ROBUST'2000, 42 – 49

ODHAD PARAMETRU CHVOSTŮ DISTRIBUČNÍ FUNKCE

ALENA FIALOVÁ

Abstrakt. Distribution of maximum of the sequence of independent identically distributed random variables with nondegenerate distribution function converges (after proper normalization) to the one of the three distribution functions, extreme value distribution function. The shape of this distribution is affected by parameter γ. Estimation of this parameter is discussed and new estimator based on sample mean and is proposed.

Функция распределения максимума последобательности независимых одинаково распределённых случайных величин с невырожденной функцией распределения сходится (после подходящей нормализации) к одной из трёх фынкций распределения, функции распределения экстремального значения. Параметр γ влияет на вид этого распределенуя. Обсуждаются и предлагаются новые оценки этого параметра основанные на выборке среднего значения.

1. Úvod

Předmětem tohoto článku je problematika odhadu parametru chvostů rozdělení s těžkými chvosty. Nejdříve uvedeme základní výsledky v tomto oboru. Budeme se zabývat teorií extrémních hodnot, jejíž výsledky využijeme.

1.1. Teorie extrémních hodnot.

Uvažujme posloupnost nezávislých náhodných veličin $X_i,\ i = 1, 2, \dots$ s nedegenerovanou distribuční funkcí $F(x)$. Označme *výběrové maximum* $M_n = \max(X_1, \dots, X_n),\ \ n \geq 1$. Základním problémem teorie extrémních hodnot je otázka, za jakých podmínek existuje asymptotické rozdělení pravděpodobností posloupnosti veličin $\{\frac{M_n - d_n}{c_n}\}$, při vhodné standardizaci posloupnostmi $\{c_n\}_{n=1}^{\infty}$, $\{d_n\}_{n=1}^{\infty}$. Odpověď dává následující věta, základní věta klasické teorie extrémních hodnot.

Věta 1.1. (Fisher-Tippett)

Nechť $\{X_n\}$ je posloupnost nezávislých stejně rozdělených náhodných veličin s nedegenerovanou distribuční funkcí $F(x)$. Pak existují posloupnosti $c_n > 0$, d_n tak, že pro $M_n = \max(X_1, \dots, X_n),\ n \geq 1$, platí

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} H_\gamma \ , \tag{1.1}$$

$$H_\gamma = \begin{cases} \Phi_{1/\gamma} & \gamma > 0 \ , \\ \Lambda & \gamma = 0 \ , \\ \Psi_{-1/\gamma} & \gamma < 0 \ . \end{cases} \tag{1.2}$$

H_γ je jedním ze tří typů:
Frèchet: $\Phi_m(x) = \exp\{-x^{-m}\},\ x > 0,\ m > 0$

2000 *Mathematics Subject Classification.* Primary 62F10; Secondary 62P05.
Klíčová slova. Rozdělení s těžkými chvosty, Hillův odhad, extrémních hodnoty.
Tato práce vznikla za podpory grantu GAČR č. 201/00/0770 a grantu MSM 113200008.

Gumbel: $\Lambda(x) = \exp\{-e^{-x}\},\ x \in \mathbb{R}$
Weibull: $\Psi_m(x) = \exp\{-(-x^m)\},\ x \leq 0,\ m > 0$

Důkaz. Důkaz věty je poněkud komplikovaný a technický, ale ve zjednodušené formě lze říci, že jde o řešení funkcionálních rovnic. Volbou $c_n = 1$ pro vš. n se dokáže, že F je typu Λ. Jestliže existuje n_0, takové, že $c_n < 1$ pro všechna n, pak F je typu Ψ_m pro nějaké $m > 0$. Pokud ex. n_0, takové, že $c_n > 1$ pro všechna n, pak F je typu Φ_m pro nějaké $m > 0$. Náznak důkazu lze nalézt v Embrechts [4] nebo podrobněji de Haan v [5]. □

