

股票收益率非正态性的蒙特卡罗模拟检验

曹志广¹, 王安兴¹, 杨军敏²

(1. 上海财经大学金融学院, 上海 200433; 2. 复旦大学管理学院, 上海 200433)

摘要:现实金融数据的分布通常表现为厚尾性和不对称性, 因此用正态分布拟合实际金融数据的分布有很大的局限性。文章利用广义双曲线分布的厚尾性和不对称性对1997年1月2日~2003年9月19日的上证综指日收益率分布分别做了正态分布、广义双曲线分布、正态逆高斯分布和双曲线分布的拟合及蒙特卡罗模拟检验, 结果表明广义双曲线分布和正态逆高斯分布可以较好地拟合上证综指日收益率分布。另外, 文章还建立了一个带噪声干扰的线性系统, 对实际的股票收益率并不服从正态分布, 而表现出尖峰厚尾的特征做出了一种可能的解释。

关键词:厚尾性; 不对称性; 广义双曲线分布; 正态逆高斯分布; 蒙特卡罗模拟

中图分类号:F830.91 **文献标识码:**A **文章编号:**1001-9952(2005)10-0034-09

一、引言

金融资产收益率的分布, 对金融资产投资、风险管理等具有重要意义, 吸引了众多学者研究这个问题。现实金融数据的分布通常表现为厚尾性和不对称性(R Mantegna, 1995), 因此用正态分布拟合实际金融数据的分布有很大的局限性, 如在VAR的计算中, 由于金融数据分布的厚尾性, 在正态分布的假设条件下计算VAR会带来较大的误差(José Santiago, 2002)。为此许多学者开始寻求更为合理的分布假设。Mandelbrot提出了用稳定分布代替金融数据正态分布的假设(Mandelbrot, 1963), 但稳定分布的尾部通常比实际分布要更厚, 国内学者对中国证券市场收益率分布的研究也证实了稳定分布的尾部通常比实际分布要更厚(何建敏, 2003)。又有学者提出用截尾的稳定分布作为证券收益率的分布(M Yu Romanovsky, 2002), 截尾的稳定分布实际上是中间部分仍用稳定分布, 两个尾部用比负幂律分布瘦的指数分布来代

收稿日期: 2005-05-08

作者简介: 曹志广(1973-), 男, 湖南永州人, 上海财经大学金融学院讲师;

王安兴(1963-), 男, 湖北襄樊人, 上海财经大学金融学院副教授;

杨军敏(1975-), 女, 安徽合肥人, 复旦大学管理学院博士生。

替的一种混合分布,但运用截尾的稳定分布假设时,应当在何处截尾是一个问题。因此,又有许多学者开始转向广义双曲线分布(Generalized Hyperbolic Distribution),提出了广义双曲线分布(Barndorff-Nielsen, 1977),后来又有学者将其应用到了金融领域(Eberlein, Keller; 1985),由于广义双曲线分布的尾部要比稳定分布的尾部要“薄”,因此广义双曲线分布在金融领域中得到了迅速发展。目前国内对稳定分布在金融领域中的应用研究比较多,但广义双曲线分布在金融领域中的应用研究尚不多见。

二、广义双曲线分布

广义双曲线分布的概率密度函数定义如下: $GH(x, \lambda, \alpha, \beta, \delta, \mu) = g(\lambda, \alpha, \beta, \delta, \mu) (\delta^2 + (x - \mu)^2)^{\lambda/2 - 0.25} K_{\lambda - 0.5}(\alpha \sqrt{\delta^2 + (x - \mu)^2}) e^{\beta(x - \mu)}$, 其中, $K_\nu(\cdot)$ 为第二类变型 Bessel 函数, $g(\lambda, \alpha, \beta, \delta, \mu) = \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi} \alpha^{(\lambda - 0.5)} \delta^\lambda k_\lambda(\delta \sqrt{\alpha^2 - \beta^2})}$ 。并且广义双曲线分布的概率密度函数中的各参数受到以下条件的约束: 当 $\lambda > 0$ 时, $\delta \geq 0$, $|\beta| < \alpha$; 当 $\lambda < 0$ 时, $\delta > 0$, $|\beta| \leq \alpha$; 当 $\lambda = 0$ 时, $\delta > 0$, $|\beta| < \alpha$ 。广义双曲线分布可以派生出正态逆高斯分布(当 $\lambda = -0.5$)和双曲线分布(当 $\lambda = 1$)两个分布。

