August 1990

Non-Monotonic Decision Rules for Sensor Fusion

Raymond McKendall  
*University of Pennsylvania*

Max L. Mintz  
*University of Pennsylvania, mintz@cis.upenn.edu*

Follow this and additional works at: [http://repository.upenn.edu/cis_reports](http://repository.upenn.edu/cis_reports)

Recommended Citation


This paper is posted at ScholarlyCommons, [http://repository.upenn.edu/cis_reports/574](http://repository.upenn.edu/cis_reports/574)
For more information, please contact repository@pobox.upenn.edu.
Non-Monotonic Decision Rules for Sensor Fusion

Abstract
This article describes non-monotonic estimators of a location parameter from a noisy measurement \( Z = \Theta + V \) when the possible values of \( \epsilon \) have the form \( \{0, \pm 1, \pm 2, \ldots, \pm n\} \). If the noise \( V \) is Cauchy, then the estimator is a non-monotonic step function. The shape of this rule reflects the non-monotonic shape of the likelihood ratio of a Cauchy random variable. If the noise \( V \) is Gaussian with one of two possible scales, then the estimator is also a nonmonotonic step function. The shape this rule reflects the non-monotonic shape of the likelihood ratio of the marginal distribution of \( Z \) given \( \Theta \) under a least-favorable prior distribution.

Comments

This technical report is available at ScholarlyCommons: http://repository.upenn.edu/cis_reports/574
Non-Monotonic Decision Rules
For Sensor Fusion

MS-CIS-90-56
GRASP LAB 228

Raymond McKendall
Max Mintz

Department of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104

August 1990
Non-Monotonic Decision Rules for Sensor Fusion

Raymond McKendall and Max Mintz*
GRASP Laboratory
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104-6389

Abstract
This article describes non-monotonic estimators of a location parameter \( \theta \) from a noisy measurement \( Z = \theta + V \) when the possible values of \( \theta \) have the form \( \{0, \pm 1, \pm 2, \ldots, \pm n\} \). If the noise \( V \) is Cauchy, then the estimator is a non-monotonic step function. The shape of this rule reflects the non-monotonic shape of the likelihood ratio of a Cauchy random variable. If the noise \( V \) is Gaussian with one of two possible scales, then the estimator is also a non-monotonic step function. The shape this rule reflects the non-monotonic shape of the likelihood ratio of the marginal distribution of \( Z \) given \( \theta \) under a least-favorable prior distribution.

1 Introduction
This article describes non-monotonic estimators in decision problems motivated by sensor fusion. It finds minimax rules under zero-one (0) loss for the location parameter \( \theta \) in two problems of the fusion paradigm \( Z = \theta + V \). The statistical background for this research is reviewed in the article Statistical Decision Theory for Sensor Fusion [McKenall, 1990b] of these Proceedings, which also defines notation and terminology.

The first problem is a standard-estimation problem in which \( \theta \in \{0, \pm 1, \pm 2, \ldots, \pm n\} \), for a given integer \( n \), and in which the noise \( V \) has the standard Cauchy distribution. A motivation for these assumptions is extension of the results of [Zeytinoglu and Mintz, 1984] and [McKendall, 1990a] that assume the distribution of \( V \) has a monotone likelihood ratio. The noise distributions in most practical applications do not have monotone likelihood ratios; the Cauchy distribution is a simple distribution that does not have a monotone likelihood ratio. The minimax rule for this problem is a non-monotonic function. In contrast, the decision rules corresponding to a noise distribution with a monotone likelihood ratio are monotonic functions.

The second problem is a robust-estimation problem in which \( \theta \in \{-1, 0, 1\} \) and the noise \( V \) has either the \( \mathcal{N}(0, \sigma^2) \) or the \( \mathcal{N}(0, \sigma^2_2) \) distribution. If the maximum allowable scale is not too large, the robust-estimation problems of [Zeytinoglu and Mintz, 1988] and [McKendall, 1990a] reduce to standard-estimation problems. The underlying distributions in these problems have a monotone likelihood ratio (in the location parameter), and so their minimax rules are monotonic. In contrast, this problem has a non-monotonic minimax rule because the maximum scale is too large. (A similar problem in which the possible locations are an interval has a randomized minimax rule. [Martin, 1987].)

Section 2 discusses the standard-estimation problem with the Cauchy noise distribution. Section 3 discusses the robust-estimation problem with uncertain noise distribution. The results listed here are a synopsis of results in [McKendall, 1990a], which gives the underlying analysis and the proofs.

