日本統計学会誌
第 44 巻, 第 1 号, 2014 年 9 月
61 頁 ～ 74 頁

特　集

レベル・シフトの検定と検出力の非単調性

山崎　大輔*，黒住　英司*

Tests for a Level Shift and the Non-Monotonic Power Problem

Daisuke Yamazaki* and Eiji Kurozumi*

本稿では，誤差項に系列相関が存在する場合の定数項シフトの検定について考察する．先行研究により，LM タイプの検定の検出力は，誤差項に系列相関がある場合には，構造変化の度合いが大きくなるほど検出力が下がってしまうという「検出力の非単調性問題」が存在することが知られている．一方，ワルド・タイプの検定では，検出力の非単調性問題は存在しないものの，検定のサイズが大幅に歪む場合がある．先行研究ではこれらの問題を回避する方法が提案されており，本稿でも，既存の方法をさらに修正することにより，有限標本特性がより優れている検定方法を提案する．新たな手法の有限標本特性は，モンテ・カルロ実験で分析され，サイズ・検出力の両面で，既存の検定よりも優れていてることが確認された．

This paper investigates tests for a shift in mean for serially correlated time series. It is known in the literature that LM-type tests suffer from the so-called non-monotonic power problem in the sense that the power decreases even if the magnitude of break gets larger. On the other hand, Wald-type tests do not have the non-monotonic power problem but they tend to suffer from over-size distortion. Several methods to overcome these problems have been developed in the literature and we propose a new test with a good finite sample property by modifying the existing test. We investigate the finite sample property of our new method and it turns out that our test performs better than the other tests in view of both size and power.

キーワード: 構造変化，ワルド検定，LM 検定

1. はじめに

長期の時系列データの分析を行う場合，構造変化の有無を検討する必要があるため，計量経済学および統計学の分野において，数多くの構造変化の検定が提案されてきている．古典的な文献においては，構造変化点を予め設定しているものも多いが，実証分析においては，変化点を予め特定化することが困難な場合も多い．そのため，Brown *et al.* (1975) で提案された CUSUM 検定や CUSUM-SQ 検定のように，構造変化点を未知とした検定方法が必要となってくる．また，経済時系列モデルにおいては，説明変数を確率変数として扱

* 一橋大学大学院経済学研究科：〒 186-8601 東京都国立市中 2-1 (E-mail: kurozumi@stat.hit-u.ac.jp).

う必要があり，誤差項に関しては一般的には系列相関が存在するため，これらを考慮した構造変化の検定手法が開発されてきた．中でも，Andrews (1993) による sup タイプの検定は，GMM 推定量に基づく検定であるため，線形モデルばかりでなく，非線形モデルにも応用できるという点で非常に有用であり，多くの実証研究で使われている．また，Andrews and Ploberger (1994) や Andrews *et al.* (1996) による exponential タイプの検定や average タイプの検定は，線形モデルというフレームワークにおいてある種の最適性を満たしているため，これらの検定もしばしば実際に使われている．なお，Sowell (1996) は，この 2 つの検定について，GMM 推定量を用いて検定統計量を構築した場合にも，検定の最適性が保持されることを示している．

これら sup タイプ，exponential タイプ，average タイプの検定は，基本的には構造変化が一度起きているという対立仮説を想定して構築されているが，構造変化が複数回起きている可能性がある場合には，そのような対立仮説を想定した検定統計量の構築が望ましい．Bai and Perron (1998) では，1 変量時系列モデルにおいて，構造変化が高々 m 回起きたという対立仮説を想定した sup タイプの検定を考案しており，Qu and Perron (2007) では，この検定を多変量モデルへ拡張している．また，Elliott and Müller (2006) では，パラメータがランダム・ウォークのように変化するモデルを用いて，パラメータの安定性に関する最適検定を導出している．一方で，一連の仮説検定で構造変化が起きたという帰無仮説が棄却された場合，実際には構造変化が何度起きたのか，その回数を推定する必要がある．Bai (1997) では対立仮説に一度の構造変化を想定した検定を繰り返し利用する方法を，また， Bai and Perron (1998) や Qu and Perron (2007) では，複数の構造変化を対立仮説に想定した検定を逐次的に利用する方法を提案しており，Kurozumi and Tuvaandorj (2011) は，修正 AIC や修正 BIC といった情報量規準による推定方法を提案している．構造変化に関する統計的推測に関する概説は，Csörgő and Horváth (1997), Perron (2006) や Aue and Horváth (2013) などを参照されたい．

構造変化の検定においては，変化の大きさが大きければ大きいほど，検定の検出力が高くなることが期待されるし，漸近的にもそのようになることがある一定の仮定の下では証明されている．しかしながら，Vogelsang (1999) は，モンテ・カルロ実験により，誤差項に系列相関の存在を想定した検定や，ラグつき内生変数が説明変数に含まれる場合には，有限標本において CUSUM 検定や LM タイプの検定では，変化の大きさが大きくなると逆に検出力が下がってしまうという，検出力の非単調性問題があることを指摘した．このような検出力の非単調性問題は，系列相関が存在するときの長期分散の推定が問題であることが知られており，実際，Crainiceanu and Vogelsang (2007), Deng and Perron (2008) や Perron and Yamamoto (2014) では，この点をモンテ・カルロ実験及び理論的に分析している．一方，ワルド・タイプの検定では検出力の非単調性問題は生じないものの，誤差項

