A NEW MEASURE OF QUANTITATIVE ROBUSTNESS

by

SONIA V.T. MAZZI

Lic., Universidad Nacional de Córdoba, Argentina, 1989

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF STATISTICS

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

December 1991

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of STATISTICS

The University of British Columbia
Vancouver, Canada

Date DECEMBER 13th, 1991

Abstract

The Gross-Error Sensitivity (GES) and the Breakdown Point (BP) are two measures of quantitative robustness which have played a key role in the development of the theory of robustness. Both can be derived from the maximum bias function $B(\epsilon)$ and constitute a two-number summary of this function. The GES is the derivative of $B(\epsilon)$ at the origin whereas the BP determines the asymptote of the curve $(\epsilon, B(\epsilon))$.

Since $GES\epsilon \approx B(\epsilon)$ for ϵ near zero, the GES summarizes the behavior of $B(\epsilon)$ near the origin. On the other hand, the BP does not provide an approximation for $B(\epsilon)$ for ϵ large and, consequently, estimates with strikingly different bias performance when ϵ is large may have the same BP.

A new robustness quantifier, the breakdown rate (BR), that summarizes the behavior of $B(\epsilon)$ for ϵ near BP will be introduced. The BR for several families of robust estimates of regression will be presented and the increased usefulness of the three-number summary (GES,BP,BR) for comparing robust estimates will be illustrated by several examples.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

To quantify the large sample properties of an estimate representable as a functional T, the study of its asymptotic behavior is usually performed on some neighborhood of the model.

We will concentrate on the study of the asymptotic bias of T and consider ϵ-contamination neighborhoods of a central or ideal model F_0. Following this criterion, robust estimates (in their asymptotic version) should change as little as possible, uniformly over some neighborhood of the model. An ϵ-neighborhood of F_0 is a set of distribution functions

$$\mathcal{V}_\epsilon(F_0) = \{F : F = (1-\epsilon)F_0 + \epsilon H;\ H \text{ is a cdf}\}.$$

If $F \in \mathcal{V}_\epsilon(F_0)$ then $F = (1-\epsilon)F_0 + \epsilon H$ for some cdf H which can be interpreted as some unspecified distribution function generating outliers and ϵ can be viewed as the fraction of outliers.

The maximum asymptotic bias of an estimate T over an ϵ-neighborhood, $B_T(\epsilon)$, is an established concept and an important measure of the quantitative and global robustness of T (see Section 2.3.1). $B_T(\epsilon)$ measures the maximum possible perturbation of the value of $T(F)$ when F ranges over $\mathcal{V}_\epsilon(F_0)$.

Naturally, when the amount ϵ of contamination increases so does $B_T(\epsilon)$ and it eventually becomes infinity. The smallest value of ϵ such that the maximum asymptotic bias is infinite is called the *breakdown point* of the estimate and indicates the amount of distortion in the model needed to make the estimate take on arbitrarily large aberrant values. The concept of breakdown point was first introduced by Hodges (1967) for one-dimensional estimates of location. Hampel (1971) gave a much more general definition of an asymptotic nature and Donoho and Huber (1983) introduced a finite sample version of the breakdown point.

Hampel (1968, 1974a) introduced a robustness quantifier called the *influence curve* which measures the speed of change of the value of an estimate when the central model is contaminated with a single observation (see Section 2.3.2). The maximum absolute value of the influence curve is called the *gross-error sensitivity* and this single number summarizes the behavior of the maximum bias curve in a neighborhood of $\epsilon = 0$. In many cases, like in the following example, the gross-error sensitivity is the derivative of the maximum bias curve at the origin.

The concepts of maximum bias curve, gross-error sensitivity and breakdown point are illustrated in Figure 1.1. In this case we consider the one dimensional Gaussian location model and the sample median. It can be shown that the maximum bias of the median is

$$B_m(\epsilon) = \Phi^{-1}\left(1/(2(1-\epsilon))\right)$$

and that its influence curve is

$$IC_m(x) = \mathrm{sgn}(x)/[2\varphi(0)].$$

It easily follows then, that the breakdown point of the sample median is $\epsilon^* = 0.5$ and that the gross error sensitivity is $\gamma^* = 1/[2\varphi(0)] \approx 1.253$ (see for instance Huber, 1981).

The breakdown point and the gross-error-sensitivity are two "one-number-summaries" of the maximum bias curve and they carry important information about this function. These two quantities are now routinely computed and characterize the performance of an estimate.

The breakdown point has proved to be very helpful for understanding the robustness properties of estimates. For example Hampel (1974b,1976) analyzed data from a Monte Carlo study of rejection rules followed by the sample mean, concluding that the performance of the different statistics considered could be ranked in terms of their breakdown points.

As another example, in the Princeton robustness study (Andrews et al, 1972, p.253) two estimates of location with similar asymptotic properties for all symmetric distributions were

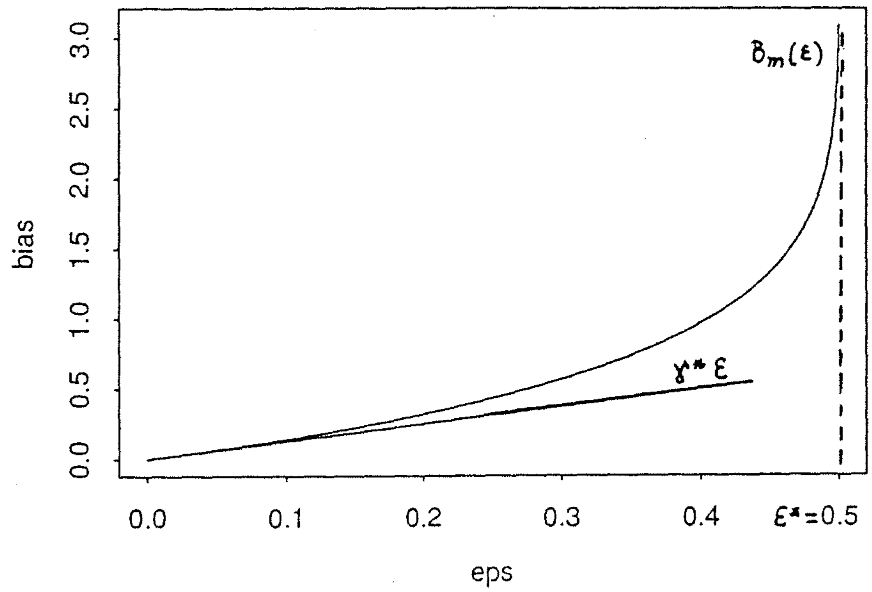

Figure 1.1: Maximum bias curve, BP and GES of the sample median

studied, among others. These location estimates used auxiliary estimates of scale and the difference in their performance was explained in terms of the breakdown points of their corresponding scale estimates.

In the regression setup, the problem of constructing an estimate with non-null breakdown point, i.e. an estimate that can deal with a certain percentage of outliers and that is efficient for a model with Gaussian errors, was a serious concern for many statisticians.

Until 1984, several efforts were made towards obtaining an affine equivariant estimate with maximal breakdown point of 50%.

In 1984, Rousseeuw and Yohai introduced the *S-estimates* , which are defined implicitly by minimizing a robust M-estimate of the scale of the residuals (see Section 3.2). S-estimates can attain a 50% breakdown point, they are affine equivariant and asymptotically normal at the usual rate of $\sqrt{n}$. But these estimates cannot combine the property of high breakdown point with high efficiency at the model with Gaussian errors.

Finally, the *MM-estimates* proposed by Yohai (1987) and the *τ-estimates* proposed by Yohai and Zamar (1988) have the three desired properties: high breakdown point, affine equivariance

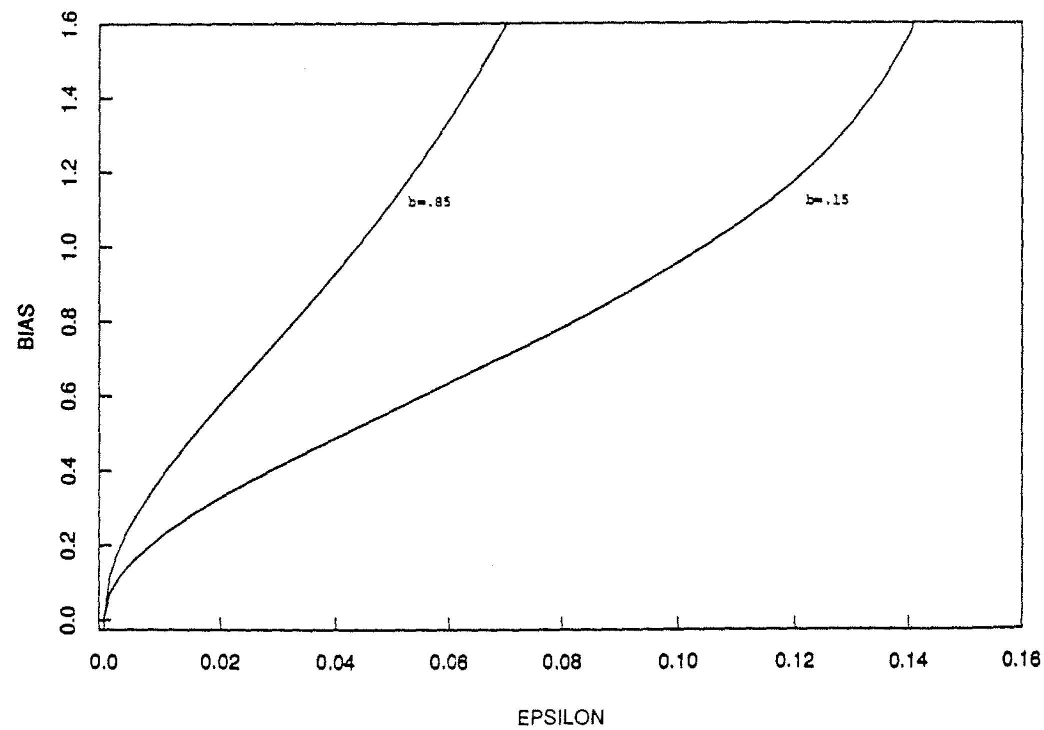

Figure 1.2: Maximum bias curves of S_b for $b = 0.85$ and $b = 0.15$

and high efficiency at the Gaussian model (see Section 3.3 and 3.4).

We see how the concept of breakdown-point (combined with other classical asymptotic concepts) inspired a fruitful search for estimates which are robust in a very precise way and also possess other desirable properties. However, the following example illustrates the fact that robust estimates with the same breakdown point can have strikingly different bias performances for large ϵ.

<u>EXAMPLE</u>: Let b_1, b_2 be such that $b_1 < 0.5$, $b_1 = 1 - b_2$. Consider the S-estimates of regression S_{b_1} and S_{b_2} based on jump functions (see Section 5.1). Since $b_1 = 1 - b_2$, these two estimates have the same breakdown point (see section 3.2). By graphing their maximum bias functions (see Figure 1.2) we notice that $B_{S_{b_2}}$ diverges much more rapidly than $B_{S_{b_1}}$, where $B_{S_{b_i}}$ denotes the maximum bias curve of S_{b_i}. This indicates that S_{b_2} is prone to take on large aberrant values much more rapidly than S_{b_1} and this fact can be formalized by computing the

following limit:

$$RBR(S_{b_1}, S_{b_2}) = \lim_{\epsilon \to BP} \frac{B^2_{S_{b_1}}(\epsilon)}{B^2_{S_{b_2}}(\epsilon)}.$$

An easy calculation shows that $RBR(S_{b_1}, S_{b_2}) = 0$, providing a formal justification of what we inferred from Figure 1.2.

The reason why the breakdown point classification fails to distinguish between rather different estimates, is that the breakdown point indicates only the location of the asymptote of the maximum bias curve but not how the curve actually behaves near this point. That is, the BP does not distinguish among estimates with maximum bias curves tending to infinity at different rates.

Therefore the gross-error sensitivity should be considered a more complete single-number description since it tells us about the behavior of the maximum bias curve in a neighborhood of the origin.

In this thesis we introduce a new measure to quantify robustness in terms of the asymptotic bias, which is fairly easy to compute and to interpret and which, in conjunction with the GES and BP criteria, helps in classifying robust estimates. This quantity is called the *Breakdown Rate* (BR).

The breakdown rate is based on another newly introduced concept called the *Relative Breakdown Rate* (RBR). Given two estimates, say T_1 and T_2 with the same breakdown point, ϵ^*, we compute their relative breakdown rate as the limit of the ratio of the square of their maximum bias curves, $B_1(\epsilon)$ and $B_2(\epsilon)$, as $\epsilon \to \epsilon^*$. If $0 < RBR(T_1, T_2) < \infty$ then for ϵ near ϵ^*,

$$B_1^2(\epsilon) \approx RBR(T_1, T_2)\, B_2^2(\epsilon).$$

If $RBR(T_1, T_2) = 0$ then there is no doubt we would prefer T_1 to T_2 and if $RBR(T_1, T_2) = \infty$ then T_1 would be inadmissible from a robust point of view with respect to T_2.

We work with the specific model of linear regression. The estimates considered are Rousseeuw-Yohai's S-estimates, Yohai's MM-estimates and Yohai-Zamar's τ-estimates.

The breakdown rate of an estimate in the just mentioned families is defined as the relative breakdown rate with respect to a baseline estimate, namely the min-max bias S-estimate among all S-estimates with the same breakdown point.

The breakdown rate together with the breakdown point concept gives a more complete description of the robustness properties of an estimate, because it not only points to the asymptote of the bias curve but also characterizes the way in which the curve goes to infinity. Observe that the gross-error sensitivity and the breakdown rate describe the maximum bias curve near the boundary of its domain, $(0, BP)$.

We will show how the triplet (GES, BP, BR) allows a finer classification of robust estimates.

Chapter 2

Quantitative Robustness

2.1 Estimates Defined by Functionals

Hampel (1968) introduced a way to define an estimate which proved to be quite fruitful since it enabled formalization of a very important aspect of robustness (qualitative robustness). It also made easier the study of the asymptotic properties of estimates, linking theoretical results of functional analysis with those of statistics.

