第 48 卷 第 4 期　西南师范大学学报(自然科学版)　2023 年 4 月
Vol. 48 No. 4　Journal of Southwest China Normal University (Natural Science Edition)　Apr. 2023

DOI:10.13718/j.cnki.xsxb.2023.04.010

一类位置不变的极值指数估计量①

蒲钰瑶，　陈守全

西南大学 数学与统计学院，重庆 400715

摘要：本文提出了一类位置不变的重尾极值估计量，

$$\hat{\gamma}_n(k_0, k, r)=\frac{\frac{1}{k_0}\sum_{i=1}^{k_0}\left(\frac{1}{r}\left(\left(\frac{X_{n-i,n}-X_{n-k,n}}{X_{n-k0,n}-X_{n-k,n}}\right)^r-1\right)\right)}{1+\frac{r}{k_0}\sum_{i=1}^{k_0}\left(\frac{1}{r}\left(\left(\frac{X_{n-i,n}-X_{n-k,n}}{X_{n-k0,n}-X_{n-k,n}}\right)^r-1\right)\right)}$$

其中：$\gamma>0$，k_0 是小于 k 的正整数. 得到了此位置不变极值估计量的弱相合性和渐近正态性，并根据其渐近展开式得到 k_0 的最优选择.

关　键　词：重尾估计；二阶正则变换；位置不变；渐近正态性

中图分类号：O211.4　　**文献标志码**：A　　**文章编号**：1000-5471(2023)04-0075-05

A Class of Location-Invariant Extreme Value Index Estimators

PU Yuyao，　CHEN Shouquan

School of Mathematics and Statistics, Southwest University, Chongqing 400715, China

Abstract: In this paper, we propose a class of location-invariant heavy-tailed extreme value estimators

$$\hat{\gamma}_n(k_0, k, r)=\frac{\frac{1}{k_0}\sum_{i=1}^{k_0}\left(\frac{1}{r}\left(\left(\frac{X_{n-i,n}-X_{n-k,n}}{X_{n-k_0,n}-X_{n-k,n}}\right)^r-1\right)\right)}{1+\frac{r}{k_0}\sum_{i=1}^{k_0}\left(\frac{1}{r}\left(\left(\frac{X_{n-i,n}-X_{n-k,n}}{X_{n-k_0,n}-X_{n-k,n}}\right)^r-1\right)\right)}$$

where $\gamma>0$, k_0 is a positive integer less than k. The weak conjunction and asymptotic normality of this location invariant extreme value estimator are obtained, and the optimal choice of k_0 is obtained according to its asymptotic expansion.

Key words: heavy-tailed estimator; second-order regular transform; position invariant; asymptotic normality

设$\{X_n, n\geqslant 1\}$是一列独立同分布的随机变量序列，其共同的分布函数为 $F(x)$，$X_{1,n}\leqslant\cdots\leqslant X_{n,n}$ 为 $X_1,\cdots,X_n$ 的顺序统计量. 如果 F 属于极值吸引场[1]：

$$G_\gamma(x)=\exp(-(1+\gamma x)^{-\frac{1}{\gamma}})$$

其中 $\gamma\in\mathbb{R}$，$1+\gamma x\geqslant 0$. 这意味着，如果存在规范化常数 $a_n>0$，$b_n\in\mathbb{R}$，使得当 $n\to\infty$，对所有 $x\in$

① 收稿日期：2022-03-02
作者简介：蒲钰瑶，硕士研究生，主要从事极值统计的研究.
通信作者：陈守全，博士，副教授.

$\mathbb{R}$，都有

$$F^n(a_n x + b_n) \to G_\gamma(x) \tag{1}$$

则可以记为 $F \in D(G_\gamma)$[2]，这等价于 $U(t):=F^{\leftarrow}\left(1-\left(\frac{1}{t}\right)\right)$ 是指数为 γ 的正则变化函数. 本文主要讨论极值指数 $\gamma > 0$ 的分布函数，即分布函数为重尾分布函数. 对所有的 $x > 0$，

$$\lim_{t\to\infty}\frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\frac{1}{\gamma}} \Leftrightarrow \lim_{t\to\infty}\frac{U(tx)}{U(t)} = x^{\gamma} \tag{2}$$