の系列相関を考慮して検定統計量を構築すると，帰無仮説の下で検定のサイズが大幅に歪むことも知られている．このことから，最近では，検定のサイズが安定しており，かつ，検出力の非単調性問題が生じない検定方法が提案されている．例えば，Kejriwal (2009) は，帰無仮説と対立仮説の両仮説下での推定残差を用いて長期分散の推定量を修正することにより，問題の回避を試みている．同様に，Juhl and Xiao (2009) においては，ノンパラメトリック法を用いて長期分散の推定精度の向上を提案しているが，残念ながら検定のサイズや検出力は，ノンパラメトリック回帰を行う際のバンド幅に強く依存してしまうことが分かっている．一方，Shao and Zhang (2010) では，self-normalization 法を用いて，長期分散そのものの推定を避けた検定を考案しており，Sayginsoy and Vogensang (2011) では，長期分散を推定する際のバンド幅を標本の大きさ T に比例させるという，fixed-b 漸近近似法という手法により，検出力の非単調性問題を克服した検定方法を提案している．

本稿では，これらの手法を概観したうえで，既存のレベル・シフトの検定方法の修正を提案し，その有限標本の特性を向上させることを提案する．具体的には，Kejriwal (2009) の方法をさらに修正することで，検定のサイズが安定しており，かつ，検出力の非単調性問題を回避できる検定統計量を構築する．検定統計量の作成自体は比較的シンプルであり，実証分析に応用することも容易である．

以下の節では，まず，第 2 節において，モンテ・カルロ実験により，レベル・シフトの検定において，誤差項に系列相関がある場合には，検出力の非単調性問題やサイズの歪みが生じることを確認し，その原因についても考察する．第 3 節では，これらの問題を回避するための既存の検定方法を概観したうえで，新たな検定方法を提案する．第 4 節では，モンテ・カルロ実験により，既存の検定方法と新たな方法の有限標本特性を比較し，各検定統計量の特性を分析する．第 5 節では，今後の課題について検討を行う．

2. sup タイプ検定の問題点

本節では，実証研究で広く使われている sup タイプの検定の問題点について考察する．まずはじめに，我々は観測値のレベル・シフトの有無に興味を持っており，以下のようなモデルを考える．

$$y_t = \mu_1 + \mu_2 DU_t(T_B^0) + u_t, \quad (t = 1, \ldots, T), \tag{2.1}$$

ただし，$DU_t(T_B^0) = 0$ $(t = 1, \ldots, T_B^0)$，$DU_t(T_B^0) = 1$ $(t = T_B^0 + 1, \ldots, T)$ で，u_t は平均 0 の確率過程であるとする．この場合，y_t の平均は $t = 1, \ldots, T_B^0$ までは μ_1 であるが，$T_B^0 + 1$ 期以降は $\mu_1 + \mu_2$ となるので，モデル (2.1) では，レベル・シフトが起きている．ただし，$\mu_2 = 0$ の場合は，レベル・シフトが起きてない，平均が一定のモデルとなるので，我々が興味を持つ検定問題は，

$$H_0: \ \mu_2 = 0 \quad \text{vs.} \quad H_1: \ \mu_2 \neq 0$$

と表すことができる．なお，構造変化点 T_B^0 は未知であるとする．

ここで，誤差項と変化点については，以下を仮定する．

仮定 1 (i) $\{u_t\}$ は平均 0 で，$T \to \infty$ の時，以下の汎関数中心極限定理が成り立つとする．

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} u_t \xrightarrow{d} \sigma W(r) \quad (0 \le r \le 1). \tag{2.2}$$

ただし，$W(r)$ は 1 次元標準ブラウン運動とする．

(ii) $T \to \infty$ の時，$T_B^0/T \to \lambda^0$ が成立つ $(0 < \lambda^0 < 1)$．

仮定 1(i) では，u_t に特定のモデルは想定せず，その部分和について汎関数中心極限定理が成り立つという，比較的ゆるやかな仮定を想定しているが，u_t が単位根過程である可能性は排除している．(ii) は構造変化に関する文献では標準的な仮定で，構造変化点と標本の大きさの比 (構造変化点比率) が一定であるというものである．この仮定は，構造変化点前・後それぞれの期間において，十分な大きさの標本が得られる，ということを意味している．

構造変化点が未知の場合の検定として，ここでは実証分析でもしばしば用いられる Andrews (1993) の sup タイプのワルド検定と LM 検定を考える．今，構造変化点を T_B(ただし，$T_B/T = \lambda$ とする) と想定した時の (2.1) 式の最小 2 乗推定を

$$y_t = \hat{\mu}_1 + \hat{\mu}_2 DU_t(T_B) + \hat{u}_t(\lambda)$$

とし，この時の残差平方和を $SSR_1(\lambda)$ とする．一方，帰無仮説の下での最小 2 乗推定を

$$y_t = \tilde{\mu}_1 + \tilde{u}_t$$

とし，この時の残差平方和を SSR_0 とする．すると，ワルド検定統計量と LM 検定統計量はそれぞれ，

$$W_T(\lambda) = \frac{SSR_0 - SSR_1(\lambda)}{\hat{\sigma}^2(\lambda)}, \quad LM_T(\lambda) = \frac{SSR_0 - SSR_1(\lambda)}{\tilde{\sigma}^2}$$

