ROBUSTNÍ METODY.

Jaromír ANTOCH

0. Úvod. Robustní metody tvoří v posledních letech široký proud v statistické literatuře. Některým otázkám s ním spojeným byla též věnována letní škola JČSMF ROBUST 80. Tento příspěvek se snaží ukázat některé historické souvislosti v robustních metodách, přinést základní definice robustnosti a shrnout základní postupy v teorii robustního odhadu.

I. Historie. Vezmeme-li si mnohé "klasické" statistické postupy, např. v teorii odhadu (ale nejen tam), vidíme, že jsou doprovázeny známými předpoklady o normálním rozdělení chyb, jejich nezávislosti, nekorelovanosti apod. Vezmeme-li si pak potencionálního uživatele např. metody nejmenších čtverců, zjistíme, že o svých měřeních vesměs předpokládá dokonce mnohem více (a čím je model složitější, tím ochotněji). Bohužel, jak se ukazuje, narušení základních předpokladů značně snižuje sílu jednotlivých procedur.

A protože je to právě normální rozdělení táhnoucí se jako červená nit statistikou, bude možná zajímavé zde uvést v doslovné citaci Gaussovy důvody , jež jej vedly k jeho zavedení(9).

"... Der Verfasser gegenwartiger Abhandlung, welcher im Jahr 1797 diese Aufgabe nach der Wahrscheinlichkeitsrechnung zuerst untersuchte, fand bald, dass die Ausmittelung der wahrscheinlichsten Werthe der unbekannten Grosse unmoglich sei, wenn nicht die Function, die Wahrscheinlichkeit der Fehler darstellt, bekannt ist. In so fern sie dies aber nicht is, bleibt nichts ubrigt, als hypothetisch eine solche Function anzunehmen. Es schien ihm, das naturlichste, zuerst den umgekehrten Weg einzuschlagen und die Function zusuchen, die zum Grunde gelegt werden muss, wenn eine algemein als gut anerkannte Regel fur den einfachsten aller Falle herforgehen soll, die nemlich, dass das aritmetische Mittel aus mehrerer fur eine und dieselbe Grosse durch Beobachtungen von gleicher Zuverlassigkeit gefundenen Werthen als der wahrscheinlichste betrachtet werden musse. Es ergab sich daraus, dass die Wahrscheinlichkeit eines Fehler X, einer Exponentiagrosse von der Form exp(-hhXX) proportional angemommen werden musse, und dass dann gerade die jenige methode, auf die er schon einige Jahre zuvor durch andere Betrachtungen gekommen war, allgemein nothwendig werde. Diese Methode, welche er nachher besonders seit 1801 bei allerlei astronomische Rechnungen fast taglich anzuwenden Gelegenheit hatte, und auf welche auch Legendre inzwischen gekommen war, ist jetz unter dem Namen Methode der kleinsten Quadrate im allgemeinen Gebraucht...."

Je velmi zajímavé, že Gauss vlastně zavádí normální rozdělění tak, aby vyhovovalo aritmetickému průměru. Vezmeme-li navíc Gauss-Markovovu větu a centrální limitní teorém, není divu, že aritmetický průměr a dogma normality se postupně stávají téměř nedotknutelné.

Nicméně již v minulém století se objevují lidé, především z řad astronomů a geodetů, jimž tyto předpoklady ne vždy vyhovují. A nejde jen o problém odlehlých pozorování (outliers), jež se do jejich měření dostaly nedopatřením. Některá měření vykazují podstatně těžší chvosty, jiná menší hodnotu koeficientu špičatosti než odpovídající normální rozdělení... Přitom, z hlediska možností tehdejších přístrojů a metod, nebylo možno prohlásit měření za špatná. Mezi jinými jmenujme na tomto místě např. Newcomba, Smithe, Pierce ... R.C.Smith (viz (27)) mj. píše:

"...One usually has or ought to have a general notion of the quantity sought for. I've been dissatisfied with the method of taking the aritmetic mean as the most probable value of a comparatively few direct observations of quantity..."

Někteří autoři na celou věc reagovali hledáním nových postupů, adekvatnějších daným skutečnostem. Ne vždy ale byla jejich cesta hladká. V mnoha případech dochází k sporům, diskusím až osobním atd. Převážná většina těchto návrhů zmizela postupem času v propadle dějin a pouze nepatrná část se stává vlastnictvím klasické statistiky. Značně k tomu též přispěl vliv škol Pearsonových, Fischerových ... Mnoho zajímavého o této prehistorii robustní statistiky (z anglosaské oblasti) lze nalézt především v pracech Stiglerových (25), (27).