2 Cauchy Noise Distribution
This section constructs a ziggurat minimax rule \( \delta^* \) for the location parameter in a standard-estimation problem \((\Theta_n, \Theta_0, L_0, Z)\) in which \( Z \) has a Cauchy distribution. A ziggurat decision rule is a non-monotonic step function with range \( \Theta_n \). The non-monotonicity of \( \delta^* \) reflects the non-monotonicity of the likelihood ratio of a Cauchy distribution. The range of \( \delta^* \) reflects the structure of the zero-one (\( \epsilon \)) loss function.

Section 2.1 reviews the Cauchy distribution. Section 2.2 summarizes the main results. The remaining sections develop these results in more detail. Their organization follows the strategy for finding a minimax decision rule by finding a Bayes equalizer rule. Section 2.3 defines ziggurat decision rules. Section 2.4 discusses Bayes analysis of a ziggurat decision rule. Sections 2.5, 2.6, and 2.7 give the risk analysis of a ziggurat decision rule. Section 2.8 combines the conclusions of this chapter to find an admissible minimax estimator.

2.1 Cauchy Distribution
A continuous random variable \( V \) has the Cauchy distribution with location parameter \( \mu \) and unit scale, written

$V \sim C(\mu, 1)$, if its density function $f$ is

$$f(v|\mu) = \frac{1}{\pi (1 + (v - \mu)^2)}.$$ 

The distribution function of a $C(\mu, 1)$ random variable is

$$F(v|\mu) = \frac{1}{\pi} \arctan(\frac{v - \mu}{1}) + \frac{1}{2}.$$ 

The $C(0, 1)$ distribution is the standard Cauchy distribution. An important property of a Cauchy distribution is that it does not have a monotone likelihood ratio. Figure 1 illustrates the shape of these ratios.

**2.2 Introduction**

This section introduces and summarizes the results through an example. In particular, it shows how to construct a minimax rule $\delta^*$ and a least-favorable probability function $\pi^*$ on $\Theta_n$ for the standard-estimation problem $(\Theta_n, \Theta_n, L_0, Z)$ in which $n = 2$ and $F$ is the $C(0, 1)$ distribution. The general results have arbitrary $n$.

The decision rule $\delta^*$, defined by figure 2, is the ziggurat decision rule over a partition $\{x_i\}_0^5$ of $\mathbb{R}^+$ onto $\Theta_2$: It is an even, non-monotonic step function with range $\Theta_2$ and with steps of unit height occurring at points of $\{x_i\}$. The points $x_1$ and $x_2$ are chosen so that $\delta^*$ is an equalizer rule. The points $x_3$ and $x_4$ and the positive probability function $\pi^*$ are constructed from $x_1$ and $x_2$ so that $\delta^*$ is Bayes against $\pi^*$. Consequently, the rule $\delta^*$ is admissible and minimax, and the probability function $\pi^*$ is least favorable.

The partition $\{x_i\}$ requires solution of the ziggurat-equalizer equations:

$$2h_0(y_1) = g_1(y_1) + h_1(y_2) = g_3(y_3)$$

The functions $g_i$ and $h_i$ are these:

$$g_i(x) := F(x-i) + F(i-\mu_i(x)), \quad i = 1, 2$$

$$h_i(x) := F(\mu_i+1(x) - i) + F(x-i), \quad i = 0, 1$$

The function $\mu_i$ is this:

$$\mu_i(x) := \begin{cases} 
  i - \frac{1}{2} & \text{if } x = i - \frac{1}{2} \\
  (i - \frac{1}{2})x - (i - \frac{1}{2})^2 + v_i^2 & \text{if } x \neq i - \frac{1}{2}
\end{cases}$$

$$v := \frac{1}{2}\sqrt{5}$$

These equations have unique solution $y_1$, $y_2$ such that

$$y_1 \in (\frac{3}{2}, \frac{1}{2} + v_1) \text{ and } y_2 \in (\frac{3}{2}, \frac{3}{2} + v_1).$$

Furthermore, $y_1 < y_2$. (The solution may be computed numerically by the Newton-Raphson method.) The partition $\{x_i\}$ is defined in terms of this solution:

$$x_0 := 0$$

$$x_1 := y_1$$

$$x_2 := y_2$$

$$x_3 := \mu_2(y_2)$$

$$x_4 := \mu_1(y_1)$$

$$x_5 := \infty$$

This partition is a $\mu_i$-constrained partition of $\mathbb{R}^+$.

The probability function $\pi^*$ is this:

$$\pi^*(\pm 1) = \pi^*(0)/\rho(1)$$

$$\pi^*(\pm 2) = \pi^*(0)/(\rho(1)\rho(2))$$

The factors $\rho(\pm l)$ connect $\pi^*$ to $\{x_i\}$ and thus to $\delta^*$:

$$\rho(l) := \frac{f_Z(x_i|l)}{f_Z(x_i|l-1)} = 1/\rho(-l)$$

The probability function $\pi^*$ is positive and unique.