Lze ukázat, že distribuční funkce H_0 je rovna limitě H_γ pro $\gamma \to 0$, proto ji můžeme psát ve tvaru

$$H_\gamma(y) = \exp\{-(1+\gamma y)^{-1/\gamma}\}, \quad y > -\gamma^{-1},\ \gamma \in \mathbb{R}\ . \tag{1.3}$$

Funkce $H_\gamma(x)$ se nazývá *distribuční funkce standardního rozdělení extrémních hodnot* a parametr γ se nazývá *parametr rozdělení extrémních hodnot.* Jelikož funkce H_γ je jedinou možnou nedegenerovanou limitní distribuční funkcí rozdělení standardizovaných extrémních hodnot nazývá se také *max-stable.*

Říkáme, že distribuční funkce F náleží do *sféry přitažlivosti* $H\gamma$, a píšeme $F \in D(H\gamma)$, jakmile platí (1.1) pro nějaké posloupnosti $\{c_n\}$, $\{d_n\}$.

Hodnoty parametru $\gamma = 0$, $\gamma > 0$, $\gamma < 0$ distribuční funkce H_γ odpovídají kvalitativně naprosto odlišným rozdělením pravděpodobností. Proto odhad na základě pozorování je základním problémem nejen z matematického hlediska, ale i z hlediska aplikací. Zejména je důležité rozlišit mezi $\gamma = 0$ (exponenciální pravý chvost) a $\gamma > 0$ (algebraický, těžký pravý chvost).

1.2. **Charakteristika tří sfér přitažlivosti.**

V tomto odstavci nejprve připomeneme některé základní definice a zaměříme se na charakterizaci tří typů limitního rozdělení maxima.

Definice 1.2. Nechť F je distribuční funkce, pak funkce $U(t)$ daná vztahem

$$U(t) := F^{\leftarrow}(1-1/t) = \inf\{y : F(y) \geq 1-1/t\}, \quad t > 0\ , \tag{1.4}$$

se nazývá *kvantilová funkce chvostu.*

Nyní připomeneme definici pravidelně se měnící funkce a pomalu se měnící funkce. Obě tyto definice jsou důležité pro následnou charakterizaci distribuční funkce rozdělení s těžkými chvosty.

Definice 1.3. Řekneme, že funkce $G(t)$ je *pravidelně se měnící funkce (v nekonečnu)* s indexem m a značíme $G \in R_m$, jestliže

$$\lim_{x\to\infty} \frac{G(xt)}{G(x)} = t^m, \quad t > 0\ . \tag{1.5}$$

Funkce $G(t)$, $G \in R_0$ se nazývá *pomalu se měnící funkce (v nekonečnu).*

V následujícím textu uvedeme stručnou charakteristiku distribučních funkcí, které náležejí příslušným sférám přitažlivosti. V označení budeme užívat $\gamma = \frac{1}{m},\ \gamma, m \neq 0$.

• 1. Frèchetovo rozdělení: $F \in D(H_\gamma)$, $\gamma > 0$

Nechť F patří do sféry přitažlivosti Frèchetova rozdělení s distribuční funkcí $\Phi_m(x) = \exp\{-x^{-m}\},\ x > 0,\ m > 0$. Tato sféra přitažlivosti zahrnuje rozdělení s těžkými chvosty s distribuční funkcí F, s odpovídající kvantilovou funkcí chvostu U a zprava neomezeným nosičem, která se dají charakterizovat následujícím způsobem:

1. $1-F(x)=x^{-m}L(x)$, L je pomalu se měnící funkce ($L(x)\in R_0$)
2. $F\in D(\Phi_m)\Leftrightarrow 1-F\in R_{-m}\Leftrightarrow U\in R_\gamma$
3. Postačující podmínka pro ověření, zda náhodná veličina s distribuční funkcí F náleží do Frèchetovy sféry přitažlivosti je von Mises podmínka: Nechť $F(x)$ je absolutně spojitá distribuční funkce s hustotou $f(x)$ a

$$\lim_{x\to\infty}\frac{xf(x)}{1-F(x)}=m,\quad \gamma>0\ ,$$

pak $F\in D(\Phi_m)$.

