DOI:10.13718/j.cnki.xsxb.2022.07.010

一类 Hill 型估计量分布的次渐近逼近①

陈艾谷， 彭作祥

西南大学 数学与统计学院，重庆 400715

摘要：在二阶正规变化条件下，定义了一个适当的中间分布序列，相较于正态分布，可以更好逼近一类统计量和由其导出的 Hill 型估计量的分布.

关 键 词：次渐近逼近；二阶正规变化条件；Hill 型估计量；重尾指数

中图分类号：O211.4 **文献标志码**：A **文章编号**：1000-5471(2022)07-0065-05

On Penultimate Approximation for Distribution of a Class of Hill-Type Estimator

CHEN Aigu， PENG Zuoxiang

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

Abstract：In this paper，an appropriate middle distribution sequence has been defined，which can better approximate the distribution of a class of moment statistics and the induced Hill-type estimator than the normal distribution under the second-order regular variation.

Key words：penultimate approximation；second-order regular variation；Hill-type estimator；heavy-tail index

设$\{X_n, n \geqslant 1\}$为独立同分布随机变量序列，$\{X_{i,n}, 1 \leqslant i \leqslant n\}$为其升序统计量，公共分布函数 F 满足 $1-F \in RV_{-\frac{1}{\gamma}}$，$\gamma > 0$ 被称为重尾指数[1]. 当 F 未知时，文献[2]提出如下 Hill 型估计量来估计 γ：

$$\gamma_n^{(\alpha)}(k) = \left(\frac{M_n^{(\alpha)}(k)}{\Gamma(\alpha+1)}\right)^{\frac{1}{\alpha}}, \alpha > 0 \tag{1}$$

其中矩统计量

$$M_n^{(\alpha)}(k) = \frac{1}{k}\sum_{i=1}^{k}(\log X_{n-i+1,n} - \log X_{n-k,n})^{\alpha}, 1 \leqslant k < n$$

$\Gamma(\cdot)$ 表示伽玛函数. 注意 $M_n^{(1)}(k) = \gamma_n^{(1)}(k)$ 为文献[3]提出的 Hill 估计量. 文献[4-5]研究了其相合性. 文献[6-7]讨论了其渐近正态性. 文献[8]给出 Hill 估计量分布的渐近展开. 文献[9]在下列二阶正规变化条件下，研究了 Hill 估计量分布的次渐近逼近：存在辅助函数 $A(t) \to 0$ $(t \to \infty)$ 在无穷远处符号恒定，使得对 $x > 0$

---

① 收稿日期：2021-10-09

基金项目：国家自然科学基金项目(11701469).

作者简介：陈艾谷，硕士研究生，主要从事极值统计分析研究.

通信作者：彭作祥，教授.

$$\lim_{t\to\infty}\frac{\log U(tx)-\log U(t)-\gamma\log x}{A(t)}=\frac{x^{\rho}-1}{\rho} \tag{2}$$

其中 $\rho\leqslant 0$，$A\in RV_{\rho}$，$U=\left(\frac{1}{1-F}\right)^{\leftarrow}$. 近期有关尾指数估计量的研究可参见文献[10-13].

设$\{E_i, i\geqslant 1\}$为独立同标准指数分布的随机变量序列，$\mu:=E(E_1^{\alpha})$，$\sigma^2:=\mathrm{Var}(E_1^{\alpha})$，$G_{k,\alpha}(x):=P\left\{\sum_{i=1}^{k}\frac{E_i^{\alpha}-\mu}{\sigma\sqrt{k}}\leqslant x\right\}$. 本文将基于 $G_{k,\alpha}$ 给出 $M_n^{(\alpha)}(k)$ 和 $\gamma_n^{(\alpha)}(k)$ 分布的次渐近逼近.

本文主要结论如下：

定理 1 令(2)式成立，$\sqrt{k}A\left(\frac{n}{k}\right)\to 0$，$\frac{\log n}{k}\to 0$ 以及 $k^{\frac{3}{2}}A\left(\frac{n}{k}\right)\to\infty$. 对 x 一致地有

$$\lim_{n\to\infty}\frac{1}{\sqrt{k}A\left(\frac{n}{k}\right)}\left(P\left\{\sqrt{k}\left(\frac{M_n^{(\alpha)}(k)}{\gamma^{\alpha}\mu}-1\right)\leqslant x\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)\right)=-\frac{((1-\rho)^{-\alpha}-1)\mu}{\gamma\rho\sigma}\phi\left(\frac{\mu}{\sigma}x\right) \tag{3}$$

