**Robust Hypothesis Testing with Band Models for the Probability Densities.**

**Authors:** Saleem A. Kassam

**Performing Organization:** University of Pennsylvania

**Contract or Grant Number:** AFOSR-77-3154

**Abstract:**

Huber's robust probability ratio test for classes of density functions described by contamination and bounded total-variation models are extended for band-models for the densities. A connection between the least favorable (risk) pair of densities and distance-measure based robustness is made explicit.
Abstract

Huber's robust probability ratio tests for classes of density functions described by contamination and bounded total-variation models are extended for band-models for the densities. A connection between the least favorable (risk) pair of densities and distance-measure based robustness is made explicit.

1. Introduction

In [1], Huber developed the ideas of robust tests for binary hypothesis testing problems. Robust tests were obtained as saddle point solutions with respect to performance criteria based on the risk. Huber was mainly interested in classes of probability measures (under the two hypotheses) which contain contaminated nominal probability measures. In [1], classes of probability measures with prescribed maximum total-variation from nominal measures were also considered.

Here we will show that we can extend Huber's analyses for the contaminated nominals classes to classes of density functions in given bands of density functions lying within known upper and lower bounds for each hypothesis. We will also point out an interesting distance-measure robustness property of the least-favorable pairs of probability density functions. This property connects the hypothesis-testing results to recent results on robust Wiener filtering by Kassam and Lim [2] and Poor [3], this recent work having been motivated to a large extent by the usefulness of the theory of robust tests for hypotheses.
It should be noted that in [4,5] Kuznetsov has also considered hypothesis testing with similar band-models for the density functions. The band-model may be more useful in some applications, and does not directly involve the notion of a nominal density.

2. PROBLEM DEFINITION

Let (\Omega,A) be a measurable space and let \( f \) be the density function, with respect to an arbitrary measure \( \mu \) on the space, of a probability measure \( P \). We consider the following classes \( F_0, F_1 \) of allowable density functions under the hypotheses \( H_0, H_1 \) respectively:

\[
F_0 = \{ f \mid f_{OL} \leq f \leq f_{OU}, \int f \mu = 1 \}
\]

\[
F_1 = \{ f \mid f_{1L} \leq f \leq f_{1U}, \int f \mu = 1 \}
\]

Here \( f_{OL}, f_{OU}, f_{1L}, f_{1U} \) are non-negative bounding functions which are such that \( F_0 \) and \( F_1 \) do not overlap, with \( f_{OL}, f_{1L} \) bounded and \( \int f_{OL} \mu, \int f_{1L} \mu < 1 \). The upper bounds are possibly extended-real-valued, and \( \int f_{OU} \mu, \int f_{1U} \mu > 1 \).

Let \( \phi \) be a test for \( F_0 \) vs. \( F_1 \) rejecting \( F_i \) with probability \( \phi_i(x) \) when \( x \in \Omega^n \) is observed, and let \( L_i \) be the loss incurred in falsely rejecting \( H_i \), \( i = 0,1 \). The risk function \( R(f_i, \phi) = L_1 E_{f_i} \{ \phi_i \}, i = 0,1 \), is the criterion of interest in various testing problems [1], and we are thus interested in obtaining a least-favorable pair of densities \( (g_0, g_1) \) in \( F_0 \times F_1 \), so that \( R(f_1, \tilde{\phi}) \leq R(g_0, \tilde{\phi}), i = 0,1 \), for any probability ratio test \( \tilde{\phi} \) for \( g_0 \) vs. \( g_1 \).

Note that the above \( f_i \) denotes an arbitrary element of \( F_i \).

In the following section we obtain the least-favorable pair \( (g_0, g_1) \) using Huber's techniques, and show that the result reduces to Huber's when \( f_{OU}, f_{1U} \) are unbounded everywhere.
3. THE LEAST-FAVORABLE DENSITIES

The introduction of upper bounds $f_{0L}$, $f_{1U}$ in the classes of allowable densities results in a modification in Huber's solution in that the robust test is not always "censored" at constant values. The specific result is the following theorem.

