应用概率统计　第二十九卷
第二期　2013年4月

Chinese Journal of Applied Probability
and Statistics Vol.29 No.2 Apr. 2013

广义Pareto分布参数的最小二乘估计 *

陈海清[1,2]　　程维虎[2]

([1]贵州师范学院数学与算计机科学学院, 贵阳, 550018;　[2]北京工业大学应用数理学院, 北京, 100124)

摘　要

传统的广义Pareto分布(Generalized Pareto Distribution, 简记GPD)的参数估计一般受分布形状参数的约束. 如: 矩估计(the Method of Moments, 简记MOM), 概率加权矩估计(the Probability Weighted Moments, 简记PWM), L矩估计(简记LM), 极大似然估计(Maximum Likelihood Estimation, 简记MLE)等. 本文利用GDP可转化成指数分布的事实及指数分布参数估计的结果, 利用最小二乘(the Least Squares, 简记LS)法, 得到了两参数和三参数GPD的参数估计; 给出了估计量具有渐近正态性的结果. 估计方法不受分布形状参数的限制. 模拟显示: 本文提出的估计在某些常用条件下优于GPD的其他参数估计, 如MOM, PWM, LM, 以及基于分位数估计(the Elemental Percentile Method, 简记EPM)等.

关 键 词: 广义Pareto分布, 矩估计, 概率加权矩估计, 最小二乘估计, L矩估计, 基于分位数估计.

学 科 分 类 号: O212.1.

§1.　引　　言

自Pickands (1975)首次提出GPD之后, 该分布就广泛地运用于超量分布(Peaks over Threshold, 简记POT)模型, 从而GPD的参数估计在实际问题中越来越重要. Smith (1984)讨论了GPD参数MLE的一些性质, 如: 在分布的形状参数$k < 0.5$时, MLE具有相合性、渐近正态性、渐近有效性等; 但估计的数值解有时难以得到, 且在小样本情形下估计效果差; 在$k > 1$时, 估计不存在等. Hosking和Wallis (1987)讨论了GPD参数的MOM和PWM, MOM和PWM具有计算简便的优点, 在一定条件下估计量具有渐近正态性, 但MOM受样本异常大观察值的影响较大, 估计不稳健. 当$k < -0.5$时, MOM不存在. 和MOM相比, PWM更适合于GPD是厚尾的情形, 当$k < -1$时, PWM估计也不存在. Hosking (1990)讨论了LM, LM不仅适合于三参数GPD的参数估计问题, 且更适合于GPD是厚尾的情形, 但LM也只适应$k > -1$的情形. 由于MOM、PWM、MLE、LM等在GPD的参数估计中都受分布形状参数k的限制, 而且在$k > 0$时, 这些估计有时无意义(估计值与真实值严重偏离). 基于此, 1997年Castillo和Hadi[5]给出了GPD参数的EPM, 该估计方法不受分布形状参数k的限制, 比较适合k, 没有先验信息情形下GPD的参数估计问题, 避免了在$k > 0$时,

*2010年北京市教委科技面上项目(KM201010005006)和贵州师范学院科研基金(12YB021)资助.
本文2011年3月11日收到, 2012年11月2日收到修改稿.

估计值与真实值明显偏离的情况, 但EPM涉及的数值计算非常繁琐, 且未充分利用样本的信息.

GPD参数的LSE最早由Moharram[6]于1993年提出, 该估计受GPD形状参数k的影响较大, 特别GPD是厚尾的情形, 估计效果很差. 然而在实际问题中常用GPD拟合厚尾数据, 故GPD参数的LSE在实际问题中很少被采用. 本文利用GDP可转化成指数分布的事实及指数分布参数估计的结果, 利用最小二乘法, 得到了两参数和三参数GPD的参数估计; 给出了估计量具有渐近正态性的结果. 估计方法不受分布形状参数的限制, 且避免了在$k>0$时, 估计值与真实值明显偏离的情况, 从而克服了传统的GPD参数估计的缺陷. 此外, 本文所提出的估计和EPM相比, 不仅充分利用了样本的信息, 而且计算过程简便, 同时也适应三参数GPD的参数估计问题.

§2. GPD及其参数估计

2.1 分布的定义

设随机变量X分布函数如下

$$F(x)=\begin{cases}1-\Big[1-\dfrac{k}{b}(x-a)\Big]^{1/k}, & k\neq 0;\\ 1-\mathrm{e}^{-(x-a)/b}, & k=0,\end{cases}\tag{2.1}$$

当$k\leq 0$时, $a\leq x<\infty$; 当$k>0$时, $a\leq x\leq a+b/k$. 其中$a\geq 0$为分布的位置参数, $b>0$为分布的刻度参数, k是分布的形状参数. 则称X服从三参数(a,b,k)的广义Pareto分布, 概率密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b}\Big[1-\dfrac{k}{b}(x-a)\Big]^{1/k-1}, & k\neq 0;\\ \dfrac{1}{b}\mathrm{e}^{-(x-a)/b}, & k=0.\end{cases}\tag{2.2}$$

由(2.1)式, 可导出分布的分位数公式

$$x(F)=\begin{cases}a+\dfrac{b}{k}[1-(1-F)^k], & k\neq 0;\\ a+b[-\ln(1-F)], & k=0.\end{cases}\tag{2.3}$$

2.2 参数估计

I. 两参数情形

在(2.1)中, 当$k\neq 0$, $a=0$时, 称$F(x)$为两参数GPD, 分布函数为

$$F(x)=1-\Big(1-\frac{k}{b}x\Big)^{1/k}.\tag{2.4}$$

设$X_1,\ldots,X_n$为两参数GPD总体的一个简单样本, 次序统计量为$X_{1:n},\ldots,X_{n:n}$. 样本与次序样本的具体观测值分别记成$x_1,\ldots,x_n$和$x_{1:n},\ldots,x_{n:n}$. 令$\theta=k/b$, $Y=-(1/k)\ln(1-\theta X)$, 记Y的分布函数和概率密度函数分别为$F_Y(y)$和$f_Y(y)$, 当$y>0$时, 有

$$F_Y(y)=\mathsf{P}(Y<y)=\mathsf{P}\Big[-\frac{1}{k}\ln(1-\theta X)<y\Big]=\mathsf{P}[-\ln(1-F(X))<y]=1-\mathrm{e}^{-y}. \quad (2.5)$$

易见Y服从标准指数分布Exp(1).