となる．ただし，$\hat{\sigma}^2(\lambda)$ は対立仮説の下での推定残差 $\hat{u}_t(\lambda)$ に基づく長期分散 σ^2 のカーネル推定量で，

$$\hat{\sigma}^2(\lambda) = \hat{\gamma}_0(\lambda) + 2\sum_{j=1}^{T-1} k(j,m)\hat{\gamma}_j(\lambda), \quad \hat{\gamma}_j(\lambda) = \frac{1}{T}\sum_{t=j+1}^{T} \hat{u}_t(\lambda)\hat{u}_{t-j}(\lambda) \tag{2.3}$$

と定義され，$k(j,m)$ はカーネル，m はバンド幅である．同様に，$\tilde{\sigma}^2$ は帰無仮説の下での推定残差 $\tilde{u}_t$ に基づく σ^2 のカーネル推定量であり，(2.3) 式において，$\hat{\gamma}_j(\lambda)$ を $\tilde{\gamma}_j =$

$T^{-1}\sum_{t=j+1}^{T}\tilde{u}_t\tilde{u}_{t-j}$ に置き換えたものである．ここで，λ のとりえる可能な範囲を $\Lambda \equiv [\underline{\lambda},\ \overline{\lambda}]$ と設定すると[1)]，sup ワルド検定統計量と supLM 検定統計量は，次のようになる．

$$\sup W_T = \sup_{\lambda\in\Lambda} W_T(\lambda), \quad \sup LM_T = \sup_{\lambda\in\Lambda} LM_T(\lambda)$$

Andrews (1993) により，これらの検定統計量の帰無仮説の下での極限分布は，

$$\sup W_T,\ \sup LM_T \ \xrightarrow{d} \sup_{\lambda\in\Lambda}\frac{\{W(r)-rW(1)\}^2}{\lambda(1-\lambda)}$$

となることが示されており，また，p-value を求めるための近似式が Hansen (1997) で求められている．

ここで，sup ワルド検定と supLM 検定の有限標本特性をモンテ・カルロ実験で確認する．データ生成過程は，

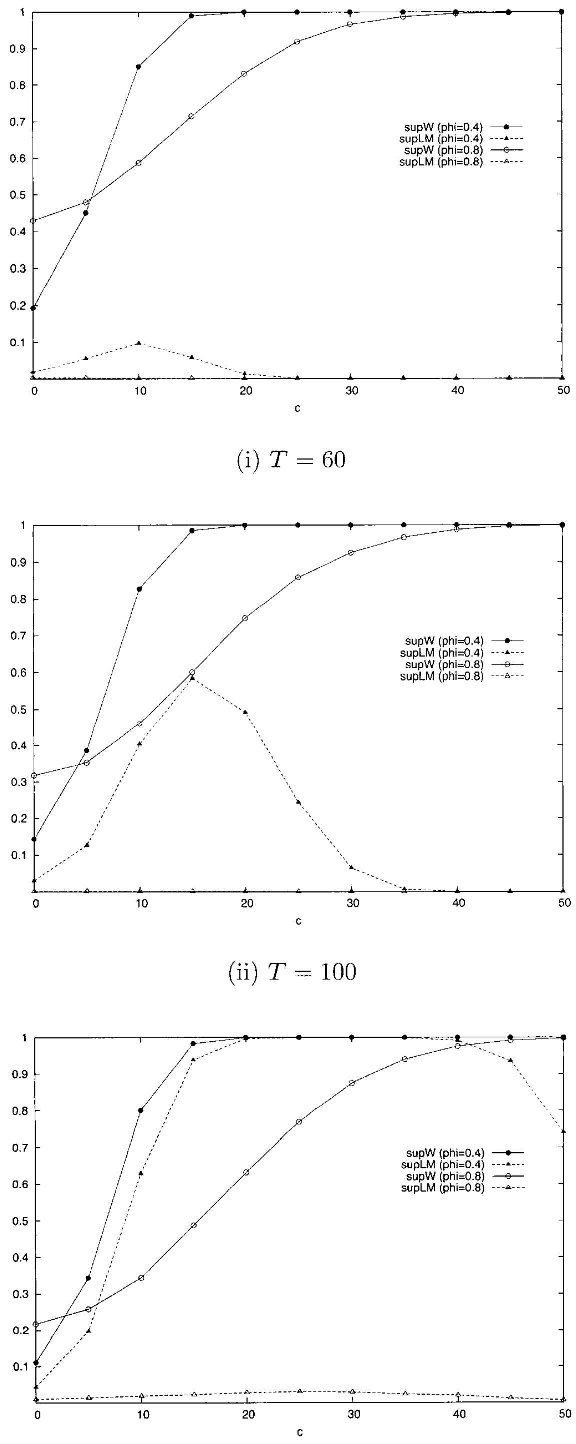
$$y_t = (c/\sqrt{T}) \times DU_t(T/2) + u_t, \quad u_t = \phi u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}N(0,1)$$

