

# Az $L_2$ -, $L_1$ - és $P$ -norma szerinti statisztikai eljárások aszimptotikus hatásfokainak összehasonlítása<sup>1</sup>

STEINER FERENC<sup>2</sup>

A statisztikai elvek és normák fogalmát, valamint kölcsönös kapcsolatukat röviden összefoglaló 1. pont után a 2. pont a hagyományos,  $L_2$ -normára épülő statisztikát jellemzi kvantitatívan az  $L_1$ -normát alapul választóval összehasonlítva, mégpedig az  $e(L_2/L_1)$  relatív statisztikai hatásfokoknak az „általánosított Gauss-szupermodell” hibátípusaira történő megadásával. Ha nincsenek kiugrók (outlierek, durva hibájú adatok), akkor a szupermodell jelentős típusartományán az  $L_2$ -norma alkalmazása kedvezőbb, mint az  $L_1$ -normáé. A 3. pont a  $P$ -normán alapuló statisztikát hasonlítja össze az  $L_1$ -norma szerintivel, megadva az  $e(P/L_1)$  relatív hatásfok értékeit és görbéjét a harang alakú hibaeloszlás-típusok  $f_c(x)$ -szel jelölt „teljes szupermodelljére”; az  $e(P/L_1)$  értékek szignifikánsan nagyobbak 100%-nál az egész  $f_c(x)$  szupermodellre. Ennek gazdaságossági következménye az, hogy a  $P$ -normára alapuló eljárások az eredmények azonos pontosságát kisebb mérési adatszámmal érik el, mint az  $L_1$ -módszerek.

**F. STEINER: Comparison of the  $L_2$ ,  $L_1$ - and  $P$ -norm based statistical procedures in respect of their asymptotic efficiencies**

After the Introduction in Section 2 the conventional ( $L_2$ -based) statistics is compared with the  $L_1$ -based one, giving the values of the relative efficiencies (i.e.,  $e(L_2/L_1)$ ) for the “generalised Gaussian supermodel” of errors. If no outliers exist,  $L_2$  works better than  $L_1$  does in a great type-domain of this supermodel. In Section 3 the  $P$ -based statistics is compared with the  $L_1$ -based one, showing the curve of the relative efficiencies  $e(P/L_1)$  being for all error types of the “complete supermodel”  $f_c(x)$  significantly greater than 100%. As economical consequence the following can be formulated: to achieve the same accuracy,  $P$ -based procedures need less data than  $L_1$ -based ones for all bell-shaped  $f_c(x)$  error distributions.

## 1. Bevezetés: statisztikai elvek és normák

Ezt az összeállítást a legszívesebben azzal kezdeném, hogy „kezdtem volna a legkisebb négyzetek elve” — de ezt aligha tehetem, ha arra gondolunk, hogy ezt az elvet alig több, mint 200 évvel ezelőtt mondták ki. Másrészt azonban amióta mérést valaha is végeztek, ezt legalább háromszor megismételte a mérést végző mérnök vagy tudós, akár kimondta (vagy nem is ismerte) azt a frappánsan tömör megállapítást, hogy „egy mérés nem mérés”; helyes érték-ként azután az  $n$  darab  $x_i$  mérési adatból azok számtani

átlagát, azaz az  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  mennyiséget fogadta el. Ez az

utóbbi (szinte mechanikusan, emberemlékezet óta) végzett művelet azonban éppen a legkisebb négyzetek elvének követelményét kielégítő értékre vezet, azaz a mért  $x_i$ -k  $\bar{x}$  számtani átlagát mint jó értéket elfogadni ekvivalens a legkisebb négyzetek elvének ösztönös gyakorlati alkalmazásával, — márpedig ez valóban ugyanolyan réginek tekinthető, mint amikor egyáltalán a mérések elkezdődtek. Ilyen szempontból ezt az elvet valóban a *kezdetektől* alkalmazták.

Mit is mond ki ez az elv?

Akár szóban, akár formulával akarjuk akármelyik statisztikai alapelvet, így a legkisebb négyzeteset is megadni, az „eltérésnek” nevezett és  $X_i$ -vel jelölt mennyiségek biztosítják a legáttekinthetőbb felírást, ill. megfogalmazást. Az  $X_i$  eltérés egyszerűen az  $x_i$  mért és a neki megfelelő,  $\xi_i$ -vel jelölt *számított* értéknek nevezett mennyiség eltérése:

$$X_i = x_i - \xi_i. \quad (1)$$

Ez a  $\xi_i$  pl. a geofizikai inverzió esetében a szó igazi értelmében számított érték: valamely adekvát geometriai modellnek a hatását adja ugyanarra az  $\bar{y}_i$  koordináta-vektorral megadott pontra, amelyben a mérés az  $x_i$  értékét eredményezte. Az  $\bar{y}_i$  vektor terepmérés esetén kétkomponensű, térbeli méréskor háromkomponensű, — a legegyszerűbb esetben, a szelvénymérés esetében viszont egykomponensű, amikor persze elegendő a mérési pont egyetlen koordinátáját vektorjel nélkül, csupán  $y_i$ -vel jelölni.

Szelvényméréssel persze csak kétdimenziós, azaz a szelvényirányra merőleges irányban elnyúlt szerkezetek (vagy ha tetszik, hatók) paramétereit tudjuk meghatározni. Az elképzelhető legegyszerűbb 2-D modell gravitációs hatók esetében a vízszintes henger,  $M$  tengelymélységgel,  $R$  sugárral, és a szelvénymérés kezdőpontját választva a vízszintes  $y$ -tengely origójának, a henger tengelyének ettől való  $Y$ -nal jelölt vízszintes távolságával jellemezve. Ha felhagyott, talajvízszint feletti, ismert  $\sigma$  sűrűségű közetben kihajtott egyetlen vágat jellemző adatait kell meghatározni, a  $\xi_i$  számított értéket jól ismeren a

$$\xi_i = -41,9 \cdot \sigma \cdot \frac{M \cdot R^2}{M^2 + (y_i - Y)^2} \quad (2)$$

formulából nyerjük ( $\mu$  Gal-ban, ha  $\sigma$  ismert értékét  $t/m^3$ -ben, a hossz dimenziójú mennyiségeket pedig  $m$ -ben helyettesítjük). Ha az eddig általánosan  $x_i$ -vel jelölt mennyiséget, mikrograviméteres mérés eredményéről lévén szó,  $g_i$ -vel jelöljük, akkor az  $X_i$  eltérés kifejezése jelen eset-

<sup>1</sup> Beérkezett: 2000. február 29-én

<sup>2</sup> Miskolci Egyetem Geofizikai Tanszék, H-3515 Miskolc, Egyetemváros

ben (1) és (2) alapján

$$X_i = g_i - \left( -41,9 \cdot \sigma \cdot \frac{M \cdot R^2}{M^2 + (y_i - Y)^2} \right) \quad (3)$$

lesz (a gravimétermérés bázisállomását vízszintesen kellően nagy távolságban felvéve, a mért  $x_i = g_i$  értékek, üreghatásról lévén szó, persze ugyanúgy negatív előjelűek lesznek, mint a  $\xi_i$  számított értékek).

