

胖尾分布及长记忆下的动态 EVT-VaR 测度研究^①

林宇^{1,2}, 黄登仕², 魏宇²

(1. 成都理工大学商学院, 成都 610059; 2. 西南交通大学经济管理学院, 成都 610031)

摘要: 针对金融收益胖尾分布特征及条件波动率长记忆性特征, 运用 FIGARCH 对条件波动率建模, 极值理论 (extreme value theory, EVT) 对标准收益序列的尾部建模, 测度出金融市场动态极值风险, 进而运用返回测试 (back-testing) 技术, 对模型在样本内的测度准确性与样本外的推广能力进行稳健性检验. 实证研究结果表明, 无论是中国新兴市场, 还是西方成熟发达市场, 金融收益与标准收益均呈现出明显的有偏胖尾分布特征; 金融收益条件波动率均展现出长记忆性特征; EVT 与 FIGARCH 模型相结合的动态极值风险测度模型不仅在样本内表现出优越的风险测度能力, 而且在样本外同样具有可靠的预测推广能力.

关键词: 胖尾分布; 长记忆性; 极值理论; 动态 VaR 测度; 稳健性检验

中图分类号: F830; F224 **文献标识码:** A **文章编号:** 1007-9807(2011)07-0071-12

0 引言

金融市场风险管理一直是政府管理当局、投资主体关注的焦点与难点. 尤其是近 20 年来爆发了一系列全球性金融经济风险危机灾难事件, 如 1997 年发生的亚洲金融危机, 2008 年出现的全球性经济危机以及最近发生的迪拜危机事件等. 这些全球性重大金融风险危机事件进一步警示我们: 只有强化金融市场风险管理, 才能有效地维护经济安全, 促进经济发展与繁荣.

在金融风险管理中最为关键的工作之一就在于准确有效地测度风险. 只有准确有效地测度出金融风险, 才可能保证风险管理效果, 维护资本市场的稳定与发展; 如果风险测度就出现偏差, 那么其后所进行的风险管理活动都可能失败, 自然也就谈不上风险的管理与控制.

长期以来, 金融风险管理都是以有效市场假说 (efficient market hypothesis, EMH) 为基石的金融理论展开的, 它假设金融收益服从正态/高斯分

布 (normal distribution/Gauss distribution), 然后根据收益的分布特征来测度金融市场的风险值 (value at risk, VaR)^[1-2]. 需要说明的是, VaR 指标仍然是目前金融市场风险管理实务所采用的最为常见的方法之一, 国际清算银行 (the bank of international settlements, BIS) 要求各大金融机构按照市场风险指标 (即 VaR) 来安排资本充足率; 许多国际金融投资机构也根据 VaR 方法来预测金融市场风险, 并以此作出金融资产投资决策制定^[3-4].

然而, 同样需要指出的是, 尽管传统的 VaR 指标方法在测度金融市场风险上目前仍然得到了相关机构的认可与使用. 但令人遗憾的是, 由于近 20 年来计算机科学与信息技术的迅猛发展, 产生了大量既不能为 EMH 所解释, 又普遍存在于不同类型与不同地区市场当中的实证成果, 如金融资产收益呈现出分布条件/非条件胖尾特征 (conditional/unconditional fat tailed distribution)、金融

① 收稿日期: 2008-12-22; 修订日期: 2010-12-15.

基金项目: 国家自然科学基金资助项目 (70771097); 教育部“新世纪优秀人才支持计划”资助项目 (NCET-08-0826); 教育部社科研究基金青年资助项目 (10YJCZH086).

作者简介: 林宇 (1973-), 男, 四川仪陇人, 博士, 副教授. Email: linyuphd@126.com

资产价格波动的长记忆性(long memory) 等重要的典型事实(stylized facts) 特征, 从而使传统的风险(VaR) 测度方法暴露出种种缺陷^[5]. 也正是由于不断涌现出来的典型事实, 使得金融收益并非服从 EMH 条件下的正态分布或学生 t 分布等单一确定的分布特征, 很可能是一个多分布的复合形式, 而有效风险测度的关键又在于风险模型能对金融时间序列分布形态等特征有一个准确可靠的描述. 因此, 如果仍然按照传统的单一分布, 又尤其是以正态分布来测度金融市场风险, 就会大大降低风险测度的准确性, 导致风险管理失效, 于金融市场安全与投资者不利^[6-7].

对金融市场管理当局与投资者来说, 尤其关注发生概率小、一旦发生就会带来巨大损失甚至是遭受灭顶之灾的极值事件风险, 因为这类风险很可能导致投资者破产, 甚至可能引发经济崩溃、社会动荡^[2]. 就金融收益分布而言, 其尾部就代表着这一类风险, 从而要求风险管理应根据金融收益的复杂分布的尾部分布进行建模分析. 而极值理论(extreme value theory, EVT) 就不必预先假设整个样本所服从的分布, 只对分布尾部进行建模^[4]. 因此, 应用 EVT 于金融风险管理与实务中来测度 VaR, 能有效克服传统风险管理的单一分布假设的缺陷, 使风险测度更加准确和容易, 从而受到实务界与学术界的欢迎^[8-12].

另外, 金融市场涌现出大量的典型事实就要求必须以这些典型事实为约束条件来开展风险测度研究工作. 一些学者运用 GARCH 簇模型以及随机波动(stochastic volatility, SV) 模型与 EVT 结合对风险测度进行建模分析, 虽然都取得了较好的研究效果^[13-14], 但由于 GARCH 模型无法捕获条件波动率的长记忆特征, 而 Bailie 等人 1996 年提出了刻画金融收益波动率长记忆性的分数积分 GARCH 模型(fractional integrated GARCH, FIGARCH)^[15]. 由于 FIGARCH 能够较好刻画条件波动率的长记忆性特征, 从而出现了以 FIGARCH 模型为基础展开的实证研究^[16-18]. 同时值得一提的是, 目前已有学者开始运用 FIGARCH 模型与正态分布、学生 t 分布(student-t distribution) 以及有偏学生 t 分布(skew student-t distribution, SKST)^[19] 相结合, 对诸如中国股市这样的新兴资

本市场的风险测度进行了实证研究^[20-24], 为以波动率长记忆性与收益率的胖尾分布为约束条件, 开展中国金融市场动态风险测度研究与实务奠定了基础.

就中国金融市场而言, 成立时间短, 从第一支股票上市到现在也只不过 20 年时间, 虽然发展迅猛, 目前已成为全球最大的新兴金融市场(emerging finance markets) 之一, 但与发达成熟的西方金融市场相比, 无论是在市场监督管理等体制建设上, 还是在投资者心理行为等方面都还表现出诸多的不成熟性. 比如说, 过剩的流动性、庄家操控以及过度的市场投机等导致市场大起大落时有发生, 从而使中国金融市场经受着更多更大的风险考验, 政府金融经济管理当局、投资者在风险管理方面承受着更多新的挑战^[24-25].