对于极值指数的研究，当 $\gamma > 0$ 时，文献[3]最早提出了著名的 Hill 估计量. 也有一些基于不同思想的估计量：将样本分块，在每个块中取两个最大值的比率[4-6]，然后将线性函数 $f(x)=x$ 而不是对数函数应用于这些比率. 文献[7]将函数族 $f_r(t)$ 引入到次序统计量中，得到估计量

$$H_n^{(l)} = \frac{1}{k}\sum_{i=0}^{k-1} f_r^l\left(\frac{X_{n-i,n}}{X_{n-k,n}}\right),\ l=1,2 \tag{3}$$

其中

$$f_r(t)=\begin{cases}\frac{1}{r}(t^r-1),\ r\neq 0\\ \ln t,\ r=0\end{cases} \tag{4}$$

借助文献[8]得到广义 Hill 估计量

$$\hat{\gamma}_n(k,r) = \frac{H_n(k,r)}{1+r\,H_n(k,r)} \tag{5}$$

本文主要讨论 $\hat{\gamma}_n(k,r)$ 的位置不变估计量[9-11]，参照文献[12]中的方法，其对应的位置不变估计量的形式为

$$\hat{\gamma}_n(k_0,k,r):=\frac{H_n(k_0,k,r)}{1+rH_n(k_0,k,r)}=\\ \frac{\frac{1}{k_0}\sum_{i=1}^{k_0}\left(\frac{1}{r}\left(\left(\frac{X_{n-i,n}-X_{n-k,n}}{X_{n-k_0,n}-X_{n-k,n}}\right)^r-1\right)\right)}{1+\frac{r}{k_0}\sum_{i=1}^{k_0}\left(\frac{1}{r}\left(\left(\frac{X_{n-i,n}-X_{n-k,n}}{X_{n-k_0,n}-X_{n-k,n}}\right)^r-1\right)\right)} \tag{6}$$

本文在二阶条件[13]下证明其渐近性质.

1 主要结果

在下文中，设 $F \in D(G_\gamma)$，$\gamma > 0$，假设存在一个函数 $A(t) > 0$，有如下二阶条件成立

$$\lim_{t\to\infty}\frac{\frac{U(tx)}{U(t)}-x^{\gamma}}{A(t)} = x^{\gamma}\frac{x^{\rho}-1}{\rho},\ x>0 \tag{7}$$

由条件(2)可得

$$\lim_{t\to\infty}\frac{U(tx)-U(t)}{U(ty)-U(t)} = \frac{x^{\gamma}-1}{y^{\gamma}-1} \tag{8}$$

其中，对任意的中间序列 k，满足

$$k=k_n,\ k_n\to\infty,\ n\to\infty,\ \frac{k_n}{n}\to 0 \tag{9}$$

定理 1 若对 $\gamma > 0$ 有二阶条件(7)成立，且当 $n\to\infty$ 时，中间序列 k 满足条件(9)，则对 $\gamma r < 1$，有

$$\hat{\gamma}_n(k_0,k,r)\xrightarrow{p}\gamma$$

定理 2 假设二阶条件(7)成立，当 $n\to\infty$，中间序列 k_0 和 k 满足 $k\to\infty$，$k_0\to\infty$，且 $\frac{k_0}{k}\to 0$，则对

$r<\dfrac{1}{2\gamma}$，渐近分布表达式为

$$\hat{\gamma}_n(k_0,k,r)\stackrel{d}{=}\gamma+\gamma\frac{\sqrt{V_r^{(\gamma)}}}{\sqrt{k_0}}P_n+b_r(-\gamma)\left(\frac{k_0}{k}\right)^{\gamma}+\frac{b_r(\rho)}{-\rho}\left(\frac{k_0}{k}\right)^{-\rho}A\left(\frac{n}{k}\right)+ o_p\left(\frac{1}{\sqrt{k_0}}\right)+o_p\left(\left(\frac{k_0}{k}\right)^{\gamma}\right)+o_p\left(A\left(\frac{n}{k}\right)\right) \tag{10}$$

其中

$$V_r^{(\gamma)}=\frac{(1-r\gamma)^2}{1-2r\gamma} \tag{11}$$

$$b_r(\rho)=\frac{-\rho(1-r\gamma)}{1-\gamma r-\rho} \tag{12}$$

$$P_n=\frac{(1-\gamma r)\sqrt{(1-2\gamma r)}\left(\frac{1}{k_0}\sum_{i=1}^{k_0}Y_i^{\gamma r}-\frac{1}{1-\gamma r}\right)}{\frac{\gamma r}{\sqrt{k_0}}}\xrightarrow{d}N(0,1) \tag{13}$$

