稳定分布中偏度参数的一个新估计

徐美萍^{1*}, 张波²

1. 北京工商大学理学院, 北京 100048;

2. 中国人民大学统计学院, 北京 100875

E-mail: xumeiping2006@163.com, mabzhang@ruc.edu.cn

收稿日期: 2015-11-22; 接受日期: 2015-12-17; 网络出版日期: 2016-08-03; * 通信作者

教育部人文社会科学重点研究基地 (批准号: 14JJD910002) 和国家自然科学基金 (批准号: 71271210 和 71471173) 资助项目

摘要 当观测到数据出现 Gauss 分布无法捕捉的厚尾和非对称特征时, 具有幂率尾行为与求和框架的稳定分布常被用作拟合模型. 考虑到偏度参数是除特征指数之外另一个度量稳定分布尾行为的重要指标, 本文首先使用逻辑函数连接偏度参数和由协变量组成的线性预测, 构成一个尾回归模型. 然后, 使用近似对数似然函数获得偏度参数回归估计并给出估计的渐近正态性质. 最后, 通过一个实例阐明本文所给的估计不仅具有一定的解释经济意义的能力, 预测表现也在预期范围内.

关键词 一般中心极限定理 稳定分布 偏度回归估计

MSC (2010) 主题分类 60E07, 60E10, 60F05, 62E20, 62F12, 62P05, 62P20

1 引言

学者们普遍认可的一个事实是, 股票收益率是大量几乎连续到来的信息和个体决策的累加结果, 因此, 在厚尾和非对称性特征出现时, 通常假定它们可以被一个非 Gauss 的稳定分布来近似. 事实上, 稳定分布族对于描述包含极值事件 (如市场对冲或自然灾害) 的数据是很有价值的模型. 到目前为止, 它们已经成功地拟合了股票收益率、超额债券收益率、外汇交易率、日用品价格收益率和房地产收益率等 (参见文献 [1, 2]).

1963 年, Mandelbrot^[3] 首次把稳定分布应用于股票收益率数据. 到目前为止已有大量文献评估稳定分布描述股票收益率的效果, 且实证分析表明, 非正态的稳定分布通常能够很好地模拟独立同分布的股票收益率并解释数据中观测到的尖峰厚尾, 此处独立同分布提法的重要性源于证券价格表现为随机游走的想法, 因为价格水平是信息的瞬时影响随机到达的结果, 参见文献 [2, 4–7]. 又由基于一般化的中心极限定理的理论依据, 使得许多学者确信使用稳定分布模型来描述股票收益率是可行的.

称随机变量 Y 服从一维参数化的 α - 稳定分布, 如果其对数特征函数可以表示为

$$\psi(t) = \log E \exp\{iYt\} = \begin{cases} i\delta t - |ct|^\alpha \left[1 - i\beta \operatorname{sign}(t) \tan\left(\frac{\pi\alpha}{2}\right) \right], & \alpha \neq 1, \\ i\delta t - |ct| \left[1 + \frac{2i\beta}{\pi} \operatorname{sign}(t) \log|t| \right], & \alpha = 1 \end{cases} \quad (1.1)$$

英文引用格式: Xu M P, Zhang B. A new estimator of skewness parameter in stable distribution (in Chinese). Sci Sin Math, 2017, 47: 423–434, doi: 10.1360/N012015-00257

对所有 $-\infty < t < \infty$, 记为 $Y \sim s(y; \alpha, \beta, c, \delta)$, 其中 $\alpha \in (0, 2]$ 称为特征指数, 决定尾部的衰减速率. 当 $\alpha < 2$ 时, 分布的方差无限; 当 $\alpha > 1$ 时, 分布的均值存在且等于 δ ; 而当 $\alpha \leq 1$ 时, 分布的尾部太厚导致其均值都不存在. 本文不考虑 $\alpha = 2$ 的情形, 因为它对应于均值为 δ 、方差为 $2c^2$ 的正态分布, 具有特殊的行为特征. $\beta \in [-1, 1]$ 是偏度参数, 说明分布非对称性的程度和符号. β 越靠近 1, 分布越偏向于右尾, 因而越可能产生 Y 的极端正值; 反之, β 越靠近 -1, 分布越偏向于左尾, 因而越可能产生 Y 的极端负值; 当 $\beta = 0$ 时, 分布关于 $y = \delta$ 对称. 另外, $c > 0$ 和 $\delta \in (-\infty, +\infty)$ 分别是尺度和位置参数 (参见文献 [8]).

对于稳定分布, 除其尾部厚度外, 我们还非常关心其分布的非对称性特征, 这主要是因为分布偏向于右尾还是左尾在股票收益率的研究中有着完全不同的经济意义. 由此可见, 偏度参数 β 是除特征指数 α 之外刻画稳定分布尾行为的另一个重要参数.

关于 β 的估计, DuMouchel^[9] 最早提出了基于用级数近似似然函数的最大似然估计 (MLE) 方法; 之后 Nolan^[10] 对其进行了优化. Koutrouvelis^[11,12] 先后提出并改进了基于样本特征函数的回归程序; 后来, Kogon 和 Williams^[13] 又作了进一步改进. McCulloch^[14] 基于稳定分布的分位点制表, 提出了通过五个事先指定的样本分位点来计算 β 的一个估计方法. 在上述提到的几种关于 β 的估计方法中, MLE 方法几乎总是最精确的.

前面提到的估计方法的基本假设都为样本是独立同分布的, 由此得到的 MLE 拥有大家熟知的渐近理论性质, 但是它们不能用于研究 β 与预测变量之间的回归关系. 而在实证分析中, β 是与预测变量密切相关的, 这就促使我们去挖掘 β 的回归估计及其理论特性.

由于左尾和右尾的相对权重决定了 β 的取值, 所以, 本文将利用已有的稳定分布尾部概率的一些结果得到 β 的一个回归估计. 基本想法如下: 当 $\alpha < 2$ 时, 若 $Y \sim s(y; \alpha, \beta, 1, 0)$, 则当 $y \rightarrow +\infty$ 时,

$$P(Y > y) \rightarrow C_\alpha(1 + \beta)y^{-\alpha}, \quad P(Y < -y) \rightarrow C_\alpha(1 - \beta)y^{-\alpha},$$

其中 $C_\alpha = \Gamma(\alpha) \sin(\frac{\pi\alpha}{2})/\pi > 0$ (参见文献 [15]).