如上所述, 金融市场呈现出收益率有偏胖尾分布与条件波动率长记忆性等诸多典型事实特征, 不仅使传统的金融市场风险管理方法面临挑战, 对新兴的中国金融市场有着更多的风险考验. 那么, 如何基于 EVT 与 FIGARCH 构建出适合金融市场实际的风险测度方法? 并且构建的风险测度模型是否又能准确测度中国金融市场的动态极值风险? 如果能, 那么在测度西方成熟发达金融市场与中国新兴市场的风险在可靠性与精度方面是否又存在明显的差异呢?

基于以上分析与认识, 以中国大陆新兴股票市场与成熟市场为对比研究对象, 引入 FIGARCH 模型捕获金融收益条件波动率的长记忆特征; 针对金融时间序列的有偏胖尾分布结构和分布的复杂性而非单一分布特征, 引入 EVT 对收益率分布的极值尾部进行建模, 以此来开展金融市场 EVT-VaR 动态风险测度; 然后运用 Kupiec^[26] 和 Christoffersen 提出的似然比率 LR_{uc} , LR_{ind} 和 LR_{cc} 对风险测度准确性进行稳健性检验^[26-28], 实证考察胖尾分布与条件波动率长记忆性下风险模型在样本内测度的准确性与精度, 以及样本外的推广能力.

最后, 就所掌握的文献而言, 目前的大多数研究仅仅运用 EVT 与 GARCH 簇模型(如 GARCH、GJR-GARCH) 结合来研究金融市场风险测度方法^[22-23, 28-30], 并依据 VaR 失败比率来确定风险模型的准确性与精度, 很少见到运用规范、完整的

返回测试 (back-testing) 技术来得到相应的检验结果, 没有见到运用能够捕获金融市场收益条件波动率长记忆性这一重要典型事实的 FIGARCH 模型与 EVT 结合, 对成熟资本市场与我国新兴市场的风险测度进行实证对比研究; 尤其是在对极值尾部门槛值 (extreme tail threshold) 的选取上, 既不是采用超均值函数 (mean excess function, MEF) 图方法^[31], 也不是运用 10% 最大损失值作为门槛值, 而是结合两种方法来确定最优门槛值. 另外, 不仅选取的样本是目前所见研究中最大的样本, 而且也得到了一些具有重要价值的实证结果, 即基于 EVT 与 FIGARCH 的风险测度模型不仅在样本内表现出准确的测度能力, 而且样本外也呈现出优越的预测推广能力, 这为典型事实条件下金融风险管理的理论与实践提供了一定的决策借鉴. 由此可见, 与其它文献相比, 所具有的差异性是显而易见的.

1 样本选取及条件收益分布的统计特征

由于研究不同类型市场的动态极值风险测度的准确性及其效果对比, 所以, 选取新兴市场的中国大陆上证综指 (SSEC) 和成熟市场的标准普尔指数 (S&P500)、纳斯达克指数 (Nasdaq), 以及美国 10 年期连续国债指数 (T-bond) 四种指数作为代表性的研究对象. 样本期为 1990 年 12 月 19 日

(上证综指编制基准日) — 2010 年 7 月 9 日. 由于不同市场开市、休市情况不一致, 对数据进行了初步的加工处理. 处理方法为: 如果某一个市场当日开市交易, 另一个市场休市, 那么对另一个市场就以上一个交易日的收盘价代替当日的收盘价补齐, 以保证数据的一致性, 共计得到 5 083 个交易日的收盘指数价格. 同时, 为了检验风险测度模型的样本外推广能力, 先留出最后 1 000 天数据作为样本外模型预测推广能力的稳健性 (robustness) 检验, 其余 4 083 天数据作为样本内风险测度模型的实证建模与分析.

在金融风险计量中, 常以金融条件损失 (条件负收益率) 开始开展金融风险管理研究. 假设金融市场第 t 日的收盘指数为 P_t , 第 t 日的条件负收益 (以下简称条件收益) R_t 定义为

$$R_t = -\ln(P_t/P_{t-1}) = -(\ln P_t - \ln P_{t-1}) \quad (1)$$

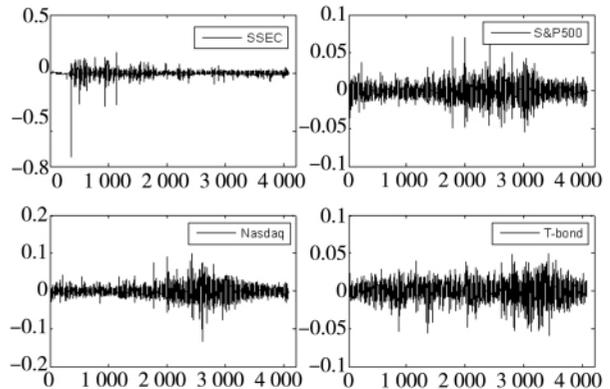


图 1 四种指数收益的波动状况图

Fig. 1 Volatility plot of conditional returns of four financial indices

表 1 四种指数条件收益的统计特征

Table 1 Descriptive statistics for returns of four indices

	SSEC	S&P500	Nasdaq	T-bond
Mean	-0.000 688	-0.000 338	-0.000 435	0.000 129
Standard deviation	0.025 999	0.009 865	0.015 165	0.010 852
Skewness	-6.240 290***	0.092 406**	0.009 521	-0.391 599***
Kurtosis (Excess)	156.968 9***	4.269 056***	5.999 766***	2.687 789***
J-B	4.216 2e + 6***	3.104 8e + 3***	6.121 1e + 3***	1.332 7e + 3***
Q(16)	65.415 7***	27.575 8**	49.716 9***	40.893 3***
No ARCH	17.849 16***	338.316 95***	565.946 28***	151.717 16***
BDS	34.920 81***	19.902 18***	37.325 57***	12.154 69***

说明: *** 表示在 1% 的显著水平下显著, J-B 是 Jarque-Bera 统计量, Q(·) 是 Ljung-Box Q 统计量, No ARCH 是 ARCH 效应的 LM 统计量检验结果.

4 种指数条件收益的波动性状况及统计特征如图 1 和表 1 所示. 从图 1 和表 1 可以观察到: 虽然 4 种指数的条件收益在分布与波动性存在较大的差别, 但是 4 种指数收益率的 J-B 值都在 1% 的显著性水平下非常显著, 都拒绝有效市场假说的正态分布假设; 所有序列的偏度 (skewness) 都在 1% 的水平下显著, 虽然在偏度方向上存在着不同, 但 4 种指数条件收益均呈现出明显的有偏胖尾分布类型特征; Q(16) 统计量检验表明 4 种指数收益都明显拒绝无自相关性; 经过 BDS 统计量检验发现 4 种指数的条件收益均拒绝独立同分布特征, 说明金融条件收益并非是一个同分布序列.

