Minimax Rules Under Zero-One Loss for a Restricted Location Parameter ⋆

G. Kamberova [1,2] M. Mintz [1]

GRASP Laboratory, Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract

Minimax Rules Under Zero-One Loss

In this paper, we obtain minimax and near-minimax nonrandomized decision rules under zero-one loss for a restricted location parameter of an absolutely continuous distribution. Two types of rules are addressed: monotone and nonmonotone. A complete-class theorem is proved for the monotone case. This theorem extends the previous work of Zeytinoglu and Mintz (1984) to the case of $2e$-MLR sampling distributions. A class of continuous monotone nondecreasing rules is defined. This class contains the monotone minimax rules developed in this paper. It is shown that each rule in this class is Bayes with respect to nondenumerably many priors. A procedure for generating these priors is presented. Nonmonotone near-minimax almost-equalizer rules are derived for problems characterized by non- $2e$-MLR distributions. The derivation is based on the evaluation of a distribution-dependent function Q_c. The methodological importance of this function is that it is used to unify the discrete- and continuous-parameter problems, and to obtain a lower bound on the minimax risk for the non- $2e$-MLR case.

1 Introduction

We address the problem of estimating a bounded location parameter θ from a noisy observation $Z = \theta + V$. We study the problem of minimax estimation of θ under zero-one loss. This problem naturally arises in robotic applications where the parameter space is defined by the workspace of the robot, where exacting tolerances on performance are demanded, and where multiple observations are not available: Kamberova, Mandelbaum and Mintz (1996).

⋆ AMS 1991 subject classifications. Primary—62C20, 62C15, 62C07, 62F10, 62F25, 62F30; secondary—62C10. Key words and phrases: minimax, location parameter, zero-one loss, MLR, nonmonotone rules, non-MLR.

[1] Supported in part by: Navy Grants N00014-92-J-1647, NOOO14-92-J-1647; AFOSR Grant 88-0296; Army/DAAL 03-89-C-0031PRI; NSF Grants CISE/CDA 88-22719, IRI 89-06770, and ASC 91 0813; and the Du Pont Corporation.

[2] NASA grant NAG 9-510.

Definition 1.1 (The minimax problem): Let $\mathbf{P}(F, \Omega, L_e)$ denote the minimax problem defined by the statistical game $(\,\Omega\,,\,D\,,\,R\,)$ and the observation model $Z = \theta + V$, where: (i) the parameter space, $\Omega \subset \boldsymbol{R}$, is a compact interval or a finite set; (ii) D is the set of all nonrandomized decision rules $\delta : \boldsymbol{R} \to \mathcal{A}$, where the action space $\mathcal{A} \subseteq \Omega$; (iii) the risk, $R(\delta,\, \theta,\, F) = \mathrm{E}_{\theta,F}[L_e(\,\delta(Z)\,,\,\theta\,)]$, is defined with respect to the zero-one loss function:

$$L_e(a,\, \theta) = \begin{cases} 0, \mid a - \theta \mid \leq\; e, \\ 1, \mid a - \theta \mid >\; e, \end{cases} \tag{1.1}$$

where $e \geq 0$; and (iv) $V \sim F$, where: F (the CDF) is independent of θ, F is absolutely continuous, and F has density f with convex support.

Definition 1.2 (The continuous-parameter problem $\mathbf{P}(F, d, L_e)$): Let $d > e > 0$, and $\mathbf{P}(F, d, L_e)$ denote an instance of the minimax problem $\mathbf{P}(F, \Omega, L_e)$, where $\Omega = [\,-d\,,\,d\,]$. The action space $\mathcal{A}$ is naturally limited to $[\,-(d-e)\,,\,(d-e)\,]$ as a consequence of the location problem and the choice of Ω and L_e.

Our goal is to find a minimax rule $\delta^* \in D$ for the continuous-parameter problem $\mathbf{P}(F, d, L_e)$, i.e., δ^* such that: $\forall\, \delta \in D,\;\; \sup_{\theta\in\Omega} R(\delta^*_,\, \theta,\, F) \leq \sup_{\theta\in\Omega} R(\delta,\, \theta,\, F)$.

Note that for $\mathbf{P}(F, d, L_e)$, the risk of a nonrandomized rule δ equals the probability of an unacceptable error: $R(\delta,\, \theta,\, F) = \Pr[\,\delta(Z) < \theta - e\,] + \Pr[\,\delta(Z) > \theta + e\,]$. Thus, if δ^* is minimax, then the interval $[\delta^*(Z) - e,\, \delta^*(Z) + e]$ can be interpreted as a confidence interval for θ with highest confidence coefficient, $\inf_\theta \Pr_{\theta,F}[\theta \in [\delta^*(Z) - e,\, \delta^*(Z) + e]\,]$, among all intervals of length $2e$.

In this paper we solve the minimax problem $\mathbf{P}(F, d, L_e)$ characterized by CDFs F which satisfy the following $2e$-MLR condition.

Definition 1.3 (The $2e$-MLR condition): Given $e > 0$, and an absolutely continuous distribution F with density f, we say that F (or f) is $2e$-MLR (strictly $2e$-MLR), if the ratio:

$$\frac{f(x+2e)}{f(x)} \tag{1.2}$$

is monotone decreasing (strictly monotone decreasing) in $x \in \boldsymbol{R}$.

The $2e$-MLR condition is equivalent to the requirement that the density f is monotone likelihood ratio (MLR) in the context of the location model, and the two-point parameter space $\{\,-2e\,,\,0\,\}$. If F is absolutely continuous and $e > 0$, then MLR implies $2e$-MLR.

When F is MLR, $\mathbf{P}(F, d, L_e)$ is a monotone estimation problem in the sense of Karlin and Rubin (1956). Consequently, the class of monotone decision rules is essentially complete:

Berger (1985). Further, if F is strictly MLR, then it follows from Brown, Cohen and Strawderman (1976) that the class of the monotone decision rules is complete. Our work regarding the monotone rules is most closely related to that of Zeytinoglu and Mintz (1984), where the authors investigated the given minimax problem for symmetric unimodal F. When F is strictly MLR, the authors obtained admissible minimax Bayes rules which are odd-symmetric monotone piecewise-linear functions. When F is not MLR, the authors proved that the rules they obtained are minimax within the class of all nonrandomized monotone procedures.

The contributions of the present paper are threefold. First, we explicitly delineate both the structure and computation of nonrandomized monotone minimax rules for $\mathbf{P}(F, d, L_e)$, where F is $2e$-MLR. Examples of such distributions include location and scale-mixtures of normal distributions, and all MLR distributions. We extend the results of Zeytinoglu and Mintz (1984) to asymmetric and multimodal F which are $2e$-MLR. Second, when F is not $2e$-MLR, we restrict our attention to the class of symmetric unimodal distributions. This class contains distributions with heavy tails, for example, the standard Cauchy. Here, we obtain nonrandomized nonmonotone almost-equalizer rules with near-minimax risk. In particular, when F is the standard Cauchy distribution, we compute a piecewise-continuous nonmonotone rule with maximum risk within 0.3% of the global minimax risk. The nonmonotone rules which we obtain in this case represent a new class of rules and provide a logical extension to the monotone rules. Third, by establishing tight bounds on the global minimax risk for the Cauchy problem, we demonstrate the near-global minimax property of the easily computed best monotone rule.

