山西大学学报(自然科学版)35(2):163～173,2012
Journal of Shanxi University(Nat. Sci. Ed.)

·山西大学 110 周年校庆特刊·

文章编号:0253-2395(2012)02-0163-11

重尾分布尾指数估计研究进展

刘维奇,邢红卫

(山西大学 管理学院,山西 太原 030006)

摘　要:气象、水文、环境、电信、保险、金融等许多领域的数据不满足正态分布假设,而是具有尖峰、重尾特征.过去三十多年间,针对重尾分布及尾指数估计的研究得到了长足发展.文章回顾了冲击正态性假设的重尾分布的发现过程,描述了重尾分布的定义,极值理论及正则变换条件,并从研究内容的阶段特征、研究方法的不同类型总结、归纳、评述了各类尾指数估计方法及重尾阈值选取方法,最后就这些估计方法的不足和应用局限性以及如何改进和深化重尾指数的估计问题作了展望.

关键词:极值理论;重尾指数估计;重尾阈值;降偏差估计

中图分类号:O212　　　文献标识码:A

0　重尾分布的发现

23 岁毕业于牛津大学、就职于爱尔兰都柏林吉尼斯酿造公司(Guinness Brewing Company)的化学技师戈塞特(Gosset),在长期从事实验和数据分析工作中发现了 T 分布,由于可能涉及公司商业机密外泄,Gosset 以“Student”的笔名发表了此项结果,这也是后人称 T 分布为“学生氏分布”的原因.然而,在当时正态分布一统天下的情形下,Gosset 的发现并没有被外界理解和接受.直到 1923 年英国统计学家费歇尔(R. A. Fisher)给出 T 分布的严格推导并于 1925 年编制分布表之后,T 分布才得到学术界承认,并得到快速的传播、发展和应用.T 分布堪称一类典型的重尾分布.1949 年,Zipf[1] 在研究美国城市规模问题时,发现第二大城市的规模只有第一大城市规模的一半,而第三大城市规模是第一大城市规模的 1/3,发现城市规模的频率分布图近似于下降的双曲线,具有重尾特征,在研究词频问题时发现了同样的特征.1963 年,Mandelbrot[2] 发现棉花的期货价格相对于正态分布也具有尖峰、重尾特征,服从一族 Levy 稳定分布.

重尾现象几乎出现于所有经济、生活领域,金融资产、保险索赔、气象观察、网络流量以及众多人类行为的数据均表现出尖峰重尾特征,由此,重尾分布得到金融、统计、环境科学等学术界和实践者的关注和重视.在金融方面,Fama[3] 发现了股票收益的尖峰重尾特征,Mandelbrot[2] 发现棉花期货价格也具有这一特征,Brodin 和 Kluppelberg[4] 介绍了重尾分布在金融资产风险管理方面的应用,Bollerslev 和 Todorov[5] 将重尾分布及重尾指数估计应用于金融资产定价模型.在气象、环境科学方面,Tessier et al.[6] 发现降雨量和河流水文数据具有尖峰重尾特征,并以重尾分布进行建模,Richard[7] 介绍了重尾分布在大气臭氧层研究、风速和降雨量预测等方面的应用,Omey et al.① 也将重尾分布应用于极端风速的研究,Allan 和 Soden[8] 以重尾分布预测了温室效应和极端降雨量.在人类社会现象及其它方面,除 Zipf[1] 在研究人类学及语言学问题时发现重尾分布并加以研究外,Barabasi[9] 在研究宇宙爆炸起源和人类动力学问题时也利用了重尾分布,Vazquez[10]、Vazquez et al.[11] 利用重尾分布对 Barabasi[9] 的人类动力学模型进行了进一步扩展.

重尾分布由于其对数据刻画更具现实性,因此激发了金融、保险、气象、环境、通讯等领域的研究需求.重

* 收稿日期:2012-03-16;修回日期:2012-03-19

基金项目:山西省高校人文社科重点研究基地项目(2011305)

作者简介:刘维奇(1963－),男,山西忻州人,博士,教授,主要从事金融工程与时间序列分析研究.Email:liuwq@sxu.edu.cn

① Omey E,Mallor F,Nualor E. An Introduction to Statistical Modeling of Extreme Values. Application to Calculate Extreme Wind Speedsl. 2009,working paper. http://ideas.repec.org/p/hub/wpecon/200936.html.

尾分布也由于其非正态表现的特质性,给统计学家在研究方法和研究难度上提出了挑战,关于重尾分布的理论研究和应用研究任重道远.

1 重尾分布的定义

重尾分布可以从分布函数的矩性质、矩母函数性质和分布的尾部性质来定义,分别阐述如下.

1.1 以四阶矩定义

重尾分布的定义相对于正态分布而言,是以其四阶中心矩为基础的.

定义 1.1 随机变量 X 服从的分布 F 是重尾的,若

$$E\left[\frac{(X-\mu_X)^4}{\sigma_X^4}\right]>3, \tag{1}$$

其中 μ_X,σ_X 分别为 X 的期望和标准差. 由于正态分布的峰度(kurtosis)为 3,因此该性质被称为尖峰,该定义只适用于 4 阶矩存在的情形(详见 Thomas et al. 2002)[12].

1.2 以矩母函数定义

定义 1.2 称分布函数 F 是重尾分布,或其随机变量 X 服从于重尾(heavy-tailed)分布,如果对 $\forall\lambda>0$,

$$Ee^{\lambda X}=\int_0^\infty e^{\lambda x}dF(x)=\infty, \tag{2}$$

或对某个 $\varepsilon>0$,

$$\lim_{x\to+\infty}\frac{\overline{F}(x)}{e^{-\varepsilon x}}=+\infty, \tag{3}$$

即不存在指数阶矩(详见 Embrechts et al.)[13].

1.3 以分布尾部的收敛速率定义

定义 1.3 如果密度函数是以幂指数速率衰减至 0,称该分布函数是重尾的;如果密度函数是以指数速率衰减至 0 的,称该分布函数是轻尾的(详见 Embrechts et al.)[13].

目前在重尾分布研究领域,大多是以正则变化理论为基础,因此分布尾部的收敛速率成为研究主题,而其中的重点和难点是尾指数估计. 考虑到这一原因,定义 1.3 更具有普遍指导意义. 一般认为,分布尾部收敛速率快于正态分布尾部收敛速率的为轻尾分布,而尾部收敛速率慢于正态分布尾部收敛速率的为重尾分布[14].

