^2)}{x + \sqrt{x^2 - 1}}\right) \;>\; \frac{1}{x}\,\ln\!\left(\frac{\exp(x^2)}{2x}\right) \;>\; 0\,. \qquad \text{(A.5)}$$

Since $x > 1$, $x + \sqrt{x^2 - 1} < 2x$. So the first inequality in (A.5) holds. Let $q(x) = \exp(x^2) - 2x$. Because $q(1) = e - 2 > 0$ and $q'(x) = 2\big(x\exp(x^2) - 1\big) > 0$ for $x > 1$, we have $q(x) > 0$, which shows that the second inequality in (A.5) also holds for $x > 1$. Hence, $p(x) > 0$ for $x > 1$ and so $0 < r(m) < m$ for $m > 1$.

(iii) Since

$$\begin{aligned} u\big(r(m)\big) \;&=\; r(m) - m\tanh\big(m\,r(m)\big) \\ &=\; \frac{1}{m}\,\ln\big(m + \sqrt{m^2 - 1}\big) - \sqrt{m^2 - 1}\,, \end{aligned}$$

we have

$$\begin{aligned} u\big(r(m)\big) < 0 \;&\Leftrightarrow\; m\sqrt{m^2 - 1} \;>\; \ln\big(m + \sqrt{m^2 - 1}\big) \\ &\Leftrightarrow\; f(m) > 0\,, \end{aligned}$$

where $f(x) = x\sqrt{x^2 - 1} - \ln\big(x + \sqrt{x^2 - 1}\big)$ for $x > 1$. Since $f(1) = 0$ and

$$\begin{aligned} f'(x) \;&=\; \sqrt{x^2 - 1} + \frac{x^2}{\sqrt{x^2 - 1}} - \frac{1}{x + \sqrt{x^2 - 1}}\left(1 + \frac{x}{\sqrt{x^2 - 1}}\right) \\ &=\; \sqrt{x^2 - 1} + \frac{x^2}{\sqrt{x^2 - 1}} - \frac{1}{\sqrt{x^2 - 1}} \\ &=\; 2\sqrt{x^2 - 1} \;>\; 0\,, \end{aligned}$$

for $x > 1$, we have $f(x)$ increases in x and $f(x) > f(1) = 0$ for $x > 1$. That is, $u\big(r(m)\big) < 0$ for $m > 1$.

(iv) By (i), $u(x)$ increases for $x > r(m)$ and for $x < -r(m)$. It decreases for $-r(m) < x < r(m)$. Since $u\big(r(m)\big) < 0$ (by (iii)) and $u(m) > 0$ by the continuity and monotonicity of $u(x)$, there exists a unique $\xi(m)$, $r(m) < \xi(m) < m$, such that $u\big(\xi(m)\big) = 0$. □

Proof of Theorem 2.2:

(i) When $0 < m \le 1$, it follows from Lemma A.1 that

$$x \le -m\tanh\big(m|x|\big) \quad\Leftrightarrow\quad x \le 0$$

and

$$x \ge m\tanh\big(m|x|\big) \quad\Leftrightarrow\quad x \ge 0$$

and this shows that, when $m \le 1$, we can rewrite (2.7) as $\delta^{\mathrm{mr}}(x) = m\tanh(mx)$ for $x \in (-\infty, \infty)$.

(ii) When $m > 1$, let $u(x) = x - m\tanh(mx)$. By Lemma A.1, $u(x)$ increases in x when $x > r(m)$ and when $x < -r(m)$. It decreases in x when $-r(m) < x < r(m)$. Moreover, $\xi(m)$ is the unique root of $u(x) = 0$ in $[r(m), m]$. Hence, $u(0) = u\big(-\xi(m)\big) = u\big(\xi(m)\big) = 0$, $u(x) > 0$ when $-\xi(m) < x < 0$ and when $x > \xi(m)$ and $u(x) < 0$ when $x < -\xi(m)$ and when $0 < x < \xi(m)$. So, when $m > 1$,