The remainder of this paper is organized as follows. In Section 2, we present: (i) the main theorem regarding the monotone rules, and (ii) give a constructive procedure for generating priors for which continuous monotone nondecreasing rules are Bayes under zero-one loss. Section 3 is devoted to the nonmonotone rules: the near-minimax performance of almost-equalizer rules, and the connection between the discrete- and continuous-parameter problems. We also compare the structure and behavior of the nonmonotone decision rules obtained in this section with the minimax rules for the corresponding decision problem under quadratic loss. In Section 4, we conclude by discussing the importance of the class of the monotone rules; such rules are easy to compute, and in many cases, although not globally minimax, have near-minimax performance. We focus on the main ideas which lead to our results. For detailed proofs, the reader is referred to Kamberova (1992), and Kamberova and Mintz (1998a).

2 Monotone Minimax Rules

We assume that the ratio d/e is an integer greater than 1, and express this by $d = (2n + 1)e + u$, where: $n \geq 0$; $u = 0$, if d/e is odd; and $u = e$, if d/e is even. The assumption that d/e is an integer merely simplifies the presentation; it is not essential to obtaining our primary results. The details of the non-integer case follow a pattern which is analogous to the non-integer case in Zeytinoglu and Mintz (1984). For $\mathbf{P}(F, d, L_e)$, where $d = (2n+1)+u$, and F is $2e$-MLR, we find a monotone nondecreasing minimax rule. This rule is a member of the family $\Delta^{d/e}$ defined below. Figure 1 illustrates the structure of the rules in Δ^3 and Δ^4.

Definition 2.1 (The family $\Delta^{d/e}$): Let $d = (2n+1)e + u$. Define $\Delta^{d/e}$ to be the set of monotone nondecreasing piecewise-linear decision rules, δ, parameterized by a $(2n + u/e)$-dimensional vector, $\boldsymbol{a}$:

$$\delta(z) = \begin{cases} d - e, & u + a_n + 2ne \leq z; \\ z - a_i, & u + a_i + 2(i-1)e \leq z < u + a_i + 2ie, \\ & i = 1, 2, \ldots, n; \\ 2(i-1)e + u, & u + 2(i-1)e + a_{i-1} \leq z < u + a_i + 2(i-1)e, \\ & i = 1, 2, \ldots, n; \\ z - a_0, & -u + a_0 < z < u + a_0; \\ -2(i-1)e - u, & -u + a_{-i} - 2(i-1)e < z \leq -u + a_{-i+1} - 2(i-1)e, \\ & i = 1, 2, \ldots, n; \\ z - a_{-i}, & -u + a_{-i} - 2ie < z \leq -u + a_{-i} - 2(i-1)e, \\ & i = 1, 2, \ldots, n; \\ -d + e, & z \leq -u + a_{-n} - 2ne, \end{cases} \tag{2.1}$$

and

$$\boldsymbol{a} = \begin{cases} (a_{-n}, a_{-n+1}, \ldots, a_{-1}, a_1, \ldots, a_{n-1}, a_n), \text{ if } u = 0; \\ (a_{-n}, a_{-n+1}, \ldots, a_{-1}, a_0, a_1, \ldots, a_{n-1}, a_n), \text{ if } u = e; \end{cases} \tag{2.2}$$

where: $-\infty < a_{-n} \leq \cdots \leq a_{-2} \leq a_{-1} \leq a_0 \leq a_1 \leq a_2 \leq \cdots \leq a_n < \infty$. Note that the parameter a_0 is used only when d/e is even, i.e., $u = e$.

Definition 2.2 (Partition of Ω): Consider $\mathbf{P}(F, d, L_e)$, where $d = (2n+1)e + u$. Define a two-set partition of Ω, $\{\Omega_0\,, \Omega_1\}$, where Ω_0 is $\{\pm(e-u), \pm(3e-u), \pm(5e-u), .., \pm(d-2e)\}$, and Ω_1 is the union of open intervals of length $2e$ complimenting Ω_0 in Ω.

Definition 2.3 We call a nonrandomized rule $\delta \in D$ an almost-equalizer rule for $\mathbf{P}(F, d, L_e)$, if $R(\delta,\, \theta,\, F) = \sup_{\theta\in\Omega} R(\delta,\, \theta,\, F)$, except on a finite subset of Ω. Thus, a rule δ such that: $R(\delta,\, \theta,\, F) = c,\ \forall\theta \in \Omega_1,$ and $R(\delta,\, \theta,\, F) < c,\ \forall\theta \in \Omega_0$, where c is a constant, is an almost-equalizer rule. We refer to c as the *essential* risk of δ.

It can be shown that any nonrandomized monotone decision rule which equalizes the risk over Ω_1 is in $\Delta^{d/e}$: Kamberova (1992). Thus, for the decision problem $\mathbf{P}(F, d, L_e)$, where

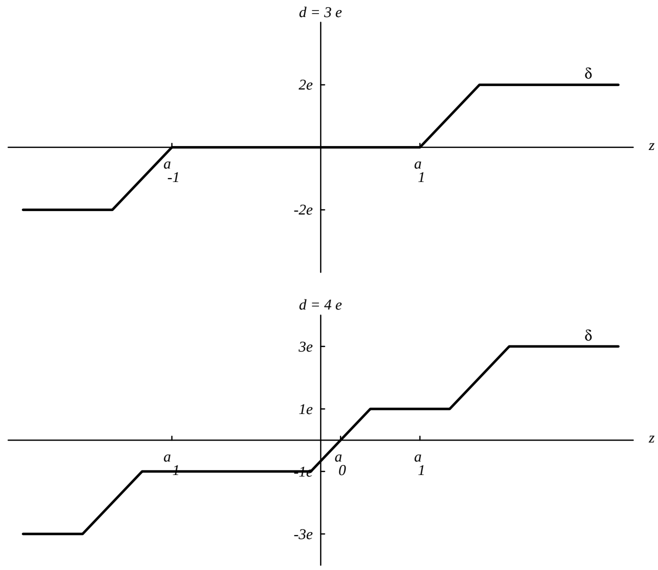

Fig. 1. Rules $\delta(z)$ in Δ^3 and Δ^4, respectively.

$d = (2n+1)e + u$ and F is symmetric, any rule with constant risk over Ω_1 must be in the class of odd-symmetric monotone decision rules obtained by Zeytinoglu and Mintz (1984). (These rules are unique up to sets of Lebesgue measure zero.)

The strategy we use to find a minimax rule for $\mathbf{P}(F, d, L_e)$, where F is $2e$-MLR, is the same as that in Zeytinoglu and Mintz (1984): an almost-equalizer Bayes rule is constructed, and this rule is minimax. The strategy is based on Observations 2.1 and 2.2.

Observation 2.1 Given the problem $\mathbf{P}(F, d, L_e)$, where $d = (2n+1)e+u$, each rule $\delta \in \Delta^{d/e}$ has a piecewise-constant risk, with maximum risk achieved over Ω_1. Thus, when $\delta \in \Delta^{d/e}$, we can restrict the risk evaluation to Ω_1.