**2.3 Ziggurat Decision Rule**

This section defines and illustrates ziggurat decision rules. A ziggurat rule is specified in terms of a partition of $\mathbb{R}^+$.
Notation: For integers \( p \leq q \), the notation \( \mathcal{I}_p^q \) means the integers from \( p \) to \( q \). For example, \( \mathcal{I}_0^p = \{0, 1, \ldots, p\} \).

Definition: partition of \( \mathbb{R}^+ \) A partition\(^2\) of \( \mathbb{R}^+ \) is a set of points \( \{x_i\}_{i}^{p+1} \) such that \( x_0 = 0 \), \( x_{p+1} = \infty \), and \( x_{i+1} > x_i \) for \( i \in \mathcal{I}_p^q \). Such a partition is abbreviated as \( \{x_i\} \).

Example 2.1 A partition of \( \mathbb{R}^+ \) with \( p = 4 \) is
\[
\{x_i\}^5_0 = \{0, 0.617, 1.912, 4.536, 11.209, \infty\}.
\]

Remark A particular partition of \( \mathbb{R}^+ \) is specified by the points \( x_i, i \in \mathcal{I}_p^q \). The specification of \( x_0 \) and \( x_{p+1} \) is implicit.

Definition: ziggurat decision rule Let \( \{x_i\}_{i}^{n+1} \) be a partition of \( \mathbb{R}^+ \). The ziggurat decision rule \( \delta \) over \( \{x_i\} \) onto \( \Theta_n \) is this:
\[
\delta(z) := \begin{cases} 
  i & \text{if } x_i \leq z < x_{i+1}, \quad i = 0, \ldots, n \\
  n+1 & \text{if } x_n+1 \leq z < x_{n+1}, \quad i = 1, \ldots, n \\
  -\delta(-z) & \text{if } z \leq 0
\end{cases}
\]

Example 2.2 Let \( n = 2 \). Define \( \delta \):
\[
\delta(z) := \begin{cases} 
  0 & \text{if } 0 \leq z < x_1 \\
  z & \text{if } x_1 \leq z < x_2 \\
  2z & \text{if } x_2 \leq z < x_3 \\
  z & \text{if } x_3 \leq z < x_4 \\
  0 & \text{if } x_4 \leq z \\
  -\delta(-z) & \text{if } z \leq 0
\end{cases}
\]

Then \( \delta \) is the ziggurat decision rule over the partition \( \{0, x_1, x_2, x_3, x_4, \infty\} \) onto \( \Theta_2 \).

Remark The ziggurat rule over \( \{x_i\}_{0}^{2n+1} \) steps between \( i-1 \) and \( i \) at \( x_i \) and between \( i \) and \( i-1 \) at \( x_{2n+1-i}, i \in \mathcal{I}_1^n \).

Remark The term ziggurat loosely describes the shape of the rule over \( \mathbb{R}^+ \): A ziggurat is a terraced pyramid.\(^2\)

2This definition differs from the set-theoretic definition of some contexts.

2.4 Bayes Rule

Notation

Bayes analysis of a ziggurat rule for a decision problem \((\Theta_n, \Theta_n, L_0, Z)\) in which \( Z \) has a Cauchy distribution requires \( \mu \)-constrained partitions of \( \mathbb{R}^+ \).

Notation \( \xi_i := (i - \frac{1}{2}, i - \frac{1}{2} + v) \)

Definition: \( \mu \)-constrained partition of \( \mathbb{R}^+ \) A \( \mu \)-constrained partition of \( \mathbb{R}^+ \) is a partition \( \{x_i\}_{0}^{2n+1} \) of \( \mathbb{R}^+ \) such that for all \( i \in \mathcal{I}_n^p \),

\[
x_i \in \xi_i
\]

and

\[
x_{2n+1-i} = \mu_1(x_i).
\]

Example 2.3 A \( \mu_1 \)-constrained partition of \( \mathbb{R}^+ \) has the following structure:

\[
\{0, x_1, x_2, \ldots, x_{n-1}, x_n, \mu_n(x_n), \mu_{n-1}(x_{n-1}), \ldots, \mu_2(x_2), \mu_1(x_1), \infty\}
\]

Furthermore, \( x_i \in \xi_i \).

