Hill-估计量的二阶 Edgeworth 展式

王晓谦[1],程士宏[2]

(1.南京师范大学数学与计算机科学学院,210097,江苏,南京) (2.北京大学数学科学学院,100871,北京)

[摘要] 具有正规变化尾的重尾分布的尾指标的估计是一个重要问题.本文利用二阶广义正规变化函数理论给出了尾指标的 Hill-估计量的分布的三阶 Edgeworth 展式.

[关键词] 二阶广义正规变化函数,Hill-估计量,Edgeworth 展式

[中图分类号]O211, [文献标识码]A, [文章编号]1001-4616(2004)03-0001-08

0 引言与结果

无论是对于极值理论,还是对于金融和保险统计,分布函数的尾部性质都具有重要意义(如[1]、[2]中关于金融中的厚尾(heavy tail)分布问题的研究).如何对分布函数的尾数性质进行估计是一个重要的统计问题[3,4].其中具有正规变化尾的分布函数是大家比较感兴趣的话题[5,6].

所谓分布函数 F 具有正规变化尾,是指它具有如下的性质:

$$\lim_{t\to\infty}\frac{1-F(tx)}{1-F(t)}=x^{-\frac{1}{\gamma}},\forall x>0 \tag{1}$$

其中 $\gamma>0$ 为由 F 惟一确定的常数,叫 F 的尾指标.对 γ 的估计是一个重要课题.

设 $X_1,X_2,\cdots$ 是 i.i.d.的随机变量序列,有共同分布 F,而 F 满足(1).Hill,B.M[7] 提出了一个 γ 的估计量:

$$\hat{H}_n=\frac{1}{k_n}\sum_{i=1}^{k_n}\log X_{n-i+1,n}-\log X_{n-k_n,n}, \tag{2}$$

其中 $X_{1,n}\leqslant X_{2,n}\leqslant\cdots\leqslant X_{n,n}$ 是$X_1,X_2,\cdots,X_n$ 的次序统计量,$\{k_n\in\mathbf{N}\}$ 满足 $k_n\to\infty,k_n/n\to 0,(n\to\infty)$,$\mathbf{N}$ 代表自然数集.文献[8—14] 对 $\hat{H}_n$ 的弱相合性、强相合性、渐近正态性、大偏差等做了广泛而深入的研究.但是,为了提高利用 Hill- 估计量的效率,还必须研究其分布的高阶展开式.

定义 1 设 f 是定义在 $\mathbf{R}^+$ 上的正值函数.若对任给 $x>0$,

$$\lim_{t\to\infty}\frac{f(tx)}{f(t)}=x^{\rho},$$

其中 $\rho\in\mathbf{R}$ 是由 $f(t)$ 惟一确定的一个常数.则称 f 为正规变化函数,记做 $f\in RV(\rho)$,并称 ρ 为f 的正规变化指标.

定义 2 设 f 是定义在 $\mathbf{R}^+$ 上的实可测函数.若存在不变号的函数 a 和实函数b,使得对任给 $x>0$,

$$\lim_{t\to\infty}\frac{f(tx)-f(t)}{a(t)}=\int_1^x s^{\rho-1}\mathrm{d}s,$$

其中$\rho\in\mathbf{R}$是由$f(t)$惟一确定的一个常数.则称f是广义正规变化函数.一般地,我们取 a 为正函数,否则以 $-f$ 作为研究对象.这时我们记做 $f\in GRV(\rho)$.

定义 3 设f是定义在 $\mathbf{R}^+$ 上的局部有界实可测函数.若存在不变号的函数 $a_1(t)$ 和 $a_2(t)$,使得对任意 $x\in(0,\infty)$,当 $t\to\infty$ 时,有

$$f(tx)=f(t)+a_1(t)h_1(x)+a_2(t)h_2(x)+o(a_2(t)),$$

其中 $a_2(t)=o(a_1(t))(t\to\infty)$,而

收稿日期:2004-06-01.
基金项目:国家自然科学基金资助项目(10071003).
作者简介:王晓谦,1966-,南京师范大学数学与计算机科学学院副教授,主要从事概率极限定理及其在金融中的应用等方面的教学与研究,E-mail:wxqmath@263.net

$$h_1(x) = \int_1^x x_1^{\alpha-1}\mathrm{d}x_1, \quad h_2(x) = \int_1^x x_1^{\alpha-1}\mathrm{d}x_1\int_1^{x_1} x_2^{\beta-1}\mathrm{d}x_2,$$

$\alpha \in \mathbf{R}, \beta \leqslant 0$ 由 f 惟一确定,则称 f 是二阶广义正规变化函数,一般地,我们取 a_1 为正函数,否则以 $-f$ 作为研究对象.这时记做 $f \in GRV_2(\alpha_1, \alpha_2)$.(关于正规变化函数的讨论参见[15,16],高阶广义正规变化函数的讨论参见[17,18]).

记 $V(t) = [1/(1-F)]^{\leftarrow}(t) := \inf\{s: 1/(1-F(s)) \geqslant t\}$,则(1)等价于 $V \in RV(\gamma)$.