として様々な c の値でシミュレーションを行い，検定のサイズと検出力を調べることにする．ただし，$\hat{\sigma}^2(\lambda)$ および $\tilde{\sigma}^2$ の推定には，$k(j,m)$ として quadratic spectral (QS) カーネルを用い，バンド幅 m は推定残差 $\hat{u}_t(\lambda)$ もしくは $\tilde{u}_t$ を用いて Andrews (1991) で提案されている方法で決めるものとする．想定する構造変化点の範囲は，$\Lambda = [0.15,\ 0.85]$ とする．有意水準は 5%，繰り返し回数は 5,000 回で，すべての計算は行列演算ソフトウェア GAUSS で行われている．

図 1(i) は，$T = 60$ で，$\phi = 0.4$ または $\phi = 0.8$ の場合の結果である．sup ワルド検定に関しては，検定のサイズの歪みが大きく，特に $\phi = 0.8$ の場合はサイズは 0.4 を超えてしまっており，検定のサイズを制御できていないことがうかがえる．一方，supLM 検定については，$\phi = 0.4$ の場合には検定のサイズは 5%に近く，c の値が比較的小さな場合は検出力が上昇するが，c が 10 を超えると検出力が下がっている．これが Vogelsang (1999) により指摘された検出力の非単調性問題であり，CUSUM 検定や Elliott and Müller (2006) の検定などにも同様の問題があることが知られている．さらに，$\phi = 0.8$ の場合には，検出力は一貫して有意水準より低く，supLM 検定では構造変化を検出することができないことが分かる．図 1(ii) および (iii) より，標本の大きさを $T = 100$ および 200 とした場合には，サイズと検出力の問題は多少緩和されるものの，やはり，sup ワルド検定のサイズは 5%を大幅に超え，supLM 検定の検出力は，構造変化の大きさに対して非単調性を持つことが確認できる．

これらの問題は，基本的に σ^2 の推定が原因で生じることが知られている．sup ワルド検定の場合は，σ^2 を対立仮説の下で推定した場合の残差を用いてカーネル推定を行っている

[1)] Λ は，$\underline{\lambda} > 0, \overline{\lambda} < 1$ を満たす区間であるとする．



(i) $T = 60$

(ii) $T = 100$

(iii) $T = 200$

図 1 sup ワルド・sup LM 検定のサイズと検出力.

が，この場合，$\hat{\sigma}^2$ には負のバイアスがあり，過小推定されやすいため，検定のサイズが大きくなってしまう．一方，supLM 検定の場合は，σ^2 を帰無仮説の下で推定した場合の残差を用いて推定しているが，対立仮説が正しい場合，帰無仮説で推定した場合の残差は分散が非常に大きくなるため，$\tilde{\sigma}^2$ も非常に大きな値となる．そのため，検出力の非単調性問題

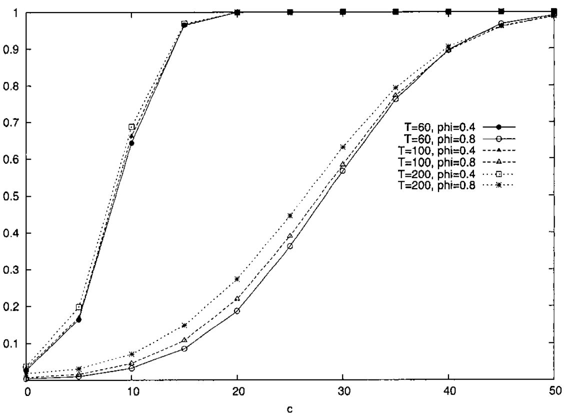


図 2　長期分散が既知の場合の sup タイプ検定のサイズと検出力.

が生じることになる．実際，真の長期分散を用いて求めた sup タイプ検定のサイズと検出力を図 2 に示す．サイズが 5%より低くなる傾向はあるものの，上記のような過大なサイズや検出力の非単調性は無くなっていることが確かめられる．

3.　非単調検出力を考慮した構造変化の検定

本節では，検定のサイズを制御しつつ，第 2 節で取り上げた検出力の非単調性問題を克服すべく考案された構造変化の検定を紹介し，さらに，その一つの検定統計量のさらなる修正方法を提案する．

3.1　長期分散のハイブリッド推定による sup ワルド検定 (Kejriwal (2009))

Kejriwal (2009) は，sup タイプ検定のサイズの歪みや検出力の非単調性の原因が長期分散の推定であることに注目し，次のような検定統計量を提案している．

$$\sup W_T^{kej} = \sup_{\lambda\in\Lambda}\frac{SSR_0 - SSR_1(\lambda)}{\hat{\sigma}_h^2}, \quad \hat{\sigma}_h^2 = \hat{\gamma}_0(\hat{\lambda}) + 2\sum_{j=1}^{T-1} k(j, m(\hat{\lambda}))\tilde{\gamma}_j.$$

