第 44 卷 第 1 期 西南师范大学学报(自然科学版) 2019 年 1 月
Vol. 44 No. 1 Journal of Southwest China Normal University (Natural Science Edition) Jan. 2019

DOI:10.13718/j.cnki.xsxb.2019.01.005

基于分块的重尾指数估计量的推广[①]

王嫣然， 彭作祥

西南大学 数学与统计学院，重庆 400715

摘要：采用分块的方法提出了一种新的重尾估计量，并讨论了这种新的估计量在正则变化条件下的相合性和渐近正态性.

关 键 词：重尾指数；分块方法；相合性；渐近正态性

中图分类号：O212 **文献标志码：A** **文章编号：**1000-5471(2019)01-0025-04

设(X_n)为独立同分布随机变量序列(简记为 i.i.d. 序列)，共同的分布函数为$F(x)$. 对任意$n \geqslant 1$，X_1，$X_2, \cdots, X_n$的顺序统计量为$X_{1,n} \geqslant X_{2,n} \geqslant \cdots \geqslant X_{n,n}$. 设$F \in D(G_\gamma)$，即存在规范常数$a_n > 0$，$b_n \in \mathbb{R}$，

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

其中极值分布$G_\gamma(x) = \mathrm{e}^{\{-(1+\gamma x)^{-\frac{1}{\gamma}}\}}$，$\gamma \in \mathbb{R}$且$1 + \gamma x > 0$. 当$\gamma = 0$时，$G_\gamma(x) = \mathrm{e}^{\{\mathrm{e}^{-x}\}}$. 我们需要研究的问题是$F$未知时，如何从有限的样本估计重尾指数$\gamma > 0$.

文献[1]提出如下的重尾指数估计量：

$$\hat{\gamma}_{n,H} = \frac{1}{k}\sum_{i=1}^{k} \log X_{n-i+1,n} - \log X_{n-k,n}$$

其中$k = k_n$满足当$n \to \infty$时，$k_n \to \infty$，$\frac{k_n}{n} \to 0$. 文献[2]讨论了 Hill 估计量的渐近性质. 文献[3]和文献[4]提出了分块的方法构造的估计量. 首先，将样本$x_1, \cdots, x_n$分成$k = k_{(n)}$块$V_1, \cdots, V_k$，每块里面包含$m = \left[\frac{n}{k}\right]$个样本观察值，其中$[x]$表示$x \geqslant 0$的整数部分. 更具体一点，$V_i = \{X_{(i-1)m+j}: 0 \leqslant j \leqslant m-1\}$ $(1 \leqslant i \leqslant k-1)$. 令顺序统计量$X_{j,m}^{(i)}$表示$V_i$里$m$个观测值中第$j$个最大值. 本文提出的估计量较文献[5]提出的估计量更具有一般性，文献[5]提出的估计量必须要知道$X_{1,m}^i, X_{2,m}^i, X_{3,m}^i$，而本文只需要知道任意相邻的 3 个顺序统计量即可.

本文基于上述分块的方法提出一种新的重尾极值指数估计量，

$$\hat{\gamma}_{(k,s)} = \frac{1}{2k}\sum_{i=1}^{k} \frac{(X_{s-1,m}^i)^{s-1} X_{s,m}^i}{(X_{s+1,m}^i)^s} \tag{2}$$

其中s是正整数且$s \geqslant 2$. 对序列$k = k(n)$，需要下列条件

$$k_n \to \infty, \ \frac{k_n}{n} \to 0 \tag{3}$$

或

$$\frac{k_n}{n} \to 0, \ \frac{k_n}{(\log \log n)^\delta} \to \infty \tag{4}$$

① 收稿日期：2018-03-20
作者简介：王嫣然(1994-)，女，硕士研究生，主要从事极值理论研究.
通信作者：彭作祥，博士，教授.