注意到稳定分布的双尾渐近等价于 Pareto 分布, 所以, 我们假定当 y 充分大时, 右尾概率和左尾概率分别由下面两式给出:

$$\begin{aligned} P(Y > y) &= (1 + \beta)y^{-\alpha}\mathcal{L}(y), \\ P(Y < -y) &= (1 - \beta)y^{-\alpha}\mathcal{L}(y), \end{aligned} \tag{1.2}$$

此处 $\mathcal{L}(y)$ 称为缓变函数, 满足对任意 $t > 0$, 当 $y \rightarrow +\infty$ 时, $\mathcal{L}(yt)/\mathcal{L}(y) \rightarrow 1$ (参见文献 [16]). 需要指出的是, 此处缓变函数 $\mathcal{L}(y)$ 的精确形式并未指定. 接下来, 我们引进一个逻辑函数把 β 与协变量向量 $\mathbf{X} \in \mathbb{R}^p$ 通过式子 $(1 + \beta(\mathbf{X}))/2 = \exp(\mathbf{X}^T\boldsymbol{\theta})/(1 + \exp(\mathbf{X}^T\boldsymbol{\theta}))$ 联系起来, 其中 $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)^T$ 是 p 维的未知回归系数向量. 为了估计 $\boldsymbol{\theta}$ 的具体取值, 我们采用 Hall^[16] 的处理方法, 假设当 $y \rightarrow \infty$ 时, 缓变函数 $\mathcal{L}(y)$ 以相当快的速度收敛于某个常数. 在此假设下, 定义于 (1.2) 中的条件示性函数 $I(Y > 0 | |Y| > y)$ 的真实分布能够被一个参数为 $\exp(\mathbf{X}^T\boldsymbol{\theta})/(1 + \exp(\mathbf{X}^T\boldsymbol{\theta}))$ 的 Bernoulli 分布很好地近似, 从而可以得到 $\boldsymbol{\theta}$ 的一个近似最大似然估计 (AMLE).

与已经存在的三种估计 β 的方法相比, 我们的方法首次考虑了协变量对偏度参数的影响, 是 MLE 方法与回归程序的一种结合, 可以看作是 β 的一个简单一致估计的推广. 其次, 在估计过程中, 我们提出了一种差异度量来选取极值的阈值. 本文将证明我们所给的估计不仅拥有大家熟知的渐近理论性质, 而且在实证分析中表现也不错.

2 偏度参数回归估计

2.1 稳定分布模型

假设 $(\mathbf{X}_i, Y_i)$ 是来自第 i 个目标的观测, 其中 $Y_i \in \mathbb{R}^1$ 是一个来自稳定分布 $s(y; \alpha, \beta, 1, 0)$ ($\alpha < 2$) 的响应变量, $\mathbf{X}_i = (X_{i1}, \dots, X_{ip})^T \in \mathbb{R}^p$ ($i = 1, 2, \dots, n$) 是相关的 p 维预测变量, 且假设各个观测 $(\mathbf{X}_i, Y_i)$ 相互独立. 此外, 令 $P(Y_i > y | \mathbf{X}_i = \mathbf{x})$ 和 $P(Y_i < -y | \mathbf{X}_i = \mathbf{x})$ ($y > 0$) 分别对应于 Y_i 在条件 $\mathbf{X}_i = \mathbf{x}$ 下的右尾概率和左尾概率. 根据 (1.2), 我们假设: 当 $y \rightarrow +\infty$ 时,

$$\begin{aligned} P(Y_i > y | \mathbf{X}_i = \mathbf{x}) &= (1 + \beta(\mathbf{x})) \cdot y^{-\alpha} \mathcal{L}(y; \mathbf{x}), \\ P(Y_i < -y | \mathbf{X}_i = \mathbf{x}) &= (1 - \beta(\mathbf{x})) \cdot y^{-\alpha} \mathcal{L}(y; \mathbf{x}), \end{aligned} \quad (2.1)$$

此处 $(1 + \beta(\mathbf{x}))/2 = \exp(\mathbf{x}^T \boldsymbol{\theta}) / (1 + \exp(\mathbf{x}^T \boldsymbol{\theta}))$, $\boldsymbol{\theta} \in \mathbb{R}^p$ 是真值为 $\boldsymbol{\theta}_0 \in \mathbb{R}^p$ 的回归系数, 且 $\mathcal{L}(y; \mathbf{x})$ 是某个依赖于协变量的缓变函数, 即当 $y \rightarrow \infty$ 时, 对任意 $t > 0$, $\mathcal{L}(yt; \mathbf{x}) / \mathcal{L}(y; \mathbf{x}) \rightarrow 1$.

我们采纳 Hall^[16] 提出的方法, 假设 $\mathcal{L}(y; \mathbf{x})$ 满足

$$\mathcal{L}(y; \mathbf{x}) = c_0(\mathbf{x}) + c_1(\mathbf{x})y^{-\alpha} + \mathcal{O}(y^{-\alpha}), \quad (2.2)$$

其中 $c_0(\mathbf{x}) > 0$ 和 $c_1(\mathbf{x})$ 是 $\mathbf{x}$ 的函数, 且 $\mathcal{O}(y^{-\alpha})$ 是高阶剩余项. 由此可见, $\mathcal{L}(y; \mathbf{x}) \rightarrow c_0(\mathbf{x})$, 且当 $y \rightarrow +\infty$ 时, $\dot{\mathcal{L}}(y; \mathbf{x}) \rightarrow 0$, 此处 $\dot{\mathcal{L}}(y; \mathbf{x}) = \partial \mathcal{L}(y; \mathbf{x}) / \partial y$.

2.2 一个近似最大似然估计

为估计偏度参数 β , 我们引进了一个阈值参数 $\omega_n > 0$ 来控制估计过程中的样本比例 n_0/n , 其中 $n_0 = \sum_{i=1}^n I(|Y_i| > \omega_n)$ 是有效样本量. 由 (2.1) 和 (2.2), 在给定 $|Y_i| > \omega_n$ 和 $\mathbf{X}_i = \mathbf{x}$ 时, $I(Y_i > \omega_n)$ 和 $I(Y_i < -\omega_n)$ 的条件概率分布可以分别由下面的两个式子来近似:

$$P(Y_i > \omega_n | |Y_i| > \omega_n, \mathbf{X}_i = \mathbf{x}) = \frac{1 + \beta(\mathbf{x})}{2}, \quad P(Y_i < -\omega_n | |Y_i| > \omega_n, \mathbf{X}_i = \mathbf{x}) = \frac{1 - \beta(\mathbf{x})}{2}.$$

从而有近似负的对数似然函数

$$\mathcal{K}_n(\boldsymbol{\theta}) = \sum_{i=1}^n \{ \log(1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta})) \cdot I(|Y_i| > \omega_n) - \mathbf{X}_i^T \boldsymbol{\theta} \cdot I(Y_i > \omega_n) \}, \quad (2.3)$$

此处 $I(\cdot)$ 是一个示性函数. 最小化 (2.3) 可以得到回归系数 $\boldsymbol{\theta}$ 的一个 AMLE, 记作 $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_p)^T$. 再利用下面的估计式可得到偏度参数 β 的一个估计

$$\hat{\beta} = \frac{1}{n_0} \sum_{i=1}^n \left\{ \frac{\exp(\mathbf{X}_i^T \hat{\boldsymbol{\theta}}) - 1}{\exp(\mathbf{X}_i^T \hat{\boldsymbol{\theta}}) + 1} \right\} \cdot I(|Y_i| > \omega_n). \quad (2.4)$$

