第 40 卷第 7 期　　西南大学学报(自然科学版)　　2018 年 7 月
Vol. 40 No. 7　　Journal of Southwest University (Natural Science Edition)　　Jul. 2018

DOI：10.13718/j.cnki.xdzk.2018.07.012

极值指标下平稳序列的风险测度[①]

蔡宗鹏，　陈守全

西南大学 数学与统计学院，重庆 400715

摘要：平稳金融序列存在相依性，不满足独立同分布假设下极值理论的条件约束. 讨论了极值指标下平稳随机序列的风险测度，通过估计极值指标 θ 并修正广义极值分布的位置参数和尺度参数，得到平稳随机序列的风险值. 实证和检验表明平稳金融序列需要极值指标修正模型，提高风险值估计的准确度.

关 键 词：极值理论；平稳序列；相依性；极值指标

中图分类号：O212.3　　**文献标志码**：A　　**文章编号**：1673－9868(2018)07－0079－06

运用极值理论对随机序列尾部刻画和建模一直是学术界关注的重要问题. 文献[1－2]都基于极值理论的 POT 模型，分别应用于巨额损失保费计算厘定和金融机构单一方面风险度量. 文献[3]基于极值理论方法给出了最准确的 VaR 估计. 文献[4]结合 Garch 模型和极值理论提出动态价值风险估计.

由于数据存在相依关系，因此有必要利用平稳随机变量序列的极值估计解决实际问题. 国外的研究中处理相依性问题大多基于文献[5]提出的方法. 该方法易造成原样本信息损失，存在一定局限. 本文引入极值指标度量数据的相依结构，给出了两种不同情况下的风险值估计并进行实证分析.

1　平稳序列的风险测度

设 $X_1, X_2, \cdots, X_n$ 是随机变量序列，若满足

(1) $E(\{X_n\})=u$；

(2) $\mathrm{cov}(X_n, X_{n-l})=\sigma_1$；

则称序列 $\{X_n\}$ 是协方差平稳的(弱平稳). 本文主要考虑弱平稳序列且 $M_n=\max(X_1, X_2, \cdots, X_n)$.

文献[6] 弱化了 Rosenblatt(1956) 提出的强混合条件，引入 $D(u_n)$ 和 $D'(u_n)$ 条件讨论平稳序列的极值理论. $D(u_n)$ 条件表示平稳随机变量序列的渐近独立性，$D'(u_n)$ 条件则表示平稳随机变量序列中极端观测值接近的概率可渐近忽略.

满足 $D(u_n)$ 和 $D'(u_n)$ 条件的平稳随机变量序列极值的渐近分布服从广义极值分布(GEV)：

$$G(x;\mu,\sigma,\gamma)=\exp\left\{-\left(1+\gamma\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\gamma}}\right\},\ 1+\gamma\frac{x-\mu}{\sigma}>0$$

其中：μ 为位置参数，σ 为尺度参数，γ 为形状参数.

$D(u_n)$ 条件实质很弱，容易满足，但是 $D'(u_n)$ 条件对多数序列并不合理，常常出现成串的极大值. 极值指标 θ 能反映序列数据的相依结构和极端情况之间的关系，下面给出极值指标的定义.

---

① 收稿日期：2017－06－20
基金项目：国家自然科学基金项目(11571283).
作者简介：蔡宗鹏(1992－)，男，硕士研究生，主要从事极值统计分析的研究.
通信作者：陈守全，副教授，硕士研究生导师.

定义1 设 $X_1, X_2, \cdots, X_n$ 为平稳随机序列，其分布函数为 $F(x)$，θ 为非负数，若对任意 $\tau > 0$ 存在序列 $\{u_n\}$ 满足

(1) $\lim\limits_{n \to +\infty} nF(u_n) = \tau$，

(2) $\lim\limits_{n \to +\infty} P(M_n \leqslant u_n) = \mathrm{e}^{-\theta\tau}$，

则称 θ 为平稳随机序列 $\{X_n\}$ 的极值指标，且 $\theta \in [0, 1]$.