2 极值理论基础

假设 $X_1,X_2,\cdots,X_n$ 是分布 F 的随机变量,

$$F(x)=\Pr(X_i\leqslant x). \tag{4}$$

令 $X_{1,n}\leqslant X_{2,n}\leqslant\cdots\leqslant X_{n,n}$ 为 $X_1,X_2,\cdots,X_n$ 的次序统计量,若存在常数序列 a_n 和 $b_n\in\mathbf{R}$,使得 $(X_{n,n}-b_n)/a_n$ 收敛于非退化极值分布 G,

$$\Pr((X_{n,n}-b_n)/a_n\leqslant x)=F(a_nx+b_n)^n\xrightarrow{d}G,n\to\infty, \tag{5}$$

则 $F\in D_M(EV_\gamma)$,即 F 属于极值分布 EV_γ 的吸引域,其中 γ 为重尾指数.

如果非退化分布 G 存在,则应属于以下三类分布之一:

(1)Fréchet 分布 $\Phi_\gamma(x)=\begin{cases}0,x\leqslant0\\ \exp(-x^{-1/\gamma}),x>0\end{cases}$;

(2) Weibull 分布 $\Psi_\gamma(x)=\begin{cases}\exp(-(-x)^{1/\gamma}),x\leqslant0\\ 1,x>0\end{cases}$;

(3)Gumbel 分布 $\Lambda(x)=\exp(-e^{-x})$,

这三类分布可以统一写为如下的极值分布

$$EV_\gamma(x)=\begin{cases}\exp(-(1+\gamma x)^{-1/\gamma}),1+\gamma x>0,\gamma\neq0\\ \exp(-\exp(-x)),x\in\mathbf{R},r=0\end{cases}. \tag{6}$$

对 $x>0$,正则变化一阶条件为

$$\bar{F}(x)=x^{-\alpha}L(x),\tag{7}$$

其中 $L(x)$ 为慢变化函数,记为 $\bar{F}\in RV_{-\alpha}$,$\alpha=1/\gamma$. 设分位数函数 $U(t)=(1/(1-F))^{\leftarrow}(t)=F^{\leftarrow}(1-1/t)$,$t>1$,其中 $F^{\leftarrow}(x)=\inf\{y:F(y)\geqslant x\}$,则具有如下等价定义.

(1)$\lim_{t\to\infty}\dfrac{\bar{F}(tx)}{\bar{F}(t)}=x^{-\alpha}$,

(2)$\lim_{t\to\infty}\dfrac{U(tx)}{U(t)}=x^{\gamma}$,即 $F\in D_M(EV_\gamma)$.

如果 $\bar{F}:=1-F$ 是以指数 α 的正则变化函数,则 U 以指数 γ 正则变化.

二阶条件为

$$\lim_{t\to\infty}\frac{\ln U(tx)-\ln U(t)-\gamma\ln x}{A(t)}=\frac{x^{\rho}-1}{\rho},x>0.\tag{8}$$

二阶参数 $\rho(\rho<0)$ 反映了一阶条件(7)的收敛速率,$|A(t)|$ 以指数 ρ 正则变化,即 $|A(t)|\in RV_\rho$. 特别对于标准的 Pareto 分布 $F(x)=1-Cx^{-1/\gamma}$,$x>C^{\gamma}$,$\ln U(tx)-\ln U(t)-\gamma\ln x\equiv 0$.

Pickands 估计[15],Hill 估计[16],矩估计[17]等一系列参数或半参数估计都密切依赖于选取的次序统计量个数 k,如果对慢变化函数 $L(x)$ 没有严格的限制条件,很难确定用于重尾指数估计的 k 值. 对于 Pareto 型分布,可以假设

$$1-F(x)=Cx^{-1/\gamma}(1+Dx^{\rho/\gamma}+o(x^{\rho/\gamma})),\tag{9}$$

$C>0$,$\rho<0$,D 为非零实数,(8)成立且可以选择 $A(t)=\gamma\rho DC^{\rho}t^{\rho}$.

为了获得二阶参数估计的性质,进一步假设三阶条件

$$\lim_{t\to\infty}\frac{\dfrac{\ln U(tx)-\ln U(t)-\gamma\ln x}{A(t)}-\dfrac{x^{\rho}-1}{\rho}}{B(t)}=\frac{x^{\rho+\rho'}-1}{\rho+\rho'},x>0.\tag{10}$$

三阶参数 $\rho'(\rho'<0)$ 反映了二阶条件(8)的收敛速率,$|B(t)|$ 以指数 ρ' 正则变化,即 $|B(t)|\in RV_{\rho'}$. 这个条件已经被 Gomes et al. [18],Fraga Alves et al. [19]用来研究 ρ 估计的渐近性质,Gomes et al. [20]用来研究降偏差尾指数估计,Caeiro 和 Gomes[21]以其为基础研究尾指数估计和高分位点.

对于 Pareto 型分布,进一步具体化 $o(x^{\rho/\gamma})$,假设

$$\bar{F}(x)=Cx^{-1/\gamma}(1+D_1x^{\rho/\gamma}+D_2x^{(\rho+\rho_1)/\gamma}+o(x^{(\rho+\rho_1)/\gamma})),\tag{11}$$

$C>0$,D_1,D_2 为非零实数,$\rho,\rho_1<0$,$\rho'=\max(\rho,\rho_1)>\rho$.

对于大多数重尾分布,三阶参数 ρ' 等于二阶参数 ρ,即

$$\bar{F}(x)=Cx^{-1/\gamma}(1+D_1x^{\rho/\gamma}+D_2x^{2\rho/\gamma}+o(x^{2\rho/\gamma})),\tag{12}$$

例如 Fréchet 分布 $F(x)=\exp(-x^{-1/\gamma})$,$x\geqslant 0$,$\rho'=\rho=-1$;广义 Pareto 分布 $F(x)=1-(1+\gamma x)^{-1/\gamma}$,$x\geqslant 0$,$\rho'=\rho=-\gamma$;Burr 分布 $F(x)=1-(1+x^{-\rho/\gamma})^{1/\rho}$,$x\geqslant 0$,$\rho'=\rho<0$;Student-$t_n$ 分布(n 为自由度),$\gamma=1/n$,$\rho'=\rho=-2/n$.

3 传统的重尾指数估计方法

由于对极端事件的敏感和关注,有必要尽可能了解分布尾部的全部信息,最直接的方式是得到尾分布函数. 假设了尾分布函数的参数形式后,这个问题就转化为如何估计函数当中的未知参数,因此估计重尾指数的本质是为了估计尾分布函数.

