

高次趋势单位根检验的小样本和有限样本分布

左秀霞

(江苏大学 财经学院, 江苏 镇江 212013)

摘要:文章采用蒙特卡罗仿真得到了小样本和有限样本下高次趋势DF单位根检验的统计量在常用显著性水平下的临界值,并与现有研究的结果进行了比较,在此基础上分析了其临界值在样本量和趋势项阶数变化过程中的变动规律。结果显示,临界值与现有研究非常接近。统计量的小样本和有限样本分布比渐近分布更分散,且其分散的趋势随样本量的减小而逐渐明显。随着趋势项阶数的增大,统计量的分布向左平移,但平移的幅度不断减小。

关键词:高次趋势;单位根检验;蒙特卡罗仿真

中图分类号:F224.0 **文献标识码:**A **文章编号:**1002-6487(2020)20-0033-06

0 引言

平稳性是时间序列的重要性质,近半个世纪以来,相关学者对序列的平稳性进行了大量研究,提出了ADF单位根检验、PP单位根检验和KPSS平稳性检验等检验序列平稳性的方法^[1-5]。目前,ADF检验的应用最为普遍。

上述检验主要是对检验统计量渐近理论的探讨,而实际经济数据对应的时间序列多是有限样本或小样本,其对应的检验统计量的分布可能会发生偏移,从而导致其临界值发生变化。后续研究对统计量在有限样本或小样本下的分布和临界值进行了探讨。Fuller(1996)、张晓峒等(1999)通过仿真得到了DF检验在有限样本和小样本下的临界值^[6,7]。周敏和蔺富明(2015)则对PP检验统计量在小样本下的临界值进行了仿真^[8]。张晓峒等(1999)、周敏和蔺富明(2015)还对DF和PP检验统计量在小样本下分布的有偏性进行了分析^[7,8]。

传统单位根检验是在带线性趋势项、截距项和无确定部分情形下讨论的,而多数宏观经济总量指标和企业总量指标序列都有高次趋势特征^[9]。此时,如果仍采用传统方法进行检验,势必会造成检验势的损失^[10]。后续研究在高次趋势的情形下扩展了传统单位根检验的理论,提出了高次趋势的ADF单位根检验和KPSS平稳性检验^[9,11]。那么,在有限样本或小样本下,高次趋势单位根检验的临界值有何变化呢?样本量和趋势项阶数对其临界值的影响又如何呢?本文将在小样本和有限样本下,用蒙特卡罗仿真得到带三次趋势项、二次趋势项、线性趋势项、截距项和无确定部分单位根检验在常用显著性水平下的临界值,将其与现有研究进行比较,分析其与渐近分布临界值的关系,并讨论其随样本量和趋势项阶数变化的规律,为检验有高次

趋势的序列的平稳性提供依据。由于ADF检验与DF检验有相同的检验统计量和渐近分布,且DF检验式更简单,不用考虑差分滞后阶数问题,因此本文将DF单位根检验为例进行讨论。

1 带高次趋势项的DF单位根检验理论

由左秀霞(2019)的研究可知^[9],带 $k+1$ 次趋势项的DF单位根检验的检验式设定为($k=2,1,0,-1$ 分别对应带三次、二次、线性趋势项、截距项情形):

$$\Delta y_t = c_0 + c_1 t + \dots + c_k t^k + b t^{k+1} + \gamma y_{t-1} + \varepsilon_t \quad (1)$$

无确定部分的DF单位根检验的检验式设定为:

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t \quad (2)$$

检验的原假设和备择假设分别为 $H_0: \gamma = 0$ 和 $H_1: \gamma < 0$ 。统计量及渐近分布为:

$$t_\gamma = \frac{\hat{\gamma} - 0}{\sigma_\gamma} \xrightarrow{L} \frac{\int_0^1 W^*(r) dW(r)}{(\int_0^1 W^*(r)^2 dr)^{1/2}} \quad (3)$$

其中, $\hat{\gamma}$ 和 σ_γ 分别是式(1)和式(2)中参数 γ 及其标准差 σ_γ 的OLS估计, $W^*(r) = W(r) - \int_0^1 W(s) F(s)' ds (\int_0^1 F(s) F(s)' ds)^{-1}$, $F(r)$ 是退化 $k+1$ 次趋势的布朗运动(当 $k=-1$ 时,是退化均值的布朗运动), $F(s) = (1, s, s^2, \dots, s^{k+1})'$, $F(s) = 0$ 对应无确定部分的DF检验,此时 $W^*(r)$ 是标准布朗运动 $W(r)$ 。

2 高次趋势DF单位根检验的小样本和有限样本临界值仿真

2.1 高次趋势DF单位根检验的小样本和有限样本临界值的蒙特卡罗仿真

为了与传统检验比较,本次仿真考虑了带三次趋势

基金项目:国家自然科学基金资助项目(71373105;11501254);江苏大学高级技术人才科研启动基金项目(12JDG131);镇江市软科学项目(RK201903);江苏省统计应用研究基地基金项目