三、样本描述及参数估计

考虑到涨跌停制度推出后股票收益率变动的差异,本文选取 1997 年 1 月 2 日~2003 年 9 月 19 日共 1616 个交易日的上证综指日对数收益率为样本进行了正态分布、广义双曲线分布、正态逆高斯分布和双曲线分布的拟合及检验。表 1 为上证综指日对数收益率的统计性描述。

对于广义双曲线分布的参数估计,极大似然估计是最为常用的方法,也有学者使用其他估计方法 (Prause, 1999)。本文采用极大似然估计方法对 1997 年 1 月 2 日~2003 年 9 月 19 日的上证综指日对数收益率进行了广义双曲线分布、正态逆高斯分布和双曲线分布的参数估计,结果如表 2 所示(本文所有计算使用 Matlab6.5)。

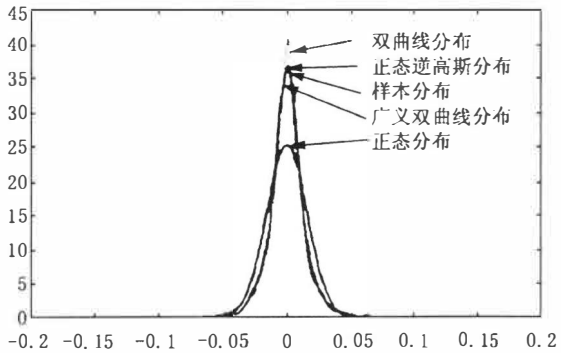


图 1 则对比了广义双曲线分布、正态逆高斯分布、双曲线分布及正态分布与样本分布的拟合情况。

表 1 上证综指日对数收益率的统计性描述

均值	0.0003	偏度	-0.1068
标准差	0.0159	峰度	6.3166
最大值	0.0940	最小值	-0.0933

表 2 参数估计

分布类型	λ	α	β	δ	μ
广义双曲线分布	-0.9899	26.9363	-1.1789	0.0148	0.0006
正态逆高斯分布	-0.5000	44.0234	-1.3787	0.0121	0.0006
双曲线分布	1.0000	96.3572	-2.2519	0.0023	0.0008

四、拟合度的统计检验

为检验广义双曲线分布、正态逆高斯分布、双曲线分布及正态分布与样本分布的拟合情况,我们分别对 Kolmogorov 距离、Kuiper 距离、Anderson-Darling 距离和 FOF 距离(José Santiago, 2002)四种衡量拟合程度的指标进行了检验。四种距离定义分别如下:

$$\text{Kolmogorov 距离: } D_{kol} = \max\{|F_{emp}(x) - F_{theo}(x)|\}$$

$$\text{Kuiper 距离: } D_{kui} = \max\{F_{emp}(x) - F_{theo}(x)\} + \max\{F_{theo}(x) - F_{emp}(x)\}$$

$$\text{Anderson-Darling 距离: } D_{ad} = \max\left\{\frac{|F_{emp}(x) - F_{theo}(x)|}{\sqrt{F_{theo}(x)(1 - F_{theo}(x))}}\right\}$$

$$\text{FOF 距离: } D_{FOF} = \max\left\{\frac{F_{emp}(x) - F_{theo}(x)}{\sqrt{F_{theo}(x)(1 - F_{theo}(x))}}\right\} + \max\left\{\frac{F_{theo}(x) - F_{emp}(x)}{\sqrt{F_{theo}(x)(1 - F_{theo}(x))}}\right\}$$