Po druhé světové válce začíná statistika pronikat i do oblastí, kde byla dříve spíše Popelkou. Často se přitom snaží opřít o metody a postupy, jež byly vymýšleny pro biologii ... Jenže ne vždy s úspěchem. Kromě jiného i proto, že klasické předpoklady, tak běžně splňované biologickými pokusy, byly zde mnohdy vážně narušeny. Objevují se odlehlá pozorování, korelovanosti v datech atd. A znovu, jako již tolikrát v minulosti, se na otázku "co nyní" objevují nové postupy, nové metody. Jednou z odpovědí je i nástup robustních metod. Začal se postupně vytvářet na základě prací Coxe, Tuckeyho, Anscomba a jejich žáků, skupiny kolem časopisu Technometrics a mnoha dalších.

Nejdůležitější metody a výsledky těchto dob lze nalézt v přehledných článcích (12), (17), jakož i leckde jinde. Některé otázky spojené s dalším vývojem a náznak dalších směrů rozvoje pak lze nalézt např. v (14). Na závěr tohoto odstavce si uveďme delší citaci charakterizující současný přístup k robustní teorii odhadu, jak jej vyjádřil Hampel v (12):

"...What do those robust estimators intend? Should we give up our familiar and simple models, such as our beautiful anylysis of variance, our powerfull regression, or our high reaching covariance matrices in multivariate statistics? The answer is no; but it may well be advantageous to modify them slightly.
In fact, good practical statisticians have done such modificationsall along in an informal way; we now only start to have a theory about them. Some likely advantages of such a formalization are a better intuitive insight into these modifications, improved applied methods (even routine methods) and the chance of having pure mathematicians contribute something to the problem. Possible disadvantages may arise along the usual transformations of a theory whe it is understood less and less by more and more people. Dogmatists, who insisted on the use of "optimal" or "admissible" procedures as long as mathematical theories contained no other criteria, may no be going to insist on "optimal robust" or "admissible robust" estimation or testing. Those who habitually try to lie with statistics, rather than seek for truth, may claim even more degrees of freedom for their wicked doings...
... Now what are the reasons for using robust procedures? There are mainly two observations which combined give an answer. Often in statistics one is using a parametric model implying

a very limited set of probability distributions thought possible, such as the common model of normally distributed errors, or that of exponencially distributed observations. Classical (parametric) statisties derives results under the assumption that these models were strictly true. However, apart of some simple models perhaps, such models are never exactly true.
We may try to distinguish three main reasons for the deviations:
(i)rounding and grouping and other "local inaccurances";
(ii)the occurence of "gross errors" such as blunders in measuring,wrong decimal points,errors in copying,inadvertent measurement of a member of a different population,or just "something went wrong";
(iii)the model may have been conceived only as an approximation anyway,e.g. by virtue of the central limit theorem..."

II. Definice robustnosti: Na otázku, co je to robustnost statistických procedur nelze dát jednotnou odpověď. Trochu by nám možná mohl pomoci Kendall-Bucklandův slovník statistických pojmů,kde pod heslem robustnost čteme: (21)

"ROBUSTNESS: Many test procedures involving probability levels depend for their exactitude on assumptions concerning the generating mechanism,e.g. that the parent variation is normal one. If the inferences are little affected by departure from those assumptions,e.g. if the significance points of a test vary little if the population departs quite substantially from the normality,the tests on the inference are said to be robust. In a rather more general sence,a statistical procedure is described as robust if it is not very sensitive to departure from the assumptions on which it depends.".

Vyjdeme-li z této"definice"robustnosti,leží před námi rázem dvě základní otázky:
(1)Jak velké je pole působnosti dané statistické procedury, neboli,je vůbec"robustní"vůči porušení základních předpokladů?
(2)Jak bychom měli navrhnout danou statistickou proceduru, aby byla necitlivá na porušení základních předpokladů a neporušili jsme přitom její dobré vlastnosti?
Různí autoři přistupovali k těmto otázkám různě,a tak si v dalším některých těchto přístupů všimneme podrobněji. Omezíme se přitom na pole teorie odhadu.