注 2.1 当 $p = 1$ 且 $X_{i1} = 1$ 时, 我们可以由 (2.3) 得到 $\boldsymbol{\theta}$ 的一个解析估计 $\hat{\boldsymbol{\theta}} = \log(n_1) - \log(n_2)$, 其中 $n_1 = \sum_{i=1}^n I(Y_i > \omega_n)$ 和 $n_2 = \sum_{i=1}^n I(Y_i < -\omega_n)$ 分别是右尾和左尾的有效样本量. 由此得到偏度参数 β 的一个估计

$$\hat{\beta} = \frac{\exp(\hat{\boldsymbol{\theta}}) - 1}{\exp(\hat{\boldsymbol{\theta}}) + 1} = \frac{n_1 - n_2}{n_1 + n_2}. \quad (2.5)$$

另一方面, 由定义式 (1.1), β 具有下述性质:

$$\beta = \lim_{y \rightarrow \infty} \frac{P(Y > y) - P(Y < -y)}{P(Y > y) + P(Y < -y)}, \quad \alpha < 2. \quad (2.6)$$

因此, (2.5) 是 β 的一个简单一致估计, 且是 (2.4) 的一种特殊情形.

注 2.2 本文使用截断双尾数据来估计偏度参数, 应用的是条件 Bernoulli 分布模型, 为此我们选择了广泛应用于分类问题的逻辑函数, 它对于 Bernoulli 分布是最佳的连接函数.

2.3 样本比例选择

在参数估计过程中, 阈值 ω_n 用于确定估计 θ_0 的有效样本量 n_0 , 下面提出一种选择 ω_n 的方法, 具体做法如下:

令 $p = \exp(\mathbf{X}^T \theta_0) / (1 + \exp(\mathbf{X}^T \theta_0))$, 在给定 $|Y| > \omega_n$ 的条件下, 事件 $\{Y > \omega_n\}$ 和 $\{Y < -\omega_n\}$ 分别以概率 p 和 $1 - p$ 发生, 即对应于一个成功概率为 p 的 Bernoulli 分布. 在预测和评估这一模型时, 我们有可能犯两类错误 (就像假设检验中的第一类和第二类错误): 一是把真实的 $Y_i > \omega_n$ 预测为 $Y_i < -\omega_n$; 二是把真实的 $Y_i < -\omega_n$ 预测为 $Y_i > \omega_n$. 要使犯两类错误的概率同时变小是不可能的. 也就是说, 当我们在减小其中一类错误的时候, 另一类错误会增加. 所以, 我们需要寻求一种合理的方案平衡二者的关系. 另一方面, 如果我们选择了一个太大的样本比例, 会出现较小的方差但是较大的渐近偏差. 相反地, 如果样本比例选择太小, 渐近偏差将较小, 但会引起较大的方差. 因此, 上述两种情形都不能满足合理预测要求.

基于上述讨论, 我们考虑一个偏差与方差的折中方案, 即选取使得两类误差之和出现的频率最小的样本比例. 为此, 需要引入一个较为合理的概率界值 a 来推断第 i 个事件 $\{|Y_i| > \omega_n\}$ 当 $\hat{p}_i > a$ 时为 $\{Y_i > \omega_n\}$, 否则为 $\{Y_i < -\omega_n\}$, 其中 $\hat{p}_i = \exp(\mathbf{X}_i^T \hat{\theta}) / (1 + \exp(\mathbf{X}_i^T \hat{\theta}))$. 一个自然的想法是, 选取事件 $\{Y_i > \omega_n\}$ 在有效样本中的比例值 n_1/n_0 作为 a . 从而, 我们给出误差度量的定义如下:

$$\hat{E}(\omega_n) = \frac{1}{n_0} \left[\sum_{i=1}^n I \left\{ \hat{p}_i \leq \frac{n_1}{n_0}, Y_i > \omega_n \right\} + \sum_{i=1}^n I \left\{ \hat{p}_i > \frac{n_1}{n_0}, Y_i < -\omega_n \right\} \right]. \quad (2.7)$$

这意味着我们可以通过 $\omega_n^* = \operatorname{argmin}_{\omega_n} \hat{E}(\omega_n)$ 来选取一个适当的样本量 $n_0^* = \sum_{i=1}^n I(|Y_i| > \omega_n^*)$. 关于样本比例的选取也许有更多更好的选择, 由于对我们的理论结果无影响及篇幅所限, 本文不作进一步探讨.

3 偏度参数回归估计的理论性质

3.1 技术条件和假设

为符号简单计, 下文假定 ω_n 为第 2.3 小节选定的阈值, n_0 为相应的有效样本量. 为考察 $\hat{\theta}$ 的渐近行为, 需要以下技术条件:

(C.1) (标准化的协方差) 记 $\mathbf{Z}_{ni} = \Sigma_{\omega_n}^{-1/2} \mathbf{X}_i$, 假设

$$\frac{1}{n_0} \sum_{i=1}^n \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \cdot I\{|Y_i| > \omega_n\} = \Sigma_{\omega_n}^{-1/2} \hat{\Sigma}_{\omega_n} \Sigma_{\omega_n}^{-1/2} \xrightarrow{p} \mathbf{I}_p,$$

其中 $\mathbf{I}_p$ 是一个 p 阶单位矩阵, $\hat{\Sigma}_{\omega_n} = \frac{1}{n_0} \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^T I\{|Y_i| > \omega_n\}$ 是修整后的协变量的协方差矩阵 $\Sigma_{\omega_n} = E[\mathbf{X} \mathbf{X}^T | |Y| > \omega_n]$ 的一个样本估计.

(C.2) (缓变函数) 假设 (2.2) 中定义的剩余项 $\mathcal{O}(y^{-\alpha})$ 满足

$$\sup_{\mathbf{x}} \{y^\alpha \mathcal{O}(y^{-\alpha})\} \rightarrow 0, \quad \text{当 } y \rightarrow +\infty \text{ 时.}$$

(C.3) (收敛率) 假设 $n_0 \rightarrow \infty$ 且存在某个非零常向量 $\mathbf{h} \in \mathbb{R}^p$, 使得

$$n_0^{-1/2} n \omega_n^{-\alpha} \cdot E\{\mathbf{Z}_{ni}(1 - \beta^2(\mathbf{X}_i))[c_0(\mathbf{X}_i) + c_1(\mathbf{X}_i)\omega_n^{-\alpha}]\} \rightarrow \mathbf{h}.$$

由定义知, $E[\mathbf{Z}_{ni} \mathbf{Z}_{ni}^T | |Y_i| > \omega_n] = \mathbf{I}_p$, 因而假设 (C.1) 加强了标准化协变量 $\mathbf{Z}_{ni}$ 的弱大数定律. 假设 (C.2) 规范化了缓变函数 $\mathcal{L}(y; \mathbf{x})$ 的极值行为. 假设 (C.3) 隐含着 ω_n 的最优收敛率. 快于假设 (C.3) 所给的收敛率会减轻 $\hat{\theta}$ 的渐近偏差, 但增加它的方差. 反之, 慢于 (C.3) 所给的收敛率可以减小 $\hat{\theta}$ 的方差但会增大其渐近偏差.

