CALL #:	**Per QA276.A1S748**
LOCATION:	**AUM :: Main Library :: INTEGRATED SCIENCE AND ENGINEERING LIBRARY**
TYPE:	Article
USER JOURNAL TITLE:	Statistics & Decisions
OCLC JOURNAL TITLE:	Statistics & decisions
AUM CATALOG TITLE:	Statistics & decisions
ARTICLE TITLE:	Minimax estimation of a constrained binomial proportion
ARTICLE AUTHOR:	Marchand, Eric and MacGibbon, Brenda
VOLUME:	18
ISSUE:	2
YEAR:	2000
PAGES:	129--167
ISSN:	0721-2631
OCLC #:	
CROSS REFERENCE ID:	157346
VERIFIED:	

BORROWER:	**ORE :: Main Library**
PATRON:	**Bulatov,Yaroslav**
PATRON ID:	
PATRON ADDRESS:	
PATRON PHONE:	
PATRON FAX:	
PATRON E-MAIL:	
PATRON DEPT:	Engineering
PATRON STATUS:	Undergraduate
PATRON NOTES:	

Statistics & Decisions 18, 129 – 167 (2000)

MINIMAX ESTIMATION OF A CONSTRAINED BINOMIAL PROPORTION[1]

Éric Marchand and Brenda MacGibbon

Received: Revised version: *February 7, 2000*

Abstract

The problem of estimating a binomial proportion constrained to lie in an interval of the form $[a,b] \neq [0,1]$ is considered. The minimax and linear minimax estimators are obtained for both quadratic and information-normalized loss functions. For parameter spaces of the form $[a,1-a]$ and $[0,b]$, the minimax estimators are used as a benchmark to compare the performance of estimators belonging to other classes such as linear minimax estimators, the MLE, and Bayes estimators associated with translated Beta priors.

1. INTRODUCTION

In many statistical problems, there is definite prior information concerning the values of a parameter, often in the form of bounds. Although such information is certainly useful, it often leads to surprising theoretical difficulties with regards to (i) the selection of an estimator, and (ii) the comparison of various possible estimators. One common, admittedly conservative, approach is, given some measure of the error, to compute the maximum risk over the bounded parameter space, and then search for the estimator that minimizes the maximum risk. The resulting estimator may or may not be an attractive estimator but, even in the failing case, the resulting minimax risk does provide (i) a benchmark against which to measure other estimators, and (ii) a measure of the prior information (by comparison with the minimax risk computed ignoring the prior information).

---

[1] AMS 1991 subject classifications: 62 C10, 62 C20, 62 F10, 62 F30

Keywords and phrases: Minimax estimation, least favourable distribution, binomial distribution, constrained parameter space, linear minimax estimation, maximum likelihood estimator, translated Beta prior

The problem considered here is that of estimating the probability parameter θ of a binomial distribution under the additional assumption that it lies in a bounded interval of the form $[a,b]\neq[0,1]$, with $0\leq a<b\leq 1$. An example of such a constraint occurring is in Warner's [28] random response model where the proportion is constrained to a symmetric interval $[a,1-a]$. Another example is given by a series system with two components C_1 and C_2 in which we are able to observe the proportion θ of successfully transmitted messages (i.e. success at both components) in a sample of size n. Suppose further that the reliability associated with C_1 is known and equal to b. Then we must have $\theta\leq b$.

The problem of estimating a bounded normal mean θ ($N(\theta,1)$ with θ such that $a\leq|\theta|\leq b$) under squared error loss has been considered by Casella and Strawderman [6], Zinzius [31], Bickel [4], Levit [20] as well as Gatsonis, MacGibbon and Strawderman [12] among others. Casella and Strawderman, as well as Zinzius, proposed exact minimax solutions for small values of b-a, Bickel and Levit found the asymptotic solution, and Gatsonis, MacGibbon and Strawderman studied other estimators. For the multivariate version, where the normal mean is constrained to lie in a sphere, Bickel [4] studied the asymptotic minimax estimator and Berry [3] gave exact solutions when the sphere's radius is small. Donoho, Liu and MacGibbon [9] compared minimax linear and non-linear risks as well as study the problem of finding minimax estimators on quadratically convex L^p bodies in L^2 space. Vidakovic and DasGupta [26] also studied the efficiency of linear and near linear rules for this problem. Johnstone and MacGibbon [14],[15] studied analogous problems for the estimation of a bounded Poisson mean θ using the information-normalized quadratic loss function.

Our problem here can be stated more precisely as follows. Let X be a binomial random variable $Bi(n,\theta)$; $0\leq a<\theta<b\leq 1$; where (n,a,b) is known; and with probability function[2]

$$P_\theta(X=x)=\binom{n}{x}\theta^x(1-\theta)^{n-x}\ [x\in\{0,1,\ldots,n\}].$$

Under squared error loss for the symmetric case a+b=1, Moors [22] and Berry [2] previously found minimax estimators for small values of n; Lehn and Rummel [19] studied Gamma-minimax

[2] Here, and throughout, we use [A] to represent the indicator function of the set A.

estimation; and the work of Charras ar
compact truncated parameter spaces c
loss $L(\theta,d)=(d-\theta)^2$, but also, in parallel,
It is well known for convex loss func
class of nonrandomized rules. The p
function, defined for $\theta\in\Theta=[a,b]$
$R^*(\theta,\delta)=E_\theta[L^*(\theta,\delta(X))]$. In the situatio
to the parameter space, we note that
be minimax) at $\theta=0$ (or $\theta=1$) if and c
An estimator δ^* is said to be minim

$$\sup_{a\leq\theta\leq b}\{R$$

where the infimum is taken over the c
is said to be linear minimax when the
of x. The associated infima will be c
respectively. Analogous definitions w
Minimax problems are often easie
Given a prior distribution G on the pa
a rule δ is the integrated risk

$$r($$

A Bayes procedure δ^G renders $r(\delta,G$
infimum $r(G)$ $(=r(\delta^G,G))$ will be refe
distribution G^* is called "least favou
Subject to the decision problem satis
favourable distribution G^* exists, **(ii)** δ
(see Wald [27] or Kempthorne [17]).
binomial case the support of G^* is dis

estimation; and the work of Charras and van Eeden [7] on Bayes and admissibility estimators in compact truncated parameter spaces can also be applied. We consider here not only quadratic loss $L(\theta,d)=(d-\theta)^2$, but also, in parallel, information-normalized loss $L^*(\theta,d)=\{(d-\theta)^2/\theta(1-\theta)\}\,[d\neq\theta]$. It is well known for convex loss functions (e.g. Ferguson [10]), that it suffices to consider the class of nonrandomized rules. The performance of a rule δ will be characterized by its risk function, defined for $\theta\in\Theta=[a,b]$, and equal to either $R(\theta,\delta)=E_\theta[L(\theta,\delta(X)]$ or $R^*(\theta,\delta)=E_\theta[L^*(\theta,\delta(X)])$. In the situations where we study normalized loss and 0 (or 1) belongs to the parameter space, we note that an estimator δ will have finite risk (and be a candidate to be minimax) at $\theta=0$ (or $\theta=1$) if and only if $\delta(0)=0$ (or $\delta(n)=1$).

An estimator δ^* is said to be minimax for quadratic loss if

$$\sup_{a\le\theta\le b}\{R(\theta,\delta^*)\}=\inf_{\delta}\sup_{a\le\theta\le b}\{R(\theta,\delta)\},$$

where the infimum is taken over the class of all nonrandomized estimators. Such an estimator is said to be linear minimax when the above infimum is taken over the class of linear estimators of x. The associated infima will be denoted as the minimax risk and the linear minimax risk respectively. Analogous definitions will be used for normalized loss.

Minimax problems are often easier to solve by considering the associated dual problems. Given a prior distribution G on the parameter space Θ, a related measure of the performance of a rule δ is the integrated risk

$$r(\delta,G)=\int_\Theta R(\theta,\delta)dG(\theta)\,.$$

A Bayes procedure δ^G renders $r(\delta,G)$ as small as possible (among all δ), and the resulting infimum $r(G)$ $(=r(\delta^G,G))$ will be referred to as the Bayes risk associated with G. A prior distribution G^* is called "least favourable" if it renders the Bayes risk as large as possible. Subject to the decision problem satisfying certain regularity conditions, **(i)** (at least one) least favourable distribution G^* exists, **(ii)** δ^{G^*} is minimax, and **(iii)** the minimax risk is given by $r(G^*)$ (see Wald [27] or Kempthorne [17]). It is also easy to deduce from Kempthorne [17] that in the binomial case the support of G^* is discrete and finite. As a consequence, we can formulate our quadratic loss problem as the following optimization problem:

$$\textit{maximize } r(G)=r(m,\theta_0,\dots,\theta_m,p_0,\dots,p_m)=\sum_{i=0}^{m} p_i R(\theta_i,\delta^G)$$

$$\textit{subject to } \; m\geq 1,\ \sum_{i=0}^{m} p_i=1;\ a\leq\theta_i\leq b,\ p_i\geq 0 \textit{ for } i=1,\dots,m;$$

$$\textit{where } R(\theta,\delta^G)=\sum_{x=0}^{n}(\delta^G(x)-\theta_i)^2\binom{n}{x}\theta_i^x(1-\theta_i)^{n-x},$$

$$\textit{and } \; \delta^G(x,n)=\frac{\sum_{i=0}^{m} p_i\theta_i^{x+1}(1-\theta_i)^{n-x}}{\sum_{i=0}^{m} p_i\theta_i^{x}(1-\theta_i)^{n-x}}.$$

In an analogous manner, the normalized problem may be cast as a similar optimization problem; with the replacement of R by R^* and δ^G by δ_*^G, with $\delta_*^G(x,n)=\delta^G(x-1,n-2)$ unless a=0,x=0, in which case $\delta_*^G(0)=0$, or unless b=1,x=n, in which case $\delta_*^G(n)=1$.

The algorithm used here is inspired by that of Kempthorne [17]. It is based on induction on m starting with m=0 or 1. We start with a prior guess for the least favourable prior $\tilde{G}$ parametrized by $(m,\tilde{\theta}_0,\dots,\tilde{\theta}_m,\tilde{p}_0,\dots,\tilde{p}_m)$. Using numerical algorithms based on subroutine EO4VCF of NAG libraries, we solve the (2m+1) dimensional optimization problem and find the "best" G supported on m points (parametrized by $(m,\theta_0,\dots,\theta_m,p_0,\dots,p_m)$). Now, we know that G will be least favourable only if it is an "equalizer",

$$i.e. \;\; \sup_{a\leq\theta\leq b} R(\theta,\delta^G)=R(\theta_0,\delta^G)=\dots=R(\theta_m,\delta^G). \tag{1.1}$$

If G is not an "equalizer", we choose new weights $\tilde{p}_0,\dots,\tilde{p}_m$ so that it is. Then it is necessary to verify whether δ^G is a minimax estimator; that is, whether $R(\theta,\delta^G)\leq R(\theta_m,\delta^G)$ for all $\theta\in\Theta$. If not, then we add one more support point and consider priors G supported on m+2 points $(\tilde{\theta}_0,\dots,\tilde{\theta}_{m+1})$ and repeat the process until a minimax estimator is achieved. It should be noted that the algorithm can be used for finding minimax estimators with respect to other loss functions such as entropy loss in Wieczorkowski and Zieliński [29].

The paper is organized as follows.
This is possible when $|\theta|\leq$b, in the c
two-point priors for normalized loss.
results are discussed in Section 5. S
yielded favourable results say in the
in the work of Donoho, Liu, and Mac
[21] among others, we determine in S
also study the loss of efficiency incu
mle is studied and compared to a cla
graphs and the figures and the adv
Appendix in Section 6 contains the t

2. SOME EXACT SOLUTIONS

2.1 *Quadratic Loss, $\theta\leq b$, Two-Point*

For the case a=0 with quadratic l
concentrated on {0,b} when b is sma

For the prior satisfying P(θ=0)=
$\delta_\varepsilon(x)$=w[x=0]+b[x≥1] with w={b(1-b)

$R(\theta,\delta_\varepsilon)=(w$

As seen by (1.1), the rule δ_ε will be
(ii) $R(\theta,\delta_\varepsilon)\leq R(0,\delta_\varepsilon)$. Condition **(i)** is
$(1-b)^{n/2}/\{1+(1-b)^{n/2}\}$ and w=w*=bε*. F

(i

From (2.1), we obtain

The paper is organized as follows. Our first objective in Section 2 is to obtain exact solutions. This is possible when $|\theta|\leq b$, in the case of two-point priors for quadratic loss and for one and two-point priors for normalized loss. For larger supports the above algorithm is used and the results are discussed in Section 5. Since the comparison of linear vs optimal, which has often yielded favourable results say in the multivariate normal case, surfaces as an underlying theme in the work of Donoho, Liu, and MacGibbon [9], Vidakovic and Das Gupta [26], and Marchand [21] among others, we determine in Section 3 the linear minimax risk for all triplets (n,a,b). We also study the loss of efficiency incurred. This too is discussed in Section 5. In Section 4, the mle is studied and compared to a class of Bayes rules for Beta priors. Section 5 discusses the graphs and the figures and the advantages and disadvantages of various estimators. The Appendix in Section 6 contains the technical proofs.

2. SOME EXACT SOLUTIONS

2.1 Quadratic Loss, $\theta\leq b$, Two-Point Priors

For the case a=0 with quadratic loss, it seems plausible that the least favourable prior is concentrated on {0,b} when b is small.

For the prior satisfying $P(\theta=0)=1-P(\theta=b)=\varepsilon$, the corresponding Bayes rule is given by $\delta_\varepsilon(x)=w[x=0]+b[x\geq1]$ with $w=\{b(1-b)^n(1-\varepsilon)\}/\{\varepsilon+(1-\varepsilon)(1-b)^n\}$ and associated risk

$$R(\theta,\delta_\varepsilon)=(w-\theta)^2(1-\theta)^n+(b-\theta)^2[1-(1-\theta)^n]. \tag{2.1}$$

As seen by (1.1), the rule δ_ε will be minimax if **(i)** $R(0,\delta_\varepsilon)=R(b,\delta_\varepsilon)$; and, for all $\theta\in[0,b]$, **(ii)** $R(\theta,\delta_\varepsilon)\leq R(0,\delta_\varepsilon)$. Condition **(i)** is equivalent to $w^2=[w-b]^2(1-b)^n$ which forces $\varepsilon=\varepsilon^*=(1-b)^{n/2}/\{1+(1-b)^{n/2}\}$ and $w=w^*=b\varepsilon^*$. For condition (ii) to be satisfied, it is necessary that

$$\textit{(iii)}\quad \frac{\partial R(\theta,\delta_{\varepsilon^*})}{\partial\theta}\Big|_{\theta=0}\leq 0.$$

From (2.1), we obtain

$$\frac{\partial R(\theta,\delta_{\varepsilon^*})}{\partial\theta}=2(\theta-b)-\{(1-\theta)^{n-1}(b-w^*)(2\theta+n-2-nb-nw^*)\}; \qquad (2.2)$$

$$\frac{\partial^2 R(\theta,\delta_{\varepsilon^*})}{\partial^2\theta}=2(1-(w^*-b)),\ \textit{for}\ n=1;$$

$$\textit{and}\ \frac{\partial^2 R(\theta,\delta_{\varepsilon^*})}{\partial^2\theta}=2+2(1-\theta)^{n-2}(b-w^*)(3\theta+n-3-nw^*),\ \textit{for}\ n>1.$$

Since the second derivative is of the form $c_1+c_2(1-\theta)^{n-2}(c_3+c_4)$ where $c_1\geq 0$, $c_2\geq 0$, and $c_3\geq 0$, it has at most one sign change on $[0,b]$ from - to +. This, in turn, now implies that condition (iii) is also sufficient, and therefore equivalent to (ii). Thus, from (2.2), (ii) is true if and only if

$$\frac{\partial R(\theta,\delta_{\varepsilon^*})}{\partial\theta}\Big|_{\theta=0}=-2b-(b-w^*)(n-2-nb-nw^*)\leq 0$$

$$\Leftrightarrow\ g(n,b)=(1-nb)(1-b)^{n/2}+(1-b)^n-\frac{nb}{2}\geq 0. \qquad (2.3)$$

We note the following properties of $g(n,\cdot)$: $g(n,\cdot)$ is decreasing on $[0,n^{-1}]$; $g(n,0)>0$; and $g(n,b)<g(n,n^{-1})<0$ for $b>n^{-1}$. These properties now imply that (2.3) is verified if and only if $b\leq b_*(n)$ where $b_*(n)$ is the unique root of the equation $g(n,b)=0$ on $[0,n^{-1}]$. To summarize, we have established the following.