Příklady:

rozdělení	hustota		parametr γ
Cauchy	$\frac{1}{\pi(1+x^2)}$,	$x\in\mathbb{R}$	1
Pareto	$\frac{m}{x^{m+1}}$,	$x\geq 1$	$1/m$
Loggamma	$\frac{\lambda^m}{\Gamma(m)}(\log x)^{\lambda-1}x^{-m-1}$,	$x>1,\ \lambda,m>0$	$1/m$

• 2. Gumbelovo rozdělení: $F\in D(H_0)$
Tato sféra přitažlivosti zahrnuje širokou škálu rozdělení včetně normálního. Tato rozdělení mohou mít nosič zprava neomezený, může však být i konečný. Její charakterizace je následující:

1. $F\in D(\Lambda)\Leftrightarrow \exists a(x)>0:\lim_{x\to x_F}\frac{1-F(x+ta(x))}{1-F(x)}=e^{-t},\ t\in\mathbb{R}$
2. $F\in D(\Lambda)\Leftrightarrow \exists c(x),a(x),g(x),\ \exists z<x_F: 1-F(x)=c(x)\exp\left\{-\int_z^x\frac{g(t)}{a(t)}\mathrm{d}t\right\}$,

$c(x)\to c>0$, $g(x)\to 1$, pro $x\to x_F$, $a(x)>0$, $\lim_{x\to x_F}a'(x)=0$.

Příklady:

rozdělení	hustota		parametr γ
Exponenciální	$\lambda\exp(-\lambda x)$,	$\lambda>0$	0
Weibullovo	$\lambda\kappa x^{\kappa-1}\exp(-\lambda x^\kappa)$,	$x>0,\ \lambda,\kappa>0$	0
Gamma	$\frac{\lambda^\kappa}{\Gamma(\kappa)}x^{\kappa-1}\exp(-\lambda x)$,	$x>0,\ \lambda,\kappa>0$	0
Normální	$\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)$,	$x\in\mathbb{R}$	0

• 3. Weibullovo rozdělení: $F\in D(H_\gamma)$, $\gamma<0$
Poslední sféra přitažlivosti obsahuje rozdělení, která mají zprava omezený nosič hustoty. Její charakterizace je obdobná jako u Frèchetovy sféry přitažlivosti. Nechť x_F je pravý koncový bod, pak

1. $1-F(x_F-1/x)=x^{-m}L(x),\ L(x)\in R_0$
2. $F\in D(\Psi_{-m})\Leftrightarrow x_F<\infty$ a $F(x_F-1/x)\in D(\Phi_{-m})$

Příklady:

rozdělení	hustota		parametr γ
Rovnoměrné	1,	$x\in(0,1)$	-1
Beta	$\frac{\Gamma(\lambda+m)}{\Gamma(\lambda)\Gamma(m)}x^{\lambda-1}(1-x)^{m-1}$,	$x\in(0,1),\ \lambda,m>0$	$-1/m$

2. Odhad parametru rozdělení extrémních hodnot γ

Tvar rozdělení extrémních hodnot je významně ovlivněn parametrem γ. Kladná hodnota γ indikuje těžké chvosty rozdělení F; pokud navíc $0<m<2$, kde $m=1/\gamma$, nemá F konečný druhý moment. S takovýmto případem se můžeme často setkávat při analýze dat z oblasti věcného neboli neživotního pojištění, příjmů obyvatel, katastrofických událostí a dalších. Proto je důležité odhadnout γ nebo zkonstruovat interval spolehlivosti pro γ, případně vhodný test o jeho hodnotě. Problémem odhadu parametru γ se zabývala řada autorů, např. Csörgő, Deheuvels a Mason [1], Dekkers a de Haan [3] a Pickands [9]. Odhady jsou vesměs založeny na pořádkových statistikách příslušných výběru $X_1, \ldots, X_n$, které označíme $X_{1:n} \leq \cdots \leq X_{n:n}$.

