GEOPHYSICAL TRANSACTIONS 1993
Vol. 38. No. 4. pp. 193- 210

PRACTICAL DEFINITION OF ROBUSTNESS

Ferenc STEINER* and Béla HAJAGOS*

The paper defines the index of robustness (*r*) as a weighted average efficiency belonging to a statistical estimating procedure. The weights are the occurrence probability densities of the various model types which can be accepted as adequate for a given discipline. The value *r* can simultaneously take very different probability distribution types into consideration. Instead of deciding categorically 'robust' — 'not robust' the examples show robustnesses in the interval from r=36 % to r=96 %. In geophysics practice quantitative comparisons are unavoidable.

Some of the figures demonstrate the original efficiency curves ($e(t)$-s), figuring in Eq. 12 given for *r*, too, thereby enabling so that the changes in the efficiencies can be analysed in detail.

Keywords: robustness, index of robustness, statistical efficiency, probability density, error distribution

1. Introduction and preliminaries

The definition of robustness by theoretical experts of mathematical statistics [see e.g. HAMPEL et al. 1986] does not result in numerical values (thereby facilitating the near-optimum choice of the statistical algorithm,) and/or it belongs to very narrow (or even infinitesimal) neighbourhood of a distribution type. Let one comment be cited from the Summary of the article of DONOHO and LIU [1988], i.e., from a paper written by mathematicians: 'Of course, this robustness is formal because μ-contamination neighbourhoods may not be large enough to contain *realistic departures from the model*' (enhancement was not made in the original text). Here we propose the acceptance of a measure of robustness which is also suitable for practical applications. The discipline of geophysics particularly needs quantitative comparisons made on the grounds of large type-intervals.

* University of Miskolc, Department of Geophysics, H-3515 Miskolc-Egyetemváros
Manuscript received: 20 July, 1993

1.1. Various estimations of the location parameter (a brief enumeration)

A chronological enumeration of different statistical procedures is given below with some comments. In every case below the task is to determine (estimate) on the grounds of a given sample the most characteristic value of the actual probability distribution (this is naturally the symmetry point if the distribution is symmetrical). — In the first and second case it is impossible to determine how old these estimations are (at least two hundred years old):

arithmetic mean		
sample median		
α-trimmed mean	1821	see e.g. FEGYVERNEKI [1992] — but may be as old as the arithmetic mean itself
Hodges-Lehmann estimate	1963	
Huber estimate	1964	
M^*-estimate	1965	this is the minimum place of the P^*-norm, see Eq. 36 in HAJAGOS and STEINER [1991]
M-estimate	1973	this is the minimum place of the P-norm, see Eq. 30 in HAJAGOS and STEINER [1991]
L_p-estimate ($p>0$, $p\neq1$, $p\neq2$)	1990	this is the minimum place of the generalized L_p-norm, see e.g. TARANTOLA [1987] (it is well known that for $p=1$ we would get the sample median and for $p=2$ the arithmetic mean). The date of L_p is given here in accordance with SOMOGYI and ZÁVOTI [1990], as the authors do not know any earlier article in applied statistics that deals in detail with a p value which is not an integer.

Where no explanation is given or no reference is cited, see e.g. the monograph HUBER [1981] or the original papers HODGES, LEHMANN [1963] and HUBER [1964] (in the present paper 'Proposal 2' of HUBER is treated). It should be mentioned that both M^*- and M-estimates are called '*m*ost *f*requent *v*alue' therefore in the case of more unknown parameters the corresponding statistical algorithm is called '*MFV* procedure' (and the simple estimate can also be called '*MFV*-value' instead of M- or M^*-estimate). Some characteristics of the M-estimate are given in a comprehensive manner in the Table at the end of the book STEINER (ed.) [1991]; in the bibliography of this book are cited the paper and thesis where M- and M^*-estimates were first defined.

1. 2. How to calculate the efficiencies

If certain conditions for the density function are fulfilled and the sample range (n) tends to infinity, the distribution type of the estimates becomes

Gaussian (see e.g. HUBER [1981]; the overwhelming majority of the following can also be found in the same monograph). This means that the dispersion can adequately be characterized by the variance ($VAR = \sigma^2$) of the estimates. To be independent of n, it is convenient to introduce the notion 'asymptotic variance' (A^2) with the equation

$$A^2 = \lim_{n \to \infty} n.\sigma^2. \tag{1}$$

It is often easy to find statistical algorithm that leads to the minimum asymptotic variance ($A^2_{\min}$) for the probability distribution in question.

The efficiency (e) of an arbitrary statistical algorithm having an asymptotic variance A^2 for a well defined probability distribution, is defined as

$$e = \frac{A^2_{\min}}{A^2}, \tag{2}$$

(where $A^2_{\min}$ obviously belongs to the same probability distribution). Often e is expressed in per cent.

Eq. 2 says that e per cent of the data would be sufficient for the same estimation accuracy if we were to use an optimum algorithm instead of the one actually used. In practice therefore, from the viewpoint of the cost of measurements it is of crucial importance that the statistical efficiency e is as great as possible.

How does one calculate the asymptotic variance A^2? If the so-called influence function $IC(x)$ is known for the statistical algorithm and for the actual probability distribution defined by the density function $f(x)$, A^2 can be determined as

$$A^2 = \int_{-\infty}^{\infty} IC^2(x) \,.\, f(x)\, \mathrm{d}x. \tag{3}$$

If primarily the $\psi(x)$-function is given (the ψ-function plays a key-role in the best elaborated part of the robust statistics), the influence function can be calculated as

$$IC(x) = \psi(x) \,.\, \left[\int_{-\infty}^{\infty} \psi'(y) \,.\, f(y)\, \mathrm{d}y \right]^{-1}. \tag{4}$$

In some cases A^2 can be calculated directly by means of a simple formula. *Table I* gives either A^2-formulas, or IC-, or ψ-functions (always choosing the simplest alternative) for the statistical procedures yet enumerated in 1.1. (for probability distributions symmetrical to the origin). The asymptotic variance A^2 can be calculated in every case without difficulty.