这些统计检验结果进一步证实, 无论是成熟资本市场, 还是新兴市场收益率序列均存在尖峰胖尾、长记忆性等典型事实特征, 也就要求对金融收益描述必须对这些特征进行分析, 从而也表明金融风险管理必须以这些典型事实为约束条件.

2 胖尾分布与波动长记忆约束下的 EVT-VaR 测度方法

2.1 金融市场动态风险测度的基本方法

在金融计量中, 常常假定金融资产条件收益率序列 R_t 满足如下波动形式

$$R_t = u_t + \varepsilon_t = u_t + \sigma_t z_t \quad (2)$$

其中 u_t 为收益率的条件均值 (conditional mean), $u_t = E(R_t | I_{t-1})$, σ_t 为条件波动率 (conditional volatility) $\sigma_t = var(R_t | I_{t-1})$, I_{t-1} 表示到 $t-1$ 日的全部信息集合 (information set), z_t 满足均值为 0, 方差为 1 的独立同分布 (independent identically distribution, *i. i. d.*) 随机变量. 进一步推导有 $z_t = (R_t - u_t) / \sigma_t$, 论文定义 z_t 为标准收益 (standardized return).

第 t 日在 100 q % 的置信水平下对第 $t+1$ 日收益的动态风险 (dynamic VaR) 测度公式^[2]

$$VaR_q^{t+1} = u_{t+1} + \sigma_{t+1} z_q \quad (3)$$

z_q 表示 z_t 分布函数 q 分位数的反函数值 $F_z^{-1}(q)$, $F_z(z)$ 是变量 z_t 序列边际分布函数. 由公式 (3) 可知, 测度金融市场动态风险的关键在于

估计出条件均值 u_t 、条件波动率 σ_t 和标准收益序列 z_t 的 q 分位数值 z_q . 下文将围绕研究主题构造出计量模型来重点讨论这三个变量的计算方法.

2.2 金融收益条件均值与条件波动率的建模方法

条件波动率的长记忆特征是金融风险管理中一个极其重要的典型事实, 有效的风险管理技术必须将其纳入约束框架之中^[5]; 同时, 大量的实证结果也表明, 金融收益的条件波动率 σ_t 时间序列表现出明显的长程记忆性特征. 因此, 引入 Baillie *et al*^[15] 提出的 FIGARCH(p, d, q) 模型, 其数学模型如下

$$\sigma_t^2 = \omega + \beta(L) \sigma_t^2 + (1 - \beta_1(L) - (1 - \varphi(L))(1 - L)^d) \varepsilon_t^2$$

L 是滞后算子, d 是分数协整阶数 ($0 \leq d \leq 1$),

$$\varphi(L) = 1 - \sum_{j=1}^p \varphi_j L^j, \quad \beta(L) = \sum_{j=1}^q \beta_j L^j.$$

分数积分阶数 d 是度量条件波动率长期记忆性参数, FIGARCH 同时具有 GARCH 与 IGARCH 两种模型的条件波动率捕获能力, 当 $d = 0$ 时, FIGARCH 退化为 GARCH 模型; 当 $d = 1$ 时, 模型就变成 IGARCH 模型.

然而, 在金融计量研究中, 一般认为使用一阶模型就能充分反应金融数据的自相关与波动性信息^[33], 因此, 在滞后阶的选取上均选取滞后一阶, 即采用条件波动率 σ_t 的 FIGARCH(1, d , 1) 模型. 其数学模型如下

$$\sigma_t^2 = \omega + \beta_1(L) \sigma_{t-1}^2 + (1 - \beta_1(L) - (1 - \varphi_1(L))(1 - L)^d) \varepsilon_t^2 \quad (5)$$

在金融收益条件均值 u_t 的建模方面, 一般假设其服从自回归的 AR(p) 与自回归移动平均的 ARMA(p, q) 过程等. 仍然假定条件收益的条件均值 \hat{u}_t 满足 AR(1) 过程

$$\hat{u}_t = \phi R_{t-1} \quad (6)$$

在对条件均值与波动模型, 即 AR(1)-FIGARCH(1, d , 1) 模型的参数估计之后, 估计出条件均值 u_t 、条件方差 σ_t , 并计算出标准收益 z_t . 由于无法确定 z_t 分布形式, 而 EVT 就不必假定 z_t 服从的分布形态, 因此考虑运用 EVT 对 z_t 的分布尾部进行建模分析.

2.3 基于 EVT 的金融收益极值尾部建模方法

EVT 作为对金融收益尾部建模的重要方法, 在尾部建模上有两种极值模型, 即传统的分块最大值模型 (the block maxima models, BMM) 和提高门限模型 (peaks over threshold, POT) [4]. 当尾部数据很少时, POT 仍然具有很高有效性, 因而考虑使用 EVT 中 POT 模型对极值尾部进行建模分析.

根据概率论原理, 给定一个阈值 u , 对于超过阈值 u 的变量的分布函数 $F_u(y)$

$$F_u(y) = P(x - u \leq y | x > u) = \frac{F(u + y) - F(u)}{1 - F(u)} \quad (0 \leq x < y + u) \quad (7)$$

根据 Haan [34] 和 Pickands [35] 的证明, 对于充分高的阈值 $u (u \rightarrow \infty)$, 超过阈值 u 的数据近似服从 EVT 中的广义帕累托分布 (generalized pareto distribution, GPD) 簇. GPD 分布簇函数的数学表达式为

$$G_{\xi, \beta} = \begin{cases} 1 - (1 + \xi y / \beta)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - \exp(-y/\beta) & \text{if } \xi = 0 \end{cases} \quad (8)$$

其中 $\beta > 0, 1 + \xi y / \beta > 0$. ξ 是 GPD 的尾部参数 (tail parameter) 或称尾部指数 (tail index), β 是标度参数 (scale parameter). 参数 ξ 描述了概率分布的尾部特征. 如果尾部参数 $\xi > 0$, 则分布为胖尾分布, 如学生 t 分布、Burr 分布以及对数 Gamma 分布等; 如果 $\xi < 0$, 则该分布属于短尾分布, 如均匀分布 (uniform distribution)、Beta 分布; 如果 $\xi = 0$, 则属于正态分布、对数正态分布以及指数分布形式等 [4].

