Maxbias of Robust Regression Estimates


In this session we will focus on the regression model where the maximum bias theory has two main possible applications:

- The comparison of competing robust regression estimates in terms of their bias behavior.
- The estimation of bias bounds for robust estimates in practical situations.

Given two estimates, the one with smaller maxbias curve is obviously more robust. In particular, its maxbias curve will have a smaller derivative at zero (called gross-error-sensitivity) and a larger asymptote (called breakdown point, BP). The comparison of maxbias functions naturally leads to the minimax bias theory also initiated by Huber’s seminal 1964 work. The minimax bias theory seeks estimates which minimize the maximum asymptotic bias in a certain class of estimates. Huber (1964, 1981) showed that the median minimizes the maxbias curve among translation equivariant location estimates. Martin et al. (1989) showed that the least median of square estimate (LMS) is nearly minimax in the class of regression M-estimates with general scale. Yohai and Zamar (1993) extend this result to the larger class of residual admissible estimates. It is precisely this large class of robust regression estimates which will be defined and further studied here.

Regarding the second application, we notice that, in general, estimates face two sources of uncertainty: sampling variability and bias. In the case of robust estimates, the sampling variability can be estimated by their standard errors (based on their empirical asymptotic variances). On the other hand, the bias caused by outliers and other departures from symmetry can be assessed using the concept of bias bounds introduced by Berrendero and Zamar (2001). Bias bounds - which we believe should be reported along with standard errors for any point estimate - can in principle be obtained starting from the maxbias curve. Bias bounds will be further described and developed in the next section.
Notation and Technical Background

Consider the linear regression model with $p$-dimensional regressors and intercept parameter

$$y_i = \alpha_0 + \beta_0'x_i + \sigma_0 u_i, \quad 1 \leq i \leq n$$

(1)

where the independent errors, $u_i$, have distribution $F_0$ and are independent of the $x_i$. We assume that the regressors $x_i$ are independent random vectors with common distribution $G_0$. The joint distribution of $(y_i, x_i)$ under this model is denoted $H_0$. It can be shown that

$$H(y, x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} F\left(\frac{y - \alpha_0 - \beta_0't}{\sigma_0}\right) dG_0(t)$$

(2)

To allow for a fraction $\epsilon$ of contamination in the data we assume that the actual true distribution $H$ of $(y_i, x_i)$ belongs to the contamination neighborhood

$$\mathcal{F}_\epsilon(H_0) = \{H : H = (1 - \epsilon)H_0 + \epsilon\tilde{H}, \quad \tilde{H} \text{ arbitrary distribution}\}.$$  

(3)

Let $T$ be an $R^p$ valued regression and affine-equivariant functional for the estimation of $\beta_0$, defined on a subset of distribution functions $H$ on $R^{p+1}$, which includes all $H$ in $\mathcal{F}_\epsilon(H_0)$ and all the empirical distributions $H_n$. A natural invariant measure of the robustness of $T$ is given by the worse case bias - maxbias function - which gives the maximum bias caused by a fraction $\epsilon$ of contamination,

$$B_T(\epsilon) = \sup_{H \in \mathcal{F}_\epsilon(H_0)} \left\{ [T(H) - \beta_0]'\Sigma_0[T(H) - \beta_0] / \sigma_0 \right\}^{1/2}$$

(4)

The matrix $\Sigma_0$ is some affine equivariant scatter matrix of the regressors under $G_0$, for instance

$$\Sigma_0 = \text{Cov}_{G_0}(\mathbf{x}).$$

(5)

In view of the equivariance of $T$ and the invariance of $B_T(\epsilon)$, we can assume without loss of generality that we have the canonical setup

$$\alpha_0 = 0, \quad \beta_0 = 0, \quad \sigma_0 = 1 \text{ and } \Sigma_0 = I,$$

2
and so

\[ B_T(\epsilon) = \sup_{H \in \mathcal{F}, (H_0)} \|T(H)\|. \]  

We can also assume without loss of generality that .

We will present a method to compute the maxbias curve for regression estimates which is very general. Although we do not find the maxbias of the intercept here (upper bounds for this bias can be found in Berrendero and Zamar 2001) we will not assume that the intercept parameter is known. Moreover, we will not require that the core distribution of the regressor, \( G_0 \), is elliptical, although the formulas will be much simpler in this case.

