August 1990

Non-Monotonic Decision Rules for Sensor Fusion

Raymond McKendall
University of Pennsylvania

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

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-90-56.

Non-Monotonic Decision Rules for Sensor Fusion

Abstract

This article describes non-monotonic estimators of a location parameter from a noisy measurement *Z = Θ + V* when the possible values of e have the form (0, ± 1, ± 2,. . . , ± *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 *Θ* under a least-favorable prior distribution.

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-90-56.

Non-Monotonic Decison 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 θ from a noisy measurement $Z = \theta + V$ when the possible values of θ 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 θ 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 θ 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* [McKendall, 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.[1] 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 $\mathcal{N}(0, \sigma_1^2)$ 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 δ^* for the location parameter in a standard-estimation problem $(\Theta_n, \Theta_n, L_0, Z)$ in which Z has a Cauchy distribution. A ziggurat decision rule is a non-monotonic step function with range Θ_n. The non-monotonicity of δ^* reflects the non-monotonicity of the likelihood ratio of a Cauchy distribution. The range of δ^* reflects the structure of the zero-one (e) 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 μ and unit scale, written

*Acknowledgement: Navy Contract N0014-88-K-0630; AFOSR Grants 88-0244, 88-0296; Army/DAAL 03-89-C-0031PRI; NSF Grants CISE/CDA 88-22719, IRI 89-06770; and the Dupont Corporation.

[1] A random variable Z with a density function $f_Z(\cdot|\theta)$, for $\theta \in \Theta$, has a *monotone likelihood ratio* if the ratio $f_Z(\cdot|\theta_1)/f_Z(\cdot|\theta_2)$ is non-decreasing for all $\theta_1 > \theta_2$.

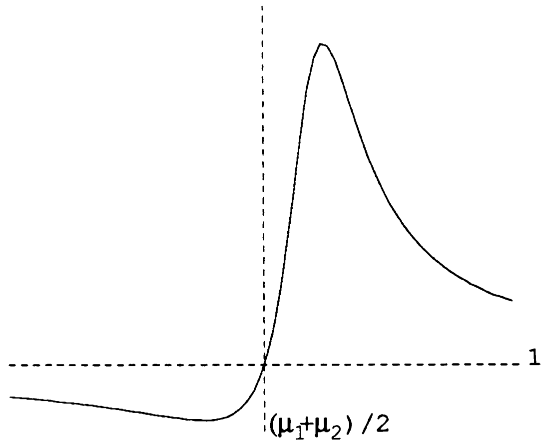

Figure 1: A likelihood ratio $f(\cdot|\mu_1)/f(\cdot|\mu_2)$ of a Cauchy distribution

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

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

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

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

The $\mathcal{C}(0,1)$ distribution is the *standard Cauchy* distribution. An important property of a Cauchy disrtibution 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 δ^* and a least-favorable probability function π^* on Θ_n for the standard-estimation problem $(\Theta_n, \Theta_n, L_0, Z)$ in which $n = 2$ and F is the $\mathcal{C}(0,1)$ distribution. The general results have arbitrary n.

The decision rule δ^*, defined by figure 2, is the *ziggurat decision rule* over a partition $\{x_i\}_0^5$ of $\Re^+$ onto Θ_2: It is an even, non-monotonic step function with range Θ_2 and with steps of unit height occurring at points of $\{x_i\}$. The points x_1 and x_2 are chosen so that δ^* is an equalizer rule. The points x_3 and x_4 and the positive probability function π^* are constructed from x_1 and x_2 so that δ^* is Bayes against π^*. Consequently, the rule δ^* is admissible and minimax, and the probability function π^* 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:

$$\begin{aligned} g_i(x) &:= F(x-i) + F(i-\mu_i(x)), & i = 1,2 \\ h_i(x) &:= F(\mu_{i+1}(x)-i) + F(x-i), & i = 0,1 \end{aligned}$$

The function μ_i is this:

$$\mu_i(x) := \begin{cases} i - \frac{1}{2} & \text{if } x = i - \frac{1}{2} \\ \dfrac{(i-\frac{1}{2})x - (i-\frac{1}{2})^2 + v_1^2}{x - (i-\frac{1}{2})} & \text{if } x \neq i - \frac{1}{2} \end{cases}$$

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

These equations have unique solution y_1, y_2 such that

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

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:

$$\begin{aligned} x_0 &:= 0 \\ x_1 &:= y_1 \\ x_2 &:= y_2 \\ x_3 &:= \mu_2(y_2) \\ x_4 &:= \mu_1(y_1) \\ x_5 &:= \infty \end{aligned}$$

This partition is a *μ_i-constrained partition* of $\Re^+$.