设 $\widetilde{X}_1, \widetilde{X}_2, \cdots, \widetilde{X}_n$ 是独立的随机变量序列，若序列 $\{\widetilde{X}_n\}$ 中每一个随机变量 $\widetilde{X}_i$ 都与平稳随机序列 $\{\widetilde{X}_n\}$ 中随机变量 X_i 相互对应且具有相同的分布，则称序列 $\{\widetilde{X}_n\}$ 为平稳随机序列 $\{X_n\}$ 的伴随独立同分布序列.

文献[7]得到了基于极值指标下平稳随机变量序列极值的渐近分布和与其相伴的独立同分布序列极值的渐近分布之间关系的结论.

定理1[7] 存在常数列 $\{a_n > 0\}$ 和 $\{b_n\}$，以及非退化分布函数 $\widetilde{F_*}(x)$ 使得 $P\{a_n(M_n - b_n) \leqslant x\} \to \widetilde{F_*}(x)$，$n \to \infty$. 若对于每个 x 有 $D(u_n)$ 条件成立，其中 $u_n = a_n x + b_n$，使得 $\widetilde{F_*}(x) > 0$，且对某个 x，$P\{a_n(M_n - b_n) \leqslant x\}$ 收敛，则对于某个常数 $\theta \in [0, 1]$，有 $P\{a_n(M_n - b_n) \leqslant x\} \to F_*(x) = \widetilde{F_*}^{\theta}(x)$，$n \to \infty$.

定理1说明通过极值指标 θ，可以将平稳随机变量序列极值的渐近分布拟合问题转化为与其相伴的独立同分布序列极值的渐近分布拟合问题. 设独立同分布序列 $\{\widetilde{X}_i\}$，$\widetilde{X}_n$ 服从位置参数为 b，尺度参数为 a，形状参数为 γ 的 GEV 分布. 根据定理1可得

$$F_*(x) = \widetilde{F_*}^{\theta}(x) = \exp\left\{-\theta\left(1 + \gamma\frac{x-b}{a}\right)^{-\frac{1}{\gamma}}\right\} = \exp\left\{-\left(1 + \gamma_*\frac{x-b_*}{a_*}\right)^{-\frac{1}{\gamma_*}}\right\}$$

其中

$$a_* = a\theta^{\gamma} \qquad b_* = b - \frac{a(1-\theta^{\gamma})}{\gamma}$$

对极值指标为 θ 且满足 $D(u_n)$ 条件的平稳随机序列 $\{X_n\}$，其样本最大值的渐近分布是形状参数为 γ 的广义极值分布，与独立同分布的随机序列相同，而位置参数 a_* 和尺度参数 b_* 均受序列数据自身相依性的影响，与极值指标有关.

平稳时间序列极值指标 θ 的估计有区组法、平均法. 区组法根据极值指标定义变形得到

$$\lim_{n \to +\infty}\frac{\ln P(M_n \leqslant u_n)}{n\ln F(u_n)} = \theta$$

利用样本信息划分为每组样本量为 k 的 $\frac{n}{k}$ 组，且 $g = \left[\frac{n}{k}\right]$. 有估计量

$$\hat{\theta}_1 = \frac{1\ln\left[1 - \frac{G(u_n)}{g}\right]}{k\ln\left[1 - \frac{N(u_n)}{n}\right]}$$

其中：$N(u_n)$ 为样本中大于门限 u_n 的样本点个数，$G(u_n)$ 为各组最大值大于门限 u_n 的组个数. 平均法即是在区组法估计量之上对分子分母分别作一阶 Taylor 展开得到估计量

$$\hat{\theta}_2 = \frac{\frac{G(u_n)}{g}}{\frac{kN(u_n)}{n}} = \frac{G(u_n)}{N(u_n)}$$

为给出极值指标下平稳随机序列的风险测度，先给出 VaR 的定义. VaR 表示在给定置信水平下，某一资产组合的最大损失值. 金融市场中 VaR 能兼顾风险发生的概率水平和损失程度，被视为风险管理主要指

标，用于度量金融机构所面临的信用风险、市场风险和操作风险这 3 类主要风险.