Jeden z prvních pokusů zavést pojem robustnosti v teorii odhadu pomocí vlastností, jež by dobrý robustní odhad měl splňovat postupně vykrystalizoval v polovině padesátých let na základě prací Boxe,Tuckeyho,Bickela,Hubera a dalších. Podle "doporučení" by měl takovýto "dobrý robustní odhad" vyhovovat alespoň některým z následujících podmínek:
(1) Mít vysokou absolutní vydatnost pro všechna dostatečně hladká rozdělení F.
(2) Mít vysokou vydatnost vzhledem k jednomu předem vybranému odhadu (např. aritmetickému průměru) a to pokud možno pro všechna rozdělení.
(3) Mít vysokou absolutní vydatnost na předem vybrané vhodné třídě rozdělení (tvořené např. normálním,exponenciálním, Gamma a Cauchyho rozděleními).
(4) Mít malý asymptotický rozptyl na okolí jednoho určitého rozdělení (např. normálního).

Nejrozšířenější se staly požadavky Huberovy, tj. směs bodů (3)-(4). Citujme přesněji z jeho práce (16):

"... A convenient measure of robustness for asymptotically normal estimators seems to be the supremum of the asymptotic variance ($n \rightarrow \infty$) when F (the distribution of the sample) ranges over a set of appropriate underlying distributions..."

V citované práci Huber zároveň definoval třídu tzv. M-odhadů, podal základní motivaci k jejich zavedení a charakterizoval přitom mezi nimi nejrobustnější třídu rozdělení (family of distributions) pro případ odhadu parametru polohy za předpokladu, že původní rozdělení je kontaminované normální. Postupem doby se pak prací mnoha autorů M-odhady staly nejpropracovanější částí celé teorie robustních odhadů.

Později se objevují definice matematicky mnohem striktnější, např. (3), (5), (11) ..., přistupující k problému z hlediska matematické anylýzy a topologie. Jednotlivé metody mají v podstatě týž přístup - definovat robustnost pomocí spojitosti funkcionálů (odpovídajících jednotlivým odhadům) na prostoru distribucí. Liší se přitom především volbou základního prostoru (jeho přesné struktury), uvažované metriky (topologie), technickými detaily a částečně i motivací. Patří sem např. Beranova snaha zavést robustnost pomocí Hellingerovy vzdálenost (viz. (3), (4)); Bickel-Lehmanova definice slabé robustnost (5) a další. Nicméně definicí pravděpodobně nejvhodnější pro daný problém zůstává definice Hampelova, využívající Prochorovových výsledků., Zastavme se proto u ní poněkud blíže.

Nechť F, G a $\varnothing$ jsou rozdělení z daného prostoru distribucí a d(./.) Prochorovova metrika na tomto prostoru. Označme $\varnothing(T_n,G)$ rozdělení odhadu T_n, založeném na výběru velikosti n (T_n záleží jak na velikosti výběru, tak na původním rozdělení) . Odhad T_n je robustní vzhledem k G tehdy a jen tehdy, když

$$\forall \varepsilon>0 \quad \exists \delta>0 \quad \forall F \quad \forall n \quad d(F,G)<\delta \quad \Rightarrow d\left[\varnothing(T_n,G),\varnothing(T_n,F)\right]<\varepsilon$$

Jak již bylo řečeno dříve, pro vyjádření robustnosti byla navržena řada jiných "vzdáleností". Většina z nich je však (na rozdíl od Prochorovovy metriky) zatížena některou z následujících nevýhod:

(1) Umožňuje měření blízkosti spojitých rozdělení, ale neumožňuje srovnání diskrétních empirických rozdělení se spojitými rozděleními výběru.
(2) Jsou vázány na jeden konkrétní prostor a jen obtížně se rozšiřují na prostory jiné.
(3) Nesplňují trojuhelníkovou nerovnost.
Technické detaily jakož i podrobnější výklad lze nalézt např. v (11), (24)...

III. Tři základní metody konstrukce robustních odhadů :

V celém odstavci budeme předpokládat, že $X_1, \ldots, X_n$ jsou nezávislé kopie téže náhodné veličiny X dané distribuční funkcí $P(X<x) = F(x-\Theta)$, F symetrická, Θ jednorozměrný parametr polohy.