3.1 重尾指数估计方法

1949 年 Zipf[1]在研究词频和城市规模等问题时发现了尖峰重尾特征的分布,并利用 Zipf-plot 给出了重尾指数的估计方法. Fama 和 Roll[22]、Press[23]和 Zolotarev[24]分别基于稳定分布($0<\alpha<2$)的特征函数给出了指数 α 的不同估计方法. Pickands[15]在假定尾部分布函数的前提下,通过求分位数给出了基于三个次序统计量的 Pickands 估计

$$\hat{\gamma}_n^P(m)=\frac{1}{\log 2}\log\frac{X_{n-m+1,n}-X_{n-2m+1,n}}{X_{n-2m+1,n}-X_{n-4m+1,n}},\tag{13}$$

其中 $X_{1,n}\leqslant X_{2,n}\leqslant\cdots\leqslant X_{n,n}$ 是随机变量 $X_1,X_2,\cdots,X_n$ 的次序统计量. Dekkers et al.[17] 研究了 Pickands 估计的渐近性质. Pickands 估计相对于其它估计而言,它是一种相合估计,然而缺点也很明显,这种估计的渐近有效性比较差.

Hill[16] 在假设 F 为 Pareto 分布的前提下,以极大似然方法给出了重尾指数 γ 的估计

$$H_n(k)=\hat{\gamma}=\frac{1}{k}\sum_{i=1}^{k}\log X_{n-i+1,n}-\log X_{n-k,n}, \tag{14}$$

其中 $X_{1,n}\leqslant X_{2,n}\leqslant\cdots\leqslant X_{n,n}$ 是随机变量 $X_1,X_2,\cdots,X_n$ 的次序统计量,k 是估计中所选取的极大次序统计量的个数. de Haan[25] 研究了 Hill 估计的渐近性质,Hill 估计具有强相合性以及渐近正态性,当随机变量不是 i. i. d 时也能得到类似的结果,但是 Hill 估计要求 $\gamma>0$.

de Haan 和 Resnick[26] 提出了

$$\hat{\gamma}=(\log X_{n,n}-\log X_{n-k,n})/\log k, \tag{15}$$

Davison[27] 提出了极大似然估计

$$\hat{\gamma}=\frac{1}{k}\sum_{i=1}^{k}\log\left(1+\hat{a}(X_{n-i+1,n}-X_{n-k,n})\right), \tag{16}$$

其中 $\hat{a}$ 是广义 Pareto 分布 $\mathrm{GP}_\gamma(ax/\gamma)$,$a>0$ 的内含规模参数 a 的极大似然估计.

Csörgö et al.[28] 提出了核估计方法

$$\hat{\gamma}=\frac{\sum_{i=1}^{k}(i/k)K(i/k)[\log X_{n-i+1,n}-\log X_{n-i,n}]}{\int_0^1 K(t)\mathrm{d}t}, \tag{17}$$

其中 K 是定义在(0,1)上的非负非增核. 当 $K(u)=1_{(0,1)}(u)$ 时就是 Hill 估计,由 de Haan 和 Resnick[26] 提出 $\hat{\gamma}=(\log X_{n,n}-\log X_{n-k,n})/\log k$,de Haan[29],Bacro 和 Brito[30] 提出的估计方法也都属于核估计,随后有 Smith[31-32] 的参数估计方法.

Dekkers et al.[17] 提出了矩估计方法

$$\hat{\gamma}=M_n^{(1)}+1-\frac{1}{2}\left(1-\frac{(M_n^{(1)})^2}{M_n^{(2)}}\right)^{-1}, \tag{18}$$

其中 $M_n^{(j)}=\frac{1}{k}\sum_{i=1}^{k}(\ln X_{n-i+1,n}-\ln X_{n-k,n})^j$,$j=1,2$. 当分布是 Pareto 分布时,矩估计是渐近无偏的,但是该估计的方差比 Hill 估计的大. 所以有 Deheuvels et al.[33] 对矩估计方法的修正,提出 peng 估计及 W 估计. 这两种估计都比矩估计的方差小,但是 W 估计在 $\gamma<1/2$ 时才是一致估计,在 $\gamma<1/4$ 时渐近正态.

1996 年 Kratz 和 Resnick[34] 提出了基于最小二乘法的 qq-plot 估计,该估计的残差所包含的信息在一定程度上能抵消偏差,但渐近方差却是 Hill 估计渐近方差的 2 倍,同年 Beirlant et al.[35] 提出 Pareto quantile plots 方法,对于小样本来说能得到很好的估计结果,所选择的权重也能克服不收敛问题. 但是该方法对所选择的权重很敏感,权重有时会对估计产生负面影响.

Fraga Alves et al.[36] 提出了混合矩估计(MM 估计)

$$\hat{\gamma}_n^{MM}(k)=\frac{\hat{\varphi}_{n,k}-1}{1+2\min(\hat{\varphi}_{n,k}-1,0)}, \tag{19}$$

其中 $\hat{\varphi}_{n,k}=\frac{M_n^{(1)}(k)-L_n^{(1)}(k)}{(L_n^{(1)}(k))^2}\xrightarrow{P}\varphi(\gamma)=\begin{cases}1+\gamma,\gamma>0\\ \frac{1-\gamma}{1-2\gamma},\gamma\leqslant 0\end{cases}$,$L_n^{(j)}(k)=\frac{1}{k}\sum_{i=1}^{k}\left(1-\frac{X_{n-k,n}}{X_{n-i+1,n}}\right)^j$.

