

V/S 长记忆检验的有效性

——以上证 50 指数及其成分股为例

管河山,王谦^①,刘春^②

(南华大学 经济管理学院,湖南 衡阳 421001)

[摘要] 文章采集上证 50 综合指数及其成分股数据进行 V/S 长记忆性检验,实证发现:V/S 方法具有较好的稳健性,但其计算复杂度较高,不适合处理海量的数据;采用日、周或月三种采样间隔所得到数据的检验结果有显著差异。增大采样时间间隔虽然可以减少数据量、简化计算,但将导致检验结论不一致;上证 50 综合指数检验结果与其成分股数据检验结果之间存在明显区别,由此可见,综指与其成分股的长记忆性判断不可混为一谈。实践中应谨慎对待 V/S 检验结果。

[关键词] 时间序列; 长记忆性; V/S; 有效性

[中图分类号] F276.6 **[文献标识码]** A **[文章编号]** 1673-0755(2015)03-0054-05

金融系统的复杂性和波动性广泛影响着社会、经济发展的各个层面。大数据时代的来临,超高频、高频及低频(以日、周、月等采集时间间隔)数据大量涌现,采样时段差异和众多的金融市场(中国、香港、美国等)都导致数据存储量的大幅增加,所采集到的数据呈现出明显的海量特征,这使得传统的数据存储和处理方式渐渐满足不了现代金融分析和决策的需求。如何在海量数据的背景下,更好地运用 V/S 方法进行金融时间序列数据的长记忆性检验,对我们了解和预测金融市场的走向、把握金融时间序列的发展规律以及更合理有效地规避市场运行过程中的风险具有重要的现实意义。

对时间序列长记忆性的研究是时间序列非线性关系研究的热点和重点,且其研究大多集中于金融股票市场。如许友传发现 R/S 检验更倾向于接受长记忆性,并得出需谨慎解释采用 R/S 方法和 MR/S 方法所得出实证结果的结论^[1];李志生采用序列相关检验和 R/S 检验法分析得出我国股票市场有显著的长记忆性并在近年来逐步增强的趋势^[2];顾巧明探讨了我国股市波动中的长记忆性效应,并对股市收益率进行了长记忆建模^[3];苑莹运用 R/S 法、MR/S 法、KPSS 检验及 Granger 因果检验方法发现中国股市收益率和成交量均具长记忆性特征^[4];肖炜麟研究了长记忆建模影响问题,发现不同模型

的定价结果证明了金融资产具有长期记忆性^[5]。

已有研究中,R/S 方法、MR/S 方法和 V/S 方法都得到了较为广泛的研究和应用。Giraitis 提出的 V/S 方法得到了广泛的应用拓展^[6],Cajueiro 比较了 V/S 方法和经典 R/S 法,通过运用蒙特卡罗模拟法发现 V/S 分析比经典 R/S 分析更为稳健^[7];何兴强^[8]和顾荣宝^[9]对我国沪深股市进行 V/S 检验,认为我国股票市场呈现反持续性特征而非长记忆性特征;王文静通过分别采用 R/S 与 V/S 方法来检验香港股市长记忆性发现两种方法短期相关性的敏感度相当^[10]。除了在金融时间序列数据分析方面,V/S 检验也运用于其他研究领域的研究中^[11],比如,孙东永将 V/S 分析引入到河川径流中进行长记忆性研究^[12]。以上研究都极大地丰富了 V/S 检验的理论研究和实践应用,然而对 V/S 检验的有效性缺乏一个系统的探讨。文章将从金融时间序列长记忆性检验角度展开分析,重点研究 V/S 检验的有效性。

一 长记忆性检验

(一)长记忆性定义

长记忆性又称长程相关性或持久性,它描述了时间序列在较远间隔处的持续依赖关系,表明时间序列具有非周期的循环特征,如果其动态相依结构表现出某种长记忆特征或长时相依关系,则时间序