The probability function π^* is this:

$$\begin{aligned} \pi^*(\pm 1) &= \pi^*(0)/\rho(1) \\ \pi^*(\pm 2) &= \pi^*(0)/(\rho(1)\rho(2)) \end{aligned}$$

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

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

The probability function π^* 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 $\Re^+$.

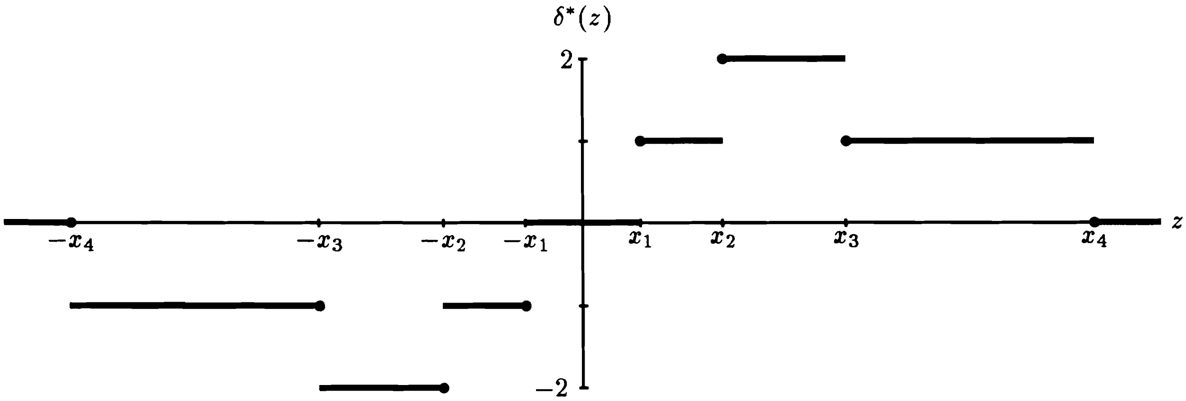

Figure 2: Ziggurat decision rule δ^*

Notation: $\mathcal{I}_p^q$ For integers $p \le q$, 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 $\Re^+$ A partition[2] of $\Re^+$ is a set of points $\{x_i\}_0^{p+1}$ such that $x_0 = 0$, $x_{p+1} = \infty$, and $x_{i+1} > x_i$ for $i \in \mathcal{I}_0^p$. Such a partition is abbreviated as $\{x_i\}$.

Example 2.1 A partition of $\Re^+$ with $p = 4$ is

$$\{x_i\}_0^5 = \{0, 0.617, 1.912, 4.536, 11.209, \infty\}. \square$$

Remark A particular partition of $\Re^+$ is specified by the points x_i, $i \in \mathcal{I}_1^p$. The specification of x_0 and x_{p+1} is implicit.

Definition: ziggurat decision rule Let $\{x_i\}_0^{2n+1}$ be a partition of $\Re^+$. The ziggurat decision rule δ over $\{x_i\}$ onto Θ_n is this:

$$\delta(z) := \begin{cases} i & \text{if} \quad x_i \le z < x_{i+1}, \quad i = 0, \ldots, n \\ n - i & \text{if} \quad x_{n+i} \le z < x_{n+i+1}, \quad i = 1, \ldots, n \\ -\delta(-z) & \text{if} \quad z \le 0 \end{cases}$$

Example 2.2 Let $n = 2$. Define δ:

$$\delta(z) := \begin{cases} 0 & \text{if} \quad 0 \le z < x_1 \\ u & \text{if} \quad x_1 \le z < x_2 \\ 2u & \text{if} \quad x_2 \le z < x_3 \\ u & \text{if} \quad x_3 \le z < x_4 \\ 0 & \text{if} \quad x_4 \le z \\ -\delta(-z) & \text{if} \quad z < 0 \end{cases}$$

Then δ is the ziggurat decision rule over the partition $\{0, x_1, x_2, x_3, x_4, \infty\}$ onto Θ_2. $\square$

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 $\Re^+$: A ziggurat is a terraced pyramid.

