

Sieve Bootstrap Test for Changes between Unit Root Process and Fractional Integrated Processes

Mingcan He, Guolong Fu

School of Mathematics and Statistics, Qinghai Normal University, Xining Qinghai
Email: 378771567@qq.com

Received: Mar. 15th, 2018; accepted: Apr. 2nd, 2018; published: Apr. 9th, 2018

Abstract

This paper aimed to test change point from unit root process to fractional integrated process as well as fractional integrated process to unit process via a Dickey-Full ratio statistic. A Sieve Bootstrap method was proposed to determine the critical values. Simulations indicate that our proposed method can control the empirical size well both under the unit root and fractional integrated process null hypotheses, and gives satisfy empirical powers under two alternative hypotheses if the change point location does not too back. Furthermore, Dickey-Full ratio statistic has better performance when detecting those changes which from unit root process to fractional integrated process.

Keywords

Unit Root Process, Fractional Integrated Process, Changes Point, Sieve Bootstrap

单位根过程与分整过程间变化变点的Sieve Bootstrap检验

何明灿, 付国龙

青海师范大学数学与统计学院, 青海 西宁
Email: 378771567@qq.com

收稿日期: 2018年3月15日; 录用日期: 2018年4月2日; 发布日期: 2018年4月9日

摘要

本文基于Dickey-Full比率统计量研究了从单位根过程向分整过程变化变点, 以及从分整过程向单位根过

程变化变点的检验问题, 并提出了用于确定检验统计量临界值的Sieve Bootstrap方法, 数值模拟结果表明本文方法在单位根及分整过程原假设下都能很好的控制检验水平, 在变点位置不是太靠后时, 对两种备择假设下的变点都有较高的检验势, 且在检验从单位根过程向分整过程变化变点时效果更显著。

关键词

单位根过程, 分整过程, 变点, Sieve Bootstrap

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

过去三十年里, 有越来越多的证据表明许多时间序列数据中存在从平稳向非平稳或从非平稳向平稳变点的变点, 这使得关于持久性变点的检验问题受到统计学和计量经济学等领域的广泛关注, 如 Hakkio 和 Rush [1]发现美国财政赤字数据中存在一个由平稳过程到单位根过程变化的持久性变点, Kim [2] [3]等提出了一种残差比率统计量来检验由平稳过程向单位根过程变化以及从单位根过程向平稳过程变化的持久性变点, Perron [4] [5]研究了具有结构变点的单位根检验问题, 并提出 Dickey-Fuller (DF)检验统计量。传统的 DF 单位根检验法应用到持久性变点检验问题中时, 由于序列中单位根部分支配渐进性导致渐进分布不会随着样本大小而发散到负无穷, 因此传统的 DF 检验无法区分从单位根过程向平稳过程变化的持久性变化变点。为此, Banerjee 等[6]在单位根过程原假设下提出了能够检验持久性变点的修正 DF 统计量。Leybourne 等[7]指出 Banerjee 等[6]构造的 Dickey-Fuller 型统计量存在不足, 即当检测序列整段为平稳过程时检测结果仍然显示序列存在持久性变点, 为此, Leybourne 等[7]提出了一种 Dickey-Fuller 比率统计量来解决此问题, 结果表明 Dickey-Fuller 比统计量不论在单位根原假设还是平稳原假设下都不会出现拒绝原假设, 认为数据中存在持久性变点的问题, 且在备择假设下能够很好的检测这两个方向变化的持久性变点。

上述这些研究都只考虑的是单位根过程和短记忆平稳过程之间变化的持久性变点检验问题。由于在短记忆平稳过程和单位根过程之间还存在一类更为广泛的分整过程(或称为长记忆过程), 这使得单位根过程、短记忆过程及分整过程间变化变点的检验问题也收到广泛关注, 如 Hassler 和 Scheithauer [8]基于 Kim [2]提出的比率统计量研究了从短记忆过程向长记忆过程变化变点的检验问题, Sibbertsen 和 Kruse [9]在长记忆过程原假设下基于一种平方 CUSUM 方法研究了持久性变点的检验问题。本文基于 Dickey-Fuller 比统计量研究从单位根过程向分整过程以及从分整过程向单位根过程变化变点的检验问题。

