

MM-ODHADY

ZDENĚK FABIÁN

Abstrakt. The recently proposed core functions of continuous probability distributions are briefly introduced and used for a generalization of the classical moment estimation method. The suggested core-moment estimators and, if necessary, the (Huber) moment M-estimators are "tailored" to the assumed distributions similarly as the maximum likelihood estimators, with which they in some cases coincide. However, these generalized moment estimates are robust. They are in multidimensional cases unexpectedly simple and their asymptotic relative efficiences seem to be reasonably near to one.

Резюме: Коре функции непрерывных распределений использованы для обобщения классической параметрической оценки методом моментов. Оценки методом коре моментов похожи на оценки методом максимального правдоподобия, но они робастны. Они чрезвычайно простой даже в многомерном случае и почти еффективны.

1. Úvod

Na Robustu'96 jsem uveřejnil příspěvek, zaznamenaný v [1]. Připomenu jeho obsah.

Budiž F distribuční funkce a f hustota rozdělení P s pravděpodobnostní mírou kladnou právě na otevřeném intervalu S_F (konečném nebo nekonečném). Rozdělení si lze představit jako vzniklé transformací $\varphi : R \to S_F$ nějakého 'vzoru' G s pravděpodobnostní mírou kladnou na celém $S_G = R$. V [1] jsem se pokusil ukázat, že rozdělení P lze výstižně popsat nejen funkcemi F a f, ale i funkcí q, danou vztahem

$$q(x) = \frac{1}{f(x)} \frac{d}{dx}(-L(x)f(x)) \tag{1}$$

kde

$$L(x) = \varphi'(\varphi^{-1}(x)) \tag{2}$$

je Jakobián určitého, na S_F závislého zobrazení $\varphi : R \to S_F$, a dále, že:

a/ Popis rozdělení pomocí q je často jednodušší a přehlednější než oba standardní: tak rozdělení s těžkými konci odpovídá q omezená (O), rozdělení ostře spadajícímu k nule q neomezená (N) a nesymetrickému rozdělení s různým typem chování na obou koncích intervalu S_F odpovídá q typu $O - N$ nebo $N - O$.

b/ Momenty funkce q jsou lepšími numerickými charakteristikami rozdělení než obvyklé momenty už proto, že existují, a zobecněná momentová metoda pro tento typ momentů poskytuje zajímavé odhady parametrů parametrických rozdělení, které sice obecně nejsou vydatné, ale někdy skoro, zato jsou jednoduché a pro rozdělení s q typu O robustní.

2000 *Mathematics Subject Classification.* Primary 62E10; Secondary 62E20.

Klíčová slova. Core function, maximum likelihood, M-estimators.

Autor děkuje Igoru Vajdovi a Josefu Machkovi za skvělé postřehy a cenné rady. Práce byla podpořena granty GA ČR 201/00/1489 a GA AV IAA1075101.

Funkci (1) nazývám v [1] geometrickou funkcí rozdělení. Že by snad nějaká zásadní funkce rozdělení unikala po celé minulé století pozornosti statistiků se ovšem zdá naprosto nepravděpodobné. Taky se to v [1] takto rezolutně netvrdí: pro G na $S_G = R$ je geometrickou funkcí rozdělení prostě skórová funkce

$$q(y) = s^G(y) = -g'(y)/g(y) \tag{3}$$

(což dostaneme z (1), vezmeme-li v tomto případě $\varphi : R \to R$ jako identické zobrazení, pak je $L(x) = 1$). Pro rozdělení F na $S_F \neq R$ vznikne podle Věty 1 z [1] funkce (1) následujícím postupem: Buď dána f. Transformací φ^{-1} najdeme hustotu g 'vzoru' G na $S_G = R$, určíme podle (3) skórovou funkci s^G, tu transformujeme zpět na S_F a dostaneme

$$q(x) = s^G(\varphi^{-1}(x)). \tag{4}$$

Neboli: problém, který neumíme vyřešit na S_F 'promítneme' na R, zde jej vyřešíme a řešení 'promítneme' zpět na S_F.

Nabízí se řada otázek: Jak to s q vypadá v novém tisíciletí ? Je ještě zajímavá nebo to byla jen duhová bublina ? A ovšem (nejčastější výhrada posluchačů mých přednášek a recenzentů nepřijatých článků: jak q může být relevantní funkcí daného rozdělení, když závisí na v podstatě libovolné (spojité rostoucí) φ ?