1. 3. The supermodel $f_a(x)$

The supermodel $f_a(x)$ was introduced by the density functions

$$f_a(x) = \Gamma\left(\frac{a}{2}\right)\cdot\Gamma^{-1}\left(\frac{a-1}{2}\right)\cdot\pi^{-\frac{1}{2}}\cdot(1+x^2)^{-a/2}\quad (a>1) \tag{5}$$

[see e.g. CSERNYÁK, STEINER 1991]; this standard form can be generalized by replacing x by $(x-T)/S$ and dividing by S (T and S are the parameter of location and parameter of scale, respectively). Here, we mention some types of this supermodel: the distribution type a=5 is called geostatistical or simply statistical having clearly the density function

$$f_{st}(x) = 0.75(1+x^2)^{-2.5}\ . \tag{6}$$

(according to DUTTER [1987] this is a very commonly occurring distribution type in geostatistics, but in the opinion of the authors its acceptance as a model is justified more generally in the practice of statistics). If short flanks are guaranteed, the so-called Jeffreys-type (a=9) can serve as an adequate model for the distribution:

statistical procedure (estimate)	characterization of the procedure from the view-point of the asymptotic variance of the estimates
arithmetic mean	$IC(x) = x$, i.e., $A^2 = VAR = \sigma^2$ (VAR means the variance, σ the scatter of the mother distribution)
sample median	$A^2 = \dfrac{1}{4\cdot f^2(0)}$
α-trimmed mean	$IC(x) = \begin{cases} \frac{1}{1-2\alpha}F^{-1}(\alpha), & \text{if } x < F^{-1}(\alpha) \\ x/(1-2\alpha), & \text{if } \lvert x\rvert \le F^{-1}(1-\alpha) \\ \frac{1}{1-2\alpha}F^{-1}(1-\alpha) & \text{if } x > F^{-1}(1-\alpha) \end{cases}$
Hodges-Lehmann estimate	$A^2 = \dfrac{1}{12\left\{\int_{-\infty}^{\infty}[f(x)]^2\,dx\right\}^2}$

Huber-estimate	$A^2 = \dfrac{\int_0^{cS} x^2 f(x)\mathrm{d}x + (cS)^2 \int_{cS}^{\infty} f(x)\mathrm{d}x}{2\left(\int_0^{cS} f(x)\mathrm{d}x\right)^2}$; the value S fulfils the condition $\frac{1}{S^2}\int_0^{cS} x^2 f(x)\mathrm{d}x + c^2 \int_{cS}^{\infty} f(x)\mathrm{d}x = \int_0^{c} f_G(x)\,\mathrm{d}x + c^2 \int_c^{\infty} f_G(x)\,\mathrm{d}x$; ($f_G(x)$ represents the Gaussian density function)
M^*-estimate	$\psi_{M^*} = \dfrac{x}{[3\,(k\varepsilon)^2 + x^2]^2}$
most frequent values	The dihesion ε fulfils in both cases the condition $\int_{-\infty}^{\infty} \dfrac{3x^2 - \varepsilon^2}{[\varepsilon^2 + x^2]^2} f(x)\,\mathrm{d}x = 0$
M-estimate	$\psi_M = \dfrac{x}{(k\varepsilon)^2 + x^2}$
L_p-estimate	$\psi_p(x) = sign\ x \,.\, \lvert x\rvert^{p-1}$

Table I. Charaterization of some statistical procedures
I. táblázat. Statisztikai eljárások jellemzése

$$f_J(x) = \frac{35}{32}(1+x^2)^{-4.5} \quad . \tag{7}$$

It can easily be shown that for the supermodel $f_a(x)$ the minimum asymptotic variance is given by the simple formula

$$A^2_{\min} = \frac{a+2}{a(a-1)} \quad . \tag{8}$$

For integer values of a we get Student distribution types characterized by $(a-1)$ degrees of freedom; the so-called Jeffreys interval of distribution types defined by $6 \le a \le 10$ was primarily given also by limits expressed as 5 and 9 degrees of freedom. Obviously $(a-3)^{-1/2} \cdot f_a\,[x \,.\, (a-3)^{-1/2}]$ tends to the standard Gaussian density function $f_G(x) = (2\pi)^{-1/2} . \exp(-x^2/2)$ if $a \to \infty$. For a=2 we trivially get the Cauchy distribution.

The probability density functions of the Cauchy-, (geo)statistical-, Jeffreys- and Gaussian type are shown in *Fig. 1*; in all four cases the probable error (i.e., the semi-interquartile range q) equals unity (choosing the parameter of scale S always appropriately). We find these curves visually very similar — although statistical procedures can behave very differently if the actually occurring error distribution type is, say, geostatistical instead of Gaussian. Some statistical procedures (first of all the classical ones) are extremely sensitive to the behaviour of the flanks but Fig. 1 (and other such commonly used visualizations, too) does not characterize these parts of the distributions very well (the small values of $f(x)$ at both ends of the $f(x)$-curve can result in misjudging the weight of the flanks measured in the occurrence probability of x of the neglected sides). The authors therefore prefer the plotting of the density function versus $F(x)$-curve since this does not depend upon the parameter of scale and, moreover, it enhances the behaviour of the tails (as usually,