定理 2 令(2)式成立，$\sqrt{k}A\left(\frac{n}{k}\right)\to 0$ 以及存在 $\eta\in(0,1)$ 使得 $k^{\eta+\frac{1}{2}}A\left(\frac{n}{k}\right)\to\lambda\in[0,\infty]$. 如果 $\lambda=\infty$，(3)式对 x 一致成立，否则对 x 一致地有

$$\lim_{n\to\infty}k^{\eta}\left(P\left\{\sqrt{k}\left(\frac{M_n^{(\alpha)}(k)}{\gamma^{\alpha}\mu}-1\right)\leqslant x\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)\right)=-\frac{((1-\rho)^{-\alpha}-1)\mu\lambda}{\gamma\rho\sigma}\phi\left(\frac{\mu}{\sigma}x\right) \tag{4}$$

定理 3 令(2)式成立，$\sqrt{k}A\left(\frac{n}{k}\right)\to 0$ 以及存在 $\eta\in\left(0,\frac{1}{2}\right]$ 使得 $k^{\eta+\frac{1}{2}}A\left(\frac{n}{k}\right)\to\lambda\in[0,\infty)$. 对 x 一致地有

$$\lim_{n\to\infty}k^{\eta}\left(P\left\{\alpha\sqrt{k}\left(\frac{\gamma_n^{(\alpha)}(k)}{\gamma}-1\right)\leqslant x\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)\right)=c_1\phi\left(\frac{\mu}{\sigma}x\right) \tag{5}$$

其中

$$c_1=\begin{cases}-\frac{((1-\rho)^{-\alpha}-1)\mu\lambda}{\gamma\rho\sigma}+\frac{(\alpha-1)\mu x^2}{2\sigma\alpha}, & \eta=\frac{1}{2}\\ -\frac{((1-\rho)^{-\alpha}-1)\mu\lambda}{\gamma\rho\sigma}. & \text{其他}\end{cases}$$

以及

$$\lim_{n\to\infty}k^{\eta}\left(P\left\{\alpha\sqrt{k}\left(\frac{\gamma_n^{(\alpha)}(k)}{\gamma}-1\right)\leqslant x\right\}-\Phi\left(\frac{\mu}{\sigma}x\right)\right)=c_2\phi\left(\frac{\mu}{\sigma}x\right) \tag{6}$$

其中

$$c_2=\begin{cases}-\frac{((1-\rho)^{-\alpha}-1)\mu\lambda}{\gamma\rho\sigma}+\frac{(\alpha-1)\mu x^2}{2\sigma\alpha}+\frac{\mu_3}{6\sigma^3}\left(1-\left(\frac{\mu}{\sigma}x\right)^2\right), & \eta=\frac{1}{2}\\ -\frac{((1-\rho)^{-\alpha}-1)\mu\lambda}{\gamma\rho\sigma}. & \text{其他}\end{cases}$$

为了证明本文结论，先给出如下 3 个引理：

引理 1 令(2)式成立. 对 $x>0$ 且 $x\neq 1$，有

$$\lim_{t\to\infty}\frac{\left(\log\frac{U(tx)}{U(t)}\right)^{\alpha}-(\gamma\log x)^{\alpha}}{\log\frac{U(tx)}{U(t)}-\gamma\log x}=\alpha(\gamma\log x)^{\alpha-1} \tag{7}$$

证 由(2)式易得

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

再结合(2)式，引理 1 得证.

引理 2　令(2)式成立. 对任意 $\varepsilon>0$，存在一个函数 $A_0\sim A$ 和充分大的 t_0 使得对 $t\geqslant t_0$，$x\geqslant 1$，一致地有

$$\left|\frac{\left(\log\frac{U(tx)}{U(t)}\right)^{\alpha}-(\gamma\log x)^{\alpha}}{A_0(t)}-\alpha(\gamma\log x)^{\alpha-1}\frac{x^{\rho}-1}{\rho}\right|\leqslant$$
$$\varepsilon x^{\rho+\varepsilon}(\log x)^{\alpha-1}+\varepsilon\,|\alpha-1|\,(1-x^{\rho})[(\log x)^{\alpha-1}+(\log x)^{\alpha-2}] \tag{8}$$

证　由(2)式易知 $\log U(x)-\gamma\log x\in ERV_{\rho}$，结合文献[14]中定理 B. 2. 18 可得，对任意 $\varepsilon>0$，存在一个函数 $A_0\sim A$ 以及充分大的 t_0 使得对 $t\geqslant t_0$，$x\geqslant 1$，一致地有

$$\left|\frac{\log\frac{U(tx)}{U(t)}-\gamma\log x}{A_0(t)}-\frac{x^{\rho}-1}{\rho}\right|\leqslant\varepsilon x^{\rho+\varepsilon}$$