V tomto článku se pokusím tyto otázky zodpovědět a odpovědi zdůvodnit. Zásadně je možno říci tolik: Uvedeným postupem skutečně kupodivu dostaneme obecně neznámou a smysluplnou funkci. V mnoha konkrétních případech je, pravda, známá a ve statistice hojně používaná, zdaleka ne však ve všech: zcela neznámá je např. pro rozdělení bez parametrů či bez toho parametru, kterému v dalším textu říkám transformovaný parametr polohy (viz tabulky 1-3 příkladu 1 v [1]). Pro různá φ můžeme samozřejmě dostat pro jedno dané rozdělení různá q, ukázuje se však, že pro analyticky zadanou f existuje zpravidla jediné φ, které dává q v matematicky uchopitelném tvaru.

Budu zde mluvit o funkci T_F, které dnes říkám *core* nebo *core funkce*. Je to 'vysvlečená' geometrická funkce rozdělení z článku [1]. Zatímco 'oblečená' q v podstatě nikoho nezajímala, 'vysvlečenou' T_F celkem ochotně přijali v poměrně slušném periodiku [4]. Upozorňuji ještě na další značení odlišné od [1]: místo $p(u)$ píšu $f(x)$ a definiční obor rozdělení F značím S_F místo T.

2. Core funkce

Uvažujme obecnou parametrickou rodinu $\mathbb{G} = \{G_\theta, \theta \in \Theta\}$ na $S_G = R$, kde

$$\theta = (\mu, \sigma, \xi) \in S_G \times (0, \infty) \times \Delta,$$

kde $\Delta = (0, \infty)^{m-2}$ kde $m \geq 2$ celé a kde μ a σ jsou parametry polohy a měřítka. Pak existuje taková H_ξ, že

$$G_\theta(y) = H_\xi\left(\frac{y-\mu}{\sigma}\right). \tag{5}$$

Hustota rozdělení G_θ je zřejmě

$$g_\theta(y) = \frac{1}{\sigma} h_\xi\left(\frac{y-\mu}{\sigma}\right), \tag{6}$$

kde $h'_\xi = H_\xi$. Definujme *core funkci* rozdělení G_θ jako

$$T_{G_\theta}(y) = -\frac{g'_\theta(y)}{g_\theta(y)}. \tag{7}$$

(7) je totéž jako (3) pro rozdělení bez parametrů nebo s parametrem polohy. Uvažujme však rozdělení (5). Pod pojmem skórová funkce parametrického rozdělení se obvykle chápe věrohodnostní skór pro parametr polohy, $s_\mu^G(y) = (\partial/\partial\mu)\log g_\theta(y)$ (log je přirozený logaritmus). Podle (6) platí

$$s_\mu^G(y) = \frac{\partial}{\partial\mu}\log\left(\frac{1}{\sigma}h_\xi\left(\frac{y-\mu}{\sigma}\right)\right) = -\frac{h'_\xi((y-\mu)/\sigma)}{h_\xi((y-\mu)/\sigma)}\frac{1}{\sigma} = -\frac{1}{\sigma}\frac{g'_\theta(y)}{g_\theta(y)} = \frac{1}{\sigma}T_{G_\theta}(y).$$

Core funkce (7) je tedy vnitřní součástí věrohodnostního skóru pro parametr polohy, s_μ^G, což se dá hezky říct anglicky jako že 'Core is the core of the (likelihood) score'.

Příklad 1. Normální rozdělení $\mathbb{N}(\mu,\sigma)$ má skórovou funkci (věrohodnostní skór pro parametr polohy) $s_\mu^{\mathbb{N}}(y) = (y-\mu)/\sigma^2$ a core funkci $T_{\mathbb{N}}(y) = (y-\mu)/\sigma$. Core funkce normálního rozdělení chápaná jako náhodná veličina je prostě normovaná náhodná veličina.