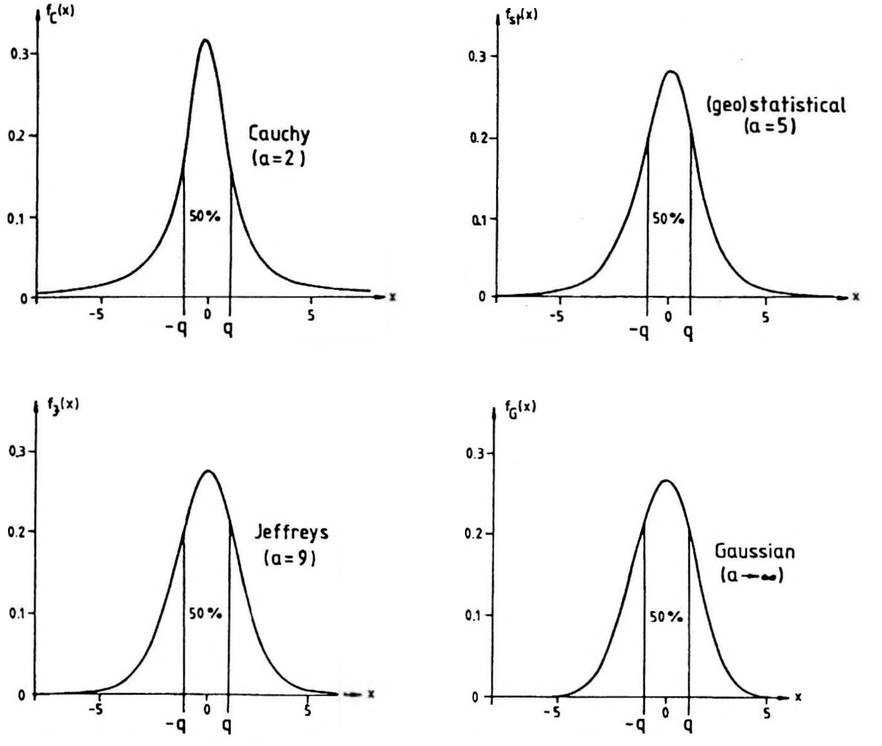

Fig. 1. Four probability density functions of x from the supermodel $f_a(x)$ (see Eqs. 5-7). With appropriately chosen parameter of scale the probable error (semi-interquartile range) q equals unity in every case

1. ábra. Az $f_a(x)$ szupermodell négy valószínűségsűrűség-függvénye (ld. az 5-7 egyenleteket). A skálaparaméter megfelelő választásával a q valószínű hiba (azaz az interkvartilis félterjedelem) egységnyi nagyságú mind a négy esetben

$F(x) = \int_{-\infty}^{x} f(x)\,dx$ represents the distribution function). It is advantageous to 'norm' the densities to their maximum value; this was done in *Fig. 2.* where the great difference between the flanks and the general features of the Cauchy-, (geo)statistical and the Gaussian type are visualized. (For Laplace- and uniform distributions the $f(x)/f_{max}$ versus $F(x)$-'curves' consist of straight lines, see the dashed lines in Fig. 2.) It should be mentioned, too, that Fig. 2 clearly shows: that there are distribution types that are characterized by much heavier flanks, than those of the Cauchy-type.

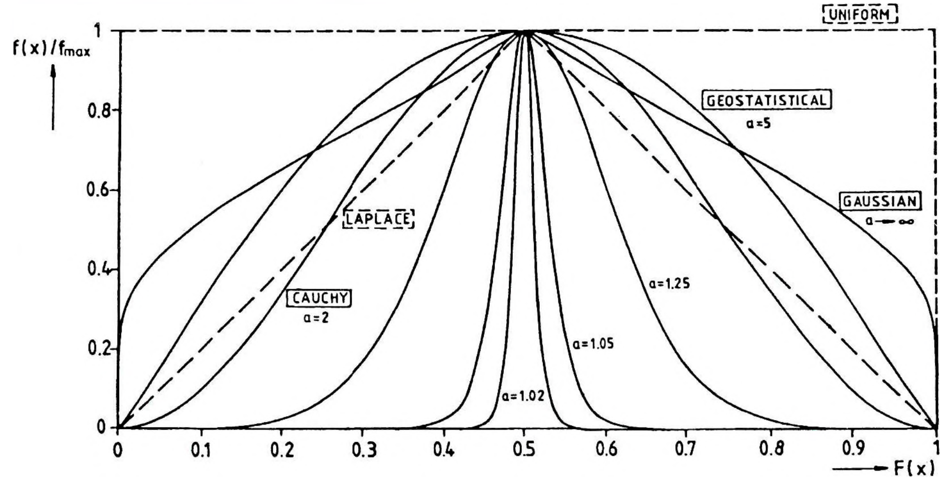

Fig. 2. Probability densities (normed to their maximum value) versus distribution function $F(x)$ as a visualization which is independent both of the parameter of scale and the parameter of location. The different behaviour of the flanks is satisfactorily accentuated here

2. ábra. Maximális értékükre normált valószínűségsűrűségek az $F(x)$ eloszlásfüggvény értékeinek a függvényében. Ily módon mind a hely-, mind a skálaparamétertől független görbéket nyerünk, amelyek jól láthatóan fejezik ki az eloszlások szárnyainak különböző viselkedését

2. Quantitative characterization of robustness

2. 1. Inherent supposition of the maximum likelihood-principle from the practical viewpoint. Occurrence probability densities ($f_J(t), f_D(t)$) of type t distribution.

Statistical procedures can be derived on the basis of the *m*aximum *l*ikelihood-principle (but these procedures are usually applied *not only* for the distribution type which was supposed in the first step). The ML-principle originally *postulates that the type of the actual distribution is a priori known* (with probability = 1). Good Heavens! Indeed, the statistician working in a practical environment *never* a priori knows the type of the actual probability distribution exactly.

Let us suppose, however, just for a moment, that this supposition is fulfilled and this a priori known type is the Jeffreys distribution (see Eq. 7). It is easy to verify that the maximum likelihood method results in the calculation of the M-estimate with k=3. This latter value is a slightly rounded one consequently the efficiency is not exactly 100 % but 'only' 99.9999 %. Obviously the practical statistician would tolerate perhaps a 'loss' of say, 2-3 %, too (and a loss of 1% would certainly be accepted as insignificant even by the most rigorous mathematician).

The question arises if other estimation procedures can approximate the maximum efficiency or not. *Fig. 3* shows the efficiencies of the L_p-estimates versus p for the Jeffreys distribution; if p=1.6 is chosen the efficiency is greater than 98 %. It can be demonstrated in a similar way that the Huber estimate has maximum efficiency for the Jeffreys distribution if c=1.4 is chosen. Briefly, the efficiencies of six estimating procedures (to an accuracy of two decimals) are summarized in *Table II*.

statistical procedure	efficiency for the Jeffreys distribution
M-estimate; k=3	100.00%
M^*-estimate; k=3	99.87%
Hodges-Lehmann estimate (H.L.)	99.86%
Huber; c=1.4	99.60%
α-trimmed mean ($\bar{x}_\alpha$); α=0.1	99.54%
L_p-estimate; p=1.6	98.19%

Table II. Efficiencies of various statistical procedures if the errors are Jeffreys-distributed
II. táblázat. Statisztikai eljárások Jeffreys-eloszlásra vonatkozó hatásfokai

From the practical viewpoint, all six procedures turned out to be equally good if the samples come from the Jeffreys distribution. It should be emphasized that the first five estimates show efficiencies even greater than 99.5 %.