[收稿日期] 2015-04-20

[基金项目] 教育部人文社科青年项目“海量金融时间序列数据平稳性检验方法研究”资助(编号:13YJJCZH044);湖南省研究生创新项目“大数据时代金融时间序列长记忆性检验方法研究”资助(编号:CX2014B397)

[作者简介] 管河山(1981-),男,湖南祁东人,南华大学经济管理学院副教授,博士(后)。

①南华大学经济管理学院硕士研究生。②南华大学研究生处副教授。

列的历史信息将有助于预测其未来变化趋势^[12]。国内外对长记忆性的定义有很多种,本文选取两种较常见的定义方式,如下所示:

第一种是根据自相关函数的定义,当平稳过程 $X_t(t = 1, 2, \dots, T)$ 的自相关函数 $\alpha(k)$ 满足:

$$\alpha(k) \rightarrow C_\alpha k^{2d-1} \quad (1)$$

其中, $k \rightarrow \infty, C_\alpha > 0, d < 1/2$, 则认为平稳时间序列 $\{X_t\}$ 具有长记忆性或长相依性。

第二种是谱密度函数定义,如果存在一个实数 d 和一个有界常数 C , 使得对于一个平稳序列 $\{X_t\}$, 它的谱密度 $f\{\rho_j\}$ 满足:

$$f\{\rho_j\} \sim C^* \rho_j^{-2d}, \rho \rightarrow 0 \quad (2)$$

其中 $d \in (0, 0.5)$, $\rho_j = 2\pi^* j/N, j = 1, 2, \dots, m$, N 为样本总数, m 为带宽, $m = g(T) \leq T$, 则认为时间序列 $\{X_t\}$ 具有长记忆性或长相依性。

长记忆性描述了时间序列在较远间隔处的持续依赖关系,表明时间序列具有非周期的循环特征,如果其动态相依结构表现出某种长记忆特征或长时相依关系,就使得利用时间序列本身的历史信息来预测其未来变化趋势成为可能,为准确掌握金融时间序列的特征、了解和预测金融市场的走向、建立符合其特征的预测模型提供了现实可能性。时间序列的记忆性作为建立预测模型的关键因素之一,其研究不仅可以为上层政策制定和宏观调控提供决策依据,还能为组织和个人投资者提供一定现实参考。

(二)R/S 和 MR/S 方法

对时间序列长记忆性的理论研究最早始于 20 世纪 50 年代。Hurst^[13] 在对水文数据的研究中发现了时间序列具有记忆性的属性,由此提出了重标极差分析的检验方法,并给出了相应的判别式。众多实证表明,R/S 方法在分析存在短期记忆性或具有异方差的时间序列时,会更多地倾向于拒绝“不存在长记忆性”的原假设^[1],而这与该方法统计量构造的数学原理有密切关系。

而 W. Lo^[14] 在其论文《股票价格市场的长记忆性》一文中描述到当序列包含短记忆性、存在异质性、具有非平稳性时,R/S 分析的结果就会大打折扣,因此 W. Lo 也提出了 MR/S 法对 R/S 方法中的分母进行了改进,通过引入调整项 $\theta_N(q)$ 可以在一定程度上修正 R/S 方法因时间序列的短记忆性和异方差等造成的过度拒绝时间序列存在长记忆性的原假设后果,使得结果的输出更加无偏有效。然而 MR/S 方法的缺陷在于调整项 $\theta_N(q)$ 中的修正阶数(参数 q)难以确定,取值太小不利于消除序列短期

记忆对检验结果的影响,取值太大又会影响该方法的检验效率和性能。

(三)V/S 方法

随着研究的不断深入,学者们发现由于 MR/S 检验统计量的构造原理上的漏洞,也使得其在使用过程中存在如何合理确定 q 值的问题。于是, Giraitis^[6] 提出了一种明显异于前两者的新的检验时间序列长记忆性的方法——重标方差法(V/S 分析法)。至此,这三种方法的提出基本上奠定了长记忆性检验方法的基础。一般认为 V/S 方法以其在构造机理上的优势有效地避免了 R/S 分析的不足,使得其在判断时间序列长记忆性上更为稳健,其检验结果也更为有效可靠。虽然 V/S 方法以其独特优势逐渐成为检验金融时间序列长记忆性的重要方法并获得广泛应用,但人们对其有效性的关注和探讨却不多。