To present Hampel's idea we need the concept of empirical distribution, which gives a way for linking a set of observations $y_1, \ldots, y_n$ to a probability distribution on R^k, $k \geq 1$.

Definition: Given a set $\{y_1, \ldots, y_n\}$ $y_i \in R^k$, the empirical distribution of $y_1, \ldots, y_n$ is the probability measure on R^k, $\mu[y_1, \ldots, y_n]$ defined by

$$\mu[y_1, \ldots, y_n](B) = \frac{1}{n} \sum_{i=1}^{n} I_B(y_i) \quad , \forall B \in \mathcal{B}^k$$

where I_B is the indicator function of the set B and $\mathcal{B}^k$ is the family of Borelian sets in R^k.

Let $\mathcal{Z}(R^k)$ denote the set of all probability measures on R^k. For each n let

$$\mathcal{F}_n = \{\mu[y_1, \ldots, y_n] : y_1, \ldots, y_n \in R^k\}$$

be the set of all empirical distributions associated with samples of size n.

Definition: an estimate T_n is given by a functional, T, defined on $\mathcal{Z}(R^k)$ if there exists a function T defined on a subset $\mathcal{D}(T) \subset \mathcal{Z}(R^k)$ such that:

$$T_n(y_1, \ldots, y_n) = T(\mu[y_1, \ldots, y_n]),$$

where $(y_1, \ldots, y_n)$ is in the domain set of T_n and $\mu[y_1, \ldots, y_n] \in \mathcal{D}(T)$.

We consider estimates which can be defined by functionals or that can be replaced by functionals. This means we assume that there exists a function $T : \mathcal{D}(T) \to R^k$ such that

$$T_n(Y_1, \ldots, Y_n) \xrightarrow{P_F} T(F), \text{ as } n \to \infty$$

when the observations are *i.i.d* according to the true distribution F. We say that $T(F)$ is the asymptotic value of T_n at F.

To illustrate the definitions, an example of how an estimate can be defined by a functional follows.

EXAMPLE: Sample mean defined by a functional.

Let

$$\mathcal{D}(T) = \{F : F \text{ is a cdf on } R \text{ and } \int |x|\, dF(x) < \infty\}$$

and

$$T(F) = \int x\, dF(x) = E_F(X).$$

Then

$$T_n(y_1, \ldots, y_n) = \frac{1}{n}\sum_{i=1}^{n} y_i = T(\mu[y_1, \ldots, y_n]).$$

2.2 ϵ-Neighborhoods

Given a functional T, we are interested in quantifying its robustness with respect to small changes in F. We want to measure the changes in $T(F)$ caused by "small" changes in F in a sense that we will define.

We need the concept of an "ideal" distribution F_0 which obtains because of physical or other reasons and which is completely known. The real data we are able to obtain have a distribution F distorted through gross errors, rounding errors or other factors beyond our control. To make a quantitative assessment of the effects of such distortions we employ a measure of such distortions in the ideal distribution, which can be a measure defined in the space of probability distributions or more generally just a discrepancy in the same space. We will work with the Huber contamination discrepancy defined as:

$$\delta_{Huber}(F; F_0) = \inf\{\zeta : F(x) \geq (1-\zeta)F_0(x), \forall\, x\}$$

Note that this is not a distance.

Let

$$\mathcal{V}_\epsilon(F_0) = \{(1-\epsilon)F_0 + \epsilon H \ : H \text{is a cdf}\};$$

then

$$\mathcal{V}_\epsilon(F_0) = \{F : \delta_{Huber}(F, F_0) \leq \epsilon\}$$

is called the ϵ-contamination neighborhood of F_0.

ϵ-contamination neighborhoods were first introduced by Huber (1964) for the location model and they provide a simple way for modeling data contaminated by outliers.

If $F \in \mathcal{V}_\epsilon$, then $F = (1-\epsilon)F_0 + \epsilon H$ where H can be interpreted as some unspecified distribution function which generates the outliers and ϵ can be viewed as the fraction of contamination.

2.3 Quantitative Robustness

For various reasons it may be useful to describe quantitatively how greatly a small change in the underlying distribution, F, changes the distribution, $d_F(T_n)$, of an estimate $T_n = T_n(x_1, \ldots, x_n)$, $x_i \in R^k$. A description by means of a few numerical quantifiers might be more effective than a detailed characterization.

For the sake of simplicity, we will assume that $k = 1$.

2.3.1 Asymptotic Bias and Asymptotic Variance

Assume that T_n is defined through a functional T, so that $T_n = T(F_n)$. In most cases of interest, T_n is strongly consistent i.e,

$$T_n \overset{a.s.[F]}{\longrightarrow} T(F)$$

and asymptotically normal,

$$\mathcal{L}_F\{\sqrt{n}[T_n - T(F)]\} \xrightarrow{\mathcal{D}} \mathcal{N}(0, A(F,T))$$

as the sample size, n, tends to infinity.

Quantitative large sample robustness is usually discussed in terms of the behavior of the *asymptotic variance* $A(F,T)$ and of the *asymptotic bias*, $T(F) - T(F_0)$, over some neighborhood $\mathcal{V}_\epsilon(F_0)$ of the model distribution (e.g. $\mathcal{V}_\epsilon(F_0)$ can be an ϵ-contamination neighborhood). In this sense, two important quantifiers are the *maximum asymptotic bias*

$$B_T(\epsilon) = \sup_{F \in V_\epsilon(F_0)} |T(F) - T(F_0)|$$

and the *maximum asymptotic variance*

$$V_T(\epsilon) = \sup_{F \in V_\epsilon(F_0)} A(F,T).$$

If we consider ϵ-contamination neighborhoods of F_0, then

$$\mathcal{V}_\epsilon(F_0) = \{F : F = (1-\epsilon)F_0 + \epsilon H \text{ , where } H \text{ is a cdf}\}.$$

Therefore, $\mathcal{V}_1(F_0) = \{H : H \text{ is a cdf}\}$ is the set of all probability measures on the sample space so that $\mathcal{V}_\epsilon(F_0) \subset \mathcal{V}_1(F_0)$, $\forall\, 0 \leq \epsilon \leq 1$ and so $B_T(\epsilon) \leq B_T(1)$. Usually $B_T(1) = \infty$.

The *asymptotic breakdown point* of T at F_0 is

$$BP(T) = \sup\{\epsilon : B_T(\epsilon) < B_T(1)\}.$$

2.3.2 The Influence Function and the Gross-Error-Sensitivity

Hampel (1968,1974a) introduced a robustness quantifier called the influence curve (IC) or influence function, defined as

$$IC(x; F, T) = \lim_{s \to 0} \frac{T((1-s)F + s\delta_x) - T(F)}{s},$$

where δ_x denotes the point mass 1 at x, $x \in R$, when the limit exists.

This quantity can be viewed as the limiting influence on the value of $T(F_n)$ of a single observation x added to the sample of size n.

The maximum absolute value of the influence curve,

$$\gamma^* = \sup_x |IC(x; F, T)|$$

is called the *gross-error-sensitivity.*

In most of the cases, when γ^* and $B'_T(0)$ are finite, it can be seen that $\gamma^* = B'_T(0)$, and so the gross-error sensitivity gives us a linear approximation of the bias curve near 0. Indeed it can be shown that under mild regularity conditions, then equality holds.

Chapter 3

Some Robust Estimates of Regression Coefficients

3.1 The Regression Model

Assume the target model is given by

$$y = \boldsymbol{x}'\theta_0 + u, \tag{3.1}$$

where $\boldsymbol{x} = (x_1, \ldots, x_p)'$ is a random vector in R^p, $\theta_0 = (\theta_{1_0}, \ldots, \theta_{p_0})'$ is the vector of true regression coefficients and the error, u, is a random variable independent of $\boldsymbol{x}$. Let F_0 be the nominal distribution function of u and G_0, the nominal distribution function of $\boldsymbol{x}$. Then the nominal distribution function, H_0, of $(y, \boldsymbol{x})$ is

$$H_0(y, \boldsymbol{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} F_0(y - \theta_0' \boldsymbol{s})\, dG_0(\boldsymbol{s}). \tag{3.2}$$

Assume G_0 is elliptical about the origin with scatter matrix $\boldsymbol{A}$. Correspondingly, we work with a zero intercept, although it can be shown that there is no loss of generality in this assumption.

Let T be an R^p valued functional defined on a ("large") subset of the space of distribution functions, H, on R^{p+1}. This subset is assumed to include all empirical distribution functions, H_n, corresponding to a sample, $(y_1, \boldsymbol{x}_1), \ldots, (y_n, \boldsymbol{x}_n)$, of size n from H. Then, $T_n = T(H_n)$ is an estimate of θ_0.

It is further assumed that T is regression invariant, i.e., if $\tilde{y} = y + \boldsymbol{x}'\boldsymbol{b}$ and $\tilde{\boldsymbol{x}} = \boldsymbol{C}^T\boldsymbol{x}$ for some full rank $p \times p$ matrix, $\boldsymbol{C}$, then $T(\tilde{H}) = \boldsymbol{C}^{-1}[T(H) + \boldsymbol{b}]$, where $\tilde{H}$ is the distribution of $(\tilde{y}, \tilde{\boldsymbol{x}})$. Correspondingly, the transformed model parameter is $\tilde{\theta}_0 = \boldsymbol{C}^{-1}[\theta_0 + \boldsymbol{b}]$.

The asymptotic bias $b^A = b_T^A(H)$ of T at H is defined as

$$b_T^A(H) = (T(H) - \theta_0)' A (T(H) - \theta_0). \tag{3.3}$$

Therefore, we can assume without loss of generality, that G_0 is spherical, i.e., $\boldsymbol{A}$ is the identity matrix, and that $\theta_0 = 0$. Accordingly, the nominal model (3.2) becomes

$$H_0(y, \boldsymbol{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} F_0(y) dG_0(||\boldsymbol{s}||) \tag{3.4}$$

and, correspondingly, the asymptotic bias of T at H is given by the euclidean norm squared of T,

$$b_{T(H)}^I = ||T(H)||^2. \tag{3.5}$$

From now on we will write $b_T(H) = b_T^I(H)$.

If the functional T is continuous at H, then $T(H)$ is the asymptotic value of the estimate when the underlying distribution of the sample is H. It is assumed that T is asymptotically unbiased at the nominal model, H_0, that is

$$T(H_0) = 0.$$

In this paper, we will assume that $(y, \boldsymbol{x}) \sim \mathcal{N}(0, \boldsymbol{I}_{p+1})$, that is H_0 is the $p+1$-dimensional multivariate standard normal distribution.

We will work with the ϵ-contamination neighborhood of the fixed nominal distribution H_0, $\mathcal{V}_\epsilon(H_0) = \{(1-\epsilon)H_0 + \epsilon H^* \,:\, H^* \text{ is any arbitrary distribution on } \mathrm{R}^{p+1}\}$. The maximum asymptotic bias of T over $\mathcal{V}_\epsilon(H_0)$ is defined as

$$B_T(\epsilon) = \sup\{||T(H)|| : H \in \mathcal{V}_\epsilon(H_0)\}. \tag{3.6}$$

Finally the asymptotic breakdown point of T is defined as

$$BP(T) = \inf\{\epsilon : B_T(\epsilon) = \infty\}. \tag{3.7}$$

The estimates of regression coefficients considered in this paper have the characteristic that their influence curves are unbounded and so their gross error sensitivity is infinite. And the derivative of their maximum asymptotic bias function at 0 is infinite but the derivative of the square of their maximum asymptotic bias function at 0 is finite. This fact and the need of a linear approximation of the maximum bias function near the origin leads us to use B_T^2 instead of B_T as a measure of maximum possible departure from the central model. Note that the breakdown point remains unaffected. We define the sensitivity of T as

$$SENS(T) = \frac{d}{d\epsilon} B_T^2(\epsilon) \mid_{\epsilon=0} . \tag{3.8}$$

In this way we can approximate $B_T^2(\epsilon) \approx \epsilon SENS(T)$ for $\epsilon \approx 0$.

Remark. Connected with the computation of the maximum asymptotic bias of the estimates considered in the next section, the following is a key result (Martin, Yohai and Zamar, 1989).

Let χ be a real-valued function on R^1 satisfying the following assumptions:

- symmetric and non-decreasing on $[0, \infty)$, with $\chi(0) = 0$;
- bounded, with $\lim_{x\to\infty} \chi(x) = 1$;
- χ has only a finite number of discontinuities.

Assume now that the target model is H_0 is given by (3.4) and that

- F_0 is absolutely continuous with density f_0 which is symmetric, continuous and strictly decreasing for $u \geq 0$ and
- G_0 is spherical and $P_{G_0}(\boldsymbol{x}'\theta = 0) = 0\,,\ \forall\, \theta \in \mathrm{R}^p$ with $\theta \neq 0$.

Under the last assumption, it is easy to see that the distribution of $\boldsymbol{x}'\theta$ depends only on $||\theta||$. Thus we set

$$h(s, ||\theta||) = E_{H_0} \chi \left(\frac{y - \boldsymbol{x}'\theta}{s} \right) .$$

Martin, Yohai and Zamar (1989) show that under the assumptions stated above on χ, F_0 and G_0, h is continuous, strictly increasing with respect to $||\theta||$ and strictly decreasing in s for $s > 0$.

If $\boldsymbol{z} = (y, \boldsymbol{x}) \sim \mathcal{N}(0, \boldsymbol{I}_{p+1})$, then

$$h(s, \gamma) = g_\chi\left(\frac{(1+\gamma^2)^{1/2}}{s}\right),$$

where

$$g_\chi(t) = E\{\chi(tZ)\} \quad \text{with} \quad Z \sim \mathcal{N}(0, 1).$$

3.2 S-Estimates

S-estimates of regression coefficients were introduced by Rousseeuw and Yohai (1984).