记$Y_{[n/2]:n}=-(1/k)\ln(1-\theta X_{[n/2]:n})$, 由标准指数分布样本分位数的渐近性质(参见文献[7], P.44), 有

$$\sqrt{n}(Y_{[n/2]:n}-y_{0.5})\xrightarrow{L}N\Big(0,\frac{1}{4[f_Y(y_{0.5})]^2}\Big),\qquad n\to\infty,$$

其中$y_{0.5}$表示总体Y的中位数, 满足$F_Y(y_{0.5})=0.5$. 易导出$y_{0.5}=\ln 2$, $1/\{4[f_Y(y_{0.5})]^2\}=1$, 故

$$\sqrt{n}(Y_{[n/2]:n}-\ln 2)\xrightarrow{L}N(0,1),\qquad n\to\infty.$$

所以$Y_{[n/2]:n}\xrightarrow{\mathsf{P}}\ln 2$, $n\to\infty$. 令

$$Y^*=\frac{\ln 2}{Y_{[n/2]:n}}Y=(\ln 2)\frac{\ln(1-\theta X)}{\ln(1-\theta X_{[n/2]:n})},$$

由Slutsky定理(参见文献[7], P.38), 知

$$Y^*\xrightarrow{L}\mathrm{Exp}(1),\qquad n\to\infty. \quad (2.6)$$

记$Y^*_{i:n}=(\ln 2)[\ln(1-\theta X_{i:n})/\ln(1-\theta X_{[n/2]:n})]$, 由(2.6)式及W. Glivenko定理, 有$Y^*_{i:n}\approx-\ln(1-p_{i:n})\triangleq c_{i,n}$, 其中$p_i=F_n(x_{i:n})$, $F_n(\cdot)$为样本的经验分布函数. 通常用$(i-0.375)/(n+0.25)$替代p_i, 使估计效果变好. 利用最小二乘法, 得到θ的一个估计$\widehat{\theta}_n(x_{1:n},\ldots,x_{n:n})=\min_\theta\sum\limits_{i=1}^n(Y^*_{i:n}-c_{i,n})^2$, 即

$$\widehat{\theta}_n(x_{1:n},\ldots,x_{n:n})=\min_\theta\sum_{i=1}^n\Big[(\ln 2)\frac{\ln(1-\theta x_{i:n})}{\ln(1-\theta x_{[n/2]:n})}+\ln(1-p_i)\Big]^2. \quad (2.7)$$

把$\widehat{\theta}_n$代入到(2.4)式, 得$F(x)\approx1-(1-\widehat{\theta}_n x)^{1/k}$, 进一步利用最小二乘法, 得到$k$的估计

$$\widehat{k}_n(x_{1:n},\ldots,x_{n:n})=\min_k\sum_{i=1}^n[F(x_{i:n})-F_n(x_{i:n})]^2,$$

即

$$\widehat{k}_n(x_{1:n},\ldots,x_{n:n})=\min_k\sum_{i=1}^n[1-(1-\widehat{\theta}_n x_{i:n})^{1/k}-p_i]. \quad (2.8)$$

再由$\theta=k/b$, 得到b的估计

$$\widehat{b}_n(x_{1:n},\ldots,x_{n:n})=\frac{\widehat{k}_n(x_{1:n},\ldots,x_{n:n})}{\widehat{\theta}_n(x_{1:n},\ldots,x_{n:n})}. \tag{2.9}$$

II. 三参数情形

在(2.1)式中, 当$k\neq 0$, $a\neq 0$时, 称$F(x)$为三参数GPD. 由(2.3)式, 知

$$x_{1:n}=a+\frac{b}{k}\big[1-(1-F(x_{1:n}))^k\big].$$

对$(1-F(x_{1:n}))^k$一阶泰勒展开, 并注意到$F(x_{1:n})=o_p(n^{-1})$, 则有

$$a=x_{1:n}-\frac{b}{k}\big[1-(1-F(x_{1:n}))^k\big]=x_{1:n}-bF(x_{1:n})+o_p(n^{-1}), \tag{2.10}$$

对$[1-(k/b)(x_{i:n}-a)]^{1/k}$一阶泰勒展开, 有

$$F(x_{i:n})=1-\Big[1-\frac{k}{b}(x_{i:n}-a)\Big]^{1/k}=\frac{1}{b}(x_{i:n}-a)+o(x_{i:n}-a),$$

所以, 当i较小时, $F_n(x_{i:n})\approx(1/b)(x_{i:n}-a)$. 取$i=1,2$时, 得到$b$的一个初步估计

$$\widetilde{b}_n=\frac{x_{2:n}-x_{1:n}}{F_n(x_{2:n})-F_n(x_{1:n})}. \tag{2.11}$$

将(2.11)代入(2.10), 得到a的估计

$$\widehat{a}_n(x_{1:n},\ldots,x_{n:n})=x_{1:n}-\widetilde{b}_nF_n(x_{1:n})=x_{1:n}-\frac{p_1}{p_2-p_1}(x_{2:n}-x_{1:n}). \tag{2.12}$$

同样, 令$\theta=k/b$, 利用最小二乘法, 类似(2.7)式, 得到θ的一个估计

$$\widehat{\theta}_n(x_{1:n},\ldots,x_{n:n})=\min_{\theta}m(\theta;x_{1:n},\ldots,x_{n:n}), \tag{2.13}$$

其中

$$m(\theta;x_{1:n},\ldots,x_{n:n})=\sum_{i=1}^{n}\Big[(\ln 2)\frac{\ln[1-\theta(x_{i:n}-\widehat{a}_n)]}{\ln[1-\theta(x_{[n/2]:n}-\widehat{a}_n)]}+\ln(1-p_i)\Big]^2. \tag{2.14}$$

再利用最小二乘法类似(2.8)式, 得到k的一个估计

$$\widehat{k}_n(x_{1:n},\ldots,x_{n:n})=\min_{k}\ell(k;x_{1:n},\ldots,x_{n:n}), \tag{2.15}$$

其中

$$\ell(k;x_{1:n},\ldots,x_{n:n})=\sum_{i=1}^{n}[1-(1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n))^{1/k}-p_i]^2. \tag{2.16}$$

由$\theta=k/b$, 得到b的估计

$$\widehat{b}_n(x_{1:n},\ldots,x_{n:n})=\frac{\widehat{k}_n(x_{1:n},\ldots,x_{n:n})}{\widehat{\theta}_n(x_{1:n},\ldots,x_{n:n})}. \tag{2.17}$$

§3. 估计量性质

3.1 基本引理

引理 3.1 若$X_1,\ldots,X_n$是抽自总体$F(x,\theta)$的简单样本, 总体有概率密度函数$f(x,\theta)$, θ为总体的分布参数. 任取m个常数$\lambda_1,\ldots,\lambda_m$满足$0<\lambda_1<\lambda_2<\cdots<\lambda_m<1$. 记$X_{1:n}$, $\ldots,X_{n:n}$为次序样本, $n_i=[n\lambda_i]+1$, $i=1,2,\ldots,m$, 其中$[n\lambda_i]$表示不超过$n\lambda_i$的最大整数, 则有

$$\sqrt{n}(X_{n_1:n}-\xi_1,X_{n_2:n}-\xi_2,\ldots,X_{n_m:n}-\xi_m)'\xrightarrow{L}N_m(0,W),$$

其中ξ_i为总体的λ_i分位点, 即ξ_i满足$\lambda_i=F(\xi_i,\theta)$, $i=1,2,\ldots,m$; 0为m维零向量, $W=(a_{ij})_{m\times m}$为m阶对称阵, 满足

$$a_{ij}=a_{ji}=\frac{\lambda_i(1-\lambda_j)}{[f(\xi_i,\theta)f(\xi_j,\theta)]},\qquad 1\le i\le j\le m.$$

注记 1 引理3.1参见文献[8], 定理1.5.