Introducing $t = (a-1)^{-1}$ as the type parameter, the assumption of the maximum likelihood-principle says nothing less than that the density function of the occurrence probabilities of various $f_a(x)$ types is

$$f_{ML}(t) = \delta(t-0.125) \tag{9}$$

(δ means Dirac-δ). For practical purposes, this is unacceptable. We can require at least that the occurrence probability density must be maximum for the type

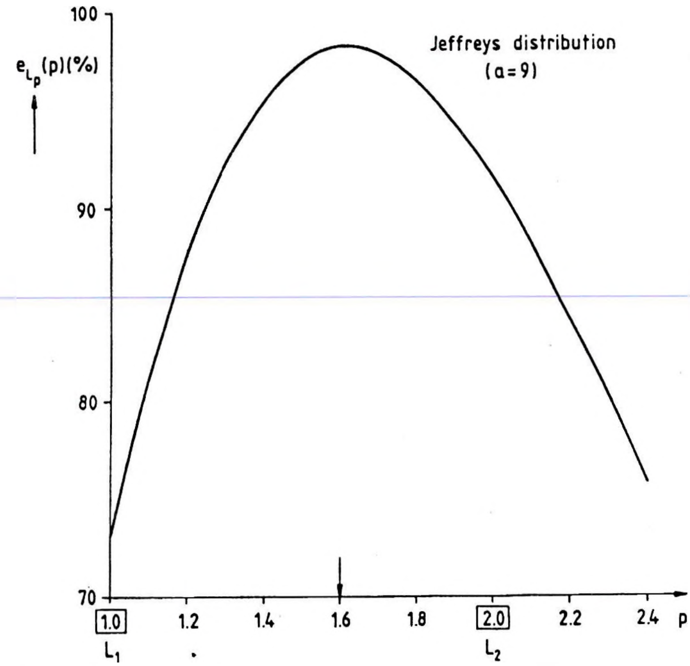

Fig. 3. Efficiency curve for different L_p-estimates for the Jeffreys distribution (see Eq. 7)
3. ábra. Különböző L_p-becslések hatásfokai a Jeffreys-eloszlásra vonatkozóan (ld. a 7 formulát)

t=0.125 (and not significantly less for the neighbouring types). If outliers seldom occur then one per cent probability density of the maximum value should be enough for the Cauchy-type to model somehow such situations, too; and finally we require $f(0)=0$ (see SZŰCS 1993 and references therein). Consequently, instead of Eq. 9 it is not only convenient but also justifiable to accept

$$f_J(t) = 64.t.e^{-8t.}, \qquad (10)$$

the letter J in the index refers to the fact that $f_J(t)$ has its maximum position at t=0.125, i.e., at the Jeffreys distribution.

(A comment seems to be appropriate here: although $f(0)=0$ holds — in agreement with the modern statistical literature — the following zero hypothesis: 'the error distribution is Gaussian' is generally accepted at the commonly used significance levels even if Eq. 10 characterizes the occurrence probabilities of each type-interval, see SZŰCS [1993].)

The so-called Jeffreys interval of probability distribution types around t=0.125 shows the shortest flanks which can realistically be hoped for in nature. For example, in geostatistics, it can be stated [after DUTTER 1987] that we can accept as the most common type an $f_a(x)$ with a=5, i.e., with t=0.25. On the other hand, STEINER (ed.) [1991] shows examples proving that in the geosciences the Cauchy-type really occurs, i.e., the probability density of the types can not be a negligible value around t=1 compared with the maximum one.

These conditions are fulfilled (and $f(0) = 0$ also) if we accept as a probability density function for the distribution type t:

$$f_D(t) = 16.t.e^{-4t} \tag{11}$$

(compare Eq. 12 in STEINER 1991). Generally speaking, it is of crucial importance that we must at least be approximately imformed about the probability densities of the types of supermodel which can be accepted for adequate modelling of the error distributions occurring in a given discipline. It is the duty of the expert of the discipline in question to give an acceptable density function formula for the types which are able to model the actual error distributions in his territory of science or application. Both $f_D(t)$ and $f_J(t)$ curves are visualized in *Fig. 4.*

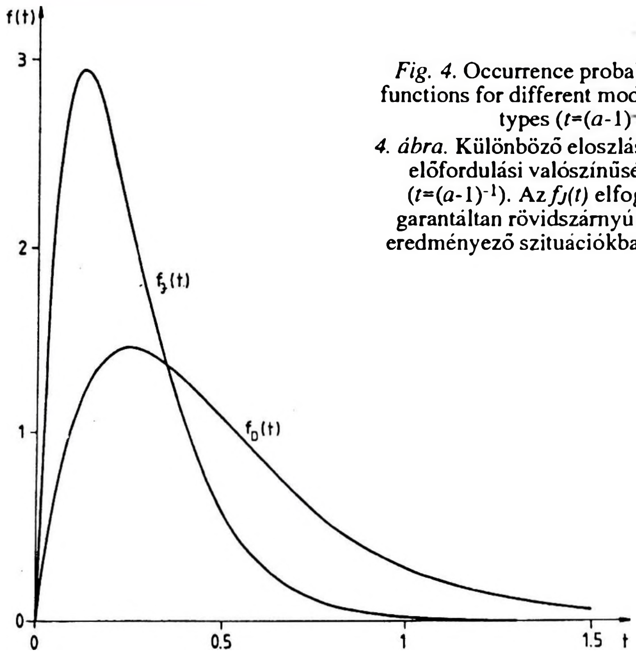

Fig. 4. Occurrence probability density functions for different model distribution types (t=(a-1)$^{-1}$)
4. ábra. Különböző eloszlástípusmodellek előfordulási valószínűségsűrűségei (t=(a-1)$^{-1}$). Az $f_J(t)$ elfogadása csak garantáltan rövidszárnyú eloszlásokat eredményező szituációkban javasolható

2.2. *Efficiency curves to visually demonstrate the different robustnesses of various statistical procedures*

One can find, in the literature of robust statistics, statements of the form: 'procedure A is robust, procedure B is not robust'. By the authors opinion such categorical distinctions are hard to justify — to say nothing about the contradiction that BOX [1953] introduced the notion 'robustness' for a method of conventional statistics (based on the L_2-norm) which letter is quite uniformly classified as 'not robust' by robust statistics (in the last three decades).