称 V 满足二阶正规变化条件,若存在不变号的函数 $a(t) \to 0, (t \to \infty)$, $A(x)$ 不恒等于一个常数,使得

$$\lim_{t\to\infty}\frac{1}{a(t)}\left\{\frac{V(tx)}{V(t)} - x^{\gamma}\right\} = x^{\gamma}A(x), \forall x > 0. \tag{3}$$

显然,(3)即 $t^{-\gamma}V(t) \in GRV(\alpha)$,而 $\alpha \leqslant 0$ 是由 F 惟一确定的一个参数.同时又等价于

$$\lim_{t\to\infty}\frac{\log V(tx) - \log V(t) - \gamma\log x}{a(t)} = A(x), \forall x > 0.$$

在此条件下,$\hat{H}_n$ 有中心极限定理[9,19].

若 $\lim\limits_{n\to\infty} k_n^{1/2}a(n/k_n) = \lambda, \lambda \in [0, \infty]$,那么当 $n \to \infty$ 时有

$$H_n = \sqrt{k_n}\,\frac{\hat{H}_n - \gamma}{\gamma}\begin{cases}\xrightarrow{d} N(0,1), & \text{若 } \lambda = 0,\\ \xrightarrow{d} N(\mu_0, 1), & \text{若 } 0 < \lambda < \infty,\\ \text{无极限分布}, & \text{若 } \lambda = \infty.\end{cases} \tag{4}$$

其中 $\mu_0 = \lambda/(\gamma(1-\alpha))$.

(4)中的收敛速度引起了大家的研究兴趣[20].在条件(3)之下,Cheng, S. & Pan, J.[21] 讨论了上述中心极限定理中 $\lambda = 0$ 的情况下 H_n 的分布的 Edgeworth 展式,得到如下结果:

若 $n \to \infty$ 时,$k_n a(n/k_n) \to \rho$,则当 $n \to \infty$ 时,在 $\mathbf{R}$ 上一致地有

$$P(H_n \leqslant x) = \Phi(x) + \frac{1}{\sqrt{k_n}}P(x)\varphi(x) + o(k_n^{-1/2}). \tag{5}$$

其中 $\Phi(x)$ 和 $\varphi(x)$ 为标准正态分布函数及其密度函数,$P(x)$ 为确定的多项式.

为了得到更好的收敛速度,Cheng, S. & de Haan. L.[22] 讨论了 Γ-分布逼近 Hill-估计量的分布的速度问题.Cuntz, A. & Häusler, E. 对上述结果做了推广[23].

设 $V(t)$ 满足

$$\frac{\log V(tx) - \log V(t) - \gamma\log x - a_1(t)A(x)}{a_2(t)} \to B(x), \forall x > 0, \tag{6}$$

其中 $B(x)$ 不是 $A(x)$ 的常数倍,$a_2(t) = o(a_1(t)), t \to \infty$.

[23]讨论了 H_n 的分布在条件(7)下的一阶 Edgeworth 展式,但是由于没有充分利用条件(6),所以并没有给出 H_n 的分布的二阶 Edgeworth 展式.

在本文中,我们在条件(6)下,讨论了 H_n 的分布的二阶 Edgeworth 展式,得到下面的定理.一方面,这是(5)的直接的二阶推广,余项的阶为 $o(k_n^{-1})$.

定理　在条件(6)之下,若对某 $\delta > 0$,当 $n \to \infty$ 时有 $k_n^{\frac{3}{2}+\delta}a_2\left(\frac{n}{k_n}\right) \to 0, k_n^{1+\delta}a_1\left(\frac{n}{k_n}\right) \to 0$,则

$$P\{H_n \leqslant x\} = \Phi(x) + \sum_{i=1}^{2}k_n^{-\frac{i}{2}}p_i(x)\varphi(x) - k_n^{\frac{1}{2}}\frac{\varphi(x)}{\gamma(1-\alpha_1)}a_1\left(\frac{n+1}{k_n}\right) + o(1/k_n)$$

在 $\mathbf{R}$ 上一致成立.其中 $p_i(x)$ 为确定的多项式.

1　引理及定理的证明

设 $U_1, U_2, \cdots, U_n$ 是服从(0,1)上均匀分布的 i.i.d. 的随机变量,$U_{1,n} \leqslant U_{2,n} \leqslant \cdots \leqslant U_{n,n}$ 为其次序统计量.记

$$a_n = 1 - \frac{k_n}{n+1},\ b_n = \sqrt{\frac{a_n(1-a_n)}{n+1}},\ (u)_n = a_n + b_n u,\ (\forall u \in \mathbf{R}),$$

并以 $\phi(x)$ 表示 $(U_{n-k_n,n} - a_n)/b_n$ 的概率密度函数.

对 $i = 1,2,\cdots,n$,取

$$Y_{n,i}(u) = \begin{cases} \log V\left(\frac{1}{1-(u)_n}\frac{1}{1-U_i}\right) - \log V\left(\frac{1}{1-(u)_n}\right), & 若 0 < (u)_n < 1, \\ 0, & 否则, \end{cases}$$

则显然对任给 $u \in \mathbf{R}$,$n \geqslant 1$,$Y_{n,1}(u),Y_{n,2}(u),\cdots,Y_{n,k_n}(u)$ 是 i.i.d. 的随机变量,记

$$H_n(u) = \frac{1}{\sqrt{k_n}\gamma}\left(\sum_{i=1}^{k_n} Y_{n,i}(u) - k_n\gamma\right).$$

Cheng, S. & Jiang, C.[24] 对二阶广义正规变化函数证明了如下基本不等式:

引理 1 若 $f \in GRV_2(\alpha_1,\alpha_2)$,则可适当选取 $a_1(t)$,$a_2(t)$,使对任给的 $\varepsilon,\delta > 0$,存在 $t_0 = t_0(\epsilon,\delta) > 0$,只要 $t,tx > t_0$,就有

$$\left|\frac{f(tx) - f(t) - a_1(t)h_{\alpha_1}(x)}{a_2(t)} - h_{\alpha_1,\alpha_2}(x)\right| \leqslant \varepsilon \max\{x^{\alpha_1+\alpha_2+\delta}, x^{\alpha_1+\alpha_2-\delta}\}.$$

引理 2 若(6)成立,则任取 $\epsilon > 0$,存在 $t_0 > 0$,$K > 0$,使得对任给 $x > 1$,只有 $t > t_0$,就有

$$|B(t,x)| \leqslant Kx^{\epsilon}.$$

其中

$$B(t,x) = \frac{\log V(tx) - \log V(t) - \gamma\log x - a_1(t)h_\alpha(x)}{a_2(t)}.$$

证明 (6)表明 $\log(t^{-\gamma}V(t)) \in GRV_2(\alpha,\beta)$,$\alpha,\beta \leqslant 0$,由引理 1 知结论成立.

引理 3

$$P\{H_n \leqslant x\} = \int_{-\infty}^{+\infty} P\{H_n(u) \leqslant x\}\phi_n(u)\mathrm{d}u. \tag{7}$$

证明 此为[21]之(2.8)式.

引理 4 对任给 $\varepsilon > 0$ 和 $r \in \{1,\cdots,n\}$,有

$$P\left\{\left|\frac{n^{1/2}}{\sigma}(U_{r,n} - \mu)\right| \geqslant \varepsilon\right\} \leqslant \exp\left(-\frac{\epsilon^2}{3\left(1+\frac{\epsilon}{\sigma\sqrt{n}}\right)}\right), \tag{8}$$

其中 $\mu = r/(n+1)$,$\sigma^2 = \mu(1-\mu)$.

证明 见文[25].

引理 5 设 $\{X_{n,i}, i = 1,\cdots,k_n; n = 1,2,\cdots\}$ 是行独立同分布连续型随机变量阵列,每行对应共同的特征函数为 $\gamma_n(t)$,$n = 1,2,\cdots$. $X_{n,i}$ 的数学期望为 0,方差为 σ_n,$\{E|X_{n,i}|^\nu, \nu = 1,2,\cdots,r+1\}$ 一致地介于两个与 n,r 无关的常数 $0 < c < C < \infty$ 之间,$r \leqslant 2$ 是确定的自然数. 存在 $T > 0$,当 $t > T$ 时,特征函数列 $\{\gamma_n(t), n = 1,2,\cdots\}$ 满足

$$\lim_{n\to\infty}|\gamma_n(t)| \leqslant \delta < 1.$$

记

$$S_{k_n} = \frac{\sum_{i=1}^{k_n} X_{n,i}}{|\sqrt{k_n}\sigma_n|},$$

则 S_{k_n} 的分布函数在 $n \to \infty$ 时一致地有如下 Edgeworth 展式:

$$F_{S_{k_n}}(x) = \Phi(x) + \sum_{i=1}^{r-2} k_n^{-\frac{i}{2}} p_{n,i}(x)\varphi(x) + o\left(k_n^{-\frac{r-2}{2}}\right).$$

其中 $\Phi(x)$ 为标准正态分布函数,$\varphi(x)$ 为标准正态密度函数,$p_{n,i}$,$i = 1,\cdots,r-2$ 即 i.i.d. 随机变量列的

规范化和的分布的 Edgeworth 展式[26] 中的 $p_i, i=1,\cdots,r-2$,但其中的系数中的各阶矩换成了 $EX^j_{n,1}, j=3,\cdots,r$.

证明　根据 Feller,W.[27] 第二卷之 541 页定理 3 知结论成立.

对任取 $u \in \mathbf{R}$,以下记

$$R_{n,i}(u) = Y_{n,i}(u) - \gamma\log\left(\frac{1}{1-U_i}\right) - a_1\left(\frac{1}{1-(u)_n}\right)h_{\alpha_1}\left(\frac{1}{1-U_i}\right), i=1,\cdots,k_n,$$

$$R_n(u) = \frac{1}{\sqrt{k_n}\gamma}\sum_{i=1}^{k_n}\{R_{n,i}(u) - E(R_{n,i}(u))\},$$

$$r_n(u) = \frac{1}{\sqrt{k_n}\gamma}\sum_{i=1}^{k_n}\left\{E(R_{n,i}(u)) + \frac{1}{1-\alpha_1}a_1\left(\frac{1}{1-(u)_n}\right)\right\},$$

$$S_n(u) = \frac{1}{\sqrt{k_n}}\sum_{i=1}^{k_n}\left\{\log\left(\frac{1}{1-U_i}\right)-1\right\} + \frac{1}{\sqrt{k_n}\gamma}\sum_{i=1}^{k_n}\left\{a_1\left(\frac{1}{1-(u)_n}\right)\left(h_{\alpha_1}\left(\frac{1}{1-U_i}\right) - \frac{1}{1-\alpha_1}\right)\right\},$$

