Acta Universitatis Carolinae. Mathematica et Physica

J. Jurečková

Robust estimators and their relations

Acta Universitatis Carolinae. Mathematica et Physica, Vol. 24 (1983), No. 1, 49--59

Persistent URL: `http://dml.cz/dmlcz/142505`

1983 ACTA UNIVERSITATIS CAROLINAE - MATHEMATICA ET PHYSICA VOL. 24. NO. 1.

Robust Estimators and Their Relations

J. JUREČKOVÁ
Department of Statistics, Charles University, Prague*)

Received 22 December 1982

Asymptotic relations of *M*-estimators and *L*-estimators of regression parameter vector are considered. Special attention is devoted to the relation of Huber's *M*-estimator and of the trimmed mean in the location case and to the relation of Huber's *M*-estimator and of the trimmed least-squares estimator, recently suggested by Koenker and Bassett, in the regression case.

Uvažují se asymptotické vztahy *M*-odhadů a *L*-odhadů vektoru regresních parametrů. Zvláštní pozornost je věnována vztahu Huberova *M*-odhadu a useknutého průměru v případě parametru polohy a vztahu Huberova odhadu a useknutého odhadu metodou nejmenších čtverců, nedávno navrženého Koenkerem a Bassettem, v regresním modelu.

В статье исследованы асимптотические отношения *M*-оценок и *L*-оценок вектора параметров регрессии. Специальное внимание посвящено отношению оценки Хубера и усеченного среднего в случае параметра положения и отношению оценки Хубера и усеченной оценки методом наименьших квадратов, недавно предложенной Кэнкером и Бассеттом, в случае параметра регрессии.

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be independent random variables, X_i distributed according to the distribution function $F(x - \sum_{j=1}^{p} c_{ij}\Theta_j)$ with c_{ij} being given constants, $i = 1, \ldots, n$; $j = 1, \ldots, p$. The problem is that of estimating the parameter $\boldsymbol{\Theta} = (\Theta_1, \ldots, \Theta_p)'$, in the situation where F is generally unspecified; it is only assumed that F belongs to an appropriate family $\mathscr{F}$ of distribution functions (e.g., in the location model, $\mathscr{F}$ is the family of absolutely continuous symmetric distributions satisfying some general regularity conditions).

The classical least-squares estimator (and the sample mean as the special case) is closely connected with the normal distribution and is highly sensitive to the deviations from normality, to the heavy-tailed distributions and to the outlying observations. This fact is illustrated in a variety of Monte Carlo studies (e.g., Andrews et al. (1972)); moreover, Kagan, Linnik and Rao (1965) proved that, under some regularity

*) 186 00 Praha 8, Sokolovská 83, Czechoslovakia.

conditions, the sample mean is admissible with respect to the quadratic loss if and only if the basic distribution F is normal; the same proposition holds also for the least squares estimator (see Kagan, Linnik, Rao (1972)).

In the light of these facts, the statisticians tried to develop alternative procedures-good not only for one model but rather insensitive to the deviations from the assumed, model – and they started speaking about robust procedures. Stigler's historical studies (1973a, 1980, among others) show that yet the earlier statisticians were aware of the picture; some intuitive alternatives to the method of least squares appeared in the 19th century. Nontheless, a big development of robust procedures started only in the 20th century, mainly in the last 20 years.

Assume that F is an unknown member of a given family of distribution functions. We are looking for an estimation procedure, alternative to the sample mean or to the least-square procedure, respectively, the behaviour of which is not too poor whatever $F \in \mathscr{F}$ we may meet. The choice of the estimation procedure then, of course, depends on $\mathscr{F}$ which may be as large as the family of all symmetric absolutely continuous distribution functions, it may be neighbourhood of a fixed distribution or only a finite set of selected distribution shapes.