当设定的阈值 u 足够高时, $F_u(y) \approx G_{\xi, \beta}(y) (u \rightarrow \infty)$. 假设 $x = u + y$, 分布的尾部累积概率分布函数 $F(x)$ 为

$$F(x) = (1 - F(u)) G_{\xi, \beta}(x - u) + F(u) \quad (x > u) \quad (9)$$

假设样本为 n , 超过阈值的样本数为 k , 那么, 用 $(n - k) / n$ 代替 $F(u)$, 对 GPD 中的参数 ξ, β 运用 QMLE 进行估计, 则有

$$\hat{F}(x) = 1 - \frac{k}{n} \left(1 + \frac{\hat{\xi}}{\hat{\beta}}(x - u) \right)^{-1/\hat{\xi}} \quad (10)$$

在估计 GPD 参数时, 先要确定出适当的 POT

阈值 u . 这里需要指出的是, u 的选择尤其重要, 根据 GPD 要求, 阈值取值必须相当的高, 以保证方差低, 但是阈值取值过高, 极值数据就很少, 会影响 GPD 估计效果; 相反, 阈值取值过小, 落入尾部的数据多, 又与 GPD 要求有差距. Neftci [8] 把 1.65σ 当作阈值, 超过 1.65σ 的值被当作极值; DuMouchel [32] 认为选择 10% 左右的数据作为极值, 与 GPD 的拟合效果较好.

在这里也需要说明的是, 到目前为止, 还没有一个确定最优阈值的数学理论模型 [9]. 目前, 也有相当一部分学者采用超均值函数 (mean excess function, MEF) 图来确定阈值 [31], MEF 数学模型如下

$$s_n(u) = \frac{\sum_{i=1}^n (Y_i - u)^+}{\sum_{i=1}^n 1_{\{Y_i > u\}}} \quad (11)$$

根据 $(u, s_n(u))$ 点得到 MEF 图, 如果 MEF 图向上倾斜, 那么超过 u 的变量具有 GPD 分布特征; 如果 MEF 图向下倾斜, 超过 u 的变量是短尾分布; 如果 MEF 图为水平线, 超过 u 的变量服从指数分布. 若标准收益的尾部服从 EVT 的 GPD, 当 $q > 1 - k/n$ 时, 得到标准收益序列 q 分位数值为

$$\hat{z}_q = u + \frac{\hat{\beta}}{\hat{\xi}} \left(\left(\frac{1 - q}{k/n} \right)^{-\hat{\xi}} - 1 \right) \quad (12)$$

综合以上推导过程, 就可以计算出第 t 日对第 $t + 1$ 日的动态极值风险 (EVT-VaR) 测度值

$$VaR_q^{t+1} = \hat{u}_{t+1} + \hat{\sigma}_{t+1} \hat{z}_q \quad (13)$$

正如前文所分析的那样, 由于标准收益序列服从的分布不同导致 z_q 取值不同, 从而引起动态风险测度结果存在着差异. 为了对比不同风险测度方法的结果, 本文也对标准收益服从正态分布、学生 t 分布 (student-t distribution) 条件下的风险进行测度分析. 若标准收益序列服从标准正态分布时, $z_q = \Phi^{-1}(q)$, $\Phi(x)$ 为标准正态分布的分布函数; 若标准收益序列服从学生 t 分布时, $z_q = \sqrt{\frac{v-2}{v}} F_T^{-1}(q)$, $F_T(t)$ 表示自由度为 v 的学生 t 分布函数.

在这里, 需要补充说明的是, 由于 SKST 具有

比学生 t 分布更能捕获金融收益分布的有偏形态. 近年来许多学者运用 SKST 分布来测度金融风险并取得了较好的研究效果,也使得 SKST 分布成为金融风险测度研究中使用的一种重要方法^[7,19]. 为了对比基于 EVT 风险测度模型与其它模型的精度与可靠性,也对金融标准收益分布的 SKST 假设进行探讨.

若标准收益序列服从 SKST 时 $\tilde{z}_q = \text{SKST}_{(q; \lambda, \rho)}$. $\text{SKST}_{(x; \lambda, \rho)}$ 概率分布函数如下

$$\text{SKST}(x; \lambda, \rho) = \begin{cases} bc \left(1 + \frac{1}{v+2} \left(\frac{bx+a}{1-\lambda} \right) \right)^{-(v+1)/2} & \text{for } x \leq -\frac{a}{b} \\ bc \left(1 + \frac{1}{v+2} \left(\frac{bx+a}{1+\lambda} \right) \right)^{-(v+1)/2} & \text{for } x > -\frac{a}{b} \end{cases} \quad (14)$$

其中

$$a = 4\lambda c \left(\frac{v-2}{v-1} \right) \quad c = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \sqrt{\pi(v+2)}}$$

$$b = \sqrt{1 + 3\lambda^2 - a^2}$$

λ 是学生 t 分布的偏度系数, ρ 是学生 t 分布的自由度, $\Gamma(\cdot)$ 为 Gamma 分布函数.

2.4 金融市场 EVT-VaR 模型准确性的 Back-testing 检验方法

金融风险管理的关键在于准确有效地测度风险,引入规范的 Back-testing 方法考察风险测度的准确性. 根据风险定义,如果 $VaR_q^{t+1} > R_{t+1}$,就风险模型测度在第 $t+1$ 天是有效的. 否则,就认为风险模型测度是失败的. 基于此,给出一个指示变量 I_t ,

$$I_t \equiv \begin{cases} 1 & \text{if } R_{t+1} > VaR_q^{t+1} \\ 0 & \text{if } R_{t+1} \leq VaR_q^{t+1} \end{cases} \quad (15)$$

根据 Kupiec^[26] 和 McNeil and Frey^[2] 的研究结果,如果 R_t 适合模型,那么 $I_t \equiv \{R_{t+1} > VaR_q^{t+1}\} \sim \text{Binoulli}(1-q)$. 因此 $\hat{\mu} = \sum_{t \in T} I_t \sim \text{Binoulli}(f, 1-q)$, 其中 f 是 n 中个数. 于是,检验零假设 H_0 : 模型测度失败比率与 $1-q$ 接近,也就是 \hat{w} 是来自于

$\text{Binoulli}(f, 1-q)$. 模型测度失败实际比率为

$$p_1 = \frac{f}{n} \quad (16)$$

式(16)中的 f 表示风险模型测度失败次数, n 表示样本总数, p_1 为测度失败比率. 假定比率为 $p_0 = 1-q$. 如果模型测度失败比率与假定的失败比率在统计意义上是一致的,即 $H_0: p_0 = p_1$, 那么就可以认为,在给定的置信水平下,模型能够准确有效地测度风险. 其似然比率的非条件检验统计量为

$$\text{LR}_{uc} = -2 \ln \left[\frac{(1-p_0)^{n-f} p_0^f}{(1-p_1)^{n-f} p_1^f} \right] \rightarrow \chi^2(1) \quad (17)$$

统计量 LR_{uc} 依概率分布服从自由度为 1 的卡方分布. 如果 LR_{uc} 的卡方检验不显著,接受零假设,说明模型可以用来测度市场风险,模型是准确可靠的.