Valamely statisztikai alapelv követelményét teljesítve kapjuk meg az ismeretlen  $p_1, p_2, \dots, p_i, \dots, p_j$  modellparaméterek (jelen esetben az  $M, R$  és  $Y$ ) valódi értékeit. (Két párhuzamos vágat esetén természetesen hat „valódi” paraméterértéket kell meghatároznunk, ld. a 10.22 ábrát a STEINER (ed.) [1997] 254. oldalán.) Ha adekvát modellt vettünk fel, ekkor a  $\xi_i$  számított értékek helyes értékeknek tekinthetők, amikor is az  $X_i$  eltérések (amelyeket ilyenkor megkülönböztetésül reziduáloknak nevezünk) a mérések hibáit szolgáltatják.

Nyilvánvaló, hogy ha a modellparamétereknek (jelen esetben az  $M, R$  és  $Y$ ) nem a valódi értékeit szerepeltetjük a számított  $\xi_i$ -ben, az  $X_i$  eltérések abszolút értéke az esetek többségében (a reziduálokhöz viszonyítva) növekedni fog, azaz ezek az  $|X_i|$  abszolút értékek *összességükben* lesznek nagyobbak. A *statisztikai alapelvek* éppen abban különböznek egymástól, hogy *más és más kifejezést adnak meg az  $|X_i|$  eltérésösszesség nagyságának a mérésére*, és ezek minimális értékére vezető modellparamétereket fogadják el a legvalószínűbb („valódi”) értékeknek. Ezen kifejezések, ill. minimum-követelmények közül két példa:

„a legkisebb négyzetek elve”: 
$$\sum_{i=1}^n X_i^2 = \min. \quad (4)$$

„a legkisebb szorzatok elve”: 
$$\prod_{i=1}^n (S^2 + X_i^2) = \min. \quad (5)$$

( $n$  a mérési adatok száma). Utóbbi elvnel standard esetben az  $\varepsilon$ -nal jelölt és az

$$\varepsilon^2 = 3 \cdot \frac{\sum_{i=1}^n \frac{X_i^2}{[\varepsilon^2 + X_i^2]^2}}{\sum_{i=1}^n \frac{1}{[\varepsilon^2 + X_i^2]^2}} \quad (6)$$

egyenletet teljesítő, „dihézió”-nak nevezett mennyiség kétszeresét alkalmazzuk  $S$ -ként, azaz ekkor a legkisebb szorzatok elvének követelménye

$$\prod_{i=1}^n (4\varepsilon^2 + X_i^2) = \min. \quad (7)$$

alakban írható fel.

Ha  $X_i$  reziduál értelmű (azaz a  $\xi_i$ -t helyes modellparaméter-értékekkel számítottuk), jogosan merül fel az az igény, hogy a statisztikai elvekben szereplő kifejezéseknek a minimumhelyet változtatlanul hagyó módosításaival hibajellemzőket adjunk meg. Ezek a módosítások akkor is cél-

szerűek, ha az  $X_i$ -knek általános eltérés értelmük van, hiszen az egyes elvek szerinti eltérésjellemezőnek nyilván  $n$ -től függetlennek és  $x_i$ -dimenziójúnak kell lennie. Ezeket az eltérésjellemezőket *normáknak* nevezik és egy csoportjukat a következő kifejezés definiálja:

$$L_p = \left[ \frac{1}{n} \sum_{i=1}^n |X_i|^p \right]^{1/p} \quad (p > 0). \quad (8)$$

Azonnal észre vesszük, hogy  $p = 2$ -re

$$L_2 = \sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2} \quad (9)$$

adódik, ami a (4) bal oldalán álló kifejezésből  $n$ -nel való osztás és négyzetgyökvonás, azaz a minimumhelyet változtatlanul hagyó (és végül  $n$ -független és  $x_i$ -dimenziójú eltérésjellemezőt eredményező) műveletek után adódik. A legkisebb négyzetek elvének követelményét tehát

$$L_2 = \min. \quad (10)$$

alakban is felírhatjuk.

A (8) kifejezés legegyszerűbb alakját nyilván a  $p = 1$  választással kapjuk:

$$L_1 = \frac{1}{n} \sum_{i=1}^n |X_i|, \quad (11)$$

és nemcsak hogy semmi akadályja nincs annak, hogy az

$$L_1 = \min. \quad (12)$$

követelményt teljesítve határozzuk meg az ismeretlen  $p_j$  paramétereket (modellparamétereket inverzió esetén, vagy pl. polinom együtthatókat regressziós feladatoknál), de jól láthatóan a kiugrók (az outlierek, azaz a durva hibájú  $x_i$ -k, amelyek extrém nagy  $|X_i|$ -kre vezethetnek,) az  $L_1$ -norma (11) kifejezését kevésbé torzítják, mint a  $L_2$ -normában szereplő összeget, így kiugrók esetén a (10) követelmény teljesítése esetleg megtévesztő  $p_j$ -ket eredményezhet akkor is, amikor ez a hatás a (12) követelmény esetén elhanyagolható.

Ami a minimumhely meghatározást illeti, a számítástechnika által nyújtott, rohamosan növekvő lehetőségek egyre inkább a „globális optimalizáció”-nak nevezett, abszolút minimumhely meghatározó algoritmusok valamelyikének, például az SA (simulated annealing) módszernek az alkalmazását teszik célszerűvé (magyar nyelvű leírása pl. KIS [1996]), akármelyik norma minimumhelyét akarjuk is meghatározni. Az idézett cikk az  $L_1$ -re és az  $L_2$ -re egyaránt mutat be példákat, — de a feladat jellegéből, a hiba típusától, valamint esetleg egyéb követelményeinktől függően a STEINER (ed.) [1997] 20. oldalán található táblázat tíz normája közül bármelyiket választhatjuk; a robusztusság és rezisztencia szempontjából pl. egyaránt előnyös (különösen földtudományi feladatoknál) az eltérések

$$P = \varepsilon \left\{ \prod_{i=1}^n \left[ 1 + (X_i / 2\varepsilon)^2 \right] \right\}^{1/2n} \quad (13)$$