其中, $F_{emp}(x)$ 、 $F_{theo}(x)$ 分别为样本经验分布函数和理论分布函数。Anderson-Darling 距离和 FOF 距离赋予尾部更多的权重,因此相对于 Kolmogorov 距离和 Kuiper 距离而言,Anderson-Darling 距离和 FOF 距离更偏重于尾部拟合程度的检验。为检验经验分布与理论分布的拟合程度,我们的原假设为经验分布与理论分布一致,并且使用 Monte Carlo 模拟方法分别对正态分布、广义双曲线分布、正态逆高斯分布和双曲线分布对四种距离在 1%、5%、10% 显著水平下的临界值进行了 10000 次模拟计算,计算过程如下:(1) 分别产生一定样本容量下(样本容量分别为 500, 1000, 1500)服从正态分布、广义双曲线分布、正态逆高斯分布和双曲线分布的随机数;(2) 分别计算模拟产生的随机数的经验分布函数值与理论分布函数的差距,分别得到 Kolmogorov 距离、Kuiper 距离、Anderson-Darling 距离和 FOF 距离的 10000 个取值;(3) 确定正态分布、广义双曲线分布、正态逆高斯分布和双曲线分布四种理论分布四种距离分别在 1%、5%、10% 显著水平下的临界值。在模拟计算过程中为了得到服从广义双曲线分布、正态逆高斯分布和双曲线分布的随机数,

我们利用服从 $[0, 1]$ 均匀分布的随机变量使用以下方法得到: $X = \int_{-\infty}^Y f(x) dx$,

其中 X 为服从 $[0, 1]$ 均匀分布的随机变量, $f(x)$ 为需要的服从某一特定分布的随机变量的密度函数,分别令 $f(x)$ 为广义双曲线分布、正态逆高斯分布和双曲线分布的密度函数,通过求解上式中的 Y ,就得到我们所需要的分别服从广义双曲线分布、正态逆高斯分布和双曲线分布的随机数。最后,得到四种理

论分布在各显著水平下的临界值(见表3~表6)。

表3 正态分布检验的临界值表

($\mu=0.0003, \sigma=0.0159$)

样本容量	D_{kol}			D_{kui}			D_{ad}			D_{Fof}		
	显著水平			显著水平			显著水平			显著水平		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
500	0.0693	0.0592	0.0532	0.0878	0.0753	0.0693	6.3011	2.2936	1.4231	6.9688	3.3208	2.4258
1000	0.0494	0.0407	0.0375	0.0627	0.0531	0.0496	4.8783	2.1965	1.3368	5.6188	3.1461	2.2875
1500	0.0411	0.0346	0.0313	0.05043	0.0445	0.0412	4.5713	2.1685	1.2589	5.4620	3.0432	2.1671

表4 广义双曲线分布检验的临界值

($\lambda=-0.9899, \alpha=26.9363, \beta=-1.1789, \delta=0.0148, \mu=0.0006$)

样本容量	D_{kol}			D_{kui}			D_{ad}			D_{Fof}		
	显著水平			显著水平			显著水平			显著水平		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
500	0.0663	0.0588	0.0532	0.0884	0.0775	0.0688	2.8293	1.9039	1.266	4.2930	3.0577	1.8669
1000	0.0442	0.0405	0.0355	0.0586	0.0503	0.0474	1.5720	1.2628	1.2136	2.4819	2.1588	1.6866
1500	0.0433	0.0366	0.0332	0.0549	0.0480	0.0440	1.4147	1.2587	1.2106	2.0762	1.8462	1.6429

表5 正态逆高斯分布检验的临界值表

($\lambda=-0.5000, \alpha=44.0234, \beta=-1.3787, \delta=0.0121, \mu=0.0006$)