THEOREM 2.1. δ_{ε^*} is minimax if $b\leq b_*(n)$. When this is the case, the two point prior such that $P(\theta=0)=1-P(\theta=b)=\varepsilon^*$ is the least favourable prior and the minimax risk is given by $R(0,\delta_{\varepsilon^*})=(w^*)^2=b^2(1-b)^n/\{1+(1-b)^{n/2}\}^2$.

REMARK 2.2. Some values of $b_*(n)$ are presented in Figure 1. We observed that the parameter space has to be quite small for the associated δ_{ε^*} of Theorem 2.1 to be minimax. Furthermore, since b_* decreases with n, let us consider $y=\lim_{n\to\infty} y_n$ where $y_n=nb_*(n)$. The reader may verify that this limit exists. As well, from the definition of b_* above, we must have $(1-y_n)(1-y_n/n)^{n/2}+(1-y_n)^n-(y_n/2)=0$, and, consequently, $(1-y)e^{-y/2}+e^{-y}-(y/2)=0$, which implies $y=y_*\approx.912955$. Thus, for large n, $b_*(n)\approx.912955/n$; and Figure 1 suggests the approximation is quite accurate.

2.2 Normalized Loss, $\theta\leq b$, One-poi

For the case a=0 with normalized
Johnstone and MacGibbon [14] that t
b is small. As was the case for the n
finite risk an estimator must satisfy δ
Bayes risk over the class of such esti
we investigate whether such a one-p
For the prior such that $P(\theta=b)=1$,
is equal to 0 at $\theta=0$ and

$$R^*(\theta,\delta_b)=\frac{}{\theta}$$

$$=(1-\theta)^{n-}$$

for $0<\theta\leq b$. δ_b will be minimax if, fo

Below, we establish that $R^*(\cdot,\delta_b)$ is co
(2.5) may hold, it is critical to compa
next result summarizes this analysis.

THEOREM 2.3. δ_b is minimax with
$b^*(n)$ is the smallest positive root of t
Proof. For n=1, all rules are linear ar
Theorem 3.2 below, it is minimax wi
$\lim_{\theta\to 0^+} R(\theta,\delta_b)=nb^2$, $R(b,\delta_b)=b(1-b)^{n-1}$ and
i.e. $b\leq b^*(n)$. This last condition turns
This last result is established in Secti

REMARK 2.4. Some values of $b^*(n)$
results for the two point prior for qua

2.2 Normalized Loss, θ≤b, One-point priors

For the case a=0 with normalized loss, it seems plausible as in the Poisson case studied by Johnstone and MacGibbon [14] that the least favourable prior is given by a point mass at b when b is small. As was the case for the normalized loss in the Poisson case, here too in order to have finite risk an estimator must satisfy $\delta(0)=0$. By a slight abuse of terminology, we minimize the Bayes risk over the class of such estimators in order to determine the Bayes rules. In this section, we investigate whether such a one-point prior is least favourable.

For the prior such that $P(\theta=b)=1$, the Bayes rule is given by $\delta_b(x)=b[x>0]$. Its risk function is equal to 0 at $\theta=0$ and

$$\begin{aligned} R^*(\theta,\delta_b) &= \frac{\theta^2}{\theta(1-\theta)}(1-\theta)^n + \frac{(b-\theta)^2}{\theta(1-\theta)}(1-(1-\theta)^n) \\ &= (1-\theta)^{n-1}\Big(2b-\frac{b^2}{\theta}\Big) + \frac{b^2}{\theta} + \frac{(1-b)^2}{1-\theta} - 1\,, \end{aligned} \tag{2.4}$$

for $0<\theta\le b$. δ_b will be minimax if, for all $0\le\theta\le b$,

$$R^*(\theta,\delta_b) \le R^*(b,\delta_b)\,. \tag{2.5}$$

Below, we establish that $R^*(\cdot,\delta_b)$ is convex on (0,b] for sufficiently small b. Thus, in order that (2.5) may hold, it is critical to compare the value $R^*(\theta,\delta_b)$ at b with its limiting value at 0. The next result summarizes this analysis.

THEOREM 2.3. δ_b is minimax with minimax risk equal to $b(1-b)^{n-1}$ as long as $b\le b^*(n)$, where $b^*(n)$ is the smallest positive root of the equation $nb-(1-b)^{n-1}=0$.

Proof. For n=1, all rules are linear and since δ_b coincides with the linear minimax rule given in Theorem 3.2 below, it is minimax with risk b for all $b\le 1$. For $n\ge 2$, we infer from (2.4) that $\lim_{\theta\to 0^+} R(\theta,\delta_b)=nb^2$, $R(b,\delta_b)=b(1-b)^{n-1}$ and that, for (2.5) to hold, it is necessary that $nb^2\le b(1-b)^{n-1}$, i.e. $b\le b^*(n)$. This last condition turns out to be also sufficient as $R(\cdot,\delta_b)$ is convex on $(0,b^*(n)]$. This last result is established in Section 6.1 and completes our proof.

REMARK 2.4. Some values of $b^*(n)$ are presented in Figure 1. We observe, analogous to the results for the two point prior for quadratic loss, that the parameter space has to be quite small for δ_b of Theorem 2.3 to be minimax and that b^* decreases with n. Let us consider $x=\lim_{n\to\infty} x_n$ where $x_n=nb^*(n)$. The reader may verify that this limit exists. As well, from the definition of b^* above, we must have $x_n=(1-x_n/n)^{n-1}$, and, consequently, $x=e^{-x}$, which implies $x=x_*\approx.567143290$. Thus, for large n, $b^*(n)\approx.567143290/n$; and Figure 1 suggests the approximation (actually an upper bound) is quite accurate.

2.3 *Normalized Loss, $\theta\leq b$, Two-point priors*

For moderate but not too large values of b (i.e. $b^*(n)<b\leq b^{**}(n)$), it seems plausible that, in essence, the least favourable prior is concentrated on $\{0,b\}$. More specifically, we consider the sequence $\{\pi_m:m=1,2,...\}$ of priors; such that $P(\theta=\gamma_m)=1-P(\theta=b)=\varepsilon$ under π_m, with $\gamma_m\leq b$, and $\lim_{m\to\infty}\gamma_m=0$; and investigate whether the associated extended Bayes rule (δ_0^* below) is minimax.

To begin our investigation, it will suffice to consider rules δ such that $\delta(0)=0$ (otherwise the risk at 0 will be infinite). Suppose that $\delta(1)=z$, in which case the Bayes risk with respect to π_m, $r_m(\delta)=\varepsilon R^*(\gamma_m,\delta)+(1-\varepsilon)R^*(b,\delta)$, may be expressed as

$$A(m,z)+\epsilon\gamma_m(1-\gamma_m)^{n-1}+(1-\epsilon)b(1-b)^{n-1}+$$
$$\sum_{x=2}^{n}\binom{n}{x}[\epsilon(\delta(x)-\gamma_m)^2\gamma_m^{x-1}(1-\gamma_m)^{n-x-1}+(1-\epsilon)(\delta(x)-b)^2b^{x-1}(1-b)^{n-x-1}],$$
$$\textit{with } A(m,z)=n\epsilon(1-\gamma_m)^{n-2}(z-\gamma_m)^2+n(1-\epsilon)(1-b)^{n-2}(z-b)^2.$$

It then follows that

$$\inf_{\delta} r_m(\delta)\geq(1-\epsilon)b(1-b)^{n-1}+\inf_{z} A(m,z).$$

Since $\inf_z A(m,z)=AB(b-\gamma_m)^2/(A+B)$, with $A=n\varepsilon(1-\gamma_m)^{n-2}$ and $B=n(1-\varepsilon)(1-b)^{n-2}$, we have

$$r=\lim_{m\to\infty}\inf_{\delta} r_m(\delta)\geq(1-\epsilon)b(1-b)^{n-1}+\frac{n\epsilon(1-\epsilon)b^2(1-b)^{n-1}}{\epsilon+(1-\epsilon)(1-b)^{n-2}}=\lim_{m\to\infty}r_m(\delta_0^*),$$

where $\delta_0^*(x)=z^*[x=1]+b[x=2]$; $z^*=\{b(1-\varepsilon)(1-b)^{n-2}\}/\{\varepsilon+(1-\varepsilon)(1-b)^{n-2}\}$. Note that δ_0^* is an extended Bayes rule with limiting Bayes risk

$$r = \epsilon \lim_{\theta \to 0^+} R^*(\theta,\delta_0^*) + (1-\epsilon) R^*(b,\delta_0^*). \tag{2.6}$$

Using a well known criteria for minimaxity (e.g. Berger [1], Theorem 18, page 350), δ_0^* will be minimax if

$$R^*(\theta,\delta_0^*) \le r, \textit{ for all } \theta \in [0,b], \tag{2.7}$$

where

$$R^*(\theta,\delta_0^*) = \theta(1-\theta)^{n-1} + n(1-\theta)^{n-2}(z^*-\theta)^2 + \frac{(b-\theta)^2}{\theta(1-\theta)}\{1-(1-\theta)^n - n\theta(1-\theta)^{n-1}\}. \tag{2.8}$$

From (2.6) and (2.7), a necessary condition for δ_0^* to be minimax is

$$\lim_{\theta \to 0^+} R^*(\theta,\delta_0^*) = R^*(b,\delta_0^*) = r, \tag{2.9}$$

which implies

$$n z^{*2} = b(1-b)^{n-1} + n(1-b)^{n-2}(z^*-b)^2. \tag{2.10}$$

Such a value of $z^* \in [0,b)$ is found to satisfy (2.10) as long as $nb^2 > b(1-b)^{n-1}$ (i.e. $b > b^*(n)$). Consequently, there exists a candidate δ_0^* satisfying (2.9) in precisely the situations where the one-point prior previously considered is not least favourable. Now, there remains to investigate whether (2.7) is verified. For n=2 and n=3, the polynomial in (2.8) may be manipulated, as shown in Section 6.2, to yield the following.

THEOREM 2.5. (a) For $n=2$ and $b>1/3$, δ_0^* is minimax with $z^*=(1+b)/4$, and minimax risk $(1+b)^2/8$.

(b) For $n=3$ and $\{(5-\sqrt{21})/2\} < b \le \{(25-2\sqrt{93})/11\}$, δ_0^* is minimax with $z^* = b-1+\{(1-b)(4-b)/3\}^{1/2}$, and minimax risk $3(z^*)^2$.

For larger n, we proceeded numerically to test (2.7) directly. We obtained for $n \le 50$ (and we conjecture for $n>50$) that δ_0^* is minimax whenever $b^*(n) < b \le b^{**}(n)$. Some values of $b^{**}(n)$ are given in Figure 1.

3. MINIMAX LINEAR RULES AND RISKS

In this section, we search for the minimax risk attained over the class of linear rules of the form $\delta_{\alpha,\beta}(x)=\beta+(\alpha x)/n$. For the problem of estimating θ without constraints, the minimax rule is linear with $\alpha=\alpha_0=\sqrt{n}/(\sqrt{n}+1)$, $\beta=(1-\alpha_0)/2(=1/(2(\sqrt{n}+1))$ for quadratic loss, and $\alpha=1$, $\beta=0$ for normalized loss. As well, the commonly used conjugate family of Beta priors yields, for both losses under study, Bayes rules that are also linear.

We determine below the linear minimax risk (and rule) for all triplets (n,a,b). Our eventual aim, given the simplicity of the minimax linear rules below, is to study the loss in efficiency incurred by restricting attention to the linear class.

3.1 Quadratic Loss

The risk function of $\delta_{\alpha,\beta}$ is given by

$$R(\theta,\delta_{\alpha,\beta})=[(\alpha-1)\theta+\beta]^2+\frac{\alpha^2}{n}\theta(1-\theta)\,; \tag{3.1}$$

which may also be written as $A\theta^2+B\theta+C$ with

$$A=(\alpha-1)^2-\frac{\alpha^2}{n},\ B=2\beta(\alpha-1)+\frac{\alpha^2}{n},\ and\ C=\beta^2\,. \tag{3.2}$$

We are searching for the minimax linear risk $S=\inf_{\alpha,\beta} S(\alpha,\beta)$ and the minimizing (α,β), sometimes called the $\text{argmin}S(\alpha,\beta)$, where $S(\alpha,\beta)=\sup_{\theta\in[a,b]} R(\theta,\delta_{\alpha,\beta})$. The two following lemmas will prove helpful.

LEMMA 3.1 Within the linear class of rules, $\delta_{\alpha,\beta}$ is inadmissible whenever $\alpha>1$ or $\beta<0$.

Proof. It is easily seen from (3.1) that $\delta_{1,0}$ dominates $\delta_{\alpha,\beta}$ whenever $\alpha>1$, and $\delta_{\alpha,0}$ dominates $\delta_{\alpha,\beta}$ whenever $\alpha\leq 1$, $\beta<0$.