2.1. Hillův odhad parametru $m = 1/\gamma > 0$.

V této práci uvedeme pouze několik řádek o Hillově odhadu. Byly navrženy i další odhady založené na podobném principu, můžeme uvést např. Pickandsův a momentový odhad. Ale Hillův odhad $H_{k,n}$ je nejužívanější a patrně nejpopulárnější odhad parametru $m = 1/\gamma$. Je založený na $k+1$ největších pozorováních $X_{n-k:n}, \ldots, X_{n:n}$:

$$H_{k,n} := (\frac{1}{k}\sum_{i=1}^{k} \log X_{n-i+1:n} - \log X_{n-k:n})^{-1}, \quad k = 1, \ldots, n-1 \ .$$

To lze přepsat do tvaru $H_{k,n} := (\frac{1}{k}\sum_{i=1}^{k} \log \frac{X_{n-i+1:n}}{X_{n-k:n}})^{-1}$, což lze lépe interpretovat. Pokud volíme k_n největších pozorováních v závislosti na n při $n \to \infty$ jako $k_n \to \infty$, $k_n/n \to 0$ pak lze dokázat slabou konzistenci $H_{k_n,n}$

$$H_{k_n,n} \xrightarrow{P} m \ .$$

Byala také dokázána silná konzistence odhadu $H_{k,n}$, tedy $H_{k_n,n} \to m \quad a.s.$ pro $n \to \infty$, za podmínky, že $k_n/\log\log n \to \infty$ pro $n \to \infty$.

Hodnota k se v praxi určuje pomocí grafu $\{(k, H_{k,n}),\ k = 1, \ldots, n-1\}$ zvaného *Hill plot*; přesněji řečeno, hodnota parametru m je odvozena ze stabilní části grafu. To se ukazuje jako největší nevýhoda této techniky, protože Hill plot se může chovat značně nestabilně a kolem m stráví jen poměrně krátký úsek. Na druhou stranu je nutno poznamenat, že Hillův odhad je nejefektivnější, když dané rozdělení je Pareto nebo blízké tomuto rozdělení.

Věta 2.1. (Asymptotická normalita Hillova odhadu)

Nechť $\{X_n\}$ je posloupnost nezávislých, stejně rozdělených náhodných veličin s distribuční funkcí $1-F(x) = = x^{-m}L(x)$, $m > 0$, kde $L(x)$ je pomalu se měnící funkce. Jestliže $\lim_{n\to\infty} k_n = \infty$ a jsou-li splněny další podmínky regularity, pak

$$\sqrt{k}_n(H_{k_n,n} - m) \xrightarrow{d} N(0, m^2) \ . \tag{2.1}$$

Postačující podmínky pro platnost (2.1) výrazně závisí na neznámé distribuční funkci $F(x)$ a dalším parametru, kterým je míra konvergence $(1-F(tx))/(1-F(x))$ k t^m. (Viz též Davis a Resnick [2]).

3. Nový odhad $m = \frac{1}{\gamma}$ založený na chvostech výběrového průměru

Navrhneme odhad parametru rozdělení extrémních hodnot založený na novém principu, na chování chvostů rozdělení průměru n náhodných veličin.