作者简介:左秀霞(1981—),女,山东滨州人,博士,讲师,研究方向:非线性、非平稳时间序列分析方法。

项、二次趋势项、线性趋势项、截距项和无确定部分五种情形。样本量T分别取10、15、20、25、30、35、40、45、50、55、100、200、250、300和500。数据生成过程是 $y_t = y_{t-1} + \varepsilon_t, y_0 = 0, \varepsilon_t \sim IIN(0, 1)$, 五种情形的DF检验的检验式分别设定为 $\Delta y_t = c_0 + c_1 t + c_2 t^2 + b t^3 + \gamma y_{t-1} + \varepsilon_t$ 、 $\Delta y_t = c_0 + c_1 t + b t^2 + \gamma y_{t-1} + \varepsilon_t$ 、 $\Delta y_t = c_0 + b t + \gamma y_{t-1} + \varepsilon_t$ 、 $\Delta y_t = b + \gamma y_{t-1} + \varepsilon_t$ 和 $\Delta y_t = \gamma y_{t-1} + \varepsilon_t$ 。对每种情形,仿真的次数与文献[9]中渐近分布临界值的仿真次数相同,都是50000次,得到的1%、2.5%、5%、10%、90%、95%、97.5%、99%分位点的分位数,就是对应临界值。

经仿真,得到了上述五种情形的DF检验在小样本和有限样本时的临界值。由于Fuller(1996)、张晓峒等(1999)已经给出了带线性趋势项、截距项和无确定部分检验统计量的小样本和有限样本临界值^[6,7],并且由后面分析可知,本文结果与其非常接近,限于篇幅,表1只给出了带三次和二次趋势项时的临界值,另外三种情形只给出了T取300时的临界值,其他临界值以图形形式给出。同时,为了与渐近分布比较,表1还给出了T取 ∞ 的临界值^[9]。

表1 带高次趋势项的DF单位根检验的 t_τ 统计量临界值

检验式类型	样本容量	高次趋势DF单位根检验 t_τ 统计量小于表中数字的概率							
		0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
带三次趋势项	10	-8.44	-6.75	-5.70	-4.79	-1.53	-1.15	-0.81	-0.41
	15	-6.28	-5.50	-4.92	-4.35	-1.81	-1.49	-1.19	-0.84
	20	-5.67	-5.07	-4.65	-4.22	-1.93	-1.62	-1.34	-1.01
	25	-5.46	-4.96	-4.56	-4.15	-1.97	-1.69	-1.42	-1.11
	30	-5.25	-4.83	-4.48	-4.09	-2.00	-1.72	-1.45	-1.12
	35	-5.18	-4.78	-4.44	-4.06	-2.02	-1.74	-1.49	-1.17
	40	-5.12	-4.72	-4.40	-4.04	-2.04	-1.77	-1.50	-1.20
	45	-5.04	-4.66	-4.35	-4.01	-2.04	-1.77	-1.52	-1.22
	50	-5.03	-4.64	-4.35	-4.02	-2.05	-1.79	-1.55	-1.26
	55	-5.03	-4.66	-4.34	-4.00	-2.06	-1.79	-1.53	-1.23
	100	-4.90	-4.56	-4.27	-3.96	-2.08	-1.84	-1.61	-1.30
	200	-4.76	-4.46	-4.21	-3.92	-2.12	-1.86	-1.64	-1.35
	250	-4.77	-4.46	-4.21	-3.93	-2.10	-1.85	-1.61	-1.30
	300	-4.76	-4.46	-4.21	-3.93	-2.11	-1.86	-1.61	-1.31
∞	-4.70	-4.42	-4.16	-3.89	-2.12	-1.87	-1.64	-1.34	
带二次趋势项	10	-6.85	-5.67	-4.88	-4.14	-1.28	-0.91	-0.58	-0.20
	15	-5.51	-4.86	-4.36	-3.88	-1.52	-1.20	-0.90	-0.54
	20	-5.17	-4.58	-4.20	-3.76	-1.57	-1.25	-0.96	-0.62
	25	-4.94	-4.48	-4.11	-3.72	-1.61	-1.31	-1.02	-0.68
	30	-4.82	-4.40	-4.06	-3.68	-1.63	-1.33	-1.06	-0.73
	35	-4.77	-4.34	-4.03	-3.68	-1.65	-1.36	-1.08	-0.75
	40	-4.71	-4.32	-3.99	-3.64	-1.67	-1.38	-1.11	-0.78
	45	-4.64	-4.28	-3.98	-3.64	-1.66	-1.37	-1.11	-0.82
	50	-4.64	-4.28	-3.97	-3.63	-1.68	-1.40	-1.13	-0.80
	55	-4.61	-4.23	-3.94	-3.62	-1.67	-1.39	-1.13	-0.81
	100	-4.50	-4.17	-3.90	-3.59	-1.71	-1.43	-1.16	-0.84
	200	-4.42	-4.10	-3.85	-3.57	-1.72	-1.44	-1.18	-0.89
	250	-4.41	-4.12	-3.86	-3.57	-1.71	-1.44	-1.18	-0.86
	300	-4.43	-4.11	-3.86	-3.57	-1.72	-1.46	-1.20	-0.88
∞	-4.38	-4.10	-3.85	-3.56	-1.73	-1.45	-1.19	-0.89	
带线性趋势项	300	-4.00	-3.68	-3.42	-3.13	-1.24	-0.93	-0.65	-0.32
∞	-3.96	-3.66	-3.40	-3.11	-1.24	-0.93	-0.64	-0.30	
带截距项	300	-3.47	-3.14	-2.87	-2.57	-0.43	-0.06	0.27	0.65
∞	-3.41	-3.11	-2.86	-2.57	-0.45	-0.09	0.24	0.61	
无确定部分	300	-2.60	-2.24	-1.95	-1.62	0.89	1.27	1.62	2.00
∞	-2.57	-2.21	-1.93	-1.61	0.88	1.29	1.62	2.01	