对GPD总体, 记$F(x)$与$f(x)$分别为总体的分布函数与概率密度函数, ξ_i为总体的p_i分位点, 即$p_i=F(\xi_i)<1$, $i=1,2,\ldots,n$. 由(2.2)和(2.3)式, 得

$$\xi_i=a+\frac{b}{k}[1-(1-p_i)^k],\qquad f(\xi_i)=\frac{1}{b}\Big[1-\frac{k}{b}(\xi_i-a)\Big]^{1/k-1}=\frac{1}{b}(1-p_i)^{1-k},$$

另记$\boldsymbol{X}=(X_{1:n},\ldots,X_{n:n})'$, $\boldsymbol{\xi}=(\xi_1,\ldots,\xi_n)'$, 由引理3.1, 知

$$\sqrt{n}(\boldsymbol{X}-\boldsymbol{\xi})\xrightarrow{L}N(0,\Sigma),\qquad n\to\infty,\tag{3.1}$$

其中$\Sigma=(\sigma_{ij})_{n\times n}$, $\sigma_{ij}=\sigma_{ji}=b^2p_i(1-p_i)^{k-1}(1-p_j)^k$, $1\le i\le j\le n$.

3.2 估计量的渐近分布

由(2.13), (2.15)式, 可通过牛顿迭代法求解非线性方程组$\partial m/\partial\theta=0$和$\partial l/\partial k=0$而获得$\widehat{\theta}_n$和$\widehat{k}_n$的数值解.

定理 3.1 方程组$\partial m/\partial\theta=0$和$\partial l/\partial k=0$在$n\to\infty$时以概率1有解, 且$\widehat{\theta}_n$和$\widehat{k}_n$分别是$\theta$和$k$的相合估计.

证明: 首先, 记θ,k的真值分别为θ_0和k_0, 由于$F(x_{i:n})=p_i+O_p(n^{-1/2})$, $i=1,2,\ldots$, n及$F(x_{i:n})=1-[1-\theta_0(x_{i;n}-a)]^{1/k_0}$, 所以

$$\ln[1-\theta_0(x_{i:n}-a)]=k_0\ln(1-F(x_{i:n}))=k_0\ln(1-p_i)+O_p(n^{-1/2}).\tag{3.2}$$

又由(2.10)和(2.12)式, 知

$$\widehat{a}_n-a=(b-\widetilde{b}_n)F(x_{1:n})+o_p(1/n)=O_p(n^{-1}),\tag{3.3}$$

由(3.2)和(3.3)式, 得$\ln[1-\theta_0(x_{i:n}-\widehat{a}_n)]=k_0\ln(1-p_i)+O_p(n^{-1/2})$, 所以

$$\frac{\ln[1-\theta_0(x_{i:n}-\widehat{a}_n)]}{\ln[1-\theta_0(x_{[n/2]:n}-\widehat{a}_n)]}=-\frac{\ln(1-p_i)}{\ln 2}+O_p(n^{-1/2}).$$

故, 对充分小的$\delta>0$, 有

$$\begin{aligned}
&m(\theta_0+\delta)-m(\theta_0)\\
=\ &\sum_{i=1}^{n}\Big\{\Big[(\ln 2)\frac{\ln[1-(\theta_0+\delta)(x_{i:n}-\widehat{a}_n)]}{\ln[1-(\theta_0+\delta)(x_{[n/2]:n}-\widehat{a}_n)]}+\ln(1-p_i)\Big]^2\\
&-\Big[(\ln 2)\frac{\ln[1-\theta_0(x_{i:n}-\widehat{a}_n)]}{\ln[1-\theta_0(x_{[n/2]:n}-\widehat{a}_n)]}+\ln(1-p_i)\Big]^2\Big\}\\
=\ &(\ln 2)^2\sum_{i=1}^{n}\Big\{\Big[\frac{\ln[1-(\theta_0+\delta)(x_{i:n}-\widehat{a}_n)]}{\ln[1-(\theta_0+\delta)(x_{[n/2]:n}-\widehat{a}_n)]}-\frac{\ln[1-\theta_0(x_{i:n}-\widehat{a}_n)]}{\ln[1-\theta_0(x_{[n/2]:n}-\widehat{a}_n)]}\Big]\\
&\times\Big[\frac{2\ln(1-p_i)}{\ln 2}+\frac{\ln[1-(\theta_0+\delta)(x_{i:n}-\widehat{a}_n)]}{\ln[1-(\theta_0+\delta)(x_{[n/2]:n}-\widehat{a}_n)]}+\frac{\ln[1-\theta_0(x_{i:n}-\widehat{a}_n)]}{\ln[1-\theta_0(x_{[n/2]:n}-\widehat{a}_n)]}\Big]\Big\}\\
=\ &(\ln 2)^2\sum_{i=1}^{n}\Big\{\Big[\frac{\ln[1-(\theta_0+\delta)(x_{i:n}-\widehat{a}_n)]}{\ln[1-(\theta_0+\delta)(x_{[n/2]:n}-\widehat{a}_n)]}-\frac{\ln[1-\theta_0(x_{i:n}-\widehat{a}_n)]}{\ln[1-\theta_0(x_{[n/2]:n}-\widehat{a}_n)]}\Big]\\
&\times\Big[\frac{\ln[1-(\theta_0+\delta)(x_{i:n}-\widehat{a}_n)]}{\ln[1-(\theta_0+\delta)(x_{[n/2]:n}-\widehat{a}_n)]}-\frac{\ln[1-\theta_0(x_{i:n}-\widehat{a}_n)]}{\ln[1-\theta_0(x_{[n/2]:n}-\widehat{a}_n)]}+O_p(n^{-1/2})\Big]\Big\}.
\end{aligned}\tag{3.4}$$

先固定δ, 让$n\to\infty$, 则上式最后一个求和符号中的每一项都为正, 故$m(\theta_0+\delta)-m(\theta_0)>0$.

同理可得: $m(\theta_0-\delta)-m(\theta_0)>0$, 由于$m(\theta)$在$[\theta_0-\delta,\theta_0+\delta]$上连续, 因而必有局部最小值点$\widehat{\theta}_n$. 由$m(\theta)$可微, 故$\partial m/\partial\theta|_{\widehat{\theta}_n}=0$. 从而, 当$n\to\infty$时, 方程$\partial m/\partial\theta=0$以概率1有解$\widehat{\theta}_n$, 且$|\widehat{\theta}_n-\theta|<\delta$. 由$\delta$的任意性知, $\widehat{\theta}_n$是θ的相合估计.

其次, 由$F(x_{i:n})=p_i+O_p(n^{-1/2})$, 得

$$1-[1-\theta_0(x_{i:n}-a)]^{1/k_0}=p_i+O_p(n^{-1/2}).$$

又$\widehat{\theta}_n\xrightarrow{\mathsf{P}}\theta$, $\widehat{a}_n\xrightarrow{\mathsf{P}}a$, 有

$$1-[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/k_0}=p_i+\varepsilon_i,$$

其中$n\to\infty$时, $\varepsilon_i\xrightarrow{\mathsf{P}}0$. 对上述充分小的$\delta>0$, 有

$$\begin{aligned}
&\ell(k_0+\delta)-\ell(k_0)\\
=\ &\sum_{i=1}^{m}\Big\{\big[1-[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/(k_0+\delta)}-p_i\big]^2-\big[1-[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/k_0}-p_i\big]^2\Big\}\\
=\ &\sum_{i=1}^{m}\Big\{\big[[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/k_0}-[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/(k_0+\delta)}\big]\\
&\times\big[2(1-p_i)-[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/k_0}-[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/(k_0+\delta)}\big]\Big\}\\
=\ &\sum_{i=1}^{m}\Big\{\big[[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/k_0}-[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/(k_0+\delta)}\big]\\
&\times\big[[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/k_0}-[1-\widehat{\theta}_n(x_{i:n}-\widehat{a}_n)]^{1/(k_0+\delta)}+2\varepsilon_i\big]\Big\}.
\end{aligned}\tag{3.5}$$

先固定δ, 让$n\to\infty$, 则上式最后一个求和符号中每一项都为正, 故$\ell(k_0+\delta)-\ell(k_0)>0$.