3.2 重尾阈值的选取

如何选取重尾阈值,或者如何选取用于估计的次序统计量个数 k,经验上的初步判断是选取次序统计量的 2%-10%. 学者们从理论上提出了选取重尾阈值 k 的许多方法. 其中一类是作图法,比如 Hill[16] 提出的 Hill-plot,Kratz 和 Resnick[34] 提出的 qq-plot,Beirlant et al.[35] 提出的 Pareto 分位数图,Resnick 和 Starica[37] 给出的对 Hill-plot 改进的 smooHill-plot,以及 Drees et al.[38] 给出的对 Hill-plot 改进的 AltHill-plot

等,这些作图法都有一定的优越性,但整体而言都不能适用于所有重尾分布族.像 Hill-plot,qq-plot,当随机变量服从 Pareto 分布时,这两种方法表现出十分优良的性质,能够容易选取重尾阈值.一旦随机变量不服从 Pareto 分布,却不能很好地选取重尾阈值,甚至无法选取重尾阈值.Pareto 分位数图,smooHill-plot 和 AltHill-plot 相对于 Hill-plot 估计精度稍高一些,但是也不能对所有的重尾分布较好地选择重尾阈值.Sousa[39]在其博士论文中提出的 Sum-plot 方法在一定程度上克服了前几种方法中选取重尾阈值所遇到的困难,而且具有比较好的性质.但是由于 Sum-plot 方法是以观察图形得到重尾阈值,因此选择重尾阈值有一定的猜测性,因而会对重尾指数估计造成一定误差.另一类方法就是以估计的最小均方误差(MSE)为标准来确定重尾阈值,最优阈值应该与最小化均方误差一致.1990 年 Hall[40]提出利用 Bootstrap 方法来选取重尾阈值,Danielsson[41]在 2001 年对 Hall 的方法作了进一步改进,提出了 Danielsson-Bootstrap 方法,然而刘维奇等[42]在 2010 年的实证研究发现,Bootstrap 方法受异常值和样本容量的影响较大,并在此基础上提出了收敛速度更好的 M-Bootstrap 方法.2010 年刘维奇和邢红卫[43]提出一种选取重尾阈值的简便优化方法,可以得到重尾阈值的解析结果.

以上这些估计方法各有特点,各有优劣,然而也存在一些共有的缺点.首先,这些方法都依赖于用于估计的次序统计量的个数 k,或者说依赖于重尾阈值,并且估计路径波动明显,对选取的 k 异常敏感;其次,Hill 类估计对于样本不具有位置变换不变性,对样本进行位置变换后,估计就会失效,对现实问题的适用性大打折扣,而 Pickands 估计具有位置变换不变性,这一点有别于 Hill 估计;第三,这些估计方法在有限样本情形具有不可忽略的偏差,处理实际问题时难以得到优良的估计结果.对于传统估计方法的这些缺点,在重尾指数估计的发展历程中又提出了一些针对性的改进方法.

4 改进的重尾指数估计方法

4.1 降偏差重尾指数估计

上世纪 90 年代末至今,研究重尾分布尾部指数估计的重点有所突破.学者们不太关注重尾阈值的选取,而是希望能够提出在有限样本情形下偏差小而且对阈值有优良稳健性的估计方法,这样既符合无法取到无限样本的实际情况,又不至于对阈值选取有苛刻的要求,于是出现了许多稳健性较好的降偏差重尾指数估计.

作为重尾指数 γ 最经典的估计方法,Hill 估计具有渐近无偏性和渐近正态性,然而在有限样本情形估计有偏差.Hill 估计的偏差目前在国外被广泛研究,例如 de Haan 和 Peng[44],Beirlant et al.[45],Feuerverger 和 Hall[46],Beirlant et al.[47],Caeiro 和 Gomes[48].这些文献在 Hill 估计的基础上都提出了二阶降偏差尾指数估计,然而渐近偏差总是大于或等于 $(\gamma(1-\rho)/\rho)^2>\gamma^2,\rho<0$.最近,Caeiro et al.[49],Gomes 和 Pestana[50],Gomes et al.[51],Gomes et al.[52]等在二阶条件下以不同的方式得到了二阶参数的估计 $\hat{\beta}$ 和 $\hat{\rho}$,在 $\sqrt{k}A(n/k)\to\lambda$ 有限,$n\to\infty$,能够降低偏差而渐近方差依然等于 Hill 估计的渐近方差 γ^2.

Gomes 和 Martins[53]提出了 Generalized Jackknife(GJ)估计

$$GJ_{\hat{\rho}}(k)=\left(\sqrt{2M_n^{(2)}(k)}-(2-\hat{\rho})M_n^{(2)}(k)/(2M_n^{(1)}(k))\right)/\hat{\rho},\tag{20}$$

其中

$$M_n^{(j)}(k)=\frac{1}{k}\sum_{i=1}^{k}V_{ik}^{j}=\frac{1}{k}\sum_{i=1}^{k}\left(\ln\frac{X_{n-i+1,n}}{X_{n-k,n}}\right)^{j},$$

GJ 估计的渐近方差为 $\gamma^2((1-\rho)/\rho)^2$,是 Drees 类估计中渐近方差最小的.

Beirlant et al.[54]提出了广义 Hill(GH)估计

$$GH_n(k)=H_n(k)+\frac{1}{k}\sum_{i=1}^{k}(\ln H_n(i)-\ln H_n(k)),\tag{21}$$

在二阶条件下,如果 $\sqrt{k}A(n/k)\to\lambda\in R$,有

$$\sqrt{k}(\hat{\gamma}_{n,k}^{GH}-\gamma)\xrightarrow{d}N(\lambda/(1-\rho),\sigma_{GH}^2),$$

$$\sigma_{GH}^{2}=\begin{cases}1+\gamma^{2},\gamma\geqslant 0\\ \dfrac{(1-\gamma)(1+\gamma+2\gamma^{2})}{1-2\gamma},\gamma<0\end{cases}.$$

Gomes 和 Pestana[50]提出了

$$\overline{H}_{\hat{\beta},\hat{\rho}}(k)=H_n(k)-\frac{\hat{\beta}}{1-\hat{\rho}}(n/k)^{\hat{\rho}}H\left(\left(\frac{(1-\hat{\rho})^{2}n^{-2\hat{\rho}}}{-2\hat{\rho}\hat{\beta}^{2}}\right)^{1/(1-2\hat{\rho})}\right),\tag{22}$$

Caeiro et al.[49]提出了修正 Hill 估计

$$CH_{\hat{\rho},\hat{\beta}}(k)=H_n(k)\left(1-\frac{\hat{\beta}}{1-\hat{\rho}}(n/k)^{\hat{\rho}}\right),\tag{23}$$

以及其渐近等价估计

$$\overline{CH}_{\hat{\rho},\hat{\beta}}(k)=H_n(k)\exp\left(-\frac{\hat{\beta}}{1-\hat{\rho}}(n/k)^{\hat{\rho}}\right).\tag{24}$$

在二阶条件及

$$A(t)=\alpha t^{\rho}=:\gamma\beta t^{\rho},B(t)=\beta' t^{\rho'},\beta,\beta'\neq 0\tag{25}$$

成立的前提下,Hill 估计的偏差 $A(n/k)/(1-\rho)=\gamma\beta(n/k)^{\rho}/(1-\rho)$ 由 $H_n(k)\hat{\beta}(n/k)^{\hat{\rho}}/(1-\hat{\rho})$ 估计得到.