结合三角不等式可得

$$\left|\frac{\left(\log\frac{U(tx)}{U(t)}\right)^{\alpha}-(\gamma\log x)^{\alpha}}{A_0(t)}-\alpha(\gamma\log x)^{\alpha-1}\frac{x^{\rho}-1}{\rho}\right|\leqslant$$
$$\varepsilon x^{\rho+\varepsilon}\left|\frac{\left(\log\frac{U(tx)}{U(t)}\right)^{\alpha}-(\gamma\log x)^{\alpha}}{\log\frac{U(tx)}{U(t)}-\gamma\log x}\right|+\frac{x^{\rho}-1}{\rho}\left|\frac{\left(\log\frac{U(tx)}{U(t)}\right)^{\alpha}-(\gamma\log x)^{\alpha}}{\log\frac{U(tx)}{U(t)}-\gamma\log x}-\alpha(\gamma\log x)^{\alpha-1}\right|$$

由拉格朗日中值定理和 Potter 界可得

$$\left|\frac{\left(\log\frac{U(tx)}{U(t)}\right)^{\alpha}-(\gamma\log x)^{\alpha}}{\log\frac{U(tx)}{U(t)}-\gamma\log x}-\alpha(\gamma\log x)^{\alpha-1}\right|\leqslant\alpha\gamma^{\alpha-2}\varepsilon\,|\alpha-1|\,((\log x)^{\alpha-1}+(\log x)^{\alpha-2})(1+o(1))$$

结合引理 1 得证.

引理 3　对任意数列 $f_k\to 0$，对 x 一致地有

$$\lim_{k\to\infty}\frac{G_{k,\alpha}(x+f_k)-G_{k,\alpha}(x)}{f_k}=\phi(x) \tag{9}$$

证　易知 $E(E_1^{\alpha})^3<\infty$，结合文献[15]中定理 2. 4. 3 可得

$$G_{k,\alpha}(x)=\Phi(x)+\frac{\mu_3}{6\sigma^3\sqrt{k}}(1-x^2)\phi(x)(1+o(1)) \tag{10}$$

其中 Φ,ϕ 分别为标准正态分布函数和概率密度函数，$\mu_3:=E(E_1^{\alpha}-\mu)^3<\infty$. 代入(9)式左边，引理 3 得证.

定理 1 的证明　令 $\{Y_i,\ i\geqslant 1\}$ 为独立同标准帕累托分布的随机变量序列，$\{Y_{i,n},\ 1\leqslant i\leqslant n\}$ 为其升序统计量. 由全概率公式和文献[9]中引理 2 可得

$$P\left\{\sqrt{k}\left(\frac{M_n^{(\alpha)}(k)}{\gamma^{\alpha}\mu}-1\right)\leqslant x\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)=$$
$$\left(P\left\{\sqrt{k}\left(\frac{M_n^{(\alpha)}(k)}{\gamma^{\alpha}\mu}-1\right)\leqslant x\,\middle|\,\left|\frac{kY_{n-k,n}}{n}-1\right|\leqslant t_n\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)\right)(1+o(1))+o\left(\sqrt{k}A_0\left(\frac{n}{k}\right)\right)$$

其中 $t_n\downarrow 0$. 注意到 $\{X_i\}_{i=1}^{n}\overset{d}{=}\{U(Y_i)\}_{i=1}^{n}$. n 充分大时可将引理 2 中的 tx,t 分别取为 $Y_{n-i+1,n},Y_{n-k,n},1\leqslant i\leqslant k$. 如果 n 充分大进一步使得 $(1-t_n)^{\rho-\varepsilon}\leqslant 1+\varepsilon$，$(1+t_n)^{\rho-\varepsilon}\geqslant 1-\varepsilon$，$(1-t_n)^{\varepsilon}\geqslant 1-\varepsilon$ 和 $(1+t_n)^{\varepsilon}\leqslant 1+\varepsilon$ 成立，由 Potter 界可得 $(1-\varepsilon)^2\leqslant\dfrac{A_0(Y_{n-k,n})}{A_0\left(\frac{n}{k}\right)}\leqslant(1+\varepsilon)^2$. 那么

$$\sqrt{k}\left(\frac{M_n^{(\alpha)}(k)}{\gamma^{\alpha}\mu}-1\right)-\frac{1}{\sqrt{k}\mu}\sum_{i=1}^{k}\left(\left(\log\frac{Y_{n-i+1,n}}{Y_{n-k,n}}\right)^{\alpha}-\mu\right)\geqslant$$