定义 2　对给定 $p\in(0,1)$，VaR 定义为：$P(\Delta X\geqslant VaR)=1-P(\Delta X<VaR)=p$，$\Delta X$ 表示资产的改变量(损失量).

利用 BMM 方法，假设每个子样本(长度为 n) 极值服从 GEV 分布，可以得到形状、位置、尺度 3 个参数估计值：$\hat{\gamma}$,$\hat{a}$ 和 $\hat{b}$. 对独立同分布收益率序列$\{\widetilde{X}_n\}$，令概率 p^* 表示一个多头头寸潜在损失超过一定限度的可能性，且 x_n^* 是在子样本最小值渐近分布为 GEV 分布条件下的$(1-p^*)$分位数，即有

$$1-p^*=\begin{cases}\exp\left\{-\left[1+\dfrac{\hat{\gamma}(x_n^*-\hat{b})}{\hat{\sigma}}\right]^{-\frac{1}{\hat{\gamma}}}\right\} & \hat{\gamma}\neq 0\\ \exp\left\{-\exp-\left(\dfrac{x_n^*-\hat{b})}{\hat{a}}\right)\right\} & \hat{\gamma}=0\end{cases}$$

取对数，反解 x_n^* 得到

$$x_n^*=\begin{cases}\hat{b}-\dfrac{\hat{a}}{\hat{\gamma}}\{1-[-\ln(1-p^*)]^{-\hat{\gamma}}\} & \hat{\gamma}\neq 0\\ \hat{b}-\hat{a}\ln[-\ln(1-p^*)] & \hat{\gamma}=0\end{cases}$$

在收益率序列样本 X_t 独立同分布假设下，样本 X_t 和其子样本$\{x_{n,n}\}$之间关系为：

$$1-p^*=P(x_{n,n}\leqslant x_n^*)=[P(X_t\leqslant x_n^*)]^n$$

则对给定概率 p，持有对数收益率 X_t 资产的 VaR 为：

$$VaR=\begin{cases}\hat{b}-\dfrac{\hat{a}}{\hat{\gamma}}\{1-[-n\ln(1-p^*)]^{-\hat{\gamma}}\} & \hat{\gamma}\neq 0\\ \hat{b}-\hat{a}\ln[-n\ln(1-p^*)] & \hat{\gamma}=0\end{cases}\tag{1}$$

在收益率序列样本 X_t 满足 $D(u_n)$ 条件且平稳的假设下，令极值指标为 $\theta(\theta\in[0,1])$，样本 X_t 和其子样本$\{x_{n,n}\}$之间关系为：

$$1-p^*=P(x_{n,n}\leqslant x_n^*)=[P(X_t\leqslant x_n^*)]^{n\theta}$$

则对给定概率 p，持有对数收益率 X_t 资产的 VaR 为：

$$VaR=\begin{cases}\hat{b}-\dfrac{\hat{a}}{\hat{\gamma}}\{1-[-n\theta\ln(1-p^*)]^{-\hat{\gamma}}\} & \hat{\gamma}\neq 0\\ \hat{b}-\hat{a}\ln[-n\theta\ln(1-p^*)] & \hat{\gamma}=0\end{cases}\tag{2}$$

由极值指标取值范围知，如果忽略序列数据相依性，平稳随机变量序列的风险值存在低估的风险.

2　实证分析

下面给出极值指标下平稳随机序列风险测度的应用.

选取我国股票市场具有代表性的上证综指作为实证对象，考虑到风险测度方法主要用途是预测和量化下一期风险程度，将样本划分为 2 个部分，其中 2010 年 1 月 4 日至 2015 年 12 月 31 日的上证综指收盘价格作为研究样本用于模型拟合和 VaR 估计，样本容量 1455；2016 年 1 月 4 日至 2016 年 12 月 1 日的上证综指收盘价格作为检验样本用于VaR 预测效果评价，样本容量 222. 样本来源于上海证券交易所公开数据，采用 R 软件进行数据处理和统计分析.