normájának az alapul választása ( $\varepsilon$  a (6)-ot kielégítő, dihézióknak nevezett mennyiség). A globális optimalizáció egyrészt nem igényel a valósághoz közel álló startmodellt (ezt a sajátját Szűcs [1997a] az A5.4 és A5.5 ábrákon nagyon meggyőzően mutatja be), az ismeretlen paraméterek nagy száma sem jelent problémát (Szűcs [1997b]-ben, mélyfúrási geofizikai példa keretében, 18 paraméter meghatározása történik SA módszerrel,) és végül: a meghatározandó paramétereknek, a számított  $\xi_i$  értékben való szereplésüket illetően, semmiféle analitikus kikötést nem kell teljesíteniük. Ez utóbbival kapcsolatban feltétlenül meg kell egy speciális esetet említeni, mégpedig azt, hogy az ismeretlen paraméterektől való lineáris függés esetén az  $L_2$ -norma minimumhelyének meghatározása egyetlen lineáris algebrai egyenletrendszer megoldását igényli, de ilyen esetekben még a legkisebb szorzatok elvének megfelelő  $P$ -norma alapul választásánál is csak lineáris algebrai egyenletrendszerek megoldása szükséges, iterációs lépésként azonban változó súlyt alkalmazva (pl. STEINER [1990] 313. oldal). Regressziós feladatoknál ez a feltétel gyakran teljesül (hogy az inverziónál szinte sohasem, azt a legegyszerűbb gravitációs modell (2) formulájából is látjuk), így az  $L_2$  alapul választása számítástechnikai kényyszerűség volt az elmúlt évtizedekig, akkor is, ha más norma alkalmazása sokkal megbízhatóbb eredményekre vezetett volna, hiszen még akár csak néhány évtizeddel ezelőtt is a fentiekben említett iterációs eljárások sem voltak rutinszerűen alkalmazhatók, a globális optimalizáció általánosan használható (tehát lineáris függést nem feltételező) módszereiről nem is beszélve. A klasszikus statisztika a fentiek miatt támaszkodott szükségképpen az  $L_2$ -normára, vagy másképpen fogalmazva: a legkisebb négyzetek elvére.

Térjünk vissza a (8)-ban definiált  $L_p$ -normacsaládhoz. Kimutatható, hogy a (8) szerinti általános  $L_p$ -norma minimumhelyét keresni akkor a legelőnyösebb, ha a méréseket terhelő hibák „általánosított Gauss-típusúak” (TARANTOLA [1987] 26. oldal), azaz standard sűrűségfüggvényüket az

$$f_p(x) = \frac{p^{1-1/p}}{2 \cdot \Gamma(1/p)} \cdot e^{-|x|^p/p} \quad (14)$$

kifejezés adja meg (persze a  $p$  indexnek az  $L_p$  indexével kell megegyeznie a legkedvezőbb esetben). Azonnal látjuk ( $\Gamma(1/2)$  azonos lévén  $\sqrt{\pi}$ -vel), hogy  $p=2$  esetén Gauss-típusú sűrűségfüggvény adódik:

$$f_G(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad (15)$$

$p=1$  esetén pedig a Laplace-típusú:

$$f_L(x) = \frac{1}{2} e^{-|x|}. \quad (16)$$

Ez utóbbinak a sűrűségfüggvénye már nem differenciálható a szimmetriapontban (az origóban), de tovább csökkentve  $p$ -t, pl.  $1/2$  értékűre, a sűrűségfüggvény már túszerűen hegyes lesz az origóban (ld. STEINER (ed.) [1997] 9.14. ábráján az  $f_{ne}(x) = e^{-2\sqrt{|x|}}$  görbét, vagy az  $f_G$ ,  $f_L$  és  $f_{ne}$  összehasonlítását a STEINER, HAJAGOS [2000a] első ábrá-

ján), így ennek előfordulása hibaeloszlásként a földtudományban már eléggé valószínűtlen. Hasonló mondható az egyenletes eloszlásra is, amely (14)-ből a  $p \rightarrow \infty$  határesetben adódik. Az alábbi, az  $L_1$ - és  $L_2$ -normát összehasonlító vizsgálatainkat mégis a (14) „általánosított Gauss-típuscsalád”  $1/2 \leq p \leq \infty$  típusparaméter-tartományára végezzük el, tehát valamivel általánosabban, mint az földtudományi szempontból szükséges lenne. Az összehasonlításához az  $f_p(x)$  szupermodell alapul választása azért kézenfekvő, mert ez a típuscsalád egyaránt tartalmazza az  $L_2$ -re optimális Gauss-típust (ld. (15)), valamint az  $L_1$ -re optimális Laplace-típust (ld. (16)).

## 2. Az $L_1$ -norma és $L_2$ -norma összehasonlítása relatív hatásfokok segítségével az általánosított Gauss-típuscsaládra

Az összehasonlítás akkor lesz a legáttekinthetőbb, ha a legegyszerűbb esetet vesszük alapul: az  $n$  darab függetlenül nyert  $x_i$  adat egyetlen ( $T$ -vel jelölt és helyparaméternek nevezett) ismeretlenre vonatkozó közvetlen mérés eredménye. Ekkor ( $X_i = x_i - T$  lévén)  $L_1$  és  $L_2$  alakja

$$L_1 = \frac{1}{n} \sum_{i=1}^n |x_i - T| \quad (17)$$

ill.

$$L_2 = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - T)^2}. \quad (18)$$

Ha nem lenne közismert, akkor is könnyen levezethetnénk, hogy az  $L_1$ -re vonatkozó (12) minimumkövetelményt a *med*-del jelölt mintamedián, az  $L_2$ -re vonatkozó (10)-et pedig az  $\bar{x}$ -sal jelölt számtani átlag teljesíti.

Két becslést (jelen esetben a  $T$ -re vonatkozó becsléseket) relatív hatásfokokkal a legcélszerűbb összehasonlítani, ezt pedig (ld. pl. STEINER [1990] (5-103) formuláját) a két becslésre vonatkozó aszimptotikus szórásnégyzetek hányadosaként célszerű számítanunk. (Az abszolút hatásfokok aránya ugyanazt a relatív hatásfok értéket szolgáltatja, — amit felidézni talán nem is felesleges, — de a számításokat az abszolút hatásfokokkal végezve felesleges többletmunkát vállalunk, amennyiben csak összehasonlítás a célunk, mint jelen esetben is.)

