



11. előadás

Kiegyenlítő számítások MSc

2018/19

Áttekintés

- Robusztus és rezisztens mérés feldolgozás
 - Bevezetés
 - Robusztus és rezisztens becslések
 - Nagy számok törvényének teljesülése
 - Korreláció és regresszió

Robusztus statisztika

- „A robusztus statisztika olyan elméleti keret, amelyen belül *gazdaságosan* oldható meg az a feladat, hogy egymástól jelentősen *eltérő eloszlástípusok* esetén is megbízható eredményt érjünk el, valamint hogy ne legyünk kitéve a *durva hibájú adatok* torzító (esetleg katasztrofális mértékben torzító) hatásának.” (Steiner, 1990)

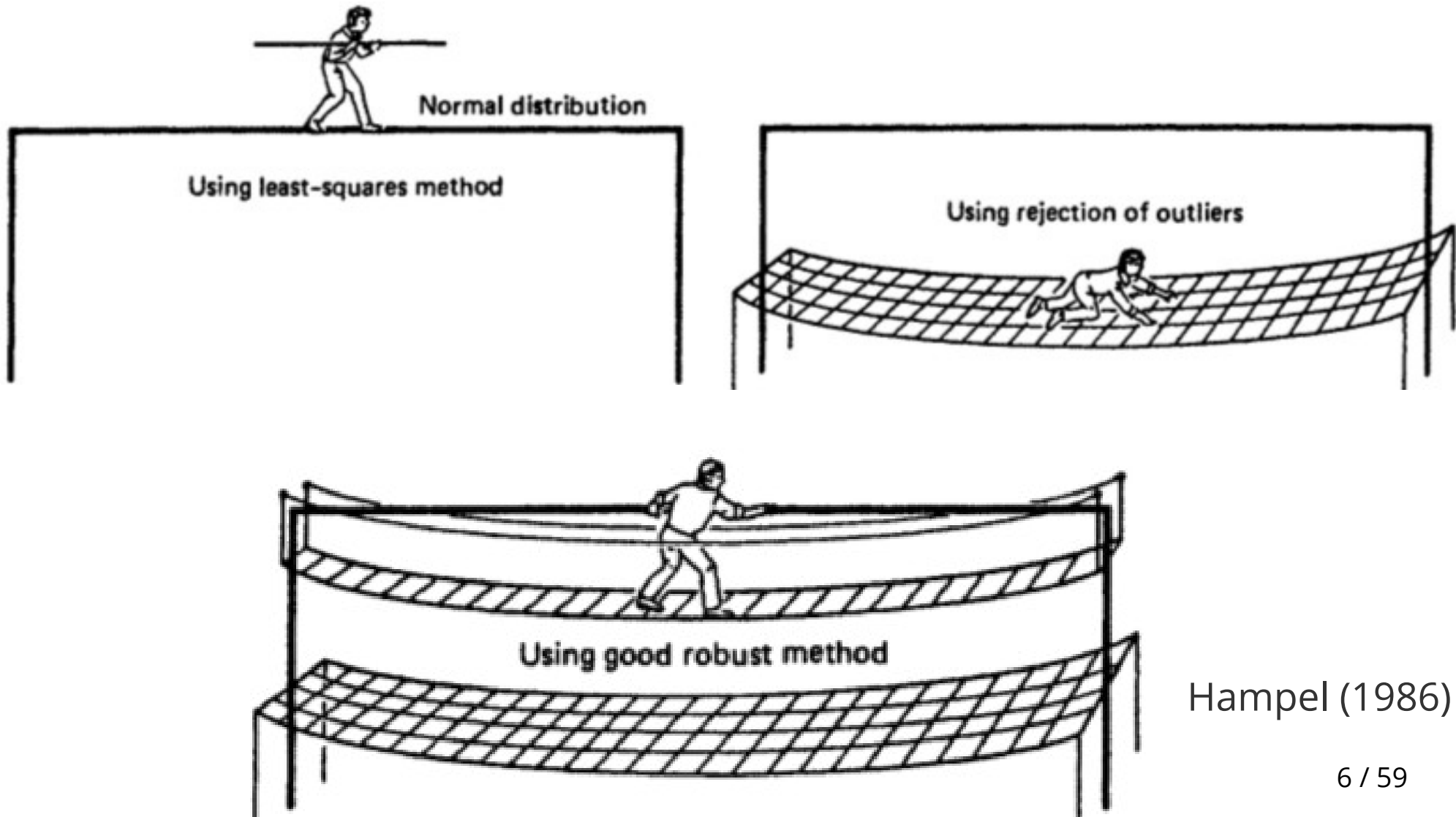
Miért fontos a robusztus mérés feldolgozás?

- A legtöbb feltevés csak közelíti a valóságot, mert
 - az adatok nem Gauss eloszlásúak
 - a modellek nem lineárisak
 - a mérések nem függetlenek
- Durva hibák is jelentkeznek, és
 - tönkreteszhetik a feldolgozás eredményét
 - szubjektív szempontok alapján történő elvetésük kérdéses

Miért fontos a robusztus mérésfeldolgozás?

- Feltételezzük a nagy számok törvényének teljesülését, DE
 - a tapasztalat szerint nem minden esetben igaz
- A klasszikus mérésfeldolgozási eljárások optimális becslést adnak jól meghatározott paraméteres modellekre, DE:
 - nem foglalkoznak azzal, ha a modell csak közelítőleg igaz
 - gyakran rosszul teljesítenek, ha akár csak kis eltérés is van a modellben

Az adatok elemzésének különböző módszerei

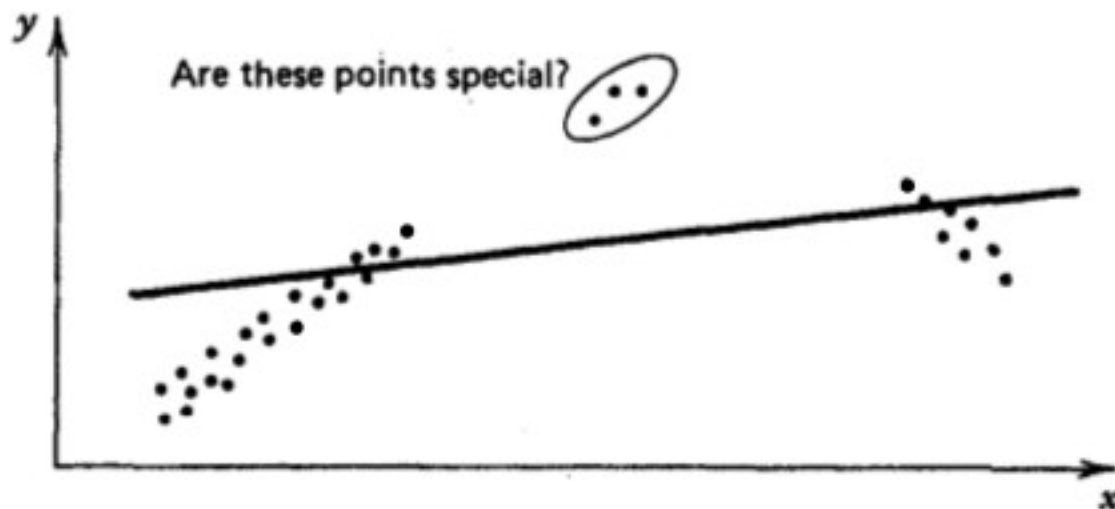


Hampel (1986)

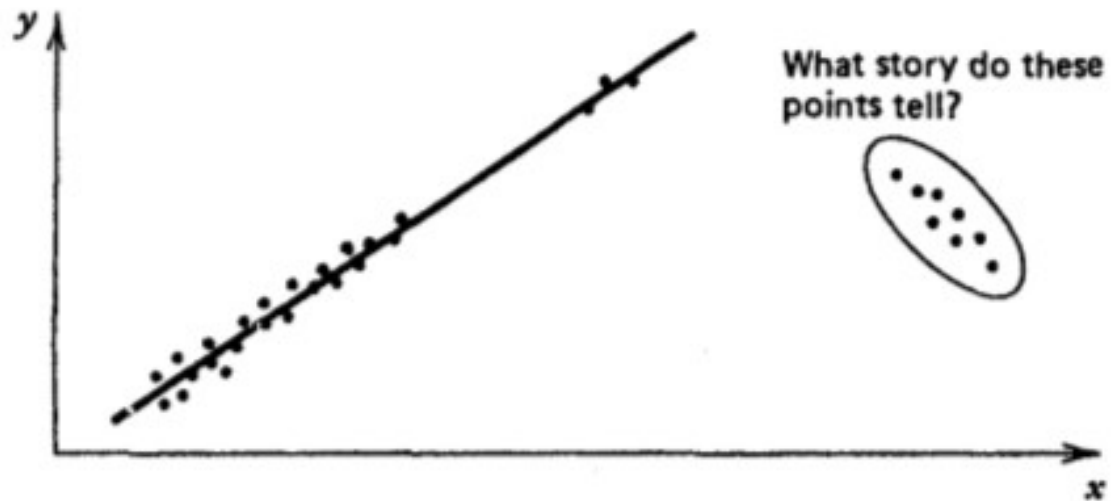
A robusztus adatfeldolgozás célja

- az adatok zöméhez legjobban illeszkedő struktúra megkeresése
- rendellenes, eltérő pontok (durva hibák) vagy részstruktúrák azonosítása
- azoknak az adatoknak az azonosítása, amelyek nagymértékben befolyásolják az eredményt

Melyik illesztést akarjuk?



