《应用概率统计》版权所有

应用简报

基于稳定分布的AR(1)模型的单位根检验 *

盛子宁　林倍羽　张世斌

(上海海事大学数学系, 上海, 201306)

摘　要

本文考虑误差项为稳定分布的一阶自回归过程$Y_t = \beta Y_{t-1} + \epsilon_t$ $(t = 1, 2, \ldots, N)$的单位根检验, 其中ϵ_t是服从稳定分布的随机误差, β是自回归参数. 若$\beta = 1$成立, 则当$N \to \infty$时, $N(b_N - 1)$的极限分布可表示为Lévy过程的一个泛函形式, 其中b_N为β的最小二乘估计. 因为该形式不依赖于除特征指数α以外的多余参数, 可把$N(b_N - 1)$作为检验原假设$H_0 : \beta = 1$的检验统计量. Chan和Tran (1989)通过直接模拟$N(b_N - 1)$, 给出$N(b_N - 1)$的经验分位数表. 但$N(b_N - 1)$的取值与Y_t有关, 给使用带来影响. 本文构造了一个与Y_t的取值无关的随机变量$E_{N,n}$, 证明了$E_{N,n}$与$N(b_N - 1)$有相同的极限分布. 通过模拟$E_{N,n}$, 得到$N(b_N - 1)$的经验分位数表. 最后, 通过三个数值例子说明了方法的有效性.

关 键 词: 一阶自回归, 单位根检验, 稳定分布, 重尾分布, 经验分位数.

学 科 分 类 号: O212.3.

§1. 引　　言

考虑一阶自回归序列

$$Y_t = \beta Y_{t-1} + \epsilon_t, \qquad Y_0 = 0, \tag{1.1}$$

其中ϵ_t $(t = 1, 2, \ldots, N)$为服从稳定分布(stable distribution)的独立同分布随机误差, β是自回归参数.

稳定分布是一类适用范围最广泛的概率分布. 除了正态分布, 其他的稳定分布都具有尖峰和重尾的特征. 它被广泛用于经济金融领域, 刻画股票价格等数据所普遍存在的高峰厚尾现象. 如果$|\beta| < 1$, 则Y_t平稳(或趋势平稳); 如果$\beta = 1$, Y_t是一阶单整非平稳序列. 因此, 判断一个序列是否平稳, 可以通过检验β是否严格小于1来实现. 即考虑原假设为$H_0 : \beta = 1$, 备择假设为$H_1 : \beta < 1$的假设检验问题.

β的最小二乘估计量为

$$b_N = \Big(\sum_{t=1}^{N} Y_{t-1}^2 \Big)^{-1} \sum_{t=1}^{N} Y_t Y_{t-1}. \tag{1.2}$$

*国家自然科学基金(10901100)资助.

本文2013年1月8日收到, 2013年2月28日收到修改稿.

当ϵ_t为重尾分布时, b_N的渐进性质已有相当多的研究. $N(b_N-1)$的极限分布可表示为Lévy过程的一个泛函形式, 且其形式不依赖于除特征指数α以外的多余参数. 参见Hannan和Kanter (1977), DuMouchel (1983), McCulloch (1978)以及Knight (1987).

Chan和Tran (1989)把$N(b_N-1)$作为检验原假设$H_0:\beta=1$的检验统计量. 通过模拟ϵ_t, 由(1.1)式, 得到Y_t的值, 利用(1.2)式计算b_N, 从而产生$N(b_N-1)$的经验分位数表. 由此产生的经验分位数表与模型(1.1)有关, 并受序列长度的影响. 故用此方法产生的经验分位数表存在一定的局限性. 在实际使用中受到限制.

本文构造了随机变量$E_{N,n}$, 证明了$E_{N,n}$与$N(b_N-1)$有相同的极限分布. 模拟计算了$E_{N,n}$的经验分位数表, 以此作为$N(b_N-1)$的经验分位数表, 此经验分位数表与模型(1.1)无关.

本文后面的安排如下: 第二节给出一些与稳定分布有关的结论. 第三节给出了$E_{N,n}$的定义及其性质, 模拟计算了$E_{N,n}$的经验分位数表. 第四节对三个实际例子做了单位根检验.