Given $u_1, \ldots, u_n$ the M-estimate of scale of these numbers, s_n, is defined as the solution of

$$\frac{1}{n}\sum_{i=1}^{n}\chi\left(\frac{u_i}{s}\right) = b$$

Where χ is bounded, even and non-decreasing on $[0, \infty)$ and b is usually taken equal to $E\{\chi(Z)\}$ with $Z \sim \mathcal{N}(0, 1)$ (see Huber, 1964). We can assume with no loss of generality that $\chi(\infty) = 1$ and $\chi(0) = 0$.

Let $(y_i, \boldsymbol{x}_i)$ be as in (3.1) and let $u_i(\theta) = y_i - \theta' \boldsymbol{x}_i$, $\theta \in \mathrm{R}^p$. The S-estimate of regression $\hat{\theta}_S$ is defined by the property of minimizing the M-estimate of scale of $\{u_i(\theta)\}_{i=1}^n$, that is

$$\hat{\theta}_S = \arg\min S_n(\theta).$$

The corresponding asymptotic version is

$$\hat{\theta}_S(H) = \arg\min S_H(\theta) \tag{3.9}$$

where $S_H(\theta)$ satisfies the equation

$$E_H \chi\left(\frac{y - \theta' \boldsymbol{x}}{S_H(\theta)}\right) = b.$$

As proved in Martin, Yohai and Zamar (1989), the maximum bias of S-estimates of regression when H_0 is Gaussian is given by

$$B_S^2(\epsilon) = \left[\frac{g^{-1}\left(\frac{b}{1-\epsilon}\right)}{g^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}\right]^2 - 1, \quad \text{with } g(t) = E_\Phi\{\chi(tZ)\} \tag{3.10}$$

where Φ denotes the standard normal distribution function.

This formula can be derived in the following way. Let us consider two situations.

1.*Residual M-Scale When the True Model $\theta = 0$ is Fitted.*

Let

$$H_{(x,y)} = (1-\epsilon)H_0 + \epsilon\delta_{(x,y)} \in \mathcal{V}_\epsilon(H_0).$$

Suppose that y is such that $\chi(y/s(\epsilon)) = 1$ where Δ_0 is the residual scale M-estimate when we fit the true model (i.e. $\theta = 0$) so that

$$(1-\epsilon)E_{H_0}\chi\left(\frac{y}{\Delta_0}\right) + \epsilon = b,$$

or equivalently

$$(1-\epsilon)g\left(\frac{1}{\Delta_0}\right) + \epsilon = b \implies \Delta_0 = \frac{1}{g^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}.$$

2.*Residual M-Scale When the Outlier $(y, \boldsymbol{x})$ is Fitted.*

Let $\Delta(||\theta||)$ be defined by the equation

$$(1-\epsilon)E_{H_0}\chi\left(\frac{y - \theta'\boldsymbol{x}}{\Delta(||\theta||)}\right) = b,$$

that is

$$(1-\epsilon)g\left(\frac{\sqrt{1+||\theta||^2}}{\Delta(||\theta||)}\right) = b \implies \Delta(||\theta||) = \frac{\sqrt{1+||\theta||^2}}{g^{-1}\left(\frac{b}{1-\epsilon}\right)}.$$

The maximum bias $B_S(\epsilon)$ is determined by the condition

$$\Delta(B_S(\epsilon)) = \Delta_0. \tag{3.11}$$

Observe that $S_H(||\theta||) \geq \Delta(||\theta||)$ and $S_H(0) \leq \Delta_0$ for all $\theta \in \mathrm{R}^p$ and all $H \in \mathcal{V}_\epsilon(H_0)$. Therefore, if $||\tilde{\theta}|| > B_S(\epsilon)$ then $S_H(\tilde{\theta}) > S_H(0)$ and $\tilde{\theta} \neq \arg\min S_H(\theta)$. Clearly then $B_S(\epsilon) \leq ||\tilde{\theta}||$.

On the other hand, following along the lines of Martin, Yohai and Zamar (1989) one can prove that given θ^* with $||\theta^*|| < B_S(\epsilon)$, there exists $H \in \mathcal{V}_\epsilon(H_0)$ such that $\theta^* = \arg\min S_F(\theta)$. Hence, $B_S(\epsilon) \leq ||\theta^*||$.

Therefore,

$$B_S(\epsilon) = \sup\{||\theta|| : \Delta(||\theta||) < \Delta_0\}$$

and so by continuity of $\Delta(\cdot)$, $B_S(\epsilon)$ must satisfy the equation $\Delta(B_S(\epsilon)) = \Delta_0$, from which (3.10) directly follows.

BREAKDOWN POINT OF S-ESTIMATES

From (3.7) and (3.10) we see that the breakdown point of an S-estimate S is

$$BP(S) = \min\{b, 1-b\}. \tag{3.12}$$

So, two distinct values of b give rise to any specified breakdown point $\epsilon^* \in (0, 0.5)$, namely, $b = \epsilon^*$ and $b = 1 - \epsilon^*$. It will be shown in chapter 4 that the S-estimates S^b for two such values of b have a strikingly different bias performance.

SENSITIVITY OF S-ESTIMATES

From (3.8) and (3.10), and if $g(t)$ is continuously differentiable in some neighborhood of $t = 1$, the sensitivity of an S-estimate, $SENS(S)$, is given by

$$SENS(S) = \frac{2}{g'(1)}. \tag{3.13}$$

More generally, suppose now that the estimate of regression coefficients, $\hat{\theta}_J(H)$, is given by

$$\hat{\theta}_J(H) = \arg\min_\theta J(F_{H,\theta}) \tag{3.14}$$

where J is a functional defined on a subset of $\mathcal{Z}(R)$ and $F_{H,\theta}$ is the distribution function under H of the residual $r(\theta) = y - \boldsymbol{x}'\boldsymbol{\theta}$. Notice that in the case of S-estimates we take $J(F_{H,\theta}) = S(F_{H,\theta})$, with $S(F_{H,\theta})$ defined by the equation

$$S(F_{H,\theta}) \;:\; E_{F_{H,\theta}}\chi\left(\frac{r}{s}\right) = b. \tag{3.15}$$

Under certain regularity conditions to be determined in future work, we conjecture that following the lines of the argument given above it can be shown that the maximum bias function for $\hat{\theta}_J(H)$ satisfies the equation (3.11) with,

$$\Delta_0 = J(F_{H,0}) \ ; \ F_{H,0}(x) = (1-\epsilon)\Phi(x) + \epsilon\delta_{+\infty}(x)$$

and

$$\Delta(||\theta||) = J(F_{\tilde{H},\theta}) \ ; \ F_{\tilde{H},\theta}(x) = (1-\epsilon)\Phi(x\sqrt{1+||\theta||^2}) + \epsilon\delta_0(x)$$

where $\delta_y(\cdot)$ is a point mass distribution at y.

3.3 τ-Estimates

A τ-estimate is given by (3.14) with $J(F_{H,\theta}) = \tau\,(F_{H,\theta})$, where

$$\tau(F_{H,\theta}) = S^2(F_{H,\theta})E_{F_{H,\theta}}\chi_2\left(\frac{r}{S(F_{H,\theta})}\right)$$

and $S(F_{H,\theta})$ is based on a function χ_1 (see Yohai and Zamar, 1988).

Let $g_i(t) = g_\chi(t)$, $i = 1,2$ and $b = E_\Phi\chi_1(Z)$. Since in this case,

$$\Delta_0 = \tau(F_{H,0}) = \frac{1}{\left[g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)\right]^2}\left\{(1-\epsilon)g_2\left(g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)\right)+\epsilon\right\},$$

and

$$\Delta(||\theta||) = (1+||\theta||^2)(1-\epsilon)\frac{g_2\left(g_1^{-1}\left(\frac{b}{1-\epsilon}\right)\right)}{\left[g_1^{-1}\left(\frac{b}{1-\epsilon}\right)\right]^2}$$

we have that from (3.11)

$$B_\tau^2(\epsilon) = \left[\frac{g_1^{-1}\left(\frac{b}{1-\epsilon}\right)}{g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}\right]^2\left\{\frac{g_2\left(g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)\right)}{g_2\left(g_1^{-1}\left(\frac{b}{1-\epsilon}\right)\right)}+\frac{\epsilon}{1-\epsilon}\frac{1}{g_2\left(g_1^{-1}\left(\frac{b}{1-\epsilon}\right)\right)}\right\}-1. \tag{3.16}$$

BREAKDOWN POINT OF τ-ESTIMATES

According to (3.16) we see that the breakdown point of a τ-estimate of regression, τ, is

$$BP(\tau) = \min\{b, 1-b\}. \tag{3.17}$$

Again, as in the case of S-estimates, two distinct values of b give rise to an specified breakdown point $\epsilon^* \in (0, 0.5)$. In chapter 5 the pronounced difference between these estimates with the same breakdown point will be shown .

SENSITIVITY OF τ-ESTIMATES

If $g_i(t)$ is continuously differentiable in a neighborhood of $t = 1$, $i = 1, 2$, the sensitivity of a τ-estimate, $SENS(\tau)$, is

$$SENS(\tau) = \frac{2}{g_1'(1)} + \frac{1}{b_2}\left(1 - \frac{g_2'(1)}{g_1'(1)}\right) \tag{3.18}$$

where $b_2 = E_\Phi \chi_2(Z)$.

3.4 MM-Estimates

Let

$$s_1 = s_1(H) = \min_\theta S_1(F_{H,\theta})$$

where $S_1(F_{H,\theta})$ is as on (3.15) and is based on a function χ_1. An MM-estimate is defined by (3.14) where the J-functional is in this case $M(F_{H,\theta}, s_1)$, with

$$M(F_{H,\theta}, s_1) = E_{F_{H,\theta}} \chi_2\left(\frac{r}{s_1}\right)$$

while χ_1 and χ_2 satisfy the conditions given in Yohai (1987) including the requirement that $\chi_1(x) \geq \chi_2(x)\ \forall\, x \in$R.

In this case

$$\Delta_0 = M(F_{H,0}, S_1(F_{H,0}))$$

and

$$\Delta(||\theta||) = M(F_{\tilde{H},\theta}, S_1(F_{H,0})).$$

Notice that

$$\sup_{H \in \mathcal{V}_\epsilon(H_0)} = S_1(F_{H,0}).$$

Let $g_i(t) = g_\chi(t)$ $i = 1,2$ and $b = E_\Phi \chi_1(Z)$. It can be easily derived that

$$\begin{aligned} \Delta_0 &= (1-\epsilon) g_2 \left(g_1^{-1} \left(\frac{b-\epsilon}{1-\epsilon} \right) \right) + \epsilon \\ \Delta(||\theta||) &= (1-\epsilon) g_2 \left(\sqrt{1+||\theta||^2} g_1^{-1} \left(\frac{b-\epsilon}{1-\epsilon} \right) \right) \end{aligned}$$

and therefore

$$B_M^2(\epsilon) = \left[\frac{g_2^{-1} \left(g_2 \left(g_1^{-1} \left(\frac{b-\epsilon}{1-\epsilon} \right) \right) + \frac{\epsilon}{1-\epsilon} \right)}{g_1^{-1} \left(\frac{b-\epsilon}{1-\epsilon} \right)} \right]^2 - 1. \tag{3.19}$$

BREAKDOWN POINT OF MM-ESTIMATES

According to (3.7) and (3.19) the breakdown point of an MM-estimate is

$$BP(M) = b. \tag{3.20}$$

SENSITIVITY OF MM-ESTIMATES

If $g_i(t)$ is continuously differentiable in a neighborhood of $t = 1$, $i = 1,2$, the sensitivity of an MM-estimate is

$$SENS(M) = \frac{2}{g_2'(1)}. \tag{3.21}$$

Chapter 4

The Relative Breakdown Rate

Given two estimates T and T' with the same breakdown point b, we define the *relative breakdown rate* of T with respect to T' as:

$$RBR(T,T') = \lim_{\epsilon \to b} \frac{B_T^2(\epsilon)}{B_{T'}^2(\epsilon)}. \tag{4.1}$$

The concept of relative breakdown rate gives a more complete description of two estimates, because it not only points to the asymptote of the bias curves but also characterizes the relative speed of divergence to infinity.

In chapter 5 we will define the Breakdown Rate of certain types of S-, τ- and MM-estimates. It will be the relative breakdown rate with respect to a baseline estimate, namely the min-max bias S-estimate of regression among all S-estimates with the same breakdown point.

We illustrate now how to compute and use the concept of the relative breakdown rate.

4.1 The Relative Breakdown Rate of S-Estimates Based on χ Functions Strictly Convex on a Neighborhood of Zero

Let S^i be an S-estimate of regression based on χ_i such that χ_i is continuous, differentiable in all but a finite number of points with $0 < \int_0^\infty \chi_i'(y) y\, dy < \infty$ and three times differentiable in some neighborhood of zero with $\chi_i''(0) \neq 0$, $i = 1, 2$. Also suppose that

$$E_\Phi \chi_1(Z) = E_\Phi \chi_2(Z) = b \ , \ 0 < b \leq 0.5.$$

According to (3.10) and (4.1), the relative breakdown rate of S^1 with respect to S^2 is

$$RBR(S^1, S^2) = \lim_{\epsilon \to b} \left[\frac{g_1^{-1}\left(\frac{b}{1-\epsilon}\right) g_2^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}{g_2^{-1}\left(\frac{b}{1-\epsilon}\right) g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)} \right]^2 \tag{4.2}$$

where $g_i = g_{\chi_i}$, $i = 1, 2$.

Note that as $\epsilon \to b$, $\frac{b-\epsilon}{1-\epsilon} \to 0$, $\frac{b}{1-\epsilon} \to \frac{b}{1-b}$ if $b < 0.5$ and $\frac{b}{1-\epsilon} \to 1$ if $b = 0.5$.