Observation 2.2 Let $\delta \in D$, $\sup_{\theta\in\Omega} R(\delta,\, \theta,\, F) = c$, and λ be a prior on Ω. If: (i) δ is Bayes with respect to λ; (ii) $\forall\, \theta \in \Omega_1,\ R(\delta,\, \theta,\, F) = c$; and (iii) $\lambda[\Omega_1] = 1$; then: δ is minimax, and λ is least favorable.

Observation 2.2 is a well-known decision-theoretic result. See Ferguson (1967).

Theorem 2.1 Consider the decision problem $\mathbf{P}(F, d, L_e)$, where $d = (2n+1)e + u$. If F is $2e$-MLR, then there exists a globally minimax Bayes rule $\delta^* \in \Delta^{d/e}$. Further, if F is strictly $2e$-MLR, then δ^* is admissible. The rule δ^* and the minimax risk c^* are the solution of the risk equation:

$$R(\delta^*,\, \theta,\, F) = c^* \qquad \forall\, \theta \in \Omega_1. \tag{2.3}$$

Moreover:

$$\frac{F^{-1}[1-c^*] - F^{-1}[c^*]}{2} \;\le\; e. \tag{2.4}$$

The proof of Theorem 2.1 follows Zeytinoglu and Mintz (1984). The extension includes: (i) allowance for the asymmetry of F; and (ii) the replacement of the MLR condition with the $2e$-MLR condition. The proof makes use of the absolute continuity of F and the convex support of f.

The $2e$-MLR condition and the zero-one loss function are sufficient to guarantee that the almost-equalizer rule is Bayes with respect to a piecewise-constant prior density λ with jumps at the points of Ω_0, i.e.:

$$\lambda(\theta) = \begin{cases} K_{i+\frac{u}{e}}, & u + (2i-1)e < \theta \leq u + (2i+1)e, \quad i = 0, \ldots, n; \\ K_{-i-\frac{u}{e}}, & -u - (2i+1)e \leq \theta < -u - (2i-1)e, \; i = 0, \ldots, n; \end{cases} \tag{2.5}$$

where the parameter vector $\boldsymbol{K} \in \boldsymbol{R}^{d/e}$ has positive components:

$$\boldsymbol{K} = \begin{cases} (K_{-n}, \ldots, K_{-1}, K_0, K_1, \ldots, K_n), & \text{if } u = 0; \\ (K_{-n-1}, \ldots, K_{-1}, K_1, \ldots, K_n, K_{n+1}), & \text{if } u = e; \end{cases} \tag{2.6}$$

and $\boldsymbol{K}$ is a solution of a nonlinear system similar to that in Zeytinoglu and Mintz (1984), and given in Kamberova and Mintz (1998a). The vector $\boldsymbol{K}$ does not have to be computed explicitly in order to determine δ^*. The existence of this prior is simply used as a tool to claim minimaxity. At the end of this section we give a constructive proof that for each rule δ, in a large class of continuous monotone nondecreasing rules, there exists a nondenumerable family of priors with respect to which δ is Bayes.

If F is MLR, the global minimax rule for $\mathbf{P}(F, d, L_e)$ is monotone. Further, there are non-MLR distributions which are $2e$-MLR, and for which, by Theorem 2.1, the global minimax rules for $\mathbf{P}(F, d, L_e)$ are also monotone. There are bimodal (thus non-MLR) and/or asymmetric distributions which satisfy the $2e$-MLR condition. For instance, by investigating the derivative of the ratio (1.2), the following observation is established.

Observation 2.3 Let φ be the standard normal density, and f be a convex combination of two normal densities:

$$f(x) = p \frac{1}{\sigma_1} \varphi\left(\frac{x-\mu_1}{\sigma_1}\right) + (1-p) \frac{1}{\sigma_2} \varphi\left(\frac{x-\mu_2}{\sigma_2}\right), \text{ where } 0 \leq p \leq 1. \tag{2.7}$$

If $\sigma_1 = \sigma_2 > 0$, and $|\mu_1 - \mu_2| < 2e$, then f is $2e$-MLR.

Example 2.1 Consider $\mathbf{P}(F, d, L_e)$, where $d = 2e$. Let f be a mixture of two normal distributions, $\mathcal{N}(-2, 1)$ and $\mathcal{N}(3, 1)$: $f(x) = 0.1\varphi(x+2) + 0.9\varphi(x-3)$. Let $e = 3.477318$. This value of e results in a minimax risk of 0.04. In this example, f is asymmetric and bimodal. The condition from Observation 2.3 is satisfied. Thus, f is $2e$-MLR. Hence, Theorem 2.1

holds. The almost-equalizer Bayes rule δ_m^e, obtained via Theorem 2.1, is:

$$\delta_m^e(z) = \begin{cases} 3.477318, & 4.701288 \le z \ ; \\ z - 1.22397, & -2.253348 \le z < \ 4.701288; \\ -3.477318, & z < -2.253348. \end{cases}$$

This rule is minimax and admissible, since f is strictly $2e$-MLR. To complete this example, we compare δ_m^e with the truncated maximum likelihood estimator (MLE) δ_{mle}. The truncated MLE has the same structure as δ_m^e. Thus, $\delta_{mle} \in \Delta^{d/e}$, where:

$$\delta_{mle}(z) = \begin{cases} 3.477318, & 6.477318 \le z \ ; \\ z - 3, & -0.477318 \le z < \ 6.477318; \\ -3.477318, & z < -0.477318. \end{cases}$$

The risk function of δ_{mle} is piecewise-constant, with maximum value: 0.093836.

Theorem 2.2 If the conditions of Theorem 2.1 are satisfied, with the exception of the $2e$-MLR condition, then the monotone almost-equalizer rule from Theorem 2.1 is not necessarily globally minimax. However, it is *monotone-minimax*, i.e., minimax within the class of all nonrandomized monotone rules.

Theorem 2.3 Consider $\mathbf{P}(F, d, L_e)$, where $d = (2n+1)e + u$. Let δ be a nonrandomized continuous rule which is strictly monotone increasing on $\delta^{-1}[\Omega_1 \cap \mathcal{A}]$. If F is $2e$-MLR, then δ is Bayes with respect to every nonatomic prior $\pi_{\alpha\beta}$ indexed by a pair of positive functions: $\alpha \in L^1[\,(-(2n+1)e - u\,,\, -2ne - u)\,]$, and $\beta \in L^1[\,(-2ne - u\,,\, -(2n-1)e + u)\,]$.

Procedure 2.1 Given α and β, we construct a prior $\pi_{\alpha\beta}$ for which δ is Bayes:

(i) Let $\mu_1 = \alpha$, and $\mu_2 = \beta$.