但是,实际上,风险测度失败情况之间也可能是相依的 (dependence),某一次出现测度失败与否可能与前一次测度失败与否的情况是非独立的. 仅使用 Kupiec 的似然比率统计量 LR_{uc} 检验也就存在明显的不足. 基于此,Christoffersen 构造出依概率分布服从自由度为 1 的卡方分布的独立性统计量 LR_{ind} ,其零假设为 H_0 : 某天风险模型测度失败与前一天是独立的.

$$\text{LR}_{ind} = -2 \ln \left[\frac{(1-p_2)^{n_{00}+n_{10}} p_2^{n_{01}+n_{11}}}{(1-p_{01})^{n_{00}} p_{01}^{n_{01}} (1-p_{11})^{n_{10}} p_{11}^{n_{11}}} \right] \rightarrow \chi^2(1) \quad (18)$$

其中 $p_{01} = \frac{n_{01}}{n_{00} + n_{01}}$, $p_{11} = \frac{n_{11}}{n_{10} + n_{11}}$, $p_2 = \frac{n_{01} + n_{11}}{n_{00} + n_{01} + n_{10} + n_{11}}$. n_{00} 表示风险模型当天测度没有失败且前一天也没有失败, n_{01} 表示风险模型当天测度失败且前一天未失败,同理类推.

在 LR_{uc} 和 LR_{ind} 检验的基础上,Christoffersen 构造出一个依概率分布服从自由度为 2 的卡方分布条件检验统计量 LR_{cc} . LR_{cc} 是 LR_{uc} 检验和 LR_{ind} 检验的综合,它不仅考察了模型测度失败比率与假定比率是否一致,而且也考察了失败情况是否是随机出现的,从而具有更强的检验能力. 如果卡方检验的统计量在某一个显著性水平显著,那么模型不能准确测度市场风险,模型缺乏准确性;相反,该模型是准确的. LR_{cc} 统计量如下

$$LR_{cc} = LR_{uc} + LR_{ind} \xrightarrow{d} \chi^2(2) \quad (19)$$

3 实证结果与分析

在 AR(1)-FIGARCH(1, d, 1) 模型的参数估计过程中, 由于无法确定标准收益 z_t 服从哪种分布形式, 因此在模型的参数估计方法上, 没有采用常规的极大似然估计方法 (maximum likelihood

estimation, MLE), 而是运用一种“让数据自己说话”的“伪极大似然估计方法” (quasi maximum likelihood estimation, QMLE) 估计模型参数. 部分参数估计结果见表 2, 从表 2 可以看到, 四种综合指数的 $d \in (0, 1)$, 并且都在 1% 的显著性水平显著不为零, 表明四种指数条件波动率均具有长期记忆性特征, 从而说明构建的模型具有一定的可靠性.

表 2 FIGARCH 模型参数的 QMLE 估计结果

Table 2 Estimation results for FIGARCH model parameters

	SSEC	S&P500	Nasdaq	T-bond
ω	-3.551 1e - 04 ***	-5.034 3e - 04 ***	-4.083 67e - 04 ***	-6.540 3e - 04 ***
d	0.600 8 ***	0.400 3 ***	0.434 6 ***	0.316 3 ***
GARCH(φ_1)	1.226 9e - 05 ***	-1.475 1e - 07	-7.450 9e - 07	0.000 0
ARCH(β_1)	0.448 6 ***	0.370 9 ***	0.396 5 ***	0.287 4 ***

注: *** 表示在 1% 的显著水平下显著.

在这里, 仅以中国沪市上证综指 (SSEC) 为例, 展示金融资产收益的条件波动率 σ_t 与标准收益 z_t 的波动状况. 图 2 为 SSEC 条件波动率与标准收益率的时间序列图. 比较图 1 与图 2 可以较为直观地看出, 标准收益 z_t 较原始收益率 R_t 序列更为平稳. 表 3 出示的是标准收益序列的描述性统计结果, 从表 3 可以看出, 所有的标准收益均值接近为 0, 方差为 1; 所有的标准收益都具有胖尾、尖峰分布特征, 标准收益与原始条件收益一样也非正态分布, 偏度方向也没有发生变化, 说明金融条件收益 R_t 的严重胖尾特征, 必然导致标准收益 z_t 分布的胖尾性; 只有 Nasdaq 指数标准收益都通过了 BDS 检验, 说明并非所有的标准收益都是 *i. i. d.* 序列. 因此, 可以认为标准收益仍然呈现出有偏胖尾且非同分布形态, 运用 EVT 对标准收益的尾部具有理论上的优越性. 基于此, 下文将运用 EVT 对其尾部进行建模分析.

正如前文所述, 采用 MEF 图方法与选择 10% 的最大值作为门槛值 (threshold) 两种方法的结合来选择门槛值. 图 3 是上证综指的标准收益的 MEF 图. 从图 3 可以直观看出, 大约从 1 左右开始, MEF 是斜向上的, 说明其近似服从 GPD 分布. 而选择 10% 的极大值为门槛值, 得到 $u = 1.06$,

也在 1 左右. 综合两种结果, 说明超过 $u = 1.06$ 的收益尾部近似服从 GPD 分布.

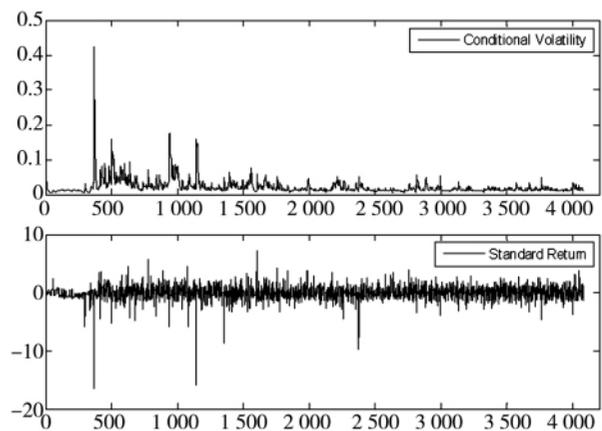


图 2 上证综指条件波动率与标准收益序列

Fig. 2 Conditional volatility and standard return of SSEC

在对金融标准收益尾部的 EVT 建模研究中, 假设尾部极值个数为 k , 门槛值为 u , 运用 QMLE 估计标准收益分布极值尾部的 GPD 参数. 图 4 是上证综指标准收益尾部经验分布与 GPD 分布拟合效果图. 从图 4 可以很直观地看出, 标准收益的 GPD 与经验分布有较好的拟合效果, 说明运用 EVT 中 GPD 分布在金融市场收益尾部的拟合上具有可靠性.