Theorem: The least-favorable pair of densities $(g_0, g_1)$ exists, and is defined by one of either (a), (b), (c) or (d) below, where $k_0$, $k_1$, $k$, $f_0$ and $f_1$ are constants:

(a) $\begin{align*}
g_1 &= \begin{cases}
        k_0 f_{0L}, & f_{1L} < k_0 f_{0L} < f_{1U} \\
        f_{1U}, & f_{1U} < k_0 f_{0L} \\
        f_{1L}, & f_{1L} > k_0 f_{0L}
          \end{cases} \\
g_0 &= \begin{cases}
        \frac{1}{k_1} f_{1L}, & f_{0L} < \frac{1}{k_1} f_{1L} < f_{0U} \\
        f_{0U}, & f_{0U} < \frac{1}{k_1} f_{1L} \\
        f_{0L}, & f_{0L} > \frac{1}{k_1} f_{1L}
          \end{cases}
\end{align*}$

where $0 < k_0 < k_1 < \infty$.

(b) $\begin{align*}
g_1 &= \begin{cases}
        k f_{0L} + kh, & f_{1L} < k f_{0L} < f_{1U} \\
        f_{1U}, & f_{1U} < k f_{0L} \\
        f_{1L} + kh, & f_{0L} < \frac{1}{k} f_{1L} < f_{0U} \\
        f_{1L}, & f_{0U} < \frac{1}{k} f_{1L}
          \end{cases} \\
g_0 &= \begin{cases}
        \frac{1}{k_1} f_{1L}, & f_{0L} < \frac{1}{k_1} f_{1L} < f_{0U} \\
        f_{0U}, & f_{0U} < \frac{1}{k_1} f_{1L} \\
        f_{0L}, & f_{0L} > \frac{1}{k_1} f_{1L}
          \end{cases}
\end{align*}$
\[
\begin{align*}
g_0 &= \begin{cases} 
\varepsilon_{0L} + h, & \varepsilon_{1L} \leq k\varepsilon_{0L} < \varepsilon_{1U} \\
\varepsilon_{0L}, & \varepsilon_{1U} < k\varepsilon_{0L} \\
\frac{1}{k}\varepsilon_{1L} + h, & \varepsilon_{0L} < \frac{1}{k}\varepsilon_{1L} < \varepsilon_{0U} \\
\varepsilon_{0U}, & \varepsilon_{0U} < \frac{1}{k}\varepsilon_{1L} \\
\frac{1}{k}\varepsilon_{1U} - h, & \varepsilon_{0L} < \frac{1}{k}\varepsilon_{1L} < \varepsilon_{0U} \\
\varepsilon_{0U}, & \varepsilon_{0U} < \frac{1}{k}\varepsilon_{1L}
\end{cases} \\
\end{align*}
\]

where

\[
k = \frac{\int\varepsilon_{1L} > k\varepsilon_{0U} \quad \int\varepsilon_{1U} < k\varepsilon_{0L}}{1 - \int\varepsilon_{0U} - \int\varepsilon_{0L}}
\]

and \(h\) is some non-negative function.

(c)

\[
\begin{align*}
g_1 &= \begin{cases} 
\ell_0\varepsilon_{ou}, & \varepsilon_{1U} > \ell_0\varepsilon_{ou} > \varepsilon_{1L} \\
\varepsilon_{1L}, & \varepsilon_{1L} > \ell_0\varepsilon_{ou} \\
\varepsilon_{1U}, & \varepsilon_{1U} < \ell_0\varepsilon_{ou} \\
\frac{1}{\ell_1}\varepsilon_{1U}, & \varepsilon_{0U} > \frac{1}{\ell_1}\varepsilon_{1U} > \varepsilon_{0L} \\
\varepsilon_{0L}, & \varepsilon_{0L} > \frac{1}{\ell_1}\varepsilon_{1L} \\
\varepsilon_{0U}, & \varepsilon_{0U} < \frac{1}{\ell_1}\varepsilon_{1U}
\end{cases} \\
\end{align*}
\]

where \(\ell_0 > \ell_0 \geq \ell_1 > 0, \ k\) as in (1), with equality not possible in both places.
(d) In (c), if \( g_1 \) is not valid with \( \ell_0 \rightarrow \infty \), then

\[
\begin{align*}
g_1 = \begin{cases} 
\text{arbitrary} & , f_{0U} = 0 \\
 f_{1U} & , f_{0U} > 0 
\end{cases}
\end{align*}
\]

