稳定分布的参数估计

顾娟 茆诗松

2012-09-12 10:25:00 来源：《应用概率统计》2002年第4期

内容提要：由于金融数据经常具有“高峰厚尾”现象，所以用稳定分布去拟合。但由于稳定分布没有密度函数显式，而且可能一阶矩或二阶矩又不存在，因此稳定分布的参数估计问题用经典方法很难处理。本文利用类似Duffie和Singleton(1993)的模拟矩法(SMM)的想法，构造了一种新的参数估计方法，并得到该估计的强相合性结果。最后举了一个实际的例子，分析了深圳成分指数的情况。

关键词：稳定分布 模拟矩法(SMM) 强相合性 特征函数

作者简介：顾娟广发证券股份有限公司博士后工作站，广州510075；茆诗松华东师范大学统计系，上海200062

引言

金融资产的收益是大量在实践上连续到来的信息和个人决策的结果。根据中心极限定理，如果经适当的平移和尺度变换，大量的独立同分布的随机变量和的极限分布一定是稳定分布族中的某一分布。因此很自然地认为，如果资产收益的积累是可加的，则应近似服从稳定分布；如果资产收益的累积是相乘的，则应近似服从对数稳定分布。

对于稳定分布的研究从本世纪初就引起了大家的注意。正态分布属于稳定分布族。同时也是大家最为熟悉也最易处理的分布，所以在处理一般问题时大都认为资产收益的分布是近似服从正态分布或对数正态分布。在处理股票价格的经验研究中，Mandelbrot[1]和Fama[2]最先发现股票价格的对数变化明显偏离了正态分布，尤其是非常小和非常大的价格变化比在正态假设出现的频率明显偏大、对这类问题用非高斯的稳定分布去拟和优于正态分布。（近年来，很多金融数据的经验研究发现，金融数据具有一些普遍的规律，其中与平稳直接相关的就是“高峰厚尾”现象，这种现象现在已得到广泛的认识）这之后，很多统计学，经济，金融学等方面的学者对稳定分布和资产收益之间的关系进行了研究。

我们考察了深圳成分指数从1994年1月3日到1998年11月2日的日收益率情况，发现经过作收益率变化后，日收益率的分布形态明显具有这种“高峰厚尾”现象。其中表1是回收益率的常用描述性统计量的估计值，其中的曲线是用正态分布拟和的结果。可以看出在上述时期，深圳市场的日平均收益率大于0，折算成年收益率大约为25%左右。考虑厚尾的衡量指标常常用峰度，即四阶矩。由于正态分布的峰度为3，而该数据的峰度为17.7618，说明该数据的峰度远大于正态分布的峰度，这正是“高峰”说法的来源。另外，数据比正态分布不仅峰度高，而且两端数据出现的频率也较大。所以我们试图用稳定分布去拟和深圳数据。

表1 深成指日收益率统计量的估计

均值	标准差	偏度系数	峰度系数
0.000671	0.027	1.3055	17.7618

对于稳定分布，没有显示的密度函数或分布函数形式。Levy(1924)得到了稳定分布特征函数的显示表达。经过Zolotarev[3]的再参数化变成极点形式，得到了目前比较常用的稳定分布的特征函数的表达式。

$$\log(\phi_x(t)) = irt - |t|^{\alpha}\delta^{\alpha}\exp\{-i\eta_{\alpha,\beta}\cdot \mathrm{sgn}(t)\}$$

其中：$\alpha \in (0,1)\cup(1,2]$，$\beta \in [-1,1]$，$r \in (-\infty,+\infty)$，$\delta \in (0,+\infty)$，$\eta_{\alpha,\beta} = \beta\cdot\min(\alpha, 2-\alpha)\cdot\pi/2$，$\mathrm{sgn}(t) = |t|/t$。

由于稳定分布没有密度函数显式，因此对它的研究比普通的密度函数要困难得多。但人们还是通过很多方法得到了它的一些性质。比如：它的密度函数是连续的，并且是单峰的；它的小于α阶矩存在，大于α阶矩不存在等等。