$$|x| < m\tanh\big(m|x|\big) \qquad \text{when} \quad |x| < \xi(m)$$

and

$$|x| > m\tanh\big(m|x|\big) \qquad \text{when} \quad |x| > \xi(m)\ .$$

This proves the result for the case where $m > 1$. □

A.3. Proofs for the Charras dominator δ^{ch}

Charras (1979) (see also Charras and van Eeden (1991)) gives conditions for estimators of the form (2.10) to dominate a boundary estimator δ, i.e. an estimator δ satisfying

$$\left.\begin{aligned} P_\theta\big(\delta(X) = a\big) &> 0 \\ P_\theta\big(\delta(X) = b\big) &> 0 \end{aligned}\right\} \qquad \text{for all} \quad \theta \in [a, b]\ . \tag{A.6}$$

Charras' conditions on $a(t)$ and $b(t)$ for (2.10) to dominate δ are

(a) $a(t)$ and $b(t)$ are continuous.

(**b**) $a(t)$ and $b(t)$ have continuous right derivatives which are bounded in absolute value on $[0,1]$.

(**c**) $a(0)=a$, $b(0)=b$ and $a(1)=b(1)$.

(**d**) For all $t\in[0,1]$, $a'_+(t)=\dfrac{da(t)}{dt+}>0$ and $b'_+<0$.

His conditions on the distributions of X and $\delta(X)$ are

(**1**) Condition (2.8) is satisfied.

(**2**) The loss function $L(\theta,d)$ has, for all θ in a neighbourhood N of $[a,b]$, a partial derivative $\partial L/\partial d$ with respect to d which is, on $N\times N$, continuous in d and in θ.
Moreover,

$$\frac{\partial L(\theta,d)}{\partial d}\begin{cases}<0 & \text{when } d<\theta\\ =0 & \text{when } d=\theta\\ >0 & \text{when } d>\theta\ .\end{cases}$$

(**3**) The estimator δ to be dominated satisfies (A.6).

(**4**) The estimator δ has, for each $\theta\in[a,b]$, a Lebesgue density on (a,b), i.e. there exists a function $f(y,\theta)$ such that, for all (α,β) with $a<\alpha<\beta<b$,

$$P_\theta\big(\alpha<\delta(X)<\beta\big)=\int_\alpha^\beta f(y,\theta)\,dy\ .$$

Moreover, that density is bounded on $(a,b)\times[a,b]$.

Clearly, these Charras conditions are satisfied for our bounded-normal-mean problem.

Remark. Charras also has results for the case where δ has a discrete distribution.

Our proof of Theorem 2.4 is a special case of Charras' proof for his general case and we need the following lemmas A.2, A.3 and A.4 for our proof. The proofs of the lemmas A.2 and A.3 are straightforward and omitted.

Lemma A.2. *Let* $u(x)=x^2\,\Phi(-x)-\Phi(-x)-x\,\phi(x)$. *Then:*

(**i**) *The risk function of* $\delta^{\rm mle}$ *is given by*

$$R(\theta,\delta^{\rm mle})\;=\;1+u(m+\theta)+u(m-\theta)\ . \tag{A.7}$$

(**ii**) *The risk function of* $\delta^{\rm ch}$ *is given by*

$$R(\theta,\delta^{\rm ch})\;=\;1+u(m-\varepsilon+\theta)+u(m-\varepsilon-\theta)\ . \tag{A.8}$$

Lemma A.3. *Let $g(x) = u'(x) = 2\,x\,\Phi(-x)$. Then $g'(x) = 2\big(\Phi(-x) - x\,\phi(x)\big)$, $g''(x) = 2(x^2-2)\,\phi(x)$ and the following properties of these functions hold:*