Example 2.4 Let \( n = 2 \). Define \( x_1, x_2, x_3, x_4 \):

\[
x_1 := 0.617, \quad x_2 := 1.912, \quad x_3 := 4.536, \quad x_4 := 11.209.
\]

Note that \( x_1 \in \xi_1 \) and \( x_2 \in \xi_2 \):

\[
\frac{1}{2} < x_1 < \frac{1}{2} + \frac{1}{2}\sqrt{5} = 1.618
\]

\[
\frac{3}{2} < x_2 < \frac{3}{2} + \frac{1}{2}\sqrt{5} = 2.618
\]

Verify that \( x_3 = \mu_2(x_2) \) and \( x_4 = \mu_1(x_1) \). Therefore, \( \{0, x_1, x_2, x_3, x_4, \infty\} \) is a \( \mu_1 \)-constrained partition of \( \mathbb{R}^+ \).

Remark Let \( \{x_i\}_{0}^{2n+1} \) be a \( \mu_1 \)-constrained partition of \( \mathbb{R}^+ \). The ziggurat rule over \( \{x_i\} \) steps between \( i-1 \) and \( i \) at \( x_i \) and between \( i \) and \( i-1 \) at \( \mu_1(x_i), i \in \mathcal{I}_1^n \).

Remark Let \( f_Z(|i|) \sim C(i, 1) \), where \( i \) is an integer. The function \( \mu_i \) satisfies the identity

\[
\frac{f_Z(\mu_i(x)|i+e)}{f_Z(\mu_i(x)|i-e-1)} = \frac{f_Z(\mu_i(x)|i+e)}{f_Z(\mu_i(x)|i-e-1)}, \quad \forall x \in \mathbb{R}.
\]

This is the functional definition of \( \mu_i \). Bayes analysis motivates this definition. The algebraic definition of \( \mu_i \) is derived from the functional definition.
Main Result
Proposition 1 shows that to any ziggurat decision rule $\delta$ over a $\mu_i$-constrained partition of $\mathbb{R}^+$, there corresponds a positive probability function $\pi$ on $\Theta_n$ such that $\delta$ is Bayes against $\pi$.

**Proposition 1** Assume $F \sim C(0,1)$. Let $\{x_i\}_{i=1}^{n+1}$ be a $\mu_i$-constrained partition of $\mathbb{R}^+$. Let $\pi$ be the even, positive probability function on $\Theta_n$ such that for all $l \in I_n^1$,

$$\pi(l - 1) = \rho(l) \pi(l).$$

The ziggurat decision rule over $\{x_i\}$ onto $\Theta_n$ is Bayes against $\pi$.

**Example 2.5** Let $n = 2$. Let $\{x_i\}_{i=1}^{5}$ be the $\mu_i$-constrained partition of $\mathbb{R}^+$ given in example 2.4:

$$\{x_i\} = \{0, 0.617, 1.912, 4.536, 11.209, \infty\}$$

Let $\delta$ be the ziggurat decision rule over $\{x_i\}$ onto $\Theta_2$:

$$\delta(z) = \begin{cases} 
0 & \text{if } 0 \leq z < 0.616 \\
1 & \text{if } 0.616 \leq z < 1.912 \\
2 & \text{if } 1.912 \leq z < 4.536 \\
1 & \text{if } 4.536 \leq z < 11.209 \\
0 & \text{if } 11.209 \leq z \\
-\delta(-z) & \text{if } z < 0
\end{cases}$$

Then $\delta$ is Bayes against some positive probability function on $\Theta_2$. \hfill $\Box$

**Example 2.6** Consider example 2.5. The conditions of proposition 1 for a probability function $\pi$ on $\Theta_2$ are these:

$$\pi(0) = \rho(1) \pi(1)$$

$$\rho(1) = \frac{f_Z(x_1)}{f_Z(x_2)} = \frac{f(0.617) - 1}{f(0.617)} = 1.204$$

$$\rho(2) = \frac{f_Z(x_2)}{f_Z(x_1)} = \frac{f(1.912) - 2}{f(1.912) - 1} = 1.818$$

Also, $\pi(-1) = \pi(1)$ and $\pi(-2) = \pi(2)$. Hence:

$$\sum_{\theta} \pi(\theta) = \pi(0) \left(1 + 2 \frac{\rho(1)}{\rho(1) + 2} + \frac{2}{\rho(1) + 2}\right) = 3.575 \pi(0)$$

Thus $\pi$ assigns these probabilities:

$$\pi(0) = 0.280$$

$$\pi(\pm 1) = 0.232$$

$$\pi(\pm 2) = 0.128$$

Therefore, the ziggurat decision rule over $\{x_i\}^5$ onto $\Theta_2$ is Bayes against the probability function $\pi$ on $\Theta_2$. \hfill $\Box$

**Example 2.7** The probability function $\pi$ of proposition 1 is given by the following equations: For all $l \in I_n^1$,

$$\pi(\pm l) = \left(\prod_{k=1}^{l} \frac{f_Z(x_k)}{f_Z(x_{k+1})}\right)^{-1} \pi(0),$$

where

$$\pi(0) = \left[1 + 2 \sum_{l=1}^{n} \left(\prod_{k=1}^{l} \frac{f_Z(x_k)}{f_Z(x_{k-1})}\right)^{-1}\right]^{-1}$$

**Remark** In proposition 1, the restriction to a $\mu_i$-constrained partition of $\mathbb{R}^+$ and the conditions on the probability function are necessary for the decision rule to minimize the posterior expected loss.