2.2 仿真过程中出现的共线性问题

在对带三次趋势项检验的仿真过程中,当 $T \geq 300$ 时,估计检验式时会出现 $X'X$ 近似奇异矩阵的情况,可能会影响计算精度,导致仿真结果扭曲。出现该现象的原因可能是:当趋势项阶数较高且样本量较大时,序列中高次趋势对应的确定性趋势部分占的比例很高,从而使序列中的确定性趋势部分起到支配作用,掩盖了其随机部分,导致检验式的自变量 X 中 t^3 与 y_{t-1} 之间出现了共线性问题。并且,高次趋势项系数或趋势项阶数越大,共线性就越严重。

2.3 仿真结果与Fuller(1996)^[6]结果的比较

为了与传统单位根检验进行比较,在带线性趋势项、截距项和无确定部分三种情形下,样本量T分别取25、50、100、250、500和 ∞ 时,将本文仿真的临界值与Fuller(1996)^[6]的结果进行比较。图1至下页图6给出了这三种情形临界值的图形。

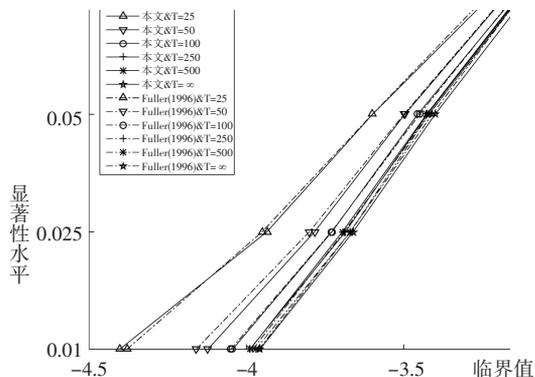


图1 本文与Fuller(1996)带线性趋势项检验左侧临界值

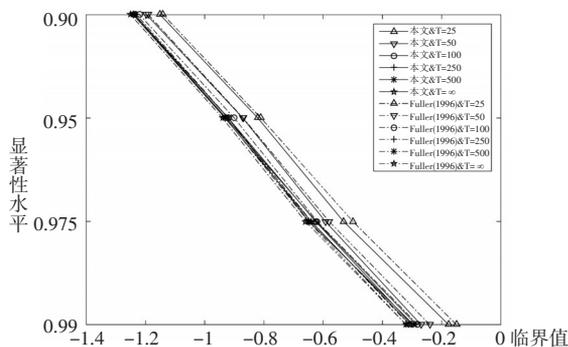


图2 本文与Fuller(1996)带线性趋势项检验右侧临界值

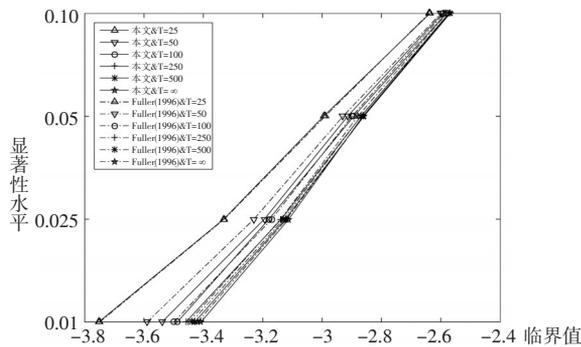


图3 本文与Fuller(1996)带截距项检验左侧临界值

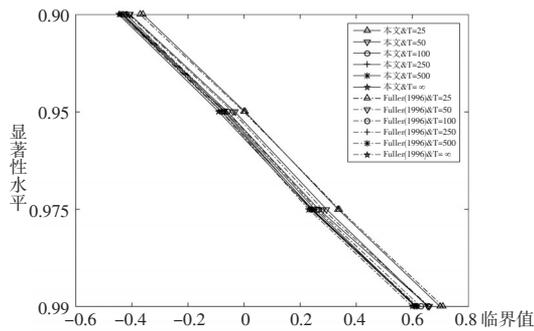


图4 本文与 Fuller(1996)带截距项检验右侧临界值

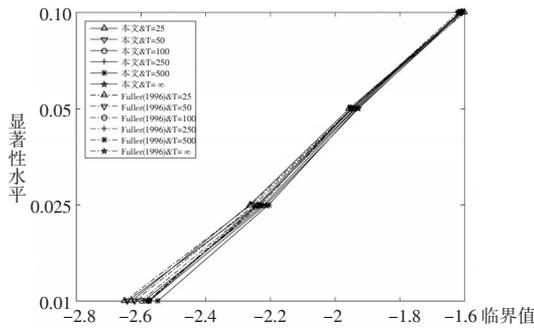