LEMMA 3.2 Let $f(x)=Ax^2+Bx+C$, $M=\sup_{x\in[a,b]} f(x)$, and $x_0=-B/2A$. Then,

$f(a$
$M=$ $f(b$
$f(x$

Lemma 3.1 permits us to restrict
3.2 permits us, with a little algebra,

LEMMA 3.3 Let $\alpha_0=n^{1/2}/(n^{1/2}+1)$, g
B are as in (3.2). Then, for $\alpha\leq 1$ an

$S(\alpha,\beta)=$

where $I_1=\{(\alpha,\beta)$: $\alpha\leq\alpha_0$, $\beta\geq g_\alpha((a+b)/2)$
$I_4=\{(\alpha,\beta)$:$\alpha>\alpha_0$,$\beta\leq g_\alpha(b)\}$, and $I_5=\{(\alpha$
To determine the miminax linear
$S_j=\inf_{(\alpha,\beta)\in I_j} S(\alpha,\beta)$ for j=1,...,5, and then S=
work, which is illustrated in Section 6
of which is deferred to the Appendix
Theorem 3.5 where the linear minim

THEOREM 3.5

Given $\alpha_*=\dfrac{n^{1/2}}{n^{1/2}+\{\frac{|a+b-1}{b-a}}$

$H(\gamma)=\dfrac{(b-a)^2}{4}\,\dfrac{K_1\gamma^4+K_2\gamma^2}{(1+\gamma)^2}$

and γ_r, *the unique* p

then under quadratic loss, the linear

$$M = \begin{cases} f(a) \text{ if } A \geq 0, x_0 \geq (a+b)/2; \text{ or } A<0, x_0<a \\ f(b) \text{ if } A \geq 0, x_0 \leq (a+b)/2; \text{ or } A<0, x_0>b \\ f(x_0) \text{ if } A<0, a \leq x_0 \leq b. \end{cases}$$

Lemma 3.1 permits us to restrict our search for the $\operatorname{argmin} S(\alpha,\beta)$ to $\alpha \leq 1$ and $\beta \geq 0$. Lemma 3.2 permits us, with a little algebra, to express $S(\alpha,\beta)$, for $\alpha \leq 1$, $\beta \geq 0$, as follows.

LEMMA 3.3 Let $\alpha_0 = n^{1/2}/(n^{1/2}+1)$, $g_\alpha(x) = x(1-\alpha) + \{\alpha^2(1-2x)\}/\{2n(1-\alpha)\}$ and $x_0 = -B/2A$ where A, B are as in (3.2). Then, for $\alpha \leq 1$ and $\beta \geq 0$,

$$S(\alpha,\beta) = \begin{cases} R(a,\delta_{\alpha,\beta}) \text{ if } (\alpha,\beta) \in I_1 \cup I_2 \\ R(b,\delta_{\alpha,\beta}) \text{ if } (\alpha,\beta) \in I_3 \cup I_4 \\ R(x_0,\delta_{\alpha,\beta}) = \beta^2 - \dfrac{B^2}{4A} \text{ if } (\alpha,\beta) \in I_5 \end{cases}$$

where $I_1 = \{(\alpha,\beta): \alpha \leq \alpha_0, \beta \geq g_\alpha((a+b)/2)\}$, $I_2 = \{(\alpha,\beta): \alpha > \alpha_0, \beta \geq g_\alpha(a)\}$, $I_3 = \{(\alpha,\beta): \alpha \leq \alpha_0, \beta \leq g_\alpha((a+b)/2)\}$, $I_4 = \{(\alpha,\beta): \alpha > \alpha_0, \beta \leq g_\alpha(b)\}$, and $I_5 = \{(\alpha,\beta): \alpha > \alpha_0, g_\alpha(b) \leq \beta \leq g_\alpha(a)\}$.

To determine the miminax linear risk S, we proceed systematically by first determining $S_j = \inf_{(\alpha,\beta)\in I_j} S(\alpha,\beta)$ for $j=1,\ldots,5$, and then $S = \min\{S_j : j=1,2,3,4,5\}$ given that $\bigcup_{j=1}^{5} I_j = \{(\alpha,\beta): \alpha \leq 1, \beta \geq 0\}$. This work, which is illustrated in Section 6.3, leads to the following lemma, Lemma 3.4, the statement of which is deferred to the Appendix because of its technical nature. It also provides the key to Theorem 3.5 where the linear minimax rule and risk are determined.

THEOREM 3.5

$$\textit{Given } \alpha_* = \frac{n^{1/2}}{n^{1/2} + \{\frac{|a+b-1|}{b-a}\}^{1/2}}, \quad K_1 = \frac{(1-a-b)^2}{n^2(b-a)^2}, \quad K_2 = \frac{2a(1-a)+2b(1-b)}{n(b-a)^2},$$

$$H(\gamma) = \frac{(b-a)^2}{4}\,\frac{K_1\gamma^4 + K_2\gamma^2 + 1}{(1+\gamma)^2}, \quad H_- = \inf_{\gamma \leq n^{1/2}} H(\gamma), \quad I(\gamma) = K_1\gamma^4 + 2K_1\gamma^2 + K_2\gamma - 1,$$

and γ_r*, the unique positive root of* $I(\cdot)$;

then under quadratic loss, the linear minimax risk is given by either:

(i) $\min\{H(\gamma_r), b(1-b)\alpha_*^2/n\}$ whenever $b \le 1/2$;

(ii) $\min\{H(\gamma_r), a(1-a)\alpha_*^2/n\}$ whenever $a \ge 1/2$;

(iii) $H(\gamma_r)$ whenever $a<1/2$, $b>1/2$ and $n<\{[(2b-1)(1-2a)]^{-1}-1\}^2$; or

(iv) $1/[4(n^{1/2}+1)^2]$ whenever $a<1/2$, $b>1/2$ and $n \ge \{[(2b-1)(1-2a)]^{-1}-1\}^2$.

The minimax linear rule $\delta_{\alpha,\beta}$ is given by either $\alpha=\alpha_r$, $\beta=g_{\alpha_r}((a+b)/2)$; $\alpha=\alpha_*, \beta=b(1-\alpha_*)$; $\alpha=\alpha_*, \beta=a(1-\alpha_*)$; or $\alpha=\alpha_0$, $\beta=(1-\alpha_0)/2$; depending whether the minimax risk is $H(\gamma_r)$, $b(1-b)\alpha_*^2/n$, $a(1-a)\alpha_*^2/n$, or $1/[4(n^{1/2}+1)^2]$ respectively.

Proof. Our proof is deferred to Section 6.3.

The above yields a complete solution for all triplets (n,a,b). As a consequence, we obtain the following results for the particular cases where a=0 or where a+b=1. (Note that we have retained the notation of Lemma 3.3 for our next two results)

COROLLARY 3.6 Whenever a=0 and the loss is quadratic, the linear minimax risk with corresponding rule $\delta_{\alpha,\beta}$ is given by:

$$(i)\ \frac{b^2}{[1+(\frac{nb}{1-b}+1)^{1/2}]^2},\ \alpha=\alpha_r=1-(\frac{1-b}{nb+1-b})^{1/2},\ \beta=g_{\alpha_r}(\frac{b}{2}),\ if\ b\le\frac{n^{1/2}+2}{2n^{1/2}+2};$$

$$or\ \ (ii)\ \frac{1}{4(n^{1/2}+1)^2},\ \alpha=\alpha_0,\ \beta=\frac{(1-\alpha_0)}{2},\ if\ b\ge\frac{n^{1/2}+2}{2n^{1/2}+2}.$$

Proof. First, whenever $b \ge (n^{1/2}+2)/(2n^{1/2}+2)$, result (ii) follows directly from (iv) of Theorem 3.5. On the other hand, whenever $b<(n^{1/2}+2)/(2n^{1/2}+2)$, we note that the function H of Lemma 3.4 simplifies to $H(\gamma)=(\{b(c\gamma^2+1)\}/\{2(\gamma+1)\})^2$ with $c=(1-b)/nb$. We may then directly establish that $H(\cdot)$ has a minimum of $H(\gamma_r)$ where $\gamma=\gamma_r=[(1+c)/c]^{1/2}-1$, which coincides with the quantities given in (i). Finally, part (i) of the Corollary follows from parts (ii) and (iii) of Theorem 3.5 as well the result $H(\gamma_r) \le \{b(1-b)\alpha_*^2\}/n=\{b(1-b)\}/\{n(1+\sqrt{c})^2\}$, which is readily verified.

COROLLARY 3.7 Whenever a+b=1 and the loss is quadratic, the linear minimax risk and rule $\delta_{\alpha,\beta}$ are given by:

(i) $\{\{(1/4)-b(1-b)\}^{-1}+n\{b(1-b)\}^{-1}\}^{-1}$, $\alpha=\alpha_r=K_2/(K_2+1)$, $\beta=(1-\alpha_r)/2$, if $b \le \{1 + (n^{1/2}+1)^{1/2}\}^{-1}/2$;

(ii) $1/[4(n^{1/2}+1)^2]$, $\alpha=\alpha_0$, $\beta=(1-\alpha_0)/2$ i

Proof. The results are obtained fron

whenever a+b=1, $K_1=0$, $K_2=\{4b(1-b$

minimax risk is given by $H(K_2^{-1})$ wh

The last three results suggest that v

the sample size parameter n is larg

parameter space coincides with the li

this sense, the linear class is not alwa

$\theta \in [a,b]$. Graphs of the minimax line

3.2 Normalized Loss

In this section, we proceed in a sim

The risk function of the linear rule δ

$$R_1\,[0<\theta<1]+R_2\,[\theta=0,\beta=0$$

$$where\ R_1=\frac{\alpha^2}{n}-(\alpha-1)^2+\frac{\beta^2}{\theta}+\frac{(1-}{(1}$$

We shall first consider the cases whe

THEOREM 3.8 Whenever (i) a=0, o

is given by (i) $\alpha=\alpha^*=(nb)/(1-b+nb)$, β=

associated minimax risk is, in both ca

Proof. Both parts are analogous and

parameter space, we are forced by (3.

$$R^*(\theta$$

which is increasing in θ on [0,b]. Th

(ii) $1/[4(n^{1/2}+1)^2]$, $\alpha=\alpha_0$, $\beta=(1-\alpha_0)/2$ if $b\geq\{1+(n^{1/2}+1)^{1/2}\}^{-1}/2$.

Proof. The results are obtained from parts (iii) and (iv) of Theorem 3.5, and by observing that, whenever $a+b=1$, $K_1=0$, $K_2=\{4b(1-b)\}/\{n(2b-1)^2\}$, $I(\gamma)=K_2\gamma-1$, $\gamma_r=1/K_2$, $\alpha_r=K_2/(K_2+1)$, and the minimax risk is given by $H(K_2^{-1})$ which coincides with the quantity in (i).

The last three results suggest that when either (i) the parameter space [a,b] is large; and/or (ii) the sample size parameter n is large, the linear minimax risk (and rule) for the restricted parameter space coincides with the linear minimax risk (and rule) for the unrestricted case. In this sense, the linear class is not always flexible enough to profitably incorporate the information $\theta\in[a,b]$. Graphs of the minimax linear risk, as well as comparisons, are presented in Section 5.

3.2 Normalized Loss

In this section, we proceed in a similar fashion as above but consider normalized loss instead. The risk function of the linear rule $\delta_{\alpha,\beta}$ is now given by $R^*(\theta,\delta_{\alpha,\beta})=$

$$R_1[0<\theta<1]+R_2[\theta=0,\beta=0 \text{ } or \text{ } \theta=1,\alpha+\beta=1]+R_3[\theta=0,\beta\neq0 \text{ } or \text{ } \theta=1,\alpha+\beta\neq1],$$

$$where\ R_1=\frac{\alpha^2}{n}-(\alpha-1)^2+\frac{\beta^2}{\theta}+\frac{(1-\alpha-\beta)^2}{(1-\theta)},\ R_2=\frac{\alpha^2}{n},\ and\ R_3=\infty. \tag{3.3}$$

We shall first consider the cases where either (i) a=0, or (ii) b=1.

THEOREM 3.8 Whenever (i) a=0, or (ii) b=1, the linear minimax rule, under normalized loss, is given by (i) $\alpha=\alpha^*=(nb)/(1-b+nb)$, $\beta=\beta^*=0$; or (ii) $\alpha=\alpha^*=\{n(1-a)\}/\{(a+n(1-a)\}$, $\beta=\beta^*=1-\alpha^*$. The associated minimax risk is, in both cases, equal to α^*/n.

Proof. Both parts are analogous and we consider only (i). Since the value $\theta=0$ belongs to the parameter space, we are forced by (3.3) to set $\beta=0$. Furthermore, again from (3.3),

$$R^*(\theta,\delta_{\alpha,0})=\frac{\alpha^2}{n}+(\alpha-1)^2\frac{\theta}{1-\theta},$$

which is increasing in θ on [0,b]. Therefore,

$$\sup_{\theta\leq b} R^*(\theta,\delta_{\alpha,0}) = \frac{\alpha^2}{n} + (\alpha-1)^2\frac{b}{1-b},$$

which has a minimum value of α^*/n attained at $\alpha=\alpha^*$.

For the cases in which a>0 and b<1, a careful analysis leads to the following and final result of this chapter, which is proven in Section 6.4. Graphs of the minimax linear risk, as well as comparisons, are presented in Section 5.

THEOREM 3.9 Whenever a>0, b<1, and the loss is normalized, the linear minimax rule $\delta_{\alpha,\beta}$ is given by $\beta=\beta_m=(m+1)/\{n(ab)^{-1}-(n-1)(m+1)^2\}$, $\alpha=\alpha_m=1-(m+1)\beta_m$; with minimax risk equal to $H/(1-H+nH)$; where $m=\{(1-a)(1-b)/(ab)\}^{1/2}$ and $H=1-\{(ab)^{1/2}+[(1-a)(1-b)]^{1/2}\}^2$.

4. THE MLE AND ITS COMPARATIVE PERFORMANCE

4.1 Introduction

For our problem, the maximum likelihood estimator may be viewed as a natural choice because it is simple to use and given by $\delta_{mle}(x)=a[nx\leq a]+(x/n)[a<nx<b]+b[nx\geq b]$; the projection of the unbiased estimator on the restricted parameter space. It has long been known, however, that maximum likelihood estimators are often inadmissible under quadratic loss for restricted parameter spaces (e.g. Sacks [24]). Charras and van Eeden [7] in a general framework of compact parameter spaces establish, in a large number of cases, the inadmissibility of the members of the so called "boundary" estimators, which so frequently includes the mle. They also deal specifically with the Binomial problem, and establish **(1)** the inadmissibility of the mle whenever a>0 and b<1, and **(2)** the admissibility of the mle whenever a=0 or $b\leq 2/n$ (or equivalently b=1, $a\geq 1-(2/n)$) for quadratic loss and normalized loss. Furthermore, Funo [11] established the inadmissibility of the mle whenever a=0, b>2/n (or equivalently b=1, a<1-(2/n)). Some of our results do apply directly to the mle, and we obtain as a consequence the following.

COROLLARY 4.1. (a) Whenever the parameter space is of the form $\theta\leq b$; with $0\leq b\leq 1/n$; the mle is unique Bayes (for a 1 point prior at b) under normalized loss. Furthermore, the mle is

E. Marchand - B. MacGibbon
minimax whenever $b\leq b^*(n)$, as defin
(b) Whenever the parameter space
Bayes (for limiting prior on {0,b}) u
and the numerical evidence whic
$b^*(n)\leq b\leq b^{**}(n)$.

Since the mle is inadmissible, for b
and b<1, further investigation is requ
(ii) to search for "good" estimators th
at some of the historical development
a N(θ,1) distribution, under quadratic
x[x>0], is inadmissible (see Brown [5
Bayes estimator with respect to the u
admissible but not a dominator of the
improvements to the mle, but the que
remains unknown.