### 2.5 Risk Function
Proposition 2 gives the risk function of a ziggurat decision rule over a $\mu_i$-constrained partition of $\mathbb{R}^+$.

**Proposition 2** Let $\{x_i\}_{i=1}^{n+1}$ be a $\mu_i$-constrained partition of $\mathbb{R}^+$, and let $\delta$ be the ziggurat decision rule over $\{x_i\}$ onto $\Theta_n$.

$$R(0, \delta) = 2h_0(x_1)$$

$$R(\pm i, \delta) = g_i(x_i) + h_i(x_{i+1}), \quad i \in I_n^1$$

$$R(\pm n, \delta) = g_n(x_n)$$

**Example 2.8** Let $n = 3$. Let $\{x_i\}_{i=1}^{7}$ be a $\mu_i$-constrained partition of $\mathbb{R}^+$, and let $\delta$ be the ziggurat decision rule over $\{x_i\}$ onto $\Theta_3$.

$$R(0, \delta) = 2h_0(x_1)$$

$$R(\pm i, \delta) = g_i(x_i) + h_i(x_{i+1}), \quad i \in I_n^1$$

$$R(\pm n, \delta) = g_n(x_n)$$

**2.6 Ziggurat-Equalizer Equations**
Equating the expressions $R(\theta, \delta)$ over $\theta \in \Theta_n$ to find a ziggurat equalizer rule leads to the ziggurat-equalizer equations. These are $n$ equations in $n$ unknowns $y_1, \ldots, y_n$. For $n = 1$, the ziggurat-equalizer equation is

$$2h_0(y_1) = g_1(y_1).$$

For $n \geq 2$, the ziggurat-equalizer equations are

$$2h_0(y_1) = g_1(y_1) + h_1(y_{i+1}) = g_n(y_n), \quad i \in I_n^1.$$

**Example 2.9** The ziggurat-equalizer equations for $n = 2$ are these:

$$2h_0(y_1) = g_1(y_1) + h_1(y_2) = g_2(y_2).$$

The ziggurat-equalizer equations for $n = 3$ are these:

$$2h_0(y_1) = g_1(y_1) + h_1(y_2) = g_2(y_2) + h_2(y_3) = g_3(y_3).$$

Proposition 3 states that the ziggurat-equalizer equations have a unique solution $y_1, \ldots, y_n$ that has certain properties. Proposition 4 uses this solution to construct an equalizer rule.

**Proposition 3** Assume $F \sim C(0,1)$. The ziggurat-equalizer equations have unique, increasing solution $y_1, \ldots, y_n$ with $y_i \in E_{l_i}$. Furthermore $y_i - y_{i-1} > 1$ for $l \in I_n^2$.

**Example 2.10** Let $F \sim C(0,1)$. The ziggurat-equalizer equations for $n = 3$ and $u = 1$ have the following solution:

$$y_1 = 0.570743$$

$$y_2 = 1.731856$$

$$y_3 = 2.979961$$

Here, $y_1 \in (0.5, 0.5 + v_1)$, $y_2 \in (1.5, 1.5 + v_1)$, and $y_3 \in (2.5, 2.5 + v_1)$. Also $y_2 - y_1 > 1$ and $y_3 - y_2 > 1$. \hfill $\Box$
2.7 Equilizer Rule

Proposition 4 gives a ziggurat equilizer rule.

**Proposition 4** Assume $F \sim C(0,1)$. Let $y_1, \ldots, y_n$ with $y_i \in \xi_i$ satisfy the ziggurat-equilizer equations. For $i \in T_n^*$, define

$$x_i := y_i \text{ and } x_{2n+1-i} := \mu_i(y_i).$$

Also, define $x_0 := 0$ and $x_{2n+1} := \infty$. If $\{x_i\}_{0}^{2n+1}$ is a partition of $\mathbb{R}^+$, then the ziggurat decision rule $\delta$ over $\{x_i\}$ onto $\Theta_n$ is an equilizer rule. Furthermore, if $\{x_i\}$ is a partition of $\mathbb{R}^+$, then the common risk of $\delta$ is $R_\delta = g_n(n) = F(-\frac{1}{2}) < R_\delta < 2F(-\frac{1}{2})$.