(ii) Let μ_m denote the m-th segment of an unnormalized density defined on the interval $(\,(m-(2n+1)+u/e)e\,,\,(m-(2n+1)+u/e+1)e\,)$. Define μ_m inductively for $3 \le m \le 2((2n+1)+u/e)+1$. For $\theta \in (\,(m-(2n+1)+u/e)e\,,\,(m-(2n+1)+u/e+1)e\,)$, μ_m is:

$$\mu_m(\theta + e) = \frac{f(\delta^{-1}(\theta) - \theta + e)}{f(\delta^{-1}(\theta) - \theta - e)}\mu_{m-2}(\theta - e)\,.$$

(iii) Let $1 \le m \le 2((2n+1)+\frac{u}{e})+1$. Define the function π^0 in terms of the μ_m over the intervals $(\,(m-(2n+1)+u/e)e\,,\,(m-(2n+1)+u/e+1)e\,)$ of length e in $\Omega_1 \cap \mathcal{A}$, i.e.,

$$\pi^0_{|(\,(m-(2n+1)+\frac{u}{e}+1)e\,,\,(m-(2n+1)+\frac{u}{e}+2)e\,)} \equiv \mu_m,\ \ 1 \le m \le 2((2n+1)+\frac{u}{e})+1.$$

(iv) Normalize π^0. The resulting density is the prior $\pi_{\alpha\beta}$. By the construction of $\pi_{\alpha\beta}$ and the fact that F is $2e$-MLR, it follows that for every fixed $z \in \boldsymbol{R}$, the posterior expected loss is minimized over $\mathcal{A}$ at $\theta = \delta(z)$. Thus, δ is Bayes with respect to $\pi_{\alpha\beta}$.

Corollary 2.1 Under the conditions of Theorem 2.1, every rule $\delta \in \Delta^{d/e}$ is Bayes with respect to nondenumerably many priors over Ω. In particular, for each monotone almost-equalizer Bayes rule δ^*, there are nondenumerably many least favorable priors.

3 Nonmonotone Almost-Equalizer Rules

We consider $\mathbf{P}(F, d, L_e)$, where F is symmetric and unimodal, and for which the almost-equalizer, Bayes, and minimax rules are nonmonotone. In order to focus on the salient issues, and simplify the presentation, we limit the study of $\mathbf{P}(F, d, L_e)$ to the case where $d = 3e$. We derive sufficient conditions for the existence of an almost-equalizer piecewise-continuous nonmonotone rule. The derivation is based on the analysis of the distribution-dependent function, Q_c, defined as follows.

Definition 3.1 Given $\mathbf{P}(F, d, L_e)$, let $Q_c : [\,0\,,\,c\,] \to [\,0\,,\,1\,]$ be defined by:

$$Q_c(x) \;=\; F\Big[2e - F^{-1}[\,c - x\,]\Big] - F\Big[2e - F^{-1}[\,1 - x\,]\Big], \quad 0 < c < 0.5. \tag{3.1}$$

The function Q_c has the following properties: (i) Q_c is differentiable; (ii) Q_c may be monotone or nonmonotone, depending on F; (iii) if F is a symmetric distribution with convex support and $e > 0$, then F is $2e$-MLR if and only if Q_c is a monotone decreasing function; and (iv) Q_c is scale-invariant. Further, for each fixed $x \in [\,\frac{c}{2}\,,\,c\,]$, $Q_c(x)$ is decreasing in c, $c \geq x$.

Definition 3.2 (Q-convex): Let F be absolutely continuous with a unimodal symmetric density f with convex support. We say that F is Q-convex, if for fixed $c, 0 < c < 0.5$, $Q_c(x)$ is convex and nonmonotone. When F is Q-convex, let (c_l, x_l) be the unique solution of the system:

$$Q_c(x) = F[\,2e - F^{-1}[c - x]\,] - F[\,2e - F^{-1}[1 - x]\,] = \frac{c}{2}\,, \tag{3.2}$$

$$\frac{dQ_c}{dx}(x) = \frac{f(\,2e - F^{-1}[c - x]\,)}{f(\,F^{-1}[c - x]\,)} - \frac{f(\,2e - F^{-1}[1 - x]\,)}{f(\,F^{-1}[1 - x]\,)} = 0. \tag{3.3}$$

Definition 3.3 (The family $\tilde{\mathcal{C}}$): Consider the family $\tilde{\mathcal{C}}$ which consists of the odd-symmetric piecewise-continuous decision rules of the form:

$$\delta(z) = \begin{cases} {\phi_1}^{-1}(z), & z \in [\,\phi_1(2e)\,,\,\phi_1(0)\,]; \\ 2e, & z \in [\,\psi_1(2e)\,,\,\phi_1(2e)\,]; \\ {\psi_1}^{-1}(z), & z \in [\,\psi_1(0)\,,\,\psi_1(2e)\,]; \\ 0, & z \in (\,-\infty\,,\,-\phi_1(0)\,) \cup (\,-\psi_1(0)\,,\,\psi_1(0)\,) \cup (\,\phi_1(0)\,,\,+\infty\,); \\ -{\psi_1}^{-1}(-z), & z \in [\,-\psi_1(2e)\,,\,-\psi_1(0)\,]; \\ -2e, & z \in [\,-\phi_1(2e)\,,\,-\psi_1(2e)\,]; \\ -{\phi_1}^{-1}(-z), & z \in [\,-\phi_1(0)\,,\,-\phi_1(2e)\,]; \end{cases} \tag{3.4}$$

where ψ_1 and ϕ_1 denote, respectively, continuous strictly monotone increasing and decreasing functions, which are defined on the closed interval $[\,0\,,\,2e\,]$, and satisfy the inequality:

$$\psi_1(2e) \leq \phi_1(2e). \tag{3.5}$$

Consider $\mathbf{P}(F, d, L_e)$. If $\delta \in \tilde{\mathcal{C}}$, then the risk, $R(\delta,\, \theta,\, F)$, is a symmetric function of $\theta \in \Omega$. For $\theta \in [\,0, e) \cup (e, 3e\,]$:

$$\begin{aligned} R(\delta,\, \theta,\, F) &= F[\,\psi_1(\theta - e) - \theta\,] + F[\,-\phi_1(\theta - e) + \theta\,], \qquad e < \theta \leq 3e; \\ R(\delta,\, \theta,\, F) &= -F[\,\psi_1(\theta + e) - \theta\,] + F[\,\phi_1(\theta + e) - \theta\,] + \\ &\qquad F[\,-\psi_1(-\theta + e) - \theta\,] - F[\,-\phi_1(-\theta + e) - \theta\,],\ 0 \leq \theta < e. \end{aligned} \tag{3.6}$$

Let:

$$\gamma(\theta) = F[\psi_1(\theta + e) - \theta - 2e], \qquad \theta \in [\,-e\,,\,e\,]. \tag{3.7}$$

We use γ in the construction of an almost-equalizer rule for $\mathbf{P}(F, d, L_e)$ when F is Q-convex. From the risk expression, (3.6), it follows that a rule $\delta \in \tilde{\mathcal{C}}$ is an almost-equalizer rule with essential risk $c < 0.5$, if $\forall\, \theta \in [-e\,,\,e\,]$, $\gamma\,(\theta)$ satisfies:

$$\begin{aligned} c = {} & F\Big[\,2e - F^{-1}[\,c - \gamma(\theta)\,]\,\Big] - F\Big[\,2e - F^{-1}[\,1 - \gamma(\theta)\,]\,\Big] + \\ & F\Big[\,2e - F^{-1}[\,c - \gamma(-\theta)\,]\,\Big] - F\Big[\,2e - F^{-1}[\,1 - \gamma(-\theta)\,]\,\Big]. \end{aligned} \tag{3.8}$$