则

$$\begin{aligned} H_n(u) &= \frac{1}{\sqrt{k_n}\gamma}\left(\sum_{i=1}^{k_n}Y_{n,i}(u) - k_n\gamma\right) \\ &= \frac{1}{\sqrt{k_n}\gamma}\sum_{i=1}^{k_n}\left\{R_{n,i}(u) + \gamma\log\left(\frac{1}{1-U_i}\right) + a_1\left(\frac{1}{1-(u)_n}\right)h_{\alpha_1}\left(\frac{1}{1-U_i}\right) - k_n\gamma\right\} \\ &= S_n(u) + R_n(u) + r_n(u). \end{aligned}$$

记

$$X = \left(\log\left(\frac{1}{1-U_1}\right)-1\right) + a_1\left(\frac{1}{1-(u)_n}\right)\left(h_{\alpha_1}\left(\frac{1}{1-U_1}\right) - \frac{1}{1-\alpha_1}\right).$$

引理 6　取 $T_n = k_n^{\delta/2}, \delta \in (0,1/2), T_n \to \infty$ 时,对 $u \in \{u: |u| \leqslant T_n\}$,在 $\mathbf{R}$ 一致地有

$$P(S_n(u) \leqslant x) = \Phi(x) + \frac{1}{\sqrt{k_n}}p_1(x)\varphi(x) + \frac{1}{k_n}p_2(x)\varphi(x) + o(1/k_n).$$

证明　对于 $S_n(u)$,考察

$$X = \left(\log\left(\frac{1}{1-U_1}\right)-1\right) + a_1\left(\frac{1}{1-(u)_n}\right)\left(h_{\alpha_1}\left(\frac{1}{1-U_1}\right) - \frac{1}{1-\alpha_1}\right).$$

记 $X_1 = \log\left(\frac{1}{1-U_1}\right), X_2 = h_\alpha\left(\frac{1}{1-U_1}\right)$,则

$$EX_1^k = \int_0^1\left(\log\frac{1}{1-u}\right)^k\mathrm{d}u = \int_1^\infty\frac{\log^k t}{t^2}\mathrm{d}t = k!, k=1,2,\cdots;$$

对 $X_2, EX_2 = \frac{1}{1-\alpha}$;$\alpha = 0$ 时,$X_1 = X_2$;$\alpha < 0$ 时,对 $k=1,2,\cdots$,

$$E|X_2|^k = \int_0^1\left|\frac{(1/(1-u))^\alpha - 1}{\alpha}\right|^k\mathrm{d}u = \left(\frac{1}{|\alpha|}\right)^k\int_1^\infty\frac{(1-t^\alpha)^k}{t^2}\mathrm{d}u < \infty.$$

所以,存在常数 $0 < c < C < \infty$,使 $E|X|^k, k=1,2,3,4,5$ 介于它们之间.记 $c_n = a_1\left(\frac{1}{1-(u)_n}\right)$,则在 $\{u: |u| \leqslant T_n\}$ 上 $c_n \to 0$.同时,我们注意到,X 的方差可以记做 $\sigma_n^2 = 1 + c_nO(1)$,事实上我们有

$$EX^j = E(X_1 - 1)^j + c_nO(1), j=1,2,\cdots.$$

当 $\alpha = 0$ 时,在 $\{u: |u| \leqslant T_n\}$ 上,

$$\begin{aligned} |\gamma_n(t)| &= |E(\exp\{tX\})| = \left|\int_0^1 \mathrm{e}^{\mathrm{i}(1+c_n)t(\log(1/(1-u))-1)}\mathrm{d}u\right| \\ &= \left|\int_0^1 \mathrm{e}^{\mathrm{i}(1+c_n)t\log(1/(1-u))}\mathrm{d}u\right| = \left|\int_1^\infty\frac{1}{x^2}\mathrm{e}^{\mathrm{i}t(1+c_n)\log x}\mathrm{d}x\right| = \left|\int_1^\infty x^{-2+\mathrm{i}t(1+c_n)}\mathrm{d}x\right| \end{aligned}$$

$$= \left| \left.\frac{x^{-1+it(1+c_n)}}{-1+it(1+c_n)}\right|_1^{\infty} \right| = \left| \frac{1}{1-it(1+c_n)} \right| \to 0,(t\to\infty);$$

当 $\alpha<0$ 时,在 $\{u:|u|\leqslant T_n\}$ 上有

$$|\gamma_n(t)| = |E(\exp\{tX\})| = \left| \int_0^1 \exp\left\{it(\log(1/(1-u))-1)+itc_n\left(h_\alpha\left(\frac{1}{1-u}\right)-\frac{1}{1-\alpha}\right)\right\}du \right|$$

$$= \left| \int_0^1 \exp\left\{it\log(1/(1-u))+itc_nh_\alpha\left(\frac{1}{1-u}\right)\right\}du \right| = \left| \int_1^{\infty} \frac{1}{x^2}\exp\{it\log x+itc_nh_\alpha(x)\}dx \right|$$

$$= \left| \int_1^{\infty} \frac{1}{x^2}\exp\{it\log x+itc_n\alpha^{-1}x^\alpha\}dx \right| = \left| \int_1^{\infty} \frac{x^{it}\exp\{tc_n\alpha^{-1}x^\alpha\}}{x^2}dx \right|$$