Nejprve zavedeme dva typy chvostů. Nechť $X_1, \ldots, X_n$ jsou nezávislé stejně rozdělené náhodné veličiny s distribuční funkcí $F(x)$, obecně neznámou. Předpokládejme, že k distribuční funkci $F(x)$ přísluší absolutně spojitá hustota $f(x)$ s nosičem zprava neomezeným. Rozlišujeme tyto dva typy chvostů:

1. *Exponenciálně klesající chvost*
je chvost typu $\exp(-bx^r)$, $b > 0$, $r \geq 1$, tedy

$$\lim_{x\to\infty} \frac{-\log(1-F(x))}{bx^r} = 1 \ . \tag{3.1}$$

Jestliže $F(x)$ má exponenciálně klesající chvost, pak F patří do sféry přitažlivosti Gumbelova rozdělení, např. volbou $a(x) = \frac{1}{br}x^{1-r}$ a

$$1 - F(x) = \exp\left\{-\int_0^x brt^{r-1}\mathrm{d}t\right\}.$$

2. *Chvost Paretova typu (rozdělení s těžkým chvostem)*
je chvost typu bx^{-m}, $b, m > 0$, tedy

$$\lim_{x\to\infty} \frac{-\log(1-F(x))}{m\log x} = 1 \ . \tag{3.2}$$

Jestliže $F(x)$ má chvost Paretova typu (a existuje-li hustota $f(x)$), pak $F(x)$ patří do sféry přitažlivosti Frèchetova rozdělení; $F \in D(H_\gamma)$, $\gamma > 0$, $m = 1/\gamma$, jak také můžeme ověřit z von Misesovy podmínky:

$$\lim_{x\to\infty} \frac{xf(x)}{1-F(x)} = m = \frac{1}{\gamma} \ .$$

Příklad. 3.1. *1. Logistické a dvojitě-exponenciální rozdělení mají exponenciálně klesající chvosty s $r = 1$. Standardizované normální rozdělení $N(0,1)$ má exponenciálně klesající chvosty s $r = 2$ a $b = \frac{1}{2}$.*
2. Cauchyho rozdělení má těžké chvosty s parametrem $m = 1$. Studentovo rozdělení má chvosty Paretova typu s parametrem m rovným počtu stupňů volnosti.

Následující věta nám dává představu o tom, kolikrát rychleji klesají chvosty rozdělení $\bar{X}_n$ k nule ve srovnání s chvosty základního rozdělení F.

Věta 3.1. *Jestliže $X_1, \dots, X_n$ je náhodný výběr z rozdělení s distribuční funkcí $F(x)$ a hustotou $f(x)$ takový, že*
a) $F(x)$ má chvosty Paretova typu pro nějaké $m > 0$,
b) $f(x) = 0$ pro $x < 0$ a $0 < f(x) < \infty$ pro $x \geq K_f \geq 0$.
Pak

$$\lim_{x\to\infty} \frac{-\log P(\bar{X}_n > x)}{-\log P(X_1 > x)} = 1 \ , \tag{3.3}$$

kde $\bar{X}_n = \frac{1}{n}\sum_{k=1}^n X_k$ je výběrový průměr.

Důkaz. Jestliže $X_1, \dots, X_n$ jsou nezáporné náhodné veličiny s distribuční funkcí $F(x)$, pak pro výběrový průměr $\bar{X}_n$ platí

$$\begin{aligned} P(\bar{X}_n > x) &= P(\textstyle\sum_{i=1}^n X_i > nx) \geq P(\max_{1\leq i\leq n} X_i > nx) = \\ &= 1 - (F(nx))^n \geq 1 - F(nx) \ . \end{aligned}$$

Pak pro rozdělení s těžkými chvosty splňující (3.2) s parametrem $m > 0$ máme

$$\limsup_{x\to\infty} \frac{-\log P(\bar{X}_n > x)}{-\log P(X_1 > x)} \leq \lim_{x\to\infty} \frac{-\log(1-F(nx))}{-\log P(X_1 > x)} = \lim_{x\to\infty} \frac{m\log nx}{m\log x} = 1 \ .$$