With regards to **(i)**, it is instructi
θ∈[a,b], where the competitor may be,
Or again, in the spirit of the minimax c
with the minimax risk, or max{R(θ,δ
a≤θ≤b, (with known variance), Casella
quadratic loss, the mle whenever b-a
functions. Gatsonis, MacGibbon, and S
estimator to the mle, and compared nu
others such as the minimax estimator a
the risk functions could cross and tha
large part of the parameter space. Ho
uniform Bayes estimator even domina
inadmissibility property of the mle is
study the mle under alternative loss

minimax whenever $b \leq b^*(n)$, as defined in Remark 2.2.

(b) Whenever the parameter space is of the form $\theta \leq b$; with $1/n \leq b \leq 2/n$; the mle is extended Bayes (for limiting prior on $\{0,b\}$) under normalized loss. Furthermore, based on Theorem 2.5 and the numerical evidence which we have obtained, the mle is minimax whenever $b^*(n) \leq b \leq b^{**}(n)$.

Since the mle is inadmissible, for both quadratic and normalized loss functions, whenever $a>0$ and $b<1$, further investigation is required: **(i)** to examine how serious this inadmissibility is; and **(ii)** to search for "good" estimators that dominate the mle. For instance, it is instructive to look at some of the historical developments for the problem of estimating the nonnegative mean θ of a $N(\theta,1)$ distribution, under quadratic loss, based on an observation x. Here, the mle, given by $x[x>0]$, is inadmissible (see Brown [5]). Katz [16] showed that both the mle and the generalized Bayes estimator with respect to the uniform prior are minimax, with the latter estimator being admissible but not a dominator of the mle. Recently, Shao and Strawderman [25] found explicit improvements to the mle, but the question of the admissibility of some of these improvements remains unknown.

With regards to **(i)**, it is instructive to compare the functions $R(\theta,\delta_{mle})$ and $R(\theta,\delta_{competitor})$, $\theta \in [a,b]$, where the competitor may be, for instance, $\delta_{minimax}$, a $\delta_{dominator}$, or δ_{Bayes} for a chosen prior. Or again, in the spirit of the minimax criterion, it is instructive to compare $\max\{R(\theta,\delta_{mle}), a \leq \theta \leq b\}$ with the minimax risk, or $\max\{R(\theta,\delta_{competitor}), a \leq \theta \leq b\}$. For the normal bounded mean problem, $a \leq \theta \leq b$, (with known variance), Casella and Strawderman [6] showed that $\delta_{minimax}$ dominates, under quadratic loss, the mle whenever $b-a \leq 2$, and gave several numerical comparisons of the risk functions. Gatsonis, MacGibbon, and Strawderman [12] compared analytically the uniform Bayes estimator to the mle, and compared numerically the risk functions of these two estimators with others such as the minimax estimator and Bickel's [4] estimators. They showed analytically that the risk functions could cross and that the uniform Bayes estimator dominates the mle over a large part of the parameter space. However, as one of their graphs show (1a), it seems that the uniform Bayes estimator even dominates the mle whenever b-a is small. Recognizing that the inadmissibility property of the mle is an attribute of the loss function, it is also of interest to study the mle under alternative loss functions. Interestingly, for the normal bounded mean problem (with known variance), Iwasa and Moritani [13] showed that the mle, which is inadmissible for quadratic loss, is a proper Bayes rule, and thus admissible, for absolute value loss.

With respect to **(ii)**, Charras and van Eeden's [7] proof of the inadmissibility of boundary estimators in compact parameter spaces, that applies to our mle, does not involve the specification of explicit dominators. Later on in their paper, dominators of the mle are exhibited. These estimators, applied to our problem, of the form $\delta_{mle}(x)+\varepsilon_1[x\leq na]+\varepsilon_2[x\geq nb]$, may not be admissible. Also, again in the spirit of **(ii)**, Kempthorne [18] presented a general algorithm that yields, among the class of dominators of the mle, an optimal estimator chosen according to the "maximin improvement" criterion.

In the next section, we give some answers to **(ii)**, and some comparisons, as suggested in **(i)**, with members of a particular class of rules. Further comparisons are presented in Section 5.

4.2 Some Improvements to the MLE and a Class of Bayes Rules for Beta Priors

Our research objective has been to search for admissible improvements to the mle, and learn about the nature of these improvements. We present below some partial results. We have focused on minimax estimators for the symmetric case $\theta\in[a,1-a]$ as did previous researchers. Berry [2] gave under quadratic loss (exact, analytical) least favourable priors for n=1,2,3 and 4, and also for n=5,6 when a is small. Moors [22] gave them for n=1,2,3 in his Table 7.9, p.129. He also has numerical results for $4\leq n\leq 16$ in his tables on p. 146-159. Further, Moors has numerical results for normalized loss for $3\leq n\leq 10$ in his tables on p. 161-178. Here we consider for comparative purposes not only minimax estimators but also the class of Bayes rules for symmetric and translated Beta priors.

Note that for a given prior G on [a,b], 0<a<b<1, the Bayes rule for quadratic loss is

$$\delta^G(x) = \frac{E^G\{\theta^{x+1}(1-\theta)^{n-x}\}}{E^G\{\theta^x(1-\theta)^{n-x}\}}, \tag{4.1}$$

which illustrates the fact that the Bayes rule depends on the prior G only through the first n+1 moments $E(\theta^r)$, r=1,...,n+1, as previously indicated by Moors [22] (p. 97, Theorem 5.11). Similarly, for normalized loss, the dependence is, as previously indicated by Moors [22] (p. 141,

formula 7.34) through $E(\theta^r)$, r=-1,1,..
the mle which are Bayes rules is repr
moments. For instance, it is instruc
symmetric (or invariant) priors.

THEOREM 4.2. For the cases in whi
where G is symmetric about 1/2, dom
and $c_2(a)=\{(1-2a)(1-2a-2a^2)\}/\{4(1+2a$
Proof. Under the conditions of the th
with a symmetric prior are equivaria
of the form $\delta_\beta(x)=\beta[x=0]+(1-\beta)[x=1]$
that $R(\theta,\delta_{mle})-R(\theta,\delta_\beta)=(a-\beta)(a+\beta-4\theta($
$a\leq\beta\leq 3a-4a^2$.[3] The result now follows
which becomes for n=1 and since E^G
rules are of the form $\delta_\beta(x)=\beta[x=0]+$
$R(\theta,\delta_{mle})-R(\theta,\delta_\beta)=(\beta-a)\{-2\theta^2(a+\beta+1)+$
as a function of θ, iff $a\leq\beta\leq\{a(1-2a^2)$
on (a,β), as well representation (4.1);
$(\Rightarrow E(\theta^3)=3/2E(\theta^2)-1/4)$, $\beta=\delta^G(0)=(4V$

We now consider a particular fami
Beta priors often used for the unres
quadratic and normalized losses, we
where $\theta\overset{d}{=}(1-2a)Y+a$ with Y~Beta(ε
associated posterior distributions, has
As a consequence of Theorem 4.1, v

[3] This was previously shown by Moors
and van Eeden [7] (p. 126-127).

formula 7.34) through $E(\theta^r)$, r=-1,1,...n. Thus, at least conceptually, the class of dominators of the mle which are Bayes rules is representable in terms of appropriate restrictions on the relevant moments. For instance, it is instructive to consider the cases $\theta \in [a,1-a]$, with n=1 or 2, and symmetric (or invariant) priors.

THEOREM 4.2. For the cases in which $\theta \in [a,1-a]$, n=1 or 2, the Bayes rule δ^G for quadratic loss, where G is symmetric about 1/2, dominates the mle iff $Var^G(\theta) \geq c_n(a)$ where $c_1(a)=(1-2a)(1-4a)/4$, and $c_2(a)=\{(1-2a)(1-2a-2a^2)\}/\{4(1+2a^2-4a^3)\}$.

Proof. Under the conditions of the theorem, both the mle and an arbitrary Bayes rule associated with a symmetric prior are equivariant (i.e. $\delta(x)=1-\delta(n-x)$). For n=1, our equivariant rules are of the form $\delta_\beta(x)=\beta[x=0]+(1-\beta)[x=1]$; $a \leq \beta \leq 1/2$. It is easy to see that $R(\theta,\delta_\beta)=\beta^2+\theta(1-\theta)(1-4\beta)$, that $R(\theta,\delta_{mle})-R(\theta,\delta_\beta)=(a-\beta)(a+\beta-4\theta(1-\theta))$. This difference is nonnegative, a function of θ, iff $a \leq \beta \leq 3a-4a^2$.[3] The result now follows from this restriction on (a,β), as well representation (4.1), which becomes for n=1 and since $E^G(\theta)=1/2$, $\beta=\delta^G(0)=(1/2)-2Var^G(\theta)$. For n=2, our equivariant rules are of the form $\delta_\beta(x)=\beta[x=0]+1/2[x=1]+(1-\beta)[x=2]$; $a \leq \beta \leq 1/2$. Here, one may verify that $R(\theta,\delta_{mle})-R(\theta,\delta_\beta)=(\beta-a)\{-2\theta^2(a+\beta+1)+2\theta(a+\beta+1)-(a+\beta)\}$, and that this difference is nonnegative as a function of θ, iff $a \leq \beta \leq \{a(1-2a^2)\}/\{1-2a+2a^2\}$. Our result now follows from this restriction on (a,β), as well representation (4.1); which becomes for n=2 and since $E^G(\theta)=1/2$, $E^G(\theta-1/2)^3=0$ $(\Rightarrow E(\theta^3)=3/2E(\theta^2)-1/4)$, $\beta=\delta^G(0)=(4Var^G(\theta)+1)^{-1}-(1/2)$.

We now consider a particular family of Bayes rules. In analogy with the conjugate family of Beta priors often used for the unrestricted problem, which yields minimax solutions for both quadratic and normalized losses, we consider the class of translated and symmetric Beta priors, where $\theta \stackrel{d}{=} (1-2a)Y+a$ with $Y \sim Beta(\varepsilon,\varepsilon)$, $\varepsilon>0$. This class of priors, and approximations to the associated posterior distributions, has been previously considered by Winkler and Franklin [29]. As a consequence of Theorem 4.1, we now obtain the following.

---

[3] This was previously shown by Moors [22] (p. 27), and is also equivalent to the condition given by Charras and van Eeden [7] (p. 126-127).

COROLLARY 4.3. For the cases where $\theta \in [a,1-a]$, n=1 or 2, the Bayes rule δ^{ε} for quadratic loss associated with a translated Beta(ε,ε) prior on [a,1-a] dominates the mle whenever $a \geq a_n$ or $\varepsilon \leq \varepsilon_n(a)$, where $a_1=1/4$, $a_2=(3^{1/2}-1)/2$, $\varepsilon_1(a)=\{\{(1-2a)/(1-4a)\}^{1/2}-1\}/2$, and $\varepsilon_2(a)=\{\{(1-2a)(1+2a^2-4a^3)/(1-2a-2a^2)\}^{1/2}-1\}/2$.

Proof. The result follows from Theorem 4.1, and the variance of our translated Beta(ε,ε) prior which is equal to $\{(1-2a)/(4\varepsilon+2)\}^2$.

Note that the values a_n are derived from the fact that for small parameter spaces, the c_n bounds in Theorem 4.2 are negative, indicating that, in these situations, all the Bayes rules considered dominate the mle. For the normalized case, the results that follow below, which are proven in Section 6.5, are analogues to Theorem 4.2 and Corollary 4.3. Note that both losses are fundamentally linked since, in fact for any parameter space with $0<a<b<1$, (i) the dominators given above for quadratic loss remain dominators under normalized loss, and vice-versa; (ii) the admissibility of a given estimator for quadratic loss implies the admissibility for normalized loss, and vice-versa; and (iii) a Bayes estimator with respect to the prior Q under normalized loss is necessarily Bayes under quadratic loss for the prior G where, for a measurable set A,

$$P^G(\theta \in A) = \int_A dG(\theta) = k \int_A \frac{dQ(\theta)}{\theta(1-\theta)} ; where\ k^{-1} = \int_\Theta \frac{dQ(\theta)}{\theta(1-\theta)} .$$

This last connection is specific to our Binomial problem and follows readily from the representation of the Bayes rule with respect to prior Q and for general loss $L(\theta,d)=w(\theta)(d-\theta)^2$, which is given by $\delta^Q(x)=E[\theta w(\theta)\,|\,x]/E[w(\theta)\,|\,x]$.

THEOREM 4.4. For the cases where $\theta \in [a,1-a]$, n=1 or 2, the Bayes rule δ^G for normalized loss, where G is symmetric about 1/2, dominates the mle iff $E^G(1/\theta) \geq d_n(a)$ where $d_1(a)=1/(3a-4a^2)$, and $d_2(a)=(1+2a^2-4a^3)/(2a-4a^3)\}$.

COROLLARY 4.5. For the cases where $\theta \in [a,1-a]$, n=1 or 2, the Bayes rule δ^{ε} for normalized loss associated with a translated Beta(ε,ε) prior on [a,1-a] dominates the mle whenever $a \geq a_n^*$ or $\varepsilon \leq \varepsilon_n^*(a)$, where $a_1^*=1/4$, $a_2^*=(3^{1/2}-1)/2$, and $\varepsilon_n^*(a)$ is the positive solution of the equation ${}_2F_1(1,\varepsilon;2\varepsilon;(1-2a)/(1-a))=(1-a)d_n$; d_n as defined in Theorem 4.4.

For the sake of completeness, we turn briefly to the comparison mle vs minimax. For a given triplet (n,Θ,loss), the minimax estimator will rarely dominate the mle, but it does seem to dominate the mle when the parameter space Θ is relatively small (our numerical results show that the larger the sample size n, the smaller the parameter space Θ has to be). However, for the cases n=1, and n=2, we have the following conditions for dominance.

COROLLARY 4.6. **(a)** Under quadratic loss, for the parameter space $\theta\in[a,1-a]$; $0<a<1/2$; the minimax estimator dominates the mle whenever n=1 and $a\geq(3-5^{1/2})/8(\approx.0954915)$, and n=2, $a>a_0\approx.159671)$. **(b)** Under normalized loss, and for n=1 or n=2, the minimax estimator dominates the mle for all parameter spaces $\theta\in[a,1-a]$; $0<a<1/2$.

Proof. (a) For the cases n=1 and n=2, and our symmetric parameter space, equivariant estimators are linear and, thus, the minimax and linear minimax estimators coincide. From Corollary 3.7, we obtain for n=1, as a function of a, $\beta=\delta_{minimax}(0)=1/4[a<(2-2^{1/2})/4]+(1-2a)^2/2[a\geq(2-2^{1/2})/4]$. Now, for n=1, the result follows from the dominance criteria $a\leq\beta\leq3a-4a^2$ within Theorem 4.2. Similarly, for n=2, $\beta=\delta_{minimax}(0)=f/2[a<(1-f^{1/2})/2]+(1-2a)^2/\{2-4a(1-a)\}[a\geq(1-f^{1/2})/2]$; where $f=2^{1/2}-1$. In this case, the dominance criteria within Theorem 4.2 is $a\leq\beta\leq\{a(1-2a^2)\}/\{1-2a+2a^2\}$, and becomes equivalent to $2a^3+fa^2-a(f+1)+f/2\leq0\Leftrightarrow a\geq a_0$ on the range $0<a<1/2$.