由于 Dickey-Fuller 比统计量的临界值不可避免的会依赖长记忆参数值, 为方便实际应用, 本文提出通过 Sieve Bootstrap 方法来近似检验统计量的临界值。Sieve Bootstrap 方法由 Buhlmann [10]于 1997 年首次提出, Poskitt [11]指出 Sieve Bootstrap 方法在分整过程分析中有非常好的结果, Kapetanios [12]提出了适用于分析非平稳长记忆时间序列的 Sieve Bootstrap 方法。Chen 等[13]将 Sieve Bootstrap 方法应用到长记忆时间序列方差变点检验问题, Chen 等[14] [15]基于 Sieve Bootstrap 方法进一步研究了从短记忆过程向长记忆过程变化变点, 以及从平稳长记忆过程向非平稳长记忆过程变化变点的检验问题, 收到了较好的效果。本文数值模拟结果表明, 基于 Sieve Bootstrap 方法确定的临界值在单位根及分整过程原假设下都能很好的控制检验水平, 且在备择假设下对两个方向变化变点都有较高的检验势。

2. 模型与假设检验

本文考虑如下分整模型:

$$(1-L)^{d_0} X_t = \varepsilon_t, \quad t=1, 2, \dots, n,$$

其中 n 为样本容量, L 是滞后算子, d_0 为长记忆参数, ε_t 为独立同分布随机变量, 且 $E(\varepsilon_t)=0$, $\text{Var}(\varepsilon_t)=\sigma^2 < \infty$, X_t 与 d_0 间的关系可被表示如下:

$$X_t = \sum_{j=0}^{\infty} w_j(d_0) \varepsilon_{t-j}.$$

其中 $w_j(d_0) = \frac{\Gamma(d_0 + j)}{\Gamma(d_0)\Gamma(1+d_0)}$, $0 \leq d_0 \leq 1$; 当 $0 \leq d_0 < 0.5$ 时 X_t 为平稳序列, $0.5 < d_0 \leq 1$ 时序列 X_t 为非平稳序列, 特别地当 $d_0 = 1$ 时序列 X_t 是一个单位根过程。

记 $X_t \sim I(1)(d_0=1)$, $X_t \sim I(d_0)(0 < d_0 < 1, \text{且 } d_0 \neq 0)$ 。

本文研究以下两类变点检验问题:

1) 序列 X_t 从 $I(1)$ 向 $I(d_0)$ 变化的持久性问题假设检验, 即检验原假设

$$H_0^1: X_t \sim I(1), \quad t=1, 2, \dots, n.$$

备择假设

$$H_1^1: X_t \sim \begin{cases} I(1), & t=1, \dots, k^*, \\ I(d_0), & t=k^*+1, \dots, n. \end{cases}$$

2) 序列 X_t 从 $I(d_0)$ 向 $I(1)$ 变化的持久性问题假设检验, 即检验原假设

$$H_0^2: X_t \sim I(d_0), \quad t=1, 2, \dots, n,$$

备择假设

$$H_1^2: X_t \sim \begin{cases} I(d_0), & t=1, \dots, k^*, \\ I(1), & t=k^*+1, \dots, n. \end{cases}$$

其中 $k^* = [\tau * n]$ 为未知变点位置, $[\cdot]$ 表示取整数部分。

为了检验上述两类假设检验问题, 采用 Lebourne 和 Tolyor [7] 提出的 Dickey-Full 比统计量:

$$\Xi_n(\tau) = \left| \frac{DF^f}{DF^r} \right|,$$

其中 $DF^f \equiv \inf_{\tau \in \Lambda} DF^f(\tau)$, $DF^r \equiv \inf_{\tau \in \Lambda} DF^r(\tau)$, $\tau \in (0, 1)$ 且 $DF^f = [\tau n](\hat{\rho}_1 - 1)$,

$$DF^r = [(1-\tau)n+1](\hat{\rho}_2 - 1), \quad \hat{\rho}_1 = \frac{\sum_{t=1}^{[\tau n]} X_{t-1} X_t}{\sum_{t=1}^{[\tau n]} X_{t-1}^2}, \quad \hat{\rho}_2 = \frac{\sum_{t=1}^{[(1-\tau)n+1]} X_{n-t} X_{n-t+1}}{\sum_{t=1}^{[(1-\tau)n+1]} X_{n-t+1}^2}.$$