样本容量	D_{kol}			D_{kui}			D_{ad}			D_{Fof}		
	显著水平			显著水平			显著水平			显著水平		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
500	0.0744	0.0584	0.0522	0.0848	0.0746	0.0655	3.8685	3.2238	2.5790	5.1172	3.8551	3.2102
1000	0.0506	0.0431	0.0397	0.0643	0.0559	0.0533	2.5790	2.3386	2.2567	3.8345	3.2034	2.8743
1500	0.0443	0.0404	0.0365	0.0554	0.0421	0.0371	2.5471	2.2715	2.1843	3.4206	2.9879	2.6457

表6 双曲线分布检验的临界值表

($\lambda=1.0000, \alpha=96.3570, \beta=-2.2519, \delta=0.0023, \mu=0.0008$)

样本容量	D_{kol}			D_{kui}			D_{ad}			D_{Fof}		
	显著水平			显著水平			显著水平			显著水平		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
500	0.0727	0.0615	0.0572	0.0913	0.0773	0.0692	1.9871	1.8969	1.5315	3.1903	2.4192	1.9232
1000	0.0602	0.0467	0.0409	0.0703	0.0542	0.0534	1.8970	1.5642	1.3684	2.8967	2.2946	1.9026
1500	0.0443	0.0356	0.0317	0.0554	0.0458	0.0418	1.8370	1.4703	1.2646	2.6423	2.1094	1.8771

表7~表10分别是样本正态分布、广义双曲线分布、正态逆高斯分布和双曲线分布在5%显著水平下的检验结果：

表7 样本正态分布的检验结果

样本容量	K_{kol}	D_{kui}	D_{ad}	D_{Fof}
1616	0.0803	0.1606	152.4379	153.4257
5%显著水平下检验结果	拒绝	拒绝	拒绝	拒绝

表8 样本广义双曲线分布的检验结果

样本容量	K_{kol}	D_{kui}	D_{ad}	D_{Fof}
1616	0.0098	0.0193	0.9858	1.4294
5%显著水平下检验结果	不能拒绝	不能拒绝	不能拒绝	不能拒绝

表9 样本正态逆高斯分布的检验结果

样本容量	K_{kol}	D_{kui}	D_{ad}	D_{Fof}
1616	0.0123	0.0237	0.9762	1.9175
5%显著水平下检验结果	不能拒绝	不能拒绝	不能拒绝	不能拒绝

表 10 样本双曲线分布的检验结果

样本容量	K_{kol}	D_{kui}	D_{nd}	D_{FOF}
1616	0.1096	0.1767	0.9940	1.1807
5%显著水平下检验结果	拒绝	拒绝	不能拒绝	不能拒绝

以上结果表明,用广义双曲线分布和正态逆高斯分布可以较好地拟合上证综指日对数收益率分布。

五、实际股票收益率非正态性的解释

对于实际的股票收益率并不服从正态分布通常有以下几种解释:

第一,由于信息的到达并不是均匀的,比如实际的股票市场中好消息倾向于接二连三地到达市场,坏消息也倾向于“祸不单行”地到达市场,因此实际的股票收益率表现出胖尾的特点。

第二,实际的股票市场是一个带噪音干扰的趋势性的非线性系统,分形市场假说认为投资者并非都是理性的,投资者对于信息的反应不是线性的而是非线性的,过去的信息会对市场产生影响,价格并没有反映所有的信息,并不遵循随机游走而是遵循有偏的随机游走,因此实际的股票收益率并不服从正态分布,这也是分形市场理论的一个基本观点(埃德加·彼得斯,2002)。