The almost-equalizer rule is parameterized by the pair of functions ψ_1 and ϕ_1 which are expressed in terms of γ by:

$$\psi_1(t) = t + e + F^{-1}[\,\gamma(t-e)\,], \qquad t \in [\,0\,,\,2e\,]; \tag{3.9}$$
$$\phi_1(t) = t + e - F^{-1}[\,c - \gamma(t-e)\,], \qquad t \in [\,0\,,\,2e\,]. \tag{3.10}$$

In terms of Q_c, (3.8) becomes:

$$Q_c(\,\gamma(\theta)\,) \;+\; Q_c(\,\gamma(-\theta)\,) \;=\; c, \qquad \forall\,\theta \in [\,-e\,,\,e\,]. \tag{3.11}$$

The strict monotonicity of ψ_1 and ϕ_1 implies that the following double inequality holds:

$$-f\Big(\,F^{-1}[\,\gamma(\theta)\,]\,\Big) \;\le\; \tfrac{d\gamma}{d\theta}(\theta) \;\le\; -f\Big(\,F^{-1}[\,c - \gamma(\theta)\,]\,\Big), \tag{3.12}$$

where γ is differentiable. Since F^{-1} is defined on $[\,0\,,\,1\,]$, it is necessary that $0 \le \gamma(0) \le c$. Further, in order that the interval defined by (3.12) be nonempty, it is necessary that $\gamma(\theta) \ge c/2$. Hence:

$$c/2 \;\le\; \gamma(\theta) \;\le\; c, \qquad \forall\,\theta \in [\,-e\,,\,e\,]. \tag{3.13}$$

Observation 3.1 The problem of delineating an almost-equalizer rule, $\delta \in \tilde{\mathcal{C}}$, when F is Q-convex, reduces to the problem of finding a continuous function, $\gamma : [\,-e\,,\,e\,] \to [\,c/2\,,\,c\,]$, satisfying (3.11)-(3.13), where c is the essential risk of δ.

Theorem 3.1 Let $d = 3e$. Assume F is symmetric, unimodal, and Q-convex. Let c_l denote the risk bound obtained in (3.2-3.3). If $c_l < 0.5$, then there exists: (i) a minimum value of $c > c_l$; (ii) unique values x_{min}, x_0, and x_{max}, where $x_{min} \le x_0 \le x_{max}$; and (iii) a continuous function ρ, such that:

(i) $\frac{dQ_c}{dx}(x_{min}) \;= 0$;
(ii) $Q_c(x_0) \;=\; c/2,\;\; x_{min} < x_0 < c$;
(iii) $Q_c(x_{min}) \;\le\; c/2$;
(iv) $Q_c(x_{min}) + Q_c(x_{max}) = c$; and
(v) $y = \rho(x)$ is the solution of the equation $Q_c(\,x\,) \;+\; Q_c(\,y\,) \;=\; c$ in the neighborhood $[\,x_{min}\,,\,x_{max}\,]$ of x_0, and $\forall\,x \in [x_{min}, x_{max}], \rho$ satisfies:

$$B_{low} \le B_{up}, \tag{3.14}$$

where:

$$B_{low} = \max\Big\{\,\frac{f\,(\,F^{-1}[\rho(x)]\,)}{\frac{d\rho}{dx}\,(x)}\;,\; -f\,\Big(\,F^{-1}[x]\,\Big)\,\Big\}, \tag{3.15}$$

$$B_{up} = \min\Big\{\,\frac{f\,(\,F^{-1}[c - \rho(x)]\,)}{\frac{d\rho}{dx}\,(x)}\;,\; -f\,\Big(\,F^{-1}[c - x]\,\Big)\,\Big\}. \tag{3.16}$$

The case $c_l \ge 0.5$ is trivial, since this contradicts the requirement that $c < 0.5$. In this case, the error-tolerance e is too tight.

The following algorithm for constructing the nonmonotone almost-equalizer rule with minimum essential risk c is based on Theorem 3.1.

Algorithm 3.1 A procedure for constructing the best nonmonotone almost-equalizer rule δ^e_{nm}:

1. input $d,\ d = 3e$.
compute c_l as in (3.2-3.3)
if $c_l \geq 0.5$ **quit**
else continue steps 2-6.
2. compute $c,\ x_{min},\ x_0,\ x_{max}$, and ρ as specified by Theorem 3.1.
3. define $\tilde{\gamma} : [-e, 0] \to [x_0, x_{max}]$ by: $\tilde{\gamma}(t) \ = \ -F[\, t + F^{-1}[c - x_0]\,] + c$.
4. set $\gamma(\theta) \ = \ \tilde{\gamma}(\theta), \qquad -e \leq \theta \leq 0;$
$\gamma(\theta) \ = \ \rho(\, \tilde{\gamma}(-\theta)\,), \quad 0 < \theta \leq e.$
5. set $\psi_1(t) \ = \ t + e + F^{-1}[\, \gamma(t-e)\,], \qquad 0 \leq t \leq 2e;$
$\phi_1(t) \ = \ t + e - F^{-1}[\, c - \gamma(t-e)\,], \ \ 0 \leq t \leq 2e.$
6. set δ^e_{nm} , as specified by (3.4).

When F is not $2e$-MLR, but is Q-convex, we obtain a relation between the minimax risks for the discrete- and the continuous-parameter problems, and obtain the monotone-minimax risk in terms of Q_c.

Let J_e be the family of nonmonotone piecewise-constant odd-symmetric decision rules parameterized by (a, b), $0 \leq a \leq b$:

$$\delta(z) = \begin{cases} 0, & b \leq z; \\ 2e, & a \leq z < b; \\ 0, & 0 \leq z < a. \end{cases} \tag{3.17}$$

Consider the discrete-parameter problem $\mathbf{P}(F, \{-2e, 0, 2e\}, L_0)$ which is an instance of the minimax problem $\mathbf{P}(F, \Omega, L_e)$, where $\Omega = \{2e, 0, 2e\}$ and L_0 is the zero-one loss function with error-tolerance $e = 0$. The minimax risk for the discrete-parameter problem is c_l. Utilizing the minimax rule for the discrete-parameter problem and a sequence of discrete priors, we obtain a sequence of Bayes rules for the continuous-parameter problem, such that the corresponding Bayes risk sequence converges to the minimax risk, c_l, for the discrete-parameter problem.

Let $\Pi_e = \{\pi_n\}_{n=1}^{\infty}$ denote the sequence of discrete symmetric priors on $\Omega = [\,-3e\,,\,3e\,]$:

$$\pi_n(\theta) = \begin{cases} K_0, & \theta = 0; \\ K_1, & \theta = \pm(2e + \frac{e}{n}); \\ 0, & \theta \in \Omega - \{\,0\,,\,\pm(2e + \frac{e}{n})\,\}; \end{cases} \tag{3.18}$$

where: $K_0 f(2e - F^{-1}[1 - x_l]) = K_1 f(F^{-1}[x_l])$, and $K_0 + 2K_1 = 1$.