The efficiency curves versus t are shown in *Figs. 5–8* for all six statistical procedures figuring in Table II (in Figs. 7 and 8 the $e(t)$-curve for the median is also given). The speed of the decrease of e is different for increasing t from the nearly equal maximum value: it is most rapid for L_p p=1.6; at $t \geq 0.8333$ even e=0 holds. (It is easy to demonstrate also for the general case that $e > 0$ can hold only if $t < (2p-2)^{-1}$.) It is curious that two pairs of estimates behave similarly (M and M^* both for k=3; Huber c=1.4 and $\bar{x}_\alpha$ $\alpha = 0.1$; see Figs. 6 and 7) though the definitions of the corresponding statistical procedures are different.

Qualitatively the order concerning the robustness of the six procedures seems to be the following: L_p p=1.6; $\bar{x}_\alpha$ $\alpha = 0.1$ and Huber c=1.4; Hodges-Lehmann estimate; M and M^* both for k=3. The interesting behaviour of the latter $e(t)$-curves is that for $t \to \infty$ $(a \to 1)$ the efficiency seems to tend to an asymptotic value of 33-34% (see Fig. 8); Fig. 2 shows that these distributions have extremely heavy flanks. In *Figs. 9* and *10* also for k=2 the efficiency curves are shown both for M and M^*; the corresponding asymptotic values here are 48 and 50%, respectively. It should be mentioned that k=2 is accepted as the 'standard version' of the most frequent value (*MFV*-) calculations, in full agreement with the fact that maximum efficiencies are to be obtained very near to t=0.25 (i.e., to a=5) where $f_D(t)$ reaches its maximum (see Eq. 11).

The asymptotic behaviour of the $e(t)$ curves is a hint that *MFV*-procedures are not only robust to a high degree but are also extremely outlier-resistant. The two notions robustness and resistance, must be distinguished although there exists some interconnection between them. The oft occurring opinion, however, that robustness = outlier-resistance, is misleading and unacceptable.

2. 3. Average efficiencies as adequate indices of robustness in practice

Definition. Let us take the probability density function $\varphi(t; x)$ for t-values in the inverval $T_1 \leq t \leq T_2$ and let it be supposed that the probability density function of the type parameter t (i.e., $f(t)$) is also given. The index of the robustness of an estimation procedure according to $f(t)$ is defined as

$$r = \int_{T_1}^{T_2} e(t).f(t)\mathrm{d}t \qquad (12)$$

where $e(t)$ is the efficiency of the estimation procedure in question if the data are distributed according to $\varphi(t; x)$.

Comment 1. The existence of $e(t)$ anticipates the existence of the Fisher-information of $\varphi(t; x)$ to the fixed value t, on the one hand and, on the other, it

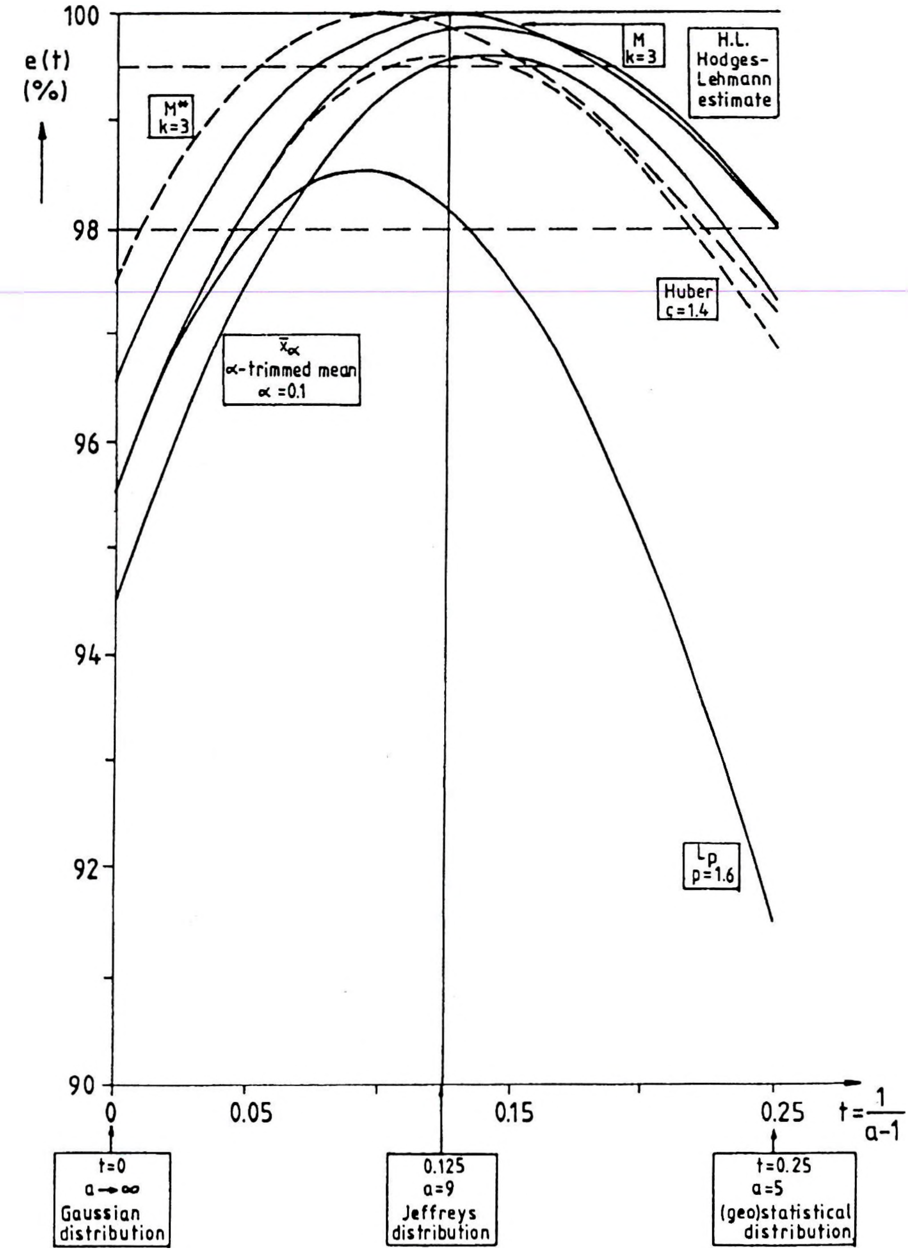

Fig. 5. Efficiency curves for six estimating procedures in the type interval $0 \leq t \leq 0.25$