Therefore, as $\epsilon \to b$, $g_i^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right) \to 0$, $g_i^{-1}\left(\frac{b}{1-\epsilon}\right) \to g_i^{-1}\left(\frac{b}{1-b}\right)$ if $b < 0.5$ and $g_i^{-1}\left(\frac{b}{1-\epsilon}\right) \to \infty$ if $b = 0.5$, $i = 1, 2$.

We will compute $L_1 = \lim_{t\to 0} \frac{g_2^{-1}(t)}{g_1^{-1}(t)}$ and $L_2 = \lim_{t \to 1} \frac{g_1^{-1}(t)}{g_2^{-1}(t)}$ using L'Hôspital's rule.

Computation of L_1. It is easier to compute

$$L_1^2 = \lim_{t \to 0} \frac{\left[g_2^{-1}(t)\right]^2}{\left[g_1^{-1}(t)\right]^2}.$$

Let

$$L_1^* = \lim_{t\to 0} \frac{\frac{d}{dt}\left[g_2^{-1}(t)\right]^2}{\frac{d}{dt}\left[g_1^{-1}(t)\right]^2} = \lim_{t \to 0} \frac{g_1'(g_1^{-1}(t))}{g_1^{-1}(t)} \frac{g_2^{-1}(t)}{g_2'(g_2^{-1}(t))}.$$

Then, if L_1^* exists (contemplating also the possibility of L_1^* being infinity), $L_1^2 = L_1^*$.

Now, for $i = 1, 2$

$$g_i(t) = 2 \int_0^\infty \chi_i(ty) \varphi(y) \, dy$$

and

$$g_i'(t) = 2 \int_0^\infty \chi_i'(ty) y \varphi(y) \, dy,$$

where $\varphi = \Phi'$.

Let $y > 0$; then a Taylor's series expansion of order 1 around 0 of χ_i' gives, $\chi_i'(ty) = \chi_i''(0)ty + o(ty)$ as $t \to 0$, so that

$$\frac{1}{t} g_i'(t) = \chi_i''(0) + 2 \int_0^\infty \frac{o(ty)}{t} y \varphi(y) \, dy \ , \ t > 0.$$

Hence,

$$L_1^* = \lim_{t \to 0} \frac{g_1'(g_1^{-1}(t))}{g_1^{-1}(t)} \frac{g_2^{-1}(t)}{g_2'(g_2^{-1}(t))} \tag{4.3}$$

$$
(4.4) \qquad = \frac{\chi_1''(0) + 2\int_0^\infty \frac{o(g_1^{-1}(t)y)}{g_1^{-1}(t)} y\varphi(y)\,dy}{\chi_2''(0) + 2\int_0^\infty \frac{o(g_2^{-1}(t)y)}{g_2^{-1}(t)} y\varphi(y)\,dy}
$$

$$
(4.5) \qquad = \frac{\chi_1''(0)}{\chi_2''(0)}.
$$

Computation of L_2. We have that

$$
L_2 = \lim_{t\to 1} \frac{g_1^{-1}(t)}{g_2^{-1}(t)} = \lim_{t\to 1} \frac{\frac{1}{g_2^{-1}(t)}}{\frac{1}{g_1^{-1}(t)}}.
$$

Let

$$
L_2^* = \lim_{t\to 1} \frac{\frac{d}{dt}\frac{1}{g_2^{-1}(t)}}{\frac{d}{dt}\frac{1}{g_1^{-1}(t)}} = \lim_{t\to 1} \frac{[g_1^{-1}(t)]^2 g_1'(g_1^{-1}(t))}{[g_2^{-1}(t)]^2 g_2'(g_2^{-1}(t))}.
$$

Then, if L_2^* exists, $L_2 = L_2^*$.

Now, for $i = 1, 2$ and $t > 0$

$$
t^2 g_i'(t) = 2\int_0^\infty \chi_i'(y) y \varphi\left(\frac{y}{t}\right)\,dy.
$$

For each $y > 0$, a Taylor's series expansion of order 1 around 0 of φ gives around 0 we can write

$$
\varphi\left(\frac{y}{t}\right) = \varphi(0) + o(\tfrac{y}{t}) \quad \text{, as } t \to \infty
$$

and

$$
t^2 g_i'(t) = 2\varphi(0)\int_0^\infty \chi_i'(y) y\,dy + 2\int_0^\infty \chi_i'(y) y o\left(\frac{y}{t}\right)\,dy
$$

so that

$$
(4.6) \qquad L_2^* = \frac{\int_0^\infty \chi_1'(y) y\,dy}{\int_0^\infty \chi_2'(y) y\,dy}.
$$

Then, by (4.2),(4.5) and (4.6), if $0 < b < 0.5$

$$
RBR(S^1, S^2) = \left[\frac{g_1^{-1}\left(\frac{b}{1-b}\right)}{g_2^{-1}\left(\frac{b}{1-b}\right)}\right]^2 \frac{\chi_1''(0)}{\chi_2''(0)};
$$

and if $b = 0.5$

$$
(4.7) \qquad RBR(S^1, S^2) = \left[\frac{\int_0^\infty \chi_1'(y) y\,dy}{\int_0^\infty \chi_2'(y) y\,dy}\right]^2 \frac{\chi_1''(0)}{\chi_2''(0)}.
$$

EXAMPLE 1: We will compute the RBR of two commonly used "smooth" initial S-estimates of regression with breakdown point equal to 0.5. The first one, S^A, is based on the function

$$\chi_A(x) = \begin{cases} \frac{1}{A^2}x^2, & \text{if } 0 \le |x| \le A \\ 1, & \text{if } |x| > A \end{cases}, \quad A > 0 \tag{4.8}$$

which is a simple truncation of the classical square loss function. The choice $A = 1.041$ gives $BP(S^A) = \int_{-\infty}^{\infty} \chi_A(x)\, d\Phi(x) = 0.5$.

The second S-estimate, S^B is based on the integrated Tukey's bisquare score function

$$\chi_B(x) = \begin{cases} \frac{3}{B^2}x^2 - \frac{3}{B^4}x^4 + \frac{1}{B^6}x^6, & \text{if } 0 \le |x| \le B \\ 1, & \text{if } |x| > B \end{cases}, \quad B > 0 \tag{4.9}$$

which is three times continuously differentiable. The choice $B = 1.547$ gives $BP(S^B) = \int_{-\infty}^{\infty} \chi_B(x)\, d\Phi(x) = 0.5$.

We have that

$$\int_0^\infty \chi_A'(y) y\, dy = \frac{2}{3}A \quad ; \quad \int_0^\infty \chi_B'(y) y\, dy = \frac{16}{35}B \tag{4.10}$$

and

$$\chi_A''(0) = \frac{2}{A^2} \quad ; \quad \chi_B''(0) = \frac{6}{B^2} \tag{4.11}$$

so that by (4.7)

$$RBR(S^A, S^B) = 0.709.$$

Also,

$$g_A'(1) = \frac{4}{A^2}[\Phi(A) - A\varphi(A) - 0.5] \tag{4.12}$$

and

$$g_B'(1) = 12\varphi(B)\left(\frac{2}{B} - \frac{5}{B^3}\right) + 12[\Phi(B) - B\varphi(B) - 0.5]\left(\frac{1}{B^2} - \frac{6}{B^4} + \frac{15}{B^6}\right). \tag{4.13}$$

For $A = 1.041$ and $B = 1.547$ we get

$$g_A'(1) = 0.404 \quad ; \quad g_B'(1) = 0.389$$

and so from (3.13)

$$SENS(S^A) = 4.950 \quad ; \quad SENS(S^B) = 5.141.$$

<table>
<tr><th>estimate</th><th>A,B</th><th>$SENS$</th><th>BP</th><th>efficiency</th><th>RBR</th></tr>
<tr><td>S^A</td><td>1.041</td><td>4.950</td><td>0.5</td><td>0.219</td><td rowspan="2">0.709</td></tr>
<tr><td>S^B</td><td>1.547</td><td>5.141</td><td>0.5</td><td>0.287</td></tr>
</table>

Table 4.1: Comparison of two S-estimates with the same BP

The asymptotic efficiency at the model with Gaussian errors of S^A is given by

$$e_A = 2[\Phi(A) - A\varphi(A) - 0.5]$$

and for $A = 1.041$ we have that $e_A = 0.219$. The efficiency of S^B for $B = 1.547$ is $e_B = 0.287$ (see Rousseeuw and Yohai, 1984).

All these computations are summarized in Table 4.1.

Based on these figures, the S-estimate based on χ_A can be expected to perform approximately the same as that based on χ_B for Gaussian or approximately Gaussian data and can be expected to perform better in the presence of a large fraction of outliers. However, this should be confirmed by extensive Monte Carlo simulation.

4.2 The Relative Breakdown Rate of MM- and S-Estimates Based on χ Functions Strictly Convex on a Neighborhood of Zero

Let M be an MM-estimate of regression based on χ_1 and χ_2 with χ_1 and χ_2 three times continuously differentiable, $\chi_i''(0) \neq 0$ for $i = 1,2$ and $E_\Phi \chi_1(Z) = b$, $0 < b \leq 0.5$.

Further, let S be an S-estimate based χ_1. Then $BP(S) = BP(M)$ and we want to compute the relative breakdown rate of M with respect to S. By (4.1) and (3.19) we have that

$$\begin{aligned} RBR(M,S) &= \lim_{\epsilon \to b} \frac{B_M^2(\epsilon)}{B_S^2(\epsilon)} \\ &= \lim_{\epsilon \to b} \left[\frac{g_2^{-1}\left(g_2\left(g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right)}{g_1^{-1}\left(\frac{b}{1-\epsilon}\right)} \right]^2 . \end{aligned}$$

To compute the limit we can use L'Hôspital's rule. Following a similar reasoning as in Section 4.1 we see that

$$RBR(M,S) = \begin{cases} \left[\frac{g_2^{-1}\left(\frac{b}{1-b}\right)}{g_1^{-1}\left(\frac{b}{1-b}\right)}\right]^2, & \text{if } 0 < b < 0.5 \\ \left[\frac{\int_0^\infty \chi_2'(y) y\, dy}{\int_0^\infty \chi_1'(y) y\, dy}\left(2 - \frac{\chi_2''(0)}{\chi_1''(0)}\right)^{-1}\right]^2 & \text{if } b = 0.5 \end{cases}$$

4.3 The Relative Breakdown Rate of τ- and S-Estimates Based on χ Functions Strictly Convex on a Neighborhood of Zero

Let τ be a τ-estimate of regression based on χ_1 and χ_2 such that $E_\Phi \chi_1(Z) = b$, $0 < b \le 0.5$. Further, let S be the S-estimate of regression based on χ_1. Then,

$$\begin{aligned} RBR(\tau,S) &= \lim_{\epsilon \to b} \frac{g_2\left(g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)\right)}{g_2\left(g_1^{-1}\left(\frac{b}{1-\epsilon}\right)\right)} + \frac{\epsilon}{1-\epsilon} \frac{1}{g_1^{-1}\left(\frac{b}{1-\epsilon}\right)} \\ &= \begin{cases} \frac{b}{1-b} \frac{1}{g_2\left(g_1^{-1}\left(\frac{b}{1-b}\right)\right)}, & \text{if } 0 < b < 0.5 \\ 1, & \text{if } b = 0.5. \end{cases} \end{aligned}$$

Note that given χ_1, if τ and S are based on χ_1 and $b = 0.5$, then $RBR(\tau, S) = 1$, no matter how we choose χ_2.

<u>EXAMPLE 2</u>: Let M be an MM-estimate based on χ_1 and χ_2 such that χ_1 and χ_2 satisfy the assumptions made in example 3. Now let τ be based on χ_1 and some other function χ_3. Then, noting that $RBR(M,\tau) = \frac{RBR(M,S)}{RBR(\tau,S)}$ where S is the S-estimate based on χ_1 we have that

$$RBR(M,\tau) = \begin{cases} \left[\frac{g_2^{-1}\left(\frac{b}{1-b}\right)}{g_1^{-1}\left(\frac{b}{1-b}\right)}\right]^2 \frac{1-b}{b} g_2\left(g_1^{-1}\left(\frac{b}{1-b}\right)\right) & \text{if } 0 < b < 0.5 \\ \left[\frac{\int_0^\infty \chi_2'(y) y\, dy}{\int_0^\infty \chi_1'(y) y\, dy}\left(2 - \frac{\chi_2''(0)}{\chi_1''(0)}\right)^{-1}\right]^2 & \text{if } b = 0.5 \end{cases} \tag{4.14}$$

Let $\mathcal{F}$ be the family of functions χ_B, $B > 0$ where $\chi_B(x)$ is given by (4.9). $\mathcal{F}$ is known as Tukey's family of χ functions.

If we take $\chi_i = \chi_{B_i}$ $i = 1,2,3$ such that $E_\Phi \chi_1(Z) = 0.5$, by (4.14), (4.10) and (4.11) we have that

$$RBR(M,\tau) = \left[\frac{B_2}{B_1}\left(2 - \frac{B_1^2}{B_2^2}\right)^{-1}\right]^2. \tag{4.15}$$

estimate	B_1	B_2	BP	efficiency	$SENS$	RBR
M	1.56	4.68	0.5	0.95	9.639	2.58
τ	1.56	6.08	0.5	0.95	13.480	

Table 4.2: Comparison of an MM- and a τ-estimate with the same BP and efficiency

estimate	B_1	B_2	BP	$SENS$	RBR
M	1.56	4.680	0.5	9.639	2.58
τ	1.56	5.025	0.5	9.639	

Table 4.3: Comparison of an MM- and a τ-estimate with the same BP and $SENS$

The choice of $B_1 = 1.56$ gives us two estimates with $BP = 0.5$ and if we choose $B_2 = 4.68$ and $B_3 = 6.08$, both estimates have 95% asymptotic efficiency at the model with Gaussian errors (see Yohai, 1985 and Yohai and Zamar, 1988).