[2] This 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 μ_i-*constrained* partitions of $\Re^+$.

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

Definition: μ_i-constrained partition of $\Re^+$ A μ_i-constrained partition of $\Re^+$ is a partition $\{x_i\}_0^{2n+1}$ of $\Re^+$ such that for all $i \in \mathcal{I}_1^n$,

$$x_i \in \xi_i$$

and

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

Example 2.3 A μ_i-constrained partition of $\Re^+$ 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, $x_i \in \xi_i$. $\square$

Example 2.4 Let $n = 2$. Define x_1, x_2, x_3, x_4:

$$x_1 := 0.617,\ x_2 := 1.912,\ x_3 := 4.536,\ x_4 := 11.209.$$

Note that $x_1 \in \xi_1$ and $x_2 \in \xi_2$:

$$\tfrac{1}{2} < x_1 < \tfrac{1}{2} + \tfrac{1}{2}\sqrt{5} = 1.618$$

$$\tfrac{3}{2} < x_2 < \tfrac{3}{2} + \tfrac{1}{2}\sqrt{5} = 2.618$$

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 μ_i-constrained partition of $\Re^+$. $\square$

Remark Let $\{x_i\}_0^{2n+1}$ be a μ_i-constrained partition of $\Re^+$. The ziggurat rule over $\{x_i\}$ steps between $i-1$ and i at x_i and between i and $i-1$ at $\mu_i(x_i)$, $i \in \mathcal{I}_1^n$.

Remark Let $f_Z(\cdot|i) \sim \mathcal{C}(i, 1)$, where i is an integer. The function μ_i satisfies the identity

$$\frac{f_Z(\mu_i(x)|i+e)}{f_Z(\mu_i(x)|i-e-1)} = \frac{f_Z(x|i+e)}{f_Z(x|i-e-1)}, \qquad \forall x \in \Re.$$

This is the functional definition of μ_i. Bayes analysis motivates this definition. The algebraic definition of μ_i is derived from the functional definition.

Main Result

Proposition 1 shows that to any ziggurat decision rule δ over a μ_i-constrained partition of $\Re^+$, there corresponds a positive probability function π on Θ_n such that δ is Bayes against π.

Proposition 1 *Assume $F \sim \mathcal{C}(0,1)$. Let $\{x_i\}_0^{2n+1}$ be a μ_i-constrained partition of $\Re^+$. Let π be the even, positive probability function on Θ_n such that for all $l \in \mathcal{I}_1^m$,*

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

The ziggurat decision rule over $\{x_i\}$ onto Θ_n is Bayes against π.

Example 2.5 Let $n = 2$. Let $\{x_i\}_0^5$ be the μ_i-constrained partition of $\Re^+$ given in example 2.4:

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

Let δ be the ziggurat decision rule over $\{x_i\}$ onto Θ_2:

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

Then δ is Bayes against some positive probability function on Θ_2. □

Example 2.6 Consider example 2.5. The conditions of proposition 1 for a probability function π on Θ_2 are these:

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

$$\rho(1) := \frac{f_Z(x_1|1)}{f_Z(x_1|0)} = \frac{f(0.617-1)}{f(0.617)} = 1.204$$

$$\pi(1) = \rho(2)\,\pi(2)$$

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

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

$$\begin{aligned} \sum_\theta \pi(\theta) &= \pi(0)\left(1 + \tfrac{2}{\rho(1)} + \tfrac{2}{\rho(1)\rho(2)}\right) \\ &= 3.575\pi(0) \end{aligned}$$

Thus π assigns these probabilities:

$$\begin{aligned} \pi(0) &= 0.280 \\ \pi(\pm 1) &= 0.232 \\ \pi(\pm 2) &= 0.128 \end{aligned}$$

Therefore, the ziggurat decision rule over $\{x_i\}_0^5$ onto Θ_2 is Bayes against the probability function π on Θ_2. □

Example 2.7 The probability function π of proposition 1 is given by the following equations: For all $l \in \mathcal{I}_1^n$,

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

where

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

Remark In proposition 1, the restriction to a μ_i-constrained partition of $\Re^+$ 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 μ_i-constrained partition of $\Re^+$.