$$= \left| \int_0^1 \frac{y^{it/\alpha}\exp\{tc_n\alpha^{-1}y\}}{y^{2/\alpha}}\frac{1}{\alpha}y^{\frac{1}{\alpha}-1}dy \right| = \left| \frac{1}{\alpha}\int_0^1 y^{\frac{it}{\alpha}-\frac{1}{\alpha}-1}\exp\{tc_n\alpha^{-1}y\}dy \right|$$

$$= \left| \frac{1}{\alpha}\int_0^1 \frac{\exp\{tc_n\alpha^{-1}y\}}{\frac{it}{\alpha}-\frac{1}{\alpha}}dy^{\frac{it}{\alpha}-\frac{1}{\alpha}} \right|$$

$$= \left| \left.\frac{1}{it-1}y^{\frac{it}{\alpha}-\frac{1}{\alpha}}\exp\{tc_n\alpha^{-1}y\}\right|_0^1 - \frac{itc_n\alpha^{-1}}{it-1}\int_0^1 y^{\frac{it}{\alpha}-\frac{1}{\alpha}}\exp\{tc_n\alpha^{-1}y\}dy \right| =:|A_n(t)+B_n(t)|,$$

则

$$|A_n(t)| \leqslant \left| \frac{1}{it-1}\exp\{tc_n\alpha^{-1}\} \right| \to 0,(t\to\infty);$$

而对任意 $T>0$ 时,只要 n 充分大,就有

$$|B_n(t)| \leqslant \left| \frac{itc_n\alpha^{-1}}{it-1} \right| \left| \int_0^1 y^{-\frac{1}{\alpha}}dy \right| = c_n \left| \frac{it\alpha^{-1}}{it-1} \right| < \delta < 1.$$

从而存在 $0<\delta<1$,当 $t>T$ 时,只要 n 充分大,就有

$$|\gamma_n(t)| < \delta < 1.$$

所以 $\frac{1}{\sqrt{\sigma_n}}S_n(u)$ 满足引理 5 的条件,从而记 $x_n=\frac{x}{\sqrt{\sigma_n}}$,利用引理 5,对 $u\in\{u:|u|\leqslant T_n\}$,注意到 $k_nc_n\to 0,(n\to\infty)$,在 R 一致地有

$$P(S_n(u)\leqslant x) = \Phi(x_n)+\frac{1}{\sqrt{k_n}}p_{n,1}(x_n)\varphi(x_n)+\frac{1}{k_n}p_{n,2}(x_n)\varphi(x_n)+o(1/k_n)$$

$$= \Phi(x_n)+\frac{1}{\sqrt{k_n}}p_{n,1}(x)\varphi(x)+\frac{1}{k_n}p_{n,2}(x)\varphi(x)+o(1/k_n)$$

$$= \Phi(x)+\frac{1}{\sqrt{k_n}}p_1(x)\varphi(x)+\frac{1}{k_n}p_2(x)\varphi(x)+o(1/k_n).$$

上面最后两个等式利用了对任意多项式 $p(x)$,$p(x)\varphi(x)$ 的各阶导数的有界性.

引理 7 若对某 $\delta\in(0,1/2)$,有 $k_n^{\frac{3}{2}+\delta}a_2\left(\frac{n}{k_n}\right)\to 0$,$k_n^{1+\delta}a_1\left(\frac{n}{k_n}\right)\to c<\infty$,$(n\to\infty)$,则对 $T_n=k_n^{\delta/2}$,当 $n\to\infty$ 时,在 $\{u:|u|\leqslant T_n\}$ 上一致地有

$$P\{H_n(u)\leqslant x\} = \Phi(x)+\sum_{i=1}^{2}k_n^{-\frac{i}{2}}p_i(x)\varphi(x)-\sqrt{k_n}a_1\left(\frac{n+1}{k_n}\right)\frac{\varphi(x)}{\gamma(1-\alpha_1)}+o(1/k_n).$$

证明 由 $|u|\leqslant T_n$,$T_n/\sqrt{k_n}\to 0,(n\to\infty)$ 知,$n\to\infty$ 时,

$$\left| \frac{1-(u)_n}{1-a_n}-1 \right| = \left| -\frac{\sqrt{a_n}}{\sqrt{k_n}}u \right| \leqslant \frac{\sqrt{a_n}}{\sqrt{k_n}}T_n \to 0,$$

所以当 n 充分大时,在 $\{u:|u|\leqslant T_n\}$ 上,由正规变化函数的局部一致收敛性知:

$$\sup_{|u|\leqslant T_n}\left| \frac{a_2((1-(u)_n)^{-1})}{a_2((1-a_n)^{-1})}-1 \right| \to 0. \tag{9}$$