对于给定区间 $\Lambda = [\tau_1, \tau_2] \subset (0, 1)$, 统计量的值定义为: $\Xi_n(X) = \left| \frac{\inf_{\tau \in \Lambda} DF^f(\tau)}{\inf_{\tau \in \Lambda} DF^r(\tau)} \right|$, 在原假设 H_0^1 下,

当 $\Xi_n(X)$ 的值小于临界值时拒绝原假设 H_0^1 , 认为数据中存在从 $I(1)$ 向 $I(d_0)$ 变化的持久性变点; 在原假设 H_0^2 下, 当 $\Xi_n(X)$ 大于给定临界值时拒绝原假设 H_0^2 , 认为数据中存在从 $I(d_0)$ 向 $I(1)$ 变化的持久性变点。

3. Bootstrap 方法

本文采用 Rambaccussing [16]提出的分数阶差分 Sieve Bootstrap 方法来近似临界值, 其步骤如下:

1) 估计 X_t 的长记忆参数 d_0 , 并记估计值为 \hat{d} 。

2) 对 X_t 进行 \hat{d} 阶差分, $\hat{\varepsilon}_t = (1-L)^{\hat{d}}(X_t - X_{t-1}), t=1, \dots, n$ 。其中 $(1-L)^{\hat{d}} = \sum_{j=0}^{\hat{d}} \hat{a}_j L^j, a_0 = 1,$

$$a_j = a_{j-1} (j - \hat{d} - 1) / j, j \geq 1。$$

3) 对 $\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_t$ 进行重抽样, 得到新的序列 $\hat{\varepsilon}_{ij}^*, \dots, \hat{\varepsilon}_{nj}^*$ 。

4) 生成 Bootstrap 样本 $\hat{Y}_{ij}^* = (1-L)^{-\hat{d}} \hat{\varepsilon}_{ij}^* + \hat{z}_{ij}$, 其中 $(1-L)^{-\hat{d}} = \sum_{s=0}^{\hat{d}} b_s L^s, b_0 = 1, b_j = b_{j-1} (j + \hat{d} - 1) / j, j \geq 1,$

\hat{z}_{ij} 是独立同分布的标准正态分布随机序列。

5) 计算统计量 $\Xi_n(\hat{Y}_{ij}^*)$, 重复步骤 3~5 B 次, 取 $\Xi_n(\hat{Y}_{ij}^*)$ 的 $1-\alpha$ 分位数作为检验统计量在显著水平 α 下的临界值。

4. 数值模拟

本节通过数值模拟检验所提方法的有效性, 所有模拟通过 R 语言程序 3.4.0 版本下实现。数据由 ARFIMA(0, d, 0) 模型生成, 样本量 n 取 200 和 500, 长记忆参数 d_0 分别取 0.2, 0.4, 0.6 和 0.8, 变点位置 $k^* = [\tau * n]$, τ 分别取 0.25, 0.5 和 0.75, Sieve Bootstrap 循环次数 $B = 199$, 检验水平取 10% 和 5%, 长记忆参数的估计方法使用 Robinson [17] 提出的 Local Whittle 估计, 所得模拟结果都经过 2000 次循环得到。

表 1 给出了统计量 $\Xi_n(X)$ 在 200 和 500 样本量下的经验水平, 由结果可以看出: 统计量在 d_0 取 0.2, 0.4 和 1 时可以很好的控制经验水平, d_0 取 0.6 和 0.8 时略高于检验水平, 在 500 样本量下所得经验水平更加接近检验水平, 结果说明 Sieve Bootstrap 方法可以有效的近似统计量的临界值。