(b) Here, as well, all equivariant rules are linear and, thus, the minimax and linear minimax estimators coincide. We have, from Theorem 3.9, as a function of a, $\beta=\delta_{minimax}(0)=2a(1-a)$ for n=1, and $\beta=\delta_{minimax}(0)=\{a(1-a)\}/\{1-(2a(1-a)\}$ for n=2. It is readily verified that the conditions for dominance within Theorem 4.2 hold for all $0<a<1/2$.

5. DISCUSSION AND CONCLUSIONS

As shown in Figure 1, if n and b are sufficiently small, then the least favourable distribution on [0,b] is a two-point prior under quadratic loss, and a one-or two-point prior under normalized loss. For higher values of n and b, however, the specification of the minimax estimator quickly becomes unwieldly and alternative estimation procedures must be considered. The resulting risks of the estimators obtained must then be compared to the minimax risk. Figures 2,3,4, and 5 look at the ratios of linear minimax risk and the maximum risk of the mle to the minimax risk in various situations. For both quadratic and normalized loss, the linear minimax estimator tends to outperform the mle on both the intervals [a,1-a] and [0,b]. For small n, small a, and normalized loss the situation is reversed. On the interval [0,b], Figure 3 indicates that when b is not too large and as n increases the three estimators become indistinguishable. However, for $n \geq 20$ there are significant differences for quadratic loss between the mle and the minimax estimators if $b \geq 0.5$. These differences are even more striking on the interval [a,1-a] (see Figure 2). Figures 4 and 5 examine the inverse of these ratios as a function of a and b respectively.

In Figures 6 to 10, the mle is compared with other estimators. The minimax estimator usually dominates the mle over the middle part of the parameter space while the mle excels on the boundary. The linear minimax rule[4], a clearly usable rule and much less complex than the minimax estimator, also dominates the mle in the middle of the parameter space, although its behaviour on the boundary is worse than that of the minimax estimator. As previously established in Section 3, these figures clearly indicate that if either the sample size or the interval is too large then the linear minimax risks and rules coincide for the constrained and unconstrained parameter spaces. A more flexible class of estimators are the Bayes rules associated with symmetric and translated Beta priors on [a,1-a]. Figures 6,7, and 10 indicate that the uniform Bayes estimator δ^1 exhibits a substantial gain over a large and central part of the parameter space. The mle performs better however when θ is near the boundary of the parameter space in the symmetric case and near one of the two boundaries in the asymmetric case. The estimator $\delta^{1/2}$, based on a more disperse Beta prior than δ^1, performs more like the mle, and seems very attractive in comparison, with a gain over the middle part of the parameter space, and a less severe loss near the extremes. By the choice of an even more disperse prior, we actually have identified some members of our translated Beta class, such as $\delta^{.19}$, that dominate the mle for a=0.25 and n=10 under quadratic loss (and consequently normalized loss). Finally, using an overly dispersed prior, such as $\varepsilon=.10$, does not lead to dominance. Let $\mathfrak{C}_{a,n}=\{\varepsilon:\delta^\varepsilon$ dominates δ_{mle} under quadratic loss$\}$. Numerically, we have explored these sets and obtained $\mathfrak{C}_{.35,10}\approx[0,.114]$, $\mathfrak{C}_{.30,10}\approx[.036,.091]$, $\mathfrak{C}_{.25,.10}\approx[.1834,.1924]$, $\mathfrak{C}_{.20,5}\approx[.178,.180]$, and $\mathfrak{C}_{a,10}=\varnothing$ for $a \leq .20$. In part, these relationships are illustrated in Figure 10. If either the parameter space or sample size n is large, this tends to shrink the set $\mathfrak{C}_{a,n}$, and it is empty in many cases. These numerical

[4] In fact, we have used its truncated version in Figures 6-10.

results suggest that the class of trans

of finding a dominator of the mle. N

δ^ε seems quite attractive in oppositio

as its study under normalized loss, w

In conclusion, minimax estimators

of a constrained binomial proportion

large, the class of linear minimax es

with symmetric translated Beta prio

estimate.

6. APPENDIX

6.1. Theorem 2.3 (Proof of convexi

Note that $b^*(n) \leq n^{-1}$, and it thus suff

we may express $R(\theta,\delta_b)$ as $A(\theta)+B(\theta)$

^{1}], it will suffice to show that $A(\cdot)$ is

for n>2, we expand $(1-\theta)^{n-1}$ and rearr

$$A(\theta) = \frac{b^2}{\theta}[1-(1$$

$$with \;\; c_k = b^2 \Big(n- \; k+$$

$$and \;\; , \frac{\partial^2 A(\theta)}{\partial^2\theta} = \sum$$

$$with \;\; a_k = k(k-2)c$$

Comparing two consecutive terms in

we obtain for $\theta \in (0,n^{-1}]$

results suggest that the class of translated beta priors is often not large enough for the purpose of finding a dominator of the mle. Nevertheless, there are several instances where a given rule δ^{ε} seems quite attractive in opposition with the mle. A more extensive analysis of $\mathfrak{C}_{a,n}$, as well as its study under normalized loss, would be of interest.

In conclusion, minimax estimators serve as a useful benchmark for measuring other estimators of a constrained binomial proportion. If n or the interval containing the parameter is not too large, the class of linear minimax estimators (truncated) or the class of Bayes rules associated with symmetric translated Beta priors is recommended. Otherwise, the mle is a very good estimate.

6. APPENDIX

6.1. Theorem 2.3 (Proof of convexity of $R(\cdot;\delta_b)$ on $(0,b^*(n)]$)

Note that $b^*(n) \leq n^{-1}$, and it thus suffices to verify that $R(\cdot,\delta_b)$ is convex on $(0,n^{-1}]$. From (2.4), we may express $R(\theta,\delta_b)$ as $A(\theta)+B(\theta)$ where $B(\theta)=-1+(1-b)^2/(1-\theta)$. Since $B(\cdot)$ is convex on $(0,n^{-1}]$, it will suffice to show that $A(\cdot)$ is convex on $(0,n^{-1}]$. For n=2, $A(\cdot)$ is linear and thus convex; for n>2, we expand $(1-\theta)^{n-1}$ and rearrange terms to obtain

$$A(\theta) = \frac{b^2}{\theta}[1-(1-\theta)^{n-1}] + 2b(1-\theta)^{n-1} = \sum_{k=0}^{n-1}(-1)^k c_k \theta^k,$$

$$\textit{with} \quad c_k = b^2\binom{n-1}{k+1} + 2b\binom{n-1}{k},\ k=0,\dots,n-2,\ c_{n-1}=2b;$$

$$\textit{and} \ ,\frac{\partial^2 A(\theta)}{\partial^2\theta} = \sum_{k=2}^{n-1}(-1)^{k-2} a_k \theta^{k-2}, \tag{6.1}$$

$$\textit{with} \ a_k = k(k-2)c_k,\ k=2,\dots n-1.$$

Comparing two consecutive terms in this last sum and using the fact that $a_{k+1} \leq n a_k$; k=2,...,n-2; we obtain for $\theta \in (0,n^{-1}]$

$$|(-1)^{k-1}a_{k+1}\theta^{k-1}| \leq \frac{a_{k+1}}{n}\theta^{k-2} \leq |(-1)^{k-2}a_k\theta^{k-2}|.$$

Finally, starting with the first term (i.e. k=2) of the sum in (6.1) which is positive, the series telecospes into a sum of positive terms and this establishes the desired result.

6.2 *Proof of Theorem 2.5*

(a) For n=2, equation (2.10) implies $z^*=(1+b)/4$. Furthermore $R^*(\theta,\delta_0^*)$, given in (2.8) with n=2 and $z^*=(1+b)/4$, is easily seen to be convex thus establishing (2.7) and the desired result.

(b) From (2.8) with n=3, we obtain

$$\frac{\partial R^*(\theta,\delta_0^*)}{\partial\theta}\Big|_{\theta=0} = 1-6z^*+3b^2-3z^{*2},$$

$$\text{and}\quad \frac{\partial^2 R^*(\theta,\delta_0^*)}{\partial^2\theta} = 12z^*-8b+\frac{2(1-b)^2}{(1-\theta)^3}.$$

We first observe that the second derivative is increasing on [0,b], and positive at $\theta=b$ as for $0\leq b\leq 1$, $8b\leq 2/(1-b)$. Given this behaviour and condition (2.9), a necessary and sufficient condition for δ_0^* to be minimax is the non-positivity of the first derivative of $R(\theta,\delta_0^*)$ with respect to θ evaluated at $\theta=0$, or equivalently $3(z^*)^2+6z^*+1\geq 3b^2$. Finally, by using the only admissible candidate implied by (2.10), that is $z^*=-(1-b)+\{(1-b)(4-b)/3\}^{1/2}$, this last inequality is verified for $b\leq(25-2\sqrt{93})/11$ only. Since we are only investigating values of $b>b^*(3)=(5-\sqrt{21})/2$, the proof is complete.

6.3 *Proof and statement of Lemma 3.4; proof of Theorem 3.5*

LEMMA 3.4. With α_*, H, H_-, I, and γ_I defined in Theorem 3.5, then

(i) $S_1 = H\_[a\leq$

(ii) $S_2 = \frac{1}{4(\sqrt{n}+}$

(iii) $S_3 = H\_[b\geq$

(iv) $S_4 = \frac{1}{4(\sqrt{n}+}$

and, (*v*)

Proof. Our proof is somewhat tedio

(i) For $(\alpha,\beta)\in I_1$, we have from (3.1

$S(\alpha,\beta)$

For fixed α, $S(\alpha.\cdot)$ is minimized on

$a(1-\alpha)\geq g$

we have, for $a\geq 1/2$,

$S_1 = \min\{\ in$

On the other hand, for $a\leq 1/2$,

By setting $\gamma=\alpha/(1-\alpha)$ and by rea

$S(\alpha,g_\alpha(a+b/2))$ as $H(\gamma)$ and, as wel

$I(\cdot\cdot)$ varies from - to + on $[0,\infty)$, the

$\alpha=\min(\alpha_r,\alpha_0)$ respectively with $\alpha_r=\gamma_r$

$$(i)\ S_1 = H_-[a\le 1/2] + \min\{a(1-a)\frac{\alpha_*^{\,2}}{n}, H(\gamma_r)\}[a>1/2];$$

$$(ii)\ S_2 = \frac{1}{4(\sqrt{n}+1)^2}[a\le 1/2] + \frac{a(1-a)}{(\sqrt{n}+1)^2}[a>1/2];$$

$$(iii)\ S_3 = H_-[b\ge 1/2] + \min\{b(1-b)\frac{\alpha_*^{\,2}}{n}, H(\gamma_r)\}[b<1/2];$$

$$(iv)\ S_4 = \frac{1}{4(\sqrt{n}+1)^2}[b\ge 1/2] + \frac{b(1-b)}{(\sqrt{n}+1)^2}[b<1/2];$$

$$and,\ (v)\ S_5 \ge \frac{1}{4(\sqrt{n}+1)^2}.$$

Proof. Our proof is somewhat tedious and, hence, we sketch the main lines only.

(i) For $(\alpha,\beta)\in I_1$, we have from (3.1) and Lemma 3.3,

$$S(\alpha,\beta) = [\beta - a(1-\alpha)]^2 + \frac{\alpha^2}{n}a(1-a). \qquad (6.2)$$

For fixed α, $S(\alpha,\cdot)$ is minimized on I_1 by choosing $\beta=\max\{a(1-\alpha),g_\alpha((a+b)/2)\}$. Since

$$a(1-\alpha)\ge g_\alpha(\frac{a+b}{2}) \Leftrightarrow a\ge 1/2 \ and\ \alpha_*\le\alpha\le\alpha_0,$$

we have, for $a\ge 1/2$,

$$S_1 = \min\{\inf_{\alpha_*\le\alpha\le\alpha_0} S(\alpha,a(1-\alpha)),\ \inf_{\alpha\le\alpha_*} S(\alpha,g_\alpha(\frac{a+b}{2})\}. \qquad (6.3)$$

On the other hand, for $a\le 1/2$,

$$S_1 = \inf_{\alpha\le\alpha_0} S(\alpha,g_\alpha(\frac{a+b}{2}). \qquad (6.4)$$

By setting $\gamma=\alpha/(1-\alpha)$ and by rearranging terms, we may express the pivotal quantity $S(\alpha,g_\alpha(a+b/2))$ as $H(\gamma)$ and, as well, establish that $\mathrm{sgn}\{\frac{\partial}{\partial\alpha}S(\alpha,g_\alpha((a+b)/2)\}=\mathrm{sgn}\{I(\gamma)\}$. Since $I(\cdot)$ varies from - to + on $[0,\infty)$, the infima in (6.3) and (6.4) are attained at $\alpha=\min(\alpha_r,\alpha_*)$ and $\alpha=\min(\alpha_r,\alpha_0)$ respectively with $\alpha_r=\gamma_r/(1+\gamma_r)$. As well, for $a\ge 1/2$ and $\gamma_*=\alpha_*/(1-\alpha_*)$, $I(\gamma_*)=I(K_1^{-1/4})>0$, which implies, in this case, that $\min(\alpha_r,\alpha_*)=\alpha_r$. Finally, by noting that $S(\alpha,a(1-\alpha))$ is increasing in α on $[\alpha_*,\alpha_0]$ whenever $a\geq 1/2$, expressions (6.3) and (6.4) now lead part (i) of the lemma.

(ii) For $(\alpha,\beta)\in I_2$, $S(\alpha,\beta)$ is given by (6.2) and is, for fixed α, minimized in β by choosing $\beta=\max\{a(1-\alpha),g_\alpha(a)\}$ if $a\leq 1/2$; and $\beta=a(1-\alpha)$ if $a>1/2$. Furthermore, we infer that, $S_2=\inf_{\alpha\geq\alpha_0} S(\alpha,g_\alpha(a))=S(\alpha_0,g_{\alpha_0}(a))=1/\{4(n^{1/2}+1)^2\}$, whenever $a\leq 1/2$; and, $S_2=\inf_{\alpha\geq\alpha_0} S(\alpha,a(1-\alpha))=S(\alpha_0,a(1-\alpha_0))=a(1-a)/\{4(n^{1/2}+1)^2\}$, whenever $a\geq 1/2$; as both $S(\alpha,g_\alpha(a))$ and $S(\alpha,a(1-\alpha))$ are increasing in α for $\alpha\geq\alpha_0$.

(iii) The proof is similar to the one in (i). Note that the form of $S(\alpha,\beta)$ differs but that $S(\alpha,g_\alpha(a+b/2))$ may still be expressed as $H(\gamma)$ with $\gamma=\alpha/(1-\alpha)$.

(iv) Similar to (ii).

(v) From Lemma 3.3, we may write

$$S(\alpha,\beta)=\frac{\alpha^2}{-nA}[(\beta-(\frac{1-\alpha}{2})^2]+\frac{\alpha^2}{4n}.$$

Since $\alpha\geq\alpha_0$ for $(\alpha,\beta)\in I_5$, $A<0$ and $S(\alpha,\beta)\geq(\alpha_0^2/4n)=1/\{4(n^{1/2}+1)^2\}$, which implies $S_5\geq 1/\{4(n^{1/2}+1)^2\}$.