此外，当 $n\to\infty$ 时，如果存在 $\lambda_1\in\mathbb{R}$，$\lambda_2\in\mathbb{R}$ 使得 $\sqrt{k_0}\left(\frac{k_0}{k}\right)^{\gamma}\to\lambda_1$，$\sqrt{k_0}A\left(\frac{n}{k}\right)\to\lambda_2$，那么有

$$\sqrt{k_0}(\hat{\gamma}_n(k_0,k,r)-\gamma)\xrightarrow{d}N(\lambda_1 b_r(-\gamma),\gamma^2V_r(\gamma)) \tag{14}$$

定理 3　假设二阶条件(7)成立，$A(t)\sim ct^{\rho}$，其中 $\rho<0$，$c\neq 0$，令

$$k_0^{(1)}=\left[\frac{\gamma V_r^{(\gamma)}}{2b_r^2(-\gamma)}\right]^{\frac{1}{2\gamma+1}}k^{\frac{2\gamma}{2\gamma+1}}$$

$$k_0^{(2)}=\left[\frac{\gamma^2\rho V_r^{(\gamma)}}{-2c^2b_r^2(\gamma)}\right]^{\frac{1}{-2\rho+1}}n^{\frac{-2\rho}{-2\rho+1}}$$

$$k_0^{(3)}=\left[\frac{cb_r(\rho)}{\rho b_r(-\gamma)}\right]^{\frac{1}{\gamma+\rho}}k^{\frac{\gamma}{\gamma+\rho}}n^{\frac{\rho}{\gamma+\rho}}$$

中间序列 k_0^{opt} 是使得 $AMSE(\hat{\gamma}_n(k_0,k,r))$ 最小的 k_0：

(i) 如果 $\gamma\leqslant-\rho$，则 $k_0^{\mathrm{opt}}=k_0^{(1)}$；

(ii) 如果 $\gamma\geqslant-\rho$，

(a) 若 $k\ll n^{-\rho(2\gamma+1)/(\gamma(-2\rho+1))}$，则 $k_o^{\mathrm{opt}}=k_0^{(1)}$；

(b) $k\gg n^{-\rho(2\gamma+1)/(\gamma(-2\rho+1))}$，若 $c<0$，则 $k_0^{\mathrm{opt}}=k_0^{(3)}$，若 $c>0$，则 $k^{\mathrm{opt}}=k_0^{(2)}$；

(c) 若 $k\sim Dn^{-\rho(2\gamma+1)/(\gamma(-2\rho+1))}$，$D\neq 0$，那么 $k_0^{\mathrm{opt}}\sim D_1 n^{\frac{-2\rho}{-2\rho+1}}$，$D_1\equiv D_1(\gamma,\rho,r,c,D)$ 满足

$$a_1D_1^{2\gamma+1}+a_2D_1^{\gamma-\rho+1}+a_3D_1^{-2\rho+1}=\gamma^2V_r^{(\gamma)}$$

其中

$$a_1=2\gamma b_r^2(-\gamma)D^{-2\gamma},\ a_2=\left[\frac{2c(\rho-\gamma)}{\rho}\right]b_r(-\gamma)b_r(\rho)D^{-\gamma},\ a_3=\frac{-2c^2b_r{}^2(\rho)}{\rho} \tag{15}$$

2　定理的证明

设$\{Y_n,n\geqslant 1\}$是分布函数为 $F_Y(y)=1-\dfrac{1}{y}$，$y\geqslant 1$ 的独立同分布的随机变量序列，$Y_{1,n}\leqslant\cdots\leqslant Y_{n,n}$ 为 $Y_1,\cdots,Y_n$ 的顺序统计量，对于任意中间序列 k，由于$\{X_i\}_{i=1}^n\stackrel{d}{=}\{U(Y_i)\}_{i=1}^n$ 且 $\dfrac{n}{k}Y_{n-k,n}\xrightarrow{p}1$，由 Renyi's 表达式[14] 可得

$$\left\{\frac{Y_{n-i,n}}{Y_{n-k,n}}\right\}_{i=0}^{k-1}\stackrel{d}{=}\{Y_{k-i,k}\}_{i=0}^{k-1} \tag{16}$$

此外，对于序列$\{Y_n, n \geqslant 1\}$，有

$$\begin{cases} E(Y^a) = \dfrac{1}{1-a},\ a < 1 \\ Var(Y^a) = \dfrac{a^2}{(1-a)^2(1-2a)},\ a < \dfrac{1}{2} \end{cases} \tag{17}$$