Az  $L_1$  norma minimumát kielégítő *med* aszimptotikus szórása unimodális, szimmetrikus  $f(x)$  sűrűségfüggvény esetén

$$A_{med} = \frac{1}{2 \cdot f(\text{med})}. \quad (19)$$

Itt az  $f$  argumentumában a mintamediánok által becsült,  $\int_{-\infty}^{\text{med}} f(x) dx = 0,5$  által definiált „valódi értéket” kell értenünk, — de ez szimmetrikus esetben úgyszólamint ismert: egybeesik a szimmetriaponttal, azaz standard ( $T=0$  és  $S=1$ ) esetben az origóval egyezik meg:

$$A_{med} = \frac{1}{2 \cdot f(0)} \quad (20)$$

A (19) formulát pl. STEINER [1990] 5.6 táblázatából olvashatjuk ki, s ugyanott (a második sorban) látjuk, hogy az  $\bar{x}$  számtani átlagok aszimptotikus szórása (amelyet most  $A_{\bar{x}}$ -sal jelölünk), megegyezik az anyaeloszlás szórásával. Az  $A_{\bar{x}}^2$  aszimptotikus szórásnégyzet tehát a (14)-ben definiált  $f_p(x)$  általánosított Gauss-típusú eloszlásra

$$A_{\bar{x}}^2 = p^p \frac{\Gamma(3/p)}{\Gamma(1/p)} \quad (21)$$

[STEINER, HAJAGOS 2000a (4) formula], míg (20)-ból és (14)-ből azonnal beláthatjuk, hogy ugyanerre az eloszlás-családra

$$A_{med}^2 = \frac{p^p \cdot \Gamma^2(1/p)}{p^2} \quad (22)$$

a mintamedianok aszimptotikus varianciája. Ha tehát  $e(L_2/L_1)$ -gyel jelöljük a hagyományosan alkalmazott  $L_2$ -norma hatásfokát az  $L_1$ -norma hatásfokához viszonyítva az egyes,  $p$  típusparaméterekkel jellemzett hibaeloszlásokra, a relatív hatásfok kifejezése  $A_{med}^2$  és  $A_{\bar{x}}^2$  hányadosa lesz:

$$e(L_2/L_1) = \frac{\Gamma^3(1/p)}{p^2 \cdot \Gamma(3/p)} \quad (23)$$

(Triviális ugyan, de megemlíthető, hogy az  $L_1$ -norma  $L_2$ -höz viszonyított relatív hatásfokát,  $e(L_1/L_2)$ -t a (23) kifejezés reciproka szolgáltatja.) Az abszolút statisztikai hatásfokokhoz hasonlóan a relatív hatásfokokat is százalékokban szokás megadni.

A  $p \rightarrow \infty$ -hez tartozó egyenletes eloszlásra vonatkozóan  $L_2$ -bázisú hagyományos statisztika háromszor akkora hatásfokú, mint az  $L_1$ -et alapul választó eljárások, azaz  $e(L_2/L_1) = 300\%$ . Az 1. táblázat a  $100 \geq p \geq 0,5$  tartományban 18-féle hibaeloszlás-típusra adja meg az  $e(L_2/L_1)$  és  $e(L_1/L_2)$  relatív hatásfokok számértékeit.

A relatív hatásfokok változásait azonban szemléletesebben mutatják ezen relatív hatásfokok görbéi, amikor is célszerű abszcisszaként az  $f_p$  eloszlástípusnak az ( $f_u$ -val jelölt) egyenletes eloszlástól mért  $D(F_u, F_p)$  típus-távolságának választani. A típus-távolság általános definícióját STEINER (ed.) [1997] 43. oldalán levő formulák adják meg, amely jelen esetben a

$$D(F_u, F_p) = \min_S \left\{ \max_x |F_p(0, S; x) - F_u(0, 1; x)| \right\} \quad (24)$$

egyszerű alakban írható (ld. még STEINER, HAJAGOS [2000a] formuláját). A felvett  $p$  típusparaméterekhez tartozó  $D(F_u, F_p)$  értékeket az 1. táblázat a negyedik oszlopban adja meg.

Az  $e(L_2/L_1)$  és  $e(L_1/L_2)$  relatív hatásfokgörbéket tehát az 1. ábra az egyenletes eloszlástól a tú-eloszlásig ( $p=0,5$ ) az egyenletes eloszlástól mért  $D(F_u, F_p)$  típus-

távolságok függvényében mutatja be. A két görbe metszéspontja a  $p=1,407$  típusparaméternél van (1. táblázat), azaz az egyenletes eloszlástól eddig a hibátípusig a hagyományos statisztika alkalmazása előnyösebb kiugrómentes esetekben, mint az  $L_1$ -bázisú statisztikai eljárásoké.

$p$	$e(L_2/L_1)$ (%)	$e(L_1/L_2)$ (%)	$D(F_u, F_p)$
100	299,86	33,35	0,00159
50	299,43	33,40	0,00311
20	296,64	33,71	0,00742
10	287,82	34,74	0,01388
5	259,89	38,74	0,02479
3	213,62	46,81	0,03647
2	157,08	63,66	0,04804
1,75	135,64	73,72	0,05226
1,5	110,35	90,62	0,05735
1,407	99,95	100,05	0,05960
1,25	81,30	122,99	0,06364
1,0	50,00	200,00	0,07168
0,9	37,74	265,00	0,07558
0,8	26,31	380,13	0,07999
0,75	21,10	473,97	0,08243
0,667	13,47	742,57	0,08691
0,55	5,58	1792,97	0,09428
0,5	3,33	3000,00	0,09794

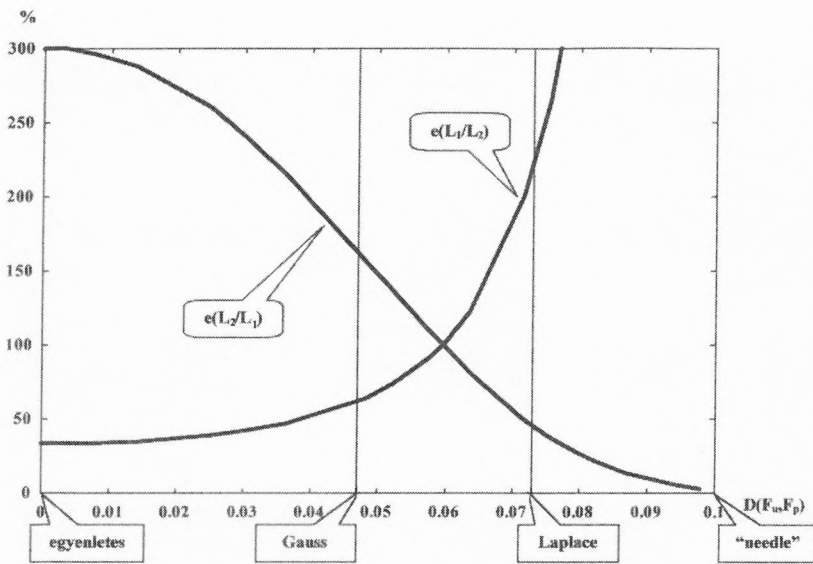
1. táblázat. Az ún. „általánosított Gauss-szupermodell” (ld. a sűrűségfüggvények (14) formuláját) 18 hibaeloszlás-típusára a táblázat megadja az egyenletes eloszlástól mért  $D(F_u, F_p)$  típus-távolságokat és az  $e(L_2/L_1)$  és  $e(L_1/L_2)$  relatív hatásfokokat. Ez utóbbi értékeket (a  $D(F_u, F_p)$  függvényében) az 1. ábra görbék-ként mutatja be