§2. 预备知识

定义 2.1 (武东, 汤银才, 2007) 称随机变量X服从稳定分布$S(\alpha,\beta,\gamma,\delta)$, 若$X$的特征函数为

$$\phi(u)=\begin{cases}\exp\Big(-\gamma^\alpha|u|^\alpha\Big[1-i\beta\tan\dfrac{\pi\alpha}{2}(\mathrm{sign}u)\Big]+i\delta u\Big), & \alpha\neq 1;\\ \exp\Big(-\gamma^\alpha|u|^\alpha\Big[1+i\beta\dfrac{2}{\pi}(\mathrm{sign}u)\log(|u|)\Big]+i\delta u\Big), & \alpha=1.\end{cases}$$

一个一般的稳定分布需要四个参数来描述: 特征指数$\alpha\in(0,2]$用来衡量概率密度函数尾部的厚度. α的值越小, 其尾部就越厚, 当$\alpha=2$时, 该分布即为正态分布; 不对称参数$\beta\in[-1,1]$, 当$\beta=0$时, 该分布是对称的稳定分布; 尺度参数$\gamma>0$; 位置参数$\delta\in R$.

这里, 仅讨论标准化的稳定分布$S(\alpha,0,1,0)$, 此时的随机数只与稳定分布四个参数中的特征指数α有关, 非标准化的稳定分布的具体内容参见Chambers, Mallows和Stuck (1976).

下面先介绍几个本文后面用到的结果.

引理 2.1 (Chan和Tran, 1989) 令

$$(U_N(t),V_N(t))=\Big(c_\alpha^{-1}\sum_{i=1}^{[Nt]}\varepsilon_i,c_\alpha^{-2}\sum_{i=1}^{[Nt]}\varepsilon_i^2\Big),\tag{2.1}$$

其中$\varepsilon_t\,(t=1,2,\ldots,N)$为服从稳定分布$S(\alpha,0,1,0)$的独立同分布随机变量, $c_\alpha=\sin(\pi\alpha/2)\cdot\Gamma(\alpha)/\pi$. 令$D[0,1]=D([0,1],R)$为有右连左极的实值函数空间. 令$n\geq 2$是一个正整数. 定义$D([0,1],R^n)=\{(x_1(t),\ldots,x_n(t)):x_i(t)\in D[0,1],1\leq i\leq n\}$. 则在$D((0,1),R^2)$中, 当$N\to\infty$时, $(U_N(t),V_N(t))\to(U(t),V(t))$, 其中$(U(t),V(t))$是有Lévy测度$\widetilde{v}(A)$的一个Lévy过程, $\widetilde{v}(A)=v\{x\in R:(x,x^2)\in A\}$. 对于$x>0$定义$v(x,\infty)=px^{-\alpha}$, $v(-\infty,-x]=qx^{-\alpha}$, 其中$0<\alpha<2$, $q=1-p$.

《应用概率统计》版权所有

引理 2.2 (Chan和Tran, 1989)

$$\lim_{N\to\infty} N(b_N-1)=\frac{1}{2}[U^2(1)-V(1)]/\int_0^1 U^2(t)\mathrm{d}t, \qquad \text{a.s.,} \tag{2.2}$$

其中, b_N见(1.2)式.

引理 2.3 (Chambers, Mallows和Stuck, 1976) 设随机变量Θ与W相互独立, Θ服从$(-\pi/2,\pi/2)$上的均匀分布. 令W服从均值为1的指数分布. $0<\alpha<2$, 令

$$Z=\begin{cases}\dfrac{\sin\alpha\Theta}{(\cos\Theta)^{1/\alpha}}\left[\dfrac{\cos((\alpha-1)\Theta)}{W}\right], & \alpha\neq 1;\\ \tan\Theta, & \alpha=1,\end{cases} \tag{2.3}$$

则$Z\sim S(\alpha,0,1,0)$.

引理2.3给出了产生服从稳定分布$S(\alpha,0,1,0)$的随机变量的方法.

§3. $E_{N,n}$的定义、性质及其经验分位数表

以下定理给出$E_{N,n}$的定义及性质.

定理 3.1 设

$$E_{N,n}=\frac{n}{2}\frac{\left(\sum\limits_{i=1}^{N}\varepsilon_i\right)^2-\sum\limits_{i=1}^{N}\varepsilon_i^2}{\sum\limits_{j=1}^{n}\left(\sum\limits_{i=1}^{[N\times(j/n)]}\varepsilon_i\right)^2}, \tag{3.1}$$

其中, N,n为正整数, $\varepsilon_i\sim S(\alpha,0,1,0)$, $i=1,2,\ldots,N$. 则

$$\lim_{N,n\to\infty} E_{N,n}=\frac{1}{2}[U^2(1)-V(1)]/\int_0^1 U^2(t)\mathrm{d}t, \qquad \text{a.s..} \tag{3.2}$$

证明: 由引理2.1

$$\left(c_\alpha^{-1}\sum_{i=1}^{N}\varepsilon_i, c_\alpha^{-2}\sum_{i=1}^{N}\varepsilon_i^2\right)=(U_N(1),V_N(1)),$$