我们知道，当 $\gamma \geqslant 0$ 时，$U \in RV_{\gamma}$，即

$$\lim_{t\to\infty}\frac{U(tx)}{U(t)} = x^{\gamma},\ x > 0 \tag{5}$$

为讨论 $\hat{\gamma}_{(k,s)}$ 的渐近分布，本文假设存在函数 $A(t) \in RV_{\rho}$，$\rho < 0$，使得

$$\lim_{t\to\infty}\frac{\frac{U(tx)}{U(t)} - x^{\gamma}}{A(t)} = x^{\gamma}\frac{x^{\rho}-1}{\rho} \tag{6}$$

对所有 $x > 0$ 成立.

1 主要结论

定理 1 假设(3)和(5)式成立，那么当 $n \to \infty$ 时，$\hat{\gamma}(k,s) \xrightarrow{P} \gamma$.

定理 2 假设(4)和(5)式成立，则 $n \to \infty$ 时，$\hat{\gamma}(k,s) \xrightarrow{a.\ s.} \gamma$.

定理 3 假设(6)式成立，且 $k = k_n \to \infty$，$\frac{k}{n} \to 0$，及

$$\sqrt{k}A\left(\frac{k}{n}\right) \to \lambda \in (-\infty,\infty) \tag{7}$$

那么有

$$\sqrt{k}(\hat{\gamma}_{(k,s)} - \gamma) \xrightarrow{d} N\left(S_{\lambda,\rho},\frac{\gamma^2}{2}\right)$$

其中 $S_{\lambda,\rho} = \frac{2s-\rho-2}{2(s-1)!}\lambda\Gamma(s-\rho-1)$.

2 定理的证明

为了证明主要结论，需要以下引理.

引理 1 $T_m^{(1)}$ 和 $R_m^{(1)}$ 独立，$S_m^{(1)}$ 和 $R_m^{(1)}$ 也独立.

证 假设 $E_1^{(i)},\cdots,E_m^{(i)}$ 是一组随机变量，独立同分布于标准指数分布. $E_{1,m}^{(i)} \geqslant E_{2,m}^{(i)} \geqslant \cdots \geqslant E_{2,m}^{(i)}$ 是它的一组顺序统计量. 显然，$U(\mathrm{e}^{E_{j,m}^{(i)}})_{i=1}^{\infty} \stackrel{d}{=} (X_{j,m}^{i})_{i=1}^{\infty}$. 不失一般性，当 $j \geqslant 1$ 时，我们假设 $X_j = U(\mathrm{e}^{E_j})$. 注意到当 $i = 1,2,\cdots,k$ 时，$(X_{s-1}^{(i)},X_s^{(i)},X_{s+1}^{(i)}) = (U(\mathrm{e}^{E_{s-1,m}^{(i)}}),U(\mathrm{e}^{E_{s,m}^{(i)}}),U(\mathrm{e}^{E_{s+1,m}^{(i)}}))$ 是独立同分布的随机向量序列. 令

$$T_m^{(i)} = E_{s-1,m}^{(i)} - E_{s+1,m}^{(i)},\ S_m^{(i)} = E_{s,m}^{(i)} - E_{s+1,m}^{(i)},\ R_m^{(i)} = E_{s+1,m}^{(i)} \tag{8}$$

由文献[6]易得结论.

引理 2 $R_m^{(1)}$ 和 m 沿用前文的定义. 若条件(6)成立，则 $n \to \infty$ 时，

$$\frac{E(A(\mathrm{e}^{R_m^{(1)}}))}{A(m)} \to \frac{1}{S!}\Gamma(s+1-\rho)$$

其中 $\Gamma(x) = \int_0^{\infty} t^{x-1}\mathrm{e}^{-t}\mathrm{d}t$ 为伽马函数.

定理 1 的证明 由 $\hat{\gamma}_{(k,s)}$ 的定义有，

$$\hat{\gamma}_{(k,s)} \stackrel{d}{=} \frac{1}{2k}\sum_{i=1}^{k}\log\frac{(U(\mathrm{e}^{E_{s-1,m}^{(i)}}))^{s-1}U(\mathrm{e}^{E_{s,m}^{(i)}})}{(U(\mathrm{e}^{E_{s+1,m}^{(i)}}))^{s}} =$$

$$\frac{1}{2k}\sum_{i=1}^{k}\left\{(s-1)\log\frac{U(\mathrm{e}^{E_{s-1,m}^{(i)}})}{U(\mathrm{e}^{E_{s+1,m}^{(i)}})} + \log\frac{U(\mathrm{e}^{E_{s,m}^{(i)}})}{U(\mathrm{e}^{E_{s+1,m}^{(i)}})}\right\}$$