Proposition 2 *Let $\{x_i\}_0^{2n+1}$ be a μ_i-constrained partition of $\Re^+$, and let δ be the ziggurat decision rule over $\{x_i\}$ onto Θ_n.*

$$\begin{aligned} R(0,\delta) &= 2h_0(x_1) \\ R(\pm i,\delta) &= g_i(x_i) + h_i(x_{i+1}), \qquad i \in \mathcal{I}_1^{n-1} \\ R(\pm n,\delta) &= g_n(x_n) \end{aligned}$$

Example 2.8 Let $n = 3$. Let $\{x_i\}_0^7$ be a μ_i-constrained partition of $\Re^+$, and let δ be the ziggurat decision rule over $\{x_i\}$ onto Θ_3.

$$\begin{aligned} R(0,\delta) &= 2h_0(x_1) \\ R(\pm u,\delta) &= g_1(x_1) + h_1(x_2) \\ R(\pm 2u,\delta) &= g_2(x_2) + h_2(x_3) \\ R(\pm 3u,\delta) &= g_3(x_3) \; \square \end{aligned}$$

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 $n \ge 2$, the ziggurat-equalizer equations are

$$2h_0(y_1) = g_l(y_l) + h_l(y_{l+1}) = g_n(y_n), \qquad l \in \mathcal{I}_1^n.$$

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:

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

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 $F \sim \mathcal{C}(0,1)$. The ziggurat-equalizer equations have unique, increasing solution $y_1, \ldots, y_n$ with $y_l \in \xi_l$. Furthermore $y_l - y_{l-1} > 1$ for $l \in \mathcal{I}_2^n$.*

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

$$\begin{aligned} y_1 &= 0.570743 \\ y_2 &= 1.731856 \\ y_3 &= 2.979961 \end{aligned}$$

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 $y_3 - y_2 > 1$. □

2.7 Equalizer Rule

Proposition 4 gives a ziggurat equalizer rule.

Proposition 4 *Assume $F \sim \mathcal{C}(0,1)$. Let $y_1, \ldots, y_n$ with $y_i \in \xi_i$ satisfy the ziggurat-equalizer equations. For $i \in \mathcal{I}_1^n$, 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 $\Re^+$, then the ziggurat decision rule δ over $\{x_i\}$ onto Θ_n is an equalizer rule. Furthermore, if $\{x_i\}$ is a partition of $\Re^+$, then the common risk of δ is $R_\delta = g_n(x_n)$ and $F(-\frac{1}{2}) < R_\delta < 2F(-\frac{1}{2})$.

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

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

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

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

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

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

Thus, the ziggurat decision rule over $\{x_i\}$ onto Θ_3 is an equalizer. Its risk is $R_\delta = g_3(x_3)$:

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

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-equalizer equations for $n = 2$. Thus, since $\{x_i\}$ is a μ_i-constrained constrained partition of $\Re^+$, the ziggurat rule over $\{x_i\}$ is an equalizer rule. □

Remark Proposition 3 asserts that $x_1, \ldots, x_n$ exist and that $x_i > x_{i-1}$, $i \in \mathcal{I}_2^n$. There is no guarantee, however, that $\{x_i\}_0^{2n+1}$ is a partition of $\Re^+$; it is necessary to verify that $\mu_{i-1}(x_{i-1}) > \mu_i(x_i)$, $i \in \mathcal{I}_2^n$. If $\{x_i\}$ is a partition of $\Re^+$, then it is a μ_i-constrained partition by construction. Numerical computations suggest that $\{x_i\}$ is in fact a partition of $\Re^+$, 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 θ for a decision problem $(\Theta_n, \Theta_n, L_0, Z)$ in which Z has a Cauchy distribution.

Theorem 1 *Assume $F \sim \mathcal{C}(0,1)$. Let $y_1, \ldots, y_n$ with $y_i \in \xi_i$ satisfy the ziggurat-equalizer equations. For $i \in \mathcal{I}_1^n$, 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 $\Re^+$, and let δ^* be the ziggurat decision rule over $\{x_i\}$ onto Θ_n.

Let π^* be the positive probability function on Θ_n defined by the following conditions: For $i \in \mathcal{I}_1^n$,

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

where

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

Then δ^* and π^* have the following properties:

1. *δ^* is Bayes against π^*.*
2. *δ^* is an equalizer rule.*
3. *δ^* is minimax.*
4. *δ^* is admissible.*
5. *π^* is least favorable.*

Example 2.13 Refer to example 2.11: The ziggurat decision rule over $\{x_i\}$ onto Θ_3 is an admissible minimax rule. □

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-equalizer equations for $n = 2$, and note that $\{x_i\}$ is a μ_i-constrained constrained partition of $\Re^+$. Thus δ is minimax and π is least favorable. □

Corollary 2 *In theorem 1, define*

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

Then

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

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

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

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 $\{\mathcal{N}(0, \sigma_1^2),\ \mathcal{N}(0, \sigma_2^2)\}$. The larger scale σ_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.