第三,协同市场假说认为(T Vaga, 1990),资本市场的价格、收益率或风险与收益状况是由“外部的基本经济环境(基本偏倚)”和“投资者的集体情绪(群众行为水平)”两个因素结合所决定的。根据这两者的不同结合,证券市场的状态也随之发生变化。市场可能出现四种情况:一是真正的随机游动状态,即有效市场状态,这时收益率表现为正态分布;二是不稳定过渡状态,即无效市场状态,随着外部环境的逐渐变化,投资者在情绪上的倾斜可能使信息的影响长期延续,这时收益率的概率密度函数变得更宽、更平;三是混沌市场状态,这时基本经济环境是不确定的或没有明显偏倚,但投资者情绪对于群体思维有很强的传导力,群众行为有可能出现大的波动,市场传言或被曲解的信息可能会引起投资者的恐慌,收益率的概率密度函数出现双重底部;四是出现协同市场状态,即强烈的正面(或反面)的基本环境状况与强烈的投资者情绪相吻合,产生巨大的市场协同效应,这时收益率的概率密度函数会出现厚尾。

另外,20世纪90年代兴起的行为金融引入投资者的心理特征对此也提出了相应的解释,认为当市场处于热情高涨的牛市行情时,由于人们的过度自信心理导致对信息的吸收和处理存在倾向性,这时人们往往会忽视“坏消息”,对“好消息”给予更多的关注;当市场处于低落行情时,人们又会出现反应不足的现象,投资者对股票市场上的“好消息”表现出反映滞后的现象,因此实际的股票收益率表现出胖尾的特点(Debon, 1985)。我们认为,实际的股票市场可以用一个带随机干扰的线性系统来近似地刻画(曹志广,2005),并且该系统

能对实际的股票收益率并不服从正态分布而表现出尖峰厚尾的现象做出解释。

考虑只存在一只股票的股票市场,假定:(1)股票的需求量随价格连续地反方向变化;(2)股票的供给量随价格连续地同方向变化;(3)股票价格的变化由股票的需求量和供给量之差决定;(4)每个时刻存在一个股票的期望价格,股票的需求量和供给量由实际价格和期望价格之间差额的大小进行调整;(5)股票的期望价格按照适应性预期变化;(6)随着时间的推移新的信息对市场的需求的影响用一个服从均值为零的正态分布随机变量来表示。记 $D(t)$ 为股票在 t 时刻的需求量, $S(t)$ 为股票在 t 时刻的供给量, $P(t)$ 为股票在 t 时刻的价格, $P_e(t)$ 为股票在 t 时刻的期望价格, $\xi(t)$ 表示 t 时刻的信息对市场的需求的影响。则系统可以表达成为:

$$D' = a(P_e - P) + \xi, S' = b(P - P_e), P' = c(D - S), P'_e = d(P - P_e) \quad (1)$$

其中, $a > 0, b > 0, c > 0, 0 < d < 1$; D', S', P', P'_e 分别表示 t 时刻市场需求、供给、股票价格和期望价格的变化率,即它们各自对时间的导数; $\xi \sim N(0, 1)$ 。

在 $P(0) = 10, P_e(0) = 11, D(0) = 10, S(0) = 9$ 以及 $a = b = 0.5, c = 0.1, d = 0.6, \xi \sim N(0, 1)$ 的条件下,考虑离散时间 $t = 0, 1, 2, \dots, 500$, 则该系统的输出结果及收益率分布如图 2 和图 3 所示,其中图 3 的结果是根据系统模拟产生的 2000 个样本的价格序列得到的。