$\overline{H}$ 估计和 CH 估计有同样的构造方法,不同之处在于 $\overline{H}$ 估计对于 Hill 估计的偏差 $\gamma\beta(n/k)^{\rho/(1-\rho)}$ 由 $H_n(\hat{k}_0)\hat{\beta}(n/k)^{\hat{\rho}}/(1-\hat{\rho})$ 估计得到,其中 $\hat{k}_0=\left(\frac{(1-\hat{\rho})^{2}n^{-2\hat{\rho}}}{-2\hat{\rho}\hat{\beta}^{2}}\right)^{1/(1-\hat{\rho})}$ 是基于 Hill 估计均方误差最小得到的.

Gomes et al.[55]提出了修正 CM 估计

$$CM_{\hat{\rho},\hat{\beta}}(k)=\hat{\gamma}_n^{M}(k)\left(1-\frac{\hat{\beta}}{1-\hat{\rho}}(n/k)^{\hat{\rho}}\right)-\frac{\hat{\beta}\hat{\rho}}{(1-\hat{\rho})^{2}}(n/k)^{\hat{\rho}}.\tag{26}$$

Gomes et al.[51]提出了

$$ML_{\hat{\beta},\hat{\rho}}(k)=H_n(k)-\hat{\beta}(n/k)^{\hat{\rho}}\left(\frac{1}{k}\sum_{i=1}^{k}(i/k)^{-\rho}U_i\right),\tag{27}$$

由 $U_i\approx\gamma\exp(\beta(n/i)^{\rho})E_i,1\leqslant i\leqslant k$,得到 γ 的极大似然估计

$$\overline{ML}_{\hat{\beta},\hat{\rho}}(k)=\sum_{i=1}^{k}\exp(-\hat{\beta}(n/i)^{\hat{\rho}})U_i/k.\tag{28}$$

Gomes et al.[52]提出了

$$\overline{WH}_{\hat{\beta},\hat{\rho}}(k)=\frac{1}{k}\sum_{i=1}^{k}\exp(\hat{\beta}(n/k)^{\hat{\rho}}\Psi_{ik}(\hat{\rho}))V_{ik},\Psi_{ik}=\Psi_{ik}(\hat{\rho})=\frac{(i/k)^{-\hat{\rho}}-1}{\hat{\rho}\ln(i/k)},\tag{29}$$

WH 估计被称为加权 Hill 估计.

Gomes et al.[51]通过比较 ML 和$\overline{ML}$的渐近均方误差,发现 ML 的表现要优于$\overline{ML}$. 对于大多数重尾分布族,ML 比$\overline{ML}$有更好的表现,因此对于$\overline{WH}_{\beta,\rho}(k)$而言,更应选择

$$WH_{\hat{\beta},\hat{\rho}}(k)=H_n(k)-\hat{\beta}(n/k)^{\hat{\rho}}\left(\frac{1}{k}\sum_{i=1}^{k}\Psi_{ik}(\hat{\rho})V_{ik}\right).\tag{30}$$

4.2　位置变换不变性重尾指数估计

大部分 Hill 类估计,如矩估计、核估计、混合矩估计,对样本进行规模变换后满足不变性,对样本进行位置变换后却会失效,然而在实际问题中往往需要对样本进行位置变换. 为此,除了上文提到的 Pickands 估计和极大似然估计等位置变换不变性估计外,一些学者在考虑估计特性的同时也提出了另外的位置变换不变性估计方法,并使这一问题成为今后研究重尾分布尾指数估计的热点之一.

Falk[56]提出了类似于 Hill 估计的 NH 估计($\gamma<0$),

$$\hat{\gamma}_n^{NH}(k)=\frac{1}{k-1}\sum_{i=2}^{k}\ln\frac{X_{n,n}-X_{n-i+1,n}}{X_{n,n}-X_{n-k,n}},\tag{31}$$

Müller 和 Husler[57]改进了 NH 估计,

$$\hat{\gamma}_n^{IH}(q,k)=\frac{1}{k-1}\sum_{i=2}^{k}\ln\frac{\hat{w}-X_{n-i+1,n}}{\hat{w}-X_{n-k,n}},\tag{32}$$

其中 $\hat{w}=\sum_{j=1}^{q}a_j(\hat{\gamma}_1)X_{n-j+1,n}$,$\hat{\gamma}_1$ 是 γ 的相合估计,a_j 为相应的权重.

Weiss[58]、Peng 和 Nadarajah[59]研究了 Weiss 类 Pickands 估计($\gamma<0$),

$$\hat{\gamma}_n^W(m)=\frac{1}{\log 2}\log\frac{X_{n,n}-X_{n-m+1,n}}{X_{n,n}-X_{n-2m+1,n}}. \tag{33}$$

Fraga Alves[60]研究了 Weiss 类估计和 Hill 估计的组合形式,

$$\hat{\gamma}_n^{WH}(k)=\hat{\gamma}_n^H(k)+\hat{\gamma}_n^W(m),m=[(k+1)/2]. \tag{34}$$

Araújo Santos et al.[61]也提出了几类位置变换不变的重尾指数估计方法,称为 PORT(Peaks Over Random Threshold)方法:

$$\hat{\gamma}_n^{PH}(p,k)=\hat{\gamma}_n^H(k,X_i^*,1\leqslant i\leqslant n), \tag{35}$$

$$\hat{\gamma}_n^{PM}(p,k)=\hat{\gamma}_n^M(k,X_i^*,1\leqslant i\leqslant n), \tag{36}$$

$$\hat{\gamma}_n^{MM}(p,k)=\hat{\gamma}_n^{MM}(k,X_i^*,1\leqslant i\leqslant n), \tag{37}$$

其中 $X_i^*=X_i-X_{[np]+1,n}$,$0<p<1$,$1\leqslant i\leqslant n$.

Fraga Alves[62]提出了位置变换不变的 Hill 类估计,

$$\hat{\gamma}_n^H(k_0,k)=\frac{1}{k_0}\sum_{i=0}^{k_0-1}\log\frac{X_{n-i,n}-X_{n-k,n}}{X_{n-k_0,n}-X_{n-k,n}}, \tag{38}$$

其中 $k_0=k_0(k)\to\infty$,$k=k(n)\to\infty$,$k_0/k\to 0$,$k/n\to 0$,当 $n\to\infty$.