ここで，$\hat{\gamma}_0(\hat{\lambda})$ は，$\hat{\lambda} = \arg\min SSR_1(\lambda)$ を真の構造変化点比率とみなして (2.1) 式を推定した場合の残差から構築されており，バンド幅 $m(\hat{\lambda})$ もこの残差を用いて Andrews (1991) の方法で決められるものとする．この場合，特徴的なのは，長期分散の推定の際，誤差項の分散の推定には対立仮説の下での推定残差に基づく $\hat{\gamma}_0(\hat{\lambda})$ を，自己共分散の推定には帰無仮説の下での推定残差に基づく $\tilde{\gamma}_j$ を用いていることである．Kejriwal (2009) はこの長期分散推定量を，ハイブリッド推定量と呼んでいる．前節で確認したように，$\hat{u}_t(\lambda)$ に基づく長期分散推定量を用いるとサイズの歪みが生じ，$\tilde{u}_t$ に基づく推定量だと検出力の非単調問題に直面するので，直観的には，両者を併用すれば二つの問題が緩和されると期待できる．Kejriwal (2009) の sup ワルド検定は長期分散の一致推定量の推定方法を変更しただけなので，帰無仮説の下では Andrews (1993) の sup ワルド検定と同じ極限分布をもつ．

なお，Yang and Vogelsang (2011) は，Kejriwal の検定に対して，次のような解釈を与えている．まず，長期分散のハイブリッド推定量については，

$$\hat{\gamma}_0(\hat{\lambda}) = \frac{1}{T}SSR_1(\hat{\lambda}), \quad \tilde{\gamma}_0 = \frac{1}{T}SSR_0$$

という関係を用いると，

$$\begin{aligned} \hat{\sigma}_h^2 &= \hat{\gamma}_0(\hat{\lambda}) + 2\sum_{j=1}^{T-1} k(j, m(\hat{\lambda}))\tilde{\gamma}_j \\ &= \left[\tilde{\gamma}_0 + 2\sum_{j=1}^{T-1} k(j, m(\hat{\lambda}))\tilde{\gamma}_j\right] - \left[\tilde{\gamma}_0 - \hat{\gamma}_0(\hat{\lambda})\right] \\ &= \tilde{\sigma}^{*2} - \frac{1}{T}(SSR_0 - SSR_1(\hat{\lambda})) \\ &= \tilde{\sigma}^{*2}\left(1 - \frac{1}{T}\sup LM^*\right) \end{aligned}$$

という表現が得られる．ただし，

$$\sup LM^* = \frac{SSR_0 - SSR_1(\hat{\lambda})}{\tilde{\sigma}^{*2}}$$

で，$\tilde{\sigma}^{*2}$ は帰無仮説の下での推定残差に基づく長期分散の推定量であるが，バンド幅は $\hat{u}_t(\hat{\lambda})$ に基づいて決められたものであるため，$\tilde{u}_t$ に基づくバンド幅を用いた長期分散推定量 $\tilde{\sigma}^2$ とは，一般的には異なる．これより，

$$\sup W_T^{kej} = \sup_{\lambda \in \Lambda} \frac{SSR_0 - SSR_1(\lambda)}{\tilde{\sigma}^{*2}\left(1 - \frac{1}{T}\sup LM^*\right)} = \frac{\sup LM^*}{1 - \frac{1}{T}\sup LM^*}$$

となるので，Kejriwal の検定は，supLM 検定を変換したものであると解釈できる．

3.2 self-normalization 法に基づく検定 (Shao and Zhang (2010))

Shao and Zhang (2010) は，Kejriwal (2009) と異なり，長期分散の推定は行わず，検定統計量の分子・分母で長期分散を相殺する方法を考案している．すなわち，2 つの統計量の分布がそれぞれ長期分散に比例するのならば，その比をとれば長期分散には依存しない統計量となる，という原理に基づいた検定を考えている．例えば，構造変化がない場合，仮定 1(i) より，以下の 2 つの分布収束が得られる．

$$\begin{aligned} T(\bar{y} - \mu)^2 &\xrightarrow{d} \sigma^2 W^2(1), \\ \frac{1}{T^2}\sum_{t=1}^{T}\left\{\sum_{s=1}^{t}(y_s - \bar{y})\right\}^2 &\xrightarrow{d} \sigma^2\int_0^1 (W(s) - sW(1))^2 ds. \end{aligned}$$

したがって，両者の比をとれば，

$$\frac{T(\bar{y} - \mu)^2}{\frac{1}{T^2}\sum_{t=1}^{T}\left\{\sum_{s=1}^{t}(y_s - \bar{y})\right\}^2} \xrightarrow{d} \frac{W^2(1)}{\int_0^1 (W(s) - sW(1))^2 ds}$$

となるため，長期分散 σ^2 には依存しなくなる．

実際の構造変化の検定として，Shao and Zhang (2010) では，次のような検定統計量を提案している．

$$G_T = \sup_{\lambda\in[0,1]} \frac{\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{[T\lambda]}(y_t-\bar{y})\right\}^2}{\frac{1}{T^2}\left\{\sum_{t=1}^{[T\lambda]}\left(S_{1,t}-\frac{t}{[T\lambda]}S_{1,[T\lambda]}\right)^2+\sum_{t=[T\lambda]+1}^{T}\left(S_{t,T}-\frac{T-t+1}{T-[T\lambda]}S_{[T\lambda]+1,T}\right)^2\right\}}.$$