Table 1. For 18 distribution types of the so-called “generalised Gaussian supermodel” (see Eq. 14) there are given the  $D(F_u, F_p)$  type-distances from the uniform distribution and the relative efficiencies  $e(L_2/L_1)$  and  $e(L_1/L_2)$ . These latter values are shown vs  $D(F_u, F_p)$  as curves in Fig. 1

A geofizikai gyakorlatban (és az ahhoz csatlakozó statisztikai elméleti vizsgálatokban) azonban az  $f_p$  hibátípusal összehasonlítva számos előnnyel bír az

$$f_a(x) = \frac{\Gamma(a/2)}{\sqrt{\pi} \cdot \Gamma[(a-1)/2]} \cdot (1+x^2)^{-a/2} \quad (a < 1) \quad (25)$$

szupermodell típusainak hibamodellként való alkalmazása, kézenfekvő tehát felvetni azt a kérdést, hogy melyik  $f_a$ -típus vehető a  $p=1,407$ -hez tartozó  $f_p$  típusal gyakorlatilag azonosnak? Nyilván a  $D(F_a, F_{p=1,407})$  típus-távolság minimumát szolgáltató  $a$  típusparamétert keressük, amely  $a=5,56$ -nak adódik (maga a minimális típus-távolság érték 0,003666). Érdekességgel bírhat az az összehasonlítás is, hogy mennyire közeli értékek az egyenletes eloszlástól mért típus-távolságok:  $D(F_{a=5,56}, F_u) = 0,0595504$  és  $D(F_{p=1,407}, F_u) = 0,0595264$ . Kimondhatjuk tehát, hogy a konvencionális statisztika nagyobb hatásfokú az  $L_1$ -re alapulónál az  $a > 5,56$  típusparaméterekkel jellemzett  $f_a(x)$ -eloszlásokra; ez a típus-távolság a teljes Jeffreys-



1. ábra. Az  $L_2$ -eltérésnormán alapuló hagyományos statisztika relatív hatásfokai az  $L_1$ -bázisú statisztikához viszonyítva, azaz az  $e(L_2/L_1)$  görbe, reciprokának, az  $e(L_1/L_2)$  relatív hatásfoknak a görbéjével együtt. Az abszcissa az ún. „általánosított Gauss-szupermodell” típusainak a távolsága az egyenletes eloszlástól (1. táblázat)

Fig. 1. Relative efficiencies of the  $L_2$ -based (conventional) statistics compared with the  $L_1$ -based one, i.e. the  $e(L_2/L_1)$ -curve; the also shown  $e(L_1/L_2)$ -curve corresponds to the reciprocal values. On the abscissa there are given the type-distances of the error-distribution types of the so-called “generalised Gaussian supermodel” from the uniform distribution (Table 1)

intervallumot ( $10 \geq a \geq 6$ ) tartalmazza és közel van  $a = 5$ -höz, amely  $f_a(x)$  hibatípus DUTTER [1986/87] szerint igen gyakran fordul elő a geostatistikában. A szerzők egyéb irányú vizsgálatai szerint (pl. geodézia, csillagászat) ez a hibatípus egyéb diszciplínák esetén is gyakori, és ezért az  $a = 5$ -höz tartozó  $f_a(x)$ -et újabban inkább nevezük „statistikai”, mint „geostatistikai” eloszlásnak, és ennek megfelelően  $f_{st}(x)$ -szel jelöljük, amelynek standard alakja (az általános (25)-ből következően)

$$f_{st}(x) = 0,75 \cdot (1 + x^2)^{-5/2}. \quad (26)$$

Kiegészítésül emlékeztetünk arra (STEINER (ed.) [1997] 366. oldal), hogy  $a = 2$  a Cauchy-eloszlást definiálja:

$$f_C(x) = \frac{1}{\pi} \frac{1}{1 + x^2}, \quad (27)$$

az  $a = 9$ -hez tartozó  $f_a(x)$ -et Jeffreys-típusúnak mondjuk:

$$f_J(x) = \frac{35}{32} \cdot (1 + x^2)^{-9/2} \quad (28)$$

s végül  $a \rightarrow \infty$  esetén a (15)-beli,  $f_G(x)$ -szel jelölt típus-hoz jutunk, azaz az  $f_a(x)$  szupermodell határesetként a Gauss-típust is tartalmazza.

### 3. Az $e(P/L_1)$ relatív aszimptotikus hatásfokok a hibák teljes $f_c(x)$ szupermodelljére

A gyakorlatban sima, harang alakú sűrűségfüggvénnyel jellemezhető hibaeloszlást várunk a földtudományokban (de feltehetően nagyon sok egyéb tudományágban is). A 2.

ponthoz hasonlóan valamivel nagyobb általánosságra törekszünk, mint amelyet a földtudományi praxis okvetlenül megkívánna, így vizsgálatainkat a harang alakú hibaeloszlások teljes  $f_c(x)$  szupermodelljére végezzük el (itt  $c$  a „complete” kezdőbetűje), amelyet STEINER, HAJAGOS [1998] definiál:

$$f_c(x) = \begin{cases} f_p(x) & (\infty > p \geq 2) \\ f_a(x) & (\infty > a > 1) \end{cases} \quad (29)$$

(ld. a (14) és (25) definíciókat). Talán szokatlan két különböző formulával definiálni egyetlen szupermodellt, de gondoljunk a 2. pont utolsó mondatára (hogy ti.  $a \rightarrow \infty$  esetén Gauss-típushoz jutunk), ugyanakkor  $p = 2$ -nél  $f_p(x)$  is ugyanaz a Gauss-típus, így az  $f_c(x)$  két különböző formulája folyamatosan változó tulajdonságú típus-sorozatot definiál (ld. STEINER, HAJAGOS [1998]-nak a szárnysúlyokat valószínűségekkel mérő táblázatát, — vagy akár a 2. ábra görbét).

A nagy  $p$ -knél kicsiny a szárnytartomány és lapos a maximumzóna, míg a típustartomány másik végén a szárnyak súlyait (az  $a = 1$ -hez köze-

ledve) már nagy valószínűségértékek jellemzik. A szupermodell nagyobb részén azonban valóban a szó köznapi értelmében is harang alakúak az  $f_c(x)$  görbék.