V/S 方法的基本思想是:将总长度为 M 的时间序列 $\{X_t\}$ 划分为 A 个长度为 N 的等长子区间,记作 $D_a, a = 1, 2, \dots, A$ 。子序列 D_a 对应的元素为 $x_{i,a}, i = 1, 2, \dots, N$,子序列均值为 \bar{x}_a ,子序列标准差为 S_N ,V/S 方法的判别式如下:

$$(V/S)_N = \frac{1}{N \cdot S_N^2} \left[\sum_{k=1}^N \left(\sum_{i=1}^k (x_{i,a} - \bar{x}_a) \right)^2 - \frac{1}{N} \left[\sum_{k=1}^N \sum_{i=1}^k (x_{i,a} - \bar{x}_a)^2 \right] \right] \quad (3)$$

其中, $S_N = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{i,a} - \bar{x}_a)^2}$ 为子序列标准差; $\bar{x}_a = \frac{1}{N} \sum_{i=1}^N x_{i,a}$ 为子序列均值。对于不同区间划分 N 有与其对应的子序列的 V/S 值,而 V/S 统计量满足于 $(V/S)_N = kN^{2H}$ 的函数关系,可通过对 $\log(N) \sim \log(V/S)_N$ 进行最小二乘估计来求得直线的斜率, H 值作为判断序列是否具有长记忆性的关键指标,在 V/S 方法中, H 值即为回归斜率的一半。利用 Hurst 指数(即 H 值)进行时间序列长记忆性检验的判别,其准则如下:

当 $0 \leq H \leq 0.5$ 时,时间序列是反持久性的分形布朗运动。 H 距离越接近于零,这种反持久性就越强,此时其比随机序列具有更强的突变性和易变性。

当 $H = 0.5$ 时,时间序列为标准的随机过程。不同时间的值毫不相关。

当 $0.5 < H < 1$ 时,时间序列的持续性越强,即具有长记忆性。过去的值对现在和未来的值有重要影响,且记忆增强的趋势随着 H 值越接近于 1 就越显著。

王文静^[10]曾提出:同一金融时间序列的不同时间区间的划分将对长记忆性的最终检验结果产生影响。那么,V/S 检验方法对不同数据采集时段(日、周、月三个时段)的金融时间序列的检验结果是否会一致呢?股市综合指数及其成分股数据的检验结果又是否存在一致性?为此,本文采集了近十年来上证 50 指数及其 50 支成分股的日、周、月数据(收益率序列)进行实证。

二 实证分析

本文利用 Matlab 软件编程,采集了 CSMAR 数据库中上证 50 综合指数(代码 000016.SH)2004 年 1 月 1 日—2014 年 2 月 28 日的收盘价数据进行实证。需说明的是:上证 50 指数自 2004 年 1 月 2 日起正式发布,由于市场环境和公司经营的差异,上证 50 的成分股基本上每年都有变动和更新。因此,每个成分的股票交易数据量和在上市时间都会有差异。由于平稳性是数据进行长记忆性分析的可靠前提,因此本文仅对收益率数据进行 V/S 长记忆性检验,然后再对其检验结果和计算效率进行对比分析。其中,收益率序列由收盘价序列经对数差分法计算而得,具体公式如(4)所示:

$$R_i = \ln(x_{i+1}) - \ln(x_i) \quad (4)$$

其中, x_i 表示第 i 天的收盘价, x_{i+1} 表示第 $(i+1)$ 天的收盘价。

(一) V/S 长记忆性检验结果

为了便于统计和区分,将上证 50 指数所选取的 50 支股票的代码按升序排列并分别用序列号 1~50 来表示,如表 1 所示:

1、上证 50 成分股收益率序列的 V/S 检验

本文对上证 50 指数成分股的收益率序列进行

V/S 检验,结果如图 1 所示。

表 1 上证 50 公布的最新成分股代码及编号

编号	股票代码	编号	股票代码	编号	股票代码	编号	股票代码
1	600000	14	600256	28	600999	42	601668
2	600010	15	600332	29	601006	43	601688
3	600015	16	600362	30	601088	44	601699
4	600016	18	600406	31	601117	45	601766
5	600018	19	600489	32	601166	46	601818
6	600028	20	600518	33	601169	47	601857
7	600030	21	600519	35	601299	48	601899
8	600031	22	600547	36	601318	49	601901
9	600036	23	600549	37	601328	50	601989
10	600048	24	600585	38	601336	—	—
11	600050	25	600637	39	601398	—	—
12	600104	26	600837	40	601601	—	—
13	600111	27	600887	41	601628	—	—

不难发现:在不同数据采集时段(日、周、月)下,收益率数据的长记忆性检验结果不完全相同。其次,在日、周、月三种不同的数据采集方式下,50 支成分股中具有长记忆性的序列数分别为 34,25 和 18;从 H 值数值大小上看,日数据 H 值一般大于其他类型数据,且更倾向于接受长记忆性,而周、月数据更倾向于具有反记忆性特征;再次,收益率检验结果中的长记忆性并不十分突出,且大约十只股票的 H 值是接近于 0.5 的弱记忆性。这可能与收益率采取了非线性转换方式的计算原理有关,经过转换后的时间序列与原序列通常呈现出不同的平稳性特征,进而改变其在长记忆性上的特征。总之,由图 1 可知收益率的日数据 H 值一般大于周、月数据的 H 值,更倾向于具有长记忆性。

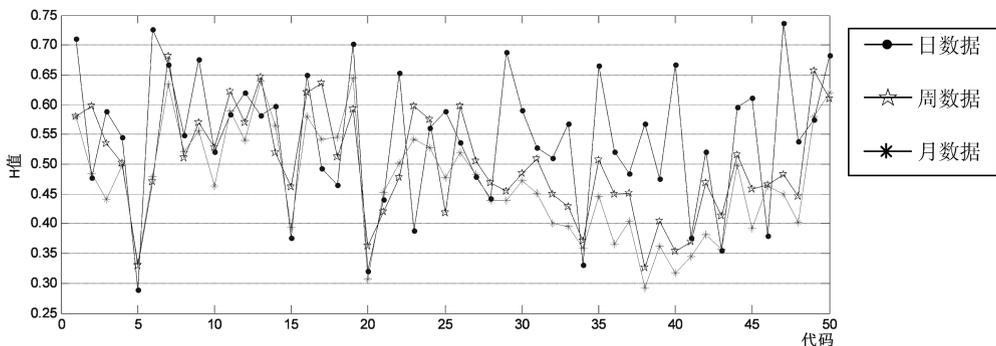


图 1 上证 50 成分股的收盘价数据 V/S 检验 H 值

2、上证 50 综指的 V/S 检验

通过对上证 50 指数的收益率序列进行长记忆

性检验,结果如表 2 所示。

表2 上证50综合指数V/S检验情况表

代码000016.SH	V/S 检验		数据量/个
	收益率H值	耗时s	
日序列数据	0.7281	17475.1194	2483
周序列数据	0.7231	79	510
月序列数据	0.6962	1.19	122

由表2可知,上证50指数的收益率序列呈现出较强的记忆性特征。可得出基本结论:收益率序列倾向于接受长记忆性,且日数据H值均大于周和月数据,日数据序列更倾向于接受长记忆性。上证50股指与其成分股的长记忆性结果比较结果如表3所示。

表3 上证50指数及成分股V/S长记忆性检验结果

数据类型	是否具有V/S长记忆性		
	日序列数据	周序列数据	月序列数据
上证50指数	√	√	√
50只成分股	34√	25√	18√
	16×	25×	32×