Proof of Theorem 3.5. Following the strategy outlined after Lemma 3.3, two additional observations suffice in order to transform the findings of Lemma 3.4 into the claims of Theorem 3.5. First, note that whenever $a\geq 1/2$, $\alpha_*\leq\alpha_0$ and, as a consequence, $S_2=(a(1-a)\alpha_0^2)/n\geq(a(1-a)\alpha_*^2)/n\geq S_1$. Similarly, whenever $b\leq 1/2$, $\alpha_*\leq\alpha_0$ and $S_4\geq S_3$. For the second observation, it follows from the results of part (i) of Section 6.3, as well as the definition of H_-, that $H_-\leq H(\gamma_0)$; $\gamma_0=\alpha_0/(1-\alpha_0)$; and, whenever $a<1/2<b$, $H_-=H(\gamma_r)$ (or $=H(\gamma_0)=1/[4(n^{1/2}+1)^2]$) depending whether $\alpha_r<\alpha_0$ (or not) $\Leftrightarrow$ $\gamma_r<n^{1/2}$ (or not) $\Leftrightarrow$ $I(n^{1/2})>0$ (or not) $\Leftrightarrow$ $n<\{[(2b-1)(1-2a)]^{-1}-1\}^2$ (or not).

6.4 *Proof of Theorem 3.9*

From (3.3), it is immediate that $R^*(\theta,\delta_{\alpha,\beta})$ is convex in θ on $[a,b]$. This implies that

$$\sup_{a\leq\theta\leq b} R^*(\theta,\delta_{\alpha,\beta}$$

irregardless of (α,β). A further analys

$$R^*(a,\delta_{\alpha,\beta})\geq R^*(b,\delta_{\alpha,\beta})\Leftarrow$$

where $y=1-\alpha-\beta$. Consequently, (6.5) i

$$\inf_{\alpha,\beta\geq 0}\ \sup_{a\leq\theta\leq b} R^*(\theta,\delta_{\alpha,\beta})=$$

with $C_1=\{(\alpha,\beta):\beta\geq 0,\ |y|\leq m\beta\}$ and C_2

being valid due to Lemma 3.1. In ter

$$R^*(x,\delta_{\alpha,\beta})=f(y,\beta)=\beta^2(\frac{1}{n}+\frac{1}{x}$$

This function admits a critical point an

in C_1 nor C_2 for x=a and x=b respecti

be attained on the boundaries of C_1 an

$f(y,\beta)$ is unbounded for large β or larg

to $\{(y,\beta):y=m\beta\}$ for C_1 and $\{(y,\beta):y=$

express either $R^*(a,\delta_{\alpha,\beta})$ or $R^*(b,\delta_{\alpha,\beta})$ a

$$\beta^2\{(ab)^{-1}-($$

Now, $(ab)^{-1}\geq(m+1)^2$, which implies tha

of $H(1-H+nH)^{-1}$. The result of our T

$R^*(b,\delta_{\alpha,0})\geq H(1-H+nH)^{-1}$. But, as a con

$b+nb)^{-1}$. Finally, we observe that $H\leq$b

the proof of the theorem.

$$\sup_{a \le \theta \le b} R^*(\theta,\delta_{\alpha,\beta}) = \max\{R^*(a,\delta_{\alpha,\beta}), R^*(b,\delta_{\alpha,\beta})\}, \tag{6.5}$$

irregardless of (α,β). A further analysis based on (3.3) leads to the comparison

$$R^*(a,\delta_{\alpha,\beta}) \ge R^*(b,\delta_{\alpha,\beta}) \Leftrightarrow \frac{\beta^2}{a}+\frac{(1-\alpha-\beta)^2}{1-a} \ge \frac{\beta^2}{b}+\frac{(1-\alpha-\beta)^2}{1-b} \Leftrightarrow |y| \le m\beta\,;$$

where $y=1-\alpha-\beta$. Consequently, (6.5) implies that the minimax risk is given by

$$\inf_{\alpha,\beta \ge 0}\sup_{a \le \theta \le b} R^*(\theta,\delta_{\alpha,\beta}) = \min\{\inf_{(\alpha,\beta)\in C_1} R^*(a,\delta_{\alpha,\beta}),\ \inf_{(\alpha,\beta)\in C_2} R^*(b,\delta_{\alpha,\beta})\},$$

with $C_1=\{(\alpha,\beta):\beta\ge 0,\ |y|\le m\beta\}$ and $C_2=\{(\alpha,\beta):\beta\ge 0,\ |y|\ge m\beta\}$; the restriction to values of $\beta\ge 0$ being valid due to Lemma 3.1. In terms of (y,β), we have from (3.3) and for $x=a$ or b,

$$R^*(x,\delta_{\alpha,\beta}) = f(y,\beta) = \beta^2\left(\frac{1}{n}+\frac{1}{x}-1\right)+y^2\left(\frac{1}{n}+\frac{1}{1-x}-1\right)+2\beta y\left(1-\frac{1}{n}\right)-2\left(\frac{\beta+y}{n}\right)+\frac{1}{n}. \tag{6.6}$$

This function admits a critical point and absolute minimum at $\beta=1-y=x$ but this point lies neither in C_1 nor C_2 for $x=a$ and $x=b$ respectively. Therefore, both $\inf_{(\alpha,\beta)\in C_1} R^*(a,\delta_{\alpha,\beta})$ and $\inf_{(\alpha,\beta)\in C_2} R^*(b,\delta_{\alpha,\beta})$ will be attained on the boundaries of C_1 and C_2 respectively. Also, since $f(y,\beta)\le f(-y,\beta)$ for $y\ge 0$, and $f(y,\beta)$ is unbounded for large β or large y, we may restrict our search for the minimizing value to $\{(y,\beta):y=m\beta\}$ for C_1 and $\{(y,\beta):y=m\beta$ or $y\ge 0,\ \beta=0\}$ for C_2. Setting $y=m\beta$ in (6.6), we may express either $R^*(a,\delta_{\alpha,\beta})$ or $R^*(b,\delta_{\alpha,\beta})$ as

$$\beta^2\{(ab)^{-1}-(m+1)^2+\frac{(m+1)^2}{n}\}-\frac{2\beta(m+1)}{n}+\frac{1}{n}.$$

Now, $(ab)^{-1}\ge(m+1)^2$, which implies that $f(y,\beta)$ above is convex in β with a minimum of $f(y,\beta_m)$ of $H(1-H+nH)^{-1}$. The result of our Theorem will now hold as long as we can show that $\inf_\alpha R^*(b,\delta_{\alpha,0})\ge H(1-H+nH)^{-1}$. But, as a consequence of Theorem 3.8, we may write $\inf_\alpha R(b,\delta_{\alpha,0})=b(1-b+nb)^{-1}$. Finally, we observe that $H\le b$, which implies $H(1-H+nH)^{-1}\le b(1-b+nb)^{-1}$ and completes the proof of the theorem.

6.5 Proofs of Theorem 4.4 and Corollary 4.5

The proof of Theorem 4.4 follows the lines of Theorem 4.2, with the dominance criteria of an equivariant estimator over the mle still valid. The difference resides in the representation of Bayes rules for normalized loss. In fact, for a given prior G on [a,b], 0<a<b<1, the Bayes rule for normalized loss is

$$\delta^G(x) = \frac{E^G\{\theta^x(1-\theta)^{n-x-1}\}}{E^G\{\theta^{x-1}(1-\theta)^{n-x-1}\}}.$$

Therefore the pivotal quantities $\beta=\delta^G(0)$ become $E^{-1}(1/\theta)$ for n=1, and $\{2E(1/\theta)-2\}^{-1}$ for n=2, which yield, in conjunction with the conditions $a\leq\beta\leq 3a-4a^2$ (n=1) and $a\leq\beta\leq\{a(1-2a^2)\}/\{1-2a+2a^2\}$ (n=2), the desired result.

The proof of Corollary 4.5 follows from Theorem 4.4, the representation $E(1/\theta)=(1-a)^{-1}$ ${}_2F_1(1,\varepsilon;2\varepsilon;(1-2a)/(1-a))$; which we prove in the following lemma; and the fact that, for 0<a<1/2, ${}_2F_1(1,\varepsilon;2\varepsilon;(1-2a)/(1-a))$ is a decreasing function in ε, $\varepsilon>0$. Note that, for 0<a<1/2, $\lim_{\varepsilon\to 0}$ ${}_2F_1(1,\varepsilon;2\varepsilon;(1-2a)/(1-a))=1/2a\geq(1-a)d_n$ which guarantees that $\varepsilon_n^*(a)>0$ and that the set of dominating Bayes rules associated with translated Beta priors is not empty.

LEMMA 6.1. When the prior distribution for θ is a translated Beta$(\varepsilon,\varepsilon)$ on [a,1-a], we have $E(1/\theta)=(1-a)^{-1}{}_2F_1(1,\varepsilon;2\varepsilon;(1-2a)/(1-a))$, where ${}_2F_1$ is the Gauss hypergeometric function given by

$${}_2F_1(a,b;c;y)=\sum_{x=0}^{\infty}\frac{(a)_x(b)_x}{(c)_x}\frac{y^x}{x!},\ with\ (\gamma)_x=\prod_{i=0}^{x-1}(a+i),\ (\gamma)_0=1.$$

Proof. The definition of our translated Beta, and its symmetry with respect to 1/2, imply $\theta \overset{d}{=} (1-\theta) \overset{d}{=} (1-2a)Y+a$ with $Y\sim$Beta$(\varepsilon,\varepsilon)$. Coupled with some algebra, we obtain

$$E(\frac{1}{\theta})=E(\frac{1}{1-\theta})=\sum_{k=}$$

$$=\sum_{z=0}^{\infty}\frac{(\epsilon)_z}{(2\epsilon)_z}(1-2a)^z\sum_{k=z}^{\infty}\binom{k}{z}a$$

using the binomial expansion, the mome

a negative binomial identity.

Acknowledgements The authors are inc
in an earlier version of the manuscript,
version of the paper. This research wa
would like to acknowledge the work
numerical algorithms. Finally, the
presentation of our work.

References

[1] Berger, J. (1985). *Statistical Deci*
NewYork.
[2] Berry, C. (1989). Bayes minimax
space. *Communications in Statist*
[3] Berry, C. (1990). Minimax estima
Multivariate Analysis, 35, 130-13
[4] Bickel, P. (1981). Minimax estim
parameter space is restricted. *Ann*
[5] Brown, L.D. (1971). Admissible
value problems. *Annals of Mathe*
[6] Casella, G., and Strawderman, W.
Statistics, 9, 870-878.
[7] Charras, A., and van Eeden, C. (1

$$E(\frac{1}{\theta}) = E(\frac{1}{1-\theta}) = \sum_{k=0}^{\infty} E(\theta^k) = \sum_{k=0}^{\infty} \sum_{z=0}^{k} \binom{k}{z} (1-2a)^z a^{k-z} E(Y^z)$$

$$= \sum_{z=0}^{\infty} \frac{(\epsilon)_z}{(2\epsilon)_z} (1-2a)^z \sum_{k=z}^{\infty} \binom{k}{z} a^{k-z} = \sum_{z=0}^{\infty} \frac{(\epsilon)_z}{(2\epsilon)_z} \frac{(1-2a)^z}{(1-a)^{z+1}} = {}_2F_1(1,\epsilon;2\epsilon;\frac{(1-2a)}{(1-a)});$$

using the binomial expansion, the moments of a Beta(ε,ε) distribution, a change of variables, and a negative binomial identity.

Acknowledgements The authors are indebted to Constance van Eeden for pointing out a mistake in an earlier version of the manuscript, and for several other comments which led to an improved version of the paper. This research was partially supported by NSERC of Canada. The authors would like to acknowledge the work of Claude Beauchamp in the implementation of the numerical algorithms. Finally, the comments and suggestions of a referee improved the presentation of our work.

References

[1] Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, NewYork.

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

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

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

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

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

[7] Charras, A., and van Eeden, C. (1991). Bayes and admissibility properties of estimators in

truncated parameter spaces. *Canadian Journal of Statistics,* 19, 121-134.

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

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

[10] Ferguson, T.S. (1967). *Mathematical Statistics.* Academic Press, New York.

[11] 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.

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

[13] 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.

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

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

[16] Katz, M.W. (1961). Admissible and minimax estimates of parameters in truncated spaces. *Annals of Mathematical Statistics,* 32, 136-142.

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

[18] Kempthorne, P.J. (1988). Dominating inadmissible procedures using compromise decision theory. *Statistical Decision Theory and Related Topics IV,* Springer-Verlag, NewYork, 381-396.

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

[20] Levit, B.Y.A. (1980). On asymptotic minimax estimation of the second order (translation). *Theory Proba. Appl. 25,* 552-568.

[21] Marchand, E. (1993). Estimatio

Canadian Journal of Statistics, 2

[22] Moors, J.J.A. (1985). Estimatio

University.

[23] Moors, J.J.A. (1981). Inadmissi

parameter spaces. *Journal of the*

[24] Sacks, J. (1963). Generalized B

Mathematical Statistics, 34, 751

[25] Shao, P.Y-S., and Strawderman,

mean. *Statistica Sinica,* 6, 275-2

[26] Vidakovic, B., and DasGupta, A.

normal mean. *Sankhyā,* 58, 81-1

[27] Wald, A. (1950). *Statistical De*

[28] Warner, S.L. (1965). Randomiz

answer bias. *Journal of the Am*

[29] Wieczorkowski, R. and Zieliński,

with entropy loss function. *Stat*

[30] Winkler, R.L., and Franklin, L.A.

approach. *Journal of the Ameri*

[31] Zinzius, E. (1981). Minimaxscha

mit bekannter Varianz bei vorge

Operationsforsch. Statist. Ser. S

Éric Marchand
University of New Brunswick
Department of Mathematics and Stati
P.O. Box 4400
Fredericton, N.B. CANADA E3B 5A

Brenda MacGibbon
GERAD and UQAM
Département de mathématiques
C.P. 8888, Succursale Centre-Ville, M

[21] Marchand, E. (1993). Estimation of a multivariate mean with constraints on the norm. *Canadian Journal of Statistics,* 21, 359-366.

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

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

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

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

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

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

[28] Warner, S.L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60, 63-69.

[29] Wieczorkowski, R. and Zieliński, R. (1992). Minimax estimation of a binomial probability with entropy loss function. *Statistics & Decisions,* 10, 39-44.

[30] Winkler, R.L., and Franklin, L.A. (1979). Warner's randomized response model: a Bayesian approach. *Journal of the American Statistical Association,* 74, 207-214.

[31] Zinzius, E. (1981). Minimaxschatzer für den Mittelwert υ einer normalverteilten Zufallgröße mit bekannter Varianz bei vorgegebener oberer und unterer Scranke für υ. *Math. Operationsforsch. Statist. Ser. Statist.,* 12, 551-557.