Among various types of robust estimators, three broad classes play the most important role: M-estimators (estimators of maximum-likelihood type; originated by Huber (1964)); R-estimators (estimators based on the rank tests; originated by Hodges and Lehmann (1963)) and L-estimators (estimators based on the ordered observations; e.g., the linear combinations of order statistics in the location model). All these estimators follow the same idea; to reduce the influence of the extreme values of observations and yet to estimate well the parameter Θ even if the assumption of normality happens to be right.

Jaeckel (1971) observed the close relations of these three types of estimators in the location model which appear when the number n of observations tends to infinity. He proved that, if the weight-functions generating the respective estimators are smooth and related in some special way, the difference of L- and M-estimators of location is of the stochastic order $O_p(n^{-1})$. Jaeckel also expressed a conjecture of a close asymptotic relation of R-estimators and M-estimators. This conjecture was later proved as true by Jurečková (1977) for the linear regression model and by Riedel (1979) for the location model.

We are interested not only in the asymptotic relations of various kinds of estimators, but also in the orders of these relations which indicated how precise is the approximation of an estimator by another one. While the problem of this order has not yet been completely solved, one fact appears in all cases: The orders of the asymptotic relations of various types of robust estimators are either $O_p(n^{-1})$ or $O_p(n^{-3/4})$ depending on whether the weight functions generating the estimators are smooth or whether they have jump-discontinuities. Let us mention some results illustrating this fact.

Either of three types of estimators is asymptotically normally distributed under various conditions. The asymptotic normality is usually proved by approximating the estimator by a sum of independent random variables. In this context, the author (1980) observed that, under some regularity conditions, the order of such approximation of an M-estimator is $O_p(n^{-1})$, provided the weight -function ψ (described in Section 2) is smooth while it is $O_p(n^{-3/4})$ only provided ψ has at least one jump-discontinuity; both orders are exact. This result was extended by Jurečková and Sen (1981, a, b) to the linear regression model.

It follows from Hušková and Jurečková (1981) that a similar fact appears also in the case of R-estimators, and moreover, if the weight-function ψ of an M-estimator M_n and the score-generating function φ of an R-estimator R_n are in the correspondence $\psi(x) = a\,\varphi(F(x)) + b,\ x \in R^1;\ a > 0,\ b \in R^1$, then $(M_n - R_n) = O_p(n^{-1})$ provided ψ (and φ) is smooth, while it is $O_p(n^{-3/4})$ only if ψ (and φ) has some jump-discontinuities.

An analogous thing may be observed in the case of L-estimators: If the L-estimator L_n is a linear combination of a finite set of single sample quantiles (so called systematic statistic) then there exists an M-estimator M_n (generated by a step-function ψ) such that $(L_n - M_n) = O_p(n^{-3/4})$; if L_n is a linear combination of order statistics generated by a smooth weight-function J then there exists an M-estimator M_n such that $(L_n - M_n) = O_p(n^{-1})$.

In the subsequent text, we shall describe some results on the orders of asymptotic relations (i.e., on the second order asymptotic relations) of various robust estimators in more details. Our attention will be devoted to the relations of M-estimators and L-estimators, in the location as well as in the regression case. We shall consider, among others, the relation of the famous pair of Huber's estimator and of the trimmed mean in the location case and that of Huber's estimator and of trimmed least-squares estimator, recently introduced by Koenker and Bassett (1978), in the regression case. We shall also touch the systematic statistics, their extension to the linear model and their M-estimator counterparts.

2. M- and L-estimation alternatives to the method of least squares

Let $X_1, \ldots, X_n$ be independent random variables, X_i distributed according to the distribution function (d.f.) $F(x - \sum_{j=1}^{p} c_{ij}\Theta_j)'$, $i = 1, \ldots, n$; $\boldsymbol{\Theta} = (\Theta_1, \ldots, \Theta_p)'$ is an unknown parameter and $\boldsymbol{c}_j = (c_{1j}, \ldots, c_{nj})$, $j = 1, \ldots, p$ are known vectors. F is generally unspecified; it is only assumed that it belongs to an appropriate family $\mathscr{F}$ of distributions.