注:“-”表示无法判别其长记忆性;“√”代表时间序列具有长记忆性;“×”代表时间序列不具有长记忆性。例如,“38√,12×”代表在上证50指数50只成分股的收盘价日序列数据中有38只具有长记忆性,12只不具有长记忆性。

由表3可知:成分股中关于记忆性特征的内部差异明显,同时结合图1分析结果发现,50支成分股的收益率序列进行检验后所得到的长记忆性结果不尽相同,有的强记忆性特征明显,有的则不能体现;其次,对同一成分股而言,不同的采样时段(日、

周、月)得到的数据的长记忆性检验结果不一致,若无法证明V/S方法确实受时间采集间距的影响,那么V/S方法的有效性和稳健性还有待考证;其三,上证50指数的收益率序列基本呈现出强记忆性特征,而成分股却因采样时段的不同表现出不一致的记忆性特征。

以上成分股和上证50股指在长记忆性检验上的结果差异提示我们:市场综合指数并不能完全刻画其成分股的数学特征和统计属性。市场指数具有强长记忆性时,不能说明其成分股也具有相应特征。本文研究金融时间序列的长记忆性并非为了进行宏观经济研究和支撑政策制定,而是为了更好地了解和预测金融市场,把握其运行规律,从而更好地进行金融投资决策。因此,想要更加全面地了解和预测未来的金融序列的价格走势,单纯地研究某一市场综合指数,可能并不能解决这一问题,反而深入地研究目标股票的历史行情数据,并以此为基础进行统计建模,可能更有利于我们预测其未来趋势。

(二)V/S检验耗时情况

尽管V/S方法被广泛应用于时间序列长记忆性检验,Giraitis等人将R/S检验中的样本累积离差的极差用序列方差来代替,发现在检验时间序列长记忆性的方法上,V/S分析更具有有效性和稳健性^[15]。但正是由于V/S方法使用的是序列方差而非极差运算才使得其计算复杂度大大高于R/S方法和MR/S方法,达到 $o(n^2)$ 。在样本量较大时,计算速率将出现明显的缓慢现象,本文以成分股收益率V/S检验的耗时情况用表4来说明。

表4 成分股收盘价V/S检验的耗时

统计量	日序列		周序列		月序列				
	H值	耗时s	数据量	H值	数据量	H值			
max	0.7363	18228.3019	2421	0.6815	151.9597	515	0.6453	1.1946	122
min	0.2896	61.4849	516	0.3252	0.8409	112	0.2925	0.0177	27

注:max表示最大值;min表示最小值。

从表4不难发现:对周、月数据而言,由于数据量较小,V/S方法处理速度相对较快,一般只需几十秒;但对于日数据而言,最大数据量时其运行速度需要5个小时左右,耗时较长。随着计算机技术的发展和存储设备的更新,金融领域已经具备采集海量数据的能力。比如采集每五分钟一个的高频数据,一天可产生48个股票数据,此时若用V/S方法来验证时间序列的长记忆性,其复杂度将非常高,其可行性将引起“质疑”。正是因为V/S方法的计算复杂度高,使得它可能并不适合处理海量的金融时间

序列数据。

三 结论

通过对上证50指数及其50支成分股近十年的金融市场股票数据进行长记忆性检验,我们得出结论如下:

(一)V/S方法对金融时间序列长记忆性的有效性和可靠性判断还有待深究。

同一序列数据样本(或同一股票)可能会因时间采集间隔的差异而导致V/S方法长记忆性检验判断结果的差异。这可能是由于样本量多时更适合

进行长记忆性检验,而样本量太少会直接影响长记忆性的判别结果。有研究指出,三至四年的日数据量更有利于检测时间序列本身的长记忆性^[10]。

(二)V/S方法在判断时间序列长记忆性时的判断标准存在模糊性。Giraitis提出了V/S方法及其数学判别式,而对其长记忆性的判别标准并没有进行严格的划分,如在计算过程中多次出现了大于1的情况,这该如何判别?仍需要在后续研究中寻求答案。