表 2 给出了检验统计量在 H_1^1 和 H_1^2 下的经验势, 由结果可以看出: 在 H_1^1 下随着 d_0 取值的增大检验势变低, d_0 取 0.2 时更容易检测到变点, 随着 d_0 增大到 0.8 时检验势递减到最低, 说明当长记忆参数间隔越大越容易检测到变点。例如 $n = 200, d_0 = 0.2$, 5% 检验水平下统计量在不同变点位置的检验势分别为 96.7%, 98.4%, 32.3%, 而 $d_0 = 0.8$ 时检验势分别为 14.9%, 24.8% 和 13%。变点位置同样对检验结果有显著影响, 当 $k^* = 0.5 * n$ 时检验势要高于 $k^* = 0.25 * n$ 和 $k^* = 0.75 * n$, 但在 $k^* = 0.25 * n$ 和 $k^* = 0.5 * n$ 时的检验结果较为接近, 而当变点位置 $k^* = 0.75 * n$ 时检验势较低, 这说明统计量 $\Xi_n(X)$ 对位置比较靠前的变点具有较好的检验效果, 而当变点位置靠后时检测效果较差。在 H_1^2 下, 统计量 $\Xi_n(X)$ 检验势结果同

Table 1. Empirical sizes of $\Xi_n(X)$ (%)

表 1. $\Xi_n(X)$ 的经验水平 (%)

d_0	$n = 200$		$n = 500$	
	10	5	10	5
1	9.1	4.6	10.4	4.9
0.8	12.6	7.9	12	7.1
0.6	12.8	8.1	12.4	7.6
0.4	9.7	5.6	10.3	5.7
0.2	9.9	5.8	9.6	4.8

Table 2. Power values of $\Xi_n(X)$ (%)**表 2.** $\Xi_n(X)$ 的检验势(%)

d_0	n	$\alpha\%$	$I(1) \sim I(d_0)$			$I(d_0) \sim I(1)$		
			$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$
0.2	200	10	97.8	99	41.8	81.5	90.2	74.3
		5	96.7	98.4	32.3	70.2	80.6	61
	500	10	98.9	99.7	46.4	96.1	99.2	93.1
		5	97.1	99.1	35.8	91.5	93.8	83.4
0.4	200	10	85	93.4	39.2	61	63.4	34.8
		5	77.1	90.5	29.9	46.8	53.2	20.1
	500	10	98.2	99	43.5	79.2	82.7	59.4
		5	96	98.5	34.9	68.6	74.3	48
0.6	200	10	48.9	69	32.5	32.6	39.9	11.1
		5	37.3	58	21.2	23	25	4.4
	500	10	74.6	87.9	36.7	44.4	54.4	17.6
		5	66.4	79.8	25	36	39.1	9.2
0.8	200	10	23.9	35.3	20.4	10.8	19.2	3.2
		5	14.9	24.8	13	5.2	11.8	1.4
	500	10	34	51	28.2	13.2	22.8	4.2
		5	20.7	37.2	17	6.9	17.3	1.9

样具有上述性质,但在 d_0 取 0.6 和 0.8 且变点位置靠后时几乎检测不到变点。此外,随着样本容量的增大检验统计量在所有情况下的检验势都显著提高,这说明该方法是检验这两类变点问题的一致方法。比较两种备择假设下的检验势可以发现,对于相同跳跃度的变点统计量 $\Xi_n(X)$ 更容易检测到从单位根过程向分整过程变化的变点,这也符合 DF 统计量最初的设计是为了在单位根原假设下做检验的特点。

5. 总结

本文基于 Dickey-Fuller 比率统计量研究了分别以单位根过程和分整过程为原假设下的持久性变点的检验问题,并提出通过 Sieve Bootstrap 方法近似检验统计量的临界值来方便实际应用。数值模拟结果说明: Sieve Bootstrap 方法在单位根过程和分整过程两种原假设下近似的临界值结果都能够很好的控制经验水平;在两种备择假设下,当长记忆参数间隔越大检验统计量越容易检验到变点,比较靠前的变点具有较好的检验效果,而当变点位置靠后时检验效果较差。而对于相同跳跃度的变点,DF 统计量在检验从单位根过程向分整过程变化的持久性变点时的效果要优于检验从分整过程向单位根过程变化的持久性变点时的效果。