(1) L - odhady:

Historicky nejstarší ucelená a rigorózní definice L-odhadů parametru polohy byla zřejmě podána Daniellem roku 1920 (viz (25)). Autor zároveň zkoumal základní vlastnosti těchto odhadů a určil příslušné podmínky pro asymptotickou normalitu těchto odhadů. Práce však zůstala naprosto nepovšimnuta a bez ohlasu, takže Daniellovy výsledky byly znovu odvozeny Jungem (20). Přístup těchto autorů lze vyjádřit následovně:

Def 1: Označme si pořádkové statistiky příslušné našemu výběru pomocí $X_{(1)},\ldots,X_{(n)}$. Nechť $a_1,\ldots,a_n$ jsou reálné konstanty takové, že

$$\Sigma_{i=1}^{n} a_i = 1 \quad \& \quad a_i \geq 0 \quad \forall i . \qquad \text{(III.1)}$$

Potom L-odhadem parametru θ nazvu statistiku $L_n^* = \Sigma_{i=1}^{n} a_i X_{(i)}$.

Vyslovme si tuto definici ještě v poněkud upraveném tvaru, jak se dnes častěji používá.

Def 2: Nechť $h(t)$ je reálná funkce na $(0,1)$ taková, že $h(t) = h(1-t)$ $\forall\, t \in (0,1)$, nechť $X_{(1)},\ldots,X_{(n)}$ jsou pořádkové statistiky příslušné našemu výběru a $(\varkappa) \int_0^1 h(t)\,dt = 1$.

Potom L-odhadem parametru θ nazvu statistiku

$$L_n = \Sigma_{i=1}^{n} (1/n).\, h(i/n+1).X_{(i)} \quad .$$

Pozn.: (1) V podstatě je tedy L-odhadem parametru θ každá vážená kombinace pozorování, u níž volba vah záleží pouze na pořadí pozorování.

(2) Někteří autoři se neomezují podmínkou (III.1), nýbrž dovolují i záporné váhy apod.

Tvrzení 1: Za značně obecných podmínek regularity (viz. (20)) je $\mathcal{L}\{n^{1/2}(T_n - \theta)\}$ asymptoticky normální $N(0,\sigma_L^2)$, kde

$$\sigma_L^2\,(F) = \int_0^1 U(t)^2 dt - \left(\int_0^1 U(t)\, dt \right)^2, \text{ kde}$$

$U(t)$ je neurčitý integrál z $U'(t) = (h(t) / f(F^{-1}(t)))$.

Pozn.: (3) Je-li $\psi_0'(x)$ derivace $\psi_0(x) = -f_0'(x)/ f_0(x)$, $I(F_0) = \int \psi_0^2(x)\, dF_0(x)$ Fischerova míra informace a zvolíme-li

$$h(t) = \psi_0'(F_0^{-1}(t))/ I(F_0) \quad ,$$

pak L_n je asymptoticky vydatný pro F_0.

(4) Uveďme si několik nejtypičtějších příkladů L-odhadů:

(i) Aritmetický průměr.
(ii) Medián.
(iii) α - useknutý průměr -

$$\bar{x}(\alpha) = (1/ n - 2[n\alpha]) \sum_{i=1+[n\alpha]}^{n-[n\alpha]} X_{(i)} .$$

(iv) Winsorizovaný průměr:

$$\bar{x}_w(\alpha) = \bar{x}(\alpha) + [\alpha n] X_{([\alpha n])} + (n+1-[n\alpha]) X_{(n+1-[n\alpha])} .$$

(v) Různé typy vážených vybraných kombinací pořádkových statistik,např. mediánu a kvrtilů s týmiž váhami apod.

(5) Velkou nevýhodou L-odhadů je potřeba uspořádání výběru, kterážto operace i na velkých počítačích je pro větší výběry značně zdlouhavá a nedovoluje používání L-odhadů při výpočtech v reálném čase.

(2) R- odhady.:

R - odhady byly odvozeny na základě analogií s pořadovými testy,odkud pochází i název (estimates derived from rank tests). Ukažme si pro názornost případ odvození takového odhadu jednorozměrného parametru polohy Θ pomocí dvouvýběrového pořadového testu.

Mějme 2 výběry $X_1,\dots,X_n$ a $Z_1,\dots,Z_n$ z rozdělení $F(x)$, resp. $F(x-\Theta)$. Vytvořme si nový výběr velikosti 2n ve tvaru $X_1,\dots,X_n,Z_1,\dots,Z_n$, označme jej $Y_1,\dots,Y_{2n}$ a uspořádejme jej. Dále nechť $J(t)$ je reálná funkce na $(0,1)$ taková, že $-J(t)=J(1-t)$ ($J(t)$ nám zde hraje opět roli váhové funkce). Potom pro test hypotézy $H:\Theta = 0$ proti alternativě $A:\Theta > 0$ lze užít testovou statistiku

$$W(Y_1,\dots,Y_{2n}) = \sum_{i=1}^{2n} J(i/2n+1) \cdot V_i \quad , \text{ kde}$$

$$V_i = \begin{cases} 0 & Y_{(i)} \sim Z_j \\ 1 & Y_{(i)} \sim X_k \end{cases} .$$