Rodinu $\mathbb{G}$ můžeme považovat za *vzor* indukované (transformované) rodiny $\mathbb{F}$ na $S_F = (\varphi(-\infty), \varphi(\infty))$. Pro $F_\theta \in \mathbb{F}$, $F_\theta = G_\theta\varphi^{-1}$ platí

$$F_\theta(x) = G_\theta(\varphi^{-1}(x)) = H_\xi\left(\frac{\varphi^{-1}(x)-\mu}{\sigma}\right),$$

což můžeme přepsat jako

$$F_\theta(x) = H_\xi(w),$$

kde

$$w = \frac{\varphi^{-1}(x)-\varphi^{-1}(\tau)}{\sigma}. \tag{8}$$

Zde jsme položili $\tau = \varphi(\mu)$. τ je transformovaný parametr polohy známý z [1]. Parametr θ teď značí vektor

$$\theta = (\tau,\sigma,\xi) \in S_F \times (0,\infty) \times \Delta.$$

Hustota *indukovaného* (transformovaného) rozdělení F_θ je pak

$$f_\theta(x) = \frac{d}{dx}H_\xi(w) = h_\xi(w)\frac{dw}{dx} = \frac{1}{\sigma L(x)}h_\xi(w) \tag{9}$$

kde $h_\xi = H'_\xi$ a $L(x)$ je dáno vztahem (2).

Teď už je možné podle (4) napsat výraz pro core funkci rozdělení F_θ:

$$T_{F_\theta}(x) = T_{G_\theta}(\varphi^{-1}(x)) = T_{H_\xi}(w) = -\frac{h'_\xi(w)}{h_\xi(w)}. \tag{10}$$

Je o něco jednodušší než původní q, protože vzniká místo ze 'score' z jednoduššího 'core' vzoru. Tím se trochu pozmění i obecný vzorec (1) pro její vyjádření bez znalosti konkrétního vzoru:

Věta 1. *Core funkce rozdělení F_θ indukovaná zobrazením $\varphi : R \to S_F$ je*

$$T_{F_\theta}(x) = \frac{1}{f_\theta(x)}\frac{d}{dx}(-\sigma L(x) f_\theta(x)). \tag{11}$$

Důkaz. Podle (10) a (9) je

$$T_{F_\theta}(x) = -\frac{h'_\xi(w)}{h_\xi(w)} = \frac{1}{\sigma L(x) f_\theta(x)} \frac{d}{dx}(-\sigma L(x) f_\theta(x)) \frac{dx}{dw}$$

a $dw/dx = 1/(\sigma L(x))$ podle (8) a (2). □

Následující věta podává interpretaci core funkce rozdělení na S_F. Poznamenejme, že vše, co platí pro rozdělení F, platí samozřejmě i pro rozdělení G na $S_G = R$ při aplikaci identické transformace φ. Tedy: τ v dalším znamená i μ.

Věta 2. *Věrohodnostní skóry rozdělení F_θ pro transformovaný parametr polohy a pro parametr měřítka lze vyjádřit pomocí core funkce jako*

(12)
$$\begin{aligned} s_\tau^F(x) &= \frac{1}{\sigma L(\tau)} T_{F_\theta}(x) \\ s_\sigma^F(x) &= \frac{1}{\sigma}(w T_{F_\theta}(x) - 1). \end{aligned}$$

Důkaz. Přímým derivováním a použitím (9) dostaneme

$$s_\tau^F(x) = \frac{\partial}{\partial \tau} \log(\frac{1}{\sigma L(x)} h_\xi(w)) = \frac{h'_\xi(w)}{h_\xi(w)} \frac{\partial w}{\partial \tau} = \frac{1}{\sigma L(\tau)} T_{H_\xi}(w)$$

$$s_\sigma^F(x) = \frac{\partial}{\partial \sigma} \log(\frac{1}{\sigma L(x)} h_\xi(w)) = -\frac{1}{\sigma} + \frac{h'_\xi(w)}{h_\xi(w)} \frac{\partial w}{\partial \sigma} = \frac{1}{\sigma}(w T_{H_\xi}(w) - 1),$$

kde w je dáno vztahem (8). Tvrzení pak získáme užitím (10). □

$T_{F_\theta}(x)$ indukovaného modelu je tedy, podobně jako v případě vzoru G_θ, *vnitřní součástí věrohodnostního skóru pro transformovaný parametr polohy.*