- **(i)** *$g''(x) \geq 0$ if and only if $|x| \geq \sqrt{2}$ and $g''(x) \to 0$ as $x \to \pm\infty$.*
- **(ii)** *$g'(x)$ increases in x if and only if $|x| > \sqrt{2}$; $g'(x)$ attains its maximum at $x = -\sqrt{2}$, its minimum at $x = \sqrt{2}$ and $g'(0) = 1$. There is a unique root η_0 of $g'(x) = 0$, $\eta_0 \in (0, \sqrt{2})$, $g'(x) \to 2$ as $x \to -\infty$ and $g'(x) \to 0$ as $x \to \infty$.*
- **(iii)** *$g(x)$ has the same sign as x for $x \in (-\infty, \infty)$; $g(x)$ increases in x if $x < \eta_0$ and decreases otherwise; $g(x)$ attains its maximum at $x = \eta_0$ and the unique root of $g(x) = 0$ is $x = 0$; $g(x) \to -\infty$ as $x \to -\infty$ and $g(x) \to 0$ as $x \to \infty$.*

Lemma A.4. *Let $h(x,\theta) = g(x+\theta) + g(x-\theta)$, where (see the lemmas A.2 and A.3) $g(x) = u'(x) = 2x\,\Phi(-x)$ and $u(x) = x^2\,\Phi(-x) - \phi(-x) - x\,\phi(x)$. Then*

- **(i)** *For fixed $\varepsilon \in (0, m)$*

$$\min_{\theta \in [m-\varepsilon, m]} h(m-\varepsilon, \theta) \;=\; h(m-\varepsilon, m) \;=\; g(2m-\varepsilon) + g(\varepsilon) - 2\,\varepsilon \;.$$

- **(ii)** *Let $\psi_m(x) = g(2m-x) + g(x) - 2x$ for $x \in [0, m]$. Then $\psi'_m(x) < 0$, $\psi_m(0) > 0$ and $\psi_m(m) < 0$, so there exists a unique root $\varepsilon_m \in (0, m)$ of $\psi_m(x) = 0$ with $\psi_m(x) > 0$ for $0 \leq x < \varepsilon_m$ and $\psi_m(x) < 0$ for $\varepsilon_m < x \leq m$.*

Proof:

- **(i)** Consider

$$\frac{\partial}{\partial\theta} h(m-\varepsilon, \theta) \;=\; g'(m-\varepsilon+\theta) - g'(m-\varepsilon-\theta) \;.$$

For $\theta \in (m-\varepsilon, m]$ we have $m-\varepsilon+\theta > 0$ and $m-\varepsilon-\theta < 0$. So (see Lemma A.3) $g'(m-\varepsilon+\theta) < g'(0) = 1$ and $g'(m-\varepsilon-\theta) > g'(0) = 1$. Hence $g'(m-\varepsilon+\theta) - g'(m-\varepsilon-\theta) < 0$ and so $\frac{\partial}{\partial\theta} h(m-\varepsilon, \theta) < 0$. In other words, $h(m-\varepsilon, \theta)$ decreases as θ increases in $(m-\varepsilon, m]$, which implies that

$$\min_{\theta \in [m-\varepsilon, m]} h(m-\varepsilon, \theta) \;=\; h(m-\varepsilon, m) \;.$$

- **(ii)** Note that $h(m-\varepsilon, m) = g(2m-\varepsilon) + g(\varepsilon) - 2\varepsilon = \psi_m(\varepsilon)$. Since $\psi'_m(x) = -2 - g'(2m-x) + g'(x)$, with (see Lemma A.3) $g'(2m-x) > g'(\sqrt{2}) > -1$ and $g'(x) < 1$, we have $\psi'_m(x) < 0$ for $x \in [0, m]$. □

Proof of Theorem 2.4: First of all it is clear that, for all $\varepsilon \in (0,m]$, $\delta^{\rm ch}$ dominates $\delta^{\rm mle}$ on $[-m+\varepsilon, m-\varepsilon]$. Further, by symmetry, it is sufficient to look at the behaviour of the risk functions on $(m-\varepsilon, m]$.