Ling et al.[63-64],① 提出了位置变换不变的矩类估计($\gamma\in R_1$)

$$\hat{\gamma}_n^{Inum}(k_0,k)=\hat{\gamma}_n^+(k_0,k)+\hat{\gamma}_n^-(k_0,k), \tag{39}$$

其中 $\hat{\gamma}_n^+(k_0,k)=\hat{\gamma}_n^H(k_0,k)$,

$$\hat{\gamma}_n^-(k_0,k)=1-\frac{1}{2}\left(1-\frac{(M_n^{(1)}(k_0,k))^2}{M_n^{(2)}(k_0,k)}\right)^{-1},$$

$$M_n^{(j)}(k_0,k)=\frac{1}{k_0}\sum_{i=1}^{k_0-1}\left(\log\frac{X_{n-i,n}-X_{n-k,n}}{X_{n-k_0,n}-X_{n-k,n}}\right)^j,j=1,2.$$

4.3 其它重尾指数估计方法

Gray 和 Schucany[65]提出的 Jackknife 方法主要目的也是降低估计偏差,为此出现了许多基于 Jackknife 方法的重尾指数估计. Falk[66]构造了 Pickands 估计的线性组合,并发现其优于 Pickands 估计. 然而,Martins et al.[67]的模拟结果发现,Hill 估计的线性组合并没有改进 Hill 估计. Peng[68]以 Jackknife 方法得到一类重尾指数的估计方法,Gomes et al.[69]比较研究了基于 Vries 估计和不同水平下 Hill 估计的广义 Jackknife 估计,得到的几类方法在二阶位置参数 ρ 的不同范围内各有千秋,且模拟结果都依赖于样本容量. Gomes et al.[70]研究了基于 Hill 类的 Jackknife 估计,模拟发现对于 Hall 类分布有较好的均方误差(MSE)结果.

实际中有时只能获得样本中较大值的信息,即只有极大次序统计量可以被用来进行统计推断,这种情形下将无法应用上文提到的基于全样本的估计方法. 对此,Davydov et al.[71]提出一类基于对样本分块的估计方法(DPR),只需每块中的第一大次序统计量和第二大次序统计量进行估计. 将样本 $X_1,X_2,\cdots,X_n$ 分成 k_n 组 $V_1,V_2,\cdots,V_{k_n}$,每组包含 $m=m(n)=[n/k_n]$ 个样本,$V_i=\{X_{(i-1)m+1},\cdots,X_{im}\}$,$1\leqslant i\leqslant k_n$,以 $X_{m,1}^{(i)}\geqslant\cdots\geqslant X_{m,m}^{(i)}$ 表示第 m 组的次序统计量,

$$S_{k_n}=\sum_{i=1}^{k_n}\frac{X_{m,2}^{(i)}}{X_{m,1}^{(i)}},$$

$$\hat{\gamma}_{DPR}(k_n)=\frac{k_n-S_{k_n}}{S_{k_n}}.$$

Paulauskas[72]研究发现 DPR 估计在大样本情形有很好的估计性质,然而渐近方差大于 Hill 估计的渐

① Ling C,Peng Z,Nadarajah S. Selecting the Optimal Sample Fraction in Location Invariant Moment-type Estimators[J]. *Theory Probab Appl*,2009,in press.

近方差. Qi[73]以类似于 Davydov et al. 的思路提出一种新的估计方法,

$$\hat{\gamma}_N(k_n,r)=\frac{1}{k_n r}\sum_{i=1}^{k_n}\sum_{j=1}^{r}(\ln X_{m,j}^{(i)}-\ln X_{m,r+1}^{(i)}).$$

Qi 估计与 DPR 估计相比有较小的渐近方差,然而并非渐近无偏,在正则变换二阶条件下其渐近偏差与每块中用于估计的最大次序统计量个数有关,这也将是今后研究的另一热点.

5 展望

自从 Levy,Mandelbrot,Gosset,Zipf 等先驱者们发现较正态分布具有尖峰重尾特征的分布以来,关于重尾分布及重尾指数估计的研究就一直没有停止过. 由于实际中极端值所带来的突出影响,如何根据实际数据推断其尾分布的形状,这是学界非常关注的问题. 学者们根据正则变化理论假设了重尾分布尾部的参数形式,通过估计其中内含的参数,把问题转化为了重尾指数估计.

上个世纪七十年代,学者们在假设尾分布参数形式的前提下,给出了许多重尾指数的估计方法,并且具有非常优良的统计性质. 同时在正则变化条件下从分布尾部的收敛速率出发,也对重尾分布的尾部参数形式做了更为详细的分析和假设. 然而这些估计方法都涉及一个无法回避的问题,即有多少个极值统计量满足所假设的尾分布函数,或者该用多少个极值统计量来估计重尾指数. 上个世纪八十年代以后,学者们对重尾阈值的选取兴致高涨,提出了许多选取方法,并研究了阈值选取对估计结果的影响. 本世纪以来,重尾分布的研究又有新突破,不太关注于重尾阈值的选取,尝试提出对重尾阈值有优良稳健性的最小方差渐近无偏估计,以及改进这些方法更广泛的适用性.

参考文献:

[1] Zipf G K. Human Behavior and the Principle of Least Effort:an Introduction to Human Ecology [M]. Cambridge,Mass,1949.

[2] Mandelbrot B. The Variation of Certain Speculative Prices[J]. *J Business*,1963,**36**:394-419.

[3] Fama E. The Behavior of Stock Market Prices[J]. *J Business*,1965,**38**:34-105.

[4] Brodin E,Kluppelberg C. Extreme Value Theory in Finance[C]//Everitt,B. and Meiinick,E.(Eds.) Encyclopedia of Quantitive Risk Assessment,2007.

[5] Bollerslev T,Todorov V. Tail,Fears and Risk Premia[J]. *J Finance*,2011,**66**:2165-2211.

[6] Tessier Y,Lovejoy S,Hubert P,*et al*. Multifractal Analysis and Modeling of Rainfall and River Flows and Scaling,Causal Transfer Functions[J]. *J Geophys Res*,1996,**101**:427-440.

[7] Richard L. Smith. Statistics of Extremes,with Applications in Enviroment,Insurance and Finance[C]//Extreme Values in Finance,Telecommunications and the Environment,edited by B. Finkenstadt and H. Rootzen,London:Chapman and Hall/CRC Press,2003:1-78.