3.2 理论性质

注意到 $\hat{\theta}$ 是 $\mathcal{K}_n(\theta)$ 的全局最小值点, 从而有规范方程

$$0 = \dot{\mathcal{K}}_n(\hat{\theta}) = \sum_{i=1}^n \left[\frac{\exp(\mathbf{X}_i^T \hat{\theta})}{1 + \exp(\mathbf{X}_i^T \hat{\theta})} \mathbf{X}_i \cdot I(Y_i < -\omega_n) - \frac{1}{1 + \exp(\mathbf{X}_i^T \hat{\theta})} \mathbf{X}_i \cdot I(Y_i > \omega_n) \right], \quad (3.1)$$

其中 $\dot{\mathcal{K}}_n(\cdot)$ 是 $\mathcal{K}_n(\cdot)$ 的一阶导数. 在建立 $\hat{\theta}$ 的理论性质之前, 需要先考察 $\dot{\mathcal{K}}_n(\cdot)$ 和 $\mathcal{K}_n(\cdot)$ 的 Hesse 矩阵 $\ddot{\mathcal{K}}_n(\cdot)$ 的渐近行为, 它们在研究最大似然估计的渐近行为时通常需要考虑, 而且需要通过有效样本量 n_0 来计算 $\hat{\theta}$, 所以, 接下来首先考虑样本比例 n_0/n 的渐近性质.

定理 3.1 在假设 (C.2) 和 (C.3) 下, 有 $n_0/n = \xi_n\{1 + o_p(1)\}$, 其中

$$\xi_n = 2\omega_n^{-\alpha} E[c_0(\mathbf{X}_i) + c_1(\mathbf{X}_i)\omega_n^{-\alpha}].$$

证明 由 (2.1), 得 $E(n_0/n) = P(|Y_i| > \omega_n) = E[2\omega_n^{-\alpha} \mathcal{L}(\omega_n; \mathbf{X}_i)]$. 应用 (2.2) 和假设 (C.2), 可进一步得到

$$E\left(\frac{n_0}{n}\right) = 2\omega_n^{-\alpha} E[c_0(\mathbf{X}_i) + c_1(\mathbf{X}_i)\omega_n^{-\alpha}]\{1 + o(1)\} = \xi_n\{1 + o(1)\}.$$

而且当 $\omega_n \rightarrow \infty$ 时,

$$\text{Var}\left(\frac{n_0}{n}\right) = \frac{1}{n} P(|Y_i| > \omega_n) [1 - P(|Y_i| > \omega_n)] \leq \frac{1}{n} P(|Y_i| > \omega_n) = \frac{\xi_n}{n} \{1 + o(1)\}.$$

上述结果结合假设 (C.3) 可以推导出

$$E\left[\frac{n_0}{n} \frac{1}{\xi_n} - 1\right]^2 = \frac{1}{\xi_n^2} \text{Var}\left(\frac{n_0}{n}\right) \{1 + o(1)\} \leq \frac{1}{n \xi_n} \{1 + o(1)\} \rightarrow 0.$$

这意味着 $n_0/n = \xi_n\{1 + o_p(1)\}$, 从而完成了定理的证明. $\square$

定理 3.1 刻画了 n_0/n (趋于 0) 是以数量级 $\omega_n^{-\alpha}$ 呈几何收敛的. 对应地, 用于似然函数 (2.3) 中的有效观测值最有可能来自厚尾部分, 而较轻的尾部将以大的概率被示性函数 $I\{|Y_i| > \omega_n\}$ 截去了.

定理 3.2 在假设 (C.1)–(C.3) 下, 有

$$n_0^{-1/2} \Sigma_{\omega_n}^{-1/2} \dot{\mathcal{K}}_n(\theta_0) \xrightarrow{d} N(\mathbf{0}, \mathbf{C}_p),$$

其中记号 $\xrightarrow{d}$ 表示 “依分布收敛”, 且

$$\mathbf{C}_p = E \left[\frac{\exp(\mathbf{X}_i^T \theta_0)}{(1 + \exp(\mathbf{X}_i^T \theta_0))^2} \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \mid |Y_i| > \omega_n \right]$$

是一个正定矩阵.

证明 因为 $n_0^{-1/2} \Sigma_{\omega_n}^{-1/2} \dot{\mathcal{K}}_n(\theta_0)$ 是 n 个独立同分布随机变量的和, 其渐近正态性可以通过以下两步来建立. 第 1 步, 证明 $E[n_0^{-1/2} \Sigma_{\omega_n}^{-1/2} \dot{\mathcal{K}}_n(\theta_0)] \rightarrow \mathbf{0}$; 第 2 步, 证明当 $n_0 \rightarrow \infty$ 时, $\text{cov}(n_0^{-1/2} \Sigma_{\omega_n}^{-1/2} \dot{\mathcal{K}}_n(\theta_0)) \rightarrow \mathbf{C}_p$; 然后, 运用中心极限定理, 立即得到 $n_0^{-1/2} \Sigma_{\omega_n}^{-1/2} \dot{\mathcal{K}}_n(\theta_0) \xrightarrow{d} N(\mathbf{0}, \mathbf{C}_p)$.

第 1 步 定义

$$\begin{aligned} \mathbf{Q}_{n1} &= n_0^{-1/2} n \cdot E \left[\mathbf{Z}_{ni} \frac{\exp(\mathbf{X}_i^T \theta_0)}{1 + \exp(\mathbf{X}_i^T \theta_0)} \cdot I(Y_i < -\omega_n) \right], \\ \mathbf{Q}_{n2} &= n_0^{-1/2} n \cdot E \left[\mathbf{Z}_{ni} \frac{1}{1 + \exp(\mathbf{X}_i^T \theta_0)} \cdot I(Y_i > \omega_n) \right], \end{aligned}$$

那么,

$$E[n_0^{-1/2} \Sigma_{\omega_n}^{-1/2} \dot{\mathcal{K}}_n(\theta_0)] = \mathbf{Q}_{n1} - \mathbf{Q}_{n2}.$$

接下来依次估计 $\mathbf{Q}_{n1}$ 和 $\mathbf{Q}_{n2}$. 我们有

$$\mathbf{Q}_{n1} = n_0^{-1/2} n \cdot E \left[\mathbf{Z}_{ni} \frac{\exp(\mathbf{X}_i^T \theta_0)}{1 + \exp(\mathbf{X}_i^T \theta_0)} \cdot P(Y_i < -\omega_n) \right],$$