Éric Marchand
University of New Brunswick
Department of Mathematics and Statistics
P.O. Box 4400
Fredericton, N.B. CANADA E3B 5A3

Brenda MacGibbon
GERAD and UQAM
Département de mathématiques
C.P. 8888, Succursale Centre-Ville, Montréal (Québec) H3C 3P8

Figure 1 Cutoff points for minimax one and two point priors on [0,b]

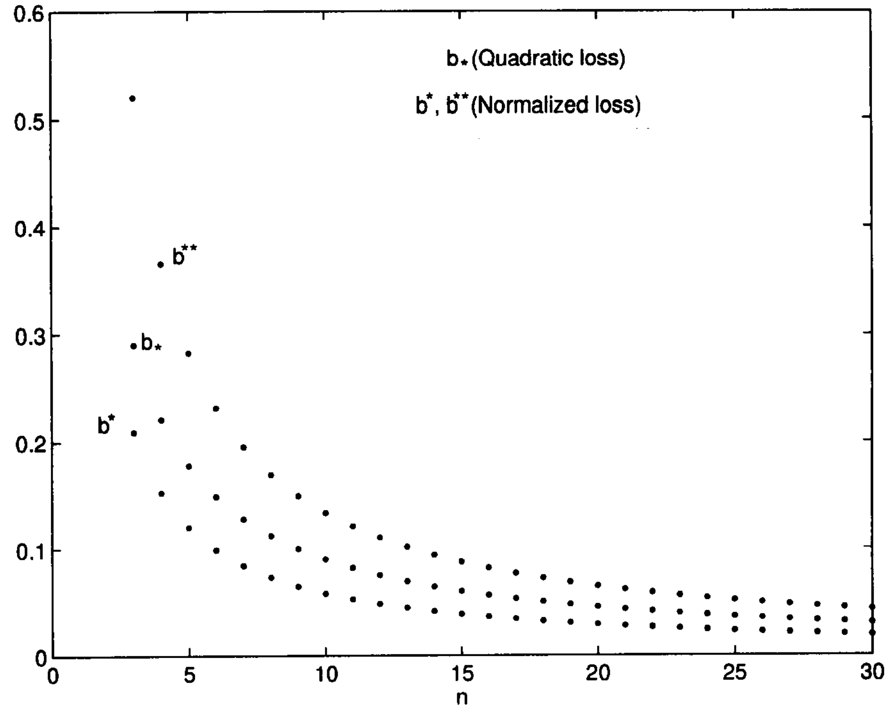

Figure 2 R

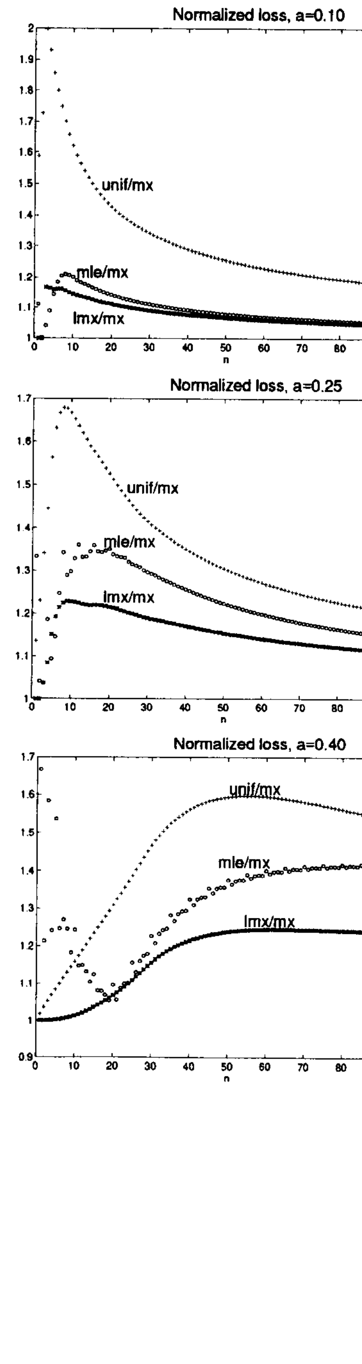

Figure 2 Ratio of maximum risks on [a,1-a]

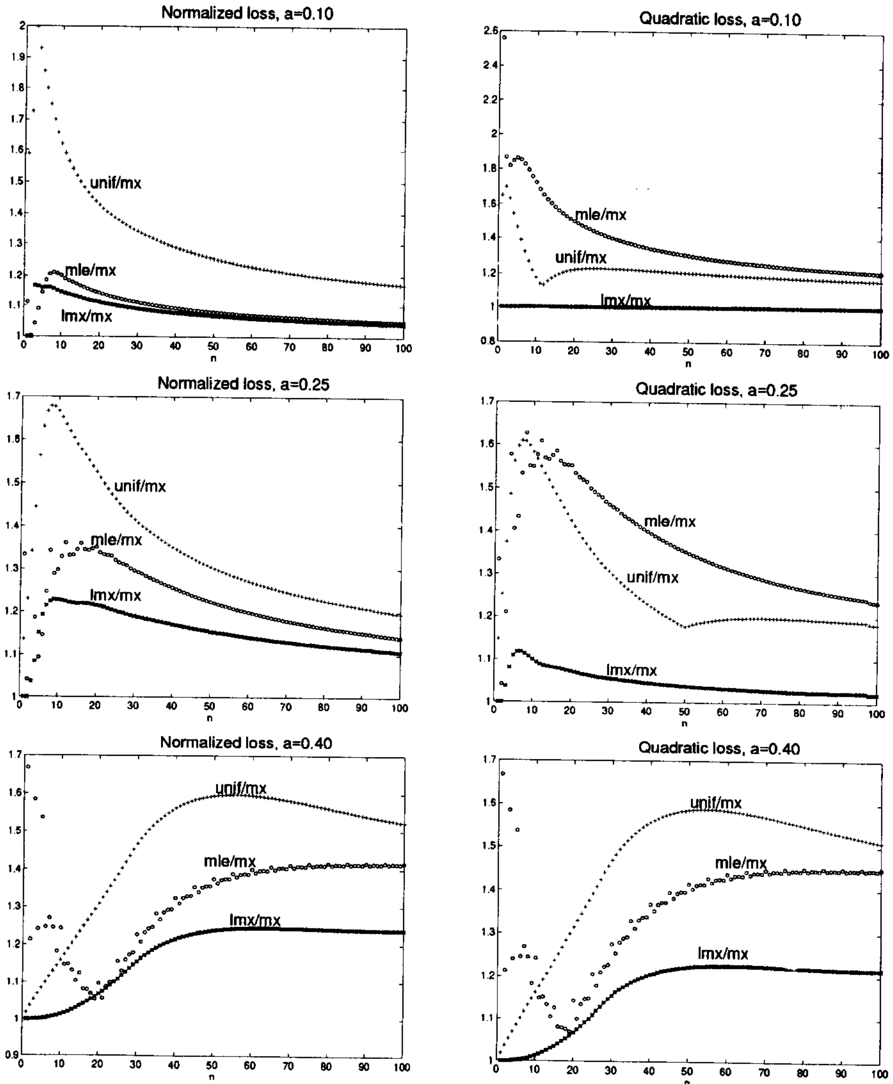

Figure 3 Ratio of maximum risks on [0,b]