图5 本文与 Fuller(1996)无确定部分检验左侧临界值

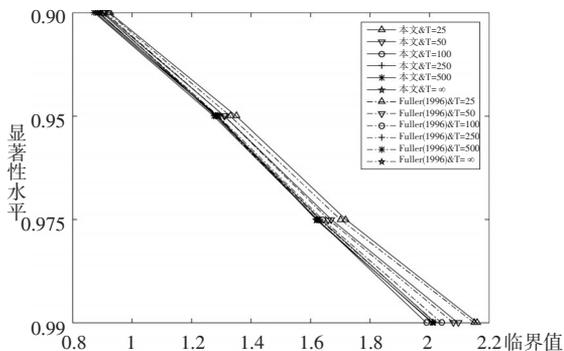


图6 本文与 Fuller(1996)无确定部分检验右侧临界值

图1至图6显示,在上述样本量和趋势项阶数相同时,本文与Fuller(1996)临界值非常接近;并且,趋势项阶数越小,样本量越大,本文与Fuller(1996)的临界值越接近。

2.4 仿真结果与张晓峒等(1999)结果的比较

在带线性趋势项、截距项和无确定部分三种情形下,样本量T分别取10、15、20、25、30、35、40、45、50和55时,将本文的临界值与张晓峒等(1999)的结果进行比较。图7至图9给出了这三种情形下临界值的图形。

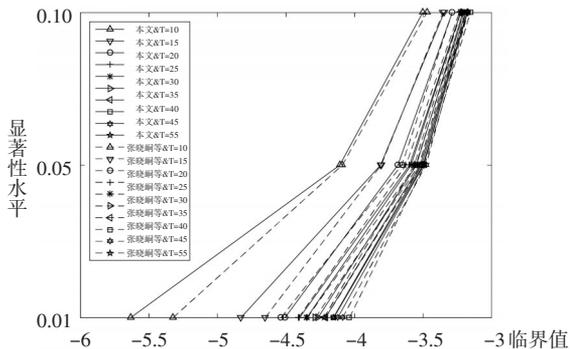


图7 本文与张晓峒等(1999)带线性趋势项检验左侧临界值

图7至图9显示,在上述样本量和趋势项阶数相同时,本文与张晓峒等(1999)的临界值非常接近。并且,趋势项

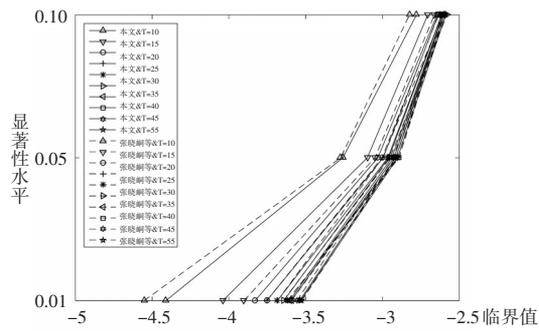


图8 本文与张晓峒等(1999)带截距项检验左侧临界值

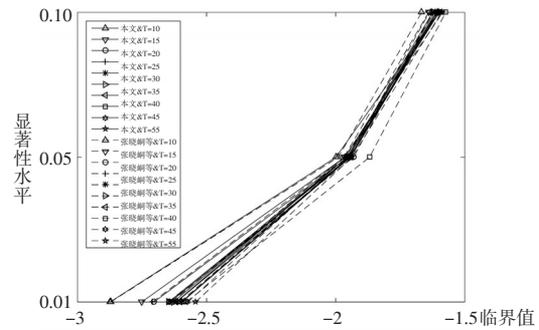


图9 本文与张晓峒等(1999)无确定部分检验左侧临界值

阶数越小,样本量越大,本文与张晓峒等(1999)的临界值越接近。

3 仿真结果分析

3.1 小样本和有限样本临界值与渐近分布临界值的比较

为了对小样本和有限样本下统计量的临界值与渐近分布进行比较,图10至图19分别给出了带三次趋势项、二次趋势项、线性趋势项、截距项和无确定部分情形下,不同样本量对应的DF单位根检验临界值的图形。