同理可得: $\ell(k_0-\delta)-\ell(k_0)>0$. 由于$\ell(k)$在$[k_0-\delta,k_0+\delta]$上连续, 因而必有局部最小值点$\widehat{k}_n$. 由$\ell(k)$可微, 知$\partial\ell/\partial k|_{\widehat{k}_n}=0$. 从而$n\to\infty$时, 方程$\partial\ell/\partial k=0$以概率1有解$\widehat{k}_n$, 且$|\widehat{k}_n-k|<\delta$. 再由$\delta$的任意性知, $\widehat{k}_n$是k的相合估计. 证毕. □

定理 3.2 记$\widehat{\theta}_n(x_{1:n},\ldots,x_{n:n})$和$\widehat{k}_n(x_{1:n},\ldots,x_{n:n})$分别为定理3.1中方程$\partial m/\partial\theta=0$和$\partial\ell/\partial k=0$的解, 则当$x_{i:n}=\xi_i$, $i=1,2,\ldots,n$时, 有$\widehat{\theta}_n(\xi_1,\ldots,\xi_n)=\theta_0+o(n^{-1/2})$和$\widehat{k}_n(\xi_1,\ldots,\xi_n)=k_0+o(n^{-1/2})$.

证明: 首先, 由$p_i=F(\xi_i)$, 得

$$p_i=1-[1-\theta_0(\xi_i-a)]^{1/k_0}, \tag{3.6}$$

$$\ln[1-\theta_0(\xi_i-a)]=k_0\ln(1-p_i). \tag{3.7}$$

又, 类似(3.3)式, 得

$$\widehat{a}_n(\xi_1,\ldots,\xi_n)-a=O(n^{-1}). \tag{3.8}$$

由(3.7)与(3.8)式, 得

$$\ln[1-\theta_0(\xi_i-\widehat{a}_n(\xi_1,\ldots,\xi_n))]=k_0\ln(1-p_i)+O(n^{-1}),$$

故

$$\frac{\ln[1-\theta_0(\xi_i-\widehat{a}_n(\xi_1,\ldots,\xi_n))]}{\ln[1-\theta_0(\xi_{[n/2]}-\widehat{a}_n(\xi_1,\ldots,\xi_n))]}=-\frac{\ln(1-p_i)}{\ln 2}+O(n^{-1}).$$

取$\delta=n^{-\alpha}$, $0.5<\alpha<1$, 则

$$\begin{aligned}\Delta_i &\triangleq \frac{\ln[1-(\theta_0+\delta)(\xi_i-\widehat{a}_n(\xi_1,\ldots,\xi_n))]}{\ln[1-(\theta_0+\delta)(\xi_{[n/2]}-\widehat{a}_n(\xi_1,\ldots,\xi_n))]}-\frac{\ln[1-\theta_0(\xi_i-\widehat{a}_n(\xi_1,\ldots,\xi_n))]}{\ln[1-\theta_0(\xi_{[n/2]}-\widehat{a}_n(\xi_1,\ldots,\xi_n))]}\\ &= O(n^{-\alpha}).\end{aligned} \tag{3.9}$$

类似(3.4)式, 有

$$m(\theta_0+\delta;\xi_1,\ldots,\xi_n)-m(\theta_0;\xi_1,\ldots,\xi_n)=(\ln 2)^2\sum_{i=1}^{n}\Delta_i[\Delta_i+O(n^{-1})].$$

由上式与(3.9)式知: 当n充分大时, Δ_i与$\Delta_i+O(n^{-1})$同号, 故

$$m(\theta_0+\delta;\xi_1,\ldots,\xi_n)-m(\theta_0;\xi_1,\ldots,\xi_n)>0.$$

同理可得: $m(\theta_0-\delta;\xi_1,\ldots,\xi_n)-m(\theta_0;\xi_1,\ldots,\xi_n)>0$.

类似定理3.1, 得$|\widehat{\theta}_n(\xi_1,\ldots,\xi_n)-\theta_0|<\delta=n^{-\alpha}$, 即$\widehat{\theta}_n(\xi_1,\ldots,\xi_n)=\theta_0+O(n^{-\alpha})$.

其次, 由$\widehat{\theta}_n(\xi_1,\ldots,\xi_n)=\theta_0+O(n^{-\alpha})$, $\widehat{a}_n(\xi_1,\ldots,\xi_n)=a+O(n^{-1})$及(3.6)式, 得

$$1-[1-\widehat{\theta}_n(\xi_1,\ldots,\xi_n)(\xi_i-\widehat{a}_n(\xi_1,\ldots,\xi_n))]^{1/k_0}=p_i+O(n^{-\alpha}).$$

取$\delta=n^{-\beta}$, $0.5<\beta<\alpha<1$, 则

$$\begin{aligned}\Delta_i^* &\triangleq [1-\widehat{\theta}_n(\xi_1,\ldots,\xi_n)(\xi_i-\widehat{a}_n(\xi_1,\ldots,\xi_n))]^{1/k_0}\\ &\quad-[1-\widehat{\theta}_n(\xi_1,\ldots,\xi_n)(\xi_i-\widehat{a}_n(\xi_1,\ldots,\xi_n))]^{1/(k_0+\delta)}\\ &= O(n^{-\beta}).\end{aligned}\tag{3.10}$$

类似(3.5)式, 得

$$\ell(k_0+\delta;\xi_1,\ldots,\xi_n)-\ell(k_0;\xi_1,\ldots,\xi_n)=\sum_{i=1}^n\Delta_i^*[\Delta_i^*+O(n^{-\alpha})].$$

由上式与(3.10)式知, 当n充分大时, Δ_i^*与$\Delta_i^*+O(n^{-\alpha})$同号, 故

$$\ell(k_0+\delta;\xi_1,\ldots,\xi_n)-\ell(k_0;\xi_1,\ldots,\xi_n)>0.$$

同理可得: $\ell(k_0-\delta;\xi_1,\ldots,\xi_n)-\ell(k_0;\xi_1,\ldots,\xi_n)>0$.