A klasszikus statisztikában a „haranggörbe” alatt túlnyomóan (vagy kizárólag) az  $f_G(x)$  sűrűségfüggvény görbéit értik, holott a „Gauss-féle haranggörbe” csak egyetlen lehetőség a végtelen sok haranggörbe közül. Bizonyos értelemben mi is megtartjuk a Gauss-féle haranggörbe kitüntetett szerepét, ezúttal azonban inkább csak azért, mert ennél a hibatípusnál csatlakoznak egymáshoz az  $f_c(x)$  szupermodell különböző formulával definiált típusai (ld. a (29) definíciót). A következő két típustávolság-jelölést fogjuk alkalmazni:

$$D \equiv D(F_a, F_G) = \min_S \left\{ \max_x |F_a(0, S; x) - F_G(0, 1; x)| \right\} \quad (30)$$

és

$$D^- \equiv D(F_p, F_G) = \min_S \left\{ \max_x |F_p(0, S; x) - F_G(0, 1; x)| \right\}. \quad (31)$$

Az utóbbi esetben a kitevő helyén levő mínusz jel arra utal, hogy a Gauss-típustól távolodva egyre kisebbek a szárnyak súlyai; a (30) szerinti  $D$  típustávolság viszont csökkenő  $a$  típusparaméter-értékeknél, azaz a szárnyak súlyainak növekedésekor lesz egyre nagyobb és nagyobb értékű. Ami a  $D^-$  és  $D$  típustávolságok számszerű értékeit illeti (a típusparaméter-értékek függvényében), arra nézve a 2. táblázat második oszlopa nyújt felvilágosítást.

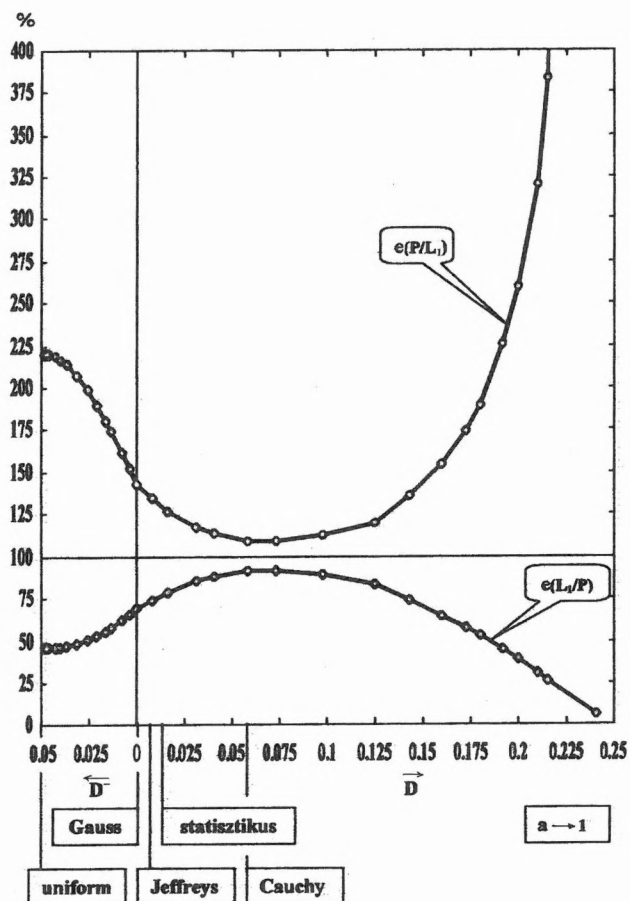
Ennyi előkészítés után már ideje rátérni a  $P$ -norma minimumhelyének meghatározására épülő statisztikai eljárások relatív hatásfokának a meghatározására, mégpedig a

modern statisztika gyakorlatában mind ez ideig leginkább alkalmazott,  $L_1$ -bázisú procedúrákhoz viszonyítva. Ahogyan a 2. pontban is, itt is célszerű feltételeznünk, hogy  $n$  db  $x_i$  adatunk egyetlen ismeretlen értékre vonatkozó közvetlen mérés eredménye. Ebben az esetben a

$$P = \min . \quad (32)$$

általános követelményt teljesítő értéket leggyakoribb értéknek nevezzük és  $M$ -mel jelöljük (*most frequent value*). Mivel a (32) követelményben szereplő  $P$ -norma (13)-beli alakja a minimumhelyet változtatlanul hagyó átalakításokkal egyszerűsödik a (7) bal oldalán álló szorzat-kifejezésre, a jelenleg  $(x_i - M)$  alakú  $X_i$ -vel a következő követelményt kell teljesítenünk:

$$\prod_{i=1}^n [4\varepsilon^2 + (x_i - M)^2] = \min . \quad (33)$$



2. ábra. Az  $e(P/L_1)$  és  $e(L_1/P)$  relatív hatásfok görbéi a harang alakú hibaeloszlás-típusok teljes  $f_c(x)$  szupermodelljére, a Gauss-eloszlástól mért típus-távolságok függvényében. A  $D^-$  és  $D$  típus-távolságokat a (30) és (31) formula definiálja, számszerű értékeit a 2. táblázat második oszlopa tartalmazza

Fig. 2. Curves of the relative efficiencies  $e(P/L_1)$  and  $e(L_1/P)$  for the complete supermodel  $f_c(x)$  of the bell-shaped error distributions, vs the type-distances from the Gaussian type. The type-distances  $D^-$  and  $D$  are defined in Eqs 30 and 31, their numerical values are given in the second column of the Table 2

A bal oldal logaritmusát véve továbbra is változatlan marad a minimumhely, így  $M$ -re vonatkozó követelésünk a

$$\frac{\partial}{\partial M} \sum_{i=1}^n \ln[4\varepsilon^2 + (x_i - M)^2] = 0 \quad (33)$$

egyenletet kielégítő  $M$ -mel teljesül. Deriválás és átrendezés után (iteratív utasításként is felfogható) egyenletet kapunk az  $M$ -mel jelölt leggyakoribb értékre:

$$M = \frac{\sum_{i=1}^n \frac{x_i}{4\varepsilon^2 + (x_i - M)^2}}{\sum_{i=1}^n \frac{1}{4\varepsilon^2 + (x_i - M)^2}} . \quad (35)$$

Az $f_c(x)$ szupermodell típus-paraméterei	A Gauss-típustól mért típus-távolságok ( $D^-$ és $D$ )	$e(P/L_1)$ (%)	$e(L_1/P)$ (%)	Az $\varepsilon$ dihézió értékei standard esetben
$p=\infty$	0,04804	219,92	45,47	0,7185223
$p=200$	0,04745	219,90	45,48	0,7356970
$p=100$	0,04686	219,85	45,49	0,7481034
$p=50$	0,04569	219,64	45,53	0,7682812
$p=20$	0,04222	218,28	45,82	0,8122672
$p=15$	0,04031	217,02	46,08	0,8307641
$p=10$	0,03656	213,70	46,79	0,8600029
$p=20/3$	0,03119	207,00	48,31	0,8907265
$p=5$	0,02584	198,85	50,29	0,9108802
$p=4$	0,02088	189,82	52,68	0,9236502
$p=10/3$	0,01619	180,37	55,44	0,9309372
$p=3$	0,01322	173,99	57,48	0,9333825
$p=5/2$	0,00762	161,37	61,97	0,9336774
$p=20/9$	0,00370	152,20	65,70	0,9306420
$p=2$	0,00000	143,41	69,73	0,9253681
$a=9$	0,00804	134,26	74,48	0,3346663
$a=5$	0,01601	126,96	78,77	0,4817797
$a=3$	0,03130	117,24	85,30	0,6972796
$a=2.5$	0,04082	113,41	88,17	0,8119323
$a=2$	0,05809	109,66	91,19	1,0000000
$a=1,75$	0,07309	109,04	91,71	1,1502030
$a=1,5$	0,09746	112,46	88,92	1,3776428
$a=1,25$	0,14249	135,60	73,74	1,7665380
$a=1,15$	0,17225	174,20	57,41	2,0133731
$a=1,1$	0,19150	225,38	44,37	2,1708114
$a=1,05$	0,21536	383,16	26,10	2,3601228
$a=1,01$	0,24077	1661,6	6,2	2,5415764
$a \rightarrow 1$	0,25000	$\infty$	0,0	2,5920401