Let

$$\Delta(\theta,\varepsilon) \,=\, R(\theta,\delta^{\rm mle}) - R(\theta,\delta^{\rm ch}) \ ,$$

then, by Lemma A.2,

$$\begin{aligned}\frac{\partial}{\partial\varepsilon}\Delta(\theta,\varepsilon) \,&=\, -\frac{\partial}{\partial\varepsilon}\Big[u(m-\varepsilon+\theta)+u(m-\varepsilon-\theta)\Big]\\ &=\, u'(m-\varepsilon+\theta)+u'(m-\varepsilon-\theta)\\ &=\, g(m-\varepsilon+\theta)+g(m-\varepsilon-\theta)\\ &=\, h(m-\varepsilon,\theta)\ ,\end{aligned}$$

where $h(x,\theta)$ and $g(x)$ are defined in Lemma A.4.

Then, by Lemma A.4 (i), we have

$$\min_{\theta\in(m-\varepsilon,m]} h(m-\varepsilon,\theta) \,=\, h(m-\varepsilon,m) \,=\, \psi(\varepsilon) \,>\, 0\ ,$$

for $\varepsilon \in (0,\varepsilon_m)$, where ε_m is given by (ii) in Lemma A.4, implying that, for $0 \le \varepsilon \le \varepsilon_m$,

$$\frac{\partial}{\partial\varepsilon}\Delta(\theta,\varepsilon) \ge h(m-\varepsilon,m) \,=\, \psi(\varepsilon) \ge 0\ .$$

But, $\Delta(m-\varepsilon,\theta) > 0$ for all $\varepsilon \in (0,m]$, which proves the theorem. □

A.4. Proof of Theorem 2.5

By Lemma A.2

$$\Delta_o(\theta,m) \,=\, R(\theta,\delta^{\rm mle}) - R(\theta,\delta_o) \,=\, u(m+\theta)+u(m-\theta)+1-\theta^2\,.$$

So, it needs to be shown that $u(2m)+1/2-m^2=0$ has a unique positive root m_1 and that

$$u(m+\theta)+u(m-\theta)+1-\theta^2 \begin{cases}\ge 0 & \text{for all } \ \theta\in[0,m]\\ > 0 & \text{for some } \ \theta\in[0,m]\end{cases}$$

if and only if $0 < m < m_1$.

First note that (see Lemma A.3)

$$\Delta_o(0,m) \,=\, 2\,u(m)+1 \,>\, 0 \qquad \text{for } \ m>0\ .$$

Further, with $g(x) = u'(x) = 2x\,\Phi(-x)$,

$$\frac{\partial}{\partial\theta}\Delta_o(\theta, m) \;=\; g(m+\theta) - g(m-\theta) - 2\,\theta$$

and

$$\frac{\partial^2}{\partial\theta^2}\Delta_o(\theta, m) \;=\; g'(m+\theta) + g'(m-\theta) - 2\;,$$

so that

$$\frac{\partial}{\partial\theta}\Delta_o(\theta, m)\big|_{\theta=0} = 0 \qquad \text{for all} \quad m > 0$$

and (see Lemma A.3)

$$\frac{\partial^2}{\partial\theta^2}\Delta_o(\theta, m) < 0 \qquad \text{for all} \quad 0 \le \theta \le m, \quad m > 0\;,$$

implying that $\Delta_o(\theta, m)$ is, for each $m > 0$, decreasing in $\theta \in [0, m]$.

A necessary and sufficient condition for δ_o to dominate δ^{mle} for a given $m > 0$ is therefore that $\Delta_o(m, m) \ge 0$. But

$$\Delta_o(m, m) \;=\; u(2m) + u(0) + 1 - m^2 \;=\; u(2m) + 1/2 - m^2$$

and this function has the following properties:

(**1**) $\Delta_o(0, 0) = u(0) + 1/2 = 0$;

(**2**) $\dfrac{d}{dm}\Delta_o(m, m) = 2\big(g(2m) - m\big) = 2m\big(4\Phi(-2m) - 1\big)$.