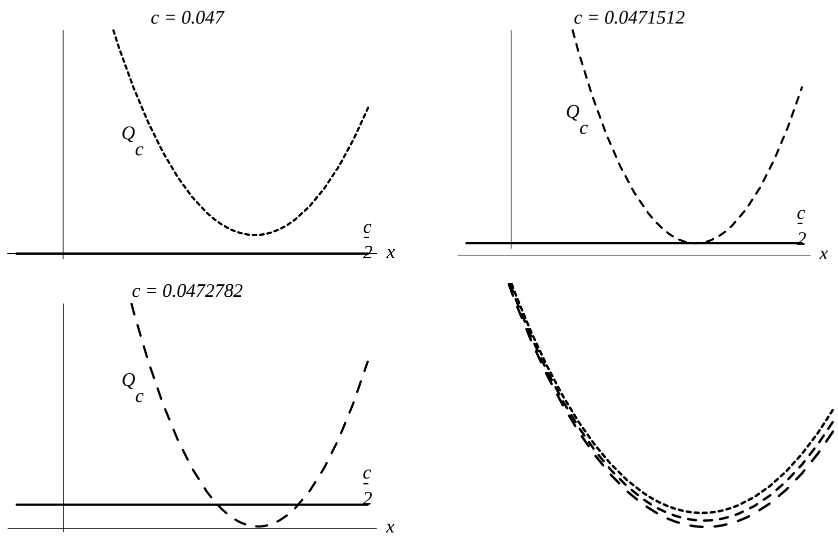

Fig. 2. The function $Q_c(x)$ for several values of c.

Theorem 3.2 Consider $\mathbf{P}(F, d, L_e)$. If F is symmetric, unimodal, and Q-convex, then there exists a sequence of rules $\{\delta_n\}_{n=1}^{\infty}$, such that: (i) $\forall n \geq 1$, $\delta_n \in J_e$, and δ_n is Bayes with respect to the prior π_n (3.18); and (ii) $\lim_{n\to\infty} r(\pi_n, \delta_n) = c_l$.

Corollary 3.1 If F is symmetric, unimodal, and Q-convex, then the minimax risk, c_l, for the discrete-parameter problem, $\mathbf{P}(F, \{-2e, 0, 2e\}, L_0)$, is a lower bound on the minimax risk for the continuous-parameter problem, $\mathbf{P}(F, d, L_e)$.

Example 3.1 Let F be the standard Cauchy distribution, $\mathcal{C}(0, 1)$. Thus,

$$Q_c(x) = \frac{1}{\pi}\Big(\arctan\{\, 2e - \tan(\pi(c - x - 0.5))\,\} - \arctan\{\, 2e - \tan(\pi(1 - x - 0.5))\,\}\Big),$$

and F is Q-convex. Let $e = 9.50998$. Figure 2 depicts Q_c for several values of c. In reference to Definition 3.2 and Theorem 3.1, the computed values of: x_l, c_l, x_{min}, x_0, and x_{max} are: $x_l = 0.0402499$, $c_l = 0.0471512$, $x_{min} = 0.0403134$, $x_0 = 0.0420462$, and $x_{max} = 0.0427575$. The computed value of the minimum essential risk c is: $c = 0.0472782$. This is the lowest value of the essential risk for which an almost-equalizer rule exists. These results are obtained by: (i) numerically solving for ρ in the neighborhood, $[\,x_{min}\,,\, x_{max}\,]$, where $x_{max} = \rho(x_{min})$, and $B_{low} \leq B_{up}$ (see Figure 3); and (ii) defining $\tilde{\gamma}$ and computing γ, ψ_1, and ϕ_1 as specified in Algorithm 3.1:

$$\tilde{\gamma}(t) = -0.5 - \frac{1}{\pi}\arctan\{t + \tan(\pi(c - 0.5 - x_0)) - c\}.$$

Figures 4 and 5 depict the computed functions γ, ψ_1 and ϕ_1. The best nonmonotone almost-

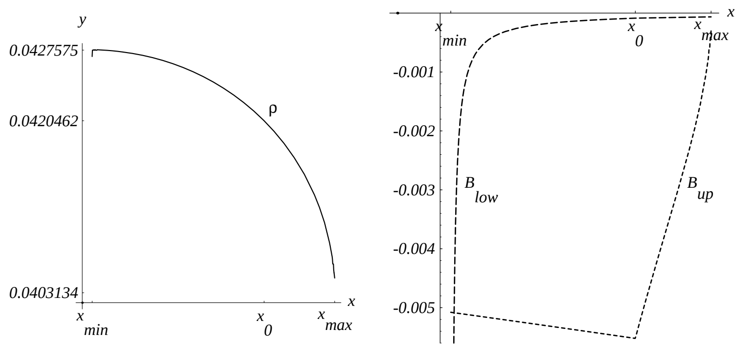

Fig. 3. The function $\rho(x)$, and the bounds B_{low} and B_{up}.

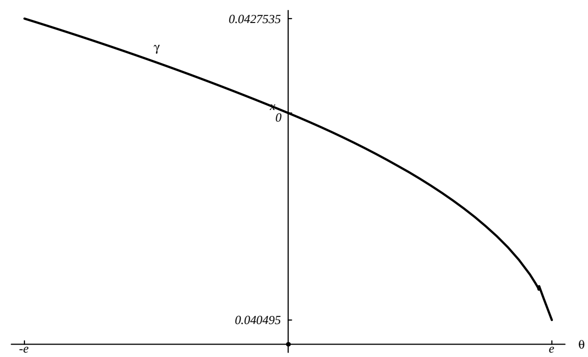

Fig. 4. The function $\gamma(\theta)$

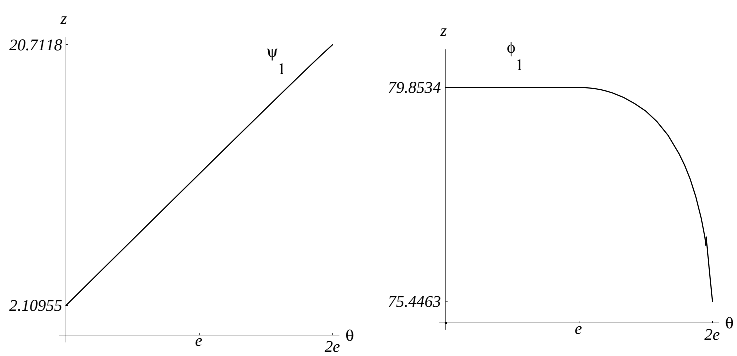

Fig. 5. The functions $\psi_1(\theta)$ and $\phi_1(\theta)$.

equalizer rule δ^e_{nm} is uniquely determined by ψ_1 and ϕ_1:

$$\delta^e_{nm}(z) = \begin{cases} 0, & 79.8541 \leq z; \\ {\phi_1}^{-1}(z), & 75.4488 \leq z < 79.8541; \\ 2e, & 20.8239 \leq z < 75.4488; \\ {\psi_1}^{-1}(z), & 2.10956 \leq z < 20.8239; \\ 0, & 0 \leq z < 2.10956. \end{cases} \tag{3.19}$$

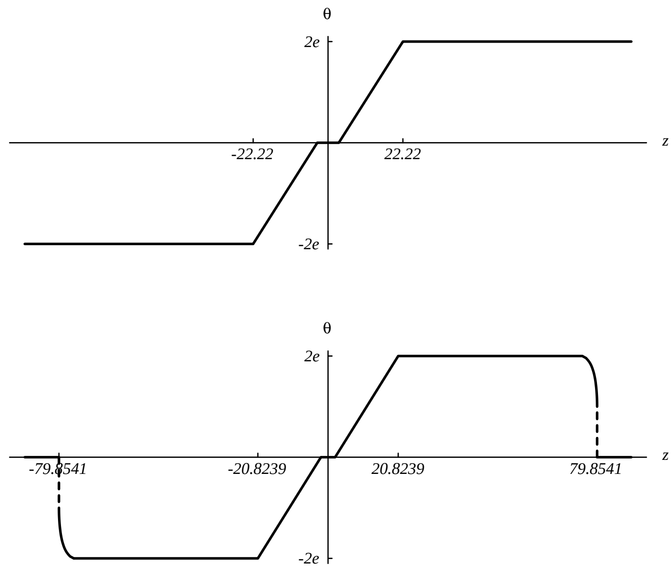

Fig. 6. Comparison of the monotone and the nonmonotone almost-equalizer rules, δ_m^e and δ_{nm}^e.