The classical method of estimating $\boldsymbol{\Theta}$ is that of minimizing the sum of squares

$$\sum_{i=1}^{n} (X_i - \sum_{j=1}^{p} c_{ij} t_j)^2 := \min \tag{2.1}$$

or, equivalently, solving the system of equations

$$\sum_{i=1}^{n} c_{ij}(X_i - \sum_{k=1}^{p} c_{ik}t_k) = 0\,, \quad j = 1, \ldots, p \tag{2.2}$$

with respect to $\boldsymbol{t} = (t_1, \ldots, t_p)'$.

We get the M-estimator of $\boldsymbol{\Theta}$ if we minimize, instead of (2.1),

$$\sum_{i=1}^{n} \varrho(X_i - \sum_{j=1}^{p} c_{ij}t_j) := \min \tag{2.3}$$

where ϱ is some (usually convex) function, less sensitive to the extreme values of the argument. Differentiating (2.3), we obtain (with $\psi = \varrho'$) the system of equations equivalent to (2.3):

$$\sum_{i=1}^{n} c_{ij}\, \psi(X_i - \sum_{k=1}^{p} c_{ik}t_k) = 0\,, \quad j = 1, \ldots, p\,. \tag{2.4}$$

The M-estimator $\boldsymbol{M}_n = (M_{n1}, \ldots, M_{np})'$ of $\boldsymbol{\Theta}$ is defined as the root of the system (2.4); if it is not uniquely determined, then it may be defined as the root of (2.4) nearest to some initial consistent estimator $\boldsymbol{T}_n$ of $\boldsymbol{\Theta}$ ($\boldsymbol{M}_n = \boldsymbol{0}$ if there is no root).

The class of M-estimators was originated by Huber (1964) (see also Huber (1972, 1973, 1977); the properties of M-estimators are thoroughly studied in Huber's recent monograph (1981)).

L-estimators are defined in a straightforward way in the location model.

Denote $X_{n:1} \leqq \ldots \leqq X_{n:n}$ the order statistics corresponding to $X_1, \ldots, X_n$, where X_i's are independent, identically distributed according to d.f. $F(x - \Theta)$ such that $F(x) + F(-x) = 1$, $x \in R^1$. The L-estimator L_n of Θ is of the form

$$L_n = \sum_{i=1}^{n} d_{in} X_{n:i} \tag{2.5}$$

where $d_{1n}, \ldots, d_{nn}$ are given constants, usually such that $d_{in} = d_{n-i+1,n} \geqq 0$, $i = 1, \ldots, n$, $\sum_{i=1}^{n} d_{in} = 1$. Simple examples are the sample mean $\bar{X}_n$ and the sample median $\tilde{X}_n$. If we wish to get a robust L-estimator, insensitive to the extreme observations, we must put $d_{in} = 0$ for $i \leqq k_n$ and $i \geqq n - k_n + 1$ for a proper k_n. Typical examples of such estimators are the α-trimmed mean,

$$L_n = \frac{1}{n - 2[\alpha n]} \sum_{i=[n\alpha]+1}^{n-[n\alpha]} X_{n:i} \tag{2.6}$$

where $0 < \alpha < \frac{1}{2}$ and $[x]$ denotes the largest integer k satisfying $k \leqq x$; and the α-Winsorized mean,

$$L_n = \frac{1}{n} \{[n\alpha]\, X_{n:[n\alpha]} + \sum_{i=[n\alpha]+1}^{n-[n\alpha]} X_{n:i} + [n\alpha]\, X_{n:n-[n\alpha]+1}\}\,. \tag{2.7}$$