由(5)式和 Potter 界[7]，对 $\forall \varepsilon > 0$，$\exists t_0 > 0$，使得对所有 $x > 1$ 和 $t > t_0$，有

$$\log(1-\varepsilon) + (\gamma-\varepsilon)\log x \leqslant \log\frac{U(tx)}{U(t)} \leqslant \log(1+\varepsilon) + (\gamma+\varepsilon)\log x \tag{9}$$

由(3)式易证 $n\to\infty$ 时，$e^{R_m^{(i)}}\xrightarrow{P}\infty$. 用 t 替换 $e^{E_{s+1,m}^{(i)}}$，用 x 分别替换 $e^{E_{s-1,m}^{(i)}-E_{s+1,m}^{(i)}}$ 和 $e^{E_{s,m}^{(i)}-E_{s+1,m}^{(i)}}$. 对充分大的 n，考虑 $\hat{\gamma}_{(k,s)}$ 的上界

$$\hat{\gamma}_{(k,s)}\leqslant\frac{1}{2k}\sum_{i=1}^{k}\{(s-1)[\log(1+\varepsilon)+(\gamma+\varepsilon)\log(e^{T_m^{(i)}})]+\log(1+\varepsilon)+(\gamma+\varepsilon)\log(e^{S_m^{(i)}})\}=$$

$$\frac{s}{2}\log(1+\varepsilon)+\frac{1}{2k}\sum_{i=1}^{k}(\gamma+\varepsilon)[(s-1)(E_{s-1,m}^{(i)}-E_{s,m}^{(i)})+s(E_{s,m}^{(i)}-E_{s+1,m}^{(i)})]$$

注意到 $\{l(E_{l,m}^{(i)}-E_{l+1,m}^{(i)})\}(l\geqslant 1,\ i=1,2,\cdots k)$ 是独立同分布并服从于标准指数分布的[6]. 运用大数定律，$\hat{\gamma}_{(k,s)}$ 的上界依概率收敛到 $\frac{s}{2}\log(1+\varepsilon)+(\gamma+\varepsilon)$. 用类似的方法可得到 $\hat{\gamma}_{(k,s)}$ 的下界，由 ε 具有任意性，定理得证.

定理 2 的证明　运用条件(4)和柯尔莫哥洛夫强大数定律，可以得到

$$e^{R_m^{(i)}}\xrightarrow{a.\ s.}\infty$$

对所有 $i-1,2,\cdots,k$ 都成立. 用与定理 1 相似的证明方法，可得结论.

定理 3 的证明　令 $u_n=E\left(\frac{1}{2}\log\frac{(X_{s-1,m}^{(1)})^{s-1}X_{s,m}^{(1)}}{(X_{s+1,m}^{(1)})^s}\right)$，由正则变化，得

$$Var\left(\frac{1}{2}\log\frac{(X_{s-1,m}^{(1)})^{s-1}X_{s,m}^{(1)}}{(X_{s+1,m}^{(1)})^s}\right)=\frac{1}{4}Var\left(\log\frac{[U(e^{E_{s-1,m}^{(1)}})]^{s-1}}{[U(e^{E_{s+1,m}^{(1)}})]^{s-1}}+\log\frac{U(e^{E_{s,m}^{(1)}})}{U(e^{E_{s+1,m}^{(1)}})}\right)=$$

$$\frac{1}{4}Var\left(\log\frac{[U(e^{T_m^{(1)}}\ e^{R_m^{(1)}})]^{s-1}}{[U(e^{R_m^{(1)}})]^{s-1}}+\log\frac{U(e^{S_m^{(1)}}\ e^{R_m^{(1)}})}{U(e^{R_m^{(1)}})}\right)=$$

$$\frac{1}{4}Var((s-1)\gamma\, T_m^{(1)}(1+O_p(1))+\gamma\, S_m^{(1)}(1+O_p(1)))$$