Příklad 2. Jedním z možných rozdělení na $S_F = (0, \infty)$ s transformovaným parametrem polohy a parametrem měřítka je Weibullovo rozdělení s hustotou

$$f_{\tau,\beta}(x) = \beta x^{-1} (x/\tau)^\beta e^{-(x/\tau)^\beta}.$$

Zvolme zobrazení $x = \varphi(y) = e^y$. Hustotu rozdělení upravíme na tvar

$$f_{\tau,\beta}(x) = \beta x^{-1} e^w e^{-e^w},$$

kde $w = \log(x/\tau)^\beta = \sigma^{-1}(\log x - \log \tau)$ a $\beta = \sigma^{-1}$. Vzorem pro F je tedy dvojitě exponenciální rozdělení, věrohodnostní skór pro τ je

$$s_\tau^F(x) = \beta \tau^{-1}((x/\tau)^\beta - 1),$$

a core spočteme pomocí (11) jako

$$T_{F_{\tau,\beta}}(x) = \frac{1}{f_{\tau,\beta}(x)} \frac{d}{dx}\left(-\frac{x}{\beta} f_{\tau,\beta}(x)\right) = (x/\tau)^\beta - 1.$$

Core je vnitřní součástí věrohodnostního skóru pro τ.

Příklad 3. V [1] je v příkladu 6 je uvažován speciální případ *Lomaxova* rozdělení na
$S_F = (0, \infty)$ s hustotou $f_\alpha(x) = \alpha/(x+1)^{\alpha+1}$. Zde α není transformovaným parametrem polohy. Věrohodnostní skór pro α je $s_\alpha^F(x) = 1/\alpha - \ln(x+1)$ a core funkce (opět při $\varphi(y) = e^y$) je $T_{F_\alpha}(x) = (\alpha x - 1)/(x+1)$, což je zbrusu nová funkce, která podle mne lépe popisuje Lomaxovo rozdělení než s_α. Nová, ale namíchaná podle starého receptu: je to vnitřní součást věrohodnostního skóru pro 'latentní' parametr polohy $\tau = 1$. Je to jen funkce, která v případech 'nedokonalých' rozdělení bez transformovaného parametru polohy prostě nebyla dosud rozpoznána.

3. Zobrazení φ

Explicitní tvar core funkce rozdělení na $S_F \neq R$ získáme po specifikaci zobrazení φ. V [1] bylo zobrazení

$$x = \varphi(y) = \frac{a(b-c) + b(c-a)e^y}{b - c + (c-a)e^y} \tag{13}$$

chybně považováno za jediné 'správné' (dodejme že v [1] je uvedena pouze jeho inverze).

Stojí zato si povšimnout, že v [12] platí $\varphi(-\infty) = a$, $\varphi(+\infty) = b$, $\varphi(0) = c$, a že $\lim_{b\to\infty} \varphi(y) = a + (c-a)e^y$. Zobrazení (13) poskytuje možnost získat širokou třídu rozdělení a jeho zvláštními případy jsou známé Johnsonovy transformace (viz [9] nebo [10]):

$\varphi : R \to (0, \infty)$ ve tvaru $(a = 0, b = \infty, c = 1)$:

$$x = \varphi(y) = e^y$$

$\varphi : R \to (0, 1)$ ve tvaru $(a = 0, b = 1, c = 1/2)$:

$$x = \varphi(y) = \frac{e^y}{1 + e^y}.$$

Rodinám rozdělení, které vzniknou z rodin 'vzorů' transformací (13) budeme říkat 'johnsonovské'. Jak se ukázalo, drtivá většina modelových rozdělení používaných ve statistice je právě 'johnsonovských'. Explicitní tvary jejich core funkcí na polopřímce $S_F = (0, \infty)$ (kde Jakobiján $L(x) = x$) a intervalu $(0, 1)$ (kde $L(x) = x(1-x)$) jsou podle Věty 1

$$T_F(x) = \sigma[-1 - xf'(x)/f(x)] \quad \text{na } (0, \infty) \tag{14}$$

$$T_F(x) = \sigma[-1 + 2x - x(1-x)f'(x)/f(x)] \quad \text{na } (0, 1). \tag{15}$$

Uvažujme však rozdělení F na $S_F = (-\frac{\pi}{2}, \frac{\pi}{2})$ s hustotou

$$f(x) = \frac{1}{\sqrt{2\pi}\cos^2 x} e^{-\frac{1}{2}\mathrm{tg}^2 x} \tag{16}$$

a zvolme $\varphi : R \to S_F$ ve tvaru $x = \varphi(y) = \mathrm{arctg}\, y$. Jacobián inverzní transformace $\varphi^{-1} : S_F \to R$ je $d\varphi^{-1}(x)/dx = 1/\cos^2 x$ a vzor G rozdělení F je zřejmě standardní normální rozdělení, $g(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$, s core funkcí $T_G(y) = y$. Podle definice (10) je arctg-indukovaná core funkce rozdělení (16) dána jako $T_F(x) = T_G(\varphi^{-1}(x)) = \mathrm{tg}\, x$.