The best nonmonotone almost-equalizer rule δ_{nm}^e and the monotone-minimax almost-equalizer rule δ_m^e are depicted in Figure 6. The essential risks of $\delta_{nm}^e$and $\delta_m^e$are, respectively, 0.0472782 and 0.05. Further, $\forall\,\theta \in [\,-d, d\,]$, $R(\delta_{nm}^e, \theta, F) \,<\, R(\delta_m^e, \theta, F)$.

By Theorem 3.2 and Corollary 3.1, $c_l = 0.0471512$ is a lower bound on the global minimax risk. In this example, the error in this approximation is less than 0.3%. With reference to Theorem 3.2, the computed values of: a, b, K_0, K_1, a^n, b^n, and K are: $a = 11.1538$, $b = 64.1357$, $K_0 = 0.499314$, $K_1 = 0.250343$,

$$a^n = \frac{2e + \frac{e}{n} - \sqrt{(2e + \frac{e}{n})^2\, K - (1-K)^2}}{1-K},$$
$$b^n = \frac{2e + \frac{e}{n} + \sqrt{(2e + \frac{e}{n})^2\, K - (1-K)^2}}{1-K},$$

and $K = 0.501375$.

Observation 3.2 Consider $\mathbf{P}(F, d, L_e)$. Let M denote the essential risk of the monotone-minimax rule δ_m^e given by Theorem 2.2. Let Q_M denote Q_c, where $c = M$; thus, $Q_M(M) = M/2$. In terms of γ (3.7), the monotone-minimax risk is $M = \gamma(0)$.

Example 3.2 We compare the structure and behavior of the nonmonotone decision rules with the minimax rules for the corresponding decision problem under quadratic loss. Thus, we consider the decision problem $\mathbf{P}(F, d, L)$, where: $F = \mathcal{C}(0, 1)$, $L(a, \theta) = (a-\theta)^2$, and $|\theta| \le d$. The salient features of the solutions are: (i) each minimax rule is a continuous, antisymmetric, nonmonotone function of z; (ii) each minimax rule is defined by the conditional mean, $E[\Theta|z]$, with respect to a symmetric m-point discrete prior which is least favorable; and (iii) the minimax rule δ^* and the least favorable prior λ^* are each parameterized by d. We delineate solutions for $m \in \{2,3\}$:

Case 1: $m = 2$ and $0 < d \leq 2.82843$.

If $\lambda^*(d) = \frac{1}{2}$, and $\lambda^*(-d) = \frac{1}{2}$, then:

$$\delta^*(z) = E[\Theta \mid z] = \frac{2\,d^2\,z}{1 + d^2 + z^2}.$$

The corresponding risk function can be expressed compactly as:

$$R(\delta^*, \theta, F) = \frac{2\,d^4}{\sqrt{1+d^2}\,\left(d^2 + 2\,i\,\sqrt{1+d^2}\,\theta - \theta^2\right)} + \frac{\left(2\,i+\theta\right)^2\left(d^2-\theta^2\right)^2}{\left(d^2+\theta\;(2\,i+\theta)\right)^2} -$$

$$\frac{4\,d^2\,\left(\left(2\,\sqrt{1+d^2} + i\,\theta\right)\,\theta^2 + d^2\,\left(-\sqrt{1+d^2} + \sqrt{1+d^2}\,\theta^2 + i\,\theta^3\right)\right)}{\left(1+d^2\right)\,\left(d^2 + \left(2\,i\,\sqrt{1+d^2} - \theta\right)\,\theta\right)^2},$$

where: $i = \sqrt{-1}$.

The risk $R(\delta^*, \theta, F)$ is depicted in Figure 7 for three values of $d \in (0, 2.82843]$. For each $d \in (0, 2.82843]$, the pair (δ^*, λ^*) is, respectively, a minimax rule and a least favorable prior, since δ^* is Bayes with respect to λ^*, and λ^* assigns probability one to a two-point subset $\{-d, d\} \subset [-d, d]$ on which $R(\delta^*, \theta, F)$ is maximized.

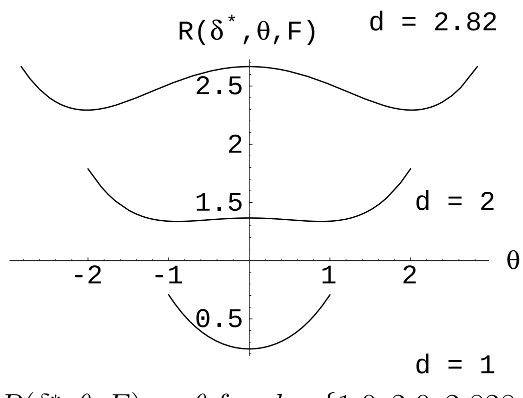

Fig. 7. $R(\delta^*, \theta, F)$ vs θ for $d \in \{1.0, 2.0, 2.82843\}$.

Case 2: $m = 3$ and $2.82843 < d \leq 3.70305$.

Let $q \in (0, 1]$. If $\lambda^*(d) = \frac{1}{2} - \frac{q}{2}$, $\lambda^*(0) = q$, and $\lambda^*(-d) = \frac{1}{2} - \frac{q}{2}$, then:

$$\delta^*(z) = E[\Theta \mid z] = \frac{2\,d^2\;(1-q)\;z\;(1+z^2)}{d^4\,q + \left(1+z^2\right)^2 + d^2\;(1 + q + z^2 - 3\,q\,z^2)}.$$

The risk $R(\delta^*, \theta, F)$ is depicted in Figure 8 for three values of $d \in (2.82843, 3.70305]$. For each $d \in (2.82843, 3.70305]$, the pair (δ^*, λ^*) is, respectively, a minimax rule and a least favorable prior, since δ^* is Bayes with respect to λ^*, and λ^* assigns probability one to a three-point subset $\{-d, 0, d\} \subset [-d, d]$ on which $R(\delta^*, \theta, F)$ is maximized.