A broad class of L-estimators (2.5) which usually appear in applications is of the form

$$L_n = \frac{1}{n} \sum_{i=1}^{n} J\left(\frac{i}{n+1}\right) X_{n:i} + \sum_{j=1}^{k} a_j X_{n:[np_j]} \tag{2.8}$$

where $J(u)$, $0 < u < 1$ is a proper smooth weight function, usually satisfying $J(u) = J(1-u) \geqq 0$, $0 < u < 1$, and $p_1, \ldots, p_k$; $a_1, \ldots, a_k$ are given constants, usually satisfying $0 < p_1 < \ldots < p_k < 1$, $p_j = 1 - p_{k-j+1}$, $a_j = a_{k-j+1} \geqq 0$, $j = 1, \ldots, k$. L_n is then of the form (2.5) with d_{in} given by $n^{-1} J(i/(n+1))$ plus an additional contribution a_j if $i = [np_j]$ for some ($j = 1, \ldots, k$). The L-estimator (2.8) is thus a combination of two special L-estimators; in many cases, the estimator under consideration is just of one type.

The L-estimators are computionally appealing and have further convenient properties in the location model (cf. Bickel and Lehmann (1975)). However, they do not extend to the linear model in straightforward way. A possible regression analog of L-estimators was suggested and studied by Bickel (1973). His estimators, defined in the two-step way with the aid of an initial estimator, have good efficiency properties but they are computionally complex.

Recently Koenker and Bassett (1978) extended the concept of quantiles to the linear model in the following way: for $\alpha \in (0, 1)$, denote

$$\varphi_\alpha(x) = \begin{cases} \alpha & \text{if } x \geqq 0 \\ \alpha - 1 & \text{if } x < 0 \end{cases} \tag{2.9}$$

and

$$\varrho_\alpha(x) = x\, \varphi_\alpha(x)\,, \quad x \in R^1\,. \tag{2.10}$$

The α-th regression quantile $\hat{\boldsymbol{\Theta}}(\alpha)$ is defined as any value of $\boldsymbol{t} = (t_1, t_2, \ldots, t_p)'$ which solves

$$\sum_{i=1}^{n} \varrho_\alpha(X_i - \sum_{j=1}^{p} c_{ij} t_j) := \min. \tag{2.11}$$

Notice that $\hat{\boldsymbol{\Theta}}(\alpha)$ is, in fact, an M-estimator. Koenker and Bassett (1978) then proposed the α-trimmed least-squares estimator $\boldsymbol{L}_n(\alpha)$, $0 < \alpha < \frac{1}{2}$, in the following way: Assume that there is a rule which selects a unique solution of (2.11). $\boldsymbol{L}_n(\alpha)$ is defined as the least-squares estimator calculated after removing all observations satisfying

$$X_i - \sum_{j=1}^{p} c_{ij}\, \hat{\Theta}_j(\alpha) < 0 \quad \text{or} \quad X_i - \sum_{j=1}^{p} c_{ij}\, \hat{\Theta}_j(1-\alpha) > 0\,, \tag{2.12}$$

$i = 1, \ldots, n$. Let $a_i = 0$ or 1 according as i satisfies (2.12) or not and denote $\boldsymbol{A}_n$ the diagonal matrix with diagonal $(a_1, \ldots, a_n)$. Then

$$\boldsymbol{L}_n(\alpha) = (\boldsymbol{C}_n' \boldsymbol{A}_n \boldsymbol{C}_n)^- (\boldsymbol{C}_n' \boldsymbol{A}_n \boldsymbol{X}_n) \tag{2.13}$$

where

(2.14) $$C_n = (c_{ij})_{i=1,\dots,n}^{j=1,\dots,p}$$

and $(C_n' A_n C_n)^-$ is a generalized inverse of $(C_n' A_n C_n)$.

The asymptotic behavior of $L_n(\alpha)$ was studied by Ruppert and Carroll (1978), (1980), who proved that $L_n(\alpha)$ is asymptotically normally distributed,

(2.15) $$n^{1/2}(L_n(\alpha) - \Theta) \xrightarrow{\mathscr{D}} N_p(0, \sigma^2(\alpha, F) \,.\, Q^{-1}), \quad \text{as} \quad n \to \infty,$$

with $\sigma^2(\alpha, F)$ being the asymptotic variance of the α-trimmed mean and $Q = \lim_{n\to\infty} (1/n)\, C_n' C_n$.