Přímým výpočtem se lze přesvědčit, že obdobou vzorců (14) a (15) je v tomto případě

$$T_F(x) = [\sin 2x - \cos^2 x f'(x)/f(x)].$$

Rozdělení (16) má ovšem i svou 'johnsonovskou core funkci', tu však nelze vyjádřit jednoduchým vzorcem. Myslím si, že obecně lze za core funkci rozdělení F považovat ten *nejjednodušší smysluplný popis rozdělení vyjádřený pomocí funkce obsahující člen* $-f'/f$, o který si matematický tvar hustoty 'řekne'. Snad bych dodal, že vím o hustotě, která si o vhodné φ 'neřekla' (lambda rozdělení, viz [7]). V dalším budeme pro jednoduchost uvažovat pouze 'johnsonovské rodiny'.

4. Momenty core funkce

Stejně jako v [1] definujeme *core momenty* vztahem

$$M_k(\theta) = \int_{S_F} T_{F_\theta}(x)^k \, dF_\theta(x), \qquad k = 1, 2, ...$$

Podle [1], Věta 1 platí $M_1 = 0$. Dále, $M_2(\theta)$ je v [1] považován za střední (Fisherovu) informaci rozdělení F_θ, což po modifikaci 'geometrické funkce' neplatí: označíme-li $I_{F_{\tau\tau}}(\theta)$ Fisherovu informaci o transformovaném parametru polohy, platí podle Věty 1 tohoto článku

$$M_2(\theta) = \sigma^2 L^2(\tau) I_{F_{\tau\tau}}(\theta). \tag{17}$$

(Poznamenejme, že proměna q v T_F se v [1] projeví už jen ve vzorcích pro momenty příkladu 5, v nichž vypadne β a jeho mocniny.)

Vztah (17) poskytuje další interpretaci pojmu core funkce. Podle (12) a (17) je

$$T^*_{F_\theta} = \frac{T_{F_\theta}(x)}{\sqrt{M_2(\theta)}} = \frac{s^F_\tau(x)}{\sqrt{I_{F_{\tau\tau}}(\theta)}},$$

čili normovaná core funkce je věrohodnostní skór pro transformovaný parametr polohy normovaný odmocninou z Fisherovy informace pro tento parametr (znovu: pokud rozdělení ten parametr má. Když ho ale nemá, nevadí. I pro rozdělení bez parametrů je T_F dána definicí (11) a M_2 je pak Fisherova informace rozdělení, viz [1]).

Příklad 4. Pro Weibullovo rozdělení (příklad 2) platí $I_{F_{\tau\tau}} = (\sigma\tau)^{-2}$ a pak $M_2 = \sigma^2\tau^2 I_{F_{\tau\tau}} = 1$.

5. Odhady metodou core momentů

Pro parametrické odhady parametru $\theta \in \Theta \subset R^m$ máme k dispozici m věrohodnostních skórů, ale jen jednu core funkci. Podle [1] však existují momenty funkce T_F, které jsou pro snad všechna jednoduchá modelová rozdělení

– vyjádřeny prostou funkcí parametrů bez různých gamma, psí a beta funkcí, kterými se hemží vzorce pro výpočet obyčejných momentů,

– a to funkcí ne všech parametrů. Core momenty (CM) nezávisí na τ a v nové verzi ani na σ (či β), takže $M_k(\theta) = M_k(\xi)$.