注意到 $T_m^{(1)}\overset{d}{=}E_{2,s}^{(1)}$，$S_m^{(1)}\overset{d}{=}E_{1,s}^{(1)}$[5]，则有

$$Var\left(\frac{1}{2}\log\frac{(X_{s-1,m}^{(1)})^{s-1}X_{s,m}^{(1)}}{(X_{s+1,m}^{(1)})^s}\right)=\frac{\gamma^2}{4}Var((s-1)T_m^{(1)}+S_m^{(1)})(1+O_p(1))=$$

$$\frac{\gamma^2}{4}Var((s-1)(E_{s-1,m}^{(1)}-E_{s,m}^{(1)})+s(E_{s,m}^{(1)}-E_{s+1,m}^{(1)}))(1+O_p(1))=$$

$$\frac{\gamma^2}{4}[Var((s-1)(E_{s-1,m}^{(1)}-E_{s,m}^{(1)}))+Var(s(E_{s,m}^{(1)}-E_{s+1,m}^{(1)}))](1+O_p(1))\to$$

$$\frac{\gamma^2}{2}$$

由中心极限定理可得

$$\sqrt{k}(\hat{\gamma}_{(k,s)}-u_n)\xrightarrow{d}N\left(0,\frac{\gamma^2}{2}\right)\tag{10}$$

注意(6)式等价于

$$\lim_{t\to\infty}\frac{\log U(tx)-\log U(t)-\gamma\log x}{A(t)}=\frac{x^\rho-1}{\rho}\tag{11}$$

由引理 1 和(11)式，并注意 $e^{T_m^{(1)}}=e^{E_{s-1,m}^{(1)}-E_{s+1,m}^{(1)}}\overset{d}{=}e^{E_{2,s}^{(1)}}$，$e^{S_m^{(1)}}=e^{E_{s,m}^{(1)}-E_{s+1,m}^{(1)}}\overset{d}{=}e^{E_{1,s}^{(1)}}$，则

$$u_n=\gamma+\frac{1}{2}EA(e^{R_m^{(1)}})E\frac{(s-1)e^{\rho T_m^{(1)}}+e^{\rho S_m^{(1)}}-s}{\rho}(1+O_p(1))=$$

$$\gamma+\frac{1}{2}EA(e^{R_m^{(1)}})\frac{\frac{s(s-1)^2}{(s-\rho)(s-\rho-1)}+\frac{s}{(s-\rho)}-s}{\rho}(1+O_p(1))=$$

$$\gamma+\frac{s(-\rho+2s-2)}{2(s-\rho)(s-\rho-1)}EA(e^{R_m^{(1)}}(1+O_p(1)))$$

运用引理 2 和(7)式，有

$$\sqrt{(k)}(u_n - \gamma) \to \frac{\lambda(-\rho + 2s - 2)}{2(s-1)!}\Gamma(s - \rho - 1) \tag{12}$$

由(10)和(12)式,定理得证.

参考文献:

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

[2] EINMAHL J H J, DEKKERS A L M, DE HAAN L. A Moment Estimator for the Index of an Extreme-value Distribution [J]. Annals of Statistics, 1989, 17(4): 1833-1855.

[3] DAVYDOV Y, PAULAUSKAS V, RACKAUSKAS. More on P-Stable Convex Sets in Banach Spaces [J]. Journal of Theoretical Probability, 2000, 13(1): 39-64.

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

[5] LI J N. Tail Index Estimator Based on Block Order Statistics [J]. 西南大学学报(自然科学版), 2008, 30(7): 6-9.

[6] REISS R D. Approximate Distributions of Order Statistics [M]. New York: Springer-Verlag, 1989.

[7] RESNICK S I. Extreme Values, Regular Variation and Point Processes [M]. New York: Springer-Verlag, 1987.

A Kind of Generalized Tail Index Estimator Based on Block Order Statistics

WANG Yan-ran, PENG Zuo-xiang

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

Abstract: In this paper, a new kind of heavy tail estimator has been proposed in block method. The consistency and asymptotic normality of the new estimator have been considered under some regular varying conditions.

Key words: heavy tail index; block method; consistency; asymptotic normality

责任编辑 张 枸