且

$$\lim_{N\to\infty}(U_N(1),V_N(1))=(U(1),V(1)), \qquad \text{a.s..}$$

即

$$\lim_{N\to\infty}\left(c_\alpha^{-1}\sum_{i=1}^{N}\varepsilon_i, c_\alpha^{-2}\sum_{i=1}^{N}\varepsilon_i^2\right)=(U(1),V(1)), \qquad \text{a.s..} \tag{3.3}$$

由引理2.1以及连续映射定理可得, 当$N\to\infty$时,

$$\int_0^1 U_N^2(t)\mathrm{d}t\to\int_0^1 U^2(t)\mathrm{d}t.$$

《应用概率统计》版权所有

由黎曼积分定义, 当$n \to \infty$时,

$$\frac{1}{n}\sum_{j=1}^{n} U_N^2\Big(\frac{j}{n}\Big) \to \int_0^1 U_N^2(t)\mathrm{d}t,$$

综合可得

$$\lim_{N,n\to\infty} \frac{1}{n}\sum_{j=1}^{n} U_N^2\Big(\frac{j}{n}\Big) = \int_0^1 U^2(t)\mathrm{d}t, \qquad \text{a.s..} \tag{3.4}$$

综合(2.1)式, (3.3)式和(3.4)式, 即得(3.2)式成立. □

由定理3.1和引理2.2可知, $E_{N,n}$和$N(b_N - 1)$有相同的极限分布.

由(3.1)式知, 我们可以利用$E_{N,n}$给出$N(b_N - 1)$的经验分位数表, 步骤如下:

1. 固定α, n, N的值, 利用引理2.3产生随机数序列ε_i, 并由(3.1)式产生$E_{N,n}$的一个值.
2. 重复上述步骤m次, 并且对固定的α产生经验分位数.

表1 经验分位数表

α	1%	2.5%	5%	10%	50%	90%	95%	9.75%	99%
0.1	-4.085	-2.257	-0.985	-0.269	0.000	0.222	0.529	1.050	2.333
0.2	-12.182	-5.504	-3.284	-1.564	0.000	0.469	0.798	1.227	1.925
0.3	-10.975	-5.959	-3.968	-2.421	0.000	0.469	0.901	1.326	2.100
0.4	-9.167	-6.549	-4.269	-2.691	-0.004	0.609	1.031	1.451	2.353
0.5	-12.328	-8.535	-4.979	-2.947	-0.016	0.695	1.158	1.691	2.870
0.6	-11.714	-7.383	-5.116	-3.285	-0.052	0.791	1.276	1.735	2.275
0.7	-10.796	-7.511	-5.641	-3.445	-0.066	0.666	1.037	1.602	2.274
0.8	-11.367	-8.754	-5.665	-3.511	-0.105	0.861	1.487	2.125	3.353
0.9	-11.273	-8.279	-6.112	-3.962	-0.177	0.956	1.447	2.154	2.875
1.0	-13.986	-9.149	-6.597	-3.797	-0.284	0.837	1.194	1.676	2.522
1.1	-11.150	-8.836	-6.239	-4.120	-0.389	0.893	1.359	1.718	2.280
1.2	-11.051	-8.208	-5.829	-4.133	-0.267	0.958	1.328	1.907	2.873
1.3	-12.074	-9.072	-6.984	-4.853	-0.438	0.814	1.254	1.586	1.997
1.4	-13.052	-8.635	-7.017	-4.972	-0.566	0.875	1.298	1.652	2.102
1.5	-14.078	-7.847	-6.064	-4.595	-0.619	0.830	1.205	1.604	2.398
1.6	-13.049	-9.516	-7.560	-5.442	-0.558	0.881	1.222	1.512	2.314
1.7	-11.783	-9.792	-7.253	-5.340	-0.614	0.867	1.162	1.563	2.344
1.8	-12.118	-9.342	-7.493	-5.318	-0.850	0.882	1.197	1.542	2.178
1.9	-12.609	-9.806	-7.915	-5.524	-0.806	0.758	1.133	1.375	1.731
2.0	-14.122	-10.957	-8.254	-6.051	-0.973	0.868	1.161	1.600	2.082

《应用概率统计》版权所有

3. 变动α的值, 重复1, 2步, 得到不同的α对应的经验分位数.

对$\alpha \in (0,2]$, 取步长0.1, 对每一个α产生$N = 1000$个随机数ε_i, 取$n = 10,000$, 循环产生$m = 1000$个$E_{N,n}$的值, 产生的经验分位数表见表1.

下面, 用本文的方法产生的经验分位数表与用Chan和Tran (1989)的经验分位数表做个对比.