**Example 2.11** Let $n = 3$. The solution $y_1, y_2, y_3$ to the ziggurat-equilizer equations specified by the proposition is

$$y_1 = 0.571, \quad y_2 = 1.732, \quad y_3 = 2.980.$$  

Let $x_1 := y_1, x_2 := y_2, \text{ and } x_3 := y_3$. Also, define $x_4, x_5, \text{ and } x_6$ as follows:

$$
\begin{align*}
x_4 & := \mu_3(x_3) = 5.104 \\
x_5 & := \mu_2(x_2) = 6.891 \\
x_6 & := \mu_1(x_1) = 18.170
\end{align*}
$$

Note that $\{x_i\}$ is a partition of $\mathbb{R}^+$:

$$\{x_i\} = \{0, 0.571, 1.732, 2.980, 5.104, 6.891, 18.170, \infty\}.$$  

Thus, the ziggurat decision rule over $\{x_i\}$ onto $\Theta_3$ is an equilizer. Its risk is $R_\delta = g_3(x_3):=

$$
\begin{align*}
g_3(x_3) & = F(x_3 - 3) + F(3 - \mu_3(x_3)) \\
& = F(x_3 - 3) + F(3 - x_4) \\
& = 0.635
\end{align*}
$$

Here, $0.352 = F(-\frac{1}{2}) < R_\delta < 2F(-\frac{1}{2})$.  

**Example 2.12** Refer to example 2.5: Verify that $y_1 := 0.617$ and $y_2 := 1.912$ satisfy the ziggurat-equilizer equations for $n = 2$. Thus, since $\{x_i\}$ is a $\mu_1$-constrained constrained partition of $\mathbb{R}^+$, the ziggurat decision rule over $\{x_i\}$ is an equilizer rule.$\square$

**Remark** Proposition 3 asserts that $x_1, \ldots, x_n$ exist and that $x_i > x_{i-1}, i \in T_n^*$. There is no guarantee, however, that $\{x_i\}_{0}^{2n+1}$ is a partition of $\mathbb{R}^+$; it is necessary to verify that $\mu_i(x_{i-1}) > \mu_i(x_i), i \in T_n^*$. If $\{x_i\}$ is a partition of $\mathbb{R}^+$, then it is a $\mu_i$-constrained partition by construction. Numerical computations suggest that $\{x_i\}$ is in fact a partition of $\mathbb{R}^+$, but there is no proof of this conjecture.

2.8 Minimax Rule

Theorem 1 combines the conclusions of this chapter to find an admissible minimax estimator of the location parameter $\theta$ for a decision problem $(\Theta_n, \Theta_n, L_0, Z)$ in which $Z$ has a Cauchy distribution.

**Theorem 1** Assume $F \sim C(0,1)$. Let $y_1, \ldots, y_n$ with $y_i \in \xi_i$ satisfy the ziggurat-equilizer equations. For $i \in T_n^*$, define

$$x_i := y_i \text{ and } x_{2n+1-i} := \mu_i(y_i).$$

Also, define $x_0 := 0$ and $x_{2n+1} := \infty$. Suppose that $\{x_i\}_{0}^{2n+1}$ is a partition of $\mathbb{R}^+$, and let $\delta^*$ be the ziggurat decision rule over $\{x_i\}$ onto $\Theta_n$.

Let $\pi^*$ be the positive probability function on $\Theta_n$ defined by the following conditions: For $i \in T_n^*$,

$$\pi^*(\pm i) = \left(\prod_{k=1}^{i} \rho(k)\right)^{-1} \pi^*(0),$$

where

$$\pi^*(0) = \left[1 + 2 \sum_{i=1}^{n} \left(\prod_{k=1}^{i} \rho(k)\right)^{-1}\right]^{-1}.$$  

Then $\delta^*$ and $\pi^*$ have the following properties:

1. $\delta^*$ is Bayes against $\pi^*$.
2. $\delta^*$ is an equilizer rule.
3. $\delta^*$ is minimax.
4. $\delta^*$ is admissible.
5. $\pi^*$ is least favorable.