定义收益率：

$$X_t=100(\ln P_t-\ln P_{t-1})$$

其中 P_t 表示上证综指第 t 日收盘价格，其基本统计特征见表 1，时序图见图 1.

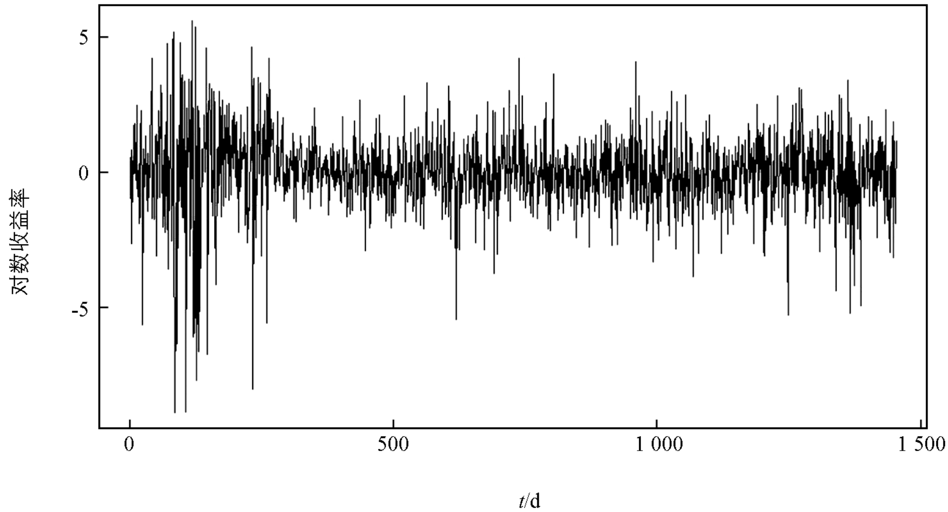

图1 收益率时序图

表1 收益率基本特征

时间序列	最大值	最小值	均值	标准差	峰度	偏度
上证综指 X_t	5.603 555	−8.873 175	0.005 991	1.485 816	5.067 688	−0.820 344

由表1可知上证综指收益序列$\{X_t\}$的峰度值和偏度值. 与正态分布比较，上证综指收益序列呈现尖峰厚尾、左偏的分布形态. 由图1可知序列还存在波动率聚集效应. 采用ADF检验和BDS统计量分别考察序列的平稳性(表2)以及序列自身数据的独立性问题(表3). ADF检验统计量为−10.876，在5%显著性水平下拒绝“序列存在单位根”的原假设，即$\{X_t\}$是平稳的. 二维、三维BDS统计量分别为2.390 8和3.299 5，在5%显著性水平下拒绝“序列数据独立”的原假设，说明序列数据自身存在相依性，在平稳序列的极值建模和VaR估计过程中不能忽略这一问题.

表2 平稳性检验

时间序列	ADF统计量	p 值	检验结果
上证综指 X_t	−10.876	0.01	平稳

表3 独立性检验

维度	BDS统计量	p 值	检验结果
2	2.390 8	0.016 8	不独立
3	3.299 5	0.001 0	不独立

选用BMM对上证综指收益序列$\{X_t\}$的极值分布进行GEV拟合，根据股票平均月度交易天数约为21天划分子样本，取$n=21$. GEV拟合的最大似然估计具有无偏、渐近正态这样的优良性质，得到估计值$\hat{\varepsilon}=0.059\,567\,76$，$\hat{a}=0.785\,868\,15$和$\hat{b}=1.984\,635\,97$. GEV拟合的残差图(图2)和对指数分布QQ图(图3)表明模型拟合合理，子样本容量取值合理.