In (c), if \( g_0 \) is not valid with \( \ell_1 \rightarrow 0 \), then

\[
\begin{align*}
g_0 = \begin{cases} 
\text{arbitrary} & , f_{1U} = 0 \\
 f_{0U} & , f_{1U} > 0 
\end{cases}
\end{align*}
\]

Proof: The proof of existence of a solution according to either (a), (b), (c) and (d) is outlined in the Appendix. To prove robustness of the probability ratio test for the \((g_0, g_1)\) pair, we have to show that Huber's Lemma 2 [1] holds in our case. We illustrate this for case (a).

As in [1], we have to show that

\[
P_{f_0 \{ \frac{g_1}{g_0} < t \}} \geq P_{f_0 \{ \frac{g_1}{g_0} < t \}} \geq P_{f_1 \{ \frac{g_1}{g_0} < t \}} \geq P_{f_1 \{ \frac{g_1}{g_0} < t \}}
\]

for any real \( t \). Suppose \( k_0 < t \leq k_1 \). Then

\[
P_{f_0 \{ \frac{g_1}{g_0} < t \}} = \int_{f_0 \frac{g_1}{g_0} < t}^{f_0} \frac{f_1}{f_0 L} \leq \int_{f_0}^{f_0 L} \frac{f_1}{f_0 L} \leq \int_{f_0}^{f_0 L} \int_{f_0 L}^{f_0 L} \frac{f_1}{f_0 L} \leq \int_{f_0}^{f_0 L} \frac{f_1}{f_0 L}
\]
Also,

\[ P_{f_1} \left( \frac{g_1}{g_0} \leq t \right) = 1 - \frac{\int_{g_0}^{f_1} f_1}{\int_{0L}^{f_1L} t} \]

\[ \leq 1 - \frac{\int_{g_0}^{f_1L} f_1}{\int_{0L}^{f_1L} t} \]

\[ = 1 - \frac{\int_{g_0}^{f_1L} f_1}{\int_{0L}^{f_1L} t} \]

Similar considerations establish the condition for \( t \leq k_0 \) and \( t > k_1 \). Robustness of the probability ratio test between \( g_0 \) and \( g_1 \) then follows directly from this result [1].

Figure 1 illustrates the form of the least favorable pair \( (g_0, g_1) \) for a simple example where (a) in Theorem 1 is assumed to be valid. Note that we get Huber's result for the \( c \)-contamination classes when the upper bounds tend to infinity.

4. DISTANCE-MEASURES AND ROBUSTNESS

It is reasonable to expect that the pair of least favorable densities we have obtained is the pair of densities "closest" together in some sense. This is indeed the case, and by employing the general formulation of a distance-measure between two hypotheses given by Ali and Silvey [6] we can show that the least favorable densities are the closest pair in \( F_0 \times F_1 \). For probability density
functions, a result by Blackwell [7] may be used to show this. This result was
used by Poor [3] to obtain a connection between robust Wiener filters and robust
tests for hypotheses. We now show more directly that the least favorable pairs
minimize the Ali-Silvey distance measure over $F_0 \times F_1$.

The Ali-Silvey distance between $f_0 \mathbb{C} F_0$ and $f_1 \mathbb{C} F_1$ can be written as

$$d(f_0, f_1) = E_0[C[L_f]]$$

where $L_f = f_1 / f_0$ and $C$ is a convex function. Note that the expectation is a
generalized expectation [6] allowing the measure corresponding to $f_1$ to have a
singular component relative to that for $f_0$.

We consider case (a) in Theorem 1 to illustrate the proof of the statement.