类似定理3.1, 得$|\widehat{k}_n(\xi_1,\ldots,\xi_n)-k_0|<\delta=n^{-\beta}$, 即$\widehat{k}_n(\xi_1,\ldots,\xi_n)=k_0+O(n^{-\beta})$. 因$0.5<\beta<\alpha<1$, 故有$\widehat{\theta}_n(\xi_1,\ldots,\xi_n)=\theta_0+o(n^{-1/2})$和$\widehat{k}_n(\xi_1,\ldots,\xi_n)=k_0+o(n^{-1/2})$. 证毕. □

定理 3.3 若$x_{1:n},x_{2:n},\ldots,x_{n:n}$为三参数GPD总体的次序样本, ξ_i为总体的$p_i=(i-0.375)/(n+0.25)$分位数, $i=1,2,\ldots,n$. 令$\theta=k/b$, $\boldsymbol{x}=(x_{1:n},\ldots,x_{n:n})'$, $\boldsymbol{\xi}=(\xi_1,\ldots,\xi_n)'$, 记$\widehat{a}_n$, $\widehat{\theta}_n$和$\widehat{k}_n$分别是由(2.12), (2.13)和(2.15)式得到的估计, 则

$$\sqrt{n}(\widehat{\theta}_n-\theta,\widehat{k}_n-k,\widehat{a}_n-a)'\xrightarrow{L}N_3(0,G\Sigma G'),\qquad n\to\infty,$$

其中$G=(\partial g_1/\partial\boldsymbol{x},\partial g_2/\partial\boldsymbol{x},\partial g_3/\partial\boldsymbol{x})'|_{\boldsymbol{x}=\boldsymbol{\xi}}$, $g_1(\boldsymbol{x})=\widehat{\theta}_n(x_{1:n},\ldots,x_{n:n})-\theta$, $g_2(\boldsymbol{x})=\widehat{k}_n(x_{1:n},\ldots,x_{n:n})-k$, $g_3(\boldsymbol{x})=\widehat{a}_n(x_{1:n},\ldots,x_{n:n})-a$.

证明: 由定理3.1知, 方程$\partial m/\partial\theta=0$有解, 记为$\widehat{\theta}_n$, 且$\widehat{\theta}_n\xrightarrow{P}\theta_0$, 其中$\theta_0$为参数真值. 将$\partial m/\partial\theta$在$\theta_0$处泰勒展开, 有

$$\frac{\partial m}{\partial\theta}=\frac{\partial m}{\partial\theta}\Big|_{\theta_0}+(\theta-\theta_0)\frac{\partial^2 m}{\partial\theta^2}\Big|_{\theta_0}+\frac{(\theta-\theta_0)^2}{2}\frac{\partial^3 m}{\partial\theta^3}\Big|_{\theta_1},$$

其中θ_1介于θ与θ_0之间. 取$\theta=\widehat{\theta}_n$, 则

$$\frac{\partial m}{\partial\theta}\Big|_{\widehat{\theta}_n}=\frac{\partial m}{\partial\theta}\Big|_{\theta_0}+(\widehat{\theta}_n-\theta_0)\frac{\partial^2 m}{\partial\theta^2}\Big|_{\theta_0}+\frac{(\widehat{\theta}_n-\theta_0)^2}{2}\frac{\partial^3 m}{\partial\theta^3}\Big|_{\theta_1},$$

其中θ_1介于θ与$\widehat{\theta}_n$之间, 又$\partial m/\partial\theta|_{\widehat{\theta}_n}=0$, 所以

$$\widehat{\theta}_n-\theta_0=-\frac{\dfrac{\partial m}{\partial\theta}\Big|_{\theta_0}}{\dfrac{\partial^2 m}{\partial\theta^2}\Big|_{\theta_0}+\dfrac{\widehat{\theta}_n-\theta_0}{2}\dfrac{\partial^3 m}{\partial\theta^3}\Big|_{\theta_1}}.\tag{3.11}$$

记$h_1(x_{1:n},\dots,x_{n:n})=\partial^2 m/\partial\theta^2|_{\theta_0}$, 则由(3.1)式和矩阵(delta)方法, 得

$$\sqrt{n}[h_1(x_{1:n},\dots,x_{n:n})-h_1(\xi_1,\dots,\xi_n)]\xrightarrow{L}N(0,\sigma_{h_1}^2),\qquad n\to\infty,$$

其中$\sigma_{h_1}^2=\sum\sum\big((\partial h_1/\partial\xi_i)(\partial h_1/\partial\xi_j)\sigma_{ij}\big)$, 所以

$$h_1(x_{1:n},\dots,x_{n:n})=h_1(\xi_1,\dots,\xi_n)+O_p(n^{-1/2}),$$

即

$$\frac{\partial^2 m}{\partial\theta^2}\Big|_{\theta_0}=h_1(\xi_1,\dots,\xi_n)+O_p(n^{-1/2}).\tag{3.12}$$

记$\psi(x_{1:n},\dots,x_{n:n})=\partial^3 m/\partial\theta^3|_{\theta_0}$, 类似(3.12)式, 得

$$\frac{\partial^3 m}{\partial\theta^3}\Big|_{\theta_0}=\psi(\xi_1,\dots,\xi_n)+O_p(n^{-1/2}).$$

由定理3.1知$\widehat{\theta}_n-\theta_0=o_p(1)$, 由于$\theta_1$介于$\widehat{\theta}_n$与$\theta_0$之间, 故$\theta_1=\theta_0+o_p(1)$, 所以

$$\frac{\partial^3 m}{\partial\theta^3}\Big|_{\theta_1}=\frac{\partial^3 m}{\partial\theta^3}\Big|_{\theta_0}+o_p(1)=\psi(\xi_1,\dots,\xi_n)+o_p(1),$$

从而

$$\frac{\widehat{\theta}_n-\theta_0}{2}\cdot\frac{\partial^3 m}{\partial\theta^3}\Big|_{\theta_1}=\frac{\widehat{\theta}_n-\theta_0}{2}[\psi(\xi_1,\dots,\xi_n)+o_p(1)]=o_p(1).\tag{3.13}$$

由(3.11), (3.12)和(3.13)式, 得

$$g_1(\boldsymbol{x})=-\frac{1}{h_1(\xi_1,\dots,\xi_n)}\frac{\partial m}{\partial\theta}\Big|_{\theta_0}+o_p(1),$$

所以

$$\frac{\partial g_1(\boldsymbol{x})}{\partial\boldsymbol{x}}=-\frac{1}{h_1(\xi_1,\dots,\xi_n)}\left(\frac{\partial\Big(\frac{\partial m}{\partial\theta}\Big|_{\theta_0}\Big)}{\partial x_{1:n}},\frac{\partial\Big(\frac{\partial m}{\partial\theta}\Big|_{\theta_0}\Big)}{\partial x_{2:n}},\dots,\frac{\partial\Big(\frac{\partial m}{\partial\theta}\Big|_{\theta_0}\Big)}{\partial x_{n:n}}\right)',\qquad n\to\infty,$$

其中$\partial(\partial m/\partial\theta|_{\theta_0})/\partial x_{i:n}$可由(2.14)式得到.