5. ábra. Hatásfokgörbék hat becslési eljárásra a $0 \leq t \leq 25$ típustartományban

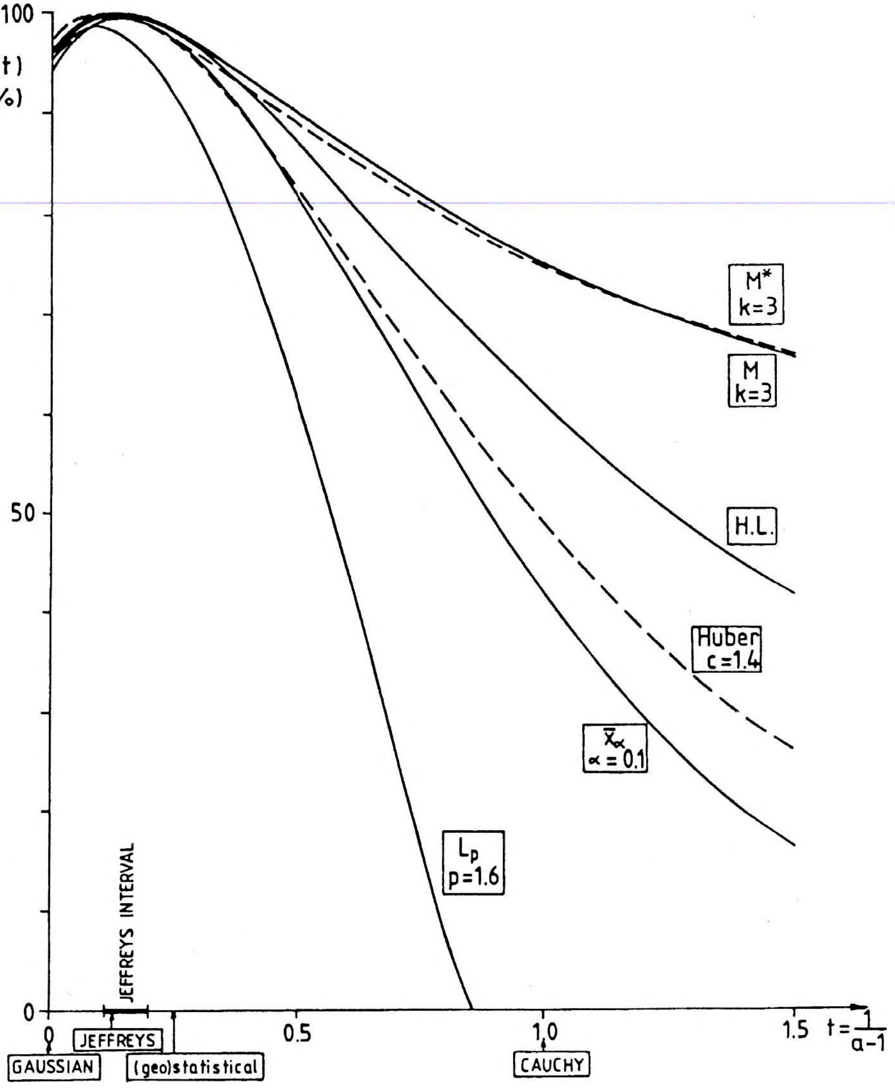

Fig. 6. Efficiency curves for six estimating procedures in the type interval $0 \leq t \leq 1.5$

6. ábra. Hatásfokgörbék hat becslési eljárásra a $0 \leq t \leq 1,5$ típustartományban

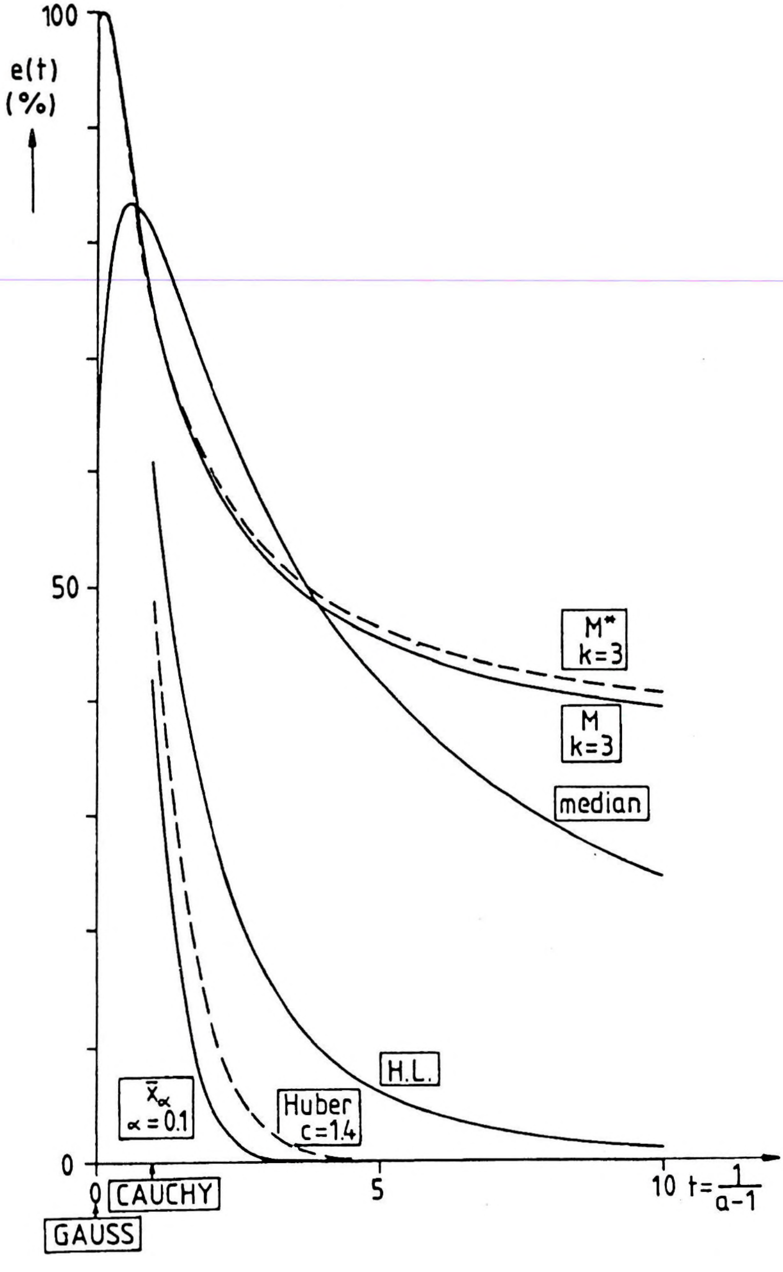

Fig. 7. Efficiency curves for six estimating procedures in the type interval $0 \leq t \leq 10$