Residual Based Regression Estimates

The class of residual admissible estimates includes, among others, LMS, S—estimates, \( \tau \)—estimates, and R—estimates (Hössjer, 1994). Some examples of estimates which are not residual admissible are GM—estimates (see for instance Hampel et al., 1986), GS—estimates (Croux et al., 1994) and P—estimates (Maronna and Yohai, 1993).

Consider the absolute regression residual

\[ r(\alpha, \beta) = |y - \alpha - \beta'x|, \]  

and its distribution under \( H \),

\[ F_{H,\alpha,\beta}(r) = P_H(|y - \alpha - \beta'x| \leq r). \]  

A good regression fit should produce relatively small absolute residuals and so regression estimates can be defined with the goal of minimizing some functional of their distribution \( F_{H,\alpha,\beta} \). We will then consider the robust loss functional \( J(F) \) which depend on the size of the majority of the absolute residuals (e.g. their 50% or 75% percentiles).

Here are some examples of robust regression loss functional:
M-scale functional

\[ S(F) = \sup\{ s > 0 : \mathbb{E}_F \left\{ \chi\left( \frac{u}{s} \right) \right\} > b \} . \]  

(9)

\(\tau\)-scale functional

\[ \tau(F) = S^2(F) \mathbb{E}_F \left\{ \rho\left( \frac{u}{S(F)} \right) \right\} \]  

(10)

with \( S(F) \) given by (9).

\(R\)-scale functional

\[ R(F) = \int_{0}^{\infty} a(F(u)) u^k dF(u) , \]  

(11)

where \( a(t) \) is some appropriate weight function, e.g.

\[ a(t) = \begin{cases} 1 & \text{for } 0 \leq t \leq c \\ 0 & \text{for } t > c \end{cases} \]  

(12)

The residual admissible estimate (functional) is now defined by the minimization problem

\[ (T_0(H), T(H)) = \arg \min_{\alpha, \beta} J(F_H, \alpha, \beta) . \]  

(13)
The main result

The characterization of the maxbias of regression estimates will be done under the following assumptions.

A1 Assumptions on the robust loss functional:

(a) **Stochastic Monotonicity**: If $F$ and $G$ are two distribution functions on $[0, \infty)$ such that

\[ F(u) \leq G(u) \text{ for all } u \geq 0, \quad (14) \]

then

\[ J(F) \geq J(G). \quad (15) \]

Note: if two distribution functions $F$ and $G$ satisfy Equation (14) then we say that $G$ is stochastically smaller than $F$. If

\[ X \sim G, \ Y \sim F, \quad (16) \]

then we also say that $X$ is stochastically smaller than $Y$ and write

\[ X \prec Y. \quad (17) \]
(b) *Epsilon Monotonicity:* Given two sequences of distribution functions on $[0, \infty)$, $F_n$ and $G_n$, which are continuous on $(0, \infty)$ and such that

$$F_n(u) \to F'(u) \text{ and } G_n(u) \to G(u), \quad (18)$$

where $F$ and $G$ are two possibly sub-stochastic and continuous distributions on $(0, \infty)$, such that

$$G(\infty) \geq 1 - \epsilon \quad (19)$$

and

$$G(u) \geq F(u), \text{ for all } u > 0, \quad (20)$$

then

$$\lim_{n \to \infty} J(F_n) \geq \lim_{n \to \infty} J(G_n). \quad (21)$$

Moreover, if (20) holds strictly, then (21) also holds strictly.
If $F$ and $G$ are two distribution functions on $[0, \infty)$, with $F$ continuous, then

$$J[(1-\epsilon)F + \epsilon G] = \lim_{n \to \infty} J[(1-\epsilon)F + \epsilon U_n] \geq J[(1-\epsilon)F + \epsilon G],$$

where $U_n$ stands for the uniform distribution function on the interval $[n-(1/n), n+(1/n)]$.

**A2 Assumptions on the distributions of the regression errors and regressors.**

(a) $F_0$ has an even and strictly unimodal density $f_0$ with $f_0(u) > 0$ for every $u \in \mathbb{R}$.

(b) $P_{G_0}(\beta' x = c) < 1$, for each $\beta \neq 0$, $c \in \mathbb{R}$.