Řešení $(\hat\theta_n)_{CM}$ soustavy

$$\frac{1}{n}\sum_{i=1}^{n} T_F(x_i|\theta)^k = M_k(\theta), \qquad k = 1, ..., m \tag{18}$$

je tedy poměrně jednoduché, zřejmě konzistentní a asymptoticky normální a pro T_F typu O (tedy právě pro rozdělení s delšími chvosty a daty obsahujícími odlehlé hodnoty) robustní. To vše lze snadno nahlédnout z obecné teorie M-estimátorů (viz [11]). Co nahlédnout nelze a je nutno počítat jsou asymptotické disperze CM odhadů resp. jejich asymptotické relativní vydatnosti (ARV),

$$\text{ARV}(\hat\theta_{\text{CM}}) = \lim_{n\to\infty} (Var(\hat\theta_n)_{ML} / Var(\hat\theta_n)_{CM}),$$

kde $(\hat\theta_n)_{ML}$ značí maximálně věrohodný (ML) odhad.

Podle [1], Věta 2, platí $M_1(\theta) = 0$. Podle Věty 2 tohoto článku pro $m = 1$ a rozdělení s transformovaným parametrem polohy je rovnice

$$\sum_{i=1}^{n} T_F(x_i|\tau) = 0 \tag{19}$$

totožná s maximálně věrohodnou rovnicí pro τ. Totožná fakticky, ne však koncepčně. Mějme dvě náhodná pozorování $\mathbf{X} = (X_1, X_2)$ z Cauchyho rozdělení s hustotou ve tvaru $g_\mu(x) = 1/\pi(1 + (x-\mu)^2)$ a s neznámým parametrem polohy μ. Potíže s Cauchyho rozdělením jsou notoricky známé: rozdělení nemá žádný 'obyčejný' moment. Věrohodnostní funkce $\sum_{i=1}^{2} \log g_\mu(X_i)$ má pro dostatečně vzdálené X_1, X_2 dva vrcholy (viz např. [6], str. 213) a rozumný odhad μ je tedy ne v maximu věrohodnostní funkce, ale v jejím *lokálním minimu*, řekněme $\hat{\mu}_{\min}$. Core funkce Cauchyho rozdělení s jediným parametrem μ je identická s jeho skórovou funkcí a rovnice (19) pro první core moment, která zní $2X_1/(X_1-\mu)^2 + 2X_2/(X_2-\mu)^2 = 0$, poskytuje odhad $\hat{\mu}_{\min}$ bez jakýchkoli koncepčních problémů.

Uvažujme rozdělení, které má pouze transformovaný parametr polohy a parametr měřítka, $F_{\tau\sigma}(x) = H_1(w)$ (tohoto typu jsou např. normální, lognormální, Johnsonovo, dvojitě exponenciální s hustotou $f(x) = e^w e^{-e^w}$, extreme value I a II, Weibullovo, logistické, log-logistické, Cauchyho a log-Cauchyho rozdělení). Označme $w_i = (\varphi^{-1}(X_i) - \varphi^{-1}(\tau))/\sigma$. Soustava rovnic pro ML odhady má tvar

$$\begin{aligned} \sum T_F(w_i) &= 0 \\ n^{-1}\sum (w_i T_F(w_i) - 1) &= 0, \end{aligned}$$

kdežto soustava rovnic (18) pro CM odhady zní

$$\begin{aligned} \sum T_F(w_i) &= 0 \\ n^{-1}\sum (T_F(w_i)^2 - 1) &= 0 \end{aligned} \tag{20}$$

kde $\sum$ značí $\sum_{i=1}^{n}$. V případě normálního rozdělení jsou obě soustavy identické.

Z tvaru rovnic je patrné, že pro rozdělení s omezenou T_F (logistické a log-logistické rozdělení) jsou momentové odhady (na rozdíl od ML odhadů) robustní. Spočtené ARV jsou $ARV(\hat{\mu}_{CM}) = 1$, $ARV(\hat{\sigma}_{CM}) = 0.874$ pro logistické rozdělení a $ARV(\hat{\tau}_{CM}) = 1$, $ARV(\hat{\sigma}_{CM}) = 0.953$ pro log-logistické rozdělení, což myslím jde.