(a) Least-squares fit: average opinion of all points (noisy)



(b) Highly robust fit: clear opinion of majority of points

Hampel (1986)

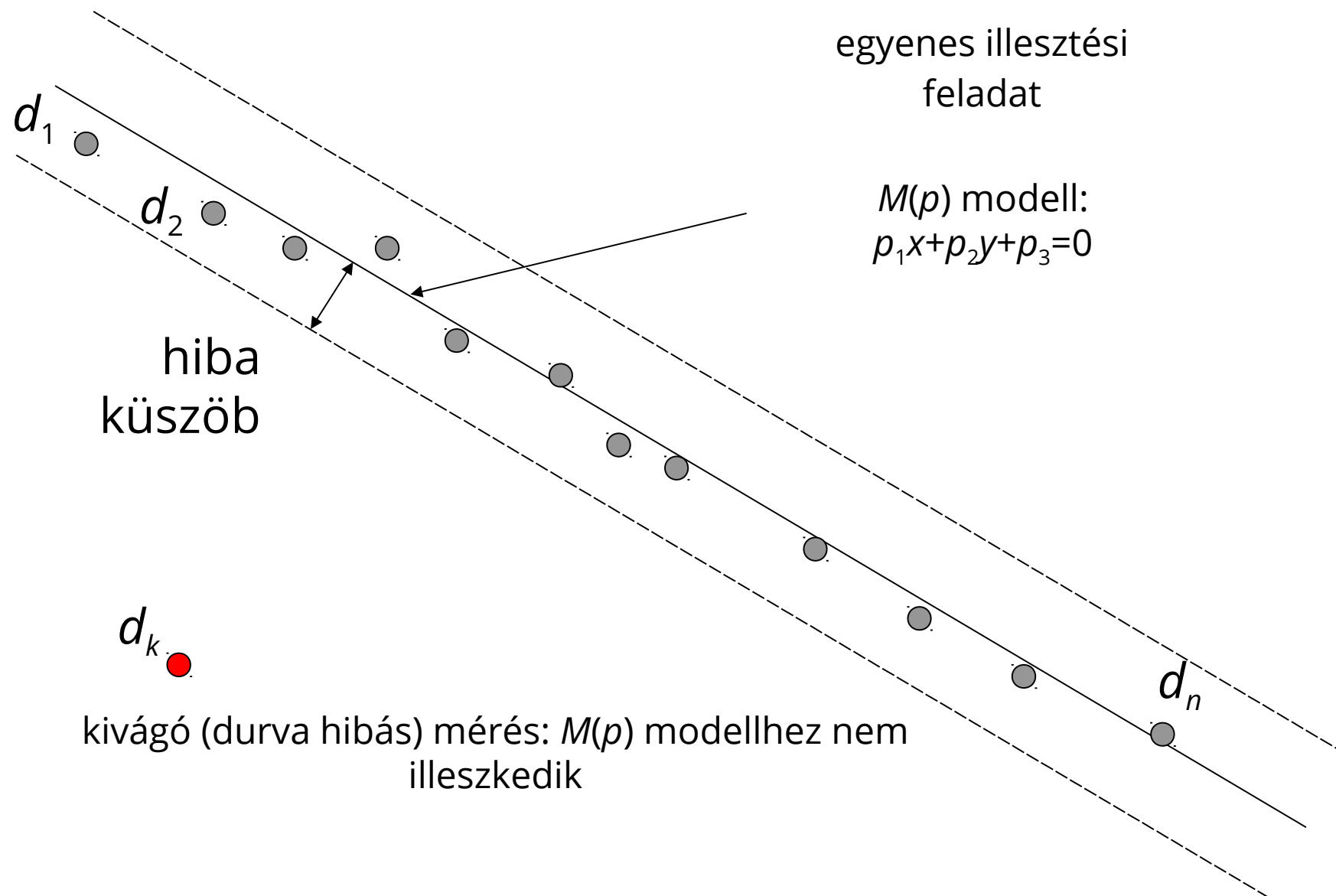
Rezisztencia és robusztusság

- Egy eljárás akkor **rezisztens**, ha indokolatlanul nem befolyásolja néhány kivágó érték (Wilks, 1995)
- Egy **robusztus** eljárás nem érzékeny az ideálistól kissé eltérő körülményekre, amely ideális körülményekre az eljárás optimális (Hampel, 1986)
 - „kissé eltérő”:
 - az összes pontban vannak kis eltérések
 - kevés pontban vannak nagy eltérések

Kivágó érték fogalma

- Adott egy M modell és annak p („valódi”) paraméterei, amely (az elkerülhetetlen mérési hibáktól eltekintve) tökéletesen képes reprodukálni a d méréseket (d : „data”)
- Egy d mérés kivágó értéknek (durva hibásnak) tekinthető, ha nem illeszkedik valamely hiba küszöbön belül M -hez.
 - Ezt a hiba küszöböt a véletlen jellegű mérési hibák hatásának tulajdonítható maximális eltérés definiálja
 - A véletlen jellegű mérési hibák *eloszlása* ismert
 - Kapcsolódik a *rezisztencia* és a *robosztusság* fogalmaihoz

Modell és mérések



A becslés torzítása

- D_I legyen a nem kivágó d mérések halmaza
- $D_{I/O}(m)$ legyen olyan d mérések halmaza, amelyben m darab nem kivágó mérést kicseréltünk kivágó (durva hibás) mérésekre
- (I : inlier, O : outlier)
- Legyen $M(p)$ a paraméteres modell
- Az $M(p)$ paraméteres modellen alapuló paraméter becslés torzítása az a maximális zavarás, amely a paraméter vektor becslésében jelentkezik akkor, ha D_I -t kicseréljük $D_{I/O}(m)$ -re

A becslés összeomlási pontja

- A becslés *összeomlási pontja* a kivágó (durva hibás) méréseknek az a *minimális* aránya (százaléka), amelyet elérve a becslés torzítása akár *tetszőlegesen nagy* is lehet
- Megmutatható, hogy a *legkisebb négyzetek* szerinti paraméter becslés összeomlási pontja 0 %
 - Akár *egyetlen* kivágó mérés is tetszőleges mértékben képes elrontani a LKN becslést!

Robusztus és rezisztens becslés

- A modelltől való eltérések nem hagyhatók figyelmen kívül a gyakorlatban
 - P.J. Huber: 5-10%-os kivágó érték (durva hiba) inkább szabálynak látszik, nem kivételnek
- Olyan becslést szeretnénk, amely nem érzékeny az $M(p)$ paraméteres modelltől való eltérésekre:
 - megváltozik a véletlen mérési hibák eloszlása, de nem változik lényegesen a becslés és pontossága: *robosztusság*
 - nem illeszkedő, kivágó adatok lépnek fel, de nem változik lényegesen a becslés: *rezisztencia*
 - a LKN becslés sem nem robusztus, sem nem rezisztens

A becslés abszolút hatásfoka

- Az e abszolút hatásfok a Cramér-Rao határhoz viszonyított aszimptotikus szórásnégyzet az adott becslési eljárásra (viszonysszámként $e \leq 1$ vagy százalékban kifejezve $e \leq 100\%$)

$$e = \frac{A_{\min}^2}{A^2}$$

- Az aktuális $f(x)$ -hez tartozó optimális eljárással csak e -szer annyi adat kell ugyanakkora pontossághoz

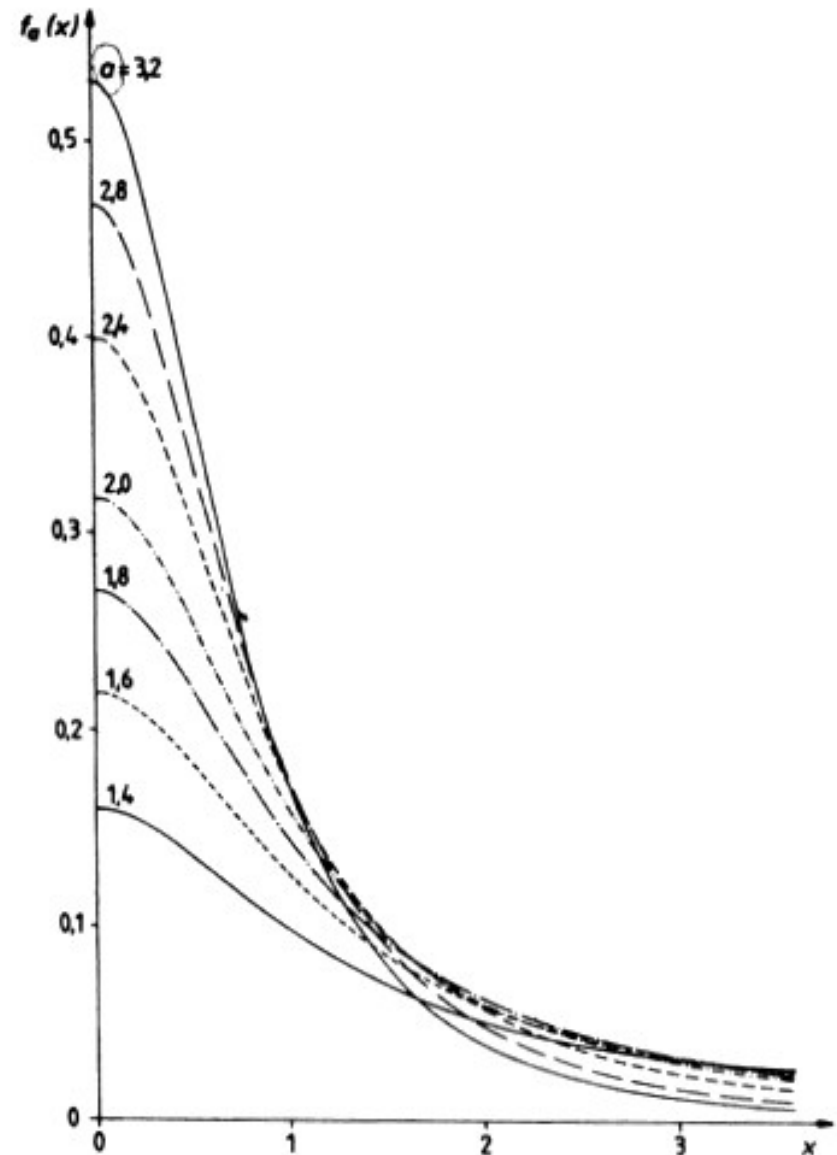
Robusztusság

- Valamely becslés akkor nevezhető robusztusnak, ha az *eloszlástípusok széles tartományán* a gazdaságossága csak jelentéktelen mértékben csökken
- Egy statisztikai algoritmus akkor robusztus, ha az *anyaeloszlásban* bekövetkező *kicsiny eltérés* a *becslések eloszlásában* is *csak kis eltérést* eredményez