类似(3.11)式, 有

$$\widehat{k}_n-k_0=-\frac{\dfrac{\partial\ell}{\partial k}\Big|_{k_0}}{\dfrac{\partial^2\ell}{\partial k^2}\Big|_{k_0}+\dfrac{\widehat{k}_n-k_0}{2}\dfrac{\partial^3\ell}{\partial k^3}\Big|_{k_1}},$$

其中k_1介于k与$\widehat{k}_n$之间. 令$h_2(x_{1:n},x_{2:n}\dots,x_{n:n})=\partial^2\ell/\partial k^2|_{k_0}$, 类似地得

$$g_2(\boldsymbol{x})=-\frac{1}{h_2(\xi_1,\dots,\xi_n)}\frac{\partial\ell}{\partial k}\Big|_{k_0}+o_p(1).$$

所以

$$\frac{\partial g_2(\boldsymbol{x})}{\partial \boldsymbol{x}} = -\frac{1}{h_2(\xi_1,\ldots,\xi_n)}\left(\frac{\partial\Big(\frac{\partial \ell}{\partial k}\Big|_{k_0}\Big)}{\partial x_{1:n}}, \frac{\partial\Big(\frac{\partial \ell}{\partial k}\Big|_{k_0}\Big)}{\partial x_{2:n}}, \ldots, \frac{\partial\Big(\frac{\partial \ell}{\partial k}\Big|_{k_0}\Big)}{\partial x_{n:n}}\right)', \qquad n\to\infty,$$

其中$\partial(\partial l/\partial k|_{k_0})/\partial x_{i:n}$可由(2.16)式得到. 由(2.12)式, 知

$$\frac{\partial g_3(\boldsymbol{x})}{\partial \boldsymbol{x}} = \Big(\frac{p_2}{p_2-p_1}, -\frac{p_1}{p_2-p_1}, 0, \ldots, 0\Big)',$$

令$g(\boldsymbol{x}) = (g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), g_3(\boldsymbol{x}))'$, 则由(3.1)式和矩阵(delta)方法, 得

$$\sqrt{n}[g(\boldsymbol{x}) - g(\boldsymbol{\xi})] \xrightarrow{L} N(0, G\Sigma G'), \qquad n\to\infty, \tag{3.14}$$

其中$G = (\partial g_1/\partial \boldsymbol{x}, \partial g_2/\partial \boldsymbol{x}, \partial g_3/\partial \boldsymbol{x})'|_{\boldsymbol{x}=\boldsymbol{\xi}}$. 由(3.8)式和定理3.2, 知

$$\begin{aligned}\sqrt{n}g(\boldsymbol{\xi}) &= (g_1(\boldsymbol{\xi}), g_2(\boldsymbol{\xi}), g_3(\boldsymbol{\xi}))' = \sqrt{n}(\widehat{\theta}_n(\xi_1,\ldots,\xi_n), \widehat{k}_n(\xi_1,\ldots,\xi_n), \widehat{a}_n(\xi_1,\ldots,\xi_n))' \\ &= (o(1), o(1), o(1))'.\end{aligned}$$

所以由(3.14)式和slutsky定理, 知$\sqrt{n}[g(\boldsymbol{x}) - g(\boldsymbol{\xi})] + \sqrt{n}g(\boldsymbol{\xi}) \xrightarrow{L} N_3(0, G\Sigma G')$, $n\to\infty$, 即

$$\sqrt{n}(\widehat{\theta}_n - \theta, \widehat{k}_n - k, \widehat{a}_n - a)' \xrightarrow{L} N_3(0, G\Sigma G'), \qquad n\to\infty.$$

证毕. $\square$

注记 2 特别地, 令$\widehat{a}_n = a = 0$, 可得两参数GPD总体分布参数估计的渐近分布

$$\sqrt{n}(\widehat{\theta}_n - \theta, \widehat{k}_n - k)' \xrightarrow{L} N_2(0, Q\Sigma Q'), \qquad n\to\infty,$$

其中$Q = (\partial g_1/\partial \boldsymbol{x}, \partial g_2/\partial \boldsymbol{x})'|_{\boldsymbol{x}=\boldsymbol{\xi}}$.

§4. 模拟研究

4.1 两参数GPD模拟结果

下面通过模拟比较参数各种估计的偏差和标准误. 由于模拟结果不受b的取值影响, 故只考虑$b=1$, k分别取-2, -1.5, -1, -0.4, -0.2, 0.2, 0.4, 1, 2时的情况, 考虑的样本容量n分别取15, 50, 100. 这里比较了MOM, PWM, EPM和本文提出的LSE的模拟结果, 模拟次数为1000. (本文并未考虑MLE, 因为Hosking和Wallis (1987)的结论告诉我们: 和MOM, PWM相比, 在$n<500$时, MLE估计效果并不理想.)

表1　两参数情形下形状参数k的4种估计的偏差和标准误的模拟

		MOM		PWM		EMP		LSE	
n	k	偏差	标准误	偏差	标准误	偏差	标准误	偏差	标准误
	-2.00	1.60	1.60	1.26	1.27	0.52	1.39	-0.14	0.91
	-1.50	1.14	1.14	0.86	0.89	0.46	0.95	-0.11	0.81
	-1.00	0.71	0.72	0.52	0.59	0.25	0.76	-0.11	0.67
	-0.40	0.30	0.38	0.28	0.45	0.16	0.55	-0.06	0.52
15	-0.20	0.19	0.30	0.23	0.40	0.12	0.47	-0.07	0.46
	0.20	0.12	0.38	0.25	0.50	0.06	0.38	-0.01	0.43
	0.40	0.09	0.39	0.25	0.52	0.03	0.37	-0.02	0.39
	1.00	0.13	0.74	0.31	0.77	-0.03	0.44	-0.00	0.49
	2.00	0.54	2.45	0.57	1.59	-0.06	0.68	0.03	0.72
	-2.00	1.53	1.53	1.09	1.09	0.39	0.90	-0.13	0.55
	-1.50	1.04	1.04	0.66	0.67	0.28	0.60	-0.10	0.45
	-1.00	0.59	0.59	0.31	0.35	0.19	0.53	-0.06	0.35
	-0.40	0.16	0.21	0.09	0.20	0.09	0.29	-0.04	0.24
50	-0.20	0.08	0.15	0.06	0.18	0.08	0.24	-0.03	0.21
	0.20	0.02	0.15	0.06	0.20	0.03	0.16	-0.02	0.17
	0.40	0.02	0.17	0.06	0.22	0.03	0.15	-0.01	0.16
	1.00	0.02	0.33	0.08	0.34	-0.01	0.19	-0.01	0.20
	2.00	0.09	0.70	0.14	0.61	-0.01	0.35	0.01	0.35
	-2.00	1.51	1.51	1.05	1.05	0.28	0.73	-0.06	0.38
	-1.50	1.02	1.02	0.60	0.61	0.16	0.47	-0.04	0.31
	-1.00	0.56	0.56	0.25	0.28	0.14	0.40	-0.04	0.24
	-0.40	0.12	0.15	0.05	0.14	0.08	0.23	-0.02	0.16
100	-0.20	0.05	0.12	0.03	0.12	0.07	0.18	-0.02	0.14
	0.20	0.01	0.10	0.03	0.13	0.03	0.11	-0.01	0.11
	0.40	0.00	0.12	0.02	0.15	0.01	0.09	-0.01	0.10
	1.00	0.01	0.22	0.03	0.22	0.00	0.12	-0.00	0.13
	2.00	0.06	0.46	0.08	0.41	0.00	0.23	-0.00	0.24