Podobně omezíme $P(\bar{X}_n > x)$ shora:

$$P(\bar{X}_n > x) \leq P(\max_{1\leq i\leq n} X_i > x) = 1 - F(x)^n \leq n(1-F(x)) \ ,$$

a pak dostáváme

$$\liminf_{x\to\infty} \frac{-\log P(\bar{X}_n > x)}{-\log P(X_1 > x)} \geq \lim_{x\to\infty} \frac{-\log n(1-F(x))}{-\log P(X_1 > x)} =$$

$$= \lim_{x\to\infty} \frac{-\log n + m \log x}{m \log x} = 1 \ .$$

Tím je limita (3.3) dokázána. □

Poznámka 3.2. Jurečková v [6] dokázala podobný výsledek, kde je uvažován pravý i levý chvost a symetrická hustota náhodné veličiny. Pokud uvažujeme exponenciálně klesající chvosty, pak chvosty výběrového průměru klesají k nule n krát rychleji než chvosty daného rozdělení, přesněji

$$\lim_{x\to\infty} \frac{-\log P(|\bar{X}_n| > x)}{-\log P(|X_1| > x)} = n \ .$$

Věta (3.1) umožňuje zkonstruovat nový odhad parametru m, založeného na výběrových průměrech. Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozdělení s těžkými chvosty, vyhovujícími (3.1). Pak podle (3.3) platí

$$\lim_{x\to\infty} \frac{-\log P(\bar{X}_n > x)}{m \log x} = 1 \tag{3.4}$$

a tedy

$$m = \lim_{x\to\infty} m_n(x) \ , \tag{3.5}$$

kde

$$m_n(x) = \frac{-\log P(\bar{X}_n > x)}{\log x} = \frac{-\log(1 - F_{\bar{X}_n}(x))}{\log x} \ . \tag{3.6}$$

To znamená, parametr m je funkcí chvostů rozdělení pravděpodobností výběrového průměru $\bar{X}_n$, tj. funkcí $-\log(1 - F_{\bar{X}_n}(x))$. Neznámou distribuční funkci $F_{\bar{X}_n}$ aproximujeme empirickou distribuční funkcí, stanovenou z k realizací výběrového průměru $\bar{X}_n$. Protože nás vzhledem k (3.5) zajímá průběh $1 - F_{\bar{X}_n}(x)$ pro velká x, musíme jako x volit vhodnou hodnotu x_k, dostatečně velkou, ale menší než $\max_{1\leq i\leq n} X_i$.

Uvažujme situaci, že na základě předběžné informace víme, že rozdělení náhodné veličiny X je Paretova typu s indexem $m \leq m_0$, $0 < m \leq m_0 < \infty$. Tento závěr můžeme učinit na základě minulých zkušeností nebo z povahy experimentu (např. jedná se o rozdělení pravděpodobností důchodů, příjmů apod.) případně na základě předběžného testu hypotézy $H : m \leq m_0$ proti alternativě $K : m > m_0$. Takový test navrhli Jurečková a Picek v [7]. V prvním přiblížení může být m_0 dost hrubé. Uvažujme k náhodných výběrů o rozsahu n, které označíme $(X_1^1, \ldots, X_n^1)$, $\ldots$, $(X_1^k, \ldots, X_n^k)$. Pak $\bar{X}_n^1, \ldots, \bar{X}_n^k$ jsou příslušné výběrové průměry; vektor $(\bar{X}_n^1, \ldots, \bar{X}_n^k)$ je nezávislý náhodný výběr z rozdělení, jehož distribuční funkci $F_{\bar{X}_n}(x) = P_m(\bar{X}_n \leq x)$ neznáme. Nechť $\hat{F}^k_{\bar{X}_n}(x) = \frac{1}{k}\sum_{i=1}^k I[\bar{X}_n^i \leq x]$ je empirická distribuční funkce založená na $(\bar{X}_n^1, \ldots, \bar{X}_n^k)$.