So

$$\frac{d}{dm}\Delta_o(m, m) \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \iff m \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} \frac{1}{2}\,\Phi^{-1}\!\left(\frac{3}{4}\right).$$

Further, $\Delta_o(\sqrt{2}/2, \sqrt{2}/2) = u(\sqrt{2}) < 0$ and thus there exists a unique $m_1 > 0$ with

$$\Delta_o(m_1, m_1) = 0 \quad \text{and} \quad \Delta_o(m, m) > 0 \qquad \text{for} \quad 1 < m < m_1\;,$$

which, together with the fact that $\Delta_o(\theta, m)$ is decreasing in θ for $\theta \in [0, m]$, proves the result. Numerically we found $m_1 \approx 0.5204372$.

B. GRAPHS FOR SECTION 3

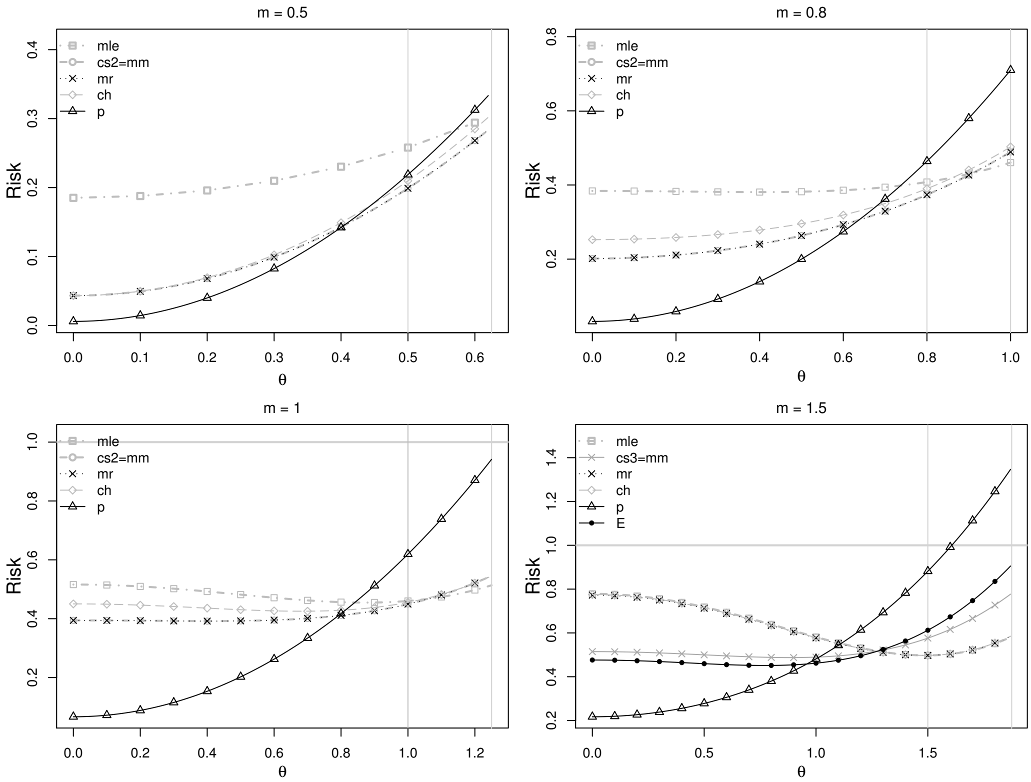

Figure 1: Risk functions of various estimators as a function of θ for $m = 0.5$, 0.8, 1.0 and 1.5.