这些对稳定分布的研究，到七十年代有很大的进展。但由于没有显示的密度函数形式，而且可能一阶矩或二阶矩又不存在，因此稳定分布的参数估计问题受到了很大的阻碍，对平稳原理的进一步推广和使用也受到了限制。八十年代对稳定分布的研究速度逐渐慢了下来。直到九十年代，由于计算机，计算方法和计算速度有了突飞猛进的发展，世界经济的发展也日益繁荣，对稳定分布的研究又引起了人们极大的兴趣。对稳定分布的估计也有了一些新方法。

本文利用类似Duffie和Singleton(1993)的模拟矩法(SMM)的想法，构造了一种新的参数估计方法，并研究了估计的一些性质。这个内容放在一、二、三给出了模拟结果，并用深圳成分指数的例子说明。

一、参数估计方法

记稳定分布为S(α，β，ɣ，δ)，由四个参数所决定。其中ɣ为位置参数，δ为尺度参数。若随机变量x～S(α，β，ɣ，δ)，则[(x-ɣ)/δ]～S(α，β，0, 1)。α，β是形状参数，其中α称为特征指数，决定分布函数尾部的厚薄，α离2越远，说明分布的尾部比正态尾部越厚，比正态峰度越大，当α=2时，就是正态分布。β称为偏度系数，当β=0时，分布函数就是对称形式。在稳定分布族中，我们最熟悉的就是当α=2时的分布，也即正态分布。另外当α=1, β=0时，稳定分布就是柯西分布。经过标准化变换后的稳定分布S(α，β，0. 1)称为标准稳定分布，简记为S(α，β)。我们下面考虑的是标准稳定分布的参数估计问题。

我们的想法非常类似模拟矩方法(Simulated Moment Method，简记为SMM)。因此首先介绍一下SMM方法。SMM方法是Duffie和Singleton(1993)提出的，其想法是：由于某些原因，直接估计参数比较困难，可以用先给定的任意参数值去产生模拟样本，而那个使模拟样本的样本矩与观察样本矩之间距离最小的那个给定的参数值就作为此分布的参数估计值。

可以看出实现SMM方法需要两个条件。一是给定参数值后要能产生来自此参数值对应的分布的样本；第二是用什么样本矩。对于稳定分布，第一个问题可以由定理1保证。

定理1[4]若随机变量ω服从标准指数分布，ω～E(1), v服从[-1/2, 1/2]上的均匀分布，v～U(-1/2, 1/2)，则随机变量x=t[, α，β](v)• ω[a-1/a]服从标准稳定分布S(α，β)。其中：

$$t_{\alpha,\beta}(v) = \left[\frac{\sin(\pi\alpha v + \eta_{\alpha,\beta})}{\cos(\pi v)}\right]\cdot\left[\frac{\cos(\pi v)}{\cos(\pi(\alpha-1)v + \eta_{\alpha,\beta})}\right]^{\alpha-1/\alpha}$$

$$\eta_{\alpha,\beta} = \beta\cdot\min(\alpha, 2-\alpha)\pi/2$$

这样根据定理1，对给定的α,β值α^*,β^*，我们可以产生来自$S(\alpha^*,\beta^*)$的样本：

$$x_1^*(\alpha,\beta), x_2^*(\alpha,\beta), \cdots, x_n^*(\alpha,\beta)$$

第二个问题是关于矩的选择。由于稳定分布的一阶，二阶矩可能不存在，因此我们不能拿一阶或二阶样本矩之间的距离来作评价参数估计好坏的标准。实际上我们用两个样本经验分布函数之间的最大距离来选择参数，即：

$$\hat{\theta}_{n_1,n_2} = \operatorname{Arg}\min_{\theta^*}\left\{\sup\left|\hat{F}_{n_1,\theta_0}(x) - \hat{F}_{n_2,\theta^*}(x^*)\right|\right\},\quad 记\ \theta = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$