表3 上证综指条件标准收益的描述性统计结果
Table 3 Descriptive statistics for standard return of SSEC

	SSEC	S&P500	Nasdaq	T-bond
Mean	0.008 160	- 0.000 583	0.001 266	- 0.010 784
Standard deviation	0.999 764	1.000 426	0.999 370	0.998 369
Skewness	- 2.499 271 ***	0.364 308 ***	0.397 614 ***	- 0.143 102 ***
Kurtosis (excess)	40.610 634 ***	1.922 121 ***	1.498 426 ***	1.659 070 ***
J-B	2.846 9e + 5 ***	7.185 0e + 2 ***	4.893 2e + 2 ***	4.819 7e + 2 ***
BDS	4.307 78 ***	- 2.675 24 ***	0.237 29	- 2.149 44 **
no ARCH	0.220 14	3.571 10	6.766 72	4.340 84

注: ***表示在1%的显著水平下显著, $Q(\cdot)$ 是Ljung-Box Q统计量, no ARCH效应是LM统计量检验结果.

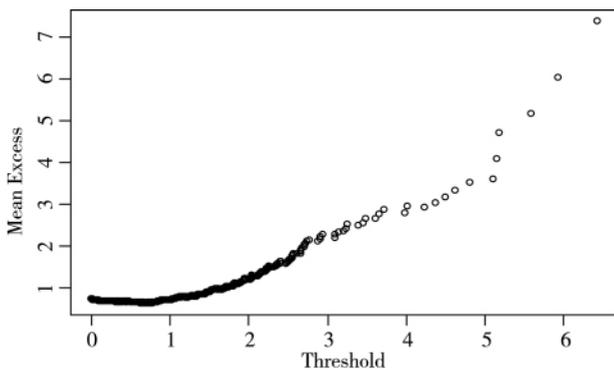


图3 上证综指标准收益的MEF图

Fig. 3 MEF plot of standard return of SSEC

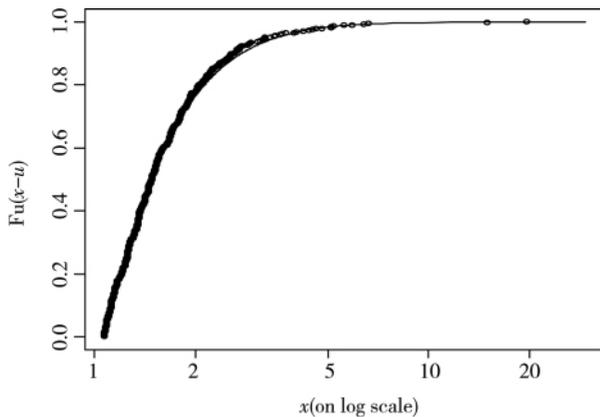


图4 上证综指标准收益尾部经验分布与GPD分布拟合效果图

Fig. 4 Empirical distribution of standard return and GPD plot

在这里,需要说明的是 q 的取值大小,由投资者对待风险的态度以及投资者的经济实力所决定^[2, 8-10]. 由于是对金融市场动态极值风险进行测度,所以选取 $q = 99\%$. 根据式(13)得到动态极值风险序列. 为了更直观地展示测度出动态 EVT 风险序列,仅仅出示第 2 600 - 2 800 天上证综指的动态风险值与条件收益直观图(见图 5). 从图 5 中可以直观看出,风险序列比条件收益序

列要大,只有很少地方收益高于风险值 VaR,直观地说明了模型在测度风险方面具有一定的可靠性.

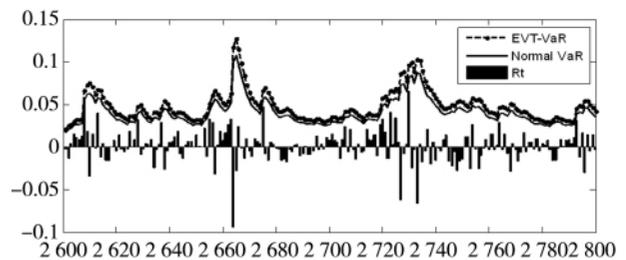


图5 上证综指收益与 VaR 时间序列图(2 600 - 2 800)

Fig. 5 VaR and return time series plot of SSEC (2 600 - 2 800)

对于风险模型的样本内测度的准确性,采用 Back-testing 方法^[4]进行检验. 表 4 是样本内风险测度准确性的 Back-testing 检验结果. 分析表 4 可以得到以下实证结论.

1) 基于金融条件收益率的胖尾分布特征与条件波动率的长记忆性构建的风险测度模型,能够准确测度中国沪市上证综指、标准普尔 500 指数、纳斯达克指数和美国 10 年期国债指数的市场风险,这个结果为金融市场管理当局、金融资产投资者测度市场风险、加强风险管理提供了实证依据.

2) 从标准收益的分布情况来看,无论是新兴市场还是成熟发达市场,所有正态分布假设下的风险模型都没有通过检验,说明正态分布不能有效刻画标准收益的分布,不能准确测度金融市场极端风险,从而说明正态分布假设下的风险测度指标体系会低估风险,导致风险管理失效.

3) 虽然学生 t 分布较正态分布具有更为肥胖的尾部,假设标准收益服从学生 t 分布时也能通

过返回测试检验;但是从 LR_{cc} 的统计量大小来看,有偏学生 t 分布 (SKST) 假设却又比学生 t 分布假设具有更加优越的风险测度能力. 分析原因可能是,与学生 t 分布相比,有偏学生 t 分布 (SKST) 更能捕获金融收益的有偏分布特征,因而表现出更高的精度. 但是,从 LR_{cc} 的统计量大小来看,无论是学生 t 分布假设还是学生 t 分布假设,EVT-GPD 比它们的统计量仍然都要小,说明 EVT-GPD 更能捕获收益尾部的分布特征.

基于以上结果不难看出,有偏胖尾分布(如

有偏学生 t 分布) 比对称正态分布、学生 t 分布条件下构建的风险测度方法更具有可行性,而极值理论 EVT 技术对极值尾部建模又比假定一个确定分布在测度风险方面表现出更大的优越性.

既然如此,在样本内 VaR 测度方面,基于条件收益率胖尾分布与条件波动率长记忆性特征的 EVT-GPD 风险测度模型,具有比本文使用的其它模型更为优越的风险测度能力. 那么,在样本外,此模型是否也具有较好的预测推广能力而使模型更具稳健性呢?