6. Momentové M-odhady

Zabývejme se rozdělením *gamma* na $S_F(0,\infty)$ s hustotou

$$f(x|\alpha,\gamma) = \frac{\gamma^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\gamma x}.$$

Core funkce rozdělení je podle (14)

$$T_F(x/\tau|\alpha) = -1 - x f'_{\alpha\gamma}(x)/f_{\alpha\gamma}(x) = \gamma x - \alpha = \alpha(x/\tau - 1)$$

kde $\tau = \alpha/\gamma$ je transformovaný parametr polohy. T_F je lineární, typu $O - N$. Ze soustavy (20) vyplývá, že

$$(\hat{\tau})_{CM} = \bar{X} \qquad (\hat{\alpha}_{CM})^{-1} = n^{-1}\sum_{i=1}^{n}(X_i - \bar{X})^2, \tag{21}$$

což jsou 'náhodou' (díky linearitě T_F) odhady shodné s odhady získané 'obyčejnou' momentovou metodou. Tyto odhady nejsou robustní a podle [10, str. 186] mají dost nízké ARV.

Robustní verzi odhadů (21) lze však najít úplně jednoduše. Definujme m-rozměrný *Huberův momentový estimátor* $\hat{\theta}_{MM}$ jako

$$(22) \qquad \hat{\theta}_{MM}: \qquad \frac{1}{n}\sum_{i=1}^{n}[T_b^*(X_i|\theta)^k - M_k(\theta)] = 0, \qquad 1 \le k \le m,$$

kde

$$T_b^*(x|\theta) = \max[-b, \min(T_F(x|\theta), b)]$$

je zobecněním Huberovy funkce $\psi(x) = \max[-b, \min(x, b)]$ a b je i v $m-$rozměrném případě jediným 'ladicím' parametrem, který je třeba určit (na základě apriorních znalostí nebo požadavků na ARV hledaného odhadu). Z teorie m-dimenzionálních M-estimátorů [11] opět plyne silná konzistence a asymptotická normalita odhadů (22). Podobně jako CM odhady, jsou i MM-odhady (neboli momentové M-odhady) výpočetně neobyčejně jednoduché.

Podívejme se nejprve na dílčí případ gamma rozdělení, na *exponenciální* rozdělení s hustotou $f_\tau(x) = \tau^{-1}e^{-x/\tau}$. Zobecněná Huberova funkce MM estimátoru je asymetrická,

$$T_b^\star(x|\tau) = \begin{cases} x/\tau - 1 & \text{pro} \quad x \le b\tilde{\tau} \\ b - 1 & \text{pro} \quad x > b\tilde{\tau}. \end{cases}$$

Zde $\tilde{\tau}$ je nějaký počáteční odhad τ (používám medián). MM-odhad parametru τ je tedy $\hat{\tau}_{MM} = \bar{X}^\star$, kde $X_i^\star = X_i$ pro $X_i \le b\tilde{\tau}$ a $X_i^\star = b-1$ pro $X_i > b\tilde{\tau}$. Podle Věty 1 v [8, str. 117] je $\hat{\tau}_{MM}$ 'optimal B-robust'. Spočtený asymptotický rozptyl odhadu $\hat{\tau}_{MM}$ je

$$Var(\hat{\tau}_{MM}) = \frac{1 - 2be^{-b}}{(1-(b+1)e^{-b})^2}\tau^2.$$

Ježto $Var(\hat{\tau}_{ML}) = \tau^2$, dostáváme následující tabulku:

b	2	2.5	3	3.5	4
$ARV(\hat{\tau}_{MM})$	0.769	0.862	0.915	0.947	0.967

Vraťme se ke gamma rozdělení. Dosadíme-li do vzorců (21) hodnoty $X_i^\star$ a $\bar{X}^\star$ namísto X_i a $\bar{X}$, dostaneme např. pro $\alpha = 2$ asymptotický rozptyl odhadu $\hat{\tau}_{MM}$ ve tvaru

$$Var(\hat{\tau}_{MM}) = \frac{1 - e^{-2b}(2b^2 + b - 1/2)}{(1 - e^{-2b}(2b^2 + 2b + 1))^2}\tau^2$$

a tabulku

b	2	2.5	3	3.5	4
$ARV(\hat{\tau}_{MM})$	0.703	0.849	0.927	0.969	0.984

takže s rostoucím α mohou klesat hodnoty 'ladicího' parametru.