2. táblázat. A harang alakú hibaeloszlás-típusok (29)-ben definiált teljes  $f_c(x)$  szupermodelljének 30-féle eloszlására vonatkozóan az  $e(P/L_1)$  és  $e(L_1/P)$  relatív hatásfok-értékek mellett a táblázat megadja a  $D^-$  és  $D$  típus-távolságokat (ld. a (30) és (31) definíciót), valamint a (42)-et kielégítő,  $\varepsilon$ -nal jelölt dihézió-értékeket is

Table 2. In this table are given for 30 different bell-shaped error distribution types from the complete supermodel  $f_c(x)$  (defined in Eq. 29) not only the relative efficiencies  $e(P/L_1)$  and  $e(L_1/P)$  but also the type-distances ( $D^-$  and  $D$ , see Eqs. 30 and 31) and the values of the dihesion (denoted by  $\varepsilon$ ) fulfilling Eq. 42

Az  $\varepsilon$  dihézióra vonatkozó általános (6) formulát is célszerű most az  $X_i = x_i - M$  figyelembevételével átírni:

$$\varepsilon^2 = 3 \cdot \frac{\sum_{i=1}^n \frac{(x_i - M)^2}{[\varepsilon^2 + (x_i - M)^2]^2}}{\sum_{i=1}^n \frac{1}{[\varepsilon^2 + (x_i - M)^2]^2}}. \quad (36)$$

Az  $x_i$  ( $i=1, \dots, n$ ) minta azt az  $M$  leggyakoribb értéket és  $\varepsilon$  dihéziót definiálja, amely értékpár egyidejűleg teljesíti (35)-öt és (36)-ot. Ha ismert az  $f(x)$  sűrűségfüggvény, azaz nem minta alapján történik az  $(M, \varepsilon)$  értékpár meghatározása, akkor a fenti összeg-formulák helyébe integrál-formulák lépnek:

$$M = \frac{\int_{-\infty}^{\infty} \frac{x}{4\varepsilon^2 + (x-M)^2} f(x) dx}{\int_{-\infty}^{\infty} \frac{1}{4\varepsilon^2 + (x-M)^2} f(x) dx} \quad (37)$$

és

$$\varepsilon^2 = 3 \cdot \frac{\int_{-\infty}^{\infty} \frac{(x-M)^2}{[\varepsilon^2 + (x-M)^2]^2} f(x) dx}{\int_{-\infty}^{\infty} \frac{1}{[\varepsilon^2 + (x-M)^2]^2} f(x) dx}. \quad (38)$$

Az e két formulát egyidejűleg teljesítő  $(M, \varepsilon)$  értékpár az  $\varepsilon$  dihézió és az  $M$  leggyakoribb érték pontos értékeit adja; a mintából (35) és (36) egyidejű teljesítésével ezen pontos értékek annál jobb közelítéseit kapjuk, minél nagyobb az  $n$  mintaelemszám. Ha az  $n$  már valóban eléggé nagy, akkor a (35) szerint meghatározott  $M$  becslések  $\sigma_M$  szórása

$$\sigma_M = \frac{A_M}{\sqrt{n}} \quad (39)$$

nagyságú, ahol  $A_M$ -et a leggyakoribb érték aszimptotikus szórásának nevezzük és a következő formula szerint számíthatjuk origóra szimmetrikus  $f(x)$ -ek esetén:

$$A_M^2 = \frac{\int_{-\infty}^{\infty} \frac{x^2}{[4\varepsilon^2 + x^2]^2} f(x) dx}{\left[ \int_{-\infty}^{\infty} \frac{4\varepsilon^2 - x^2}{[4\varepsilon^2 + x^2]^2} f(x) dx \right]^2} \quad (40)$$

(ld. pl. STEINER [1990] (5-90) formuláját). A (29) egyenletben definiált  $f_c(x)$  szupermodell mindegyik típusa szimmetrikus az origóra (hiszen standard alakban definiáltuk azokat), így (40)-ben  $f(x)$  helyébe  $f_c(x)$ -et írva, kezünkben van az az  $A_M^2$  aszimptotikus variancia, amellyel az

$$e(P/L_1) = \frac{A_{med}^2}{A_M^2} \quad (41)$$

relatív hatásfokokat bármelyik  $f_c(x)$  harang alakú hibatípusra számítani tudjuk. (A (40)-ben szereplő  $\varepsilon$ -t (38) szerint határozzuk meg, mivel azonban  $f_c(x)$  az origóra

szimmetrikus, a medián és az  $M$  leggyakoribb érték egyaránt zérussal egyenlő, így most az

$$\varepsilon^2 = 3 \cdot \frac{\int_{-\infty}^{\infty} \frac{x^2}{(\varepsilon^2 + x^2)^2} f(x) dx}{\int_{-\infty}^{\infty} \frac{1}{(\varepsilon^2 + x^2)^2} f(x) dx} \quad (42)$$

egyenletet kielégítő  $\varepsilon$ -t kell csak meghatároznunk, pl. magát (42)-t iterációs utasításként alkalmazva. Az így adódó  $\varepsilon$ -ok számszerű értékeit a 2. táblázat utolsó oszlopa tartalmazza.)