ただし，$t_1 \le t_2$ に対して，$S_{t_1,t_2}=\sum_{s=t_1}^{t_2} y_s$ とする．この場合，帰無仮説の下では，

$$G_T \xrightarrow{d} \sup_{\lambda\in[0,1]} \frac{\{W(\lambda)-rW(1)\}^2}{\int_0^{\lambda}\left\{W(s)-\frac{s}{\lambda}W(\lambda)\right\}^2 ds+\int_{\lambda}^{1}\left\{W(1)-W(s)-\frac{1-s}{1-\lambda}[W(1)-W(\lambda)]\right\}^2 ds}$$

となる．

3.3 fixed-b 漸近近似法に基づく修正 sup ワルド検定 (Sayginsoy and Vogelsang (2011))

カーネル法で長期分散を推定する場合，一致推定量を得るためには通常，バンド幅 m の発散する速度は標本の大きさ T より遅くとらなければならない．しかしながら，Sayginsoy and Vogelsang (2011) では，m を T に比例した速さで発散させるよう設定している．このように，バンド幅 m を $m=bT$ ($b\in(0,1]$) とした場合の漸近理論は「fixed-b asymptotics」と呼ばれ，Kiefer and Vogelsang (2005) により，この場合の長期分散のカーネル推定量は一致性を持たず，長期分散に比例した分布に収束することが証明されている．したがって，長期分散の処理方法については，self-normalization と似た原理に基づいていることが分かる．

Sayginsoy and Vogelsang (2011) ではさらに，確率項 u_t が単位根を持つ場合にも頑健な検定統計量として，次のようなものを提案している．

$$\sup W_T^{SV} = \sup_{\lambda\in\Lambda} \frac{SSR_0 - SSR_1(\lambda)}{\hat{\sigma}_b^2}\cdot \exp\left(-c\cdot UR\right).$$

ただし，$\hat{\sigma}_b^2$ は，(2.3) 式でバンド幅を $m=bT$ とした場合の長期分散の推定量であり，UR は Park and Choi (1988) および Park (1990) で提案された単位根検定統計量である．また，バンド幅の比例定数 b および定数 c はいわゆるチューニング・パラメータとなっており，UR の値に依存して決められるものとされている．詳しくは，ワーキング・ペーパー版の Sayginsoy and Vogelsang (2008) を参照されたい．検定統計量のうち，最初の項はバンド幅を $m=bT$ とした sup ワルド検定統計量であり，第 2 項が単位根の場合を考慮した追加的な項である．確率項 u_t が定常であれば，$\exp(-c\cdot UR)\xrightarrow{p} 1$ となるので，元のデータが定常である限り，この項は漸近的には影響を及ぼさない．

前述したように，分子と同様，分母である $\hat{\sigma}_b^2$ も帰無仮説の下では σ^2 に比例した分布に収束するので，$\sup W_T^{SV}$ は σ^2 に依存しない分布に収束することになる．ただし，分母の分

布収束先は，使用するカーネルに依存することが知られている．Sayginsoy and Vogelsang (2011) では，ダニエル・カーネルの利用を推奨している．

3.4 sup ワルド検定 (Kejriwal) の再検討

Kejriwal (2009) では，長期分散の推定に関してはハイブリッド推定量を使用したが，その際，分散推定量 $\hat{\gamma}_0(\hat{\lambda})$ およびバンド幅は，$\hat{\lambda} = \arg\min SSR_1(\lambda)$ を真の構造変化点比率と見なして (2.1) 式を推定した場合の残差から構築されていた．ここでは，そのような制約は課さずに，最大化の引数全体を λ の関数ととらえ，次のような修正ワルド検定統計量を提案することにする．[2)]

$$\sup W_T^{kej*} = \sup_{\lambda\in\Lambda} \frac{SSR_0 - SSR_1(\lambda)}{\hat{\gamma}_0(\lambda) + 2\sum_{j=1}^{T-1} k(j, m(\lambda))\tilde{\gamma}_j}.$$

ただし，$m(\lambda)$ は $T_B = \lambda T$ とした場合の (2.1) 式の残差を元に，Andrews (1991) の方法により求められたバンド幅である．長期分散の推定量は，構造変化点比率 λ の扱いを除いて，Kejriwal (2009) と同一であるので，分散推定量 $\hat{\gamma}_0(\lambda)$ は対立仮説の下での推定残差に，自己共分散推定量 $\tilde{\gamma}_j$ は帰無仮説の下での推定残差に基づいている．検定統計量が未知の構造変化点比率 λ に依存する以上，長期分散についても分子と同様，λ に依存させて検定統計量全体を λ について最大化させることは，自然な拡張であるといえる．また，$\sup W_T^{kej*}$ の分母において $\lambda = \hat{\lambda}$ としたものが $\sup W_T^{kej}$ となるので，

$$\sup W_T^{kej} \leq \sup W_T^{kej*}$$

という関係が読み取れる．すなわち，検出力の面では，修正版の方がよりパフォーマンスが良くなることが期待できる．

以上の修正は，一見，単純な拡張であるが，その有限標本での効果の大きさは，次節で確認することにする．

4. 各検定の有限標本特性

本節では，前節で紹介した先行研究による 3 つの検定と，本稿で修正を提案した 1 つの検定の有限標本特性を，モンテ・カルロ実験で考察する．モンテ・カルロ実験で用いるデー