With these values of B_1 and B_2 (note that $BRB(M, \tau)$ doesn't depend on B_3)

$$RBR(M, \tau) = 2.58.$$

By (4.13) $g_1'(1) = 1.300$, $g_2'(1) = 0.207$ and $g_3'(1) = 0.138$ and the value of b_3 such that $b_3 = E_\Phi \chi_3(Z)$ is $b_3 = 0.075$.

Therefore,

$$SENS(\tau) = 13.480 \quad ; \quad SENS(M) = 9.639.$$

We summarize the calculated quantifiers in Table 4.2.

Since (4.15) does not depend on the choice of B_3, if we take $B_1 = 1.56$, $B_2 = 4.68$ and $B_3 = 5.025$, we have that $SENS(M) = SENS(\tau) = 9.639$ and RBR remains the same as the one calculated before, i.e. $RBR(M, \tau) = 2.58$.

We can conclude from Table 4.3 that τ-estimates can be expected to outperform comparable MM-estimates for a wide range of fractions of contamination. This should also be confirmed by extensive Monte Carlo studies.

Chapter 5

The Breakdown Rate

In this chapter we will define the breakdown rate for S-, τ- and MM-estimates of regression. The min-max asymptotic bias (among all S-estimates with the same breakdown point) S-estimate will be used as a baseline estimate. In Section 5.1 we justify the choice of this baseline estimate, in Section 5.2 we give the definition of the breakdown rate and in the subsequent sections we compute the breakdown rate for certain types of S-, τ- and MM-estimates.

5.1 The Baseline Estimate

We will denote by χ_a the function

$$\chi_a(x) = \begin{cases} 1 & \text{if } |x| \leq a \\ 0 & \text{otherwise.} \end{cases}$$

We call χ_a a "jump function" with jump constant a.

Let C be the family of functions $\chi : R \to R$ such that:

- χ is even and nondecreasing in $[0, \infty)$;
- χ is either continuous or a jump function;
- χ is continuously differentiable in all but a finite number of points;
- $\chi(0) = 0$ and $\chi(x) \to 1$ as $x \to \infty$;
- $0 < \mathrm{E}_\Phi\{\chi(X)\} < 1$.

For $\chi \in \mathcal{C}$, let

$$g_\chi(t) = \mathrm{E}_\Phi\{\chi(tX)\}.$$

The following lemma was stated and proved by Martin and Zamar (1989). [1]

Lemma 1 : *Given* $0 < b < 1$, *let*

$$\mathcal{C}_b = \{\chi : \chi \in \mathcal{C} \text{ and } E_\Phi\{\chi(X)\} = b\} \tag{5.1}$$

and a *satisfying* $2[1 - \Phi(a)] = b$.

Then, for all $\chi \in \mathcal{C}_b$

$$\begin{aligned} g_{\chi_a}(t) &\geq g_\chi(t)\,,\ \forall\, t \geq 1; \\ g_{\chi_a}(t) &\leq g_\chi(t)\,,\ \forall\, t < 1. \end{aligned}$$

Proof: Since $\chi, \chi_a \in \mathcal{C}_b$ we have that

$$\int_0^a \chi(y)\varphi(y)\,dy = \int_a^\infty [1 - \chi(y)]\,\varphi(y)\,dy.$$

Now, note that $\varphi(y/t)/\varphi(y)$ is an increasing function of y if $t \geq 1$ and it is decreasing in y if $0 \leq t \leq 1$.

Then, $\forall\, t \geq 1$

$$\begin{aligned} \frac{1}{t}\int_0^a \chi(y)\varphi\left(\frac{y}{t}\right)dy &= \frac{1}{t}\int_0^a \chi(y)\varphi(y)\frac{\varphi\left(\frac{y}{t}\right)}{\varphi(y)}\,dy \\ &\leq \frac{1}{t}\frac{\varphi\left(\frac{a}{t}\right)}{\varphi(a)}\int_0^a \chi(y)\varphi(y)\,dy \\ &= \frac{1}{t}\frac{\varphi\left(\frac{a}{t}\right)}{\varphi(a)}\int_a^\infty [1 - \chi(y)]\,\varphi(y)\,dy \\ &\leq \frac{1}{t}\int_a^\infty [1 - \chi(y)]\,\varphi\left(\frac{y}{t}\right)dy \end{aligned}$$

[1] The result proved in the reference paper is more general than the one presented in Lemma 1. It is valid for any distribution function F_0 with a density f_0 symmetric about 0 and such that $f(tx)/f(x)$ is decreasing in x for $t > 1$.

Therefore,

$$
\begin{aligned}
g_\chi(t) &= \frac{1}{t}\int_0^\infty \chi(y)\varphi\left(\frac{y}{t}\right)\,dy \\
&= \frac{1}{t}\int_0^a \chi(y)\varphi\left(\frac{y}{t}\right)\,dy + \frac{1}{t}\int_a^\infty \chi(y)\varphi\left(\frac{y}{t}\right)\,dy \\
&\leq \frac{1}{t}\int_a^\infty \varphi\left(\frac{y}{t}\right)\,dy \\
&= g_{\chi_a}(t)
\end{aligned}
$$

For $t < 1$ the inequalities above are reversed and the result follows.□

The following theorem, which follows directly from Lemma 1, shows that the S-estimate of regression S_b based on χ_a with $2[1-\Phi(a)] = b$ is min-max bias over $\mathcal{C}_b$, where $\mathcal{C}_b$ is as in (5.1).

Theorem 1 : *For all* $0 < b < 1$,

$$B_S(\epsilon) \geq B_{S_b}(\epsilon) \ \ , 0 \leq \epsilon < b$$

for all S *based on* $\chi \in \mathcal{C}_b$.

Proof: From (3.10), the maximum bias function of an S-estimate based on a function $\chi \in \mathcal{C}_b$ is given by

$$B_S^2(\epsilon) = \left[\frac{g_\chi^{-1}\left(b/(1-\epsilon)\right)}{g_\chi^{-1}\left((b-\epsilon)/(1-\epsilon)\right)}\right]^2 - 1 \ \ , 0 < \epsilon < b.$$

Since $\forall\, \epsilon > 0$, $b/(1-\epsilon) > b$ and $(b-\epsilon)/(1-\epsilon) < b$, it follows that

$$g_\chi^{-1}\left(b/(1-\epsilon)\right) > 1 \ \text{ and } \ g_\chi^{-1}\left((b-\epsilon)/(1-\epsilon)\right) < 1.$$

By the preceding lemma, we have that

$$g_{\chi_a}^{-1}\left(b/(1-\epsilon)\right) \leq g_\chi^{-1}\left(b/(1-\epsilon)\right) \ \text{ and } \ g_{\chi_a}^{-1}\left((b-\epsilon)/(1-\epsilon)\right) \geq g_\chi^{-1}\left((b-\epsilon)/(1-\epsilon)\right).\square$$

Proposition 1 : *If* S *is an S-estimate of regression based on* $\chi \in \mathcal{C}_b$, $0 < b < 1$, *then*

$$\lim_{\epsilon\to b} \frac{B_S^2(\epsilon)}{B_{S_b}^2(\epsilon)} \geq 1.$$

Proof: The result follows from last theorem, since $B_S^2(\epsilon) \geq B_{S_b}^2(\epsilon)$, $0 \leq \epsilon < b$.$\square$

Proposition 2 : *Let* $0 < b < 1$ *and* τ *be a* τ*-estimate of regression based on* $\chi_1 \in C_b$ *and* $\chi_2 \in C$. *Then, if either*

- $b = 0.5$,

  or

- $0 < \min\{b, 1-b\} < 0.5$ *and* $g_2\left(g_1^{-1}\left(\frac{b}{1-b}\right)\right) \leq \frac{b}{1-b}$, [2]

$$\lim_{\epsilon \to b} \frac{B_\tau^2(\epsilon)}{B_{S_b}^2(\epsilon)} \geq 1.$$

Proof: Denote by S^1 the S-estimate of regression based on the function χ_1 and let $g_i = g_{\chi_i}$, $i = 1, 2$. Then,

$$\lim_{\epsilon \to b} \frac{B_\tau^2(\epsilon)}{B_{S_b}^2(\epsilon)} = \lim_{\epsilon \to b} \frac{B_{S^1}^2(\epsilon)}{B_{S_b}^2(\epsilon)} \left\{ \frac{g_2\left(g_1^{-1}(\frac{b-\epsilon}{1-\epsilon})\right)}{g_2\left(g_1^{-1}(\frac{b}{1-\epsilon})\right)} + \frac{\epsilon}{1-\epsilon} \, \frac{1}{g_2\left(g_1^{-1}(\frac{b}{1-\epsilon})\right)} \right\}.$$

Now,

$$\lim_{\epsilon \to b} \frac{g_2\left(g_1^{-1}(\frac{b-\epsilon}{1-\epsilon})\right)}{g_2\left(g_1^{-1}(\frac{b}{1-\epsilon})\right)} + \frac{\epsilon}{1-\epsilon} \, \frac{1}{g_2\left(g_1^{-1}(\frac{b}{1-\epsilon})\right)} = \begin{cases} 1 & \text{, if } b = 0.5 \\ \frac{b}{1-b} \frac{1}{g_2\left(g_1^{-1}(\frac{b}{1-b})\right)} & \text{, if } 0 < \min\{b, 1-b\} < 0.5 \end{cases}$$

The hypothesis $g_2\left(g_1^{-1}\left(\frac{b}{1-b}\right)\right) \leq \frac{b}{1-b}$ implies that the last limit above is greater than or equal to one $\forall\, 0 < b < 1$, and so by proposition 1 the result follows.$\square$

Proposition 3 : *Let* M *be an MM-estimate of regression based on* $\chi_1 \in C_b$ *and* $\chi_2 \in C$, $0 < b < 0.5$. *Then, if either*

- $b = 0.5$

  or

[2] If $\chi_1(x) \geq \chi_2(x) \forall x$, then $g_1(t) \geq g_2(t) \forall t$, and so $g_2(g_1^{-1}(t)) \leq t$, $\forall t$. Usually, χ_1 and χ_2 are taken in the same family of functions (e.g. Tukey's family, see section 4.1) . In this case, since χ_1 is chosen to attain a high breakdown point and χ_2 to attain high efficiency, the choice $\chi_1 \geq \chi_2$ is the natural one to do.

- $0 < b < 0.5$, $g_2\left(g_1^{-1}\left(\frac{b}{1-b}\right)\right) \leq \frac{b}{1-b}$ [3], $\exists\, c \geq 0$ *and* $d > c$ *such that* $\chi_1(x) = 0$, $\forall\, |x| \leq c$ *and* χ_1 *is strictly increasing and two times continuously differentiable on* (c, d)

$$\lim_{\epsilon \to b} \frac{B_M^2(\epsilon)}{B_{S_b}^2(\epsilon)} \geq 1.$$

Proof: Suppose first that $b = 0.5$.

For each $0 < \epsilon < 0.5$, let $f_\epsilon(b) = B_{S_b}^2(\epsilon)$ and $b(\epsilon) = \text{argmin}_{\epsilon < b < 1-\epsilon} f_\epsilon(b)$.

It was proved by Yohai and Zamar (1991) that if T is an estimate of regression depending only on the residuals, then for each $0 < \epsilon < 0.5$, $B_{S_{b(\epsilon)}}^2(\epsilon) \leq B_T^2(\epsilon)$.

This fact implies that

$$\lim_{\epsilon \to 0.5} \frac{B_T^2(\epsilon)}{B_{S_{b(\epsilon)}}^2(\epsilon)} \geq 1.$$

Now, note that since $b(\epsilon) \to 0.5$ as $\epsilon \uparrow 0.5$,

$$\lim_{\epsilon \to 0.5} \frac{B_{S_{b(\epsilon)}}^2(\epsilon)}{B_{S_{\frac{1}{2}}}^2(\epsilon)} = 1.$$

and since

$$\lim_{\epsilon \to 0.5} \frac{B_M^2(\epsilon)}{B_{S_{\frac{1}{2}}}^2(\epsilon)} = \lim_{\epsilon \to 0.5} \frac{B_M^2(\epsilon)}{B_{S_{b(\epsilon)}}^2(\epsilon)} \frac{B_{S_{b(\epsilon)}}^2(\epsilon)}{B_{S_{\frac{1}{2}}}^2(\epsilon)}$$

the assertion follows.

Now suppose that $0 < b < 0.5$, then[4]:

$$\lim_{\epsilon \to b} \frac{B_M^2(\epsilon)}{B_{S_b}^2(\epsilon)} = \begin{cases} \infty & \text{if } c = 0 \\ \frac{a^2}{c^2}\left(\frac{g_2^{-1}(b/(1-b))}{h^{-1}(b/(1-b))}\right)^2 & \text{if } c > 0 \end{cases}$$

Suppose that $c > 0$ and denote by S^1 the S-estimate based on χ_1. Since $g_1^{-1}\left(\frac{b}{1-b}\right) \leq g_2^{-1}\left(\frac{b}{1-b}\right)$

$$\lim_{\epsilon \to b} \frac{B_M^2(\epsilon)}{B_{S_b}^2(\epsilon)} \geq \lim_{\epsilon \to b} \frac{B_{S^1}^2(\epsilon)}{B_{S_b}^2(\epsilon)} \geq 1. \square$$

[3] this condition will be automatically satisfied for the MM-estimates such as we have defined them in Section 3.4 since it is required that $\chi_2 \leq \chi_1$.

[4] We delay the proof of this fact until Section 5.5.

5.2 The Definition of the Breakdown Rate

Define $\mathcal{E}_b$ as the family of S-,τ- and MM-estimates with breakdown point b.

The results of the previous section motivate us to define the *Breakdown Rate (BR)* of an estimate $T_b \in \mathcal{E}_b$ as:

$$BR^2(T_b) = \lim_{\epsilon \to b} \frac{B^2_{T_b}(\epsilon)}{B^2_{S_b}(\epsilon)}$$

Under the assumptions stated in propositions 1, 2 and 3, the BR indicates the speed of divergence to infinity of the square of the maximum asymptotic bias function of an estimate T_b with respect to that of a baseline function, namely the maximum asymptotic bias of S_b, $B^2_{S_b}(\epsilon)$.