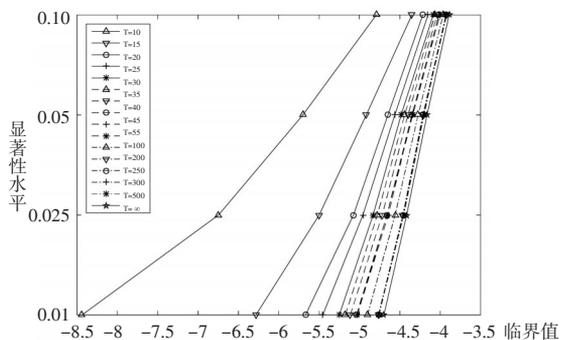


图10 不同样本量带三次趋势项检验的左侧临界值

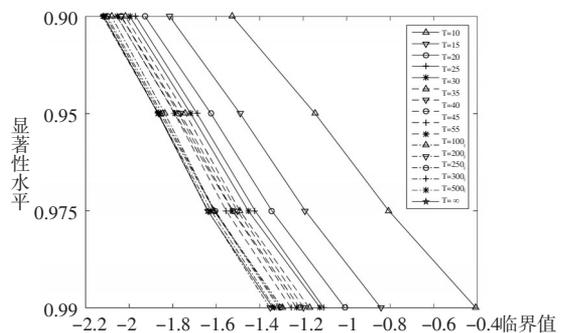


图11 不同样本量带三次趋势项检验的右侧临界值

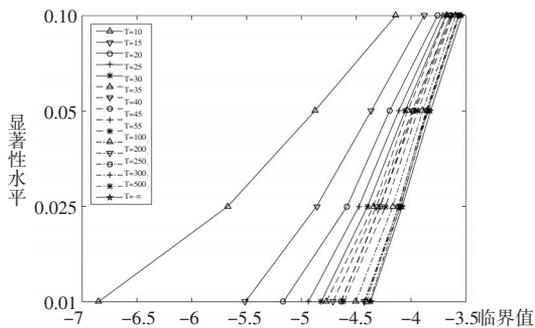


图12 不同样本量带二次趋势项检验的左侧临界值

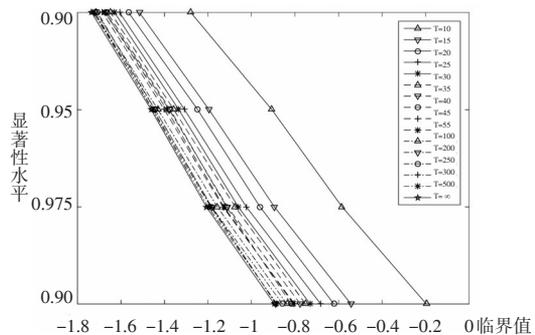


图13 不同样本量带二次趋势项检验的右侧临界值

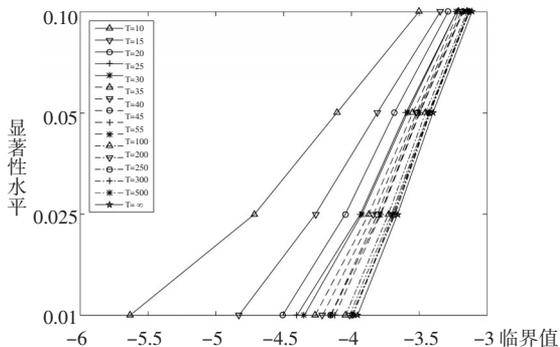


图14 不同样本量带线性趋势项检验的左侧临界值

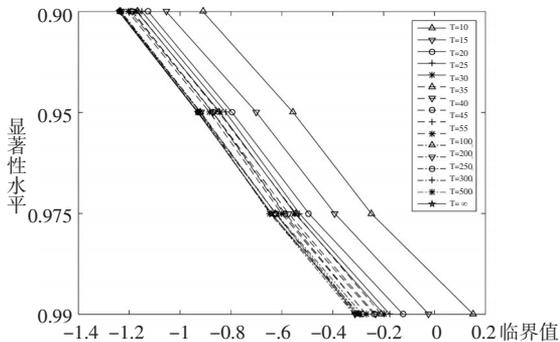


图15 不同样本量带线性趋势项检验的右侧临界值

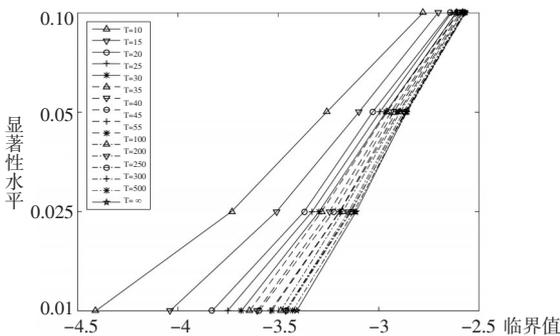


图16 不同样本量带截距项检验的左侧临界值

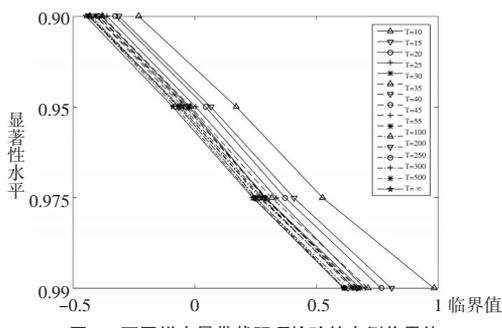


图17 不同样本量带截距项检验的右侧临界值

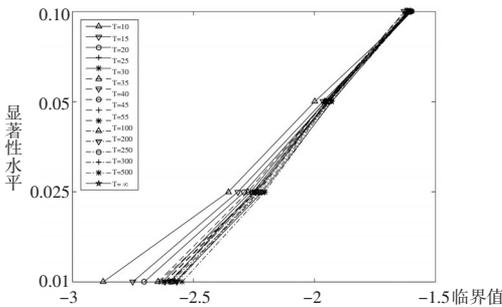


图18 不同样本量无确定部分检验的左侧临界值

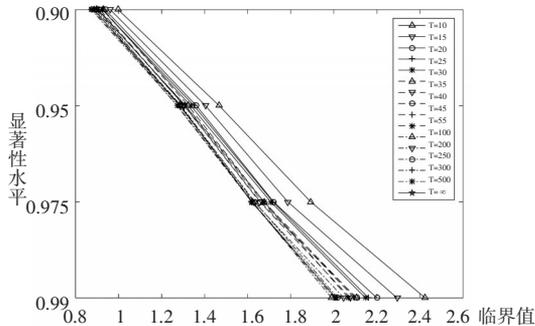


图19 不同样本量无确定部分检验的右侧临界值