定理 1 的证明

$$\frac{1}{k_0}\sum_{i=0}^{k_0-1}\left(\frac{X_{n-i,n}-X_{n-k,n}}{X_{n-k_0,n}-X_{n-k,n}}\right)^r \overset{d}{=} \frac{1}{k_0}\sum_{i=0}^{k_0-1}\left(\frac{\left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right)^\gamma - 1}{\left(\frac{Y_{n-k_0,n}}{Y_{n-k,n}}\right)^\gamma - 1}\cdot(1+o_p(1))\right)^r$$

$$\overset{d}{=} \frac{1}{k_0}\sum_{i=0}^{k_0-1}\left(\frac{Y_{k-i,k}^\gamma - 1}{Y_{k-k_0,k}^\gamma - 1}\cdot(1+o_p(1))\right)^r$$

$$\overset{d}{=} \frac{1}{k_0}\sum_{i=0}^{k_0-1}(Y_{k_0-i,k_0}^{\gamma r})(1+o_p(1)) \xrightarrow{p} \frac{1}{1-r\gamma}$$

最后一步由大数定律可得. 又因为

$$\frac{1}{\frac{1}{k_0}\sum_{i=0}^{k_0-1}\left(\left(\frac{x_{n-i,n}-x_{n-k,n}}{x_{n-k_0,n}-x_{n-k,n}}\right)^r - 1\right)} \xrightarrow{p} \frac{1}{\frac{1}{1-\gamma r}-1} = \frac{1-\gamma r}{\gamma r}$$

所以，

$$\hat{\gamma}_n(k_0, k, r) \xrightarrow{p} \left(r\cdot\frac{1-\gamma r}{\gamma r}+r\right)^{-1} = \gamma$$

其中 $r < \dfrac{1}{\gamma}$.

定理 2 的证明　首先定义

$$\hat{x}^r := \frac{1}{k_0}\sum_{i=1}^{k_0}\left(\frac{X_{n-i,n}-X_{n-k,n}}{X_{n-k_0,n}-X_{n-k,n}}\right)^r \tag{18}$$

利用泰勒展式，有

$$\frac{U(Y_{n-i,n})-U(Y_{n-k,n})}{U(Y_{n-k_0,n})-U(Y_{n-k,n})} \overset{d}{=} Y_{k_0-i,k_0}^\gamma\Bigg\{1 - Y_{k_0-i,k_0}^{-\gamma}Y_{k-k_0,k}^{-\gamma} + \frac{Y_{k_0-i,k_0}^\rho Y_{k-k_0,k}^\rho - 1}{\rho}A(Y_{n-k,n}) + Y_{k-k_0,k}^{-\gamma} - \frac{Y_{k-k_0,k}^\rho - 1}{\rho}A(Y_{n-k,n}) + Y_{k_0-i,k_0}^{-\gamma}Y_{k-k_0,k}^{-\gamma}\frac{Y_{k-k_0,k}^\rho - 1}{\rho}A(Y_{n-k,k}) + Y_{k-k_0,k}^{-\gamma}\frac{Y_{k_0-i,k_0}^\rho Y_{k-k_0,k}^\rho - 1}{\rho}A(Y_{n-k,n})\Bigg\}(1+o_p(1))$$

成立. 因此可得

$$\hat{x}^r \overset{d}{=} \frac{1}{k_0}\sum_{i=0}^{k_0-1}Y_i^{\gamma r} + r\left(\frac{k_0}{k}\right)^\gamma\frac{1}{k_0}\sum_{i=1}^{k_0}Y_i^{\gamma r}(1-Y_i^{-\gamma}) + \frac{rA\left(\frac{n}{k}\right)}{\rho}\left(\frac{k_0}{k}\right)^{-\rho}\frac{1}{k_0}\sum_{i=1}^{k_0}Y_i^{\gamma r}(Y_i^\rho - 1) + \frac{rA\left(\frac{n}{k}\right)}{\rho}\left(\frac{k_0}{k}\right)^\gamma\frac{1}{k_0}\sum_{i=1}^{k_0}Y_i^{\gamma r}\left[Y_i^{-\gamma}\left(\left(\frac{k_0}{k}\right)^{-\rho}-1\right)+Y_i^\rho\left(\frac{k_0}{k}\right)^{-\rho}-1\right] + o_p\left(\left(\frac{k_0}{k}\right)^\gamma\right) + o_p\left(A\left(\frac{n}{k}\right)\right)$$