The BR is a measure of global robustness which summarizes information contained in the last portion of the maximum asymptotic bias function of T_b. It provides a simple way of comparing robust estimates with the same breakdown point.

Note that we are comparing all estimates of $\mathcal{E}_b$ with the same estimate, namely $S_b \in \mathcal{E}_b$, the min-max asymptotic bias S-estimate of regression.

5.3 Breakdown Rate of S-Estimates of Regression

Let $0 < b < 1$, $\chi \in \mathcal{C}_b$ and S be an S-estimate of regression based on χ.

In this section we calculate the breakdown rate of S, that is

$$\begin{aligned} BR^2(S) &= \lim_{\epsilon \to b} \frac{B^2_S(\epsilon)}{B^2_{S_b}(\epsilon)} \\ &= \lim_{\epsilon \to b} \left[\frac{g^{-1}\left(\frac{b}{1-\epsilon}\right)}{h^{-1}\left(\frac{b}{1-\epsilon}\right)} \frac{h^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}{g^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)} \right]^2 \end{aligned}$$

where,

$$h(t) = 2\left(1 - \Phi\left(\frac{a}{t}\right)\right), \text{ with } a \text{ such that } b = 2\left(1 - \Phi(a)\right)$$

and

$$g(t) = \int_{-\infty}^{\infty} \chi(tx)\varphi(x)\,dx.$$

The following two results show that we can restrict our attention to the case $0 < b \le 0.5$.

Lemma 2 : *Let* $0.5 < b < 1$ *and* $\chi_1 \in \mathcal{C}_b$, $\chi_2 \in \mathcal{C}_{1-b}$. *Further assume that either:*

- $0 < \int_0^\infty \chi_1'(y) y\, dy < \infty$,

  or

- $\chi_i = \chi_{a_i}$, $i = 1, 2$ *where* $2[1 - \Phi(a_1)] = b$ *and* $2[1 - \Phi(a_2)] = 1 - b$.

  Denote by S^b *the S-estimate of regression based on* χ_1 *and* S^{1-b} *the S-estimate of regression based on* χ_2. *Then,*

$$RBR(S^b, S^{1-b}) = \infty.$$

Proof: Assume that $0 < \int_0^\infty \chi_1'(y) y\, dy < \infty$. Let $g(t) = g_{\chi_1}(t)$ and $f(t) = g_{\chi_2}(t)$.

$$RBR(S^b, S^{1-b}) = \lim_{\epsilon \to 1-b} \left[\frac{g^{-1}\left(\frac{b}{1-\epsilon}\right) f^{-1}\left(1 - \frac{b}{1-\epsilon}\right)}{g^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right) f^{-1}\left(\frac{1-b}{1-\epsilon}\right)} \right]^2.$$

We can apply L'Hôspital's rule to compute the limit:

$$\begin{aligned} \lim_{t \to 1} g^{-1}(t) f^{-1}(1-t) &= \lim_{t \to 1} \frac{f^{-1}(1-t)}{1/g^{-1}(t)} \\ &= \lim_{t \to 1} (g^{-1}(t))^2 \frac{g'(g^{-1}(t))}{f'(f^{-1}(1-t)} \end{aligned}$$

We can write,

$$\begin{aligned} t^2 g'(t) &= 2\left[\varphi(0) \int_0^\infty \chi_1'(x) x\, dx + \int_0^\infty o(x/t) \chi_1'(x) x\, dx\right] \\ f'(t) &= 2 \int_0^\infty \chi_2'(tx) x \varphi(x)\, dx. \end{aligned}$$

Thus,

$$\begin{aligned} \lim_{t \to 1} (g^{-1}(t))^2 \frac{g'(g^{-1}(t))}{f'(f^{-1}(t)} &= \lim_{t \to 1} \frac{\varphi(o) \int_0^\infty \chi_1'(x) x\, dx + \int_0^\infty o(x/(g^{-1}(t)) \chi_1'(x) x\, dx}{\int_0^\infty \chi_2'(f^{-1}(1-t)x) x \varphi(x)\, dx} \\ &= \infty. \end{aligned}$$

Now suppose that χ_i is of the jump type with jump constant a_i, $i = 1, 2$ where $2[1 - \Phi(a_1)] = b$ and $2[1 - \Phi(a_2)] = 1 - b$.

$$RBR(S^b, S^{1-b}) = \lim_{\epsilon\to 1-b}\left[\frac{\Phi^{-1}\left(1-\frac{b-\epsilon}{2(1-\epsilon)}\right)}{\Phi^{-1}\left(1-\frac{1-b-\epsilon}{2(1-\epsilon)}\right)}\frac{\Phi^{-1}\left(1-\frac{1-b}{2(1-\epsilon)}\right)}{\Phi^{-1}\left(1-\frac{b}{2(1-\epsilon)}\right)}\right]^2.$$

Now, by applying L'Hôspital's rule we get

$$\lim_{\epsilon\to 1-b}\Phi^{-1}\left(1-\frac{1-b-\epsilon}{2(1-\epsilon)}\right)\Phi^{-1}\left(1-\frac{b}{2(1-\epsilon)}\right) = \lim_{\epsilon\to 1-b}\frac{\Phi^{-1}\left(1-\frac{b}{2(1-\epsilon)}\right)}{1/\Phi^{-1}\left(1-\frac{1-b-\epsilon}{2(1-\epsilon)}\right)}$$

$$= \lim_{\epsilon\to 1-b}\left[\Phi^{-1}\left(1-\frac{1-b-\epsilon}{2(1-\epsilon)}\right)\right]^2\frac{\varphi\left(\Phi^{-1}\left(1-\frac{1-b}{2(1-\epsilon)}\right)\right)}{\varphi\left(\Phi^{-1}\left(1-\frac{b}{2(1-\epsilon)}\right)\right)}$$

$$= \frac{1}{\varphi(0)}\lim_{t\to\infty}t^2\varphi(t) = 0$$

And so, $RBR(S^b, S^{1-b}) = \infty$.□

Let $\mathcal{S}_b$ be the family of S-estimates based on χ functions such that $\chi \in \mathcal{C}_b$.

Proposition 4 : *If* $0.5 < b < 1$ *and* $S \in \mathcal{S}_b$ *then* $BR(S) = \infty$.

Proof: Let S^{1-b} be any S-estimate in $\mathcal{S}_{1-b}$ such that it is based on a function of the same type as the function on which S is based (i.e. if S is based on a jump type function then S^{1-b} should be based on a jump type function as well, and similarly if S is based on a continuously differentiable in all but a finite number of points function). Then, since

$$BR^2(S) = RBR(S, S^{1-b})BR^2(S^{1-b}) \geq RBR(S, S^{1-b})$$

the result follows as a consequence of the previous lemma.□

We will concentrate now in the case when $0 < b \leq 0.5$.

Note that if $L_1 = \lim_{\epsilon\to b}\frac{g^{-1}\left(\frac{b}{1-\epsilon}\right)}{h^{-1}\left(\frac{b}{1-\epsilon}\right)}$ and $L_2 = \lim_{\epsilon\to b}\frac{h^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}{g^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}$ both exist, then $BR^2(S) = (L_1L_2)^2$.

Lemma 3 : *Suppose that* $\chi \in \mathcal{C}_b$ *and* $0 < \int_0^\infty \chi'(y)y\,dy < \infty$. *Then*

$$L_1 = \lim_{t\to 1}\frac{g^{-1}(t)}{h^{-1}(t)} = \frac{1}{a}\int_0^\infty \chi'(y)y\,dy.$$

Proof: Since $h^{-1}(t)$ and $g^{-1}(t)$ tend to infinity as t tends to one, we can apply L'Hôspital's rule to compute L_1.

Let

$$\begin{aligned} L_1^* &= \lim_{t\to 1} \frac{\frac{d}{dt}\frac{1}{h^{-1}(t)}}{\frac{d}{dt}\frac{1}{g^{-1}(t)}} \\ &= \frac{[g^{-1}(t)]^2 g'(g^{-1}(t))}{[h^{-1}(t)]^2 h'(h^{-1}(t))}. \end{aligned}$$

Then, if L_1^* exists, $L_1 = L_1^*$.

Note that

$$\begin{aligned} h'(t) &= \frac{2}{t^2} a\varphi\left(\frac{a}{t}\right), \\ g'(t) &= \frac{2}{t^2}\int_0^\infty \chi'(y) y \varphi\left(\frac{y}{t}\right)\, dy; \end{aligned}$$

and that a Taylor's series expansion of order 1 around 0 gives

$$\varphi\left(\frac{z}{t}\right) = \varphi(0) + o\left(\frac{z}{t}\right) \;, \text{ as } \tfrac{z}{t} \to 0$$

so that we can write,

$$\begin{aligned} h'(t) &= 2\frac{a}{t^2}\left[\varphi(0) + o\left(\frac{a}{t}\right)\right] \;, \text{ as } t \to \infty; \\ g'(t) &= 2\frac{1}{t^2}\left[\varphi(0)\int_0^\infty \chi'(y) y\, dy + \int_0^\infty \chi'(y) y\, o\left(\frac{y}{t}\right)\, dy\right] \;, \text{ as } t \to \infty. \end{aligned}$$

Hence,

$$L_1^* = \lim_{t\to 1} \frac{\left(g^{-1}(t)\right)^2 g'(g^{-1}(t))}{\left(h^{-1}(t)\right)^2 h'(h^{-1}(t))} = \frac{1}{a}\int_o^\infty \chi'(y) y\, dy$$

and the result follows.□

Lemma 4 : *Let $\chi \in C_b$ and $L_2 = \lim_{t\to 0} \frac{h^{-1}(t)}{g^{-1}(t)}$. Suppose that $\exists\, c \geq 0$ and $d > c$ such that $\chi(y) = 0\, \forall y \in [-c, c]$ and χ is strictly increasing and two times continuously differentiable in (c, d).*

Then if $c = 0$, $L_2 = \infty$ and if $c > 0$,

$$L_2 = \frac{a}{c}.$$

Proof: Since $h^{-1}(t)$, $g^{-1}(t)$, $h'(t)$, $g'(t) \to 0$ as $t \to 0$, we can apply L'Hôspital's rule (two times) to compute L. Let

$$\tilde{L}_2 = \lim_{t\to 0} \frac{\frac{d}{dt}\frac{1}{[g^{-1}(t)]^2}}{\frac{d}{dt}\frac{1}{[h^{-1}(t)]^2}} = \frac{[h^{-1}(t)]^3 h'(h^{-1}(t))}{[g^{-1}(t)]^3 g'(g^{-1}(t))}$$

and

$$\begin{aligned}
\hat{L}_2 &= \lim_{t\to 0} \frac{\frac{d}{dt}\{[h^{-1}(t)]^3 h'(h^{-1}(t))\}}{\frac{d}{dt}\{[g^{-1}(t)]^3 g'(g^{-1}(t))\}} \\
&= \frac{3[h^{-1}(t)]^2 + [h^{-1}(t)]^3 \frac{h''(h^{-1}(t))}{h'(h^{-1}(t))}}{3[g^{-1}(t)]^2 + [g^{-1}(t)]^3 \frac{g''(g^{-1}(t))}{g'(g^{-1}(t))}}
\end{aligned}$$

If $\hat{L}_2$ exists, then $\tilde{L}_2 = \hat{L}_2$ and so $L_2^2 = \hat{L}_2$.

Now,

$$\begin{aligned}
h'(t) &= 2\varphi\left(\frac{a}{t}\right)\frac{a}{t^2} \\
h''(t) &= h'(t)\frac{1}{t}\left(\frac{a^2}{t^2} - 2\right) \\
g'(t) &= 2\frac{1}{t^2}\int_0^\infty \chi'(y) y \varphi\left(\frac{y}{t}\right)\, dy \\
g''(t) &= 2\frac{1}{t^5}\int_0^\infty \chi'(y) y^3 \varphi\left(\frac{y}{t}\right)\, dy - 2\frac{1}{t}g'(t)
\end{aligned}$$

so that

$$\begin{aligned}
\frac{g''(t)}{g'(t)} &= \frac{1}{t^3}\frac{\int_0^\infty \chi'(y) y^3 \varphi\left(\frac{y}{t}\right)\, dy}{\int_0^\infty \chi'(y) y \varphi\left(\frac{y}{t}\right)\, dy} - \frac{2}{t} = \frac{1}{t^3}\left\{\frac{\int_0^\infty \chi'(y) y^3 \varphi\left(\frac{y}{t}\right)\, dy}{\int_0^\infty \chi'(y) y \varphi\left(\frac{y}{t}\right)\, dy} - 2t^2\right\} \\
\frac{h''(t)}{h'(t)} &= \frac{1}{t^3}\left(a^2 - 2t^2\right)
\end{aligned}$$

and so

$$\hat{L}_2 = \lim_{t\to 0} \frac{a^2 + h^{-1}(t)}{\frac{\int_0^\infty \chi'(y) y^3 \varphi\left(\frac{y}{t}\right) dy}{\int_0^\infty \chi'(y) y \varphi\left(\frac{y}{t}\right) dy} + g^{-1}(t)}.$$

Under the stated hypothesis on χ,

$$\frac{\int_0^\infty \chi'(y) y^3 \varphi\left(\frac{y}{t}\right)\, dy}{\int_0^\infty \chi'(y) y \varphi\left(\frac{y}{t}\right)\, dy} = t^2 \frac{\int_{c/t}^\infty \chi'(ty) y^3 \varphi(y)\, dy}{\int_{c/t}^\infty \chi'(ty) y \varphi(y)\, dy}.$$

Let

$$f(t,h) = t^2 \frac{\int_{(c+h)/t}^{\infty} \chi'(ty) y^3 \varphi(y)\, dy}{\int_{(c+h)/t}^{\infty} \chi'(ty) y \varphi(y)\, dy} \quad ; t > 0 \ , \ h > 0.$$

Let $t, h > 0$ and $y \geq (c+h)/t$. Then, by performing a Taylor's series expansion of order 0 around $c + h$ we can write

$$\chi'(ty) = \chi'(c+h) + R_0(ty, c+h)$$

where $R_0(ty, c+h) = \chi''(\xi)[ty - (c+h)]$ for some $\xi \in (c+h, ty)$.