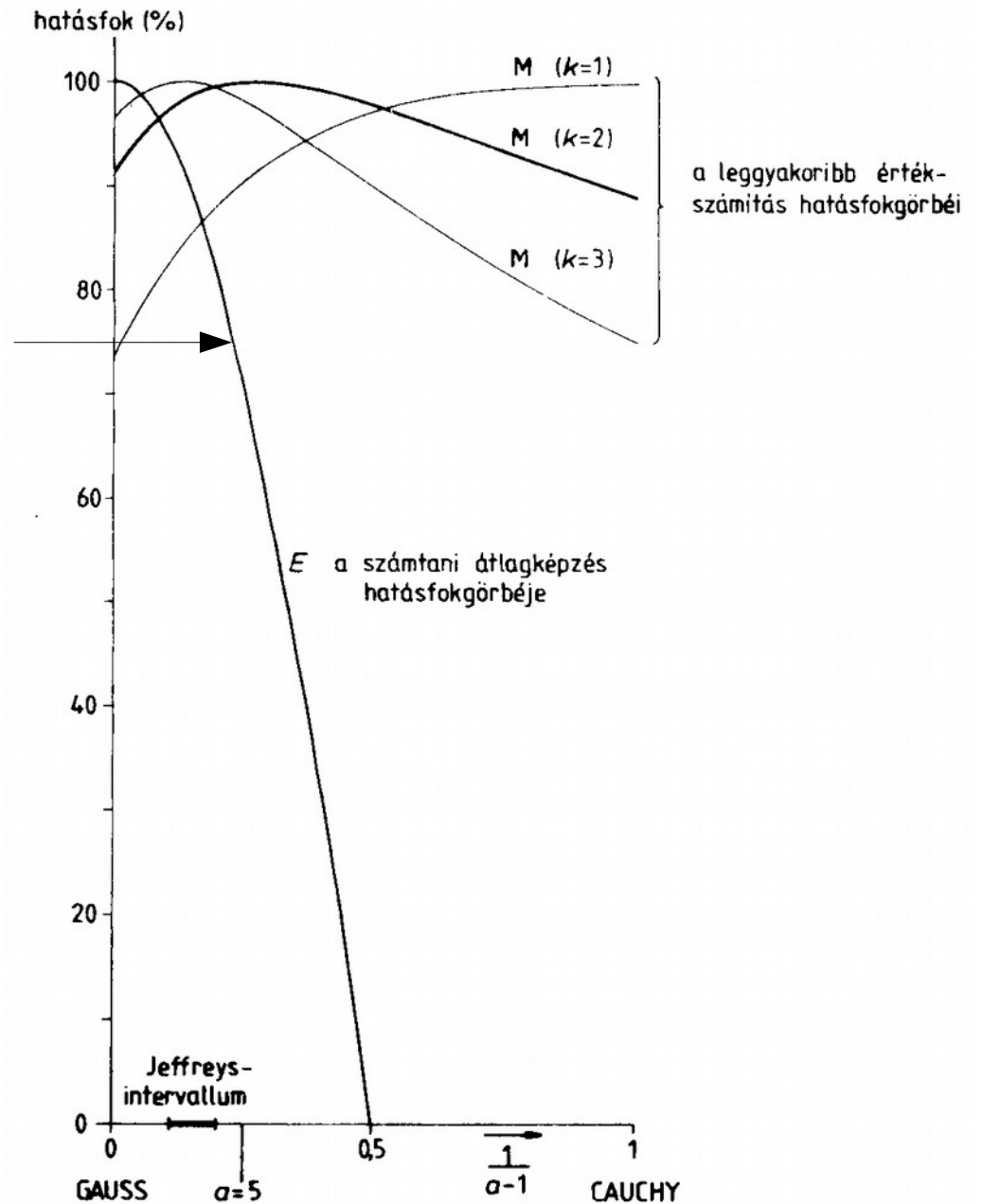
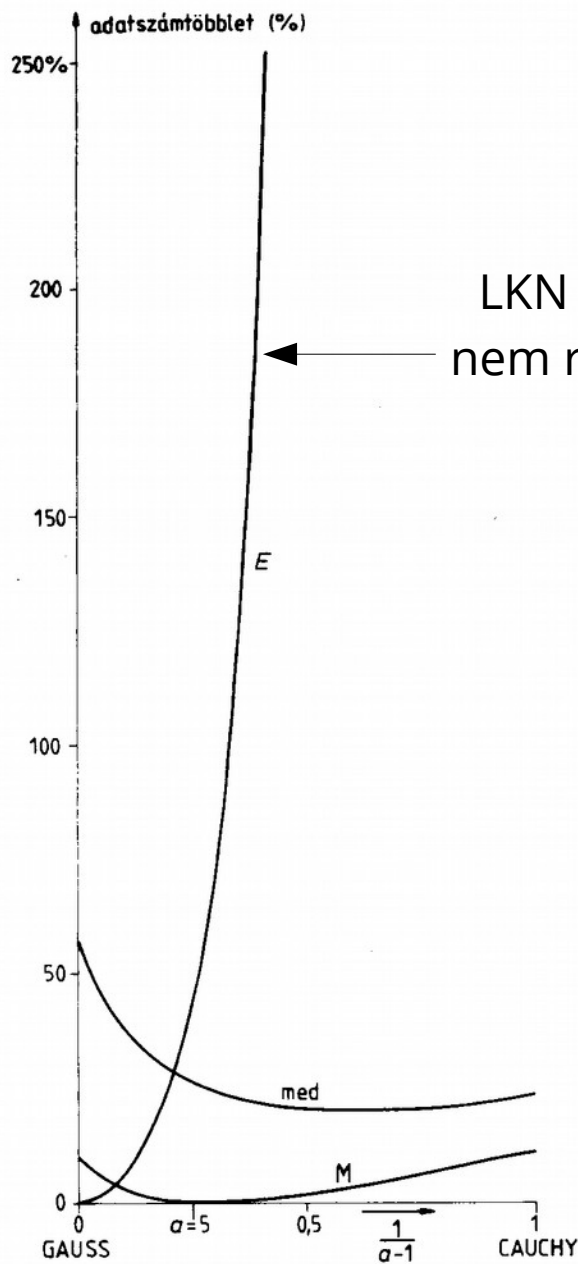
Az $f_a(x)$ szupermodell

- a : típusparaméter
- $a \rightarrow \infty$: Gauss-eloszlás
- $a = 2$: Cauchy-eloszlás
- $a = N + 1$:
 N szabadságfokú
Student-eloszlás

$$\frac{1}{a-1} = (0, 1) = (\text{Gauss}, \text{Cauchy})$$



Robusztusság különböző becslésekre az $f_a(x)$ szupermodellre



Rezisztencia számszerűsítése – *IC*-függvény

- Milyen mértékben módosítja egyetlen x adat a helyparaméter becslés T eredményét?
- a válasz függ:
 - a becslés algoritmusától (legyen ez is T)
 - az aktuális F eloszlástól
 - az adat x értékétől
- $IC(x, F, T)$ hatásgörbe (*influence curve*, *IC*-görbe, *IC*-függvény adja meg a választ

Az IC -függvény definíciója

$$IC(x, F, T) = \lim_{t \rightarrow 0} \frac{T[(1-t) \cdot F + t \cdot H(x)] - T(F)}{t}$$

$H(x)$ az egységugrás (Heaviside) függvény

- Az IC -függvény megadja egyetlen, x értékű járulékos észlelésnek a T értékváltozásában megnyilvánuló hatását
- $T(F)$ az eredeti F eloszlás alapján becsült T paraméter
- $[(1-t) \cdot F + t \cdot H(x)]$ a megváltozott eloszlás
- t a járulékos észlelés részaránya a többi adathoz képest

A becslés értékének változása

- nagy n esetén közelítőleg ΔT értékkel változik meg a T becslés értéke, ha az F -eloszlásfüggvényű eloszlásból származó n elemű mintánkhoz még egyetlen x értékű észlelést is figyelembe veszünk ($t = 1/n$):

$$IC(x, F, T) = \frac{\Delta T}{1/n}$$

$$\Delta T = \frac{IC(x, F, T)}{n}$$

Az *IC*-görbe alkalmazása

- Közvetlen számszerű információt ad arról, hogy a durva hibájú adatok milyen mértékben torzítják az adott algoritmus szerint számított hely (vagy skála) paraméter becslés értékét (**rezisztencia**)
 - a számtani átlag *IC*-függvénye x -szel egyenlő
 - nem rezisztens
- a minta egy x értékű eleme milyen mértékben vesz részt a helyparaméter becslés értékének kialakításában (**robosztusság**)

A becslés pontosságának növekedése az n mintaelemszámmal

- Az n mintaelemszám növekedésével pontosságnövekedést várunk (a bizonytalanság csökkenését)
 - Ha ez valóban teljesül, azt mondjuk, hogy „teljesül a nagy számok törvénye”
- Nem robusztus becslési algoritmusoknál előfordulhat, hogy a pontosság *romlik* n növekedésével („fordítva teljesül a nagy számok törvénye”)

Becslések határeloszlásai

- Ha a becslések eloszlása növekvő n mintaelemszám esetén egy eloszlástípust közelít, ezt *határeloszlásnak* hívjuk
- A határeloszlás kiszámításához már a legegyszerűbb becslés: a számtani átlagképzés esetén is szükség van a *karakterisztikus függvény* fogalmára
- A számtani átlagok határeloszlása véges σ esetén elvezet a *centrális határeloszlástétel*hez
- Az átlagok előbb-utóbb mindig $1/\sqrt{n}$ szerint válnak pontosabbá? Egyáltalán pontosabbá válnak minden esetben?

Számtani átlag határeloszlása

- Összeg és átlag sűrűségfüggvénye
- A karakterisztikus függvény fogalma
- Számtani átlagok határeloszlása véges σ esetén
- A centrális határeloszlástétel
- A nagy számok törvényének teljesülése és nem teljesülése

Összeg sűrűségfüggvénye

- Hogyan adódhat egy konkrét érték, pl. 0.45 két érték összegeként?

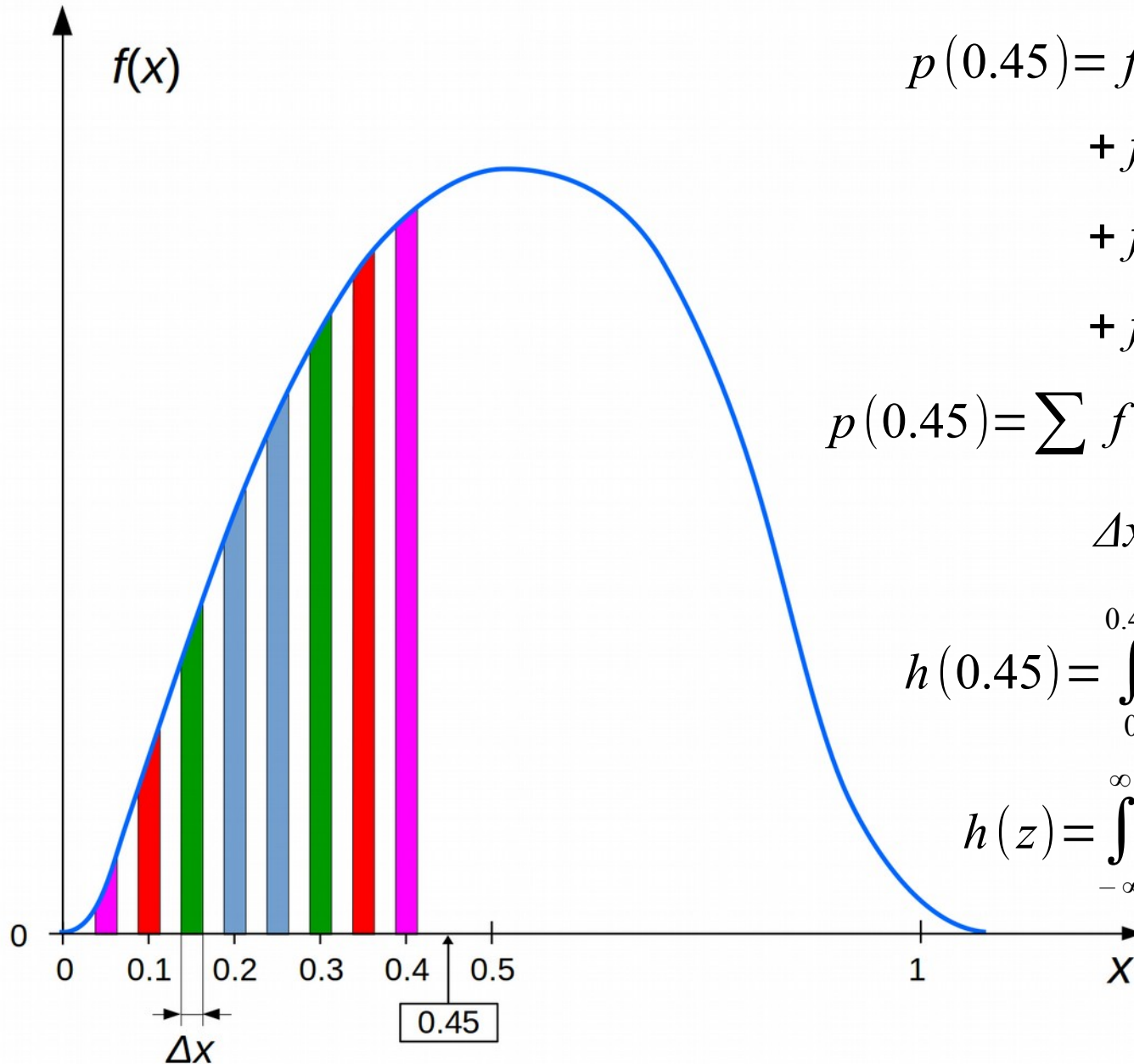
$$0.45 = 0.05 + 0.40$$

$$0.45 = 0.10 + 0.35$$

$$0.45 = 0.15 + 0.30$$

$$0.45 = 0.20 + 0.25$$

Összeg $h(x)$ sűrűségfüggvénye



$$\begin{aligned}
 p(0.45) &= f(0.05) \cdot f(0.4) \cdot (\Delta x)^2 \\
 &+ f(0.1) \cdot f(0.35) \cdot (\Delta x)^2 \\
 &+ f(0.15) \cdot f(0.3) \cdot (\Delta x)^2 \\
 &+ f(0.2) \cdot f(0.25) \cdot (\Delta x)^2
 \end{aligned}$$