**Example 2.13** Refer to example 2.11: The ziggurat decision rule over $\{x_i\}$ onto $\Theta_3$ is an admissible minimax rule.$\square$

**Example 2.14** Refer to examples 2.5 and 2.6: Verify that $y_1 := 0.617$ and $y_2 := 1.912$ satisfy the ziggurat-equilizer equations for $n = 2$, and note that $\{x_i\}$ is a $\mu_i$-constrained constrained partition of $\mathbb{R}^+$. Thus $\delta$ is minimax and $\pi$ is least favorable.$\square$

**Corollary 2** In theorem 1, define

$$\tau := F(-\frac{1}{2})/F(\frac{1}{2}).$$

Then

$$F(-\frac{1}{2}) < R_{\delta^*} \leq 1 - \left(1 + 2\tau \frac{1 - \tau N}{1 - \tau}\right)^{-1}.$$  

**Remark** The upper bound of this corollary is better than the upper bound $2F(-\frac{1}{2})$ of proposition 4:

$$1 - \left(1 + 2\tau \frac{1 - \tau N}{1 - \tau}\right)^{-1} \uparrow 2F(-\frac{1}{2})u \quad \text{as} \quad N \uparrow \infty$$

3 Uncertain Noise Distribution

This section constructs a minimax rule for the location parameter in a robust-estimation problem $(\Theta_1 \times \{\sigma_1, \sigma_2\}, \Theta_1, L_0, Z)$ in which the uncertainty class is $\{N(0, \sigma_1^2), N(0, \sigma_2^2)\}$. The larger scale $\sigma_2$ is large enough that the problem does not reduce to standard-estimation. Examples 3.1 and 3.2 give minimax rules for specific values of the scales. Example 3.3 considers a similar problem in which the scale set has more than two points. The minimax rules of these examples are not monotonic even though the nominal distribution has a monotone likelihood ratio in its location parameter. Examples 3.4 – 3.7 discuss the analysis underlying these results.
Example 3.1 Let \( \sigma_1 := 1 \) and \( \sigma_2 := 2.5 \). Define the decision rule \( \delta^* \) as follows:

\[
\begin{align*}
x_1 &:= 1.09504 \\
x_2 &:= 2.93635 \\
x_3 &:= 3.20822 \\
\end{align*}
\]

\[
\delta^*(z) := \begin{cases} 
0 & \text{if } 0 \leq z < x_1 \\
1 & \text{if } x_1 \leq z < x_2 \\
0 & \text{if } x_2 \leq z < x_3 \\
1 & \text{if } x_3 \leq z \\
\delta^*(-z) & \text{if } z < 0.
\end{cases}
\]

(See figure 3.) This rule is a minimax rule for \((\Theta_1 \times \{\sigma_1, \sigma_2\}, \Theta_1, L_0, Z)\). Let \( \pi^* \) be the following probability function on \( \Theta_1 \times \{\sigma_1, \sigma_2\} \):

\[
\begin{align*}
\pi^*(0, \sigma_1) &:= 0 \\
\pi^*(0, \sigma_2) &:= 0.40587187 \\
\pi^*(\pm1, \sigma_1) &:= 0.04816666 \\
\pi^*(\pm1, \sigma_2) &:= 0.24890241
\end{align*}
\]

Then \( \delta^* \) is a Bayes rule against \( \pi^*_1 \) and \( \pi^* \) is a least-favorable probability function. The rule \( \delta^* \) is almost an equalizer rule over \( \Theta_1 \times \{\sigma_1, \sigma_2\} \):

\[
R((0, \sigma_1), \delta^*) = 0.26453
\]

\[
R((0, \sigma_2), \delta^*) = R((\pm1, \sigma_1), \delta^*) = R((\pm1, \sigma_2), \delta^*) = 0.576597
\]

The risk for the parameter \((0, \sigma_1)\) is less than the equalized risk for the other parameters, and the probability mass for \((0, \sigma_1)\) is zero. □

Example 3.2 Let \( \sigma_1 := 1 \) and \( \sigma_2 := 2 \). The corresponding points \( x_1, x_2, x_3 \) are these:

\[
\begin{align*}
x_1 &:= 1.09504 \\
x_2 &:= 2.93635 \\
x_3 &:= 3.20822
\end{align*}
\]

Define \( \delta^* \) by definition (1). Then \( \delta^* \) is minimax. The corresponding least-favorable probability function \( \pi^* \) is this:

\[
\begin{align*}
\pi^*(0, \sigma_1) &:= 0 \\
\pi^*(0, \sigma_2) &:= 0.43414873 \\
\pi^*(\pm1, \sigma_1) &:= 0.09183446 \\
\pi^*(\pm1, \sigma_2) &:= 0.19109118
\end{align*}
\]

The risk function is this:

\[
R((0, \sigma_1), \delta^*) = 0.271514
\]

\[
R((0, \sigma_2), \delta^*) = R((\pm1, \sigma_1), \delta^*) = R((\pm1, \sigma_2), \delta^*) = 0.550656
\]

In this example, too, the risk for the parameter \((0, \sigma_1)\) is less than the equalized risk for the other parameters, and the probability mass for \((0, \sigma_1)\) is zero. □

Example 3.3 This example extends example 3.2 by allowing the scale set to have more than two points.