Then,

$$f(t,h) = t^2 \frac{\chi'(c+h) \int_{(c+h)/t}^{\infty} y^3 \varphi(y)\, dy + \int_{(c+h)/t}^{\infty} R(ty, c+h) y^3 \varphi(y)\, dy}{\chi'(c+h) \int_{(c+h)/t}^{\infty} y \varphi(y)\, dy + \int_{(c+h)/t}^{\infty} R(ty, c+h) y \varphi(y)\, dy}.$$

But,

$$t^2 \frac{\int_{(c+h)/t}^{\infty} y^3 \varphi(y)\, dy}{\int_{(c+h)/t}^{\infty} y \varphi(y)\, dy} = \frac{\varphi\left(\frac{c+h}{t}\right) [(c+h)^2 + 2t^2]}{\varphi\left(\frac{c+h}{t}\right)} = (c+h)^2 + 2t^2;$$

$$t^2 \frac{\int_{(c+h)/t}^{\infty} R_0(ty, c+h) y^3 \varphi(y)\, dy}{\int_{(c+h)/t}^{\infty} y \varphi(y)\, dy} = \int_0^{\infty} R_0(tx + c + h, c+h)(tx + c + h)^3 \frac{1}{t} \frac{\varphi\left(x + \frac{c+h}{t}\right)}{\varphi\left(\frac{c+h}{t}\right)}\, dx$$

and

$$\frac{\int_{(c+h)/t}^{\infty} R_0(ty, c+h) y \varphi(y)\, dy}{\int_{(c+h)/t}^{\infty} y \varphi(y)\, dy} = \int_0^{\infty} R_0(tx + c + h, c+h)(tx + c + h) \frac{1}{t} \frac{\varphi\left(x + \frac{c+h}{t}\right)}{\varphi\left(\frac{c+h}{t}\right)}\, dx$$

where $R_0(tx + c + h, c + h) = \chi''(\xi) tx$, for some $\xi \in (c + h, tx + c + h)$ for each $x \geq 0$.

Therefore, $f(t,h) \longrightarrow c^2$ for $(t,h) \to (0,0)^+$ and

$$\hat{L}_2 = \frac{a^2}{c^2}. \square$$

Theorem 2 : *Let* $0 < b \leq 0.5$ *and* $\chi \in C_b$. *Suppose that* $\exists\, c \geq 0$ *and* $d > c$ *such that* $\chi(y) = 0\, \forall y \in [-c, c]$ *and* χ *is strictly increasing and two times continuously differentiable in* (c, d). *Let* S *be the S-estimate based on* χ.

- *If* $c = 0$ *and* $0 < \int_0^{\infty} \chi'(y) y\, dy < \infty$, *then* $BR^2(S) = \infty\ \forall b$.

- *If* $c > 0$ *and* $0 < b < 0.5$, *then*

$$BR^2(S) = \left[\frac{1}{c}\Phi^{-1}\left(1 - \frac{b}{2(1-b)}\right) g^{-1}\left(\frac{b}{1-b}\right)\right]^2.$$

- *If* $c > 0$, $0 < \int_0^\infty \chi'(y) y \, dy < \infty$ *and* $b = 0.5$, *then*

$$BR(S)^2 = \left[\frac{1}{c}\int_c^\infty \chi'(y) y \, dy\right]^2. \tag{5.2}$$

Proof: It is a direct consequence of lemmas 3 and 4.□

EXAMPLE 3: Let

$$\chi_{C,A} = \begin{cases} 0 & \text{if } 0 \le |x| < C \\ \frac{x^2 - C^2}{A^2 - C^2} & \text{if } C \le |x| \le A \\ 1 & \text{if } |x| > A \end{cases}$$

and let $S^{C,A}$ be the S-estimate based on $\chi_{C,A}$. Let $0 < b \le 0.5$ be such that $E_\Phi \chi_{C,A}(Z) = b$ that is

$$\frac{2}{A^2 - C^2}[\Phi(A)(1 - A^2) - A\varphi(A) - \Phi(C)(1 - C^2) + C\varphi(C)] + 2 = b.$$

The choice $C = 0.202$ and $A = 1$ gives $BP(S^{C,A}) = 0.5$.

Since

$$g'_{C,A}(1) = \frac{4}{A^2 - C^2}[\Phi(A) - A\varphi(A) - \Phi(C) + C\varphi(C)]$$

we have that

$$SENS(S^{C,A}) = \frac{A^2 - C^2}{2[\Phi(A) - A\varphi(A) - \Phi(C) + C\varphi(C)]}$$

and the efficiency of $S^{C,A}$ at the Gaussian model is

$$e(S^{C,A}) = 2[\Phi(A) - A\varphi(A) - \Phi(C) + C\varphi(C)].$$

To compute the breakdown rate of $S^{C,A}$ note first that

$$\int_0^\infty \chi'_{C,A}(x) x \, dx = \frac{2}{3}\frac{A^3 - C^3}{A^2 - C^2}$$

and by(5.2)

$$BR^2(S^{C,A}) = \left[\frac{2}{3}\frac{A^3 - C^3}{A^2 - C^2}\frac{1}{C}\right]^2.$$

estimate	C	A	BP	$SENS$	efficiency	BR
$S^{C,A}$	0.202	1	0.5	4.879	0.196	3.412
S^A	0	1.041	0.5	4.950	0.219	∞

Table 5.1: Comparison of two S-estimates with the same BP but markedly different bias performance

Note that if $C = 0$, the estimate reduces to the one introduced in Section 4.1, Example 1, based on χ_A given by (4.8). We summarize the quantifiers calculated above for the specific values of C and A in Table 5.1 including S^A as well.

Notice that these two estimates have very similar asymptotic properties such as the BP, $SENS$ and efficiency. They can only be distinguished in terms of their BR.

5.4 Breakdown Rate of τ-Estimates of Regression

Let $0 < b < 1$, $\chi_1 \in \mathcal{C}_b$, $\chi_2 \in \mathcal{C}$ and τ be a τ-estimate of regression based on χ_1 and χ_2.

If $b \neq 0.5$, the breakdown rate of τ is

$$\begin{aligned} BR^2(\tau) &= \lim_{\epsilon \to b} \frac{B_\tau^2(\epsilon)}{B_{S_b}^2(\epsilon)} \\ &= \frac{b}{1-b}\left[g_2\left(g_1^{-1}\left(\frac{b}{1-b}\right)\right)\right]^{-1} \lim_{\epsilon \to b}\left[\frac{g_1^{-1}\left(\frac{b}{1-\epsilon}\right)}{h^{-1}\left(\frac{b}{1-\epsilon}\right)} \frac{h^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}{g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}\right]^2, \end{aligned}$$

and if $b = 0.5$

$$BR^2(\tau) = \lim_{\epsilon \to b}\left[\frac{g_1^{-1}\left(\frac{b}{1-\epsilon}\right)}{h^{-1}\left(\frac{b}{1-\epsilon}\right)} \frac{h^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}{g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}\right]^2;$$

where $g_i(t) = g_{\chi_i}(t)$, $i = 1, 2$ and h is the same function defined in Section 5.3.

Lemma 5 : *Let* $0.5 < b < 1$, $\chi_1^i, \chi_2^i \in \mathcal{C}$, $i = 1, 2$, $\chi_1^1 \in \mathcal{C}_b$ *and* $\chi_1^2 \in \mathcal{C}_{1-b}$. *Let* τ^i *be the* τ*-estimate based on* χ_1^i *and* χ_2^i. *Then*

$$RBR(\tau^1, \tau^2) = \infty.$$

Proof: Let $g_j(t) = g_{\chi_j^1(t)}$ and $f_j(t) = g_{\chi_j^2(t)}$, $j = 1, 2$.

$$RBR(\tau^1, \tau^2) = \lim_{\epsilon \to 1-b} \left(\frac{b}{1-b}\right)^2 \frac{\left[g_2\left(g_1^{-1}\left(\frac{b}{1-b}\right)\right)\right]^{-1}}{\left[f_2\left(f_1^{-1}\left(\frac{1-b}{b}\right)\right)\right]^{-1}} \left[\frac{g_1^{-1}\left(\frac{b}{1-\epsilon}\right) f_1^{-1}\left(1 - \frac{b}{1-\epsilon}\right)}{g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right) f_1^{-1}\left(\frac{1-b}{1-\epsilon}\right)}\right]^2,$$

and by lemma 1, this limit is equal to infinity.□

Let $\mathcal{T}_b$ be the family of τ-estimates based on functions $\chi_1 \in \mathcal{C}_b$ and $\chi_2 \in \mathcal{C}$.

The following result shows that we can restrict our attention to the case $0 < b \leq 0.5$.

Proposition 5 : *If* $0.5 < b < 1$, $\chi_1 \in \mathcal{C}_b$, $\chi_2 \in \mathcal{C}$ *and* τ *is the* τ*-estimate of regression based on* χ_1 *and* χ_2, *then*

$$BR(\tau) = \infty.$$

Proof: Let τ^{1-b} be any τ-estimate in $\mathcal{T}_{1-b}$. Then since $BR^2(\tau) \geq RBR(\tau, \tau^{1-b})$, the result follows as a consequence of Lemma 5.□

Theorem 3 : *Let* $0 < b \leq 0.5$, $\chi_1 \in \mathcal{C}_b$, $\chi_2 \in \mathcal{C}$ *and* τ *be the* τ*-estimate of regression based on* χ_1 *and* χ_2. *Suppose that* $\exists\, c \geq 0$ *and* $d > c$ *such that* $\chi_1(y) = 0\, \forall y \in [-c, c]$ *and* χ_1 *is strictly increasing and two times continuously differentiable in* (c, d). *Then,*

- *if* $c = 0$ *and* $0 < \int_0^\infty \chi_1'(y) y\, dy < \infty$, *then* $BR(\tau) = \infty$;
- *if* $c > 0$ *and* $b < 0.5$,
$$BR^2(\tau) = \frac{b}{1-b}\left[g_2\left(g_1^{-1}\left(\frac{b}{1-b}\right)\right)\right]^{-1}\left[\frac{1}{c}\Phi^{-1}\left(1 - \frac{b}{2(1-b)}\right) g_1^{-1}\left(\frac{b}{1-b}\right)\right]^2;$$
- *if* $c > 0$, $b = 0.5$ *and* $0 < \int_0^\infty \chi_1'(y) y\, dy < \infty$,

$$BR^2(\tau) = \left[\frac{1}{c}\int_c^\infty \chi_1'(y) y\, dy\right]^2. \tag{5.3}$$

Proof: It is also a consequence of lemmas 3 and 4.□

Corollary 1 : *If* $b = 0.5$, $S \in \mathcal{S}_b$ *is based on some function* χ_1 *and* $\tau \in \mathcal{T}_b$ *is based on* χ_1 *and* χ_2, *then* $BR(S) = BR(\tau)$. *If* $b < 0.5$ *then* $BR(S) \leq BR(\tau)$.

Remark: In the case $c > 0$, $b = 0.5$, the BR of the τ-estimate does not depend on the choice of χ_2.

5.5 Breakdown Rate of MM-Estimates of Regression

Let $0 < b \leq 0.5$, $\chi_1 \in \mathcal{C}_b$, $\chi_2 \in \mathcal{C}$ and M be an MM-estimate of regression based on χ_1 and χ_2.

The breakdown rate of M is

$$\begin{aligned} BR^2(M) &= \lim_{\epsilon \to b} \frac{B_M^2(\epsilon)}{B_{S_b}^2(\epsilon)} \\ &= \lim_{\epsilon \to b} \left(\frac{g_2^{-1}\left[g_2\left(g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right]}{h^{-1}\left(\frac{b}{1-\epsilon}\right)} \frac{h^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)}{g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)} \right)^2 . \end{aligned}$$

Lemma 6 : *Let* $0 < b \leq 0.5$ *and* $\chi_i \in \mathcal{C}$, $i = 1, 2$. *Then,*

- *if* $b < 0.5$,
$$\lim_{\epsilon \to b} \frac{g_2^{-1}\left[g_2\left(g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right]}{h^{-1}\left(\frac{b}{1-\epsilon}\right)} = \frac{g_2^{-1}\left(\frac{b}{1-b}\right)}{h^{-1}\left(\frac{b}{1-b}\right)};$$

- *if* $b = 0.5$ *and* $0 < \int_0^\infty \chi_2'(y) y \, dy < \infty$,
$$\lim_{\epsilon \to b} \frac{g_2^{-1}\left[g_2\left(g_1^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right]}{h^{-1}\left(\frac{b}{1-b}\right)} = \frac{1}{a} \int_0^\infty \chi_2'(y) y \, dy \left(2 - \lim_{t \to 0} \frac{g_2'(t)}{g_1'(t)}\right)^{-1} .$$

Proof: Assume that $b = 0.5$. We can apply L'Hôspital's rule to compute the limit of interest: if

$$\begin{aligned} L &= \lim_{\epsilon \to 0.5} \frac{g_2^{-1}\left(g_2\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right)}{h^{-1}\left(\frac{0.5}{1-\epsilon}\right)} \\ &= \lim_{\epsilon \to 0.5} \frac{\left[h^{-1}\left(\frac{0.5}{1-\epsilon}\right)\right]^{-1}}{\left[g_2^{-1}\left(g_2\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right)\right]^{-1}} \end{aligned}$$

and

$$L' = \lim_{\epsilon \to b} \frac{\frac{d}{dt}\left[h^{-1}\left(\frac{0.5}{1-\epsilon}\right)\right]^{-1}}{\frac{d}{dt}\left[g_2^{-1}\left(g_2\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right)\right]^{-1}}$$

or

$$\begin{aligned} L' = \; & \frac{g_2'\left(g_2^{-1}\left(g_2\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right)\right)\left[g_2^{-1}\left(g_2\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right)\right]^2}{h'\left(h^{-1}\left(\frac{0.5}{1-\epsilon}\right)\right)\left[h^{-1}\left(\frac{0.5}{1-\epsilon}\right)\right]^2} \times \\ & \left\{2 - \frac{g_2'\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right)}{g_1'\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right)}\right\}^{-1} . \end{aligned}$$

Then, if L' exists, $L = L'$.