$$p(0.45) = \sum f(x) \cdot f(0.45 - x) \cdot (\Delta x)^2$$

$$\Delta x \rightarrow 0$$

$$h(0.45) = \int_0^{0.45} f(x) \cdot f(0.45 - x) dx$$

$$h(z) = \int_{-\infty}^{\infty} f(x) \cdot f(z - x) dx$$

Összeg sűrűségfüggvénye

- Két különböző valószínűség sűrűségű eloszlásból az összeg sűrűségfüggvénye *konvolúcióval* számítható ki:

$$h(z) = \int_{-\infty}^{\infty} f(x) \cdot g(z-x) dx = f(x) * g(x)$$

- n tagú összegre n tényezős konvolúció:

$$h(z) = f(x) * f(x) * f(x) \cdots f(x) = [f(x)]^{n*}$$

Számtani átlag sűrűségfüggvénye

- n -el való osztással $1/n$ -szeresére csökken a sűrűségfüggvény szélessége, vagyis a skála paraméter $1/n$:

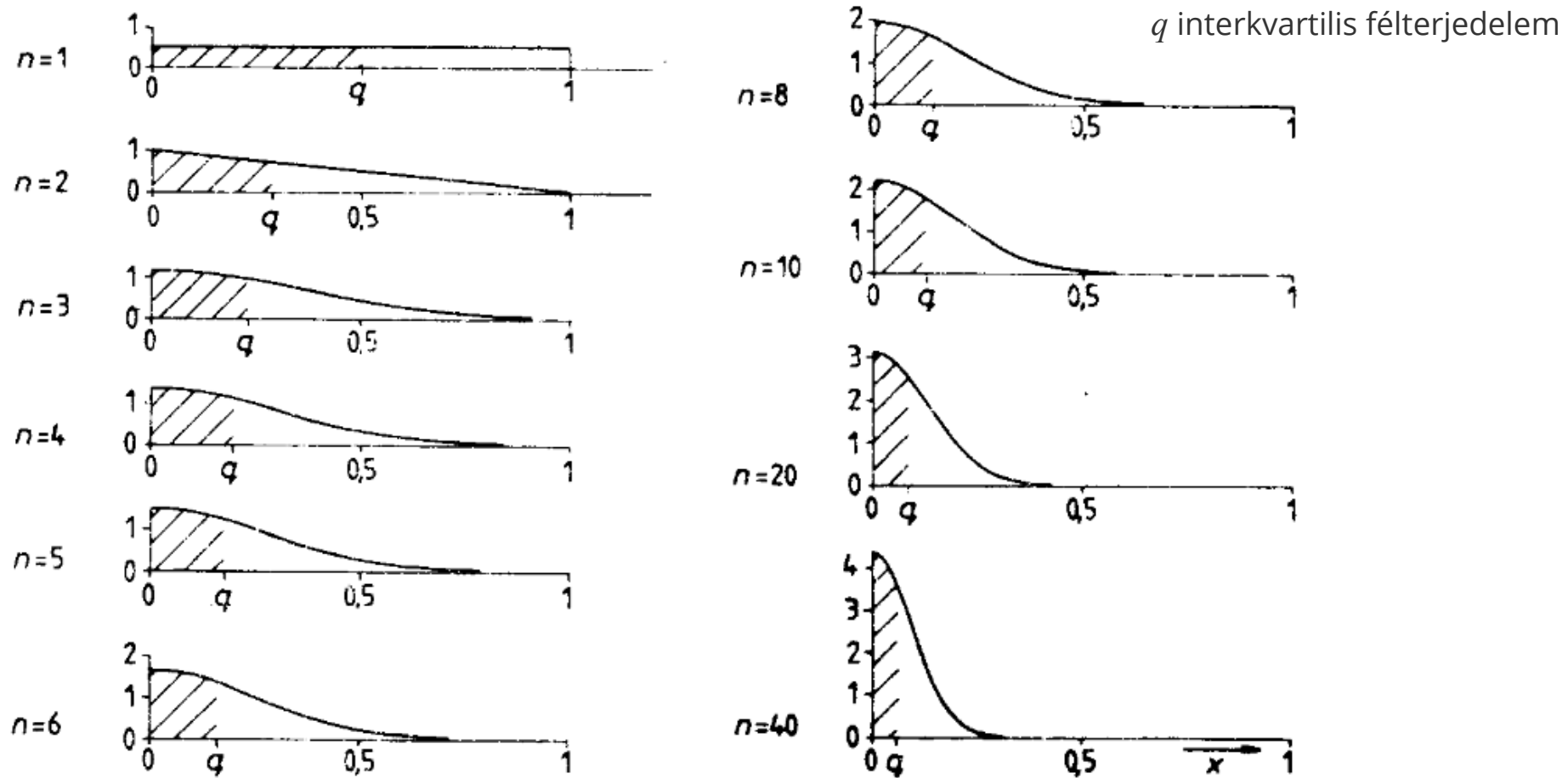
$$g^{(n)}(x) = n \cdot [f(nx)]^{n*}$$

- $[-1, 1]$ közötti egyenletes eloszlásra

$$g_u^{(n)}(x) = \frac{n}{2^n (n-1)!} \sum_{k=0}^{\text{int}\left(\frac{n+x}{2}\right)} (-1)^k \binom{n}{k} (n-2k+nx)^{n-1}$$

Számtani átlag sűrűségfüggvénye

- $[-1, 1]$ közötti egyenletes eloszlásra (Steiner, 1990)



- Gauss-eloszláshoz közeledik és q csökken

Számtani átlag sűrűségfüggvénye

- Minden esetben tapasztalható az átlagok \sqrt{n} -nel növekvő pontossága?
- Minden esetben tapasztalható az átlagok Gausshoz tartozó eloszlása?
 - A válasz mindkét kérdésre: NEM

Tanulságos példa – Tukey modell

- A kivágó értékek Tukey (1960) által bevezetett modellje
 - két Gauss-típusú sűrűségfüggvény lineáris kombinációja

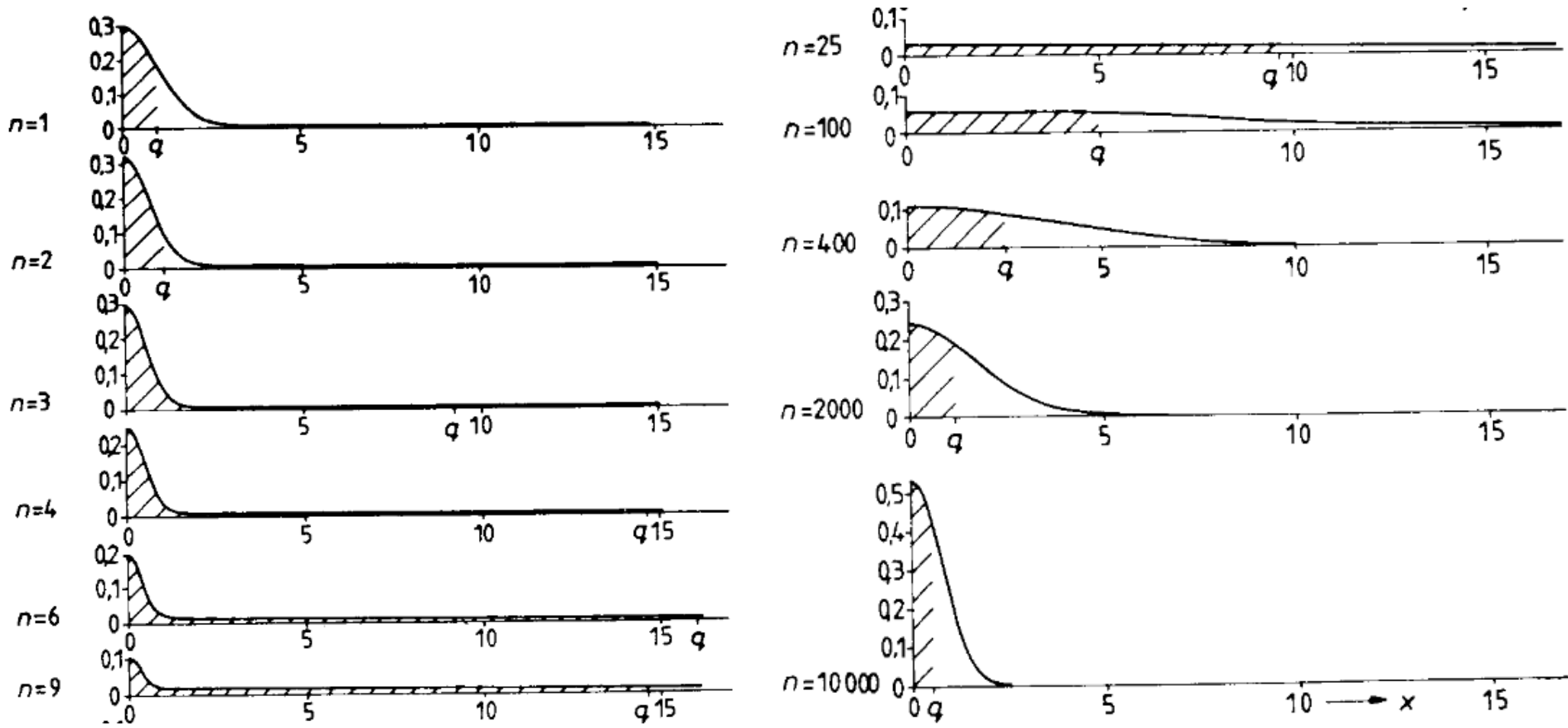
$$f_T(x) = \frac{1}{\sqrt{2\pi}} \left\{ (1-p) e^{-x^2/2} + \frac{p}{\sigma_c} \cdot e^{-x^2/(2\sigma_c^2)} \right\}$$