表2　两参数情形下刻度参数b的4种估计的偏差和标准误的模拟

		MOM		PWM		EMP		LSE	
n	k	偏差	标准误	偏差	标准误	偏差	标准误	偏差	标准误
	-2.00	/	/	/	/	-0.12	1.11	0.30	1.67
	-1.50	95.40	1.10e+03	17.52	186.31	0.26	0.87	0.23	1.00
	-1.00	7.50	69.90	1.84	11.82	-0.03	0.61	0.12	0.68
	-0.40	0.41	0.72	0.33	0.63	-0.04	0.51	0.07	0.52
15	-0.20	0.20	0.44	0.23	0.51	-0.04	0.48	0.03	0.44
	0.20	0.12	0.47	0.23	0.57	-0.03	0.42	0.05	0.44
	0.40	0.08	0.43	0.20	0.52	-0.04	0.41	0.03	0.38
	1.00	0.10	0.54	0.19	0.55	-0.05	0.37	0.02	0.36
	2.00	0.24	1.07	0.23	0.71	-0.03	0.33	0.02	0.34
	-2.00	/	/	/	/	0.05	0.48	0.05	0.50
	-1.50	259.17	4.57e+03	14.30	229.23	0.04	0.41	0.02	0.39
	-1.00	4.33	11.60	0.55	0.84	0.03	0.33	0.03	0.31
	-0.40	0.24	0.35	0.10	0.27	0.02	0.26	0.00	0.26
50	-0.20	0.09	0.23	0.07	0.24	0.02	0.23	0.00	0.23
	0.20	0.02	0.20	0.05	0.23	-0.01	0.21	0.00	0.21
	0.40	0.02	0.21	0.05	0.23	0.01	0.19	0.00	0.19
	1.00	0.02	0.24	0.04	0.24	-0.01	0.18	0.00	0.17
	2.00	0.04	0.30	0.05	0.27	-0.01	0.17	0.00	0.17
	-2.00	/	/	/	/	0.06	0.36	0.01	0.30
	-1.50	635.76	1.51e+04	17.04	377.92	0.05	0.31	0.01	0.26
	-1.00	4.47	13.92	0.38	0.56	0.03	0.25	0.00	0.22
	-0.40	0.19	0.27	0.06	0.18	0.01	0.18	0.01	0.18
100	-0.20	0.06	0.16	0.04	0.16	0.02	0.16	0.00	0.16
	0.20	0.01	0.14	0.02	0.16	0.01	0.14	0.00	0.14
	0.40	0.00	0.14	0.02	0.16	0.00	0.13	0.00	0.13
	1.00	0.01	0.16	0.02	0.16	0.00	0.12	0.00	0.12
	2.00	0.02	0.19	0.03	0.18	0.00	0.11	0.00	0.12

从表1, 表2可得以下结论:

(1) 当$-0.2 < k < 0.2$时, MOM估计效果较好;

(2) 当$-0.4 < k < -0.2$时, PWM估计效果较好;

(3) 当$k > 0.2$或$k < -0.4$时, 本文提出的LSE的估计效果较好, 且在k的上述各种取值情况下, 本文提出的LSE几乎都优于EMP, 特别在$k < -0.4$时LSE明显优于EMP;

(4) 和MOM, PWM以及EMP相比, 随形状参数k的变化, LSE的估计效果较稳定.

4.2 三参数GPD模拟结果

这里对常用的三参数GPD的参数估计方法LM, PWM和本文提出的LSE进行比较. 同样只考虑$a = 5$, $b = 1$, k分别取-2, -1.5, -1, -0.4, -0.2, 0.2, 0.4, 1, 2时的情况, 考虑的样本容量n分别取15, 50, 100. 下面的数据是1000次模拟下的结果.

从表3, 表4和表5得以下结论:

(1) 当$-0.4 < k < 0.4$时, PWM估计效果较好;

(2) 当$|k| > 0.4$时, 本文提出的LSE的估计效果较好, 且随n增大估计效果显著提高;

(3) 和LM, PWM相比, 随形状参数k的变化, LSE的估计效果较稳定.

表3 三参数情形下形状参数k的3种估计的偏差和标准误的模拟

		LM		PWM		LSE	
n	k	偏差	标准误	偏差	标准误	偏差	标准误
15	-2.00	1.44	1.47	1.66	1.67	-0.29	1.22
	-1.50	1.05	1.10	1.19	1.20	-0.31	1.20
	-1.00	0.77	0.89	0.75	0.76	-0.25	0.79
	-0.40	0.75	0.98	0.27	0.31	-0.21	0.59
	-0.20	0.87	1.11	0.13	0.19	-0.24	0.59
	0.20	1.45	1.79	-0.14	0.18	-0.23	0.53
	0.40	2.15	2.59	-0.28	0.29	-0.22	0.52
	1.00	11.22	185.03	-0.81	0.81	-0.28	0.59
	2.00	-28.97	196.01	-1.90	1.90	-0.37	0.85
50	-2.00	1.15	1.15	1.56	1.56	-0.10	0.56
	-1.50	0.72	0.73	1.08	1.08	-0.10	0.44
	-1.00	0.38	0.43	0.63	0.64	-0.08	0.36
	-0.40	0.18	0.28	0.20	0.24	-0.06	0.25
	-0.20	0.17	0.26	0.09	0.16	-0.06	0.22
	0.20	0.27	0.33	-0.03	0.11	-0.04	0.18
	0.40	0.36	0.41	-0.11	0.14	-0.05	0.18
	1.00	0.87	0.95	-0.52	0.52	-0.05	0.21
	2.00	3.93	4.37	-1.71	1.72	-0.08	0.35
100	-2.00	1.08	1.08	1.53	1.53	-0.06	0.38
	-1.50	0.65	0.65	1.04	1.05	-0.05	0.31
	-1.00	0.28	0.31	0.58	0.58	-0.06	0.25
	-0.40	0.08	0.17	0.14	0.18	-0.04	0.17
	-0.20	0.08	0.15	0.06	0.12	-0.04	0.15
	0.20	0.12	0.17	-0.01	0.09	-0.03	0.11
	0.40	0.17	0.21	-0.05	0.09	-0.02	0.10
	1.00	0.38	0.43	-0.39	0.39	-0.02	0.13
	2.00	1.30	1.41	-1.66	1.66	-0.04	0.24

表4　三参数情形下形状参数b的3种估计的偏差和标准误的模拟

n	k	LM 偏差	LM 标准误	PWM 偏差	PWM 标准误	LSE 偏差	LSE 标准误
15	-2.00	4.55e+03	6.30e+04	1.87e+04	2.60e+05	0.13	1.00
	-1.50	2.40e+03	7.36e+04	9.90e+03	3.04e+05	0.02	0.75
	-1.00	2.97	7.25	6.04	27.28	-0.06	0.59
	-0.40	1.47	1.80	0.45	0.82	-0.15	0.47
	-0.20	1.55	1.94	0.11	0.40	-0.16	0.48
	0.20	2.36	3.24	-0.24	0.31	-0.16	0.43
	0.40	3.50	4.71	-0.35	0.38	-0.19	0.39
	1.00	3.81e+03	6.25e+04	-0.60	0.60	-0.25	0.39
	2.00	3.50e+03	7.68e+04	-0.82	0.82	-0.33	0.41
50	-2.00	6.52e+05	2.05e+07	9.59e+06	3.02e+08	0.02	0.47
	-1.50	31.86	326.55	4.34	4.79e+03	-0.01	0.35
	-1.00	1.72	14.60	15.08	213.73	-0.02	0.30
	-0.40	0.32	0.46	0.41	0.55	-0.03	0.24
	-0.20	0.27	0.40	0.13	0.28	-0.05	0.24
	0.20	0.35	0.45	-0.06	0.18	-0.05	0.21
	0.40	0.41	0.50	-0.13	0.19	-0.06	0.19
	1.00	0.83	0.92	-0.40	0.40	-0.07	0.18
	2.00	3.62	4.32	-0.76	0.76	-0.10	0.19
100	-2.00	2.61e+03	5.26e+04	7.71e+04	1.56e+06	-0.01	0.29
	-1.50	315.21	8.95e+03	9.30e+03	2.66e+05	-0.01	0.26
	-1.00	0.82	2.56	9.67	72.38	-0.02	0.21
	-0.40	0.14	0.26	0.32	0.43	-0.03	0.17
	-0.20	0.12	0.22	0.09	0.21	-0.03	0.16
	0.20	0.15	0.23	-0.03	0.14	-0.03	0.14
	0.40	0.19	0.25	-0.07	0.13	-0.02	0.13
	1.00	0.33	0.90	-0.30	0.31	-0.03	0.12
	2.00	0.89	0.96	-0.74	0.74	-0.05	0.12