Máme-li nyní náhodný výběr $X_1,\dots,X_n$ z rozdělení s distribuční funkcí $F(x) = P(X < x) = F(x-\Theta)$, $J(t)$ je váhová funkce definovaná výše a $r \in R_1$ daná konstanta,vytvořme nový výběr $X_1-r,\dots,X_n-r,r-X_1,\dots,r-X_n$ délky 2n, označme jej $Y_1^*,\dots,Y_{2n}^*$, uspořádejme jej a vytvořme statistiku $W^*(r)$ analogicky dle $W(r)$ tak,že

$$W^*(r) = \sum_{i=1}^{2n} J(i/2n+1) \cdot V_i^* \quad , \text{ kde}$$

$$V_i^* = \begin{cases} 0 & Y_{(i)}^* \sim Y_{(i)}^* \sim r - X_j \\ 1 & Y_{(i)}^* \sim Y_{(i)}^* \sim X_k - r \end{cases} .$$

Def 3: R - odhadem parametru Θ rozumíme řešení rovnice - v R_n -

$$W^*(R_n) = 0.$$

Tvrzení 2: Za značně obecných podmínek regularity (viz.(17)) je $\mathcal{L}\{n^{1/2}(R_n - \theta)\}$ asymptoticky normální $N(0,\sigma_R^2(F))$ kde

$$\sigma_R^2(F) = \frac{\int J^2(t)\,dt}{\left(\int \frac{d}{dx}(J(F(x)))\,f(x)\,dx\right)^2} .$$

Pozn.: (1) Při volbě $J(t) = \psi_0(F_0^{-1}(t))$ je R_n asymptoticky vydatný pro F_0 (viz. (17)).

(2) Velkou výhodou R-odhadů z teoretického hlediska je fakt, že se můžeme opřít o velmi bohatou znalost pořadových testů, znalost jejich chování a vlastností. Na druhé straně z hlediska jejich praktického použití o nich platí totéž, co o L-odhadech. Navíc zde potřebujeme dvakrát více paměti, uspořádávání dvojnásobných souborů ...

(3) Příkladem R-odhadu nám může sloužit např. Hodges-Lehmanův odhad, jenž má po úpravách tvar

$$R_n^* = \operatorname*{med}_{1\le i\le j\le n} \frac{X_i + X_j}{2} .$$

(3) M - odhady:

Protože ve sborníku o zavedení M-odhadů, jejich vlastnostech apod. podrobně pojednává příspěvek dr. Novovičové, omezíme se pouze na základní definici, jíž budeme potřebovat v následujícím odstavci.

Def 4.: Nechť $\rho(x)$ je reálná symetrická funkce, nechť $\frac{d}{dx}\rho(x) = \psi(x)$. Nechť $X_1,\dots,X_n$ jsou nezávislé kopie téže náhodné veličiny X dané distribucí $F(x-\theta) = P(X<x)$, $\theta\in R_1$. Potom M-odhadem parametru θ nazveme statistiku $M_n = M_n(X_1,\dots,X_n)$ takovou, že buď

$$\sum_{i=1}^{n} \rho(X_i - M_n) = \inf_t \sum_{i=1}^{n} \rho(X_i - t)$$

nebo je řešením rovnice

$$\sum_{i=1}^{n} \psi(X_i - M_n) = 0 .$$

Tvrzení 3 : Za značně obecných podmínek regularity (viz.(24)) je M_n konzistentní odhad θ a $\mathcal{L}\{n^{1/2}(M_n-\theta)\}$ je asymptoticky normální $N(0,\sigma_M^2(F))$, kde

$$\sigma_M^2(F) = \frac{\int \psi^2(x)\,dF(x)}{\left(\int \psi'(x)\,dF(x)\right)^2} .$$

(4) Vztah mezi asymptotickými rozptyly L,R a M-odhadů :

V předchozích třech tvrzeních jsme měli vyjádřeny asymptotické rozptyly jednotlivých L,R a M–odhadů. Vzniká přitom přirozená otázka, v jakém jsou vzájemném vztahu. Odpověď nám dává následující tvrzení.