表 4 风险测度准确性的样本内检验结果

Table 4 Back-testing results for VaR measurement in sample

	LR_{uc}	LR_{ind}	LR_{cc}
SSEC-EVT-GPD	3.170 5	1.386 6	4.557 1
SSEC-normal	11.450 8***	0.001 7	11.452 5***
SSEC-student-t	4.248 9	1.078 6	5.327 5
SSEC-SKST	4.130 9	1.076 5	5.207 4
S&P500-EVT-GPD	4.105 0	1.467 2	5.572 2
S&P500-normal	6.589 0***	10.834 3***	17.423 3***
S&P500-student-t	3.934 0	4.588 9	8.522 9
S&P500-SKST	3.765 8	2.873 2	6.639 0
Nasdaq-EVT-GPD	0.542 7	1.080 2	1.623 9
Nasdaq-normal	16.275 2***	10.576 7***	26.852 0***
Nasdaq-student-t	4.058 0	0.022 8	4.080 8
Nasdaq-SKST	4.009 2	0.009 0	4.018 2
T-bond-EVT-GPD	5.272 0	0.392 8	5.664 8
T-bond-normal	11.439 6***	10.396 1***	21.835 7***
T-bond-student-t	6.106 6	0.000 4	6.107 0
T-bond-SKST	5.861 4	0.001 2	5.862 6

注:***表示在 1% 的显著性水平下显著,说明模型无法准确测度风险.

基于此,选择最后 1 000 天样本,仍然运用 Back-testing 中的 LR_{uc} 、 LR_{ind} 和 LR_{cc} 检验指标对 EVT-GPD 模型风险预测推广能力进行检验. 由于篇幅限制,仅出示 EVT-GPD 风险测度模型的检验结果. 结果见表 5. 从表 5 可以发现,在胖尾分布与波动率长记忆性条件下 EVT-GPD 风险测度模型,对于各个市场均在 1% 的显著性水平下都通过

Back-testing 检验,不仅说明了基于 EVT-FIGARCH 模型的 VaR 测度方法在样本外同样具有可靠的推广能力,从而也说明风险模型具有样本外的稳健性,同时还说明了基于 EVT-FIGARCH 的 VaR 模型测度西方成熟发达金融市场与中国新兴市场风险在可靠性与精度方面并未展示出实质性差异.

表5 金融市场风险测度模型样本外推广能力检验
Table 5 Back-testing results for forecasting VaR out-of-sample

	LR _{uc}	LR _{ind}	LR _{cc}
SSEC	2.172 5	1.346 6	3.519 1
S&P500	3.125 0	1.437 2	4.562 2
Nasdaq	0.342 7	1.180 2	1.522 9
T-bond	4.173 0	0.357 8	4.530 8

注: $\chi^2(1)$ 在 1% 的显著水平临界值为 6.635, $\chi^2(2)$ 在 1% 的显著水平临界值为 9.210.

4 结束语

金融资产价格条件收益率展示出有偏胖尾的复杂分布形态以及条件波动率呈现的长记忆性特征,一直是金融学理论研究的重要内容.通过运用 FIGARCH 对金融资产收益条件波动率的长期记忆性建模分析,进而运用 EVT 对具有复杂有偏胖尾的标准收益尾部进行建模来测度风险,然后运用返回测试(back-testing)方法对构建的风险测度模型在样本内测度准确性与样本外的预测推广能力进行了稳健性检验.

通过实证研究得到了一些重要的结果:(1) 金融资产收益与标准收益率,无论是发达成熟的西方市场如纳斯达克股市,还是诸如中国沪市等新兴市场,均表现比正态分布更具有尖峰、肥胖的特征,表明 EMH 条件下的正态分布假设确实与金融市场实际收益分布明显不相符合,运用传统的风险测度方法会低估风险,金融风险管理必须引入非正态分布的有偏胖尾分布形式,以提高风险管理的效果;(2) 四种指数收益条件波动率呈现出异方差与长期记忆性特征,这表明典型事实存在于不同类型与不同国家的金融市场当中,风险管理必须以这些典型事实为约束条件进行;虽然具有有偏分布形态特征的 SKST 在风险测度的精度上表现出了比学生 t 分布更为卓越的测度效果,也

通过了样本内风险测度准确性的 Back-testing 检验,但 EVT 对尾部建模的风险测度模型却表现出比 SKST 更高的准确性与精度,这充分说明了对金融资产的条件收益率更可能是复杂分布形态而非单一分布;(3) 结合 EVT 与 FIGARCH 的风险模型,在测度中国大陆沪市、标准普尔 500 指数、纳斯达克指数及美国 10 年期国债市场的风险准确性方面,无论是样本内的风险测度能力,还是样本外的预测推广能力,都表现出明显的优越性与重要性.

就政府金融市场管理当局和投资者而言,随着金融市场和计算机信息技术的发展,必将出现更多目前尚未认知的“典型事实”,典型事实的增多也必将使风险管理面临“维数灾难”问题,这也必将使金融市场风险管理面临着更大更新的挑战.因此,强化对金融市场典型事实的探讨,建立更为有效度量市场风险的计量模型与方法,采取行之有效措施,防止金融市场出现大起大落,进而防范极端风险灾难危机的发生,维护经济安全,促进经济发展与繁荣,是一项长期而又艰巨的任务.

最后需要说明的是,尽管本文仅仅以四种指数为例,探讨的是成熟市场与新兴金融市场风险测度问题,但研究技术与相关结论,仍然可以为金融市场并非完全有效条件下政府金融市场管理当局、金融机构与投资者在风险管理方面提供一定的决策借鉴.

参考文献:

- [1] Morgan J P. Inc. RiskMetrics Technical Document [M]. 4th Edition, 1996: 58-61.
- [2] McNeil A, Frey R. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach [J]. Journal of Empirical Finance, 2000, (7): 271-300.
- [3] Basle Committee. Overview of the Amendment of the Capital Accord to Incorporate Market Risk [M]. Basle Committee on Banking Supervision, 1996.