其中 $\hat{F}_{n_1,\theta_0}(x)$ 是实际观测样本的经验分布函数，$\hat{F}_{n_2,\theta^*}(x^*)$ 是给定 α^*,β^* 时模拟样本产生的经验分布函数。$\hat{\alpha} = \operatorname{Argmin} f(x,\alpha)$ 指 $\hat{\alpha}$ 是使得 $f(x,\alpha)$ 最小的参数值。我们用上述距离的理由是:判断两个经验分市函数(即两组样本)是否来自同一总体，可以用 CMNPHOB 检验[5]。若 $F_{n_1}(x),G_{n_2}(x)$ 分别是来自总体分布为 F 和 G 的经验分布函数，当原假设:$F = G$ 成立时，$F_{n_1}(x)$ 和 $G_{n_2}(x)$ 应非常接近。因此统计量:

$$S_{n_1,n_2} = \sup_{-\infty<x<\infty}\left|F_{n_1}(x) - G_{n_2}(x)\right|$$

很小时，不拒绝原假设。也就是 S_{n_1,n_2} 值越小，原假设成立的可信度越大，所以我们认为使得 S_{n_1,n_2} 值最小的模拟样本与观察样本来自同一总体。若 F 和 G 为同一分布具有不同参数的两个分布 F_{θ_1} 和 F_{θ_2}，则把上述距离记为 $S_{n_1,n_2}(\theta_1,\theta_2)$，则:

$$\hat{\theta}_{n_1,n_2} = \operatorname{Arg}\min_{\theta^*} S_{n_1,n_2}(\theta_0,\theta^*)$$

二、参数估计的大样本性质

通过上述方法得到的参数估计 $\hat{\theta}_{n_1,n_2}$，我们可以得到它的大样本性质。

定理 2 若存在 $\lambda>0$，使当 $n_1\to\infty, n_2\to\infty$ 时，$n_1/n_2<\lambda$ 和 $n_2/n_1<\lambda$，则当 $n_1\to\infty, n_2\to\infty$ 时，$\hat{\theta}_{n_1,n_2}$ 是 θ 的强相合估计。

证明：对于统计量 S_{n_1,n_2}，当 n_1,n_2 很大时，CMNPHOB 证明了下述极限定理[5]：

当 $F\equiv G$ 连续，且存在 $\lambda>0$，使当 $n_1\to\infty, n_2\to\infty$ 时，始终保持 $n_1/n_2<\lambda$ 和 $n_2/n_1<\lambda$。则当 $n_1\to\infty, n_2\to\infty$ 时，有

$$\sqrt{\frac{n_1 n_2}{n_1+n_2}}S_{n_1,n_2}\xrightarrow{L}k(x)$$

其中

$$k(x)=\begin{cases}0, & x\leqslant 0;\\ \sum\limits_{i=-\infty}^{\infty}(-1)^i\exp(-2i^2x^2) & x>0\end{cases}$$

当 $F\equiv G$ 时，记 $\sqrt{(n_1n_2)/(n_1+n_2)}S_{n_1,n_2} = T$，即 $T\xrightarrow{L}k(x)$。

可算得:

$$\begin{aligned}
\mathrm{E}(T) &= \int_0^\infty t\,\mathrm{d}(k(t))\\
&= \int_0^\infty\sum_{i=-\infty}^{\infty}(-1)^i t(-4i^2t)\mathrm{e}^{-2i^2t^2}\,\mathrm{d}t\\
&= \sum_{i=-\infty}^{\infty}(-1)^i\int_0^\infty -4i^2t^2\mathrm{e}^{-2i^2t^2}\,\mathrm{d}t\\
&= 2\sum_{i=1}^{\infty}(-1)^i\int_0^\infty -4i^2t^2\mathrm{e}^{-2i^2t^2}\,\mathrm{d}t\\
&= \sum_{i=1}^{\infty}(-1)^i\left[t\mathrm{e}^{-2i^2t^2}\Big|_0^\infty - \int_0^\infty \mathrm{e}^{-2i^2t^2}\,\mathrm{d}t\right]
\end{aligned}$$

$$= \sqrt{2\pi}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots\right) = \sqrt{2\pi}\ln 2$$

$$\mathrm{E}\,(T^2) = \int_0^\infty \sum_{i=-\infty}^{\infty} (-1)^i t^2 (-4i^2 t) \mathrm{e}^{-2i^2 t^2} dt$$

$$= 2\sum_{i=1}^{\infty} (-1)^i \left[t^2 \mathrm{e}^{-2i^2 t^2} \Big|_0^\infty - \int_0^\infty t \mathrm{e}^{-2i^2 t^2} \mathrm{d}t \right]$$