Tvrzení 4: Nechť $X_1,\ldots,X_n$ jsou nezávislé kopie téže náhodné veličiny X dané distribuční funkcí $F(x-\theta) = P(X < x)$. Nechť $\psi(x)$ je funkce definující M-odhad θ. Definujme váhové funkce pro L a R odhad následujícím způsobem

$$h(t) = \psi'(F^{-1}(t)) = \psi'(x) \quad \text{a}$$

$$J(t) = \psi(F^{-1}(t)) = \psi(x),$$

kde $\frac{d}{dx}\psi(x) = \psi'(x)$; $x = F^{-1}(t)$; $(t = F(x))$.

Potom, za předpokladu existence asymptotických rozptylů pro R, L a M odhady parametru θ definované funkcemi $J(t)$, $h(t)$ a $\psi(t)$, platí

$$\sigma^2_R(F) = \sigma^2_L(F) = \sigma^2_M(F) .$$

Důsledek: Je-li alespoň jeden z odhadů asymptoticky vydatný pro F, pak jsou pro F asymptoticky vydatné i ostatní odhady.

Literatura:

1 ANDREWS & all.: Robust Estimates of Location. Princeton University Press 1972.

2 ANSCOMBE F.J.: Rejection of Outliers, Technometrics 2(1960), 123-147.

3 BERAN R.: Robust Location Estimates, AS 5(1977), 431-444.

4 BERAN R.: Minimum Helinger Distance Estimates for Parametric Models, AS 5(1977), 445-463.

5 BICKEL P.J. & LEHMANN E.L.: Descriptive Statistics for Nonparametric Models - Location, AS 3(1975), 1045-1069.

6 BOX G.E.P.: Non-normality and Tests on Variances, Biometrika 40 (1953), 318-335.

7 BOX G.E.P. & DRAPER N.R.: Robust Designs, Biometrika 62(1975), 347-352.

8 COLLINS J.R.: Robust Estimation of Location Parameter in Prasence of Asymetry, AS 4(1976), 68-85.

9 GAUSS C.F.: článek z roku 1821 přetištěný v Werke Bd.4, str.98.

10 GNADESIKAN R. & KETTERING J.R.: Robust Estimates, Reziduals and Outlier Detection with Multiresponse Data, Biometrics 28 (1972), 81-124.

11 HAMPEL F.R.: A General Qualitative Definition of Robustness, AMS 42(1971), 1887-1896.

12 HAMPEL F.R.: Robust Estimation: A Condensed Partial Survey, Z.f. Wahrscheinlichkeitstheorie verw. Geb. 27(1973), 87-104.

13 HAMPEL F.R.:The Influence Curve and its Role in Robust Estimation,JASA 69(1974),383-393.

14 HAMPEL F.R.:Modern Trends in the Theory of Robustness, Report ETH,ZURICH 1977.

15 HARTER H.L.: The Method of Least Squares and Some Alternatives.Int. Stat. Rev. 42(1974),147-174,235-264,282;43(1975), 1-44,125-190,273-278,269-272;44(1976),113-159.

16 HUBER P.J.:Robust Estimation of Location Parameter,AMS 35 (1964),73-101.

17 HUBER P.J.:Robust Statistics:A Review,AMS 43(1972),1041-1067.

18 HUBER P.J.:Robust Regression:Asymptotics,Conjectures and Monte Carlo,AS 1(1973),799-821.

19 JAECKEL L.A.:Robust Estimates of Location:Symetric and Asymetric Contamination,AMS 42(1971),1020-1034.

20 JUNG J.:On Linear Estimates Defined by Continuous Weight Function,Arkiv f.Mat. 3,199-209.

21 KENDALL M.G.&BUCKLAND W.R.: A Dictionary of Statistical Terms,Oliver&Boyd,Edinburgh 1971,3rd Edition.

22 MILLER R.G.:The Jacknife:A Review,Biometrika 61(1974),1-15.

23 N.C.H.S.:Annoted Bibliography on Robustness Studies of Statistical Procedures,U.S.Dept.Health Educ.Welf.,Publication HSM 1972,72-1051.

24 REY W.J.J.: Robust Statistical Methods.Springer Lecture Notes 690,1980.

25 STIGLER S.M.:Simon Newcomb,Percy Daniell and the History of Robust Estimation 1885-1920,JASA 68(1973),872-879.

26 STIGLER S.M.: Do Robust Estimates Work with Real Data?, AS 5(1977),1055-1098.

27 STIGLER S.M.:R.C.Smith - A Victorian Statisticians,Biometrika 67(1980).