V další tabulce jsou zachyceny výsledky simulačního experimentu. Data byla generována z kontaminovaného rozdělení $F(x) = (1-\epsilon)F_1(x) + \epsilon F_\rho(x)$, kde F_τ značí exponenciální rozdělení. Testovány byly čtyři estimátory: maximálně věrohodný estimátor $\hat{\tau}_{\mathrm{ML}}$ není robustní, systematické vychýlení odhadů získaných pomocí Huberova H15 estimátoru (nejlepší estimátor ze studie [5]) je způsobeno nesymetrií rozdělení a α-odhady ($\hat{\tau}_\alpha : \frac{1}{n}\sum_{i=1}^{n}(X_i/\tau - 1)f_\tau^\alpha(X_i) = 0$, viz [14]) příliš závisí na hodnotě α, která opět závisí na 'neznámém' ρ. MM estimátor při volbě $b = 3.25$ a $\tilde{\tau}$ rovném mediánu dává docela slušné výsledky. V tabulce uvádíme střední hodnoty odhadů $\bar{\tau}$ z 500 experimentů a rozsahem výběru $n = 200$ pro několik hodnot ρ a $\epsilon = 0.05$.

ρ	$\bar{\tau}_{ML}$	$\bar{\tau}_{H15}$	$\bar{\tau}_{MM}$	$\bar{\tau}_{\alpha=0.1}$	$\bar{\tau}_{\alpha=0.2}$
0	1.01	0.89	0.95	0.95	0.89
2	1.11	0.92	0.97	1.04	0.97
5	1.42	0.93	1.10	1.12	1.05
8	1.74	0.95	1.08	1.37	1.04

Zobecněnou a 'robustifikovanou' momentovou metodou lze v případě gamma rozdělení získat minimálně slušné počáteční hodnoty pro jemnější, ale mnohem komplikovanější postup navržený v [13]. Existuje však řada nesymetrických rodin, pro které není robustní verze ML odhadů známa (např. beta rozdělení, obecná rodina transformovaného beta (viz [12]) nebo rodina z příkladu 5 v [1]) a pro které obvyklé symetrické robustní estimátory nefungují (podobně jako H15 pro výběry z exponenciálního rozdělení). V takových případech je navržený MM-odhad možná jedinou metodou, která může poskytnout jakési ne úplně špatné výsledky.

Literatura

(1) Fabián Z. (1997). Geometrické momenty, *Sborník Robust'96*, JČMF, 49-62.

(2) Fabián Z. (1997). On the relation between gnostical and probability theories, *Kybernetika,* 33, 259-270.

(3) Fabián, Z. (1997). Information and entropy of continuous random variables, *IEEE Trans. on Information Theory*, 43, 1080-1083.

(4) Fabián Z. (2001). Induced cores and their use in robust parametric estimation, *Communications in Statistics, Theory and methods,* 30, 3.

(5) Andrews D. F., Bickel D.J., Hampel F.R., Huber P.J., Rogers W.H. and Tukey J.W. (1972), *Robust estimates of location*, Princeton Univ. Press, Princeton.

(6) Barnett V. (1999). *Comparative statistical inference*, Wiley, New York.

(7) Guiard V. (1984). Systems of One-Dimensional Continuous Distributions and their Application in Simulation Studies. In *Robustness of Statistical Methods and Nonparametric Statistics*, Rash D. and Tiku M.L. (eds.), VEB Deutscher Verlag den Wissenschaften, Berlin.

(8) Hampel F.R., Rousseeuw P.J., Ronchetti E.M., Stahel W.A. (1987). *Robust Statistic. The Approach Based on Influence Functions,* J. Wiley, New York.

(9) Johnson N. L. (1949). Systems of Frequency Curves Generated by Methods of Translations, *Biometrika,* 36, 149.

(10) Johnson N.L., Kotz S. (1970). *Continuous univariate distributions 2*, Houghton Mifflin, Boston.

(11) Jurečková J., Sen P.K. (1996). *Robust statistical procedures: asymptotics and interrelations,* Wiley, New York.

(12) Klugmann S.A., Panjer H.H., Willmott G.E. (1998). *Loss models. From data to decisions,* (Wiley, New York).

(13) Marazzi A., Ruffieux C. (1996). Implementing M-estimators of the Gamma distribution, in *Robust Statistics, Data Analysis and Computer Intensive Methods* (H.Rieder edt.), Springer Lecture Notes in Statistics 109, 277-297.

(14) Vajda I. (1986). Efficiency and robustness control via distorted maximum likelihood estimation, *Kybernetika,* 22, 47-67.

Ústav informatiky AV ČR, Pod Vodárenskou Věží 2, 182 07 Praha 8
E-mail: `zdenek@uivt.cas.cz`