(三)上证50指数不受时间采集间距的影响,在收益率时间序列上都呈现出显著的长记忆性,而其成分股在检验过程中随时间采集间距的差异检验结果各不相同。这说明综合股指并不能完全体现其成分股的属性特征,想要真实了解成分股特征和未来作出正确的投资决策还需具体情况具体分析。同时,V/S检验时样本的数据量越大,其运算耗时就越长,计算的速度就越慢。

综上所述,实证表明尽管V/S方法在时间序列长记忆性检验方面有较好的稳健性,但也应谨慎对待其结果并与其他检验方法综合使用。本文研究结论丰富了V/S长记忆性检验方法的相关研究,为以后长记忆性检验方法的选择提供了分析依据。

[参考文献]

- [1] 许友传. 金融时间序列R/S长记忆性检验的有效性[J]. 中国管理科学, 2010(8): 204-209.
- [2] 李志生, 刘正捷. 资产收益的短记忆性与长记忆性: 我国股票市场效率的动态分析[J]. 江西财经大学学报, 2011(1): 5-12.
- [3] 顾巧明. 我国股市波动的长记忆性与货币政策的非对

称性研究[J]. 中国经济问题, 2011(1): 38-43.

- [4] 宛莹, 杜乐鹿, 庄新田. 股市收益率与交易量长记忆性实证研究[J]. 东北大学学报: 社会科学版, 2012(7): 1056-1060.
- [5] 肖炜麟, 张卫国, 徐维东. 双分式布朗运动下股本权证定价[J]. 系统工程学报, 2013, 28(3): 348-354.
- [6] Giraitis L, Kokoszka P, Leipus R, et al. Rescaled variance and related tests for long memory in volatility and levels[J]. Journal of econometrics, 2003, 112(2): 265-294.
- [7] Daniel O Cajueiro, Benjamin M Tabak. The rescaled variance statistic and the determination of the Hurst exponent [J] Mathematics and Computers in Simulation, 2005 (70): 172-179.
- [8] 何兴强, 李仲飞. 上证股市收益的长期记忆: 基于V/S的经验分析[J]. 系统工程理论与实践, 2006, 12(12): 47-54.
- [9] 顾荣宝, 陈霖霞. 基于分形V/S技术的沪深股市长记忆性研究[J]. 安徽大学学报: 自然科学版, 2008, 32(3): 18-21.
- [10] 王文静, 马军海. 基于R/S分析和V/S分析的香港股市长记忆性比较研究[J]. 经济经纬, 2009(2): 140-143.
- [11] 刘春, 王谦, 管河山. 系统协调度模型的稳健性研究[J]. 南华大学学报: 社会科学版, 2014(5): 66-70.
- [12] 孙东永, 黄强, 王义民. 基于V/S分析的河川径流长记忆性研究[J]. 水力发电学报, 2012, 30(6): 178-182.
- [13] Hurst H E. Long-term storage capacity of reservoirs[J]. Trans. Amer. Soc. Civil Eng., 1951(116): 770-808.
- [14] Lo A W. Long-Term Memory in Stock Market Prices [J]. Econometrica, 1991, 59(5): 1279-1313.

Study on the V/S Test Validity of Long Memory

—As the SSE50 index and its 50 constituent stocks for example

GUAN He-shan, WANG Qian, LIU Chun

(University of South China, Hengyang 421001, China)

Abstract: The paper did V/S Test with the data of SSE50 index and its 50 constituent stocks in Shanghai Security Exchange and also found that: The V/S method has high complexity and time-consuming in calculation, so it may not be suitable for handling massive amounts of data; under different collected ways of data (day, week and month), the results are significantly different in the V/S test of long memory; increasing the sampling interval can reduce the amount of data and simplify the calculation, but it will lead to an inconsistent conclusions of the test; the long memory test results of 50 constituent stocks of SSE50 index are not very consistent, thus, we cannot confuse the results of them. In a word, the long memory test results still should be treated seriously.

Key words: financial time series; long memory; V/S test; validity