$$\frac{A_0\left(\frac{n}{k}\right)}{\gamma^\alpha\mu\sqrt{k}}\sum_{i=1}^{k}\left((1-\varepsilon)^2\alpha\left(\gamma\log\frac{Y_{n-i+1,n}}{Y_{n-k,n}}\right)^{\alpha-1}\frac{\left(\frac{Y_{n-i+1,n}}{Y_{n-k,n}}\right)^{\rho}-1}{\rho}-\right.$$
$$(1+\varepsilon)^2\varepsilon\left(\frac{Y_{n-i+1,n}}{Y_{n-k,n}}\right)^{\rho+\varepsilon}\left(\log\frac{Y_{n-i+1,n}}{Y_{n-k,n}}\right)^{\alpha-1}-$$
$$\left.(1+\varepsilon)^2\varepsilon\mid\alpha-1\mid\left(1-\left(\frac{Y_{n-i+1,n}}{Y_{n-k,n}}\right)^{\rho}\right)\left(\left(\log\frac{Y_{n-i+1,n}}{Y_{n-k,n}}\right)^{\alpha-1}+\left(\log\frac{Y_{n-i+1,n}}{Y_{n-k,n}}\right)^{\alpha-2}\right)\right)$$

结合$\left\{\frac{Y_{n-i+1,n}}{Y_{n-k,n}}\right\}_{i=1}^{k}$和 $Y_{n-k,n}$ 的独立性，$\sum_{i=1}^{k}\frac{Y_{n-i+1,n}}{Y_{n-k,n}}\overset{d}{=}\sum_{i=1}^{k}\exp(E_i)$以及全概率公式，有

$$P\left\{\sqrt{k}\left(\frac{M_n^{(\alpha)}(k)}{\gamma^\alpha\mu}-1\right)\leqslant x\,\middle|\,\left|\frac{kY_{n-k,n}}{n}-1\right|\leqslant t_n\right\}\leqslant P\left\{H_k\leqslant x+\frac{\varepsilon\sqrt{k}A_0\left(\frac{n}{k}\right)}{\gamma^\alpha\mu}\right\}+P\{Q_k\leqslant-\varepsilon\sqrt{k}\}\tag{11}$$

其中

$$H_k=\frac{1}{\sqrt{k}\mu}\sum_{i=1}^{k}(E_i^\alpha-\mu)+\frac{\sqrt{k}A_0\left(\frac{n}{k}\right)}{\gamma^\alpha\mu}((1-\varepsilon)^2\alpha\gamma^{\alpha-1}\rho^{-1}((1-\rho)^{-\alpha}-1)\Gamma(\alpha)-$$
$$(1+\varepsilon)^2\varepsilon(1-\rho-\varepsilon)^{-\alpha}\Gamma(\alpha)-(1+\varepsilon)^2\varepsilon\mid\alpha-1\mid(1-(1-\rho)^{-\alpha})(\Gamma(\alpha)+\Gamma(\alpha-1)))=:$$
$$\frac{1}{\sqrt{k}\mu}\sum_{i=1}^{k}(E_i^\alpha-\mu)+\frac{\sqrt{k}A_0\left(\frac{n}{k}\right)}{\gamma^\alpha\mu}c_\varepsilon$$

以及

$$Q_k=\frac{1}{\sqrt{k}}\sum_{i=1}^{k}((1-\varepsilon)^2\alpha\gamma^{\alpha-1}\rho^{-1}(E_i^{\alpha-1}(\exp(\rho E_i)-1)-((1-\rho)^{-\alpha}-1)\Gamma(\alpha))-$$
$$(1+\varepsilon)^2\varepsilon\mid\alpha-1\mid((1-\exp(\rho E_i))(E_i^{\alpha-1}+E_i^{\alpha-2})-(1-(1-\rho)^{-\alpha})(\Gamma(\alpha)+\Gamma(\alpha-1)))-$$
$$(1+\varepsilon)^2\varepsilon(\exp((\rho+\varepsilon)E_i)E_i^{\alpha-1}-(1-\rho-\varepsilon)^{-\alpha}\Gamma(\alpha)))=:\frac{1}{\sqrt{k}}\sum_{i=1}^{k}(V_i-E(V_1))$$

其中 $\mathrm{Var}(V_1)<\infty$. 由引理 3 可得

$$\frac{1}{\sqrt{k}A_0\left(\frac{n}{k}\right)}\left(P\left\{H_k\leqslant x+\frac{\varepsilon\sqrt{k}A_0\left(\frac{n}{k}\right)}{\gamma^\alpha\mu}\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)\right)\to\frac{\varepsilon-c_\varepsilon}{\gamma^\alpha\sigma}\phi\left(\frac{\mu}{\sigma}x\right)\tag{12}$$