---

2) 査読者より，分子と分母の λ を別のパラメータと見なして，2 次元の方向で検定統計量を最大化させれば検出力が向上するのではないかという指摘があった．そこで，実際にシミュレーション実験を行ったところ，検出力は最大で数%ほどだが，向上することが分かった．ただし，この場合は，分母 (長期分散推定量) を λ について最小化することになるため，構造変化点比率 λ を長期分散に基づいて推定していると解釈できる．しかしながら，構造変化点については，先行研究において，(残差平方和を最小にするという意味で) いわゆる最小 2 乗法か疑似最尤法に基づく推定量について分析がなされており，長期分散に基づく推定の妥当性や解釈は，現状では良く分かっていない．このような理由から，分子と分母の λ を 2 つの引数と見なして検定統計量を作成するという方策は，今回はとらないことにし，長期分散に基づく構造変化点の推定については，今後の検討課題としたい．

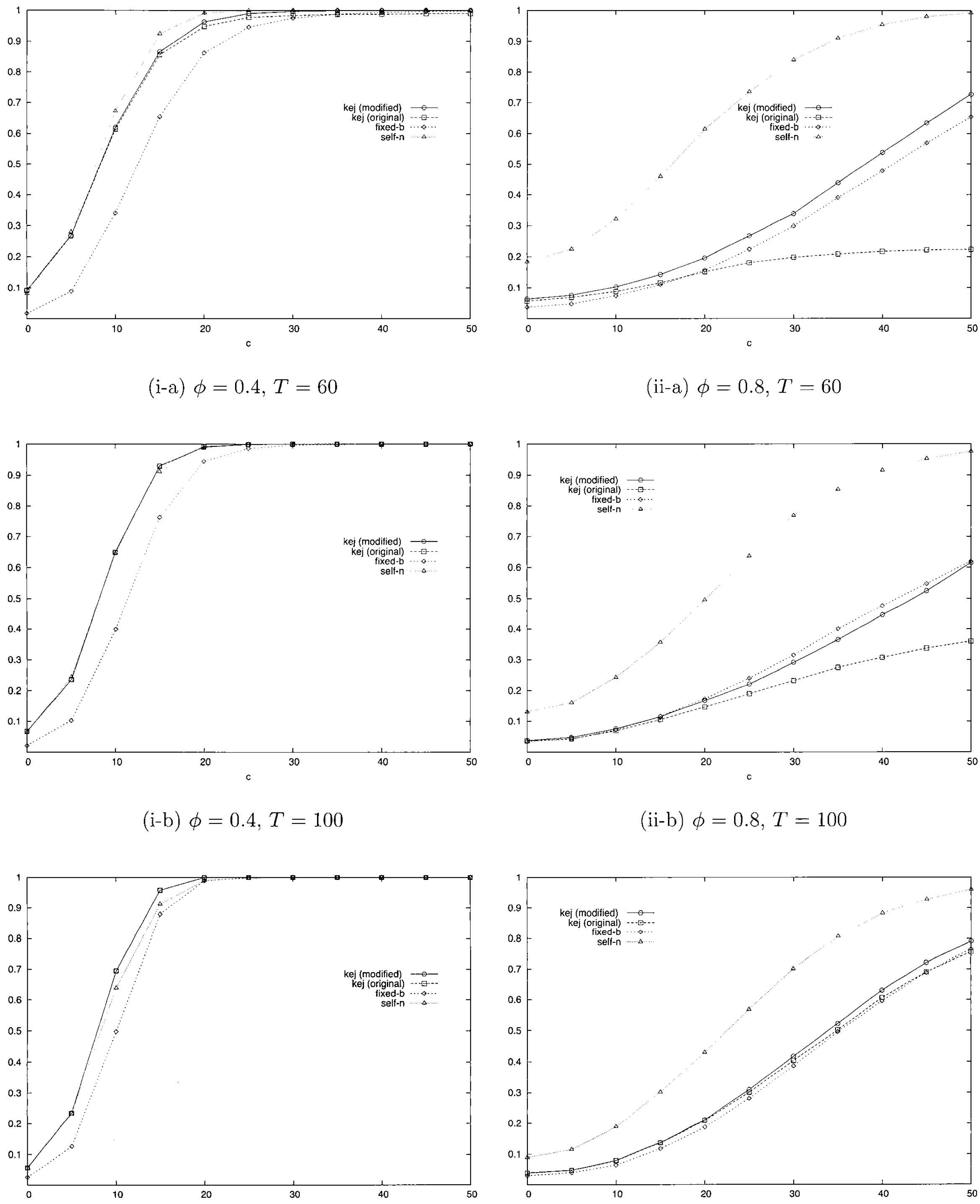
タ生成過程は，以下のとおりである．

$$y_t = (c/\sqrt{T}) \times DU_t(T/2) + u_t, \quad u_t = \phi u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.} N(0,1).$$

ただし，$\phi = 0.4$ または 0.8，$u_1 = \varepsilon_1/(1-\phi^2)^{1/2}$ である．標本の大きさ T は 60，100 または 200 で有意水準は 5%と設定する．Kejriwal (2009) のオリジナルおよび修正版の sup ワルド検定統計量については，QS カーネル及び Andrews (1991) のバンド幅の設定方法を用いて，検定統計量を算出してる．