由此可得

$$\hat{x}^r \stackrel{d}{=} \frac{1}{1-\gamma r}\left(1+\frac{\gamma r}{\sqrt{1-2\gamma r}\ \sqrt{k_0}}P_n^{(r)}+\frac{\gamma r}{1+\gamma-\gamma r}\left(\frac{k_0}{k}\right)^{\gamma}+\frac{r}{1-r\gamma-\rho}\left(\frac{k_0}{k}\right)^{-\rho}A\left(\frac{n}{k}\right)+\right.$$
$$\left. o_p\left(\left(\frac{k_0}{k}\right)^{\gamma}\right)+o_p\left(A\left(\frac{n}{k}\right)\right)\right)$$

那么

$$\hat{\gamma}_n(k_0,k,r) \stackrel{d}{=} \gamma+\frac{(1-\gamma r)\gamma}{\sqrt{1-2\gamma r}\ \sqrt{k_0}}P_n+\frac{(1-\gamma r)\gamma}{1+\gamma-\gamma r}\left(\frac{k_0}{k}\right)^{\gamma}+\frac{1-\gamma r}{1-\rho-\gamma r}\left(\frac{k_0}{k}\right)^{-\rho}A\left(\frac{n}{k}\right)+$$
$$o_p\left(\frac{1}{\sqrt{k_0}}\right)+o_p\left(\left(\frac{k_0}{k}\right)^{\gamma}\right)+o_p\left(A\left(\frac{n}{k}\right)\right)$$

这就得到$\hat{\gamma}_n(k_0,k,r)$的渐近分布表达式，由此可得$\hat{\gamma}_n(k_0,k,r)$的渐近正态性(14). 证毕.

定理 3 的证明　类似文献[15]中的定理 2.3 可得.

参考文献：

[1] JENKINSON A F. The Frequency Distribution of the Annual Maximum (or Minimum) Values of Meteorological Elements [J]. Quarterly Journal of the Royal Meteorological Society, 1955, 81(348): 158-171.

[2] RESNICK S. Extreme Values, Regular Variation, and Point Processes [M]. Colorado: Springer, 1987.

[3] HILL B M. A Simple General Approach to Inference about the Tail of a Distribution [J]. The Annals of Statistics, 1975, 3(5): 1163-1174.

[4] DE HAAN L, FERREIRA A. Extreme Value Theory [M]. New York: Springer, 2006.

[5] 王嫣然，彭作祥. 基于分块的重尾指数估计量的推广 [J]. 西南师范大学学报(自然科学版), 2019, 44(1): 25-28.

[6] 胡爽，彭作祥. 基于分块思想的 Pickands 型估计量 [J]. 西南大学学报(自然科学版), 2019, 41(5): 53-58.

[7] PAULAUSKAS V. A New Estimator for a Tail Index [J]. Acta Applicandae Mathematica, 2003, 79(1-2): 55-67.

[8] PAULAUSKAS V, VAIČIULIS M. On an Improvement of Hill and some other Estimators [J]. Lithuanian Mathematical Journal, 2013, 53(3): 336-355.

[9] ALVES M I F. A Location Invariant Hill-Type Estimator [J]. Extremes, 2001, 4(3): 199-217.

[10] LING C X, PENG Z X, NADARAJAH S. A Location Invariant Moment-Type Estimator. I [J]. Theory of Probability and Mathematical Statistics, 2008, 76: 23-31.

[11] FRAGA ALVES M I, GOMES M I, HAAN L, et al. Mixed Moment Estimator and Location Invariant Alternatives [J]. Extremes, 2009, 12(2): 149-185.

[12] LI J N, PENG Z X, NADARAJAH S. Asymptotic Normality of Location Invariant Heavy Tail Index Estimator [J]. Extremes, 2010, 13(3): 269-290.

[13] DE HAAN L, STADTMÜLLER U. Generalized Regular Variation of Second Order [J]. Journal of the Australian Mathematical Society Series A Pure Mathematics and Statistics, 1996, 61(3): 381-395.

[14] RÉNYI A. On the Theory of Order Statistics [J]. Acta Mathematica Academiae Scientiarum Hungarica, 1953, 4(3-4): 191-231.

[15] LI J N, PENG Z X, NADARAJAH S. A Class of Unbiased Location Invariant Hill-Type Estimators for Heavy Tailed Distributions [J]. Electronic Journal of Statistics, 2008(2): 829-847.

责任编辑　张　枸