根据序列相依性取$k=10$，利用组方法估计极值指标. 分别考察基于样本数据的平均超越函数图(图4)和极值指标估计图(图5)以确定门限和极值指标的估计值. 可知门限2.5附近样本平均超越量函数斜率为正且线性变化，而对应的极值指标估计值$\hat{\theta}=0.8$.

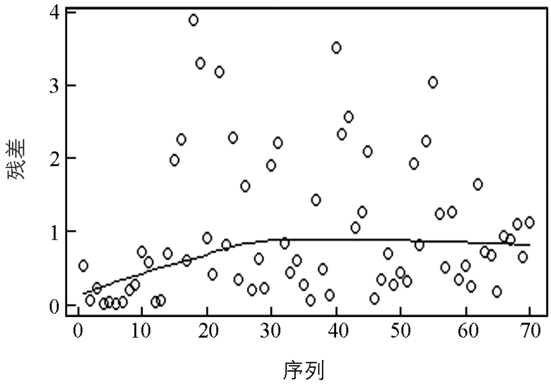

图 2　GEV 拟合残差图

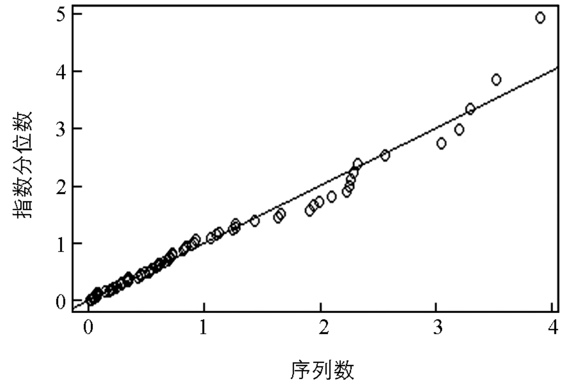

图 3　对指数分布 QQ 图

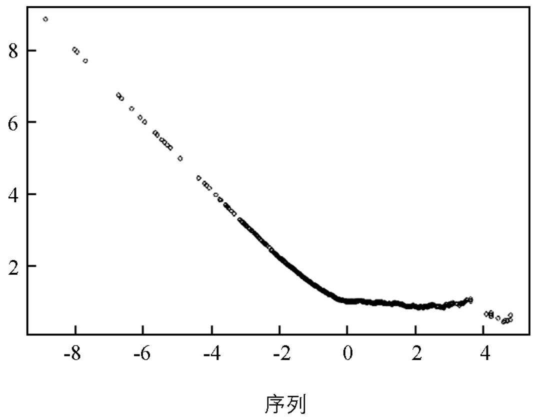

图 4　平均超越函数图

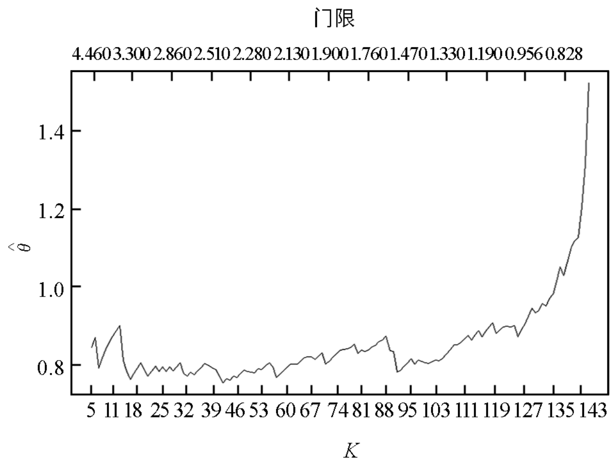

图 5　极值指标估计图

为对不同风险测度的准确性进行检验，计算在未引入极值指标和引入极值指标两种情况下上证综指收益序列$\{X_t\}$的 VaR_1 和 VaR_2. 由公式(1)得到未引入极值指标下 5% VaR_1 值为 1.926 4，表示假设上证综指资产组合的多头头寸为 100 万人民币，其相应的多头头寸在 1 天持有期内的 5% VaR 为 19 264 元. 同样，由公式(2)得到引入极值指标($\hat{\theta}=0.8$)下，5% VaR_2 为 2.102 1，表示假设上证综指资产组合的多头头寸为 100 万人民币，其相应的多头头寸在 1 天持有期内的 5% VaR 为 21 021 元.