$$= \sum_{i=1}^{\infty} (-1)^i \frac{1}{2i^2} = \frac{\pi^2}{24}$$

当 $F \equiv G$ 时，$S_{n_1,n_2} = \sqrt{(n_1+n_2)/(n_1 n_2)}\,T$，

$$\mathrm{E}\,(S_{n_1,n_2}) = \sqrt{\frac{n_1+n_2}{n_1 n_2}}\sqrt{2\pi}\ln 2$$

$$\mathrm{E}\,(S_{n_1,n_2})^2 = \frac{n_1+n_2}{n_1 n_2}\frac{\pi^2}{24}$$

根据切贝雪夫不等式

$$\mathbf{P}(\left|S_{n_1,n_2} - 0\right| > \varepsilon) < \frac{\mathrm{E}\,(S_{n_1,n_2})^2}{\varepsilon^2}$$

所以

$$\lim_{n_1,n_2\to\infty} S_{n_1,n_2} = 0 \quad \text{a.s.}$$

即

$$\lim_{n_1,n_2\to\infty} S_{n_1,n_2}(\theta_0,\theta_0) = 0 \quad \text{a.s.} \tag{1}$$

当 $F \not\equiv G$ 时，$\lim_{n_1,n_2\to\infty} S_{n_1,n_2} > 0$ a.s.，即当 $\theta^* \neq \theta_0$ 时，

$$\lim_{n_1,n_2\to\infty} S_{n_1,n_2}(\theta_0,\theta^*) > 0 \quad \text{a.s.}$$

也可以描述为：

$$\forall \delta > 0, \text{当} \left|\theta^* - \theta_0\right| > \delta, \exists\, \varepsilon > 0, \lim_{n_1,n_2\to\infty} S_{n_1,n_2}(\theta_0,\theta^*) > \varepsilon, \quad \text{a.s.} \tag{2}$$

下面证明 $\lim_{n_1,n_2\to\infty} \hat{\theta}_{n_1,n_2} = \theta$，a.s.

反证法：$\exists A, p(A) > 0$，在 A 上，有 $\lim_{n_1,n_2\to\infty} \hat{\theta}_{n_1,n_2} \neq \theta$，即 $\exists \delta_0$，对 $\forall i_1, i_2$，$\exists N_1 > i_1$，$N_2 > i_2$ 有 $\left|\hat{\theta}_{N_1,N_2} - \theta_0\right| > \delta_0$。

所以由（2）有，在 A 上 $\exists \varepsilon_0 > 0$，$\lim_{n_1,n_2\to\infty} S_{n_1,n_2}(\theta_0,\hat{\theta}_{N_1,N_2}) > \varepsilon_0$，所以对 $r_0 < \varepsilon_0$，$\exists k_1, k_2$，当 $n_1 > k_1, n_2 > k_2$，$S_{n_1,n_2}(\theta_0,\hat{\theta}_{N_1,N_2}) > \varepsilon_0 - r_0$。

又由（1）有

对 $\varepsilon_0 - r_0$，$\exists l_1, l_2$，当 $n_1 > l_1, n_2 > l_2$，$S_{n_1,n_2}(\theta_0,\theta_0) < \varepsilon_0 - r_0$

取 $m_i = \max\{k_i, l_i\}, i = 1,2$，当 $n_1 > m_1, n_2 > m_2$ 时，在 A 上

$$S_{n_1,n_2}(\theta_0,\hat{\theta}_{N_1,N_2}) > \varepsilon_0 - r_0, \quad S_{n_1,n_2}(\theta_0,\theta_0) < \varepsilon_0 - r_0$$

因为 i_1, i_2 的任意性，对 $i_1 = m_1, i_2 = m_2$ 有 $N_1 > m_1, N_2 > m_2$，所以有

$$S_{N_1,N_2}(\theta_0,\hat{\theta}_{N_1,N_2}) > \varepsilon_0 - r_0, \quad S_{N_1,N_2}(\theta_0,\theta_0) < \varepsilon_0 - r_0$$

所以 $S_{N_1,N_2}(\theta_0,\hat{\theta}_{N_1,N_2}) > S_{N_1,N_2}(\theta_0,\theta_0)$。这与 $S_{N_1,N_2}(\theta_0,\hat{\theta}_{N_1,N_2}) = \min_{\theta^*} S_{N_1,N_2}(\theta_0,\theta^*)$ 矛盾。这样就证明了 $\hat{\theta}_{n_1,n_2}$ 的强相合性。