由假设 (C.2) 及模型假设 (2.1) 和 (2.2), 上式的右边可以写成

$$\begin{aligned} & n_0^{-1/2} n \cdot E \left[\mathbf{Z}_{ni} \frac{\exp(\mathbf{X}_i^T \theta_0)}{1 + \exp(\mathbf{X}_i^T \theta_0)} (1 - \beta(\mathbf{X}_i)) \omega_n^{-\alpha} \mathcal{L}(\omega_n, \mathbf{X}_i) \right] \\ &= 2^{-1} n_0^{-1/2} n \omega_n^{-\alpha} \cdot E[\mathbf{Z}_{ni} (1 - \beta^2(\mathbf{X}_i)) [c_0(\mathbf{X}_i) + c_1(\mathbf{X}_i) \omega_n^{-\alpha}]] \{1 + o(1)\}. \end{aligned} \quad (3.2)$$

类似地, 有

$$\begin{aligned} \mathbf{Q}_{n2} &= n_0^{-1/2} n \cdot E \left[\mathbf{Z}_{ni} \frac{1}{1 + \exp(\mathbf{X}_i^T \theta_0)} (1 + \beta(\mathbf{X}_i)) \omega_n^{-\alpha} \mathcal{L}(\omega_n, \mathbf{X}_i) \right] \\ &= 2^{-1} n_0^{-1/2} n \omega_n^{-\alpha} \cdot E[\mathbf{Z}_{ni} (1 - \beta^2(\mathbf{X}_i)) [c_0(\mathbf{X}_i) + c_1(\mathbf{X}_i) \omega_n^{-\alpha}]] \{1 + o(1)\}. \end{aligned} \quad (3.3)$$

由此可见, 假设 (C.3) 与 (3.2) 和 (3.3) 一起蕴含着 $\mathbf{Q}_{n1} - \mathbf{Q}_{n2} \rightarrow \mathbf{0}$. 这就完成了第 1 步的证明.

第 2 步 估计 $\text{cov}(n_0^{-1/2} \Sigma_{\omega_n}^{-1/2} \dot{\mathcal{K}}_n(\theta_0))$.

定义 $\mathbf{h}_{n1} = n_0^{1/2} n^{-1} \mathbf{Q}_{n1}$, $\mathbf{h}_{n2} = n_0^{1/2} n^{-1} \mathbf{Q}_{n2}$. 由第 1 步的结果, 有 $n_0^{-1/2} n \cdot \mathbf{h}_{n1} \rightarrow \mathbf{h}$, 它意味着 $nn_0^{-1} \cdot \mathbf{h}_{n1} \mathbf{h}_{n1}^T \rightarrow 0$. 同理, $nn_0^{-1} \cdot \mathbf{h}_{n2} \mathbf{h}_{n2}^T \rightarrow 0$.

对应地,

$$\text{cov}(n_0^{-1/2} \Sigma_{\omega_n}^{-1/2} \dot{\mathcal{K}}_n(\theta_0))$$

$$\begin{aligned}
&= \frac{n}{n_0} \cdot \left[\text{cov} \left(\mathbf{Z}_{ni} \frac{\exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)} \cdot I(Y_i < -\omega_n) \right) + \text{cov} \left(\mathbf{Z}_{ni} \frac{1}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)} \cdot I(Y_i > \omega_n) \right) \right] \\
&= \frac{n}{n_0} \cdot E \left[\left(\frac{\exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)} \right)^2 \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \cdot I(Y_i < -\omega_n) \right] - \frac{n}{n_0} \cdot \mathbf{h}_{n1} \mathbf{h}_{n1}^T \\
&\quad + \frac{n}{n_0} \cdot E \left[\left(\frac{1}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)} \right)^2 \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \cdot I(Y_i > \omega_n) \right] - \frac{n}{n_0} \cdot \mathbf{h}_{n2} \mathbf{h}_{n2}^T \\
&= \frac{n}{n_0} \cdot P(|Y_i| > \omega_n) \cdot \left\{ E \left[\left(\frac{\exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)} \right)^2 \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \cdot I(Y_i < -\omega_n) \mid |Y_i| > \omega_n \right] \right. \\
&\quad \left. + E \left[\left(\frac{1}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)} \right)^2 \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \cdot I(Y_i > \omega_n) \mid |Y_i| > \omega_n \right] \right\} + o(1) \\
&= \frac{n}{n_0} \cdot P(|Y_i| > \omega_n) \cdot \left\{ E \left[\left(\frac{\exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)} \right)^2 \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \cdot \frac{1}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)} \mid |Y_i| > \omega_n \right] \right. \\
&\quad \left. + E \left[\left(\frac{1}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)} \right)^2 \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \cdot \frac{\exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)} \mid |Y_i| > \omega_n \right] \right\} + o(1) \\
&= \frac{n}{n_0} \cdot P(|Y_i| > \omega_n) E \left[\frac{\exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)}{(1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0))^2} \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \mid |Y_i| > \omega_n \right] + o(1). \tag{3.4}
\end{aligned}$$

运用定理 3.1 证明中的结果, 有 $nn_0^{-1} \cdot P(|Y_i| > \omega_n) \rightarrow 1$. 再利用假设 (C.1), (3.4) 右边可以写成

$$E \left[\frac{\exp(\mathbf{X}_i^T \boldsymbol{\theta}_0)}{(1 + \exp(\mathbf{X}_i^T \boldsymbol{\theta}_0))^2} \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \mid |Y_i| > \omega_n \right] + o(1) \rightarrow \mathbf{C}_p, \quad \text{当 } \omega_n \rightarrow \infty \text{ 时,}$$

从而完成了定理 3.2 的证明. □

定理 3.2 指出 $n_0^{-1/2} \boldsymbol{\Sigma}_{\omega_n}^{-1/2} \dot{\mathcal{K}}_n(\boldsymbol{\theta}_0)$ 渐近于一个无偏正态分布.

定理 3.3 在假设 (C.1)–(C.3) 下, 有

$$\frac{1}{n_0} \boldsymbol{\Sigma}_{\omega_n}^{-1/2} \ddot{\mathcal{K}}_n(\boldsymbol{\theta}_0) \boldsymbol{\Sigma}_{\omega_n}^{-1/2} \xrightarrow{p} \mathbf{C}_p,$$

此处 $\xrightarrow{p}$ 表示“依概率收敛”.

定理 3.3 的证明类似于定理 3.2 的证明, 因而略去. 下一个定理将建立 $\hat{\boldsymbol{\theta}}$ 的渐近正态性.