Define \( \sigma_0 = 0.9073846 \). Let \( \Sigma \) be a scale set that includes \( \sigma_1, \sigma_2, \) and any finite number of points between \( \sigma_0 \) and \( \sigma_1 \). Then \( \delta^* \) is robust minimax for the decision problem \((\Theta_1 \times \Sigma, \Theta_1, L_0, Z)\). The probability function of example 3.2 is extended as follows: If \( \sigma \neq \sigma_1 \) or \( \sigma \neq \sigma_2 \), then \( \pi^*(\sigma, \sigma) := 0 \) for all \( \sigma \). Here, too, \( \delta^* \) is Bayes against \( \pi^*_1 \) and \( \pi^* \) is least favorable. □

Example 3.4 In the standard-estimation problems of [McKendall, 1990a], the likelihood ratio of the sampling density \( f_Z(\cdot | \sigma) \) is important to Bayes analysis. If \( Z \) has a monotone likelihood ratio, for example, the corresponding Bayes rule is monotonic. Alternatively, if \( Z \) has a Cauchy distribution, the non-monotonic shape of a Bayes rule mimics the non-monotonic shape of a Cauchy likelihood ratio. In this robust-estimation problem, however, it is the likelihood ratio of the marginal density of \( Z \) given \( \theta \) under the least-favorable distribution \( \pi^*_1 \), denoted \( \beta_Z(z | \theta) \), that is important to Bayes analysis:

\[
\beta_Z(z | \theta) := \sum_{\sigma} f_Z(z | \theta, \sigma) \pi(\theta, \sigma), \quad z \in \mathbb{R}
\]

Figure 4 plots a likelihood ratio of \( \beta_Z(z | \theta) \) for the robust-estimation problem of example 3.1. The non-monotonic shape of \( \delta^* \) mimics the shape of this ratio. □
Example 3.5 The probability function $\pi^*$ of example 3.1 or 3.2 satisfies the following linear system of equations:

$$\beta_Z(x_i | 1) = \beta_Z(x_i | 0), \quad i = 1, 2, 3$$

$$\sum_{\theta} \sum_{\sigma} \pi^*(\theta, \sigma) = 1$$

Define $y_0, y_1, y_2, \text{and } y_3$:

$$y_0 := \pi^*(0, \sigma_1)$$
$$y_1 := \pi^*(0, \sigma_2)$$
$$y_2 := \pi^*(1, \sigma_1)$$
$$y_3 := \pi^*(2, \sigma_2)$$

The equations are these ($i = 1, 2, 3$):

$$\frac{1}{\sigma_1} f(x_1 / \sigma_1) y_0 + \frac{1}{\sigma_2} f(x_1 / \sigma_2) y_1$$
$$- \frac{1}{\sigma_1} f(x_1 - 1 / \sigma_1) y_2 - \frac{1}{\sigma_2} f(x_1 - 1 / \sigma_2) y_3 = 0$$

$$y_0 + y_1 + 2y_2 + 2y_3 = 1$$

When $x_1$, $x_2$, and $x_3$ are known, these are four equations in four variables.

These constraints on the probability function are analogous to those of proposition 1. \(\square\)

Example 3.6 The results of examples 3.1, and 3.2 are computed from the following nonlinear system of equations with the assumption that $\pi^*(0, \sigma_1) = 0$ (or $y_0 = 0$):

$$y_1 + 2y_2 + 2y_3 = 1$$

$$\beta_Z(x_i | 1) = \beta_Z(x_i | 0), \quad i = 1, 2, 3$$

$$R(1, \sigma_j, \delta^*) = R(0, \sigma_j, \delta^*), \quad j = 1, 2$$

These are six equations in the six unknowns $x_1$, $x_2$, $x_3$, $y_1$, $y_2$, $y_3$. It must be verified for any solution that $x_1 \leq x_2 \leq x_3$, that $y_1$, $y_2$, and $y_3$ are non-negative, that $\delta^*$ is Bayes against $\pi^*$, and that $R((0, \sigma_1), \delta^*) \leq R((0, \sigma_2), \delta^*)$. \(\square\)

Example 3.7 This example lists the risk function of a decision rule $\delta^*$ of definition (1).

$$R((0, \sigma), \delta^*) = -2F(x_1 / \sigma) + 2F(x_2 / \sigma) + 2F(-x_3 / \sigma)$$

$$R((1, \sigma), \delta^*) = F((x_1 - 1) / \sigma) - F((x_2 - 1) / \sigma) + F((x_3 - 1) / \sigma)$$

$$R((-1, \sigma), \delta^*) = R((1, \sigma), \delta^*)$$

References