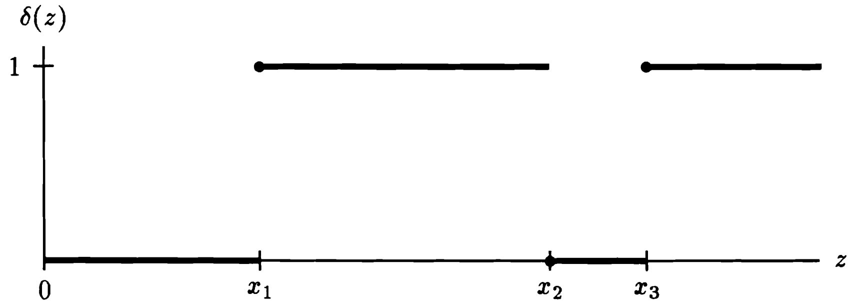

Figure 3: A minimax rule for $(\Theta_1 \times \{\sigma_1, \sigma_2\}, \Theta_1, L_0, Z)$ $(z \geq 0)$

Example 3.1 Let $\sigma_1 := 1$ and $\sigma_2 := 2.5$. Define the decision rule δ^* as follows:

$$\begin{aligned} x_1 &:= 1.09833 \\ x_2 &:= 2.59355 \\ x_3 &:= 3.095 \end{aligned}$$

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

(See figure 3.) This rule is a minimax rule for $(\Theta_1 \times \{\sigma_1, \sigma_2\}, \Theta_1, L_0, Z)$.

Let π^* be the following probability function on $\Theta_1 \times \{\sigma_1, \sigma_2\}$:

$$\begin{aligned} \pi^*(0, \sigma_1) &:= 0 \\ \pi^*(0, \sigma_2) &:= 0.40587187 \\ \pi^*(\pm 1, \sigma_1) &:= 0.048166 \\ \pi^*(\pm 1, \sigma_2) &:= 0.24890241 \end{aligned}$$

Then δ^* is a Bayes rule against π^*; and π^* is a least-favorable probability function.

The rule δ^* is almost an equalizer rule over $\Theta_1 \times \{\sigma_1, \sigma_2\}$:

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

$$R((0, \sigma_2), \delta^*) = R((\pm 1, \sigma_1), \delta^*) = R((\pm 1, \sigma_2), \delta^*) = 0.576597$$

The risk for the parameter $(0, \sigma_1)$ is less than the equalized risk for the other pairs, 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:

$$\begin{aligned} x_1 &:= 1.09504 \\ x_2 &:= 2.93635 \\ x_3 &:= 3.20822 \end{aligned}$$

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

$$\begin{aligned} \pi^*(0, \sigma_1) &:= 0 \\ \pi^*(0, \sigma_2) &:= 0.43414873 \\ \pi^*(\pm 1, \sigma_1) &:= 0.09183446 \\ \pi^*(\pm 1, \sigma_2) &:= 0.19109118 \end{aligned}$$

The risk function is this:

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

$$R((0, \sigma_2), \delta^*) = R((\pm 1, \sigma_1), \delta^*) = R((\pm 1, \sigma_2), \delta^*) = 0.550656$$

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 Σ be a scale set that includes σ_1, σ_2, and any finite number of points between σ_0 and σ_1. Then δ^* 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 $\pi^*(\theta, \sigma) := 0$ for all θ. Here, too, δ^* is Bayes against π^*; and π^* is least favorable. □

Example 3.4 In the standard-estimation problems of [McKendall, 1990a], the likelihood ratio of the sampling density $f_Z(\cdot|\theta)$ 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 θ under the least-favorable distribution π^*, denoted $\beta_Z(\cdot|\theta)$, that is important to Bayes analysis:

$$\beta_Z(z|\theta) := \sum_{\sigma} f_Z(z|(\theta, \sigma))\, \pi(\theta, \sigma), \qquad z \in \Re$$