Zvolme posloupnost $\{x_k\}_{k=1}^\infty$ takovou, že $\lim_{k\to\infty} x_k = \infty$ a $x_k = k^{\frac{1-\delta}{m_0}} L_1(k)$, kde $0 < \delta < \frac{1}{m_0}$ a $L_1(k)$ je pomalu se měnící funkce (v nekonečnu). Uvažujme odhad parametru m ve tvaru:

$$\hat{m}_n(x) = \bar{m}_n(x) I[\hat{F}^k_{\bar{X}_n}(x) < 1] + m_0 I[\hat{F}^k_{\bar{X}_n}(x) = 1] \;, \tag{3.7}$$

kde

$$\bar{m}_n(x) = \frac{-\log\bigl(1 - \hat{F}^k_{\bar{X}_n}(x)\bigr)}{\log x} \;, \quad x > 0 \;. \tag{3.8}$$

Následující věta ukazuje, že $\hat{m}_n(x_k)$ je (slabě) konzistentním odhadem m při $k \to \infty$.

Věta 3.3. *Za výše uvedených předpokladů a za podmínek a) a b) Věty 3.1 platí*

$$\lim_{k\to\infty} P_m\left(|\hat{m}_n(x_k) - m| > \varepsilon\right) = 0 \quad \textit{pro lib.} \quad \varepsilon > 0 \;. \tag{3.9}$$

Důkaz. Nejprve dokážeme, že pro posloupnost x_k platí

$$P_m(\hat{F}^k_{\bar{X}_n}(x_k) < 1) \longrightarrow 1 \quad \text{pro} \quad k \to \infty,\; x_k \to \infty \;,$$

tedy, že čitatel (3.8) má smysl s pravděpodobností blížící se k 1 a $\bar{m}_n(x_k)$ je definován. Pro rozdělení s těžkými chvosty s distribuční funkcí $F(x)$ a distribuční funkcí výběrového průměru $F_{\bar{X}_n}(x)$, víme

$$\lim_{x\to\infty} \frac{-\log(1 - F(x))}{m \log x} = \lim_{x\to\infty} \frac{-\log(1 - F_{\bar{X}_n}(x))}{m \log x} = 1 \;. \tag{3.10}$$

Pokud existují příslušné hustoty $f(x)$ a $f_{\bar{X}_n}(x)$, pak $F \in D(H_{1/m})$ a také $F_{\bar{X}_n} \in D(H_{1/m})$. Označme $\bar{X}_n^{(k)} = \max(\bar{X}_n^1, \ldots, \bar{X}_n^k)$. Pak z (3.10) a s použitím von Misesovy podmínky rozdělení maxima náleží do $D(\Phi_m)$, sféry přitažlivosti Frèchetova rozdělelní s disribuční funkcí $\Phi_m(x) = \exp\{-x^{-m}\},\; x > 0$. Tedy můžeme psát

$$P_m\Bigl(\frac{\bar{X}_n^{(k)}}{c_k} \le x\Bigr) \longrightarrow \Phi_m(x) \;, \quad k \to \infty \;,$$

kde $c_k = k^{\frac{1}{m}} L_2(k)$ a $L_2(k)$ je pomalu se měnící funkce (v nekonečnu). Z předpokladů pro posloupnost $\{x_k\}_{k=1}^{\infty}$ dostáváme

$$P_m(\bar{X}_n^{(k)} \le x_k) = P_m\left(\frac{\bar{X}_n^{(k)}}{c_k} \le \frac{x_k}{c_k}\right) \longrightarrow 0 \;, \quad k \to \infty \;, \tag{3.11}$$

protože

$$0 \le \frac{x_k}{c_k} = \frac{k^{\frac{1-\delta}{m_0}} L_1(k)}{k^{\frac{1}{m}} L_2(k)} \le k^{-\frac{\delta}{m_0}} \frac{L_1(k)}{L_2(k)} \to 0 \quad \text{pro} \quad k \to \infty$$

a $\Phi_m(0) = 0$.
Konečně spolu s (3.11), dostáváme $\lim_{k\to\infty} P_m(\hat{F}^k_{\bar{X}_n}(x_k) < 1) = 1$.