7. ábra. Hatásfokgörbék hat becslési eljárásra a $0 \leq t \leq 10$ típustartományban

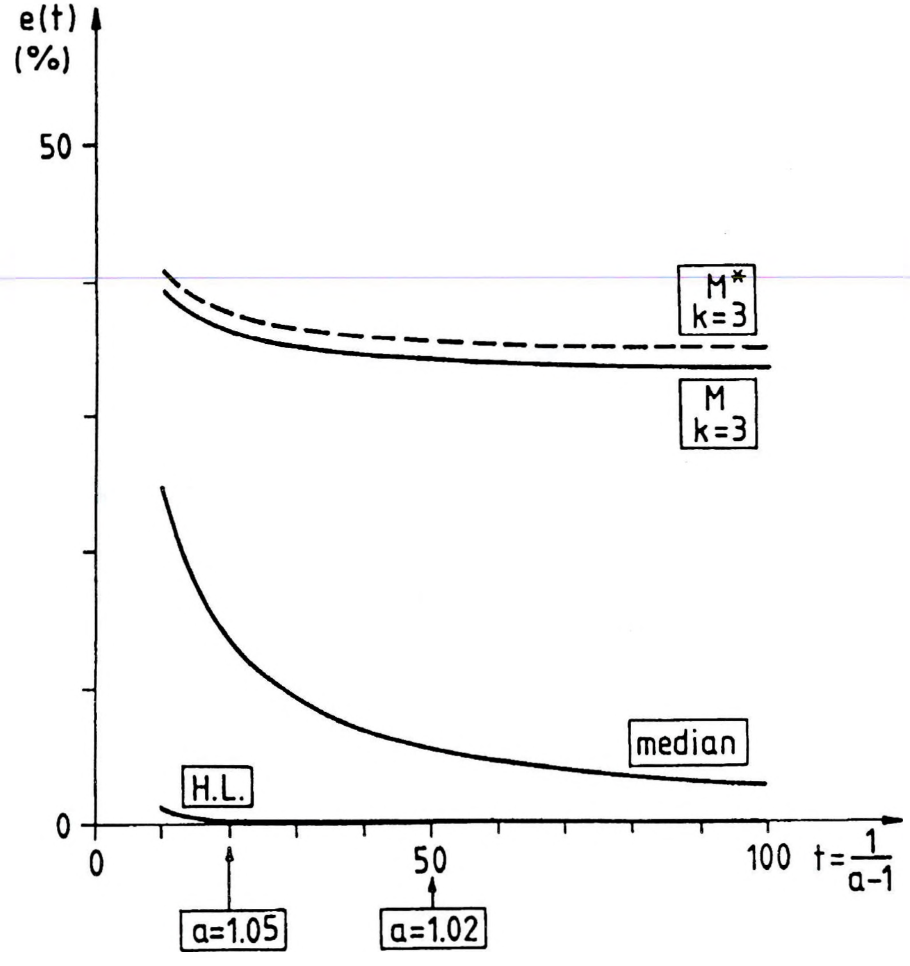

Fig. 8. Efficiency curves for four estimating procedures in the type interval $0 \leq t \leq 100$
8. ábra. Hatásfokgörbék négy becslési eljárásra a $0 \leq t \leq 100$ típustartományban

also anticipates the existence of the asymptotic variance of the estimates if the data are distributed according to $\varphi(t;x)$ ($T_1 < t < T_2$).

Comment 2. It is the task of the expert of a discipline (and not the task of the mathematician) to define a function $f(t)$ which can be accepted as an adequate one for the discipline in question. The choice $f(t)=f_D(t)$ (see Eq. 11) seems to be an adequate one in the geosciences (but the authors of the present paper suppose that this choice may be all right in other territories of statistics, too). The choice $f(t)=f_J(t)$ (see Eq. 10) seems to be a 'quasi-classical' one as

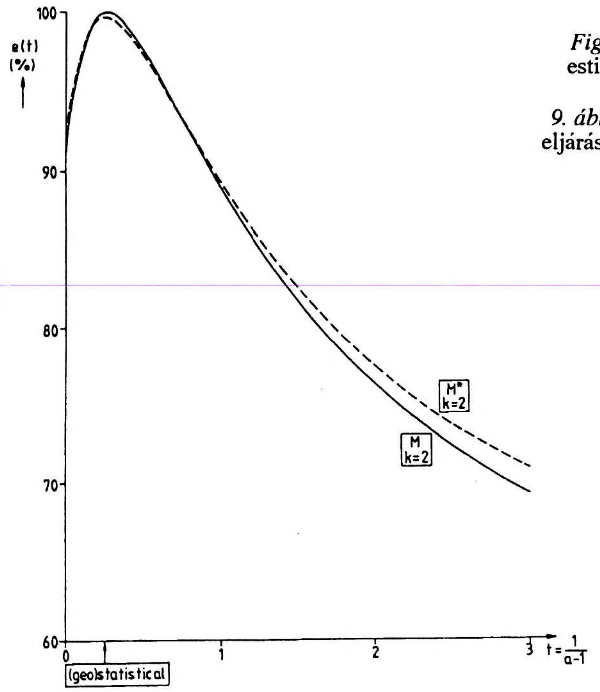

Fig. 9. Efficiency curves for two estimating procedures in the type interval $0 \leq t \leq 3$

9. ábra. Hatásfokgörbék két becslési eljárásra a $0 \leq t \leq 3$ típustartományban