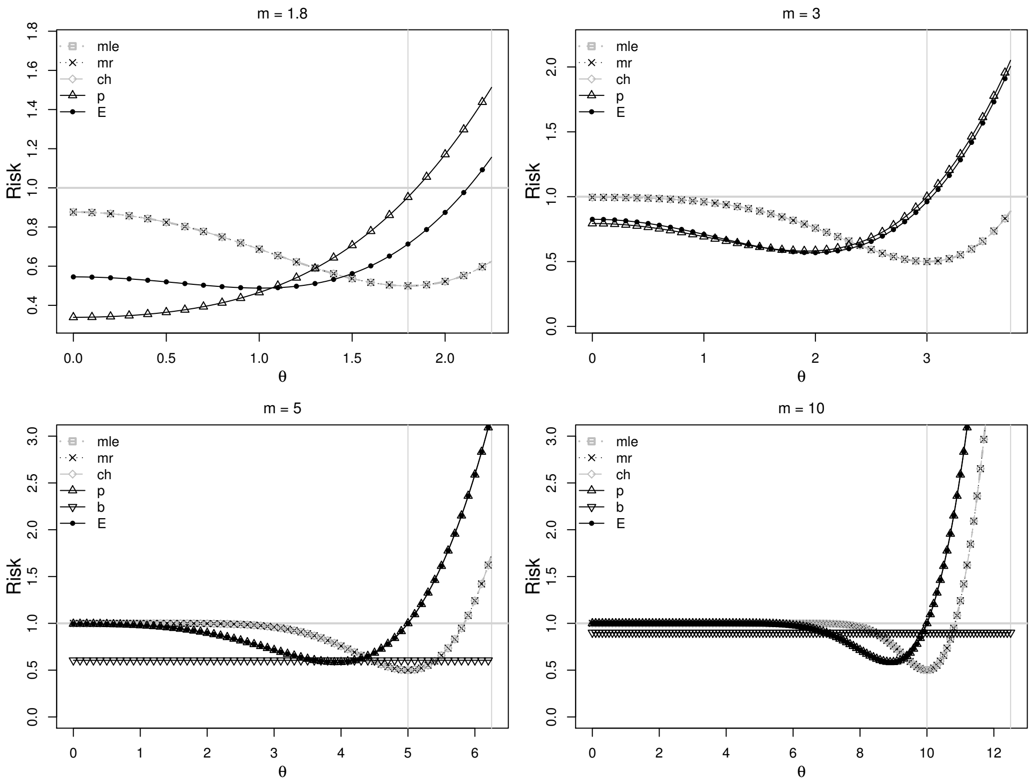

Figure 2: Risk functions of various estimators as a function of θ for $m = 1.8$, 3, 5 and 10.

ACKNOWLEDGMENTS

The authors thank the referee for his very useful comments and Éric Marchand for pointing out the result in Marchand and Perron (2001). Yiping Dou highly appreciates Professor James V. Zidek's generous financial support for this work.

REFERENCES

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

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

[3] Brown, L.D. (1986). *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory*, IMS Lecture Notes-Monograph Series, Vol. 9, Institute of Mathematical Statistics, Hayward, California.

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

[5] Charras, A. (1979). *Propriété Baysienne et Admissibilité d'Estimateurs dans un Sous-ensemble Convex de* $\mathbb{R}^p$, PhD thesis, Université de Montréal, Montréal, Canada.

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

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

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

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

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

[11] Perron, F. (2003). *Improving on the MLE of* p *for a binomial* (n, p) *when* p *is around* $1/2$. In: "Mathematical Statistics and Applications: Festschrift for Constance van Eeden" (M. Moore, S. Froda and C. Léger, Eds.), *IMS Lecture Notes and Monograph Series*, **42**, 45–61, Institute of Mathematical Statistics, Hayward, California, USA.

[12] Robert, C. (1990). Modified Bessel functions and their applications in probability and statistics, *Statistics and Probability Letters*, **9**, 155–161.

[13] van Eeden, C. (2006). Restricted-Parameter-Space Estimation Problems: Admissibility and Minimaxity Properties, *Lecture Notes in Statistics*, **188**, Springer.