Assumption A1 (a) can be easily verified on given examples. Assumption A1 (b) was introduced by Yohai and Zamar (1993). When $J$ is $\epsilon$-monotone, the corresponding estimate belongs to the general class of residual admissible regression estimates (He, 1990). The definition of residual admissible estimate is rather involved but, loosely speaking, one can say that the distribution of the absolute residuals produced by a residual admissible estimate cannot be improved (in the sense of stochastic dominance) by using any other set of parameters. Finally, Assumption A1 (c) is a condition that specifies the form of the contaminations that cause the greatest loss.

Assumption A2 (a) is very mild and A2 (b) allows for great generality. In particular, it does not require ellipticity nor continuity of the regressors distribution.
The following theorem characterizes the maxbias for the slope coefficients for estimates defined by (13).

**Theorem 1** Suppose that A1 and A2 hold and let \( T \) be an estimate of the slope coefficients satisfying (13). Let

\[
c = J[(1 - \epsilon)F_{H_0,0} + \epsilon\delta_{\infty}]
\]

and

\[
m(t) = \inf_{\|t\| = t} \inf_{\alpha \in \mathbb{R}} J[(1 - \epsilon)F_{H_0,\alpha,t} + \epsilon\delta_0], \tag{23}
\]

where \( \delta_0 \) and \( \delta_{\infty} \) are point mass distributions at zero and infinity, respectively. Then, the maxbias curve for \( T \) is given by

\[
B_T(\epsilon) = m^{-1}(c). \tag{24}
\]
Some general properties of the contamination sensitivity and BP of residual admissible estimates with unknown intercept can now be obtained. These properties are stated in the following corollary.

**Corollary 1** Let $T$ be a regression estimate defined by (13). With the same notations and assumptions of Theorem ???:

(a) The slope of $B_T(\epsilon)$ at zero is infinity:

$$\lim_{\epsilon \to 0} B_T(\epsilon)/\epsilon = \infty$$

(b) The BP of $T$, $\epsilon^*$, is given by

$$\epsilon^* = \inf\{\epsilon > 0 : m(t) < c, \text{ for all } t > 0\}.$$  

Part (a) shows that the contamination sensitivity of robust residual based estimates is infinity. Yohai and Zamar (1997) showed that

$$B_T(\epsilon) = K \sqrt{\epsilon} + o(\sqrt{\epsilon})$$

Part (b) shows how the BP of robust residual based estimates can be characterized by the loss function $J$ through the function $m(t)$.

---

**Gaussian and elliptical regressors**

The following theorem shows that under strong symmetry assumptions on the regressors distribution the function $m(t)$, defined in (23), can be substantially simplified.
Theorem 2 Assume A1 and A2, and that the distribution of $t'x$ is symmetric, unimodal and only depends on $\|t\|$ for all $t' \neq 0$. Then,

$$\inf_{\alpha \in \mathbb{R}} J[(1 - \epsilon)F_{H_0, \alpha, t} + \epsilon \delta_0] = J[(1 - \epsilon)F_{H_0, 0, t} + \epsilon \delta_0] = m(\|t\|). \quad (25)$$

As a consequence the infima in (23) are no longer needed and the maxbias function $B_T(\epsilon)$ satisfies

$$m(B_T(\epsilon)) = J[(1 - \epsilon)F_{H_0, 0, 0} + \epsilon \delta_\infty]. \quad (26)$$

The symmetry assumption of Theorem 2 is clearly satisfied by Gaussian regressors, and more generally, by elliptically symmetric and unimodal regressors. Moreover, in the Gaussian case the distribution of $y - t'x$ is normal with mean 0 and variance $1 + \|t\|^2$ and the function $m(t)$ then becomes particularly simple.

We will use Theorem 2 to obtain the maxbias function for S, $\tau$ and R estimates.

Maxbias curves for S-estimates

Corollary 2 We will assume that the score function $\chi$ satisfies the following mild assumptions:

(a) The function $\chi$ is even, bounded, monotone on $[0, \infty)$, continuous at 0 with $0 = \chi(0) < \chi(\infty) = 1$; and

(b) The function $\chi$ has at most a finite number of discontinuities.