[8] Allan R,Soden B. Atmospheric Warming and the Amplification of Precipitation Extremes[J]. *Science*,2008,**321**(5895):1481-1484.

[9] Barabasi A L. The Origin of Bursts and Heavy Tails in Human Dynamics[J]. *Nature*,2005,**435**(7039):207-211.

[10] Vazquez A. Exact Results for the Barabasi Model of Human Dynamics[J]. *Physical Review Letters*,2005,**95**:248701(1)-248701(4).

[11] Vazquez A,Oliveira J G,Dezsö Z,*et al*. Modeling Bursts and Heavy Tails in Human Dynamics[J]. *Physical Review E*,2006,**73**(3):036127(1)-036127(19).

[12] Werner T,Upper C. Time Variation in the Tail Behaviour of Bund Futures Returns[J]. *Journal of Futures Markets*,2004,**24**(4):387-398.

[13] Embreehts P,Kluppelberg C,Mikoseh C. Modeling Extremal Events for Insurance and Finance[M]. Berlin:Springer,1997.

[14] 邢红卫. 重尾现象、重尾分布与重尾指数估计[D]. 太原:山西大学,2010.

[15] Pickands J. Statistical Inference Using Extreme Order Statistics[J]. *Annals of Statistics*,1975,**3**:119-131.

[16] Hill B. A Simple General Approach to Inference About the Tail of a Distribution[J]. *Annals of Statistics*,1975,**3**:1163-

1174.

[17] Dekkers A, Einmahl J, De Haan L. A Moment Estimator for the Index of an Extreme-value Distribution [J]. *Annals of Statistics*, 1989, **17**: 1833-1855.

[18] Gomes M I, De Haan, Peng L. Semi-parametric Estimation of the Second Order Parameter Asymptotic and Finite Sample Behaviour[J]. *Extremes*, 2002, **5**(4): 387-414.

[19] Fraga Alves M I, Gomes M I, De Haan. A new Class of Semi-parametric Estimators of the Second Order Parameter[J]. *Portugaliae Mathematic*, 2003, **60**(1): 193-213.

[20] Gomes M I, Caeiro F, Figueiredo F. Bias Reduction of a Tail Index Estimator Trough an External Estimation of the Second Order Parameter[J]. *Statistics*, 2004, **38**(6): 497-510.

[21] Caeiro F, Gomes M I. Minimum-variance Reduced-bias Tail Index and Quantile Estimation[J]. *Revstat*, 2008, **6**(1): 1-20.

[22] Fama F, Roll R. Some Properties of Symmetric Stable Distributions[J]. *Journal of the American Statistical Association*, 1968, **63**: 817-836.

[23] Press S. Estimation in Univariate and Multivariate Stable Distributions[J]. *Journal of the American Statistical Association*, 1972, **67**(340): 842-847.

[24] Zolotarev V. One-dimensional Stable Distributions[J]. *Translations of Mathematical Monographs*, 1986, **65**.

[25] de Haan. Extreme Value Statistics[C]//Galambos, J., Lechner, J., Simiu, E. (eds.) Extreme Value Theory and Applications, Kluwer Academic Publications, Dordrecht, 1994: 93-122.

[26] de Haan, Resnick S. A Simple Asymptotic Estimate for the Index of a Stable Distribution[J]. *Journal of the Royal Statistical Society, Series B*, 1980, **42**: 83-87.

[27] Davison A C. Modelling Excesses Over High Thresholds[C]//Tiago de Oliveira, J. (Ed), Statistical Extremes and Applications. Reidel, Dordrecht, 1984: 461-482.

[28] Csörgö S, Deheuvels P, Mason D. Kernel Estimates of the Tail Index of a Distribution[J]. *Annals of Statistics*, 1985, **13**(3): 1050-1077.

[29] De Haan. Estimation of the Minimum of a Function Using Order Statistics[J]. *Journal of the American Statistical Association*, 1981, **76**: 467-469.

[30] Bacro J N, Brito M. Strong Limiting Behaviour of a Simple Tail Pareto-index Estimator[J]. *Statistics and Decision*, 1993, **3**: 133-134.

[31] Smith R. Estimating Tails of Probability Distributions[J]. *Annals of Statistics*, 1987, **15**: 1174-1207.

[32] Smith R. Extreme Value Analysis of Environmental Time Series: an Application to Trend Detection in Ground-level ozone[J]. *Statistical Science*, 1989, **4**: 367-393.

[33] Deheurels P, De Haan L, Peng L, *et al*. Comparison of Extreme Value Index Estimators[J]. NEPTUNE. T400: EUR-09.

[34] Kratz M, Resnick S. The qq-estimator and Heavy Tails[J]. *Stochastic Modelsm*, 1996, **12**(4): 699-724.

[35] Beirlant J, Vynckier P, Teugels J L. Tail Index Estimation, Pareto Quantile Plots, and Regression Diagnostics[J]. *Journal of the American Statistical Association*, 1996, **436**: 1659-1667.

[36] Fraga Alves M I, Gomes M I, De Haan L, *et al*. Mixed Moment Estimators and Location Invariant Alternatives[J]. *Extremes*, 2009, **12**: 149-185.

[37] Resnick S, Starica C. Smoothing the Hill Estimator[J]. *Advances in Applied Probability*, 1997, **29**: 271-293.

[38] Drees H, De Haan L, Resnick S. How to Make a Hill Plot[J]. *Annals of Statistics*, 2000, **28**: 254-274.

[39] Sousa B. A Contribution to the Estimation of the Tail Index of Heavy-tailed Distributions[D]. The University of Michigan, 2002.

[40] Hall P. Using the Bootstrap to Estimate Mean Square Error and Select Smoothing Parameters in Non-parametric Problems[J]. *Journal of Multivariate Analysis*, 1990, **32**: 177-203.

[41] Danielsson J. Using a Bootstrap Method Choose the Sample Fraction in Tail Index Estimation[J]. *Journal of Multivariate Analysis*, 2001, **76**: 226-248.

[42] **刘维奇, 赫英迪, 邢红卫. 选择重尾阈值 K 的 Bootstrap 方法**[J]. **山西大学学报:自然科学版**, 2010, **4**: 508-512.

[43] **刘维奇, 邢红卫. 重尾指数估计中阈值 K 的简便优化估计**[J]. **系统工程理论与实践**, 2010, **30**(8): 1465-1470.