- n -elemű minták számtani átlagainak sűrűségfüggvénye:

$$g_T^{(n)}(x) = \frac{n}{\sqrt{2\pi}} \sum_{k=0}^n \binom{n}{k} (1-p)^{n-k} \cdot p^k \cdot \frac{1}{\sqrt{n-k+k\sigma_c^2}} \exp \left\{ -\frac{(nx)^2}{2[(n-k)+k\sigma_c^2]} \right\}$$

Tukey modell számtani átlagainak sűrűségfüggvénye

- A Tukey modell paraméterei: $p = 0.25, \sigma_C = 150$ (Steiner, 1990)



- Végül Gauss-eloszláshoz közeledik és q csökken

A karakterisztikus függvény

- Az $f(x)$ sűrűségfüggvénnyel jellemzett eloszlás *karakterisztikus függvénye*:

$$\varphi(t) = \int_{-\infty}^{\infty} e^{ixt} f(x) dx = \mathcal{F}^{-1}\{f(x)\}$$

- origóra szimmetrikus $f(x)$ -ek esetén:

$$\varphi(t) = 2 \int_0^{\infty} \cos(xt) f(x) dx$$

A karakterisztikus függvény

- Az $f(x)$ sűrűségfüggvény a karakterisztikus függvény Fourier-transzformáltja:

$$f(x) = \mathcal{F}\{\varphi(t)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ixt} \varphi(t) dt$$

- origóra szimmetrikus $\varphi(t)$ -k esetén:

$$f(x) = \frac{1}{\pi} \int_0^{\infty} \cos(xt) \varphi(t) dt$$

- S skála paraméterű sűrűségfüggvényre a karakterisztikus függvény

$$\frac{1}{S} \int_{-\infty}^{\infty} e^{i\frac{x}{S}St} f\left(\frac{x}{S}\right) dx = \int_{-\infty}^{\infty} e^{iuSt} f(u) du = \varphi(St)$$

A karakterisztikus függvények táblázata

Eloszlástípus vagy szupermodell	A $\varphi(t)$ karakterisztikus függvény (a standard alakhoz)
<i>eloszlások:</i>	
egyenletes $f_u(x)$	$\frac{\sin t}{t}$
Gauss $f_G(x)$	$e^{-t^2/2}$
Laplace $f_L(x)$	$\frac{1}{1+t^2}$
Cauchy $f_C(x)$	$e^{- t }$
exponenciális $f_e(x)$	$\frac{1}{1-it}$
<i>szupermodellek:</i>	
$f_a(x)$	
speciális esetek:	
$a=2$ esetén [l. $f_C(x)$]	$e^{- t }$
$a=4$ esetén	$(1+ t) \cdot e^{- t }$
$a=6$ esetén	$(1+ t +t^2/3) \cdot e^{- t }$

(általános a esetén a karakterisztikus függvény csak harmadfajú módosított Bessel-függvényekkel írható fel)

(Steiner, 1990)

Számtani átlagok határeloszlása véges σ esetén

- Az $f(x)$ eloszlásból származó n elemű minta alapján meghatározott átlagok $f^{(n)}(x)$ sűrűségfüggvénye

$$f^{(n)}(x) = n \cdot [f(nx)]^{n-1}$$

- Az n elemű átlagok karakterisztikus függvénye

$$\varphi^{(n)}(t) = \left[\varphi\left(\frac{t}{n}\right) \right]^n$$

Számtani átlagok határeloszlása

véges σ esetén

- Az átlagok $f^{(n)}(x)$ sűrűségfüggvénye helyett azok \sqrt{n} -szeresének $\bar{f}(x)$ sűrűségfüggvényét vizsgáljuk. Ennek karakterisztikus függvénye

$$\bar{\varphi}^{(n)}(t) = \left[\varphi\left(\frac{t}{\sqrt{n}}\right) \right]^n$$

- Az origóra szimmetrikus $f(x)$ sűrűségfüggvény esetében a $\varphi(t/\sqrt{n})$ Taylor-polinomjával

$$\bar{\varphi}(t) = \left[\varphi(0) + \varphi''(0) \frac{t^2}{2n} + O\left(\frac{t^2}{n}\right) \right]^n$$

Számtani átlagok határeloszlása véges σ esetén

- A páros hatványok $f(x)$ szimmetriája miatt zérusok és $\varphi(0)=1$
- A Fourier transzformáció tulajdonságai miatt

$$\left. \frac{d^m \varphi(t)}{dt^m} \right|_{t=0} = i^m \cdot \int_{-\infty}^{\infty} x^m \cdot f(x) dx$$

tehát (ha a szórás véges)

$$\varphi''(0) = - \int_{-\infty}^{\infty} x^2 \cdot f(x) dx = -\sigma^2$$

Számtani átlagok határeloszlása véges σ esetén

- Az átlagok \sqrt{n} -szeresének karakterisztikus függvénye

$$\bar{\varphi}(t) = \left[1 - \frac{\sigma^2 t^2}{2n} + O\left(\frac{t^2}{n}\right) \right]^n$$

- A következő jelölést bevezetve:

$$c_n = -\frac{\sigma^2 t^2}{2} + n \cdot O\left(\frac{t^2}{n}\right)$$

$$\bar{\varphi}(t) = \left[1 + \frac{c_n}{n} \right]^n \quad \text{és} \quad \lim_{n \rightarrow \infty} \bar{\varphi}(t) = \lim_{n \rightarrow \infty} \left(1 + \frac{c_n}{n} \right)^n = e^{\lim_{n \rightarrow \infty} c_n}$$

Számtani átlagok határeloszlása véges σ esetén

- Az átlagok \sqrt{n} -szeresének karakterisztikus függvénye $n \rightarrow \infty$ esetén:

$$\lim_{n \rightarrow \infty} \bar{\varphi}(t) = e^{\lim_{n \rightarrow \infty} c_n} = e^{-\sigma^2 t^2 / 2}$$

amiből viszont

$$\bar{f}(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

- Tehát nagy n -eknél σ/\sqrt{n} szórású Gauss eloszlást közelít az átlagok eloszlása, $\bar{f}(x)$ definíciója miatt

Nagy számok törvénye

- Nagy n -eknél σ/\sqrt{n} szórású Gauss eloszlást közelít az átlagok eloszlása
- Ez a *nagy számok törvénye*, ami egy becslési mód, az átlagképzés \sqrt{n} -el arányos pontosságnövekedéséről tudósít nagy n -ek esetén

Centrális határeloszlástétel

- Nagy n -eknél σ/\sqrt{n} szórású Gauss eloszlást közelít az átlagok eloszlása
- Ez *határeloszlástétel*, mert az átlagok (mint becslések) eloszlásának a típusát is szolgáltatja $n \rightarrow \infty$ esetén
- A *centrális* jelző utal a hagyományos (átlag-szórás) statisztikában betöltött központi szerepére

Átlagok eloszlása Cauchy-eloszlás esetén

- Nincs véges szórása a Cauchy-eloszlásnak
 - a véges szórás *feltétele* a centrális határeloszlás tétele érvényességének
- Standard Cauchy-eloszlás esetében az átlagok karakterisztikus függvénye

$$\varphi^{(n)}(t) = \left[\varphi\left(\frac{t}{n}\right) \right]^n = [e^{-|t/n|}]^n = e^{-|t|}$$

- Bármilyen n -re az átlagok eloszlását *ugyanaz* az $f_C(x)$ sűrűségfüggvény írja le
 - *semmilyen formában nem teljesül* az átlagokra a nagy számok törvénye
 - szó sincs arról, hogy az átlagok nagy n -ekre Gauss-eloszlást közelítenek

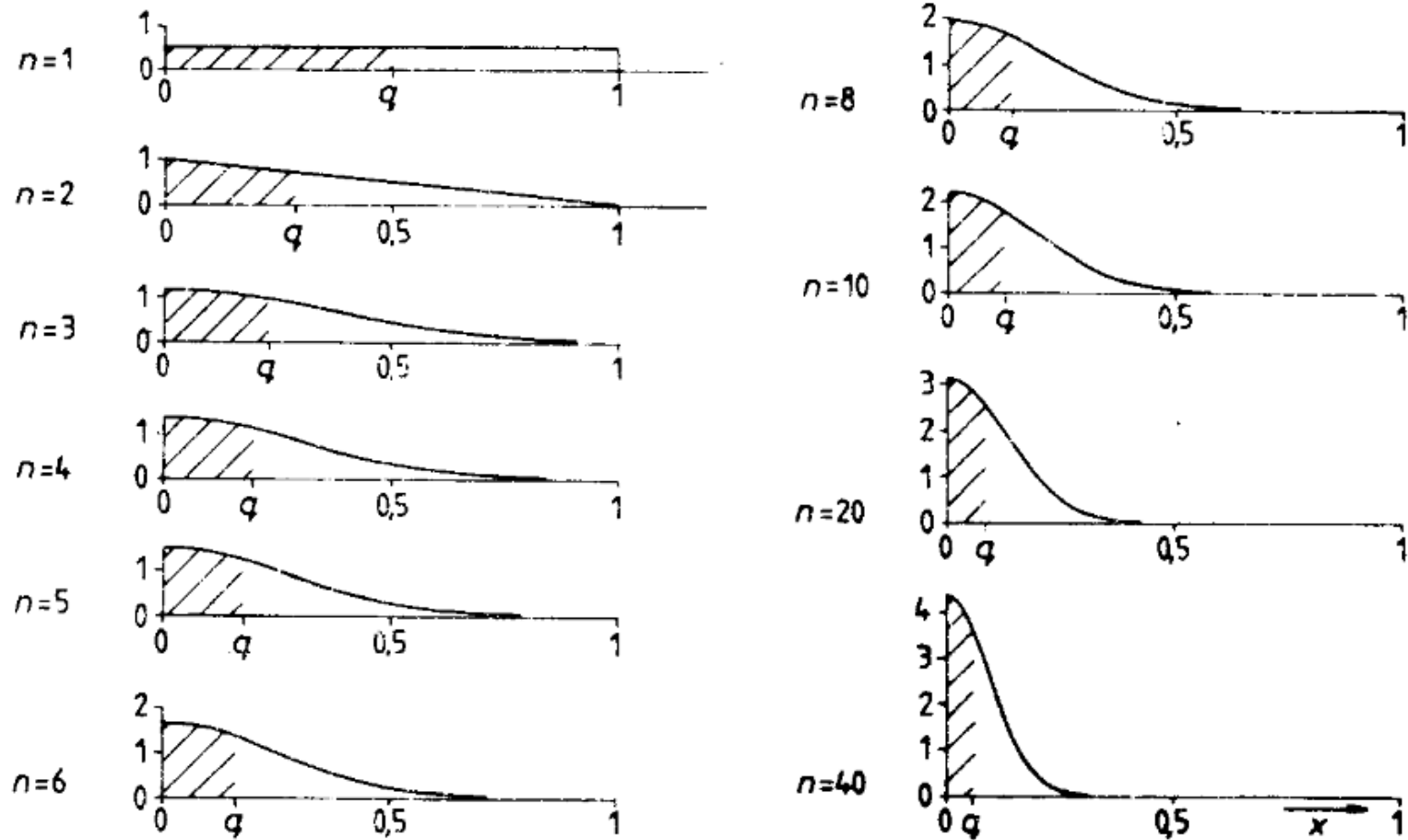
Aszimptotikus szórás

- Ha valamely becsléseloszlás Gauss-típusú A/\sqrt{n} szórással, A -t *aszimptotikus szórásnak* nevezzük
- Az *átlagképzés* A_E aszimptotikus szórása azonos az anyaeloszlás σ szórásával
- A szórás kapcsolata az interkvartilis félterjedelemmel Gauss-eloszlás esetén $q = 0.6745\sigma$, tehát az átlagképzésre a