We have

$$d(f_0, f_1) - d(g_0, g_1) = \int f_0 C[L_f] - \int g_0 C[L_g]$$

$$= \int f_0 (C[L_f] - C[L_g]) + \int (f_0 - g_0) C[L_g]$$

$$\geq \int f_0 C'(L_g) (L_f - L_g) + \int (f_0 - g_0) C[L_g]$$

$$= \int (f_0 - g_0) (C[L_g] - L_g C'[L_g]) + \int (f_1 - g_1) C'[L_g]$$

Now $\int (f_0 - g_0) (C[L_g] - L_g C'[L_g])$ is non-negative because $(C[L_g] - L_g C'[L_g])$ is
non-increasing with $L_g$ and because of the definition of $g_0$ in different sets
according to (a), Theorem 1. Specifically,

$$\int (f_0 - g_0) (C[L_g] - L_g C'[L_g]) \geq \int (f_0 - g_0) (C[k_1] - k_1 C'[k_1])$$

$$= 0$$
because whenever $L_1 < k_1$ we have $f_0 - g_0 > 0$ and whenever $L_1 > k_1$ we have $f_0 - g_0 < 0$.

Similarly, $\int (f_1 - g_1) C(L_g) \geq 0$, and we have the result.

Exactly the same conclusion is valid for the least favorable spectral densities for robust Wiener filtering [2,3]. Recently, Poor [8] has obtained similar results for a related set of problems.

5. CONCLUSION

It has been shown that Huber's robust testing results can be extended to classes of probability density functions described by band-models, which do not require an explicit notion of a nominal density function. The results reduce to Huber's result on contaminated-nominal density classes when the upper bound becomes large. We have also shown a direct connection between these results and results on robust Wiener filters through a consideration of distance-measures over the classes of allowable characteristics.

REFERENCES


8. H.V. Poor, "Robust decision design using a distance criterion," (to be published).
APPENDIX

PROOF OF EXISTENCE OF SOLUTION IN THEOREM

Let \( P_0(k_0) = \int g_1 \) and \( P_1(k_1) = \int g_0 \), with \( g_1, g_0 \) defined by (a) in Theorem 1. The \( P_0 \) and \( P_1 \) are continuous functions, \( P_0 \) non-decreasing and \( P_1 \) non-increasing.

We have \( \lim_{k \to 0} P_0(k) < 1 \); let \( \lim_{k \to 0} P_0(k) = a_0 \). Also, \( \lim_{k \to \infty} P_1(k) < 1 \) and let \( \lim_{k \to \infty} P_1(k) = a_1 \).

If \( a_1 > 1 \) and \( a_0 > 1 \), the solutions \( k_0, k_1 \) always exist. If \( k_0 > k_1 \) or if one or both equations \( P_1(k) = 1, P_0(k) = 1 \) do not have a solution, we check existence of (b).

Now we consider the solution of \( P_0(k) + \Delta = P_1(k) + \Delta/k = 1 \). The first equation requires \( P_0(k) - P_1(k) = \Delta(1-k)/k \). This equation has an infinite number of pairs of solutions \((\Delta, k)\), and one of these will also be a solution to \( P_0(k) = 1 - \Delta \). We then check if a solution for \((g_1, g_0)\) for these values \( \Delta, k \) can be obtained, as in (b). Note that \( \Delta, \Delta/k \) have to be less than unity for a solution to be possible. We would get \( \Delta/k = \int h \).

If (b) fails to give a solution, we consider \( Q_0(\ell_0) = \int g_1 \) and \( Q_1(\ell_1) = \int g_0 \) in (c). These are continuous and non-decreasing and non-increasing, respectively. If solutions \( \ell_0 \) and \( \ell_1 \) for \( Q_0(\ell_0) = 1 \) and \( Q_1(\ell_1) = 1 \) exist, than necessarily we cannot have \( \ell_0 = k = \ell_1 \) where \( k \) was the value considered above [Eq. (1)]. If (c) does not work, a solution of the form of (d) will always exist.
Figure Legend

Figure 1. Illustration of a least-favorable pair $(g_0, g_1)$ given by part (a) of Theorem