对任给 $\rho>0$,利用切比雪夫不等式和引理 2 有

$$P\{|R_n(u)|\geqslant k_n^{-\rho}\}\leqslant k_n^{2\rho}ER_n^2(u)=k_n^{2\rho}\mathrm{Var}(R_n(u))=\frac{k_n^{2\rho}}{\gamma^2}\mathrm{Var}(R_{n,1}(u))$$
$$\leqslant\frac{k_n^{2\rho}}{\gamma^2}E(R_{n,1}(u))^2=\frac{k_n^{2\rho}a_2^2((1-(u)_n)^{-1})}{\gamma^2}E\left(\frac{R_{n,1}(u)}{a_2((1-(u)_n)^{-1})}\right)^2$$
$$\leqslant\frac{K^2k_n^{2\rho}a_2{}^2((1-(u)_n)^{-1})}{\gamma^2}E(1-U_1)^{-2\varepsilon}=\frac{K^2k_n^{2\rho}a_2^2((1-(u)_n)^{-1})}{(1-2\varepsilon)\gamma^2}$$
$$=\frac{K^2k_n^{2\rho-3-2\delta}(k_n^{3/2+\delta}a_2((1-(u)_n)^{-1}))^2}{(1-2\varepsilon)\gamma^2}=o\left(\frac{1}{k_n}\right),(0<\rho\leqslant 1+\delta).$$

所以,若记 $x_{n,\pm}(u)=x-r_n(u)\pm k_n^{-\rho}$,则对 $|u|\leqslant T_n$,$n>n_0$ 时有

$$P(S_n\leqslant x_{n,-}(u))+o\left(\frac{1}{k_n}\right)\leqslant P(H_n(u)\leqslant x)\leqslant P(S_n\leqslant x_{n,+}(u))+o\left(\frac{1}{k_n}\right),\tag{10}$$

利用引理 6,对 $u\in\{u:|u|\leqslant T_n\}$,在 R 一致地有

$$P(S_n(u)\leqslant x)=\Phi(x)+\frac{1}{\sqrt{k_n}}p_1(x)\varphi(x)+\frac{1}{k_n}p_2(x)\varphi(x)+o(1/k_n).$$

从而利用(11),对 $|u|\leqslant T_n$,$n>n_0$ 时有

$$\Phi(x_{n,-}(u))+\frac{1}{\sqrt{k_n}}p_1(x_{n,-}(u))\varphi(x_{n,-}(u))+\frac{1}{k_n}p_2(x_{n,-}(u))\varphi(x_{n,-}(u))+o\left(\frac{1}{k_n}\right)\leqslant P(H_n(u)\leqslant x)\leqslant$$
$$\Phi(x_{n,+}(u))+\frac{1}{\sqrt{k_n}}p_1(x_{n,+}(u))\varphi(x_{n,+}(u))+\frac{1}{k_n}p_2(x_{n,+}(u))\varphi(x_{n,+}(u))+o\left(\frac{1}{k_n}\right).\tag{11}$$

下面考虑 $r_n(u)$:

$$k_nr_n(u)=\frac{\sqrt{k_n}}{\gamma}\sum_{i=1}^{k_n}\left\{E(R_{n,i}(u))+a_1\left(\frac{1}{1-(u)_n}\right)E\left(h_{\alpha_1}\left(\frac{1}{1-U_i}\right)\right)\right\}$$
$$=\frac{k_n^{\frac{3}{2}}}{\gamma}\left\{E(R_{n,1}(u))+\frac{1}{1-\alpha_1}a_1\left(\frac{1}{1-(u)_n}\right)\right\}$$
$$=\frac{k_n^{\frac{3}{2}}}{\gamma}a_2\left(\frac{1}{1-(u)_n}\right)E\frac{R_{n,1}(u)}{a_2(\frac{1}{1-(u)_n})}+\frac{k_n^{\frac{3}{2}}}{\gamma(1-\alpha_1)}a_1\left(\frac{1}{1-(u)_n}\right),$$

所以

$$\left|k_nr_n(u)-\frac{k_n^{\frac{3}{2}}}{\gamma(1-\alpha)}a_1\left(\frac{n+1}{k_n}\right)\right|$$
$$\leqslant\left|\frac{k_n^{\frac{3}{2}}}{\gamma}a_2\left(\frac{1}{1-(u)_n}\right)E\left(\frac{R_{n,1}(u)}{a_2\left(\frac{1}{1-(u)_n}\right)}\right)\right|+\left|\frac{k_n^{\frac{3}{2}}}{\gamma(1-\alpha)}\left(a_1\left(\frac{1}{1-(u)_n}\right)-a_1\left(\frac{n+1}{k_n}\right)\right)\right|$$
$$=:r_{n,1}(u)+r_{n,2}(u).$$

而在 $|u|\leqslant T_n$,$n>n_0$ 时一致地有

$$r_{n,1}(u)\leqslant\frac{Kk_n^{\frac{3}{2}}}{\gamma}\left|a_2\left(\frac{1}{1-(u)_n}\right)\right|E(1-U_i)^{-\epsilon}=\frac{K}{\gamma(1-\epsilon)}\left|k_n^{\frac{3}{2}}a_2\left(\frac{1}{1-(u)_n}\right)\right|\to 0,$$

$$r_{n,2}(u)=\left|\frac{k_n^{\frac{3}{2}}a_2\left(\frac{n+1}{k_n}\right)}{\gamma(1-\alpha_1)}\right|\left|\frac{a_1\left(\frac{1}{1-(u)_n}\right)-a_1\left(\frac{n+1}{k_n}\right)}{a_2\left(\frac{n+1}{k_n}\right)}\right|$$
$$\leqslant\left|\frac{k_n^{\frac{3}{2}}a_2\left(\frac{n+1}{k_n}\right)}{\gamma(1-\alpha_1)}\right|\left|\frac{\left(\frac{k_n}{(n+1)(1-(u)_n)}\right)^{-\alpha}a_1\left(\frac{n+1}{k_n}\frac{k_n}{(n+1)(1-(u)_n)}\right)-a_1\left(\frac{n+1}{k_n}\right)}{a_2\left(\frac{n+1}{k_n}\right)}\right|$$