图10至图19显示,小样本和有限样本下1%~10%的临界值比渐近分布下的临界值要小,而90%~99%的临界值则比渐近分布下临界值要大。由此可见,统计量小样本和有限样本分布向两侧延伸得更多,即统计量的小样本和有限样本分布比渐近分布更加分散。图1至图6对应的Fuller(1996)及图7至图9对应的张晓峒等(1999)的结果也体现了这一点。

3.2 统计量有限样本临界值随样本量增加发生的变化

从图10至图19中还可发现:(1)随着样本量的减小,相同趋势项阶数对应的1%~10%的临界值变小,而90%~99%的临界值则增大。可见,随着样本量的减小,统计量的分布越分散。(2)随着样本量增大,相同趋势项阶数对应1%~10%的临界值变大,而90%~99%的临界值则变小,并且相同趋势项阶数和分位点对应的临界值与渐近分布临界值越来越接近。可见,随着样本量的增大,统计量的分布向渐近分布集中,与渐近分布逐步接近。图1至图6对应的Fuller(1996)及图7至图9对应的张晓峒等(1999)的结果也体现了这两点。(3)当样本量T达到500时,相同趋势项阶数对应的临界值已经与渐近分布临界值非常接近,有的1%~10%临界值甚至比渐近分布临界值还要大,也有的90%~99%的临界值甚至比渐近分布临界值还要小。可

见,当样本量较大时,有的分布较渐近分布更集中,出现了反方向变化,即对应的左侧临界值更大,右侧临界值更小。

3.3 统计量分布随趋势项阶数的增加发生的变化

为了比较不同趋势项阶数DF单位根检验统计量的临界值,图20至图27分别给出了样本量T为(10,15,20,25)、(30,35,40,45)、(50,55,100,200)和(250,300,500,∞)情形下,带三次趋势项、二次趋势项、线性趋势项、截距项和无确定部分的DF单位根检验对应临界值的图形。