Chan和Tran的方法产生的经验分位数是基于序列长度为500由10,000次重复得到的. 为进行有效对比, 用本文的方法产生的经验分位数也是产生$N = 500$个随机数, 由10,000次重复得到. 结果见表2, 表中A表示本文的方法, B表示Chan和Tran的方法. 由表2可见, 由于在两种方法中, 产生的随机数个数N相同, 重复次数也相同, 得到的分位数差异不大. 但由于本文的方法与具体模型无关, 有更强的适应性.

表2　经验分位数表的对比表

α	方法	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%
0.5	A	-10.290	-6.589	-4.629	-2.881	-0.012	0.722	1.171	1.760	3.036
0.5	B	-10.631	-7.092	-4.794	-2.916	-0.009	0.715	1.117	1.756	2.911
1	A	-11.993	-8.851	-6.360	-4.208	-0.239	0.880	1.285	1.810	2.576
1	B	-11.903	-8.855	-6.143	-4.166	-0.271	0.847	1.260	1.677	2.683
2	A	-13.814	-10.304	-7.970	-5.646	-0.893	0.918	1.288	1.636	2.075
2	B	-14.108	-10.628	-8.083	-5.827	-0.836	0.919	1.276	1.625	2.044

§4. 实证分析

由Ruey (2010), 股票市场价格取对数后符合AR(1)模型. 下面以三组实际金融数据为例, 对AR(1)模型进行单位根检验.

第一组选取2000年1月3日到2011年5月20日, 每个工作日的标准普尔500 (S&P500)的每日收盘价. 该组数据的样本容量为2864.

第二组选取2002年1月1日到2011年5月20日, 每个工作日的上海证券综合指数(SSE)的每日收盘价. 该组数据的样本容量为2394.

第三组选取1998年10月28日到2002年9月4日, 每个工作日的英镑对美元汇率的每日收盘价. 该组数据的样本容量为1000.

对这三组数据都进行对数处理. 选定显著性水平$\gamma = 0.05$, 检验$H_0 : \beta = 1$, $H_1 : \beta \neq 1$. 取$T_N = N(b_N - 1)$作为检验统计量, 若T_N落在区间$(T_{\gamma/2}, T_{1-\gamma/2})$之外就拒绝$H_0$, 其中$T_{\gamma/2}$和$T_{1-\gamma/2}$由表1获得. 三组数据算得的$T_N$值见表3.

由表1中T_N的2.5%和97.5%分位数可知, 三组数据都不能拒绝原假设H_0, 即三组数据都存在单位根.

《应用概率统计》版权所有

表3　三组数据T_N的值

S&P500	SSE	GBP−USD
-1.05429	-1.01591	-1.3740

《应用概率统计》版权所有

参 考 文 献

[1] Hannan, E.J. and Kanter, M., Autoregressive processes with infinite variance, *Journal of Applied Probability*, **14(2)**(1977), 411–415.

[2] DuMouchel, W.H., Estimating the stable index α in order to measure tail thickness: a critique, *The Annals of Statistics*, **11(4)**(1983), 1019–1031.

[3] McCulloch, J.H., Continuous time processes with stable increments, *The Journal of Business*, **5(4)**(1978), 601–619.

[4] Knight, K., Rate of convergence of centred estimates of autoregressive parameters for infinite variance autoregressions, *Journal of Time Series Analysis*, **8(1)**(1987), 51–60.

[5] 武东, 汤银才, 稳定分布及其在金融中的应用, 应用概率统计, **23(4)**(2007), 434–445.

[6] Shin, D.W. and So, B.S., New tests for unit roots in autoregressive processes with possibly infinite variance errors, *Statistics and Probability Letters*, **44(4)**(1999), 387–397.

[7] Chan, N.H. and Tran, L.T., On the first-order autoregressive process with infinite variance, *Econometric Theory*, **5(3)**(1989), 354–362.

[8] Chambers, J.M., Mallows, C.L. and Stuck, B.W., A method for simulating stable random variables, *Journal of the American Statistical Association*, **71(354)**(1976), 340–344.

[9] Distaso, W, Testing for unit root processes in random coefficient autoregressive models, *Journal of Econometrics*, **142(1)**(2008), 581–609.

[10] Rathie, P.N., Coutinho, M., Sousa, T.R., Rodrigues, G.S. and Carrijo, T.B., Stable and generalized-t distributions and applications, *Communications in Nonlinear Science and Numerical Simulation*, **17(12)**(2012), 5088–5096.

[11] Ruey, S.T, *Analysis of Financial Time Series* (*Third Edition*), John Wiley & Sons, Inc., 2010.