表5　三参数情形下形状参数a的3种估计的偏差和标准误的模拟

n	k	LM 偏差	LM 标准误	PWM 偏差	PWM 标准误	LSE 偏差	LSE 标准误
15	-2.00	-1.96e+03	2.72e+04	-1.00e+04	1.39e+05	0.13	0.19
	-1.50	-1.03e+03	3.18e+04	-5.32e+03	1.63e+05	0.12	0.15
	-1.00	-0.93	2.96	-2.69	14.48	0.12	0.15
	-0.40	-0.32	0.36	-0.05	0.28	0.12	0.15
	-0.20	-0.32	0.36	0.03	0.15	0.11	0.14
	0.20	-0.39	0.43	0.11	0.15	0.10	0.12
	0.40	-0.48	0.52	0.13	0.17	0.10	0.13
	1.00	-1.69	20.67	0.16	0.18	0.09	0.11
	2.00	2.51	17.99	0.17	0.18	0.09	0.11
50	-2.00	-3.12e+05	9.84e+06	-5.99e+06	1.89e+08	0.03	0.04
	-1.50	-14.87	156.26	-269.63	2.99e+03	0.03	0.04
	-1.00	-0.66	6.97	-8.89	133.51	0.03	0.04
	-0.40	-0.08	0.11	-0.13	0.20	0.03	0.04
	-0.20	-0.07	0.09	-0.02	0.08	0.03	0.03
	0.20	-0.07	0.09	0.03	0.06	0.03	0.04
	0.40	-0.08	0.09	0.04	0.07	0.03	0.03
	1.00	-0.12	0.13	0.09	0.10	0.03	0.03
	2.00	-0.30	0.31	0.14	0.15	0.03	0.03
100	-2.00	-1.28e+03	2.57e+04	-4.98e+04	1.00e+06	0.01	0.02
	-1.50	-153.90	4.38e+03	-6.00e+03	1.71e+05	0.01	0.02
	-1.00	-0.29	1.22	-5.77	46.65	0.01	0.02
	-0.40	-0.03	0.06	-0.12	0.18	0.01	0.02
	-0.20	-0.03	0.05	-0.02	0.06	0.01	0.02
	0.20	-0.03	0.05	0.01	0.04	0.01	0.01
	0.40	-0.03	0.05	0.02	0.04	0.01	0.01
	1.00	-0.05	0.06	0.07	0.07	0.01	0.02
	2.00	-0.10	0.10	0.14	0.14	0.01	0.01

4.3 小结

通过模拟可以看出GPD参数的各种估计方法效果都随着样本容量n和分布形状参数k大小变化而变化, 虽不存在一致最优的估计方法, 但从4.1和4.2节的模拟结果可得以下结论:

(1) 当形状参数$|k|$较大时, 本文提出的LSE较优, 特别在$k < -0.4$时(此时GPD为厚尾分布), LSE在上述几种估计中最优. 实际生活中我们常用GPD拟合厚尾数据, 所以此时本文提出的LSE要优于MOM、PWM、LM、EMP以及MLE.

(2) 不论形状参数k在何范围, 本文提出的LSE均存在且估计效果相对较稳定. 在实际问题中, 时常没有k的任何先验信息, 因此MOM、PWM、LM以及MLE都缺少稳健性, 甚至可能不存在, 而EMP只适应两参数GPD的参数估计问题且当GPD是厚尾分布时估计的效果较差, 故此时本文提出的LSE在上述估计中最优.

参 考 文 献

[1] Pickands, J., Statistical inference using extreme value order statistivs, *The Annals of Statistics*, **3(1)**(1975), 119–131.

[2] Smith, R.L., Threshold methods for sample extremsh, *Statistical Extremes and Applications*, de Oliveira, J.T., Reidel, Dordrecht, 1984, 621–638.

[3] Hosking, J.R.M. and Wallis, J.R., Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics*, **29(3)**(1987), 339–349.

[4] Hosking, J.R.M., L-moments: analysis and estimation of distributions using linear combinationsof order statistics, *Journal of the Royal Statistical Society: Series B*, **52**(1990), 105–124.

[5] Castillo, E. and Hadi, A.S, Fitting the generalized Pareto distribution to data, *Journal of the American Statistical Association*, **92**(1997), 1609–1620.

[6] Moharram, S.H., Gosain, A.K. and Kapoor, P.N., A comparative study for the estimators of the generalized Pareto ditribution, *Journal of Hydrology*, **150**(1993), 169–185.

[7] 茆诗松, 王静龙, 濮晓龙, 高等数理统计, 北京: 高等教育出版社, 1999.

[8] David, H.A., *Order Statistics*, Second edition, New York: Wiley, 1981.

[9] Salvadori, G., Linear combinations of order statistics to estimate the position and scale parameters of the generalized Pareto distribution, *Stochastic Environmental Research and Risk Assessment*, **16**(2002), 1–17.

[10] Luceño, A., Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators, *Computational Statistics and Data Analysis*, **51**(2006), 904–917.

Estimation of Parameters of the Generalized Pareto Distribution by the Least Squares

Chen Haiqing[1,2] Cheng Weihu[2]

([1] *College of Mathematics and Computer Science, Guizhou Normal College, Guiyang, 550018*)

([2] *College of Applied Science, Beijing University of Technology, Beijing, 100124*)

Traditional estimations of parameters of the generalized Pareto distribution (GPD) are generally constrained by the shape parameter of GPD. Such as: the method-of-moments (MOM), the probability-weighted moments (PWM), L-moments (LM), the maximum likelihood estimation (MLE) and so on. In this paper we use the fact that GPD can be transformed into the exponential distribution and use the results of parameters estimation for the exponential distribution, than we propose parameters estimators of the two-parameter or three-parameter GPD by the least squares method. Some asymptotic results are provided and the proposed method not constrained by the shape parameter of GPD. A simulation study is carried out to evaluate the performance of the proposed method and to compare them with other methods suggested in this paper. The simulation results indicate that the proposed method performs better than others in some common situation.

Keywords: Generalized Pareto distribution, the method-of-moments, probability-weighted moments, the least squares, L-moments, elemental percentile method.

AMS Subject Classification: 62F12, 62F10.