$$+\left|\frac{k_n^{\frac{3}{2}}a_2\left(\frac{n+1}{k_n}\right)}{\gamma(1-\alpha_1)}\right|\left|\frac{a_1\left(\frac{1}{1-(u)_n}\right)\left(1-\left(\frac{k_n}{(n+1)(1-(u)_n)}\right)^{-\alpha}\right)}{a_2\left(\frac{n+1}{k_n}\right)}\right|$$

$$=\left|\frac{k_n^{\frac{3}{2}}a_1\left(\frac{1}{1-(u)_n}\right)\left(1-\left(1-\frac{\sqrt{a_n}(u)_n}{\sqrt{k_n}}\right)^{\alpha}\right)}{\gamma(1-\alpha_1)}\right|+o(1)$$

$$\leqslant\left|\frac{k_n^{\frac{3}{2}}a_1\left(\frac{1}{1-(u)_n}\right)\left(\alpha\left(1-\theta\frac{\sqrt{a_n}(u)_n}{\sqrt{k_n}}\right)^{\alpha-1}\frac{\sqrt{a_n}(u)_n}{\sqrt{k_n}}\right)}{\gamma(1-\alpha_1)}\right|$$

$\leqslant Ck_n^{1+\delta/2}a_1\left(\frac{1}{1-(u)_n}\right)\to 0$. 其中 $\theta\in(0,1)$，C 为充分大的正数. 所以 $n\to\infty$ 时

$$\sup_{|u|\leqslant T_n}\left|k_n r_n(u)-\frac{k_n^{\frac{3}{2}}}{\gamma(1-\alpha_1)}a_1\left(\frac{n+1}{k_n}\right)\right|\to 0. \tag{12}$$

由 $\varphi''(x)$，$(p_i(x)\varphi(x))'$，$(i=1,2)$ 是 R 上的有界函数，即存在 $K>0$ 使

$$|\varphi''(x)|,\ |(p_i(x)\varphi(x))'|<K,$$

所以取 $1<\rho\leqslant 1+\delta$，$n\to\infty$ 时有

$$k_n\sup_{|u|\leqslant T_n}\left|\frac{p_1(x_{n,\pm}(u))\varphi(x_{n,\pm}(u))-p_1(x)\varphi(x)}{\sqrt{k_n}}\right|$$
$$\leqslant K\sqrt{k_n}\sup_{|u|\leqslant T_n}|x_{n,\pm}(u)-x|=K\sup_{|u|\leqslant T_n}\left|\sqrt{k_n}r_n(u)\pm k_n^{\frac{1}{2}-\rho}\right|\to 0; \tag{13}$$

$$k_n\sup_{|u|\leqslant T_n}\left|\frac{p_2(x_{n,\pm}(u))\varphi(x_{n,\pm}(u))-p_2(x)\varphi(x)}{k_n}\right|$$
$$\leqslant K\sup_{|u|\leqslant T_n}|x_{n,\pm}(u)-x|=K\sup_{|u|\leqslant T_n}\left|r_n(u)\pm k_n^{-\rho}\right|\to 0; \tag{14}$$

$$k_n\sup_{|u|\leqslant T_n}\left|\Phi(x_{n,\pm}(u))-\left\{\Phi(x)-k_n^{\frac{1}{2}}\frac{\varphi(x)}{\gamma(1-\alpha_1)}a_1\left(\frac{n+1}{k_n}\right)\right\}\right|$$
$$\leqslant k_n\sup_{|u|\leqslant T_n}\left|\Phi(x_{n,\pm}(u))-\{\Phi(x)+\varphi(x)[x_{n,\pm}(u)-x]\}\right|$$
$$+k_n\sup_{|u|\leqslant T_n}\left|\varphi(x)[x_{n,\pm}(u)-x]+k_n^{\frac{1}{2}}\frac{\varphi(x)}{\gamma(1-\alpha_1)}a_1\left(\frac{n+1}{k_n}\right)\right|$$
$$\leqslant K(\sup_{|u|\leqslant T_n}(k_n(r_n^2(u)+2k_n^{-\rho}|r_n(u)|))+k_n^{1-2\rho})$$
$$+\varphi(x)\left[k_n^{1-\rho}+\sup_{|u|\leqslant T_n}\left|k_n r_n(u)-\frac{k_n^{3/2}}{(\gamma(1-\alpha_1))}a_1\left(\frac{n+1}{k_n}\right)\right|\right]\to 0. \tag{15}$$

结合(11)、(13)、(14)、(15)，得引理 7.