Ami az  $A_{med}^2$ -et illeti, a medián elméleti értéke zérus lévén minden (29) szerinti  $f_c(x)$ -re, a (20) változatlanul fennáll, azaz  $A_{med} = 1/[2 \cdot f_c(0)]$ . Az  $f_a(x)$  szupermodellre (amely (29) szerint része a teljes  $f_c(x)$  típuscsaládnak,) a (25) egyenlet szerint

$$A_{med}^2 = \frac{\pi \cdot \Gamma^2[(a-1)/2]}{4 \cdot \Gamma^2(a/2)}; \quad (43)$$

az  $f_p(x)$ -re pedig már (22)-ben megismertük az  $A_{med}^2$  kifejezését (a (29) szerint a  $\infty > p \geq 2$  típusparaméter-tartományra az  $f_p(x)$ -ek adják az  $f_c(x)$  szupermodell típusainak másik részét). Semmi akadályja nincs tehát annak, hogy az  $e(P/L_1)$  relatív hatásfokot (vagy annak reciprokát,  $e(L_1/P)$ -t) a (41) szerint meghatározhassuk. A relatív statisztikai hatásfokok számértékeit (százalékokban) a 2. táblázat harmadik és negyedik oszlopa tartalmazza; szemléltetésül az  $e(P/L_1)$  és  $e(L_1/P)$  görbéit a Gauss-féltől mért típusávolság ( $D^-$  és  $D$ ) függvényében (ld. a 2. táblázat második oszlopát és a (30) és (31) definíciókat) a 2. ábra mutatja be. Az  $e(P/L_1)$  relatív hatásfokok mindenütt (a teljes  $f_c(x)$  szupermodellre) határozottan nagyobbak 100%-nál: *a P-norma alkalmazása bármely  $f_c(x)$  hibatípus esetén előnyösebb az  $L_1$ -norma alapul vételénél.* Mivel a relatív hatásfokok azonosak az azonos pontosságot szolgáltató mérési adatszámok arányával (ld. pl. STEINER [1990] 191. oldalán az (5-103) definíciót követő szöveges részt), ez nem jelent kevesebbet, mint pl. azt, hogy a DUTTER [1986/87] szerint a földtudományokban leginkább várható,  $a=5$ -höz tartozó  $f_{sr}(x)$  hibaeloszlás esetén (standard esetre vonatkozó képletére vonatkozóan ld. (26)-ot,) az  $L_1$ -norma 27%-kal több adatot igényel azonos pontossághoz, mint a  $P$ -norma alapulvétele. Ugyanez az igényelt adatszám-többlet Gauss-típusú hibánál már 43,4%, és a hibaeloszlás szárnyainak további csökkenésével (azaz  $D^-$  növekedésével) egyre nagyobb lesz (a hatásesethez közelítve a 100%-ot is meghaladja). Hasonló a helyzet az extrém nagy súlyú szárnyak (nagy  $D$  értékek) esetén, ez azonban nem meglepő, mert  $L_1$  alkalmazásakor az  $e_{L_1}$  abszolút hatásfok is zérushoz tart, ha  $D \rightarrow 0,25$ , azaz  $a \rightarrow 1$  (ld. STEINER (ed.) [1997] 128. oldalán a (6-24) formulát), míg ugyanebben az esetben az  $e_p$  abszolút hatásfok határértéke még mindig 48% (ld. az imént idézett oldalon a (6-23)-at követő sorokat).

A fenti, lényegében a  $P$ -normán alapuló eljárásoknak az  $L_1$ -bázisúaknál gazdaságosabb voltát igazoló vizsgálataink során nem voltunk tekintettel a kiugrókkal (outlierekkel) szembeni viselkedés esetleges eltéréseire, azaz a rezisztenciára. A kiugrók előfordulása azonban végtelenül sokféle lehet, így kvantitatív összehasonlítás csak konkrét szituációkra végezhető el. Egy példa: TARANTOLA [1987] egy megjegyzése nyomán (ld. az idézett könyv 613. oldalát) a STEINER (ed.) [1997] 175. oldali ábráján különböző helyparaméter-meghatározási eljárások  $T$ -vel jelölt torzulásainak görbéi láthatók  $p$  függvényében arra az esetre, ha a kiugrók aszimmetrikusan, egyetlen  $p/(1-p)$  súlyú Cauchy-szárnyat képezve jelentkeznek. Az „ $M k = 2$ ” jelű görbe felel meg a  $P$ -norma alkalmazásának, a „med” jelű nyilván az  $L_1$ -ének; az utóbbi esetben (az ábrázolt  $0 \leq p \leq 0,1$  tartományon) a torzulások minden esetben nagyobbak voltak, mint a  $P$ -norma alkalmazásakor. A STEINER (ed.) [1997] 246. oldalán levő ábra azt mutatja, hogy a kiugrók extrém nagy aránya esetén az  $L_1$ -gyel kapott eredmények teljesen irreálisak lehetnek, míg a  $P$ -normával kapottak torzulása 10% alatt marad a vizsgálat tárgyát képező inverziós feladatnál. Ismételten le kell azonban szögezni, hogy a rezisztenciára felhozott akárhány példa sem szolgáltatna kvantitatív összehasonlítást. Az  $L_1$ - és  $P$ -eljárások általában egyaránt rezisztenseknek tekinthetők: *a  $P$  és  $L_1$  közötti kvantitatív összehasonlítás alapjául csak a relatív hatásokok szolgálhatnak*, márpedig ezek az  $e(P/L_1)$ -gyel jelölt értékek (ld. még egyszer a 2. táblázatot, ill. a 2. ábrát,) egyértelműen a  $P$ -norma alapul vételének szignifikáns előnyeit bizonyítják.

- DUTTER R. 1986/87: *Mathematische Methoden in der Montangeologie*. Vorlesungsnotizen, Manuscript, Leoben
- KIS M. 1996: Geofizikai adatok globális optimalizációja a simulated annealing módszer alkalmazásával. *Magyar Geofizika* **37**, 3, 170–181
- STEINER F. 1990: *A geostatisztika alapjai*. Tankönyvkiadó, Budapest, 363 p.
- STEINER, F. (ed.) 1997: *Optimum Methods in Statistics*. Akadémiai Kiadó, Budapest, 370 p.
- STEINER F., HAJAGOS B. 1998: Error-types characterized by arbitrary short or heavy flanks. *Acta Geod. Geoph. Hung.* **33**, 2–4, 259–277
- STEINER F., HAJAGOS B. 2000a: Relative efficiency of the conventional statistics compared to the  $L_1$ -norm based statistical procedures, tested on distribution types of the generalized Gaussian supermodel (közlésre elfogadva az *Acta Geod. Geoph. Acad. Sci. Hung.* folyóiratban)
- STEINER F., HAJAGOS B. 2000b:  $P$ -norm based statistical procedures are more efficient than the  $L_1$ -based ones for all error-types of the complete supermodel  $f_c$  (közlésre elfogadva az *Acta Geod. Geoph. Acad. Sci. Hung.* folyóiratban)
- SZÜCS P. 1997a: Theoretical and practical consequences of the global optimization methods. Appendix V. In: STEINER (ed.): *Optimum Methods in Statistics* Akadémiai Kiadó, Budapest, 303–311
- SZÜCS P. 1997b: Joint inversion of well log data minimizing the  $P$ -norm. 10.3 alfejezet. In: STEINER (ed.): *Optimum Methods in Statistics*, Akadémiai Kiadó, 257–270
- TARANTOLA A. 1987: *Inverse Problem Theory*. Elsevier, Amsterdam, 613 p.