3. Huber's estimator and the trimmed mean in the location case

Let $X_1, X_2, \dots$ be independent random variables, identically distributed according to the d.f. $F(x - \Theta)$; assume that F is absolutely continuous and symmetric, i.e. $F(x) + F(-x) = 1$, $x \in R^1$.

Let $L_n(\alpha)$ denote the α-trimmed mean, defined in (2.6). Let M_n be Huber's M-estimator of Θ, which could be written as

(3.1) $$M_n = \tfrac{1}{2}(M_n^+ + M_n^-)$$

with

(3.2) $$M_n^+ = \inf \{t : \sum_{i=1}^{n} \psi(X_i - t) < 0\}$$

$$M_n^- = \sup \{t : \sum_{i=1}^{n} \psi(X_i - t) > 0\}$$

and

(3.3) $$\psi(x) = \begin{cases} x & \text{if } |x| \leqq c \\ c \,.\, \operatorname{sign} x & \text{if } |x| > c\,; \end{cases}$$

$c > 0$ is constant.

There has been some confusion in the history of the problem of asymptotic relations of $L_n(\alpha)$ and M_n. One might intuitively expect that the Winsorized mean rather than the trimmed mean resembles Huber's estimator (cf. Huber (1964)). Bickel (1965) was apparently the first who recognized the close connection betwenn Huber's estimator and the trimmed mean. Jackel (1971) studied the asymptotic relations of M- and L-estimators but, in fact, his theorem does not cover this special case.

The following theorem shows that, with a proper choice of the constant c in (3.3), the difference of both estimators is of the order $O_p(n^{-1})$ as $n \to \infty$.

Theorem 3.1. *Let $X_1, X_2, \ldots$ be independent random variables, identically distributed according to the* d.f. $F(x - \Theta)$ *such that* $F(x) + F(-x) = 1$, $x \in R^1$ *and which satisfies the following conditions*:

(i) *F has an absolutely continuous density f and finite Fisher's information*, i.e. $\mathscr{I}(f) = \int (f'/f)^2 \, dF < \infty$.

(ii) *$f(x) > a > 0$ for all x such that*

$$\alpha - \varepsilon \leqq F(x) \leqq 1 - \alpha + \varepsilon, \quad 0 < \alpha < \tfrac{1}{2}, \quad \varepsilon > 0.$$

(iii) *f' exists in the interval* $(F^{-1}(\alpha - \varepsilon), F^{-1}(1 - \alpha + \varepsilon))$.

Then

$$L_n(\alpha) - M_n = O_p(n^{-1}), \quad \text{as} \quad n \to \infty, \tag{3.4}$$

where $L_n(\alpha)$ is the α-trimmed mean (2.6) *and M_n is the M-estimator defined in* (3.1)–(3.3) *with* $c = F^{-1}(1 - \alpha)$.

Proof. The theorem was proved in Jurečková (1983).

4. Huber's estimator and the trimmed least squares estimator in the regression case

Let $X_{n1}, \ldots, X_{nn}$ be independent observations, X_{ni} distributed according to the distribution function $F(x - \sum_{j=1}^{p} c_{ij}\Theta_j)$ such that $F(-x) + F(x) = 1$, $x \in R^1$ and F is absolutely continuous with the density f. Let $\boldsymbol{M}_n = (M_{n1}, \ldots, M_{np})'$ be Huber's estimator of $\boldsymbol{\Theta}$, defined as a solution of (2.4) with ψ given in (3.3); let $\boldsymbol{L}_n(\alpha)$ be the α-trimmed least-squares estimator defined in (2.9)–(2.14); furthermore, we shall assume that the design matrix satisfies the following regularity conditions (A):