定理 3.4 在假设 (C.1)–(C.3) 下, 有

$$n_0^{1/2} \boldsymbol{\Sigma}_{\omega_n}^{1/2} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \mathbf{C}_p^{-1}).$$

证明 定义 $\boldsymbol{\gamma}_n \doteq \boldsymbol{\Sigma}_{\omega_n}^{1/2} (\boldsymbol{\theta} - \boldsymbol{\theta}_0)$, $\boldsymbol{\gamma}_n^* \doteq \boldsymbol{\Sigma}_{\omega_n}^{1/2} \boldsymbol{\theta}_0$, 那么, $\mathbf{X}_i^T \boldsymbol{\theta} = \mathbf{Z}_{ni}^T (\boldsymbol{\gamma}_n + \boldsymbol{\gamma}_n^*)$, 且目标函数 $\mathcal{K}_n(\boldsymbol{\theta})$ 可以写成

$$\begin{aligned}
\mathcal{K}_n^*(\boldsymbol{\gamma}_n) &\doteq \sum_{i=1}^n \{ \log(1 + \exp\{\mathbf{Z}_{ni}^T (\boldsymbol{\gamma}_n + \boldsymbol{\gamma}_n^*)\}) \cdot I(Y_i < -\omega_n) \\
&\quad - [\mathbf{Z}_{ni}^T (\boldsymbol{\gamma}_n + \boldsymbol{\gamma}_n^*) - \log(1 + \exp\{\mathbf{Z}_{ni}^T (\boldsymbol{\gamma}_n + \boldsymbol{\gamma}_n^*)\})] \cdot I(Y_i > \omega_n) \}.
\end{aligned}$$

设上述目标函数的最小值点为 $\hat{\boldsymbol{\gamma}}_n$. 接下来通过下面两个步骤来证明该定理. 第 1 步证明

$$\hat{\boldsymbol{\gamma}}_n \doteq \boldsymbol{\Sigma}_{\omega_n}^{1/2} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = O_p(n_0^{-1/2});$$

第 2 步建立 $\hat{\boldsymbol{\gamma}}_n$ 的渐近正态性.

第 1 步 $\mathcal{K}_n^*(\gamma_n)$ 的 Hesse 矩阵为

$$\ddot{\mathcal{K}}_n^*(\gamma_n) = \sum_{i=1}^n \frac{\exp(\mathbf{Z}_{ni}^T(\gamma_n + \gamma_n^*))}{[1 + \exp(\mathbf{Z}_{ni}^T(\gamma_n + \gamma_n^*))]^2} \mathbf{Z}_{ni} \mathbf{Z}_{ni}^T \cdot I(|Y_i| > \omega_n),$$

对任意 $\gamma_n \in \mathbb{R}^p$, 这是一个正定矩阵. 因此, $\mathcal{K}_n^*(\gamma_n)$ 是关于 γ_n 的严格凸函数. 接下来, 如果能够证明 $\mathcal{K}_n^*(\gamma_n)$ 至少有一个阶为 $O_p(n_0^{-1/2})$ 的局部最小值, 那么这个局部最小值一定就是其全局最小值 $\hat{\gamma}_n$.

考虑 $\mathcal{K}_n^*(n_0^{-1/2}\mathbf{u}) - \mathcal{K}_n^*(\mathbf{0})$, 其中 $\mathbf{u} \in \mathbb{R}^p$ 是一个长度为 L 的任意非随机向量. 利用二阶 Taylor 展开式, 可以得到

$$\mathcal{K}_n^*(n_0^{-1/2}\mathbf{u}) - \mathcal{K}_n^*(\mathbf{0}) = n_0^{-1/2} \mathbf{u}^T \dot{\mathcal{K}}_n^*(\mathbf{0}) + \mathbf{u}^T \left\{ \frac{1}{n_0} \ddot{\mathcal{K}}_n^*(\mathbf{0}) \right\} \frac{\mathbf{u}}{2} + o_p(1). \quad (3.5)$$

再应用定理 3.2 和 3.3 的结论, 我们得到

$$\begin{aligned} n_0^{-1/2} \dot{\mathcal{K}}_n^*(\mathbf{0}) &= n_0^{-1/2} \Sigma_{\omega_n}^{-1/2} \dot{\mathcal{K}}_n^*(\boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \mathbf{C}_p), \\ \frac{1}{n_0} \ddot{\mathcal{K}}_n^*(\mathbf{0}) &= \Sigma_{\omega_n}^{-1/2} \left\{ \frac{1}{n_0} \ddot{\mathcal{K}}_n^*(\boldsymbol{\theta}_0) \right\} \Sigma_{\omega_n}^{-1/2} \xrightarrow{p} \mathbf{C}_p. \end{aligned}$$

然后, 注意到只要常数 L 足够大, 以任意大的概率, (3.5) 中的二次项比线性项占优势. 特别地, 对任意 $\epsilon > 0$, 总存在常数 L , 使得

$$\liminf_n P\left(\inf_{\mathbf{u} \in \mathbb{R}^p: \|\mathbf{u}\|=L} \mathcal{K}_n^*(n_0^{-1/2}\mathbf{u}) > \mathcal{K}_n^*(\mathbf{0})\right) > 1 - \epsilon.$$

这意味着 $\mathcal{K}_n^*(\gamma_n)$ 至少有一个局部最小值, 即 $O_p(n_0^{-1/2})$ (参见文献 [17]). 正如在第 1 步开始时提到的那样, 此局部最小值一定是 $\hat{\gamma}_n = O_p(n_0^{-1/2})$.

第 2 步 由第 1 步, 我们知道 $\hat{\gamma}_n$ 满足规范方程 $\dot{\mathcal{K}}_n^*(\hat{\gamma}_n) = 0$, 且 $\hat{\gamma}_n = O_p(n_0^{-1/2})$. 运用下面的一阶 Taylor 展开式:

$$n_0^{-1/2} \dot{\mathcal{K}}_n^*(\hat{\gamma}_n) = n_0^{-1/2} \dot{\mathcal{K}}_n^*(\mathbf{0}) + \frac{1}{n_0} \ddot{\mathcal{K}}_n^*(\mathbf{0}) \{n_0^{1/2} \hat{\gamma}_n\} \cdot \{1 + o_p(1)\},$$

并再一次利用结果 $n_0^{-1/2} \dot{\mathcal{K}}_n^*(\mathbf{0}) \xrightarrow{d} N(\mathbf{0}, \mathbf{C}_p)$ 和 $\frac{1}{n_0} \ddot{\mathcal{K}}_n^*(\hat{\gamma}_n) \xrightarrow{p} \mathbf{C}_p$, 可得

$$n_0^{1/2} \hat{\gamma}_n = - \left\{ \frac{1}{n_0} \ddot{\mathcal{K}}_n^*(\hat{\gamma}_n) \right\}^{-1} \{n_0^{-1/2} \dot{\mathcal{K}}_n^*(\mathbf{0})\} \cdot \{1 + o_p(1)\} \xrightarrow{d} N(\mathbf{0}, \mathbf{C}_p^{-1}).$$

从而完成了定理 3.4 的证明. $\square$

定理 3.4 表明, $n_0^{1/2} \Sigma_{\omega_n}^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ 也渐近于一个无偏正态分布, 且较大的 ω_n 会导致 $\hat{\boldsymbol{\theta}}$ 较大的方差, 因为它正比于 $1/n_0$. 此外, 该定理还蕴含着 $\hat{\boldsymbol{\theta}}$ 是 $\boldsymbol{\theta}_0$ 的一个一致无偏估计.