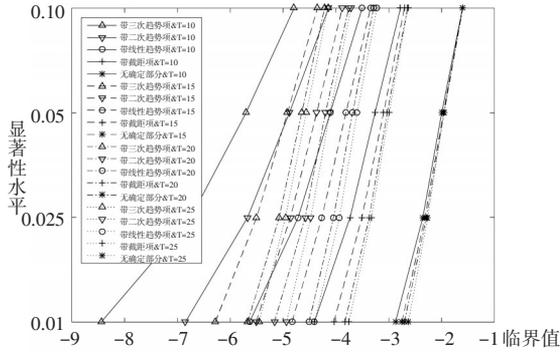


图20 T=10、15、20、25 检验的左侧临界值

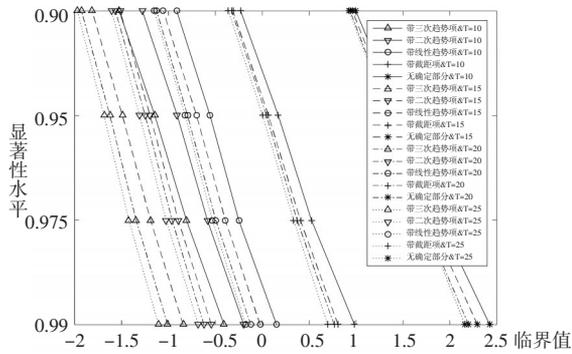


图21 T=10、15、20、25 检验的右侧临界值

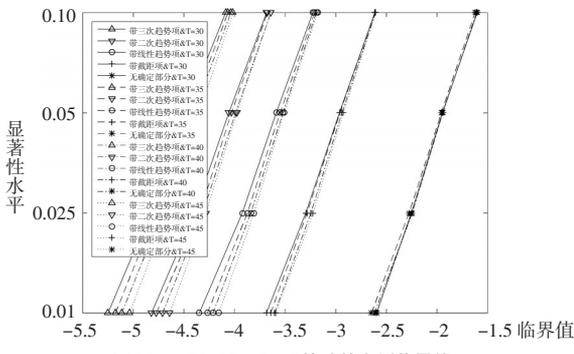


图22 T=30、35、40、45 检验的左侧临界值

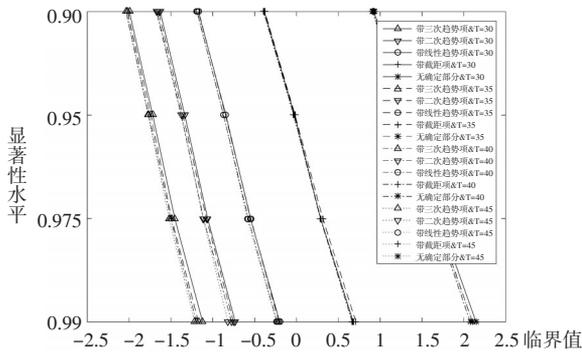


图23 T=30、35、40、45 检验的右侧临界值

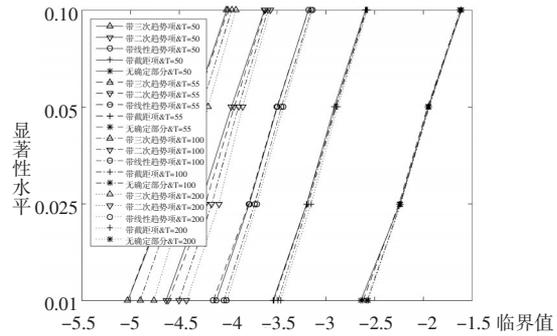


图24 T=50、55、100、200 检验的左侧临界值

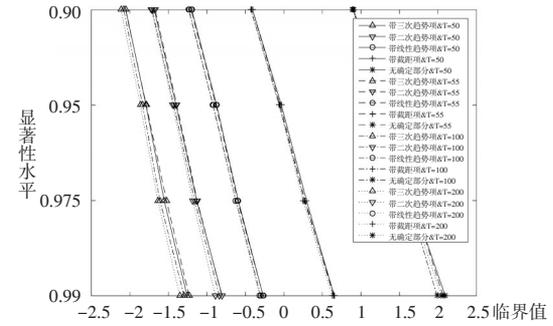


图25 T=50、55、100、200 检验的右侧临界值

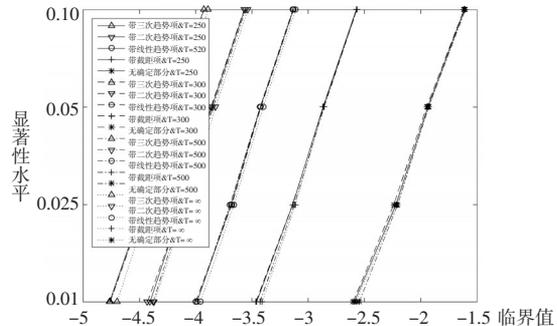


图26 T=250、300、500、∞ 检验的左侧临界值

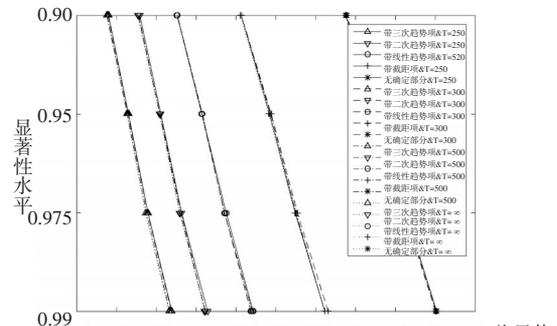


图27 T=250、300、500、∞ 检验的右侧临界值

从图20至图27可以发现:(1)随着趋势项阶数的增大,1%~10%临界值和90%~99%临界值都变小。可见,随着趋势项阶数的增大,统计量的小样本、有限样本和渐近分布都向左平移。(2)随着趋势项阶数的增大,同一样本量和分位点的临界值越接近。可见,随着趋势项阶数的增大,统计量的小样本、有限样本和渐近分布向左平移的幅度在减小。图1至图6对应的 Fuller(1996)及图7至图9对应的张晓峒等(1999)的结果也体现了这两点。(3)比较图20和图21、图22和图23、图24和图25以及图26和图27可以发现,随着趋势项阶数的减小和样本量的增大,不同样本量、相同趋势项阶数和相同分位点对应的临界值越接

近。这说明统计量的小样本、有限样本和渐近分布随着趋势项阶数的减小而不断接近,随着样本量的增大越集中。

4 结论

本文仿真得到了小样本和有限样本的高次趋势DF单位根检验在常用显著性水平下的临界值,并将其与现有研究进行了比较,还分析了该临界值随样本量和趋势项阶数变化的变动规律。结论如下:

(1)在带线性趋势项、截距项和无确定部分情形下,本文与Fuller(1996)、张晓峒等(1999)的临界值非常接近。

(2)统计量的小样本和有限样本分布比渐近分布更加分散。随着样本量的减小,统计量的分布越分散;随着样本量的增大,统计量的分布越向渐近分布集中,与渐近分布不断接近。因此,在有限样本或小样本时,如果采用渐近分布的临界值检验,就更容易拒绝原假设,即将一个序列误判为平稳序列,从而降低了检验的实际水平,导致检验出现水平扭曲。

(3)随着趋势项阶数的增大,统计量的小样本、有限样本和渐近分布都向左平移,但其平移的幅度在减小。随着趋势项阶数的减小,统计量的有限样本、小样本都与渐近分布不断接近。因此,如果对高次趋势的序列采用传统检验式进行检验,就更容易拒绝原假设,即将一个序列误判为平稳序列,从而降低了检验的实际水平,导致检验出现水平扭曲。

(4)为了尽可能降低检验的水平扭曲,对样本量有限的高次趋势序列进行单位根检验时,应该采用对应高次趋

势单位根检验的有限样本或小样本临界值。

参考文献:

- [1]Dickey D A, Fuller W A. Distribution of the Estimators for Autoregressive Time Series With a Unit Root [J].Journal of the American Statistical Association, 1979, 74(366).
- [2]Dickey D A, Fuller W A. Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root [J].Econometrica, 1981, 49(4).
- [3]Phillips P C B. Time Series Regression With a Unit Root [J].Econometrica, 1987, 55(2).