Under the assumptions of Theorem 2 and assuming that the errors and the regressors are Gaussian, the maxbias curves for S-estimates, $B_S(\epsilon)$, are given by

$$B_S(\epsilon) = \sqrt{\left[ g^{-1} \left( \frac{b - \epsilon}{1 - \epsilon} \right) \right]^2} - 1 ,$$

with

$$g(s) = E_{\Phi} \left\{ \chi \left( \frac{u}{s} \right) \right\} .$$
Proof:

It is easy to verify that in the case of S-estimates the functional $J$ is given by

$$J(F_{H,\alpha,t}) = S(F_{H,\alpha,t}),$$

with $S(F)$ given by (9). It can be easily checked that under these conditions
A1 (a) and (c) hold. Moreover, Yohai and Zamar (1993) showed that under
these conditions $S(F)$ is $\epsilon$-monotone for all $\epsilon > 0$. It is also clear that under
the Gaussian case A2 (a) and (b) hold.

In the Gaussian case, $(x,y)$ are jointly normal with mean $0 \in \mathbb{R}^{p+1}$ and co-
variance matrix $I$, where $I$ is the identity matrix of order $(p+1) \times (p+1)$. In
this case

$$F_{H_0,0,t} (u) = P (|y - x't| \leq u)$$

$$= P \left( |N \left( 0, 1 + \|t\|^2 \right) \left| \leq u \right. \right)$$

$$= 2\Phi \left( \frac{u}{\sqrt{1 + \|t\|^2}} \right) - 1. \quad (27)$$

To apply (26), let

$$g(s) = E_{\Phi} \left\{ \chi \left( \frac{u}{s} \right) \right\} \quad (28)$$

and notice that the scale functional

$$S[(1 - \epsilon)F_{H_0,0,0} + \epsilon\delta_{\infty}] = s$$

satisfies the equation

$$(1 - \epsilon)g(s) + \epsilon = b,$$
and therefore
\[
c = S[(1-\epsilon)F_{H_0,0,0} + \epsilon\delta_\infty]
\]
\[
= g^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right).
\] (29)

On the other hand, by (25) and (27), we can restrict attention to the case \(\alpha_0 = 0\) and, under \(H_0\), the residuals
\[
y - t'x
\]
are normally distributed with mean 0 and variance \(1 + ||t||^2\). Hence,
\[
S[(1-\epsilon)F_{H_0,t,0} + \epsilon\delta_0] = m(||t||)
\] (30)
satisfies the equation
\[
(1-\epsilon)g \left(\frac{m(||t||)}{\sqrt{1 + ||t||^2}}\right) = b.
\]
Therefore,
\[
m(||t||) = S[(1-\epsilon)F_{H_0,t,0} + \epsilon\delta_0]
\]
\[
= \sqrt{1 + ||t||^2} g^{-1}\left(\frac{b}{1-\epsilon}\right).
\] (31)

From (26), (29) and (31), the condition
\[
m(B_S(\epsilon)) = c
\]
implies
\[
\sqrt{1 + B_S(\epsilon)^2} g^{-1}\left(\frac{b}{1-\epsilon}\right) = g^{-1}\left(\frac{b-\epsilon}{1-\epsilon}\right),
\]
from which follows

\[
BS(\epsilon) = \sqrt{\left[ g^{-1}\left( \frac{b-\epsilon}{1-\epsilon} \right) \right]^2 \left[ g^{-1}\left( \frac{b}{1-\epsilon} \right) \right]^2 - 1.
\]

(32)

Maxbias curves for \(\tau\)-estimates.

The loss functional \(J\) in (13) for the case of \(\tau\)-estimates is given by

\[
J(F_{H,\alpha,t}) = \tau^2(F_{H,\alpha,t})
\]

with

\[
\tau^2(F) = S^2(F)E_{F}\left\{ \rho\left( \frac{u}{S(F)} \right) \right\},
\]

and \(S(F)\) is defined as in (9).

If \(\chi\) and \(\rho\) satisfy the conditions (a) and (b) above and \(\rho\) is differentiable with

\[
2\rho(u) - \rho'(u)u \geq 0,
\]

then A1 (a) and (c) hold. Moreover, Yohai and Zamar (1993) showed that under these conditions \(\tau(F)\) is \(\epsilon\)-monotone for all \(\epsilon > 0\).