$$c_{i1} = 1, \quad i = 1, \ldots, n; \quad \sum_{i=1}^{n} c_{ij} = 0, \quad j = 2, \ldots, p; \tag{A}$$

$$\lim_{n \to \infty} \frac{1}{n} \boldsymbol{C}_n' \boldsymbol{C}_n = \boldsymbol{Q}, \quad \boldsymbol{Q} \text{ is a positive matrix};$$

$$\max_{1 \leqq j \leqq p} n^{-1} \sum_{i=1}^{n} |c_{ij}|^3 = O(1), \quad \max_{1 \leqq j \leqq p} n^{-1} \sum_{i=1}^{n} |c_{ij}|^4 = O(1).$$

Theorem 4.1. *Let $X_{n1}, \ldots, X_{nn}$, $n = 1, 2, \ldots$ be the triangular array of independent observations with X_{ni} distributed according to the* d.f. $F(x - \sum_{j=1}^{p} c_{ij}\Theta_j)$, $i = 1, \ldots, n$. *Then, provided $\boldsymbol{C}_n$ satisfies the conditions* (A) *and F satisfies the conditions of Theorem* 3.1, (4.1) $$\|L_n(\alpha) - M_n\| = O_p(n^{-3/4}), \quad \text{as} \quad n \to \infty,$$

where $L_n(\alpha)$ *is the* α*-trimmed least-squares estimator and* M_n *is Huber's estimator generated by* ψ *of* (3.3) *with* $c = F^{-1}(1 - \alpha)$; $0 < \alpha < \frac{1}{2}$.

Proof. The proof is based on a refinement of that of Theorem 3 of Ruppert and Carroll (1978) and on Corollary 3.2 of Jurečková and Sen (1981b). See Jurečková (1983) for details.

Remark. The order (4.1) differs from that proved in the location case and the proof does not indicate that it is the best possible; but the author surmises that it is the case due to the fact that $L_n(\alpha)$ is, unlike in the location case, a two-step estimator.

5. Systematic statistics and their M-estimator conterparts

Let us turn back to the location model. Let L_n be the systematic statistic defined as

(5.1) $$L_n = \sum_{j=1}^{k} a_j X_{n:[np_j]}$$

where $0 < p_1 < \ldots < p_k < 1$; $a_1, \ldots, a_k$ are given constants; $p_j = 1 - p_{k-j+1}$, $a_j = a_{k-j+1} > 0$, $j = 1, \ldots, k$; $\sum_{j=1}^{k} a_j = 1$. The following theorem shows that, under some mild conditions on F, there exists an M-estimator M_n such that $L_n - M_n = O_p(n^{-3/4})$.

Theorem 5.1. *Let* $X_1, X_2, \ldots$ *be independent observations, identically distributed according to the* d.f. $F(x - \Theta)$ *such that* $F(x) + F(-x) = 1$, $x \in R^1$, *the inverze* F^{-1} *is strictly increasing at* $p_1, \ldots, p_k$ *and* F *has two bounded derivatives* f, f' *in neighbourhoods of* $F^{-1}(p_1), \ldots, F^{-1}(p_k)$. *Then*

(5.2) $$L_n - M_n = O_p(n^{-3/4})$$

holds for the L-estimator L_n *of* (5.1) *and for the M-estimator* M_n *generated by the function*

(5.3) $$\psi(x) = \sum_{j=1}^{k} \frac{a_j}{f(F^{-1}(p_j))} (I[F(x) \leq p_j] - p_j), \quad x \in R^1;$$

$I[A]$ *is the indicator of the set* A.

Proof. See Jurečková (1982).

Remark. It follows from (5.3) that $\psi(x)$ is a nondecreasing step-function with the jumps at the points $F^{-1}(p_1), \ldots, F^{-1}(p_k)$. The theorem, as well as other analogous theorems, does not enable to evaluate numerically the asymptotically equivalent L-estimator, once we have calculated the M-estimator, because the corresponding L-estimator depends on given ψ through the unknown distribution F. The results rather indicate the close relations of both types of estimators; namely, (5.3) shows that the linear combinations of single sample quantiles are asymptotically close to M-estimators generated by the step-functions.