三、模拟结果

我们从均匀分布U(0,1)，指数分布E(1)中分别产生了n个随机数。令α=1.7，β=0.9，根据定理1产生了n个来自稳定分布S(1.7,0.9)的随机数，这样就得到了经验分布函数

$F_{n_1,\theta_0}(x)$。然后令 $\alpha=1,\beta=-1$ 开始，每次产生一个经验分布函数 $F_{n_1,\theta^*}(x)$，求出 $S_{n_1,n_2}(\theta_0,\theta^*)$；按照搜索步长为0.01，使 S_{n_1,n_2} 最小的原则搜索，直到 $\alpha=2,\beta=1$ 结束。当 $n=500$ 时，第一次得到结果为 $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1.67\\ 0.92\end{pmatrix}$，第二次得到结果为 $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1.65\\ 0.85\end{pmatrix}$，第三次得到结果为 $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1.7\\ 0.885\end{pmatrix}$，第四次得到结果为 $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1.73\\ 0.89\end{pmatrix}$。取四次模拟结果的平均值作为估计，得到最后模拟的估计值为 $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1.688\\ 0.886\end{pmatrix}$。

最后对深圳成分指数从1994年1月3日到1998年11月2日的1187个数据进行分析。首先对数据作收益率变换，得到1185个数据，根据Fama(1971)的方法的得到尺度参数的估计，然后用均值和尺度估计作标准化变化，这时得到的数据可认为是来自某一标准稳定分析。用标准化后的数据得到经验分布函数

$F_{1185,\theta_0}(x)$。然后令 $\alpha=1,\beta=-1$ 开始，n_1 取为数据的样本量 1185，每次产生一个经验分布函数，$F_{n_1,\theta^*}(x)$，同样方法求出 $S_{n_1,n_2}(\theta_0,\theta^*)$，按照搜索步长为0.01，使 S_{n_1,n_2} 最小的原则搜索，直到 $\alpha=2,\beta=1$ 结束。按上述方法我们作了三次模拟，第一次得到结果为 $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1.31\\ -0.11\end{pmatrix}$，第二次得到结果为 $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1.33\\ -0.1\end{pmatrix}$，第三次得到结果为 $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1.30\\ -0.12\end{pmatrix}$，第四次得到结果为 $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1.32\\ -0.06\end{pmatrix}$，取四次模拟结果的平均值作为估计，得到最后模拟的估计值为 $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1.315\\ -0.098\end{pmatrix}$。

Peters(1993)曾对从1989年到1992年标准普尔S&P500指数作过计算，发现尾部系数α在1.6至1.9之间。这说明深圳成指收益率的变化与国外股市之间有明显不同之处，差别主要在深成指的峰度更高，尾部更厚。这说明出现收益率暴涨与暴跌的概率比国外更大。另外α估计值明显小于2，其显然不是正态分布，这说明出现收益率暴涨与暴跌的概率比正态场合出现的概率也大。而且分布略右偏，说明深圳市场出现收益率较大跌幅的概率比出现较大升幅的概率大。

参考文献：

1Mandelbrot,B.,The Pareto-Levy law and the distribution of income,Internat.Econom.Rev.,1(1960),79-106.

2Fama,E.F.,Mandelbrot and the stable Paretian Hypothesis,J.Business,36(1963),420-429.

3Zolotarev,V.M.,Mellin-Stietjes transforms in probability theory,Theory Probab.Appl.,2(1957),433-460.

4Luc Devroye,Non-Uniform Random Variate Generation,Springer-Verlag New York Inc.,1986.

5陈希孺，柴根象编著，非参数统计教程，华东师范大学出版社，1986.

6Peters,E.E.,Fractal Market Analysis,A Willy Finance Edition,1993.

7Fama,E.F and Roll,R.,Parameter estimates for symmetric stable distribution,J.A.S.A.,67(334)(1971),331-338.

责任编辑：夏雨

文档附件：

隐藏评论

用户昵称： (您填写的昵称将出现在评论列表中) 匿名

请遵纪守法并注意语言文明。发言最多为2000字符（每个汉字相当于两个字符）

2351

发表