- [4]Phillips P C B, Perron P. Testing for a Unit Root in Time Series Regression [J]. Biometrika, 1988, 75(2).
- [5]Kwiatkowski D, Phillips P C B, Schmidt P, et al. Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root [J]. Journal of Econometrics, 1992, (54).
- [6]Fuller W A. Introduction to Statistical Time Series (Second Edition) [M]. New York: John Wiley, 1996.
- [7]张晓峒, 大川勉, 张世英.小样本DF统计量的分布特征[J].系统工程理论与实践, 1999,(3).
- [8]周敏, 蔺富明.小样本PP检验统计量的分布特征[J].西南大学学报(自然科学版), 2015, 37(5).
- [9]左秀霞.带高次趋势项的ADF单位根检验[J].数量经济技术经济研究, 2019, 36(1).
- [10]刘田. ADF与PP单位根检验法对非线性趋势平稳序列的伪检验[J].数量经济技术经济研究, 2008,(6).
- [11]Busetti F, Taylor A M R. Variance Shifts, Structural Breaks, and Stationarity Tests [J].Journal of Business & Economic Statistics, 2003, 21(4).

(责任编辑/浩 天)

Distribution of Small and Finite Samples of Unit Root Test With High-order Trend

Zuo Xiuxia

(School of Finance & Economics, Jiangsu University, Zhenjiang Jiangsu 212013, China)

Abstract: By using Monte Carlo simulation, this paper obtains the critical value of the DF unit root test statistics with high-order trend in small and finite samples at the commonly used significance level, and also compares it with the results of existing researches. On this basis, the paper analyzes the change rule of its critical value in the process of sample size and trend term order change. The results show that the critical value is very close to that in existing researches, that the small and finite samples distribution of the statistics are more dispersed than the asymptotic distribution, that the trend of dispersion becomes more obvious with the decrease of sample size, and that the distribution of statistics moves towards the left and the amplitude decreases continuously with the increase of the trend term order.

Key words: high-order trend; unit root test; Monte Carlo simulation