采用 Kupiec 失败频率检验法对检验样本进行测试. 记损失量超过 VaR 估计量为失败，记损失量低于 VaR 估计量为成功，设考察总天数为 T，失败天数为 N，失败概率为 $p=\frac{N}{T}$，而在置信水平 q 下失败的期望概率为 $\tilde{p}=1-q$. Kupiec 提出了原假设为 $p=\tilde{p}$ 的似然比检验，其中 LR 统计量：

$$LR=-2\ln(\tilde{p}^N\ (1-\tilde{p})^{T-N})+2\ln\left(\left(\frac{N}{T}\right)^N\left(1-\frac{N}{T}\right)^{T-N}\right)$$

q 为 VaR 的分位数. 检验结果如下所示(表 4).

表 4　检验结果

	估计方法总天数/d	失败天数/d	失败率	检验 p 值
未引入极值指标 VaR_1	222	13	0.058 5	0.568 4
引入极值指标 VaR_2	222	11	0.049 5	0.975 4

从表 4 中可以看出引入极值指标计算的 VaR_2 失败天数更少，失败率更小且更接近 5%的概率设定，p 值也表明引入极值指标计算的 VaR_2 能更准确度量风险，效果更好.

3 结 论

在对近7年上证综指收益序列数据的实证研究中，计算并比较未引入极值指标和引入极值指标两种情况下的风险值，检验样本证明引入极值指标计算风险值是更准确的风险测度. 理论和实证均表明忽略序列数据自身相依性会导致风险被低估，而极值指标能够修正风险值，引入极值指标下的平稳序列的风险值更准确.

参考文献:

[1] 赵 智，李兴绪. 非寿险中巨额损失数据的拟合与精算 [J]. 数理统计与管理，2010，29(5)：336－347.

[2] 宋 坤，陈野华. 基于变点理论的 POT 模型阈值确定方法——对操作风险经济资本的度量 [J]. 统计与信息论坛，2011，26(7)：23－27.

[3] TOTIĈ S，BULAJIĈ M，VLASTELICA T. Empirical Comparison of Conventional Methods and Extreme Value Theory Approach in Value-at-Risk Assessment [J]. Business Management，2011，5(33)：12810－12818.

[4] BHATTACHARYYA M，SIDDARTH M. A Comparison of VaR Estimation Procedures for Leptokurtic Equity Index Returns [J]. Journal of Mathematical Finance，2012，2(1)：13－30.

[5] ALEXANDER M N，RUDIGER F. 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] LOYNES R. Extreme Values in Uniformly Mixing Stationary Stochastic Processes [J]. Annals of Mathematical Statistics，1965，36：993－999.

[7] LEADBETTER M R，LINDGREN G，ROOTZEN H. Extremes and Related Properties of Random Sequences and Processes [M]. Berlin：Springer，1983.

[8] RUEY S T. Analysis of Financial Time Series [M]. 3rd ed. Hoboken：John Wiley & Sons，2012.

Risk Measurement of Stationary Sequence Under Extreme Value Index

CAI Zong-peng， CHEN Shou-quan

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

Abstract：The stationary financial sequence is not independent while it does not satisfy the requirement of the conditional constraint of the extreme value theory under the consumption of modified independent identical distribution (i.i.d.). In this paper，we discuss the risk measurement of the stationary random sequence under the extremal index，and derive the valve at risk of the stationary random sequence by estimating the extremal index and modifying the position parameter and the scale parameter of the GEV. Empirical and test results show that the stationary financial sequence needs the extremal index to improve the accuracy of the VaR estimation.

Key words：extreme value theory；stationary sequence；independence；extremal index

责任编辑 张 枸