6. Systematic statistics and M-estimators in linear model

Let $X_{n1}, \ldots, X_{nn}$ be independent, X_{ni} distributed according to the distribution function $F(x - \sum_{j=1}^{p} c_{ij}\Theta_j)$, F symmetric. Let $0 < p_1 < \ldots < p_k < 1$ be given numbers such that $p_{k-m+1} = 1 - p_m$, $m = 1, \ldots, k$ and assume F has two bounded derivatives in neighborhoods of $F^{-1}(p_1), \ldots, F^{-1}(p_k)$. Denote $\hat{\Theta}(p_m)$, $m = 1, \ldots, k$, the p_m-regression quantile defined in (2.9)–(2.11). Then, we may consider an analogue of the systematic statistics (5.1) defined as

$$L_n = \sum_{m=1}^{k} a_m \hat{\Theta}(p_m) \tag{6.1}$$

where $a_m = a_{k-m+1} > 0$, $m = 1, \ldots, k$; $\sum_{m=1}^{k} a_m = 1$. Such statistics were studied by Koenker and Bassett (1978) who also derived their asymptotic distribution. We shall show that the function ψ of (5.3) generates the M-estimator M_n of Θ which is asymptotically equivalent to L_n with the order $O_p(n^{-3/4})$. This is expressed in the following theorem.

Theorem 6.1. *Let C_n satisfy the assumption* (A) *of Section* 4. *Then, under the above assumptions on F, p_m and a_m, $m = 1, \ldots, k$,*

$$\|L_n - M_n\| = O_p(n^{-3/4}) \tag{6.2}$$

holds for the L-estimator L_n of (6.1) *and for the M-estimator M_n of Θ generated by the step-function* (5.3).

Proof. The theorem follows by combining the representation of regression quantiles given in Theorem 2 of Ruppert and Carroll (1978), and the representations of M-estimators (and thus of regression quantiles) given in Theorem 3.3 of Jurečková (1980) and in Corollary 3.2 of Jurečková and Sen (1981b).

References

[1] Adichie, J. N. (1967): Estimate of regression parameters based on rank test. *Ann. Math. Statist.* 38, 894—904.

[2] ANDREWS, D. F., BICKEL, P. J., HAMPEL, F. R., HUBER, P. J., ROGERS, W. H., TUKEY, J. W. (1972): Robust Estimates of Location: Survey and Advances. *Princeton Univ. Press.*

[3] BICKEL, P. J. (1965): On some robust estimates of location. *Ann. Math. Statist.* 36, 847—858.

[4] BICKEL, P. J. (1973): On some analogous to linear combinations of order statistics in the linear model. *Ann. Statist.* 1, 597—616.

[5] BICKEL, P. J. and LEHMANN, E. L. (1975): Descriptive statistics for nonparametric model. II. Location. *Ann. Statist.* 3, 1045—1069.

[6] HODGES, J. L. and LEHMANN, E. L. (1963): Estimates of location based on rank tests. *Ann. Math. Statist.* 34, 598—611.

[7] HUBER, P. J. (1964): Robust estimation of a location parameter. *Ann. Math. Statist.* 35, 73—101.

[8] HUBER, P. J. (1972): Robust statistics: A review. *Ann. Math. Statist.* 43, 1041—1067.

[9] HUBER, P. J. (1973): Robust regression: Asymptotics, conjectures and Monte Carlo. *Ann. Statist.* 1, 799—821.

[10] HUBER, P. J. (1977): Robust Statistical Procedures. SIAM, Philadelphia.

[11] HUBER, P. J. (1981): Robust Statistics. *J. Wiley,* N. York.

[12] HUŠKOVÁ, M. (1982): On bounded length sequential confidence interval for parameter in regression model based on ranks. Coll. Math. Soc. J. Bolyai 32 (Proc. Intern. Colloquium on Nonparametrics, Budapest 1980), pp. 435—463.