定理的证明　由引理 3 有

$$P\{H_n\leqslant x\}=\int_{-\infty}^{+\infty}P\{H_n(u)\leqslant x\}\phi_n(u)\mathrm{d}u. \tag{16}$$

由引理 7，

$$\int_{-T_n}^{T_n}P\{H_n(u)\leqslant x\}\phi_n(u)\mathrm{d}u=\Phi(x)+\sum_{i=1}^{2}k_n^{-\frac{i}{2}}p_i(x)\varphi(x)-k_n^{\frac{1}{2}}a_1\left(\frac{n+1}{k_n}\right)\frac{\varphi(x)}{\gamma(1-\alpha_1)}+o(1/k_n). \tag{17}$$

在 R 上一致成立. 由 $T_n=k_n^{\frac{1}{2}-\delta}$，$\delta$ 为 $(0,\frac{1}{2})$ 上的一个确定的数，利用引理 4，就有

$$\int_{|u|>T_n}P\{H_n(u)\leqslant x\}\phi_n(u)\mathrm{d}u\leqslant P\left\{\left|\frac{U_{n-k_n,n}-a_n}{b_n}\right|>T_n\right\}\leqslant\exp\left(-\frac{T_n^2}{3\left(1+\frac{T_n}{b_n\sqrt{n}}\right)}\right)=o\left(\frac{1}{k_n}\right). \tag{18}$$

结合(16)、(17)、(18)，即得所要结果，定理证毕.

[参考文献]

[1] Müller U A, Dacorogna M M, Pictet O V. A Practical Guide to Heavy Tails[M]. Boston: Birkhäuser, 1998.55—77.

[2] Resnick S I. A Practical Guide to Heavy Tails[M]. Boston: Birkhäuser, 1998.219—239.

[3] Reiss R D. Estimating the tail index of the claim size distribution[J]. Blätter der DGVM, 1987, 18: 21—25.

[4] Smith R L. Estimating tails of probability distributions[J]. Ann Statist, 1987, 15: 1174—1207.

[5] Mcneil A J, Frey R. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach[J]. Journal of Empirical Finance, 2000, 7: 271—300.

[6] Mikosch T, Nagaev A V. Large deviations of heavy-tailed sums with applications in insurance[J]. Extremes, 1998, 1: 81—110.

[7] Hill B M. A simple approach to inference about the tails of a distribution[J]. Ann Statist, 1975, 3: 1163—1174.

[8] Hall P. On the rate of convergence of normal extremes[J]. J Appl Probab, 1979, 16: 433—439.

[9] Häusler E, Teugels J L. On asymptotic normality of Hill's estimator for the exponent of regular variation[J]. Ann Statist, 1985, 13: 743—756.

[10] Csörgö S, Deheuvels P, Mason D. Kernel estimates of the tail index of a distribution[J]. Ann Statist, 1985, 13: 1050—1077.

[11] Beirlant J, Teugels J L. Asymptotics of Hill's estimator[J]. Theory Probab Appl, 1986, 31: 463—469.

[12] de Haan L. Extreme value statistics, in Extreme value theory and applications[M]. Dordrecht: Kluwer Academic, 1994.93—122.

[13] Cheng S, Pan J. Asymptotic expansions of estimators in extreme statistics[J]. in Proc 50th Sess Int Statist Inst IP, 1995, 15(1): 593—605.

[14] Cheng S. Large deviation theorem for Hill's estimator[J]. Acta Mathematica Sinica, New Series, 1992, 8(3): 243—254.

[15] de Haan L. On regular variation and its application to the weak convergence of sample extremes[M]. Amsterdam: Math Centrum, 1970.

[16] Bingham N H, Goldie C M, Teugels J L. Regular Variation[M]. Cambridge: Cambridge University Press, 1987.

[17] Wang Xiaoqian. Problems of the Actuarial Science and Extreme Value Theory[C]. Beijing: Peking University, 2002

[18] de Haan L, Stadtmiiller U. Generalized regular variation of second- order[J]. J Austral Math Soc Ser A, 1996, 61: 381—395.

[19] de Haan L, Peng L. Comparison of tail index estimators[J]. Statist Neerl, 1998, 52: 60—70.

[20] Häusler E, Hall P, Welsh A H. Best attainable rates of convergence for estimates of parameters of regular variation[J]. Ann Statist, 1984, 12: 1079—1084.

[21] Cheng S, Pan J. Asymptotic expansions of estimators for the tail index with applications[J]. Scand J Statist, 1998, 25: 717—728.

[22] Cheng S, de Haan L. Penultimate Approximation for Hill's estimator[J]. Scand J Statist, 2001, 28: 569—575.

[23] Cuntz A, Häusler E. On Edgeworth expansions for the distribution function of the Hill estimator[Z]. Germany: University of Giessen, 2001.

[24] Cheng S, Jiang C. The edgeworth expansion for distributions of extreme values[J]. Science in China (Series A), 2001, 44: 427—437.

[25] Reiss R D. Approximate distributions of order Statistics[M]. New York: Springer-Verlag, 1989.

[26] Petrov V V. Sums of independent random variables[M]. New York: Springer-Verlag, 1975.

[27] Feller W. An introduction to probability theory and its applications II[M]. New York: Wiley, 1966.

The Second Order Edgeworth Expansion of the Distribution of the Hill Hstimator

Wang Xiaoqian[1], Cheng Shihong[2]

(1. School of Mathematics and Computer Science, Nanjing Normal University, 210097, Nanjing, China)

(2. College of Mathematics Science, Peking University, 100871, Beijing, China)

Abstract: The estimation of the tail index of the heavy tailed distribution with regular variation tail is an important problem. In this paper we study the second order Edgeworth expansion of the distribution of the Hill estimator of the tail index by second order regular variation function.

Key words: second order regular variation function, Hill-estimator, Edgeworth expansion

[责任编辑: 陆炳新]