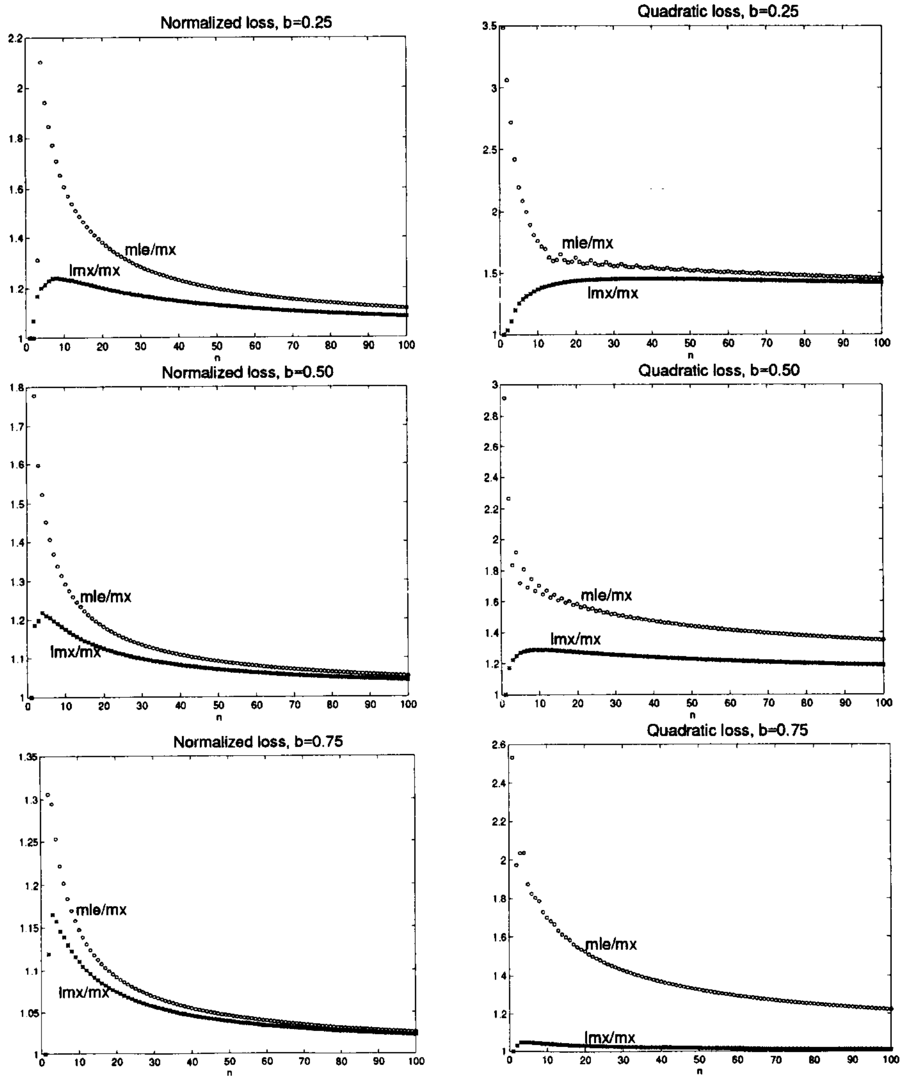

Figure 4 Ratios of

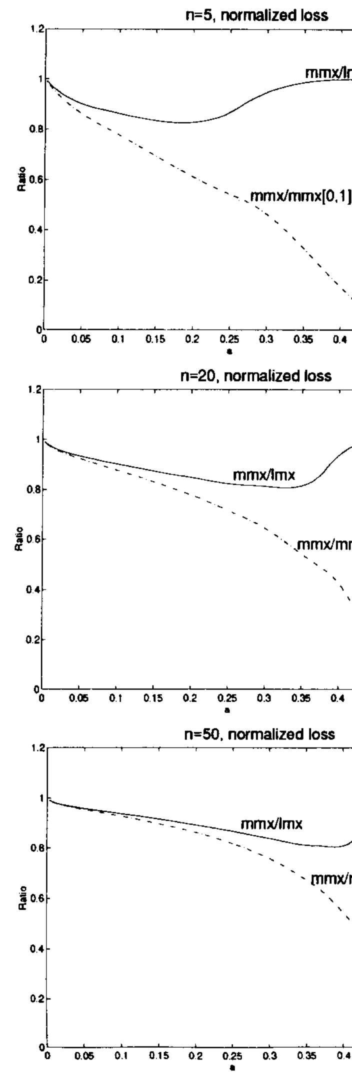

Figure 4 Ratios of minimax risks on [a,1-a] as a function of a

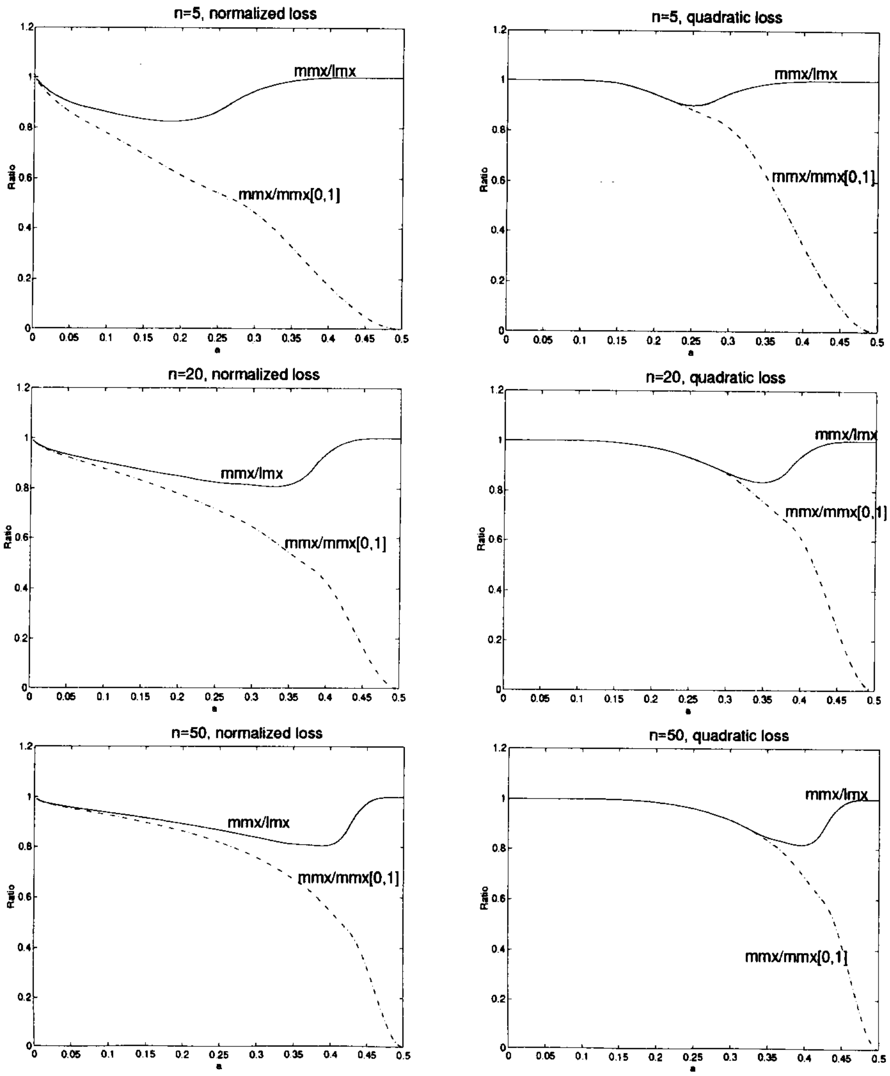

Figure 5 Ratios of minimax risks on [0,b] as a function of b

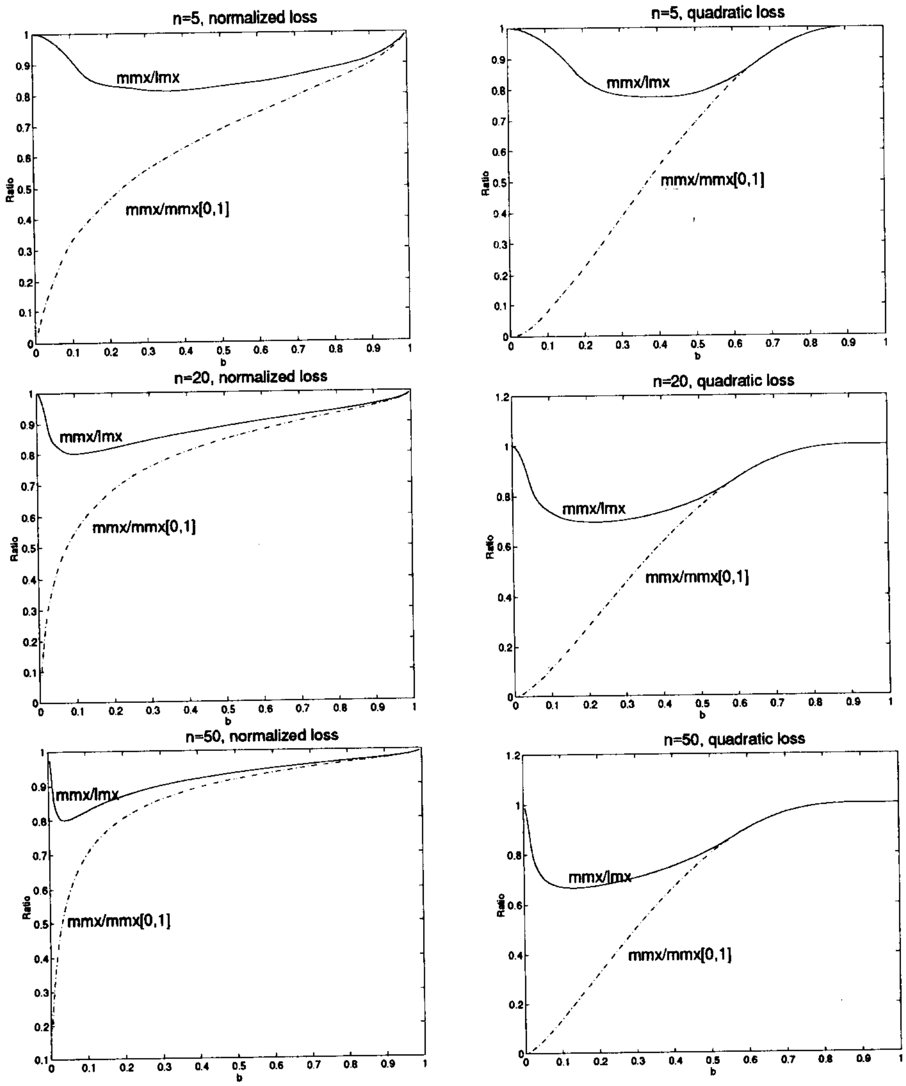

Figure 6 Compa

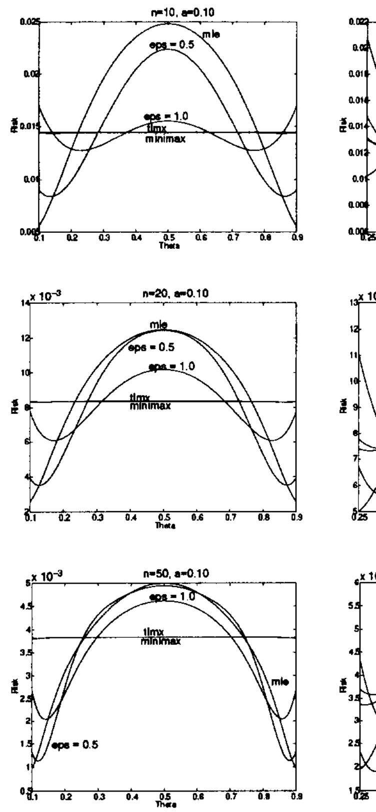

Figure 6 Comparative risks on [a,1-a], quadratic loss

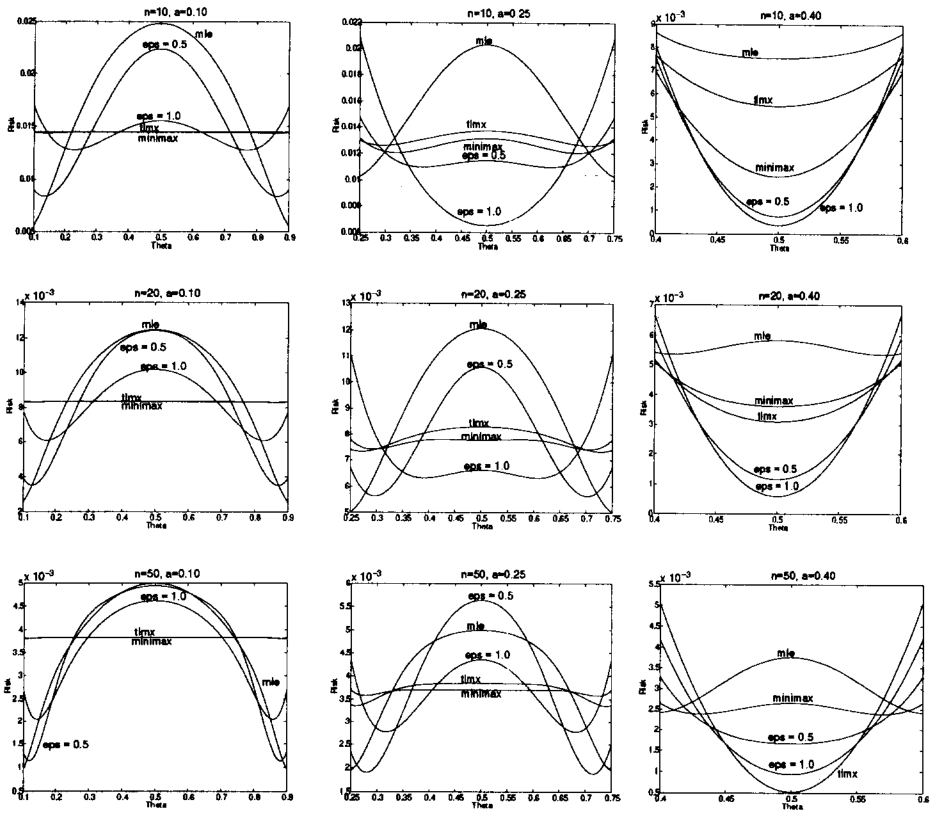

Figure 7 Comparative risks on [a,1-a], normalized loss

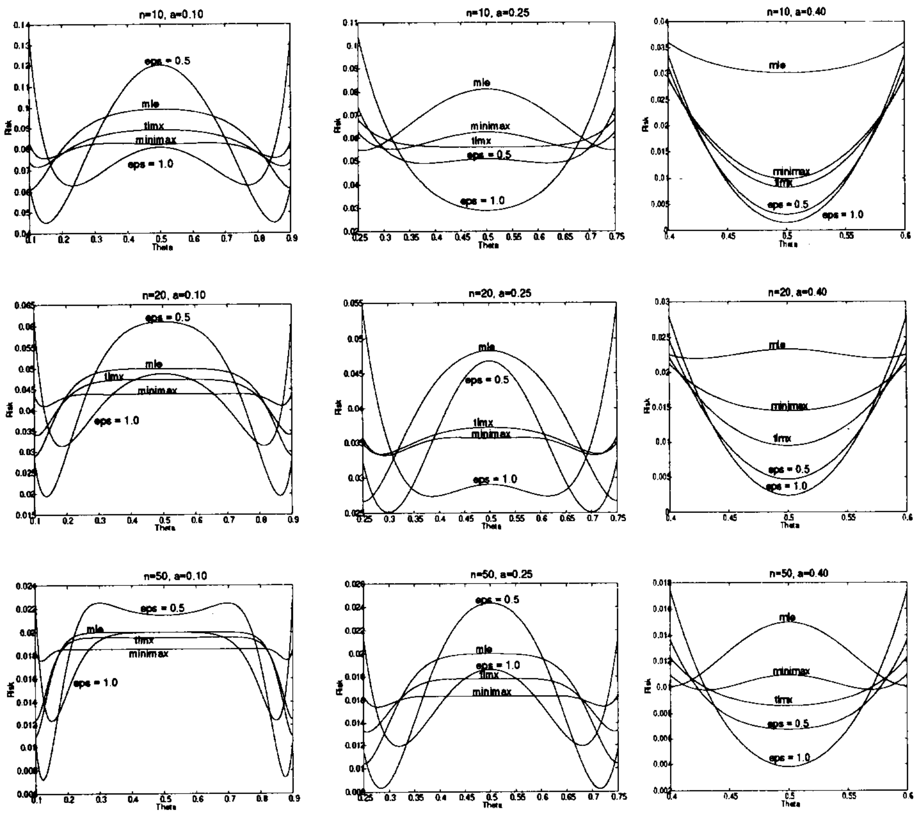

Figure 8 Compa

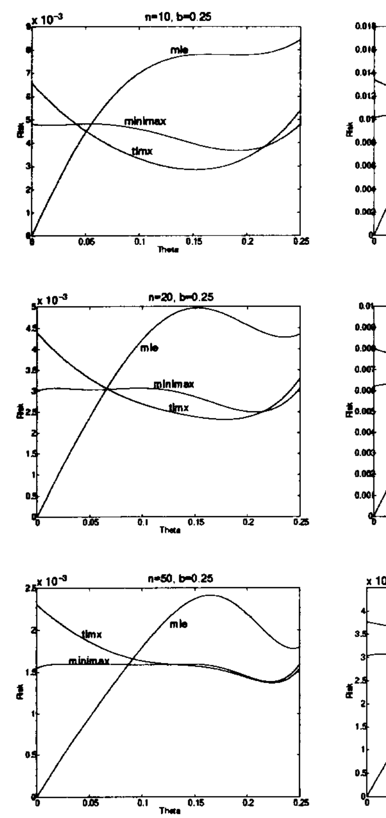

Figure 8 Comparative risks on [0,b], quadratic loss

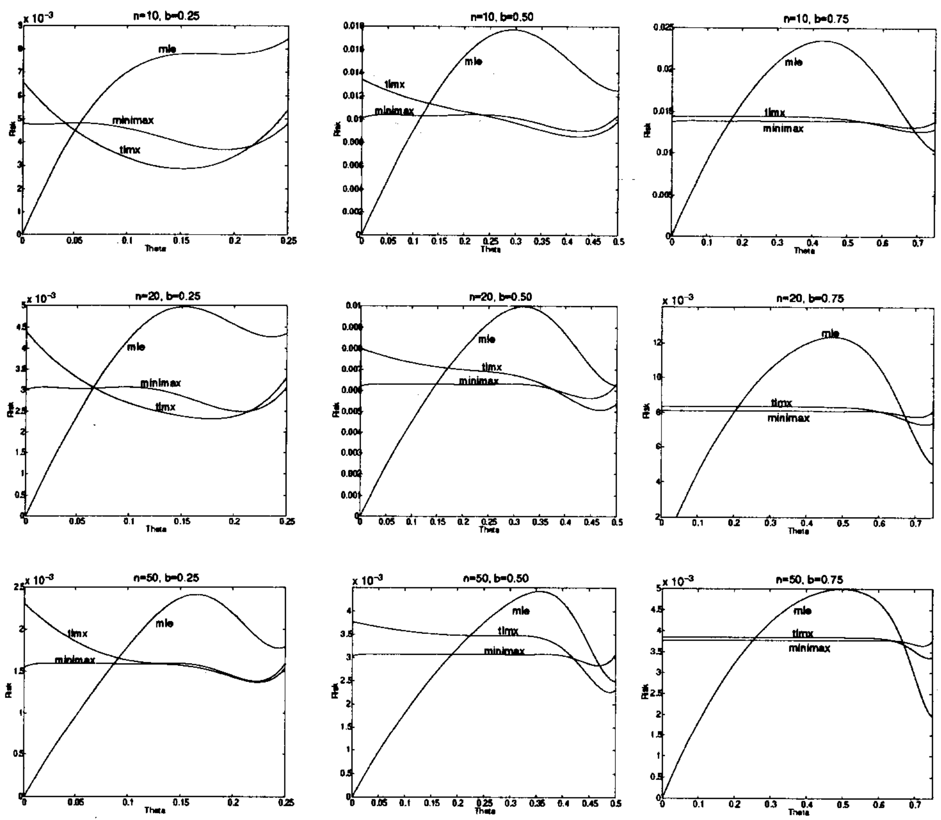

Figure 9 Comparative risks on [0,b], normalized loss

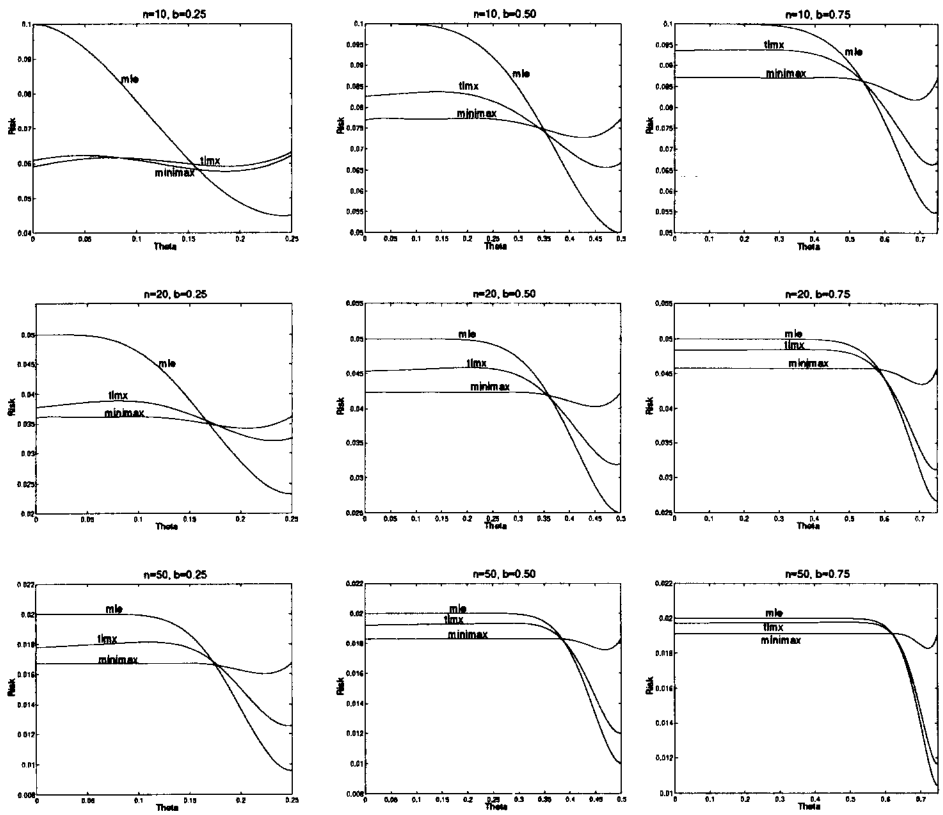

Figure 10 Comparativ

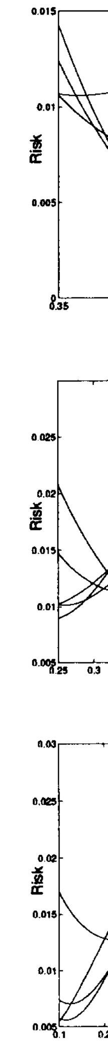

Figure 10 Comparative risks on [a,1-a], mle vs Bayes, quadratic loss