Předpokládáme, že $\{\bar{X}_n^i\}_{i=1}^k$ jsou nezávislé stejně rozdělené náhodné veličiny s distribuční funkcí $F_{\bar{X}_n}$, pak z Glivenko-Cantelliho věty plyne

$$\sup_{x\in\mathbb{R}} |\hat{F}^k_{\bar{X}_n}(x) - F_{\bar{X}_n}(x)| \longrightarrow 0 \quad \text{a.s.} \quad \text{pro} \quad k \to \infty \ . \tag{3.12}$$

Dle (3.8) můžeme psát

$$\begin{aligned} \bar{m}_n(x_k) &= \frac{-\log\left(1-\hat{F}^k_{\bar{X}_n}(x_k)\right)}{\log x_k} = \\ &= \frac{-\log \frac{1-\hat{F}^k_{\bar{X}_n}(x_k)}{1-F_{\bar{X}_n}(x_k)}}{\log x_k} + \frac{-\log\left(1-F_{\bar{X}_n}(x_k)\right)}{m \log x_k}\, m \ . \end{aligned}$$

Pro každé $\varepsilon > 0$ existuje k_0 takové, že $\forall k \geq k_0$ je x_k dostatečně velké a oba členy mohou být omezeny

$$\left| \frac{-\log \frac{1-\hat{F}^k_{\bar{X}_n}(x_k)}{1-F_{\bar{X}_n}(x_k)}}{\log x_k} \right| \leq \varepsilon \quad \text{a} \quad \left| 1 - \frac{-\log\left(1-F_{\bar{X}_n}(x_k)\right)}{m\log x_k} \right| \leq \varepsilon \ .$$

Nyní užijeme obě meze pro $\bar{m}_n(x_k)$, dostáváme

$$-\varepsilon + (1-\varepsilon)m \leq \bar{m}_n(x_k) \leq \varepsilon + (1+\varepsilon)m$$

a odtud

$$|\bar{m}_n(x_k) - m| \leq \varepsilon + \varepsilon m \longrightarrow 0 \quad \text{pro} \quad x_k \to \infty \ .$$

Tedy odhad $\hat{m}_k(x_k)$ je dobře definován a je slabě konzistentním odhadem parametru m. □

Literatura

[1] Csörgő, S., Deheuvels, P. and Mason, D. (1985). Kernel estimates of the tail index of a distributions. *Ann. Statist.*, **13**, 1050–1077

[2] Davis, R.A. and Resnick, S.T. (1984). Tail estimates motivated by extreme value theory. *Ann. Statist.*, **12**, 1467–1487

[3] Dekkers, A.L.M., Einmahl, J.H.J. and de Haan, L. (1989). A moment estimator for the index of an extreme value distribution. *Ann. Statist.*, **17**, 1833–1855

[4] Embrechts, P., Klüppelbelg, C. and Mikosch, T. (1997). *Modelling Extremal Events.* Springer-Verlag.

[5] de Haan, L. (1975). On regular variation and its application to the weak convergence of sample extremes. *Math Centre Tracts* **32**. Centre for Mathematics and Computer Science, Amsterdam.

[6] Jurečková, J. (1981). Tail-behavior of location estimators. *Ann. Statist.*, **9**, 578–585

[7] Jurečková, J. (2000). Test of tails based on extreme regression quantiles. *Statist & Probab. Letters*, **49**, 53–61

[8] Mason, D.M. (1982). Laws of large numbers for sums of extreme value. *Ann. Probab.*, **10**, 756–764

[9] Pickands, J. III (1975). Statistical inference using extreme order statistics. *Ann. Statist.*, **3**, 119–131

UK MFF, KPMS, Sokolovská 83, 186 75 Praha
E-mail: `fialova@karlin.mff.cuni.cz`