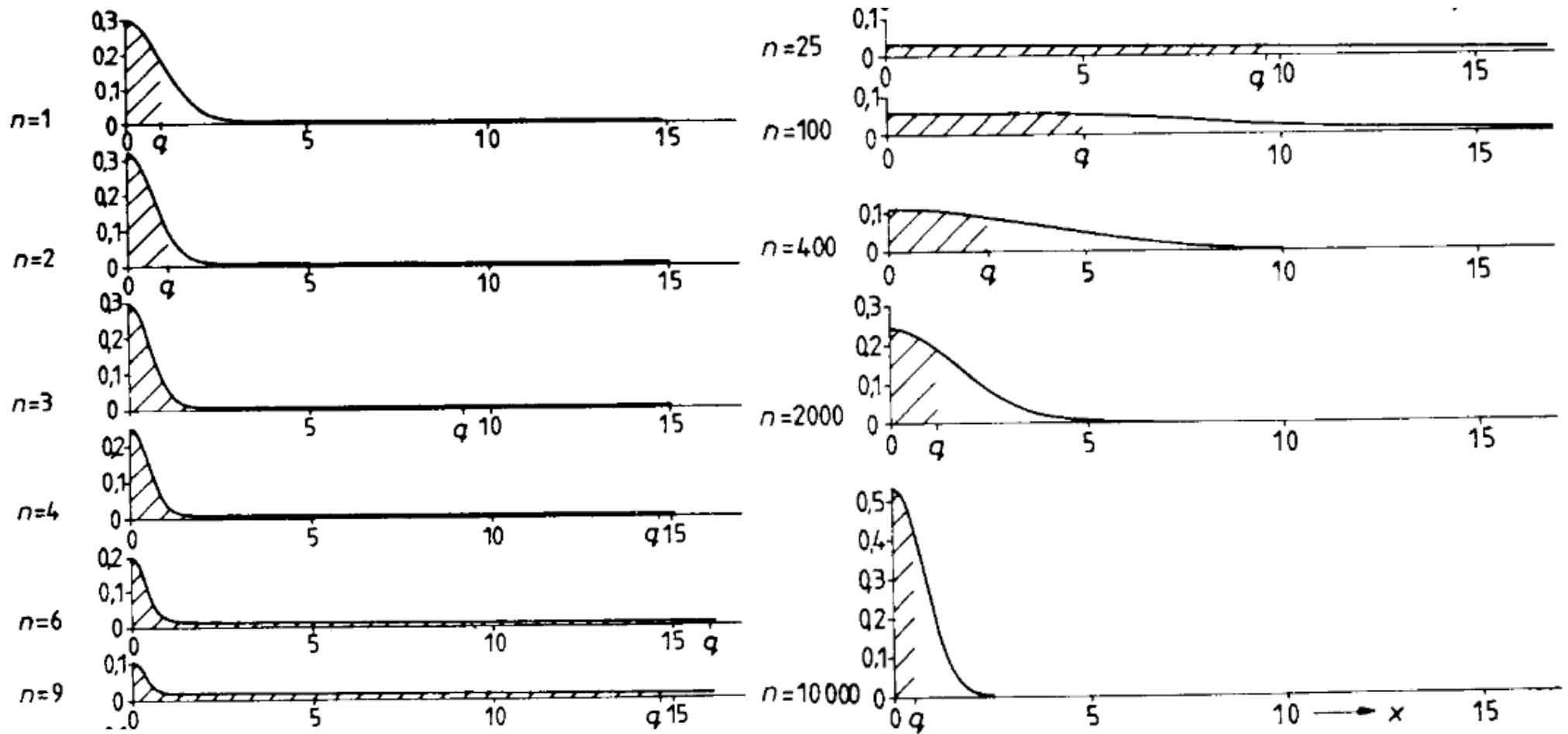
$$\frac{q_n}{A_E \cdot 0.6745}$$

hányadosokat ábrázoló pontok egy $1/\sqrt{n}$ abszcisszájú koordináta-rendszerben az origóból induló 45° -os egyenesen vannak

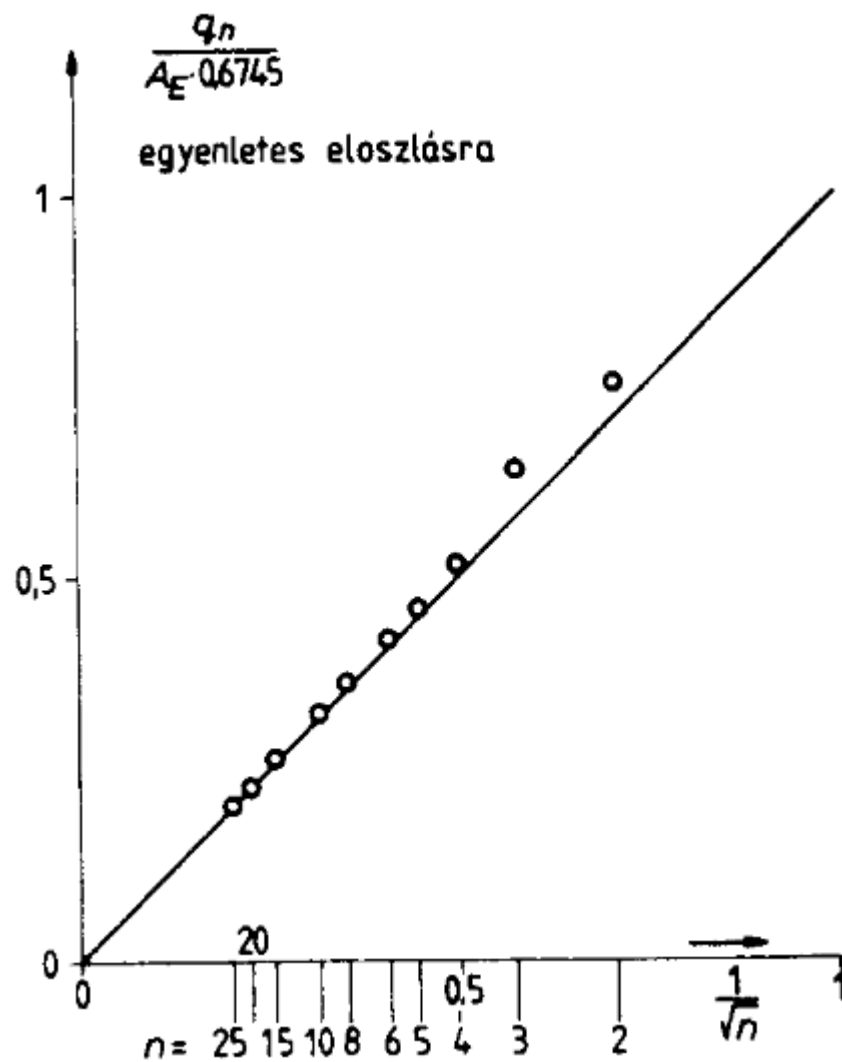
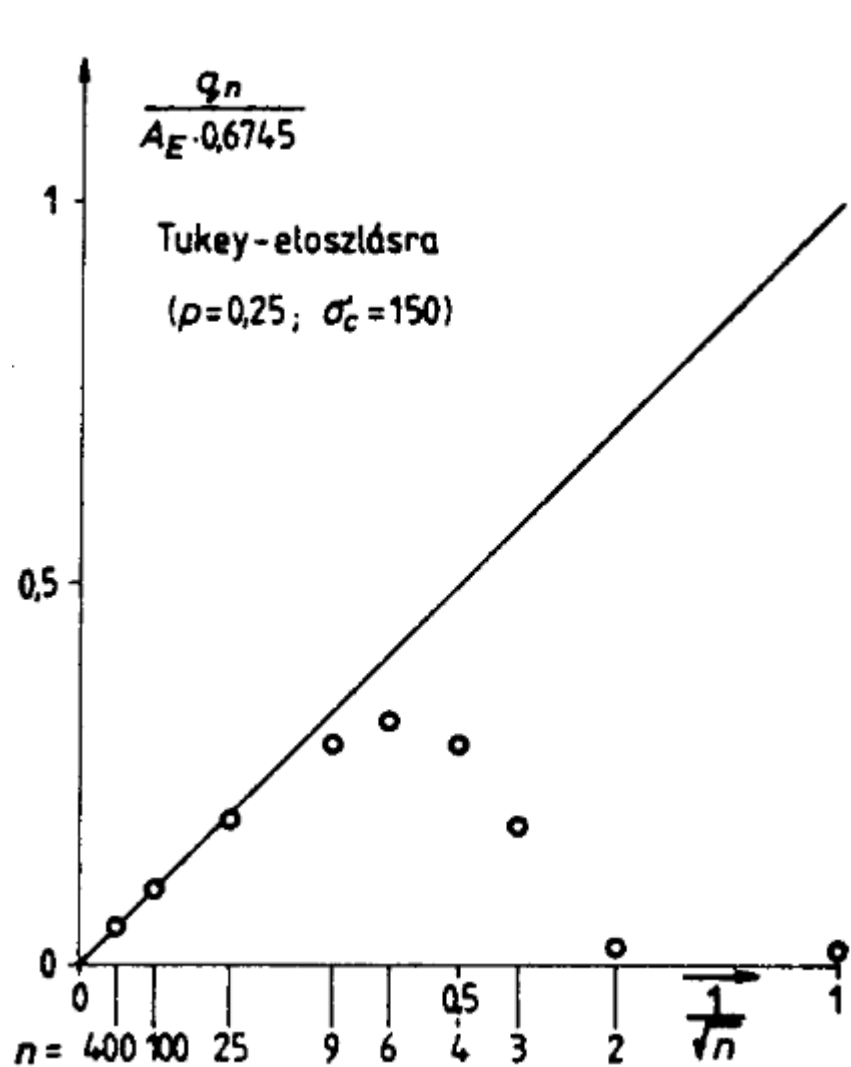
Számtani átlag sűrűségfüggvénye