由切比雪夫不等式可得

$$P\{Q_k\leqslant-\varepsilon\sqrt{k}\}=o\left(\sqrt{k}A_0\left(\frac{n}{k}\right)\right)\tag{13}$$

那么

$$\limsup_{n\to\infty}\frac{1}{\sqrt{k}A_0\left(\frac{n}{k}\right)}\left(P\left\{\sqrt{k}\left(\frac{M_n^{(\alpha)}(k)}{\gamma^\alpha\mu}-1\right)\leqslant x\,\middle|\,\left|\frac{kY_{n-k,n}}{n}-1\right|\leqslant t_n\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)\right)\leqslant\frac{\varepsilon-c_\varepsilon}{\gamma^\alpha\sigma}\phi\left(\frac{\mu}{\sigma}x\right)$$

同理可得下极限的下界. 分别令 $\varepsilon\to 0$，定理 1 得证.

定理 2 的证明 由全概率公式，文献[9]中引理 4 可得

$$P\left\{\sqrt{k}\left(\frac{M_n^{(\alpha)}(k)}{\gamma^\alpha\mu}-1\right)\leqslant x\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)=$$
$$\left(P\left\{\sqrt{k}\left(\frac{M_n^{(\alpha)}(k)}{\gamma^\alpha\mu}-1\right)\leqslant x\,\middle|\,\left|\frac{kY_{n-k,n}}{n}-1\right|\leqslant t_n\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)\right)(1+o(1))+o(k^{-\eta})$$

由(11)－(13)式，定理 2 得证.

定理 3 的证明 令 $x_k=\sqrt{k}\left(\left(1+\frac{x}{\alpha\sqrt{k}}\right)^{\alpha}-1\right)$，有

$$k^{\eta}\left(P\left\{\alpha\sqrt{k}\left(\frac{\gamma_n^{(\alpha)}(k)}{\gamma}-1\right)\leqslant x\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)\right)=$$

$$k^{\eta}\left(P\left\{\sqrt{k}\left(\frac{M_n^{(\alpha)}(k)}{\gamma^{\alpha}\mu}-1\right)\leqslant x_k\right\}-G_{k,\alpha}\left(\frac{\mu}{\sigma}x_k\right)+\frac{\mu}{\sigma}(x_k-x)\frac{G_{k,\alpha}\left(\frac{\mu}{\sigma}x_k\right)-G_{k,\alpha}\left(\frac{\mu}{\sigma}x\right)}{\frac{\mu}{\sigma}(x_k-x)}\right)$$

由定理 2 和引理 3 知，(5) 式成立. 结合(5),(10) 式知(6) 式成立.

参考文献：

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

[2] GOMES M I，MARTINS M J. Generalizations of the Hill Estimator-Asymptotic Versus Finite Sample Behaviour [J]. Journal of Statistical Planning and Inference，2001，93(1-2)：161-180.

[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] MASON D M. Laws of Large Numbers for Sums of Extreme Values [J]. The Annals of Probability，1982，10(3)：754-764.

[5] DEHEUVELS P，HAEUSLER E，MASON D M. Almost Sure Convergence of the Hill Estimator [J]. Mathematical Proceedings of the Cambridge Philosophical Society，1988，104(2)：371-381.

[6] HAEUSLER E，TEUGELS J L. On Asymptotic Normality of Hill's Estimator for the Exponent of Regular Variation [J]. The Annals of Statistics，1985，13(2)：743-756.

[7] DE HAAN L，RESNICK S. On Asymptotic Normality of the Hill Estimator [J]. Communications in Statistics Stochastic Models，1998，14(4)：849-866.

[8] CHENG S H，PAN J Z. Asymptotic Expansions of Estimators for the Tail Index with Applications [J]. Scandinavian Journal of Statistics，1998，25(4)：717-728.

[9] CHENG S，DE HAAN L. Penultimate Approximation for Hill's Estimator [J]. Scandinavian Journal of Statistics，2001，28(3)：569-575.

[10] 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.

[11] 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.

[12] 张茜，彭作祥. 不同观测点重尾参数的估计 [J]. 西南师范大学学报(自然科学版)，2019，44(1)：29-33.

[13] XIONG L，PENG Z X. Heavy Tail Index Estimation Based on Block Order Statistics [J]. Journal of Statistical Computation and Simulation，2020，90(12)：2198-2208.

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

[15] LEHMANN E L. Elements of Large-Sample Theory [M]. New York：Springer-Verlag，1999.

责任编辑　张栒