致 谢

本文是在我的导师陈占寿教授的指导下完成的,陈占寿教授本着严谨的科学态度,一丝不苟的学术精神,悉心育人。在此感谢陈占寿教授对我学业上的指导与关心,同时还要感谢国家自然科学基金、青海省自然科学基金对我的资助。

基金项目

国家自然科学基金(11301291, 11661067), 青海省自然科学基金(2015-ZJ-717)。

参考文献

- [1] Hakkio, A.W. and Rush, M. (1991) Is the Budget Deficit Too Large? *Econometrics Inquiry*, **29**, 429-445.
- [2] Kim, J.Y. (2000) Detection of Change in Persistence of Linear Times Series. *Journal of Econometrics*, **95**, 97-116. [https://doi.org/10.1016/S0304-4076\(99\)00031-7](https://doi.org/10.1016/S0304-4076(99)00031-7)
- [3] Kim, J.Y., Belaire Franch, J. and Badilli Amador, R. (2002) Corrigendum to “Detection of Change in Persistence of Linear Times Series”. *Journal of Econometrics*, **109**, 389-392. [https://doi.org/10.1016/S0304-4076\(02\)00087-8](https://doi.org/10.1016/S0304-4076(02)00087-8)
- [4] Perron, P. (1989) The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, **57**, 1361-1401. <https://doi.org/10.2307/1913712>
- [5] Perron, P. (1990) Testing for a Unit Root in a Time Series with a Changing Mean. *Journal of Business & Economic Statistics*, **8**, 153-162.
- [6] Anindya, Banerjee, Lumsdaine, R. and Stock, J. (1992) Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence. *Journal of Business & Economic Statistics*, **10**, 271-287.
- [7] Leybourne, S.J., Kim, T.H. and Taylor, A.M.R. (2006) Regression-Based Tests for a Change in Persistence. *Oxford Bulletin of Economics & Statistics*, **68**, 595-621. <https://doi.org/10.1111/j.1468-0084.2006.00179.x>
- [8] Hassler, U. and Scheithauer, J. (2011) Detecting Changes from Short to Long Memory. *Statistical Paper*, **52**, 847-870. <https://doi.org/10.1007/s00362-009-0292-y>
- [9] Sibbertsen, P. and Kruse, R. Testing for a Break in Persistence under Long-Range Dependencies.
- [10] Buhlmann, P. (1997) Sieve Bootstrap for Time Series. *Bernoulli*, **3**, 123-148. <https://doi.org/10.2307/3318584>
- [11] Poskitt, D.S. (2008) Properties of the Sieve Bootstrap for Fractionally Integrated and Non-Invertible Processes. *Journal of Time Series Analysis*, **29**, 224-250. <https://doi.org/10.1111/j.1467-9892.2007.00554.x>
- [12] Kapetanios, G. (2010) A Generalization of a Sieve Bootstrap Invariance Principle to Long Memory Processes. *Quantitative and Qualitative Analysis in Social Sciences*, **4**, 19-40.
- [13] Chen, Z., Tian, Z., Xiao, Q. and Xing, Y. (2014) Sieve Bootstrap Test for Variance Change in Long Memory Time Series. *International Joint Conference on Applied Mathematics*, **41**, 61-73.
- [14] Chen, Z., Tian, Z. and Xing, Y. (2016) Sieve Bootstrap Monitoring Persistence Change in Long Memory Process. *Statistics & Its Interface*, **9**, 37-45. <https://doi.org/10.4310/SII.2016.v9.n1.a4>
- [15] Chen, Z., Xing, Y. and Li, F. (2016) Sieve Bootstrap Monitoring for Change from Short to Long Memory. *Economics Letters*, **140**, 53-56. <https://doi.org/10.1016/j.econlet.2015.12.023>
- [16] Rambaccussing, D. (2015) A Test of the Long Memory Hypothesis Based on Self-Similarity. *Journal of Time Series Econometrics, Sire Discussion Papers*, **7**.
- [17] Robinson, P.M. (1995) Gaussian Semiparametric Estimation of Long Range Dependence. *Annals of Statistics*, **23**, 1630-1661. <https://doi.org/10.1214/aos/1176324317>

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2251, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>
期刊邮箱: sa@hanspub.org