[13] HUŠKOVÁ, M. and JUREČKOVÁ J. (1981): Second order asymptotic relations of M-estimators and R-estimators in two-sample location model. *Journ. Statist. Planning and Inference* 5, 309—328.

[14] JAECKEL, L. A. (1971): Robust estimates of location: Symmetry and asymmetric contamination. *Ann. Math. Statist.* 42, 1020—1034.

[15] JAECKEL, L. A. (1972): Estimating regression coefficients by minimizing the dispersion of the residuals. *Ann. Math. Statist.* 43, 1449—1458.

[16] JUREČKOVÁ, J. (1971): Nonparametric estimates of regression coefficients. *Ann. Math. Statist.* 42, 1328—1338.

[17] JUREČKOVÁ, J. (1977): Asymptotic relations of M-estimates and R-estimates in linear regression model. *Ann. Statist.* 5, 464—472.

[18] JUREČKOVÁ, J. (1980): Asymptotic representation of M-estimators of location. *Math. Operationsforsch. Statist.,* Ser. Statistics 11, 61—73.

[19] JUREČKOVÁ, J. (1982): Robust estimators of location and their second order asymptotic relations. To appear in *ISI Centennial Volume* (A. Atkinson, ed.).

[20] JUREČKOVÁ, J. (1983): Robust estimators of location and regression parameters and their second order asymptotic relations. *Trans. 9th Prague Conf. on Inform. Theory, Statist. Decision Functions and Random Processes, pp.* 19—32.

[21] JUREČKOVÁ, J. and SEN, P. K. (1981a): Invariance principles for some stochastic processes relating to M-estimators and their role in sequential statistical inference. *Sankhya* A 43, 190—210.

[22] JUREČKOVÁ, J. and Sen, P. K. (1981b): Sequential procedures based on M-estimators with discontinuous score functions. *Journ. Statist. Planing and Inference* 5, 253—266.

[23] KAGAN, A. M., LINNIK, JU. V. and RAO, C. R. (1965): On a characterization of the normal law based on a property of the sample average *Sankhya* A 27, 405—406.

[24] KAGAN, A. M., LINNIK, JU. V. and RAO, C. R. (1972): Characteristic Problems of Mathematical Statistics. (in Russian). *Nauka,* Moscow.

[25] KOENKER, R. and BASSETT, G. (1978): Regression quantiles. *Econometrica* 46, 33—50.

[26] KOUL, H. L. (1971): Asymptotic behavior of a class of confidence regions based on ranks in regression. *Ann. Math. Statist.* 42, 466—476.

[27] PORTNOY, S. L. (1977): Robust estimation in dependent situations. *Ann. Statist.* 5, 22—43.

[28] Riedl, M. (1979): *M*-estimators of regression and location (in Czech). Unpublished Thesis. *Charles University,* Prague.

[29] Ruppert, D. and Carroll, R. J. (1978): Robust regression by trimmed least-squares estimation. *Univ. of North Carolina,* Inst. of Statist. Mimeo Series No 1186.

[30] Ruppert, D. and Carroll, R. J. (1980): Trimmed least-squares estimation in the linear model. *J. Amer. Statist. Assoc.* 75, 828—838.

[31] Stigler, S. M. (1973a). Simon Newcomb, Percy Daniell, and the history of robust estimation. 1885—1920. *J. Amer. Statist. Assoc.* 68, 872—879.

[32] Stigler, S. M. (1973b): The asymptotic distribution of the trimmed mean. *Ann. Statist.* 1, 472—477.

[33] Stigler, S. M. (1974): Linear functions of order statistics with smooth weight functions. *Ann. Statist.* 2, 676—693.

[34] Stigler, S. M. (1980): Studies in the history of probability and statistics XXXVIII. R. H. Smith, a Victorian interested in robustness. *Biometrika* 67, 217—221.