4 一个实例

4.1 背景

一个公司的所得一般由资产收益率 (ROE) 来度量. ROE 是衡量公司盈利能力的一个重要指标, 许多资本分析家依赖它来判断股票的投资价值. 因此, 理解和预测一个公司的未来收入在现实的投资实践中非常重要, 详见文献 [18]. 中国的股票市场是世界上增长最快的资本市场之一, 它的不成熟和中

国资本市场的迅速发展等因素导致它极端不稳定, 这使得在中国股票市场运作的那些公司的资产收益非常难以预测. 在这样一个汹涌的资本市场里运作的公司所面临的不仅有机遇也有相当的风险和不确定性. 因此, 从业者迫切地想要了解为什么在正常情形下相似的公司极端情形下可能有完全不同的表现. 上述原因促使我们使用尾参数回归方法来研究一个公司的 ROE 分布的尾行为与其预测变量之间的关系.

4.2 数据和秩变换

本文采用了文献 [19] 所用的数据, 它包含了从 1997 年到 2000 年一共 2,943 个观测, 其中每个观测对应着一个上市公司的财务信息. 我们保留了 2000 年的 907 个观测用于样本外检验, 而把其余从 1997 年到 1999 年的 2,036 个观测作为训练样本用作描述性分析和模型拟合.

因为我们感兴趣的是极端正的和负的净资产收益率, 所以, 定义响应变量 Y 为下一年度的净资产收益率. 通过一个简单计算可以得到 Y 的样本峰度是 187.4412, 这意味着响应分布具有非常厚的尾部. 进一步, Y 的直方图 (图 1(a)) 显示大的样本峰度是源于响应分布的厚尾而不是一些“异常值”. 一般而言, 当年表现好的公司, 由于惯性效应, 其下一年度的表现也会趋向于较好. 所以为解释响应变量 Y 的尾行为, 我们从本年度选取了下面几个标准财务指标作为自变量. 它们是, 公司当年的净资产收益率 (ROE)、对数变换后的资产总计 (ASSET)、主营业务利润 (PM)、主营业务收入增长率 (GROWTH)、债务资本比例 (LEV) 和资产周转率 (ATO).

描述性分析透露这些预测变量具有极端厚的尾部, 因而会导致较差的估计. 为减轻预测变量中的那些极值观测, 我们采纳文献 [19] 提出的建议, 对这些厚尾预测变量进行了秩变换, 即令 R_{ij} 是 X_{ij} 在样本 $\{X_{ij} : 1 \leq i \leq n\}$ 中的秩, 重新定义 $X_{ij} := \Phi^{-1}\{(R_{ij} - 3/8)/(n + 1/4)\}$ (正态得分变换) 或 $X_{ij} := \sqrt{12}(R_{ij}/n - 0.5)$ (均匀得分变换). 由于两种变换得到的结果类似, 我们仅给出与正态得分变换相关的结果. 在此需要指出的是, 经过秩变换之后, 原来包含在原始数据中的大量信息会丢失, 而且回归估计和它们对应的 t -检验也仅可用于识别相关预测变量的重要性, 而不能作通常意义下的解释.

图 1 响应变量的直方图和误差度量曲线

4.3 估计

我们利用文献 [19] 得到 α 的估计值小于 2, 在特征指数的范围之内, 从而数据可以用稳定分布来拟合. 首先把我们的方法应用于训练样本, 图 1(b) 显示在误差度量 (2.7) 下计算得到的最优样本比例是 $n_0/n = 0.5661$. 然后, 利用 (2.4) 可得到 β 的一个估计值为 $\hat{\beta} = 0.7038$, 这意味着 ROE 的分布是右偏的, 因而更可能产生极端正的净资产收益率.

表 1 给出了偏度回归估计和它们的相关统计量. 从表 1 中可以看出所有的变量均与预测来年的 ROE 相关, 因为它们相伴 p 值都在 0.02 之下. 我们的分析进一步透露那些具有较大的 ROE (较高的本年度的净资产收益率)、PM (更好的利润率) 和 ATO (资产的更好使用) 的公司会有较大的偏度参数, 因而, 它们更可能报告极值收益而不是极值损失, 这是与这些变量的经济意义相一致的. 与它们相比, 变量 GROWTH、LEV 和 ASSET 对偏度参数的影响是很弱的.

4.4 预测

为评价上述估计结果是否对预测起作用, 我们做了样本外预测研究. 设 $\{\mathbf{X}_i^*, Y_i^*\}$ ($1 \leq i \leq 907$) 是保留的测试样本, $\tilde{\omega}_n$ 是预先指定的用以区分正常和极值情形的阈值. 例如, 当 $\{-\tilde{\omega}_n < Y_i^* < \tilde{\omega}_n\}$ 时公司被认为是正常的, 否则被认为是异常情形. 为检验预测表现, 我们对样本外预测估计如下.

首先, 利用第 4.3 小节得到的 β 的估计值定义概率界 $a^* = (1 + \hat{\beta})/2 = 0.8519$, 并且推断事件 $\{|Y_i| > \tilde{\omega}_n\}$ 当 $\hat{p}_i^* > a^*$ 时为 $\{Y_i^* > \tilde{\omega}_n\}$, 否则为 $\{Y_i^* < -\tilde{\omega}_n\}$, 此处 $\hat{p}_i^* = \exp(\mathbf{X}_i^{*\text{T}} \hat{\boldsymbol{\theta}}) / (1 + \exp(\mathbf{X}_i^{*\text{T}} \hat{\boldsymbol{\theta}}))$. 接着, 通过下面三个式子分别估计样本外预测误差 (OPE)、事件 $\{Y_i^* > \tilde{\omega}_n\}$ 被正确预测的概率 (TPR) 和事件 $\{Y_i^* < -\tilde{\omega}_n\}$ 被错误预测的概率 (FPR) 分别为

$$\begin{aligned} \text{OPE}(\rho) &= \frac{1}{\tilde{n}_1 + \tilde{n}_2} \left[\sum_{i=1}^{907} I\{\hat{p}_i^* \leq a^*, Y_i^* > \tilde{\omega}_n\} + \sum_{i=1}^{907} I\{\hat{p}_i^* > a^*, Y_i^* < -\tilde{\omega}_n\} \right], \\ \text{TPR}(\rho) &= \frac{1}{\tilde{n}_1} \sum_{i=1}^{907} I\{\hat{p}_i^* > a^*, Y_i^* > \tilde{\omega}_n\}, \quad \text{FPR}(\rho) = \frac{1}{\tilde{n}_2} \sum_{i=1}^{907} I\{\hat{p}_i^* > a^*, Y_i^* < -\tilde{\omega}_n\}, \end{aligned}$$

其中 $\tilde{n}_1 = \sum_{i=1}^{907} I(Y_i^* > \tilde{\omega}_n)$, $\tilde{n}_2 = \sum_{i=1}^{907} I(Y_i^* < -\tilde{\omega}_n)$, 且 $\rho = (\tilde{n}_1 + \tilde{n}_2)/907$ 是测试数据中用到的样本比例.