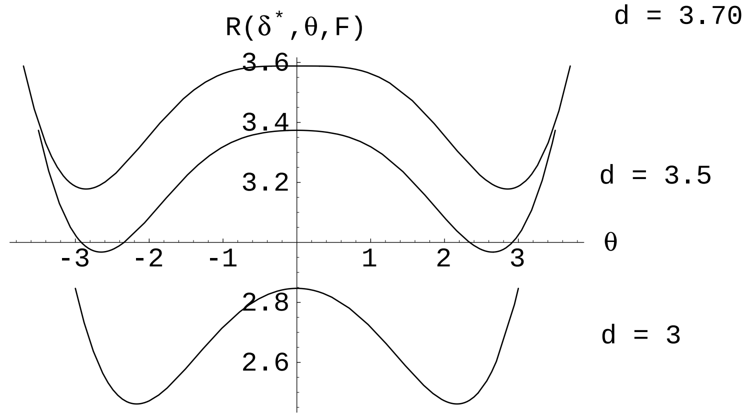

Fig. 8. $R(\delta^*, \theta, F)$ vs θ for $(d, q) \in \{(3.0, 0.0123457), (3.5, 0.0385488), (3.70305, 0.0462882)\}$.

The structure and behavior of minimax rules under quadratic loss for a restricted location parameter and absolutely continuous sampling distributions appears in Cherkassky and Mintz (1998).

4 Conclusions

For the case $d = 3e$, we summarize the salient connections between the behavior of the CDFs F and the structure of the minimax and near-minimax rules that we have obtained. We also highlight some bounds on the global minimax risk that we have derived. These connections and bounds are determined by the behavior of the distribution-dependent family of functions $\{Q_c : 0 < c < 0.5\}$.

Consider the continuous-parameter problem, $\mathbf{P}(F, d, L_e)$. Let f be symmetric and unimodal with convex support. Fix $M \in (0, 0.5)$. Let $e = -(F^{-1}[M/2] + F^{-1}[M])/2$. If $Q_M(x)$ is strictly convex and decreasing, then the monotone-minimax rule δ_m^e (Theorem 2.2) is Bayes, globally minimax, and admissible, with essential risk M.

When F is Q-convex, the family $\{Q_c : 0 < c < 0.5\}$ provides a common framework for determining: (i) a nonmonotone minimax rule for the discrete-parameter problem; and (ii) the monotone-minimax rule δ_m^e, and the best almost-equalizer rule δ_{nm}^e for the continuous-parameter problem. Let (c_l, x_l) be the unique solution of (3.2)–(3.3), then:

(a) For the discrete-parameter problem $\mathbf{P}(F, \{-2e, 0, 2e\}, L_0)$ the minimax risk is c_l.
(b) For the continuous-parameter problem $\mathbf{P}(F, d, L_e)$:
  – There is a sequence of Bayes rules which converges in Bayes risk to c_l. Therefore, c_l is a lower bound on the global minimax risk.

– Let c denote the essential risk of the best nonmonotone almost-equalizer rule δ^e_{nm}.
– Let M denote the essential risk of δ^e_m, where $Q_M(M) = M/2$.
– If c^* denotes the global minimax risk, then the following relations hold:

$$c_l \ < \ c^* \ \leq \ c \ < \ M, \text{ and } \frac{M - c_l}{2} \ = \ Q_M(M) - Q_{c_l}(x_l).$$

Thus, using the family of functions $\{Q_c : 0 < c < 0.5\}$, we obtain bounds on the minimax risk and its deviation from the essential risk of the best almost-equalizer rule. The monotone almost-equalizer rules are easy to compute compared to the nonmonotone almost-equalizer rules. In those cases for which the global minimax risk is suitably close to the monotone-minimax risk, the monotone rules have near-minimax performance, and can be used in place of the minimax rules. When the monotone almost-equalizer rule does not have a satisfactory risk, then the best nonmonotone almost-equalizer rule could be used. This nonmonotone rule has very good performance, even for such an extreme distribution as the Cauchy. As illustrated in Example 3.1, the risk of the nonmonotone almost-equalizer rule, for $\mathbf{P}(F, d, L_e)$, where F is $\mathcal{C}(0, 1)$, is within 0.3% of the global minimax risk.

The discrete-parameter problem $\mathbf{P}(F, \{-2e, 0, 2e\}, L_0)$, where F is Cauchy, was solved in McKendall (1990). By means of Q_c, we reconcile within a single framework, both the discrete- and continuous-parameter problems.

We derive minimax rules for a restricted location parameter under zero-one loss, for sampling distributions which satisfy a weak MLR condition, i.e., the $2e$-MLR distributions. The location data model and zero-one loss lead to a fixed size confidence interval with the highest confidence coefficient. Location estimation problems with restricted parameter spaces arise, for example, in decision and control problems with mobile robots. For example, it may be necessary to provide precise estimates of a mobile robot's position and orientation in a constrained workspace using limited sensor data. Examples of such applications appear in Kamberova, Mandelbaum and Mintz (1996).

Extending the class of distributions, for which minimax and near-minimax rules can be computed, has practical implications. With this larger class, the user can potentially obtain a better match between the empirical sampling distribution and the distribution model selected for the given decision problem. We extend the class of problems which can be addressed by obtaining either exact solutions or highly accurate approximations. Our solutions to the single-sample minimax problems provide the basis for the solutions to both the single-sample and the multi-sample robust minimax problems which appears in Kamberova and Mintz (1998b). We also compare the structure and behavior of the nonmonotone decision rules with the minimax rules for the corresponding decision problem under quadratic loss.

References

[1] Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis*, Springer-Verlag, New York.

[2] Brown, L., A. Cohen and W. Strawderman (1986). “A complete class theorem for strict monotone likelihood ratio with applications,” *Ann.Statist.* **4**, 712-722.

[3] Cherkassky, D. and M. Mintz (1998). “Minimax rules under quadratic loss for a restricted location parameter,” preprint.

[4] Ferguson, T. (1967). *Mathematical statistics: a decision theoretic approach*, Academic Press, Inc.

[5] Kamberova, G. (1992). *Robust location estimation for MLR and non-MLR distributions*, PhD Dissertation, Tech. report MS-CIS-92-93, Department of Computer and Information Science, University of Pennsylvania, Philadelphia.

[6] Kamberova G., R. Mandelbaum and M. Mintz (1996). “Statistical decision theory for mobile robotics: theory and application,” *Proc. IEEE/SICE/RSJ Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems, MFI'96*, 17–24, Washington DC.

[7] Kamberova, G. and M. Mintz (1998a). “Almost-equalizer, Bayes and minimax decision rules,” Tech. Report MS-CIS-98-25, Department of Computer and Information Science, University of Pennsylvania, Philadelphia. Available at:
`http://www.cis.upenn.edu/~kamberov/public_html/doc/papers.html.`

[8] Kamberova G. and M. Mintz (1998b). “Uncertainty classes for robust minimax estimation under zero-one loss,” preprint.

[9] Karlin, S., H. Rubin (1956). “The theory of decision procedures for distributions with monotone likelihood ratio,” *Ann.Math.Statist.* **27**, 272-299.

[10] McKendall, R. (1992). *Minimax estimation of a discrete location parameter for a continuous distribution*, PhD Dissertation, Department of Systems Engineering, University of Pennsylvania, Philadelphia.

[11] Zeytinoglu, M. and M. Mintz (1984). “Optimal fixed sized confidence procedures for a restricted parameter space,” *The Annals of Statistics*, **12** 945–957.

[12] Zeytinoglu, M. and M. Mintz (1988). “Robust optimal fixed sized confidence procedures for a restricted parameter space,” *The Annals of Statistics*, **16** 1241–1253.