Now, since

$$\begin{aligned} h'(t) &= 2\frac{a}{t^2}\left[\varphi(0) + o\left(\frac{a}{t}\right)\right] \text{ as } t \to \infty \\ g_2'(t) &= 2\frac{1}{t^2}\left[\varphi(0)\int_0^\infty \chi_2'(y)y\,dy + \int_0^\infty \chi_2'(y)y\,o\left(\frac{y}{t}\right)\,dy\right] \text{ as } t \to \infty; \end{aligned}$$

then,

$$\lim_{\epsilon\to 0.5} \frac{\left[g_2^{-1}\left(g_2\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right)\right]^2 g_2'\left(g_2^{-1}\left(g_2\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right)\right)}{\left[h^{-1}\left(\frac{0.5}{1-\epsilon}\right)\right]^2 h'\left(h^{-1}\left(\frac{0.5}{1-\epsilon}\right)\right)} = \frac{1}{a}\int_0^\infty \chi_2'(y)y\,dy. \square$$

Remark: Unfortunately, we were not able to compute $\lim_{t\to 0}\frac{g_2'(g^{-1}(t))}{g_1'(g^{-1}(t))}$ in general. However we will compute it for one important special case.

Let χ_1 , $\chi_2 \in \mathcal{C}$ be such that $\chi_2(x) \le \chi_1(x), \forall\, x$. Suppose that there exists $c_1 \ge 0$ such that $\chi_1(x) = 0$ if $x \in [0, c_1]$. Then there exists $c_2 \ge c_1$ such that $\chi_2(x) = 0$ if $x \in [0, c_2]$.

If we want to define an MM-estimate based on χ_1 and χ_2, the choice of χ_2 should be done to obtain efficiency at the Gaussian model and so χ_2 should be as closely as possible to $\chi(x) = x^2$. For this reason we only consider the case $c_1 = c_2 = c$.

Now,

$$\begin{aligned} \frac{g_2'(t)}{g_1'(t)} &= \frac{\int_c^\infty \chi_2'(y)y\varphi\left(\frac{y}{t}\right)\,dy}{\int_c^\infty \chi_1'(y)y\varphi\left(\frac{y}{t}\right)\,dy} \\ &= \frac{\int_{c/t}^\infty \chi_2'(ty)y\varphi(y)\,dy}{\int_{c/t}^\infty \chi_1'(ty)y\varphi(y)\,dy}. \end{aligned}$$

If we let

$$f(t,h) = \frac{\int_{(c+h)/t}^\infty \chi_2'(ty)y\varphi(y)\,dy}{\int_{(c+h)/t}^\infty \chi_1'(ty)y\varphi(y)\,dy} \text{ for } t, h > 0i$$

and assume that there exists $d > c$ such that χ_i is strictly increasing and two times continuously differentiable in (c, d). By continuity of χ_i' on (c, d), for each $y > (c+h)/t$ we have

$$\begin{aligned} \chi_2'(ty) &= \chi_2'(c+h) + \tilde{R}_0(ty, c+h), \\ \chi_1'(ty) &= \chi_1'(c+h) + R_0(ty, c+h). \end{aligned}$$

where $\tilde{R}_0(ty, c+h)$ and $R_0(ty, c+h)$ converge to zero as ty tends to $c+h$.

Then,

$$\begin{aligned} f(t,h) &= \frac{\chi_2'(c+h)\int_{(c+h)/t}^{\infty} y\varphi(y)\,dy + \int_{(c+h)/t}^{\infty} \tilde{R}_0(ty, c+h)y\varphi(y)\,dy}{\chi_1'(c+h)\int_{(c+h)/t}^{\infty} y\varphi(y)\,dy + \int_{(c+h)/t}^{\infty} R_0(ty, c+h)y\varphi(y)\,dy} \\ &= \frac{\chi_2'(c+h)\varphi\left(\frac{c+h}{t}\right) + \int_{(c+h)/t}^{\infty} \tilde{R}_0(ty, c+h)y\varphi(y)\,dy}{\chi_1'(c+h)\varphi\left(\frac{c+h}{t}\right) + \int_{(c+h)/t}^{\infty} R_0(ty, c+h)y\varphi(y)\,dy} \\ &= \frac{\frac{\chi_2'(c+h)}{\chi_1'(c+h)} + \frac{1}{\chi_1'(c+h)}\int_{(c+h)/t}^{\infty} \tilde{R}_0(ty, c+h)y\frac{\varphi(y)}{\varphi((c+h)/t)}\,dy}{1 + \frac{1}{\chi_1'(c+h)}\int_{(c+h)/t}^{\infty} R_0(ty, c+h)y\frac{\varphi(y)}{\varphi((c+h)/t)}\,dy} \end{aligned}$$

Then, it's easy to see that if χ_i is not differentiable at c, then

$$f(t,h) \longrightarrow \frac{\chi_2'(c^+)}{\chi_1'(c^+)}, \text{ as } (t,h) \to (0,0);$$

where $\chi_i'(c^+)$ denotes the right lateral derivative of χ_i at c, $i = 1, 2$.

Theorem 4 : *Let $0 < b \leq 0.5$, $\chi_1 \in \mathcal{C}_b$, $\chi_2 \in \mathcal{C}$, $\chi_2(x) \leq \chi_1(x)$, $\forall\, x$ and M be the MM-estimate of regression based on χ_1 and χ_2. Suppose that $\exists\, c \geq 0$ and $d > c$ such that $\chi_1(x) = 0\,\forall\,|x| \leq c$ and χ_1 is strictly increasing and two times continuously differentiable on (c, d). Then,*

- *if $0 < b < 0.5$ and $c > 0$,*

$$BR^2(M) = \left[\frac{1}{c}\Phi^{-1}\left(1 - \frac{b}{2(1-b)}\right) g_2^{-1}\left(\frac{b}{1-b}\right)\right]^2$$

Further assume that $0 < \int_0^{\infty} \chi_2'(y)y\,dy < \infty$. Then,

- *if $c = 0$, $BR(M) = \infty$;*
- *if $b = 0.5$, $c > 0$ and $\lim_{t\to 0}\frac{g_2'(t)}{g_1'(t)}$ exists,*

$$BR^2(M) = \left\{\frac{1}{c}\int_0^{\infty} \chi_2'(y)y\,dy\left[2 - \lim_{t\to 0}\frac{g_2'(g_1^{-1}(t))}{g_1'(g_1^{-1}(t))}\right]^{-1}\right\}^2. \tag{5.4}$$

Proof: Let

$$f(\epsilon) = \frac{g_2^{-1}\left[g_2\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right) + \frac{\epsilon}{1-\epsilon}\right]}{h^{-1}\left(\frac{0.5}{1-\epsilon}\right)}.$$

Then, since $g_2\left(g_1^{-1}\left(\frac{0.5-\epsilon}{1-\epsilon}\right)\right) > 0$, $g_2(t) \leq g_1(t)$ and g_2 is increasing,

$$\frac{g_2^{-1}\left(\frac{\epsilon}{1-\epsilon}\right)}{h^{-1}\left(\frac{0.5}{1-\epsilon}\right)} \leq f(\epsilon) \leq \frac{g_2^{-1}\left(\frac{0.5}{1-\epsilon}\right)}{h^{-1}\left(\frac{0.5}{1-\epsilon}\right)}.$$

By Lemma 3 the right hand side of the above inequality converges to $1/a \int_0^\infty \chi_2'(y)y\,dy$ as $\epsilon \to 0.5$ and by a similar reasoning the left hand side tends to $\frac{1}{2a}\int_0^\infty \chi_2'(y)y\,dy$ as $\epsilon \to 0.5$. Since by hypothesis $0 < \int_0^\infty \chi_2'(y)y\,dy < \infty$, there exist A_1, A_2 such that $0 < A_1 \leq A_2 < \infty$ and $A_1 \leq f(\epsilon) \leq A_2$.

Hence, by Lemma 4, if $c = 0$, $BR(M) = \infty$. If $c > 0$ and $b = 0.5$, the result is a consequence of Lemma 6.

5.6 Conclusions

Following are some conclusions obtained from the results proved in this chapter.

- The results of Chapter 5 can be used to choose the loss function, χ_1, which determines the breakdown point of S-, τ- and MM-estimates so that they have good bias-robustness properties. In particular, the fact that χ_1 should be constant and equal to zero on a neighborhood of zero (among other regularity conditions) was first discovered here.

- MM- and τ-estimates were developed for the purpose of achieving a high breakdown point and a high efficiency at the Gaussian model. The results in this chapter show that the breakdown rate of τ-estimates with breakdown point equal to 0.5 does not depend on the choice of the "efficiency determining" loss function χ_2. On the other hand, the condition that $\chi_2(x) \leq \chi_1(x)$, $\forall x$ for MM-estimates, forces χ_2 to be constant near the origin with ensuing loss of efficiency (see the remark to Lemma 6, Section 5.5).

- We think that the breakdown rate is a good criterion for defining optimality as in the following problem: "maximize the efficiency of an estimate subject to a constraint on its breakdown rate". If we can find an estimate that solves such a problem, it will be an adaptive estimate in the sense that if the model is Gaussian or nearly Gaussian the estimate will perform well (because of its high efficiency) and if the fraction of contamination is high, the estimate will perform well compared to other estimates with the same breakdown point.

- The breakdown rate of an estimate is a robustness quantifier of an asymptotic nature. It remains to be determined whether the good breakdown rate properties of an estimate carries over to finite sample situations. A next step in this work will be to perform an extensive Monte Carlo study.

Bibliography

Andrews, D.F.,Bickel, P.J.,Hampel, F.R.,Huber, P.J.,Rogers, W.H. and Tukey, J.W. (1972),*Robust Estimates of Location: Survey and Advances,* Princeton University Press, Princeton, N.J.

Donoho, D.L., and Huber, P.J. (1983), The notion of breakdown point, in *A Festschrift for Erich Lehmann,* edited by P. Bickel, K. Doksum, and J.L. Hodges, Jr., Wadswoth, Belmont, CA.

Hampel, F.R. (1968), Contributions to the theory of robust estimation, Ph.D. Thesis, University of California, Berkeley.

Hampel, F.R. (1971), A general qualitative definition of robustness, *Ann. Math. Stat.,* **42,** 1887-1896.

Hampel, F.R. (1974a), The influence curve and its role in robust estimation, *J. Am. Stat. Assoc.,* **69,** 389-393.

Hampel, F.R. (1974b), Rejection rules and robust estimates of location: an analysis of some Monte Carlo results, *Proc. European Meeting of Statisticians and 7^{th} Prague Conference on Information Theory, Statistical Decision Functions and Random Processes,* Prague, 1974.

Hampel, F.R. (1976), On the breakdown point of some rejection rules with mean, *Res. Rep. No. 11,* Fachgruppe für Statistik, Eidgen. Tech. Hochschule, Zurich.

Hampel, F.R. (1978), Optimally bounding the gross-error sensitivity and the influence position in factor space, in *Proceedings of the Statistical Computing Section of the American Statistical Association,* ASA, Washington, D.C., 59-64.

Hodges, J.L. (1967), Efficiency in normal samples and tolerance of extreme values for some estimates of location, *Proc. Fifth Berkeley Symp. Math. Stat. Probab.,* **1,** 163-168.

Huber, P.J. (1964), Robust estimation of a location parameter, *Ann. Math. Stat.*,**35**, 73-101.

Huber, P.J. (1973), Robust regression: Asymptotics, conjectures and Monte Carlo, *Ann. Stat.*, **1**, 799-821.

Huber, P.J. (1981), *Robust Statistics,* John Wiley & Sons, New York.

Martin, R.D., Yohai, V.J. and Zamar, R.H. (1989), Min-max bias robust regression, *Ann. Math. Stat.,* **4**, 1608-1630.

Martin, R.D.,and Zamar, R.H. (1989), Asymptotically min-max bias robust M-estimates of scale for positive random variables, *J. Am. Stat. Assoc.,* **406**, 494-501.

Rousseeuw, P.J. and Yohai, V.J. (1984), Robust regression by means of S-estimators, in *Robust and Nonlinear Time Series Analysis,* edited by J. Franke, W. Härdle, and R.D. Martin, Lecture Notes in Statistics No. 26, Springer Verlag, New York, pp 256-272.

Yohai, V.J. (1987), High breakdown point and high efficiency robust estimates for regression, *Ann. Math. Stat.,* **15**, 642-656.

Yohai, V.J., and Zamar, R.H. (1988), High breakdown point and high efficiency robust estimates for regression, *Ann. Math. Stat.,* **83**, 406-413.

Yohai, V.J., and Zamar, R.H. (1991), unpublished manuscript.

BIOGRAPHICAL INFORMATION

NAME: SONIA V. T. MAZZI

MAILING ADDRESS: SAMUEL BRETÓN 180
LOMAS DE SAN MARTÍN
5000 - CÓRDOBA
ARGENTINA

PLACE AND DATE OF BIRTH: CÓRDOBA, ARGENTINA
SEPTEMBER 24th, 1963

EDUCATION (Colleges and Universities attended, dates, and degrees):

UNIVERSIDAD NACIONAL DE CÓRDOBA - ARGENTINA
3/1983 - 3/1989
LICENCIADA EN MATEMÁTICA

POSITIONS HELD:

PUBLICATIONS (if necessary, use a second sheet):

AWARDS:

Complete one biographical form for each copy of a thesis presented to the Special Collections Division, University Library.