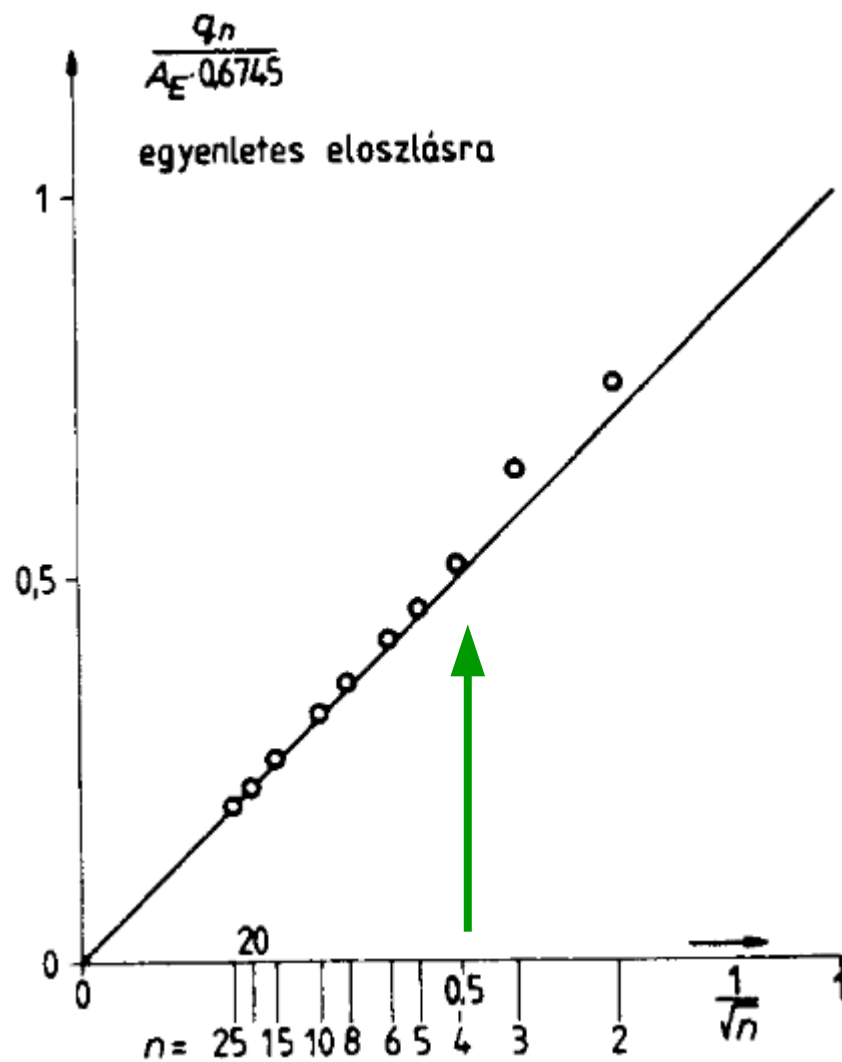
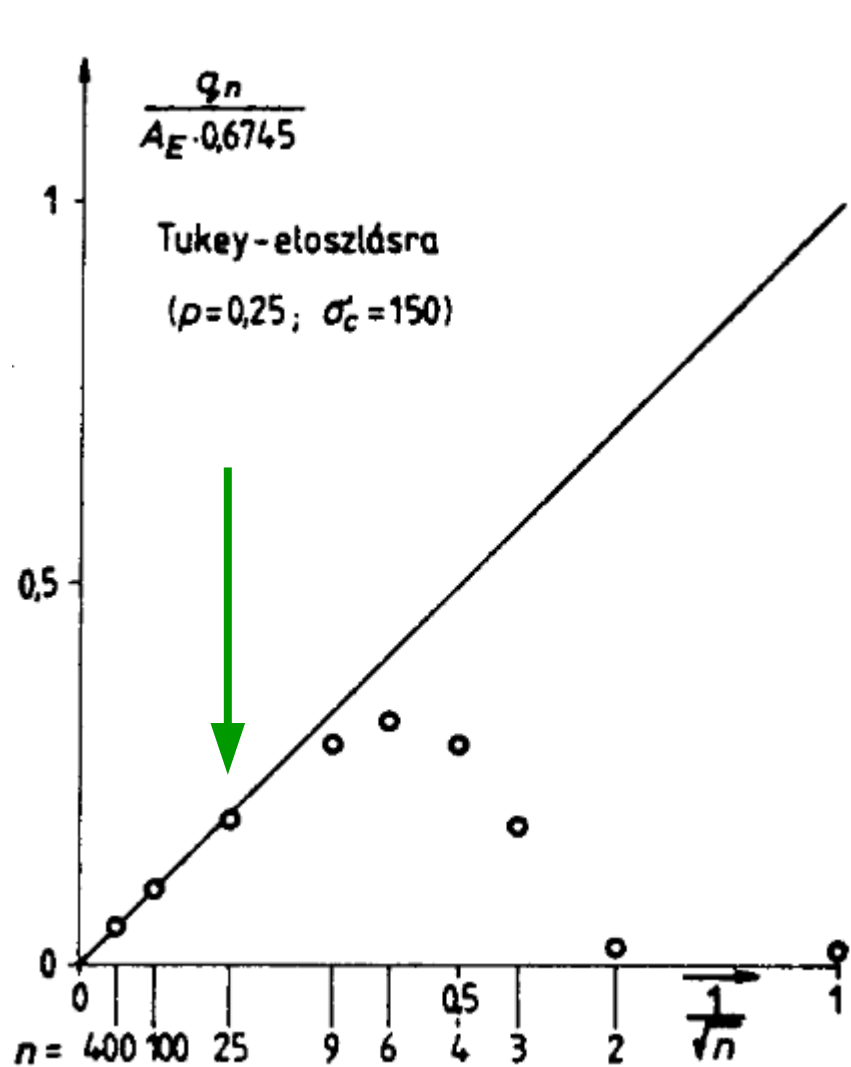
Tukey modell számtani átlagainak sűrűségfüggvénye