Hössjer (1992) finds the most efficient regression S-estimate for a given breakdown point and notices that robust S-estimates have a very low efficiency. Yohai and Zamar (1988) showed that \(\tau\)-estimates inherit the BP of the initial S-estimate defined by \(\chi\) whereas its efficiency is mainly determined by \(\rho\). Therefore \(\tau\)-estimates can have high efficiency and high BP simultaneously. A natural question then is to what extent the \(\tau\)-estimate inherits the good bias behavior of the initial S-estimate. We can answer this question using Theorem 3 below. Notice that (33) establishes a nice relationship between the maxbias curves of the \(\tau\)-estimates and their initial S-estimates.
Corollary 3 Suppose that $\chi$ and $\rho$ satisfy the conditions (a) and (b) in Corollary 2 above. In addition suppose that
(c) $\rho$ is differentiable with

$$2\rho(u) - \rho'(u)u \geq 0.$$ 

Under the assumptions of Theorem 2 and assuming that the errors and the regressors are Gaussian, the maxbias curves for $\tau$-estimates, $B_\tau(\epsilon)$, are given by

$$B_\tau(\epsilon) = \{[1 + B_S^2(\epsilon)]H(\epsilon) - 1\}^{1/2},$$

(33)

where $B_S(\epsilon)$ is the maxbias curve of the initial S-estimate based on $\chi_1$ and $H(\epsilon)$ is defined as

$$H(\epsilon) = \left[ \frac{\Phi\left(\frac{b - \epsilon}{1 - \epsilon}\right) + \epsilon}{\Phi\left(\frac{b}{1 - \epsilon}\right)} \right],$$

with

$$\Phi(s) = g_2 [g_1^{-1}(s)]$$

and

$$g_1(s) = E_\Phi \left\{ \chi\left(\frac{u}{\tau}\right) \right\},$$

$$g_2(s) = E_\Phi \left\{ \rho\left(\frac{u}{\tau}\right) \right\}.$$
Problems

Problem 1 Prove Corollary 3 using the fact that A1 (a)-(c) hold under the Corollary’s assumptions. For a proof that A1 (b) holds see Yohai and Zamar (1993).

Problem 2 Use the “influence function approach” for deriving the asymptotic covariance matrix of τ-estimates under the Gaussian regression model (normal regression errors and jointly normal explanatory variables).

Problem 3 Use the “influence function approach” for deriving the asymptotic covariance matrix of S-estimates under the Gaussian regression model.

Problem 4 Use the result of Problem 3 to find the tuning constants for the 95% efficient S-estimate with Huber’s and Tukey’s loss functions. What is the efficiency of the Huber’s and Tukey’s S-estimates with breakdown point equal to 1/2? See Problem 6 for the definition of Huber’s and Tukey’s loss functions.

Problem 5 Use the result of Problem 2 to find the tuning constants for the 95% efficient τ-estimate with breakdown point equal to 1/2 for the following combinations of loss functions: Huber-Huber, Huber-Tukey, Tukey-Huber and Tukey-Tukey. See Problem 6 for the definition of Huber’s and Tukey’s loss functions.

Problem 6 Compute and plot the Gaussian maxbias curves for the S-estimates with Huber’s loss function

\[ \chi(y) = \min \left\{ \left( \frac{y}{c} \right)^2 , 1 \right\} \]

and Tukey’s loss function

\[ \chi_T(y) = \min \left\{ 3 \left( \frac{y}{c} \right)^2 - 3 \left( \frac{y}{c} \right)^4 + \left( \frac{y}{c} \right)^6 , 1 \right\} . \]

for several values of c, including the cases with breakdown point equal to 1/2 and with efficiency equal to 95%.

Problem 7 Consider τ-estimates with loss functions ρ and χ, where χ determines the breakdown point and ρ determines the efficiency. Derive the maxbias curves for 95% efficient τ-estimates and corresponding initial S-estimates for the following cases
(a) Huber’s loss functions

\[ \chi(y) = \min \left\{ \left( \frac{y}{c_1} \right)^2, 1 \right\} \]

\[ \rho(y) = \min \left\{ \left( \frac{y}{c_2} \right)^2, 1 \right\}. \]

(b) Tukey’s loss functions

\[ \chi(y) = \min \left\{ 3 \left( \frac{y}{c_1} \right)^2 - 3 \left( \frac{y}{c_1} \right)^4 + \left( \frac{y}{c_1} \right)^6, 1 \right\} \]

\[ \rho(y) = \min \left\{ 3 \left( \frac{y}{c_2} \right)^2 - 3 \left( \frac{y}{c_2} \right)^4 + \left( \frac{y}{c_2} \right)^6, 1 \right\}. \]