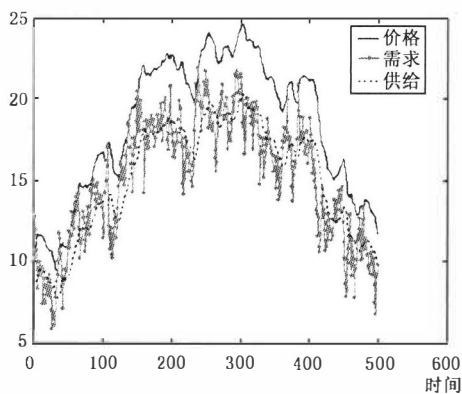


图 2 模拟系统产生的价格、需求与供给时间序列

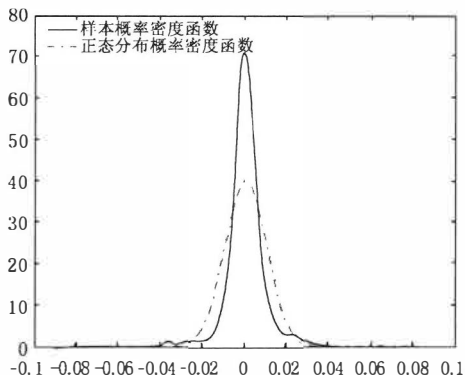


图 3 模拟系统产生的对数收益率的概率密度函数与拟合的正态分布概率密度函数

表 11 模拟系统产生的对数收益率的正态性检验

分布检验	Jarque-Bera 统计量	P 值	检验结果
Jarque-Bera 检验	24.8657	3.9854e-006	拒绝正态分布原假设

表 11 表明模拟系统产生的对数收益率并不服从正态分布,由图 3 也可以看出模拟系统产生的对数收益率与正态分布相比表现出了尖峰厚尾的特征。

对于系统(1)如果我们将去掉其中的随机项 ξ , 即没有新的信息到达系统, 则系统变成一个确定的线性系统。

为避免繁琐的符号表达, 我们同样考虑在 $P(0)=10, P_e(0)=11, D(0)=10, S(0)=9$ 以及 $a=b=0.5, c=0.1, d=0.6$ 的条件下的解的形式, 运用 Laplace 变换及逆变换, 我们得到以下形式的解:

$$P = -\frac{8}{5}e^{-0.3t}\cos(0.1t) - \frac{19}{5}e^{-0.3t}\sin(0.1t) + \frac{58}{5}$$

$$P_e = -\frac{39}{5}e^{-0.3t}\sin(0.1t) - \frac{3}{5}e^{-0.3t}\cos(0.1t) + \frac{58}{5}$$

$$D = \frac{13}{2}e^{-0.3t}\sin(0.1t) + \frac{1}{2}e^{-0.3t}\cos(0.1t) + \frac{19}{2}$$

$$S = -\frac{13}{2}e^{-0.3t}\sin(0.1t) - \frac{1}{2}e^{-0.3t}\cos(0.1t) + \frac{19}{2}$$

所以在上述条件下系统很快达到均衡状态: 供给等于需求, 期望价格与实际价格一致(见图 4)。

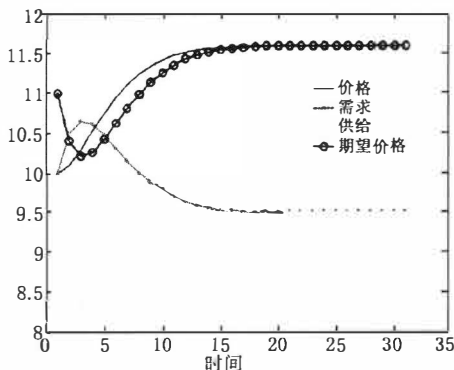


图 4 没有新信息时模拟系统的输出结果

上述系统是一个带噪声干扰的简单线性系统, 粗略地对实际的股票收益率并不服从正态分布, 而表现出尖峰厚尾的特征做出了一种可能的解释。实际的股票市场要比上述系统更复杂, 更可能是一个带噪声干扰的非线性系统。

六、结束语

本文用上证综指日收益率数据从实证角度表明实际的股票收益率并不服从正态分布, 用广义双曲线分布和正态逆高斯分布可以较好地拟合上证综指日对数收益率分布。然而对实际的股票收益率并不服从正态分布的解释学术上存在几种不同的解释, 反映了学术界对股票市场的不同认识, 这将引导我们对股票市场究竟是一个随机系统, 还是一个确定性的非线性系统, 抑或是前面两种系统某种形式的结合做出进一步的思考和求证。

参考文献:

- [1] R Mantegna, H Stanley. Scaling behavior in the dynamics of an economic index [J]. Nature, 1995, (376): 46~49.
- [2] José Santiago, Fajardo Barbachan, Aquiles Rocha de Farias, José Renato Haas Ornelas. Analyzing the use of generalized hyperbolic distributions to value at risk calculations[R]. working paper, 2002, Banco Central do Brasil.
- [3] B Mandelbrot. The variation of certain speculative prices [J]. Journal of Business, 1963, (36): 394~419.
- [5] M Yu Romanovsky. Truncated levy distribution of SP500 stock index distribution of one-share fluctuations in a model space [J]. Physica A, 2000, (287): 450~460.
- [6] Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size [A]. Proceedings of the Royal Society London A, 1977: 401~419.
- [7] E Eberlein, U Keller, K Prause. New Insights into smile, mispricing and value at risk: the hyperbolic model [J]. Journal of Business, 1998, (71): 371~405.
- [8] José Santiago, Fajardo Barbachan, Aquiles Rocha de Farias, José Renato Haas Ornelas. Goodness-of-fit tests focus on VaR estimation[R]. working paper, Banco Central do Brasil, 2002.
- [9] Prause K. The generalized hyperbolic model; Estimation, financial derivatives, and risk measures [A]. Tese de Doutorado, University of Freiburg, 1999.
- [11] T Vaga. The coherent market hypothesis [J]. Financial Analysts Journal, 1990, (46): 36~49.
- [12] Debon Werner F M, Richard H Thaler. Does the stock market overreact? [J]. Journal of Finance, 1985, (40): 793~808.
- [4] 何建敏, 朱林, 常松. 中国股票市场价格波动的尺度特性[J]. 中国管理科学, 2003, (2): 1~5.
- [10] 埃德加·彼得斯. 分形市场分析[M]. 北京: 经济科学出版社, 2002.
- [13] 曹志广, 杨军敏, 王其藩. 证券市场价格行为系统动力学研究[J]. 管理科学学报, 2005, (1): 62~72.

The Test of non-normal Distribution of Stock Returns with Monte Carlo Simulation and The Explanation

CAO Zhi-guang¹, WANG An-xing¹, YANG Jun-min²

(1. School of Finance, Shanghai University of Finance and Economics, Shanghai 200433, China; 2. School of Management, Fudan University, Shanghai 200433, China)

Abstract: The distribution of financial data is usually (下转第 52 页)

The Shock Effect of Short Selling Mechanism to Emerging Market: Evidence from Hong Kong Stock Market

LIAO Shi-guang¹, ZHANG Zong-xin²

(1. Management School, Shanghai Jiaotong University, Shanghai 200052, China;

2. Institute of Financial Studies, Fudan University, Shanghai 200433, China)

Abstract: As an important trading system in the modern securities market, the short selling mechanism plays great roles in perfecting the market functions. But may the allowance of short selling mechanism in the emerging market enhance the market volatility? The paper discusses the relationship between the market volatility and short sales with the data from Hong Kong stock market, and the result suggests that the effect of short selling on the whole market volatility is very complicated, and the final result may depend on the types of short sellers, trading strategies of investors and the openness of trading information. If only with the consideration of speculative short sales, the short sales can play an important role in buffering stock price volatility in stead of aggregating it.

Key words: short selling; trading mechanism; stock price volatility

(责任编辑 喜 雯)

(上接第 41 页)

fat-tailed and asymmetric, so it is not proper to fit the real financial data by normal distribution. This paper fits the returns of Shanghai Composite using generalized hyperbolic distribution and tests its goodness-of-fit. The results of test show that the daily log-return of Shanghai Composite can be well described by generalized hyperbolic distribution and normal inverse Gaussian distribution. The fact of non-normal distribution can be well explained by a simple linear system with noise.

Key words: fat-tailed; asymmetry; generalized hyperbolic distribution; normal inverse Gaussian distribution; Monte Carlo simulation

(责任编辑 喜 雯)