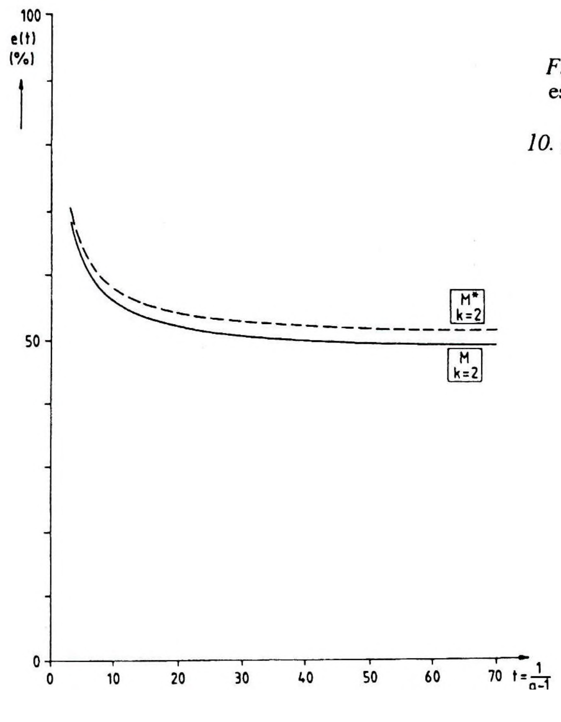

Fig. 10. Efficiency curves for two estimating procedures in the type interval $0 \leq t \leq 70$

10. ábra. Hatásfokgörbék két becslési eljárásra a $0 \leq t \leq 70$ típustartományban.

the tails of the distributions in the overwhelming majority of the cases are very short.

Comment 3. The definition of *r* given in Eq. 12 based on a supermodel $\varphi(t;x)$, i.e., for a case of only one type parameter, can be trivially generalized if more than one type parameter exist in the supermodel used.

In *Table III.* for ten statistical estimating procedures the indices of robustness are given (in per cent), calculated for both $f(t)=f_J(t)$ and $f(t)=f_D(t)$; the ordering was made according to the latter one.

statistical estimate		index of robustness (r) concerning the supermodel $f_a(x)$ if the occurrence probability of the various error distribution types are characterized by the density function	
name	symbol	$f_J(t)$ (Eq. 10)	$f_D(t)$ (Eq. 11)
arithmetic mean	$\bar{x}$ (L_p; $p = 2$)	67%	**36%**
	L_p; p=1.6	85%	**60%**
α-trimmed mean	$\bar{x}_\alpha$; $\alpha = 0.1$	93%	**79%**
sample median	med (L_p; p=1)	77%	**80%**
Huber-estimate (Proposal 2)	Huber; c=1.4	94%	**81%**
Hudges-Lehmann estimate	H. L.	96%	**85%**
most frequent value (*MFV*)	M^*; k=3	96%	**89%**
	M; k=3	97%	**90%**
	M^*; k=2	98%	**96%**
	M; k=2	98%	**96%**

Table III. Indices of robustness for various statistical estimates
III. táblázat. A robusztusság mérőszámai különböző statisztikai becsléseknél

REFERENCES

BOX G. E. P. 1953: Non normality and test on variances. Biometrika **40**, pp. 318–335

CSERNYÁK L., STEINER F. 1991: Monte Carlo studies on the rate of fulfillment of the law of large numbers. *In:* The Most Frequent Value. Introduction to a modern conception of statistics. (ed. F. STEINER) Chapter XII. pp. 173–185, Akadémiai Kiadó, Budapest

DONOHO D. L., LIU R. C. 1988: The 'automatic' robustness of minimum distance functionals. The Annals of Statistics **16**, No. 2 pp. 552–586

DUTTER R. 1987: Mathematische Methoden in der Montangeologie. Vorlesungsnotizen 1986/87. Leoben (Manuscript), 187 p

FEGYVERNEKI S. 1992: Robust estimates according to the model distribution and their applications (in Hungarian). Thesis. Manuscript. (University of Miskolc), 70 p

HAJAGOS B., STEINER F. 1991: The P*-norm. Acta Geod. Geoph. Mont. Acad. Sci. Hung. **26**, (1-4) pp. 153-182

HAMPEL F. R., RONCHETTI E. M., ROUSSEEUW P. J., STAHEL W. A. 1986: Robust Statistics. Wiley, New York, 502 p

HODGES J. L., Jr. LEHMANN E. 1963: Estimates of location on rank tests. Ann. Math. Statist. **34**, pp. 518-611

HUBER P. J. 1964: Robust estimation of a location parameter. Ann. Math. Statist. **35**, pp. 73-101

HUBER P. J. 1981: Robust Statistics. Wiley, New York, 308 p.

SOMOGYI J., ZÁVOTI J. 1990: Die Anwendung der L_p-Norm-Schätzung für Ähnlichkeitstransformationen. Zeitschrift für Vermessungswesen **115**, (1) pp. 28-36

STEINER F. 1991: Average efficiency of statistical procedures. Acta Geod. Geoph. Mont. Acad. Sci. Hung. **26**, (1-4) pp. 135-151

STEINER F. (ed.); 1991: The Most Frequent Value. Akadémiai Kiadó, Budapest (Hungary), 315 p

SZŰCS P. 1993: Comment on an old dogma: 'The data are normally distributed' (present issue)

TARANTOLA A. 1987: Inverse Problem Theory. Elsevier, Amsterdam, 613 p

A ROBUSZTUSSÁG MÉRŐSZÁMÁNAK DEFINÍCIÓJA

STEINER Ferenc és HAJAGOS Béla

A dolgozat megadja a robusztusság r-rel jelölt mérőszámának a definícióját. A definíció szerint r a szóban forgó statisztikai eljárás hatásfokainak a súlyozott átlagaként számítandó; a súlyok valamely tudományág szemszögéből adekvátnak minősülő hibaeloszlástípusoknak az előfordulási valószínűségsűrűségei. A „robusztus" — „nem robusztus" kategórikus megítélés helyett, amely ma már túlhaladottnak tekintendő, a bemutatott példák az r=36 %-tól r=96 %-ig terjedő intervallumba eső robusztusság-értékeket mutatnak. A geofizika gyakorlatának különösen szüksége van ezen a téren is arra, hogy kvantitatív összehasonlításokat tehessen.

A dolgozat hat ábrája azokat az $e(t)$ hatásfokgörbéket is bemutatja, amelyek alapján az r számítása történik. Az olvasónak így módja van arra, hogy esetleges speciális szempontok szerint is vizsgálat tárgyává tegye a különböző statisztikai eljárások hatásfokainak a hibaeloszlástípus szerinti változásait.