(c) Huber’s - Tukey’s loss functions

\[ \chi(y) = \min \left\{ \left( \frac{y}{c_1} \right)^2, 1 \right\} \]

\[ \rho(y) = \min \left\{ 3 \left( \frac{y}{c_2} \right)^2 - 3 \left( \frac{y}{c_2} \right)^4 + \left( \frac{y}{c_2} \right)^6, 1 \right\}. \]

(d) Tukey’s - Huber’s loss functions

\[ \chi(y) = \min \left\{ 3 \left( \frac{y}{c_1} \right)^2 - 3 \left( \frac{y}{c_1} \right)^4 + \left( \frac{y}{c_1} \right)^6, 1 \right\} \]

\[ \rho(y) = \min \left\{ \left( \frac{y}{c_2} \right)^2, 1 \right\}. \]

Questions: do the \( \tau \)-estimates inherit the bias robustness of the initial S-estimates? What of the four combinations above would you recommend?

Problem 8 Consider the “least trimmed of squares” estimate (LTS) defined by Rousseeuw (1984). Given the (tentative) absolute regression residuals

\[ r_i(a, b) = |y_i - a - b'x_i|, \quad i = 1, 2, \ldots, n \]
define the sorted squared residuals

\[ r^2_{(1)}(a,b) \leq r^2_{(2)}(a,b) \leq \cdots \leq r^2_{(n)}(a,b). \]

The LTS estimate is now defined as

\[ \left( \hat{\alpha}_n, \hat{\beta}_n \right) = \operatorname{arg\,min}_{a,b} \left[ n\delta \sum_{i=1}^{[n\delta]} r^2_i(a,b) \right] \]

where \([n\delta]\) is the integer part of \(n\delta\), that is,

\[ [n\delta] = \sup \{ k : k \leq n\delta, \ k \text{ is an integer} \}. \]

Use Theorem 2 to find the maxbias curve for \(\hat{\beta}_n\). Plot the curve for different values of \(\delta\) (e.g. \(\delta = 0.05, 0.10, \ldots, 0.5\)). What do you conclude?

**Problem 9** Define the sorted "power of absolute residuals"

\[ r^k_{(1)}(a,b) \leq r^k_{(2)}(a,b) \leq \cdots \leq r^k_{(n)}(a,b), \]

where \(k = 1, 2, \ldots\) The corresponding trimmed (power) residuals estimate is defined as

\[ \left( \hat{\alpha}^{(k)}_n, \hat{\beta}^{(k)}_n \right) = \operatorname{arg\,min}_{a,b} \left[ n\delta \sum_{i=1}^{[n\delta]} r^k_i(a,b) \right] \quad (34) \]

(a) Use Theorem 2 to find the maxbias curve for \(\hat{\beta}_n\). Plot maxbias curves for the case \(k = 1\) for different values of \(\delta\) (e.g. \(\delta = 0.05, 0.10, \ldots, 0.5\)). What do you conclude?

(b) Plot maxbias curves for the case \(\delta = 0.5\) for different values of \(k\) (e.g. \(k = 1, 2, 3, 4, 10\)). What do you conclude?

**Problem 10** Determine the "limiting case" of 34 when \(k \to \infty\).

**Problem 11** Suppose that \(Y \sim F\) and \(X \sim G\) are such that

\[ F \prec G, \quad (35) \]
where \( F \prec G \) means that \( F \) is stochastically smaller than \( G \) (see Equations (14), (16) and (17)).

(a) Explain why Equation (35) does not implies that \( X \leq Y \).
(b) Construct new random variables \( \tilde{X} \) and \( \tilde{Y} \) such that \( \tilde{Y} \sim F \) and \( \tilde{X} \sim G \) and

\[
\tilde{X} \leq \tilde{Y}.
\]

**Problem 12** Let \( \rho \) be an even and continuous loss function, non-decreasing on \([0, \infty)\) and that \( F \) and \( G \) are distribution functions supported on \([0, \infty)\) (that is, \( F \) and \( G \) distribution functions of non-negative random variables). If

\[
F \prec G,
\]

then

\[
E_G \{\rho(X)\} \leq E_F \{\rho(Y)\}.
\]

**Problem 13** Show that the asymptotic breakdown point of regression S-estimates with loss function \( \rho \) and

\[
b = E\Phi \{\rho(Z)\}
\]

is given by

\[
BP = \min \left\{ \frac{b}{\rho(\infty)}, 1 - \frac{b}{\rho(\infty)} \right\}.
\]
References