- [4] Kevin D. Measuring Market Risk—2nd ed [M]. Chichester: John Wiley & Sons, Ltd. 2005.
- [5] Cont R. Empirical properties of assets returns: Stylized facts and statistical issues [J]. *Quantitative Finance*, 2001, (1): 223–236.
- [6] Aas, Haff. The general hyperbolic skew student's t distribution [J]. *Journal of Financial Econometrics*, 2006, (4): 275–309.
- [7] 魏 宇. 股票市场的极值风险测度及后验分析研究 [J]. *管理科学学报*, 2008, (1): 78–88.
Wei Yu. EVT risk measures and its backtesting in stock markets [J]. *Journal of Management Sciences in China*, 2008, (1): 78–88. (in Chinese)
- [8] Neftci S. Value at risk calculations, extreme events and tail estimation [J]. *The Journal of Derivations*, 2000, (spring): 1–15.
- [9] 黄大山, 刘明军, 卢祖帝. 极值风险 E-VaR 及深圳成指实证研究 [J]. *管理评论*, 2005, 17(6): 16–24.
Huang Dashan, Liu Mingjun, Lu Zudi. Extreme VaR and its empirical analysis of Shenzhen stock index [J]. *Management Review*, 2005, 17(6): 16–24. (in Chinese)
- [10] Ramazan G, Faruk S. Extreme value theory and value-at-risk: Relative performance in emerging markets [J]. *International Journal of Forecasting*, 2004, (20): 287–303.
- [11] Bystrom. Extreme value theory and extremely large electricity price changes [J]. *International Review of Economics and Finance*, 2005, (14): 41–55.
- [12] Huang Y C, Lin B J. Value-at-risk analysis for Taiwan stock index futures: Fat tail and conditional asymmetries in return innovations [J]. *Review of Quantative Finance and Accounting*, 2004, (22): 79–95.
- [13] Fernandez. Risk management under extreme events [J]. *International Review of Financial Analysis*, 2005, (14): 113–148.
- [14] 余素红, 张世英, 宋 军. 基于 GARCH 模型和 SV 模型的 VaR 比较 [J]. *管理科学学报*, 2004, 7(5): 61–65.
Yu Suhong, Zhang Shiyong, Song Jun. Comparison of VaR based on GARCH and SV models [J]. *Journal of Management Sciences in China*, 2004, 7(5): 61–65. (in Chinese)
- [15] Baillie R T, Bollerslev T, Mikkelsen H O. Fractionally integrated generalized autoregressive conditional heteroskedasticity [J]. *Journal of Econometrics*, 1996, (74): 3–30.
- [16] Grau T M N. Modeling Daily Value-at-Risk Using FIGARCH Type Models [R]. University of Alicante, Working Paper, 2002.
- [17] Beine M, Laurent S, Lecourt C. Accounting for conditional leptokurtosis and closing days effect in FIGARCH models of daily exchange rates [J]. *Applied Financial Economics*, 2002, (12): 589–600.
- [18] Wu Ping-Tsung, Shieh Shwu-Jane. Value-at-risk analysis for long-term interest rate futures: Fat-tail and long memory in return innovations [J]. *Journal of Empirical Finance*, 2007, (14): 248–259.
- [19] Hansen B. Autoregressive conditional density estimation [J]. *International Economic Review*, 1994, 35(3), 705–730.
- [20] 王吉培, 旷志平. 偏态 t 分布下 FIGARCH 模型的动态 VaR 计算 [J]. *统计与信息论坛*, 2009, 24(5): 75–79.
Wang Jipei, Kuang Zhiping. Calculation of dynamic VaR based on the skew student t distribution of FIGARCH model [J]. *Statistics & Information Forum*, 2009, 24(5): 75–79. (in Chinese)
- [21] 徐 炜, 黄炎龙. GARCH 模型与 VaR 的度量研究 [J]. *数量经济技术经济研究*, 2008, (1): 120–132.
Xu Wei, Huang Yanlong. Empirical analysis on GARCH-type models and VaR [J]. *The Journal of Quantitative & Technical Economics*, 2008, (1): 120–132. (in Chinese)
- [22] 肖 智, 傅肖肖, 钟 波. 基于 EVT-POT-FIGARCH 的动态 VaR 风险测度 [J]. *南开管理评论*, 2008, (11): 100–104.
Xiao Zhi, Fu Xiaoxiao, Zhong Bo. Dynamic VaR risk measures based on EVT-POT-FIGARCH [J]. *Nankai Business Review*, 2008, (11): 100–104. (in Chinese)
- [23] 林 宇, 魏 宇, 黄登仕. 基于 GJR 模型的 EVT 动态风险测度研究 [J]. *系统工程学报*, 2008, 23(1): 45–51.
Lin Yu, Wei Yu, Huang Dengshi. Study on dynamic risk measurement based on GJR and EVT [J]. *Journal of Systems Engineering*, 2008, 23(1): 45–51. (in Chinese)

- [24]王春峰,张庆翠. 中国股市波动性过程中的长期记忆性实证研究[J]. 系统工程,2004,(1): 78-83.
Wang Chunfeng, Zhang Qingcui. Modeling long memory in Chinese stock market volatility[J]. System Engineering, 2004, (1): 78-83. (in Chinese)
- [25]Hong Y M, et al. Granger causality in risk and detection of extreme risk spillover between financial markets[J]. Journal of Econometrics, 2009, 150(2): 271-287.
- [26]Kupiec P. Techniques for verifying the accuracy of risk measurement models[J]. Journal of Derivatives, 1995, (3): 73-84.
- [27]Christoffersen F. Elements of Financial Risk Management[M]. McGill University and CIRANO, 2002.
- [28]Konstantinos T, Athansios K, Richard A B. Extreme risk and value at risk in the German stock market[J]. The European Journal of Finance, 2007, 13(4): 373-395.
- [29]Brooks C, Persaud G. The effect of asymmetries on stock index return value-at-risk estimates[J]. Journal of Risk Finance, 2003, (4): 29-43.
- [30]Glosten L R, Jagannathan R, Runkle D E. On the relation between the expected value and the volatility of the nominal excess return on stocks[J]. The Journal of Finance, 1993, VOL. XLVIII(5): 193-209.
- [31]Stelios D B, Dimitris A G. Estimation of value-at-risk by extreme value and conventional methods: A comparative evaluation of their predictive performance[J]. Int. Fin. Markets, Inst and Money, 2005, (15): 209-228.
- [32]DuMouchel W M. Estimating the stable index-in order to measure tail thickness: A critique [J]. Annals of Statistics, 1983, (11): 1019-1031.
- [33]Lamoureux, William. Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities[J]. Review of Financial Studies, 1993, 6(2): 293-326.
- [34]Haan L. Residual life time at great age[J]. Annals of Probability, 1974, (2): 792-804.
- [35]Pickand A. Statistical inference using extreme order statistics[J]. The Annals of Statistics, 1975, (3): 119-131.

Study on financial markets dynamic EVT-VaR measuring based on fated-tail distribution and long memory volatility

LIN Yu^{1 2}, HUANG Deng-shi², WEI Yu²

1. Business School, Chengdu University of Technology, Chengdu 610059, China;

2. School of Economics and Management, Southwest Jiaotong University, Chengdu 610031, China

Abstract: For stylized facts characteristics such as fat tail distribution and long memory of conditional volatility for financial return series, this paper used FIGARCH model to conditional volatility, and then used extreme value theory (EVT) model extreme tail of standard returns to measure dynamic risk. We find that all conditional return distributions and their standardized returns are skewed and fat-tailed; Conditional volatilities of financial returns are characters of long memory, EVT are fit extreme tails of financial standardized returns. We introduce EVT and FIGARCH model to calculate dynamic risk for different financial markets, and applied Back-testing to check measurement ability, our result show that this risk measurement method is of accuracy either in sample or out of sample.

Key words: fat-tail distribution; long memory; extreme value theory; dynamic VaR measure; robustness testing