[44] de Haan, Peng L. Comparison of Tail Index Estimators[J]. *Statistica Neerlandica*, 1998, **52**: 60-70.

[45] Beirlant J, Dierckx G, Goegebeur Y, *et al*. Tail Index Estimation and an Exponential Regression Model[J]. *Extremes*,

1999,**2**:177-200.

[46] Feuerverger A,Hall P. Estimating a Tail Exponent by Modeling Departure from a Pareto Distribution[J]. *Annals of Statistics*,1999,**27**:760-781.

[47] Beirlant J,Dierckx G,Guillou A,*et al*. On Exponential Representations of Log-spacings of Extreme Order Statistics[J]. *Extremes*,2002,**5**(2):157-180.

[48] Caeiro F,Gomes M I. A Class of Asymptotically Unbiased Semi-parametric Estimators of the Tail Index[J]. *Test*,2002,**11**(2):345-364.

[49] Caeiro F,Gomes M I,Pestana D D. Direct Reduction of Bias of the Classical Hill Estimator[J]. *Revstat*,2005,**3**(2):111-136.

[50] Gomes M I,Pestana D D. A Simple Second Order Reduced-bias Tail Index Estimator[J]. *Journal of Statistical Computation and Simulation*,2007,**77**(6):487-504.

[51] Gomes M I,Martins M J,Neves M M. Improving Second Order Reduced-bias Tail Index Estimation [J]. *Revstat*,2007,**5**(2):177-207.

[52] Gomes M I,De Haan,Henriques R L. Tail Index Estimation for Heavy-tailed Models:Accommodation of Bias in Weighted Log-excesses[J]. *Journal of the Royal Statistical Society:Series B*,2008,**70**(1):31-52.

[53] Gomes M I,Martins M J. "Asymptotically Unbiased" Estimators of the Tail Index Based on External Estimation of the Second Order Parameter [J]. *Extremes*,2002,**5**(1):5-31.

[54] Beirlant J,Dierckx G,Guillou A. Estimation of the Extreme Value Index and Generalized Quantile Plots[J]. *Bernoulli*,2005,**11**(6):949-970.

[55] Gomes M I,Fraga Alves M I,Araújo Santos P. PORT Hill and Moment Estimators for Heavy-tailed Models[J]. *Communications in Statistics Simulation and Computation*,2008,**37**:1281-1306.

[56] Falk M. Some Best Parameter Estimates for Distributions with Finite Endpoint[J]. *Statistics*,1995,**27**:115-125.

[57] Müller S,Hüsler J. Iterative Estimation of the Extreme Value Index[J]. *Methodology and Computing in Applied Probability*,2005,**2**:139-145.

[58] Weiss L. Asymptotic Inference About a Density Function at the end of Its Range[J]. *Nav Res Logist Q*,1971,**1**:111-114.

[59] Peng Z,Nadarajah S. The Pickands' Estimator of the Negative Extreme Value Index[J]. *Acta Scientiarum Naturalium Universitatis Pekinensis*,2001,**37**:12-19.

[60] Fraga Alves M I. Weiss-Hill Estimator[J]. *Test*,2001,**10**:203-224.

[61] Araújo Santos P,Fraga Alves M I,Gomes M I. Peaks Over Random Threshold Methodology for Tail Index and Quantile Estimation[J]. *Revstat*,2006,**4**:227-247.

[62] Fraga Alves M I. A Location Invariant Hill-type Estimator[J]. *Extremes*,2001,**4**:199-217.

[63] Ling C,Peng Z,Nadarajah S. A Location Invariant Moment-type Estimator I[J]. *Theory of Probability and Mathematical Statistics*,2007,**76**:23-31.

[64] Ling C,Peng Z,Nadarajah S. A Location Invariant Moment-type Estimator II[J]. *Theory of Probability and Mathematical Statistics*,2007,**77**:177-189.

[65] Gray H L,Schucany W R. The Generalized Jackknife Statistic[M]. New York:Marcel Dekker,1972.

[66] Falk M. Efficiency of Convex Combinations of Pickands Estimator of the Extreme Value Index[J]. *Nonparametric Statistics*,1994,**4**:133-147.

[67] Martins M J,Gomes M I,Neves M. Some Results on the Behavior of Hill's Estimator[J]. *Journal of Statistical Computation and Simulation*,1999,**63**:283-297.

[68] Peng L. Asymptotically Unbiased Estimators for the Extreme-value Index[J]. *Statistics and Probability Letters*,1998,**38**(2):107-115.

[69] Gomes M I,Martins M J,Neves M. Alternatives to a Semi-parametric Estimator of Parameters of Rare Events-the Jackknife Methodology[J]. *Extremes*,2000,**3**(3):207-299.

[70] Gomes M I,Pereira H,Miranda M C. Revisiting the Role of the Jackknife Methodology in the Estimation of a Positive Tail Index[J]. *Communications in Statistics Theory and Methods*,2005,**34**:319-335.

[71] Davydov Y,Paulauskas V,Rackauskas A. More on P-stable Convex Sets in Banach Spaces[J]. *Journal of Theoretical*

Probability,2000,**13**:39-64.

[72] Paulauskas V. A new Estimator for a Tail Index[J]. *Acta Applicandae Mathematicae*,2003,**79**:55-67.

[73] Qi Y. On the Tail Index of a Heavy Tailed Distribution[J]. *Annals of the Institute of Statistical Mathematics*,2010,**62**:277-298.

Advanced in Heavy-tailed Distribution Tail Index Estimation

LIU Wei-qi,XING Hong-wei

(*School of Management*,*Shanxi University*,*Taiyuan* 030006,*China*)

Abstract:Many studies show that the data in application fields,such as meteorology,hydrology,environment,telecommunications,insurance and finance,does not meet the normal distribution assumption,but illustrates spikes and heavy tail characteristics. In the last three decades,the research on heavy-tailed distributions and tail index estimation has been developed rapidly. We review the discovery process of heavy-tailed distributions challenging the normal distribution assumption,and introduce several different definitions of the heavy-tailed distribution,extreme value theory and regular variation conditions;moreover,summarize different types of tail index estimators and methods of selecting the heavy tail threshold by development periods and method types,and discuss the characteristics and shortages of these methods,finally,in term of the deficiencies and limitations of current estimators,how to deepen and improve the problem are reviewed.

Key words:extreme value theory;heavy tail index estimator;heavy tail threshold;reduce-biased estimator