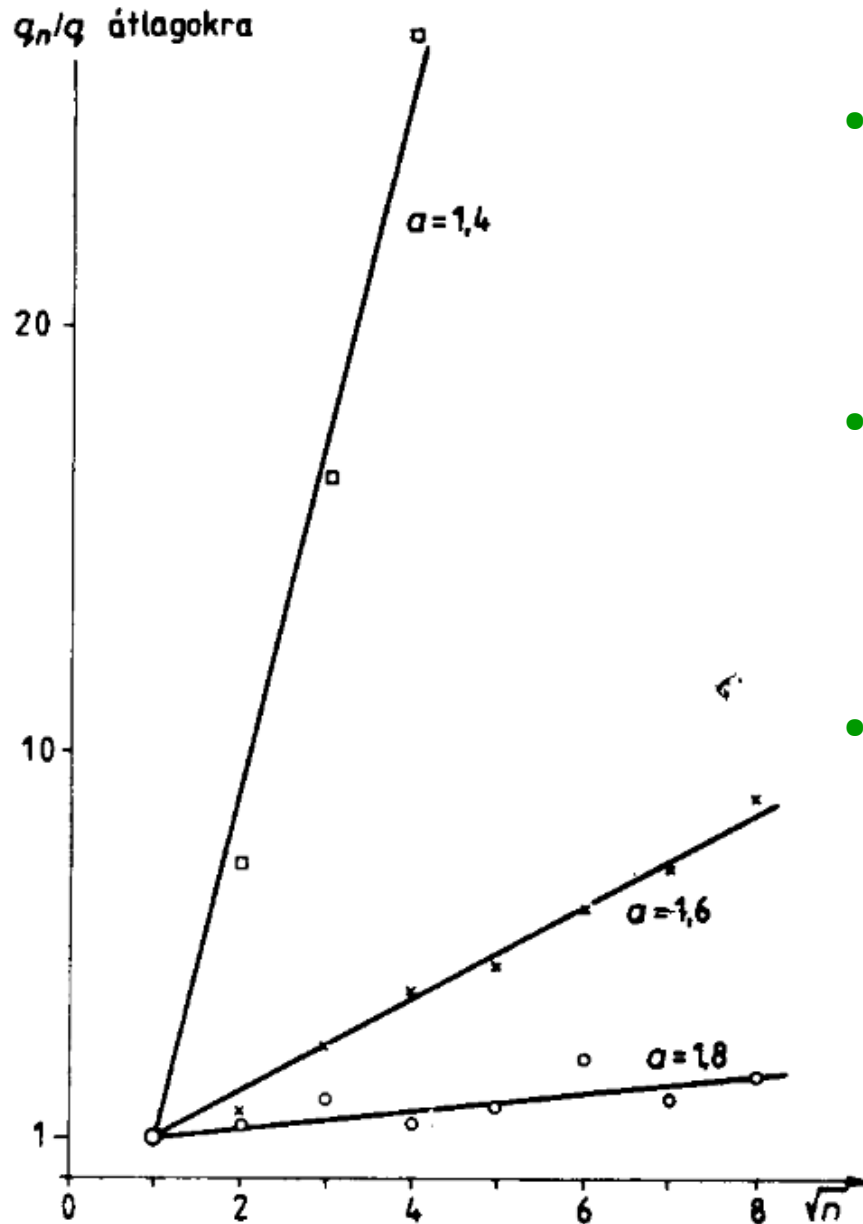
Nagy számok törvényének teljesülése



Nagy számok törvényének teljesülése

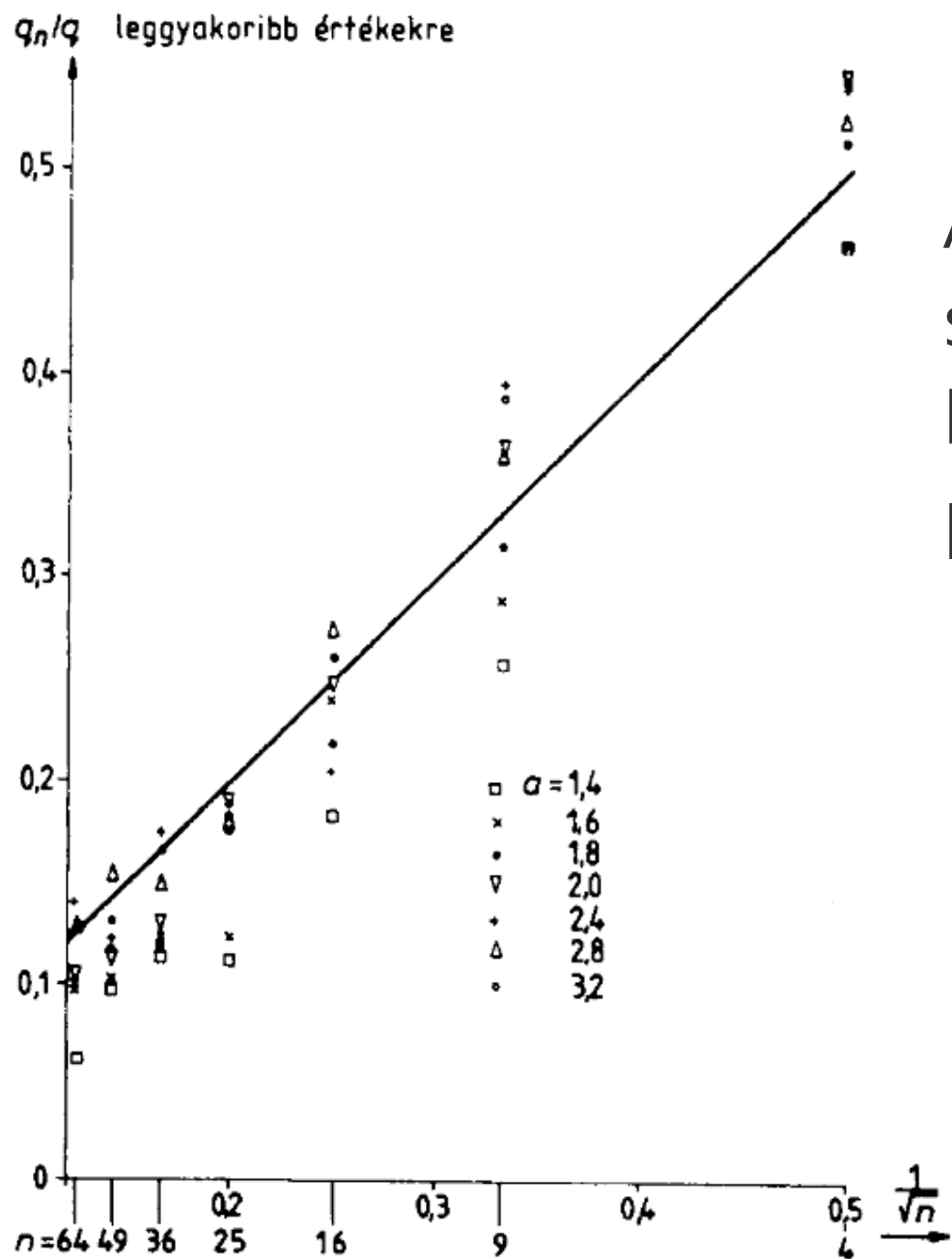


Nagy számok törvényének teljesülése az $f_a(x)$ szupermodellre



- Az átlagok \sqrt{n} -el arányosan *egyre pontatlanabb becslései* az eloszlás szimmetriapontjának
- Az átlag mint becslés romlásának üteme az α típusparaméter-érték csökkenésével egyre nagyobb lesz
- A leggyakoribb értékek ezzel szemben a becslési pontosság \sqrt{n} -el arányos *növekedését* mutatják

Leggyakoribb értékek becslése



A leggyakoribb értékek számításán alapuló becslések követik a \sqrt{n} szerinti pontosságnövekedést

Korrelációs együttható rezisztenciája

- Hagyományosan: átlag, szórás

$$r_E = \frac{\sum_{i=1}^n (x_i - E_x) \cdot (y_i - E_y)}{\sqrt{\sum_{i=1}^n (x_i - E_x)^2} \cdot \sqrt{\sum_{i=1}^n (y_i - E_y)^2}}$$

Korrelációs együttható rezisztenciája

- Rezisztensen: leggyakoribb érték, dihézió

$$r_M = \frac{\sum_{i=1}^n \frac{1}{\varepsilon_x^2 + (x_i - M_x)^2} (x_i - M_x) \cdot \frac{1}{\varepsilon_y^2 + (y_i - M_y)^2} (y_i - M_y)}{\sqrt{\sum_{i=1}^n \left(\frac{1}{\varepsilon_x^2 + (x_i - M_x)^2} \right)^2 (x_i - M_x)^2} \cdot \sqrt{\sum_{i=1}^n \left(\frac{1}{\varepsilon_y^2 + (y_i - M_y)^2} \right)^2 (y_i - M_y)^2}}$$

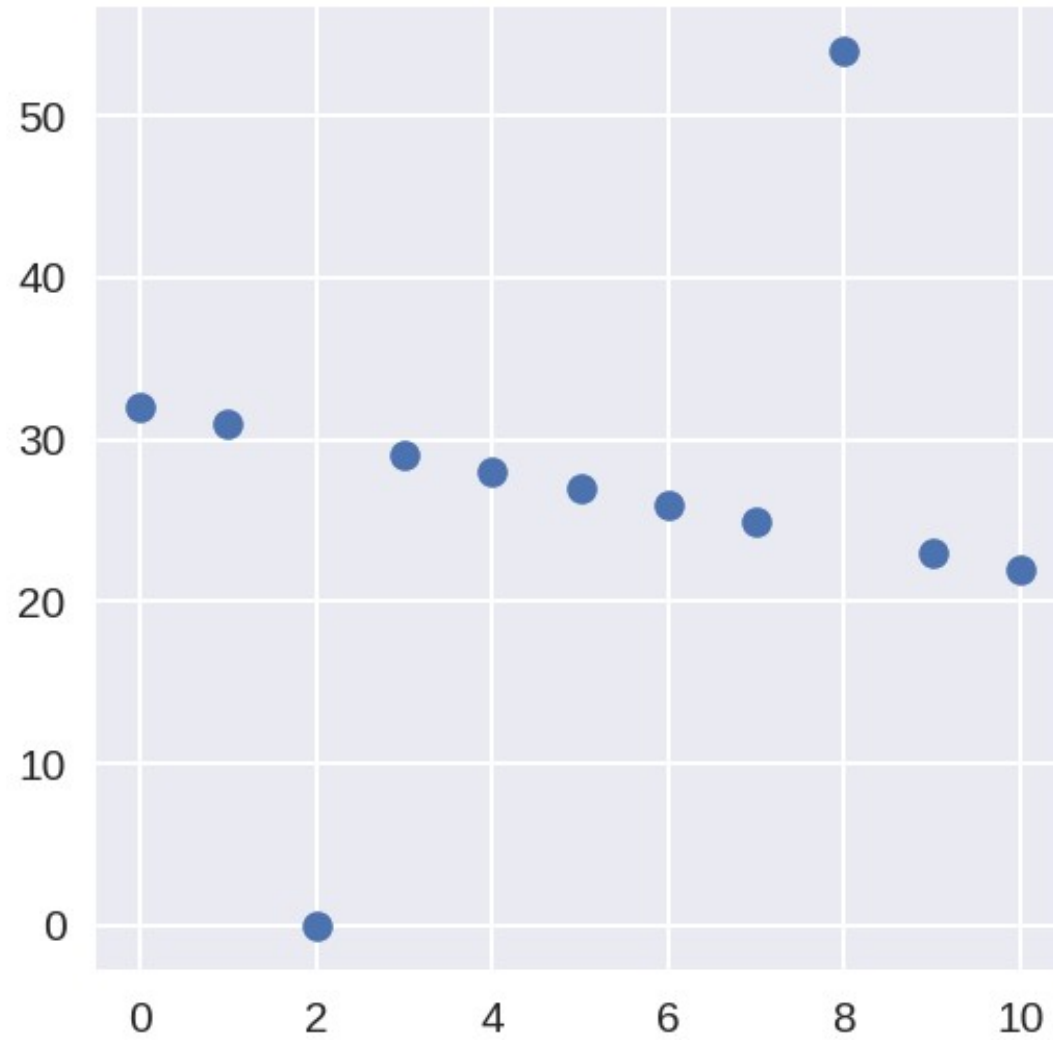
Szám példa (Steiner F.)

- Két pontot ((2, 0) és (8, 54)) kivéve az összes pont egy -1 meredekségű egyenesre esik:

x_i	0	1	2	3	4	5	6	7	8	9	10
y_i	32	31	0	29	28	27	26	25	54	23	22

- Átlag, szórás számítással $r_E = +0.1695$
- Leggyakoribb érték, dihézió számítással $r_M = -0.7833$

Adatok



Regressziós egyeneselek

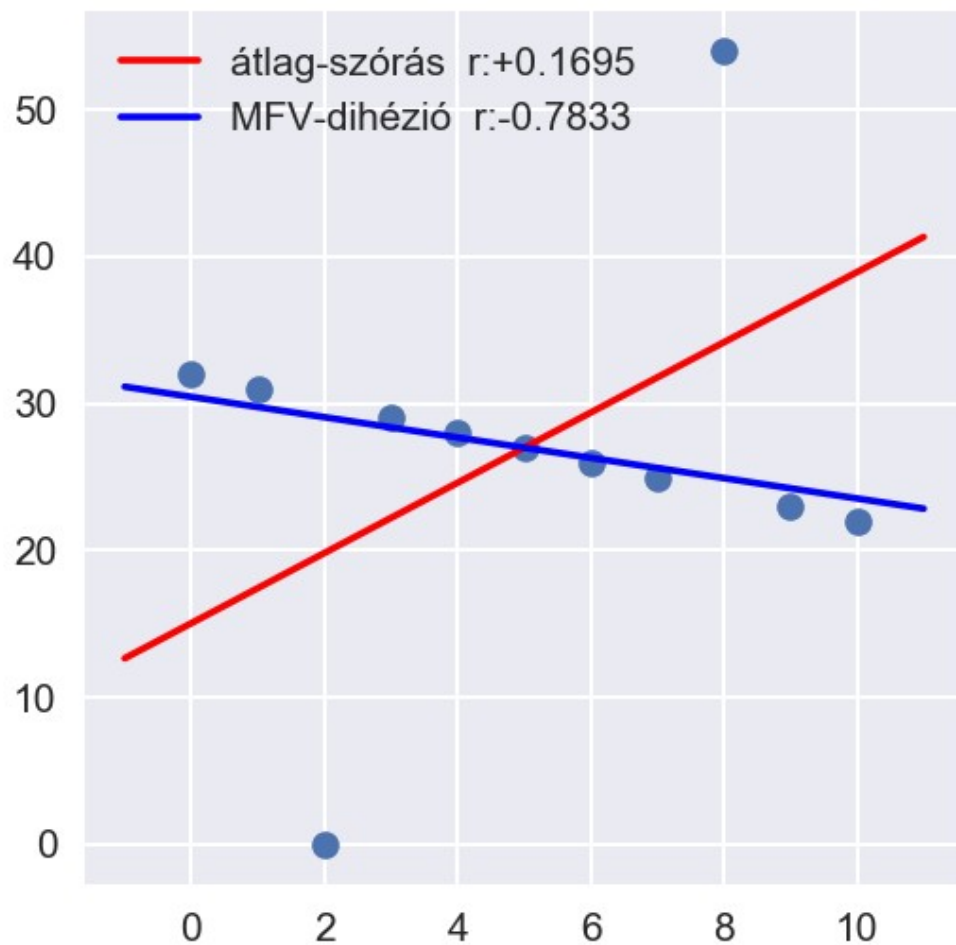
- Hagyományosan: átlag, szórás

$$y = r_E \frac{\sigma_y}{\sigma_x} x + E_y - E_x r_E \frac{\sigma_y}{\sigma_x}$$

- Rezisztensen: leggyakoribb érték, dihézió

$$y = r_M \frac{\varepsilon_y}{\varepsilon_x} x + M_y - M_x r_M \frac{\varepsilon_y}{\varepsilon_x}$$

Regressziós egyenesek



- Melyik regressziós egyenest szeretnénk?

Összefoglalás

- Robusztus és rezisztens mérésfeldolgozási eljárások felhasználásával érhetjük csak el, hogy
 - a mérések eloszlására érzéketlen legyen a számítás (becslés) eredménye
 - a kivágó értékek ne befolyásolják indokolatlan mértékben az eredményeket
 - megkapjuk az elvárt pontosságnövekedést a mérések számának növelésével

Tananyag, szakirodalom

- Steiner (1990): 5.5
- Vincze (1968): 2.5
- Hampel F.R, Ronchetti E.M, Rousseeuw P.J, Stahel W.A (1986): Robust Statistics. The Approach Based on Influence Functions. Wiley & Sons