Figure 4 plots a likelihood ratio of $\beta_Z(\cdot|\theta)$ for the robust-estimation problem of example 3.1. The non-monotonic shape of δ^* mimics the shape of this ratio. □

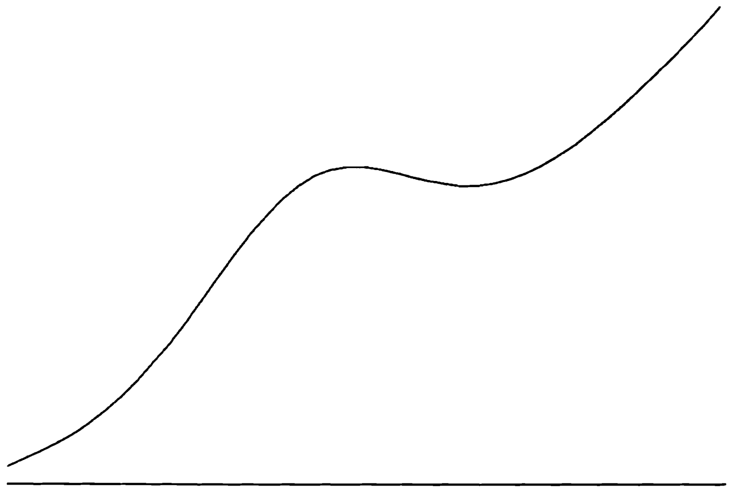

Figure 4: A likelihood ratio of $\beta_Z(\cdot|\theta)$

Example 3.5 The probability function π^* of example 3.1 or 3.2 satisfies the following linear system of equations:

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

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

Define y_0, y_1, y_2, and y_3:

$$\begin{aligned} y_0 &:= \pi^*(0,\sigma_1) \\ y_1 &:= \pi^*(0,\sigma_2) \\ y_2 &:= \pi^*(1,\sigma_1) \\ y_3 &:= \pi^*(2,\sigma_2) \end{aligned}$$

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

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

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

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. □

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 $y_0 = 0$):

$$\begin{array}{ll} y_1 + 2y_2 + 2y_3 = 1 & \\ \beta_Z(x_i|1) = \beta_Z(x_i|0), & i = 1,2,3 \\ R((1,\sigma_j),\delta^*) = R(0,\sigma_2),\delta^*), & j = 1,2 \end{array}$$

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 $x_1 \le x_2 \le x_3$, that y_1, y_2, and y_3 are non-negative, that δ^* is Bayes against π^*_i and that $R((0,\sigma_1),\delta^*) \le R((0,\sigma_2),\delta^*)$. □

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

$$\begin{aligned} R((0,\sigma),\delta^*) &= -2F(x_1/\sigma) + 2F(x_2/\sigma) \\ &\quad + 2F(-x_3/\sigma) \\ R((1,\sigma),\delta^*) &= F((x_1-1)/\sigma) - F((x_2-1)/\sigma) \\ &\quad + F((x_3-1)/\sigma) \\ R((-1,\sigma),\delta^*) &= R((1,\sigma),\delta^*) \end{aligned}$$ □

References

[Martin, 1987] K.E. Martin. *Randomized Robust Confidence Procedures.* PhD thesis, Department of Systems Engineering, University of Pennsylvania, December 1987.

[Martin and Mintz, 1984] K.E. Martin and M. Mintz. Randomized robust confidence procedures. In *Proceedings of the Twenty-Second Annual Allerton Conference on Communication, Control, and Computing*, pages 309–317, University of Illinois, October 1984.

[McKendall, 1990a] R. McKendall. *Minimax Estimation of a Discrete Location Parameter for a Continuous Distribution.* PhD dissertation. GRASP Lab technical report MIS-CS-90-28, Department of Computer and Information Science, University of Pennsylvania, May 1990.

[McKendall, 1990b] R. McKendall. Statistical decision theory for sensor fusion. In these Proceedings, October 1990.

[Zeytinoglu and Mintz, 1984] M. Zeytinoglu and M. Mintz. Optimal fixed sized confidence procedures for a restricted parameter space. *The Annals of Statistics*, 12(3):945–957, September 1984.

[Zeytinoglu and Mintz, 1988] M. Zeytinoglu and M. Mintz. Robust optimal fixed sized confidence procedures for a restricted parameter space. *The Annals of Statistics*, 16(3):1241–1253, September 1988.