经计算知测试数据中用到的样本比例 $\rho \in [0.1852, 0.6108]$. 从图 2 可以看出 OPE 对所有 ρ 的取值一致地小于 0.21. 图 3 显示, 当 TPR 的值较大时相应 FPR 的值也较大, 这意味着要获得一定的 TPR 精度, 就需要牺牲一些 FPR 值. 当样本比例大于 0.4 时, TPR 精度逐渐降低, FPR 值也随之减少. 注

表 1 ROE 数据的估计结果

预测变量	回归估计	标准误	检验统计量	显著性水平
INTERCEPT	1.9559	0.0319	61.2377	0.0000
ROE	0.7790	0.0435	17.9274	0.0000
ASSET	-0.0953	0.0305	3.1234	0.0018
PM	0.5720	0.0478	11.9785	0.0000
GROWTH	0.0781	0.0327	2.3912	0.0168
LEV	0.1572	0.0376	4.1762	0.0000
ATO	0.4639	0.0397	11.6922	0.0000

图 2 样本外预测误差

图 3 样本外预测图

意到, 我们希望对所有样本比例都产生大的 TPR 和小的 FPR. 图 3(a) 显示对所有 ρ 的取值 TPR 都大于 0.77, 图 3(b) 显示对所有 ρ 的取值 FPR 一致小于 0.165, 预测结果基本在我们可以接受的范围之内. 我们也对使用其他截尾方法获取的样本比例做了实证分析, 得到的结果类似.

5 结论与评注

本文研究了缓变函数满足 (2.2) 的稳定分布 (1.1) 中的偏度参数的估计问题. 借鉴假设检验中两类错误的判别方法提出了一种选择极值阈值的方法, 尝试把偏度参数与协变量通过逻辑函数联系在一起, 利用尾回归模型和最大似然方法给出了偏度参数的一个新估计, 并得到了偏度参数回归估计的理论性质, 从而方便把它应用到实际问题中. 实证分析表明, 本文所给的估计不仅具有一定的解释经济意义能力, 预测表现也在预期范围内.

后续, 我们可以从以下四个方向推进当前的研究. 第一, 寻求一个“最优”样本比例的选择方法, 使其从理论上来说更合理且可以更高效地选取样本比例; 第二, 把偏度参数回归方法应用于几个相互独立的分区组数据的情形; 第三, 引进其他半参变换连接偏度参数与协变量; 第四, 提出一种方法, 同时考虑响应变量和预测变量都是厚尾分布的情形. 我们相信这些努力会进一步加强回归方法在极值数据分析中的应用.

致谢 衷心感谢审稿人提出的修改意见和建议.

参考文献

- 1 McCulloch J H. Financial applications of stable distributions. In: Handbook of Statistics, vol. 14. New York: North-Holland, 1996, 393–425
- 2 Rachev S T, Mittnik S. Stable Paretian Models in Finance. New York: Wiley, 2000
- 3 Mandelbrot B. The variation of certain speculative prices. J Bus, 1963, 36: 394–419
- 4 Embrechts P, Klüppelberg C, Mikosch T. Modelling Extremal Events for Insurance and Finance. Berlin: Springer-Verlag, 1997
- 5 Fama E. The behavior of stock market prices. J Bus, 1965, 38: 34–105
- 6 Roll R. The Behavior of Interest Rates: The Application of the Efficient Market Model to US Treasury Bills. New York: Basic Books, 1970
- 7 Samuelson P. Rational theory of warrant pricing. Ind Manage Rev, 1965, 6: 13–32
- 8 Zolotarev V M, Vladimir M. Mellin-Stieltjes transforms in probability theory. Theory Probab Appl, 1957, 2: 433–460
- 9 DuMouchel W H. On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution. Ann Statist, 1973, 1: 948–957
- 10 Nolan J P. Maximum likelihood estimation of stable parameters. In: Levy Processes: Theory and Applications. Boston: Birkhäuser, 2001, 379–400
- 11 Koutrouvelis I A. Regression-type estimation of the parameters of stable laws. J Amer Statist Assoc, 1980, 75: 918–928
- 12 Koutrouvelis I A. An iterative procedure for the estimation of the parameters of stable laws. Comm Statist Simulation Comput, 1981, 10: 17–28
- 13 Kogon S M, Williams D B. Characteristic function based estimation of stable parameters. In: A Practical Guide to Heavy Tailed Data. Boston: Birkhäuser, 1998, 311–335
- 14 McCulloch J H. Simple consistent estimators of stable distribution parameters. Comm Statist Simulation Comput, 1986, 15: 1109–1136
- 15 Lévy P. Théorie des erreurs la loi de Gauss et les lois exceptionnelles. Bull Soc Math France, 1925, 52: 49–85
- 16 Hall P. On some simple estimates of an exponent of regular variation. J R Stat Soc Ser B Stat Methodol, 1982, 44: 37–42
- 17 Fan J, Li R. Variable selection via nonconcave penalized likelihood and its oracle properties. J Amer Statist Assoc, 2001, 96: 1348–1360
- 18 Harrison W T, Horngren C T. Financial Accounting. New York: Prentice Hall, 2001
- 19 Wang H, Tsai C L. Tail index regression. J Amer Statist Assoc, 2009, 104: 1233–1240

A new estimator of skewness parameter in stable distribution

XU MeiPing & ZHANG Bo

Abstract Power-law tail behavior and the summation scheme of stable laws are the basis for their frequent use as models for data when fat tails and asymmetric feature above a Gaussian distribution are observed. Considering the skewness parameter to be another important index to gauge the tail behavior of stable laws besides characteristic exponent, the logistic function is employed to link the skewness parameter to a linear predictor induced by covariates, which constitutes a tail regression model. Then an approximate log-likelihood function is used to obtain skewness parameter regression estimators, and the asymptotic normality of these estimators is proved. Finally, A real example is presented to illustrate that the estimators provided by the paper have some abilities to explain economic meaning and forecast performances as expected too.

Keywords generalized central limit theorem, stable laws, skewness regression estimation

MSC(2010) 60E07, 60E10, 60F05, 62E20, 62F12, 62P05, 62P20

doi: 10.1360/N012015-00257