図 3 は，各検定のサイズおよび検出力を描いたグラフである．系列相関の強さが中程度である $\phi = 0.4$ の場合には，fixed-b 漸近近似法に基づく修正ワルド検定のサイズは 5%よりも低くなりがちであり，結果として，検出力が他の検定より大幅に劣る．他の 3 つの検定に関してはあまり大きな差はないものの，標本の大きさが小さな場合は，self-normalization 法に基づく検定の検出力が若干高く，標本の大きさが大きくなるほど，Kejriwal のオリジナル及び修正版 sup ワルド検定の方が，パフォーマンスは良くなる．一方，系列相関が強い $\phi = 0.8$ の場合には，self-normalization 法に基づく検定のサイズは 5%を大幅に上回っており，検出力は単調であるものの，第 2 節で議論した Andrews (1991) の sup ワルド検定と同じ問題を抱えている事が分かる．また，Kejriwal (2009) で提案された sup ワルド検定は，検出力はかろうじて単調に見えるものの，標本の大きさが小さな場合は，大きな構造変化があっても，検出力がなかなか上昇しないのに対し，本稿で提案した修正を施せば，パフォーマンスが大幅に改善することが分かる．ただし，標本の大きさが大きな場合，オリジナル版と修正版の差は小さくなる．また，修正版 Kejriwal の sup ワルド検定と fixed-b 漸近近似法に基づく修正 sup ワルド検定の差は，$T = 60$ の場合には前者の方が良いように見えるが，サイズの差も考慮すべきであり，総じて，$\phi = 0.8$ の場合には両者の差は小さいことがうかがえる．

Kejriwal の sup ワルド検定の検出力がオリジナル版と修正版で大きく異なる要因として，長期分散の推定精度の違いが挙げられる．どちらも長期分散は過大に推定されがちなのだが，オリジナル版の方がより過大に推定される．例えば，$\phi = 0.8$ の場合，真の長期分散は 25 であるが，$c = 40$ の場合のオリジナル版／修正版の長期分散推定値の平均はそれぞれ，67.4 ／ 46.9 ($T = 60$)，59.6 ／ 52.1 ($T = 100$)，49.5 ／ 47.9 である．一方，第 2 節で論じた本来の sup ワルド検定統計量 $\sup W_T$ では帰無仮説の下でのサイズの歪みが大きかったが，Kejriwal の修正版 sup ワルド検定ではそのような過剰な棄却は生じていない．理由はオリジナルの検定の場合と同じで，前者では長期分散が過少推定されがちであるが，後者ではバイアスが小さくなることが実験により確認されている．結果として，修正版の sup ワルド検定は，帰無仮説の下での Kejriwal オリジナル版の良い点を継承しつつ，対立仮説の下でのオリジナル版の欠点を補ったものとなっている．



(i-a) $\phi = 0.4$, $T = 60$　(ii-a) $\phi = 0.8$, $T = 60$

(i-b) $\phi = 0.4$, $T = 100$　(ii-b) $\phi = 0.8$, $T = 100$

(i-c) $\phi = 0.4$, $T = 200$　(ii-c) $\phi = 0.8$, $T = 200$

図 3　各検定のサイズと検出力.

総じて評価すると，実際の経済分析では系列相関の強さは未知であり，系列相関の強弱に依存して有限標本特性が大きく左右されるような検定はあまり望ましくないということを考慮すると，今回提案した修正版 Kejriwal の sup ワルド検定の総合的なパフォーマンス

は，既存の検定に匹敵するか，それ以上のものを持っているといえる．

5. 今後の展望

確率項に系列相関がある場合のレベル・シフトの検定では，サイズの歪みや検出力の非単調性などの問題があることが知られている．そこで本稿ではまず，これまで先行研究で提案されてきた対処策を整理したうえで，その対処策の一つである Kejriwal (2009) の方法を修正した，修正 sup ワルド検定を提案した．モンテ・カルロ実験により，修正 sup ワルド検定の有限標本特性は，サイズと検出力の両面で，先行研究で提案されてきた対処策と同程度か，それ以上のものであることが確認された．

しかしながら，これで問題がすべて解決されたわけではない．現段階では，先行研究と比較して有限標本特性が改善されただけであり，理論的に可能な最大のパフォーマンスからの乖離は少なからず存在すると思われる．すなわち，検出力を向上させる余地はまだあると同時に，検定のサイズも，設定された有意水準へより近づけるべく制御されなければならない．したがって，モデルは非常にシンプルではあるものの，レベル・シフトの検定の有限標本特性の向上は，今後も追及されるべき研究テーマである．一つの方向性としては，確率項の系列相関の構造をモデル化することにより，問題の根幹を理論的に解明し，それを修正していくということが考えられる．この場合，有限標本におけるバイアス修正というのが一つの可能性であり，それがどのような系列相関の構造に適応可能であるのか，その限界を探ることも興味深い．

また，本稿ではレベル・シフトの検定のみに焦点を当てたが，一般の回帰係数の構造変化の検定はより複雑なものとなるため，今回の手法が有効であるかは明らかではない．したがって，今後の方向性の一つとして，レベル・シフトの検定だけではなく，線形・非線形の一般モデルにおいて，サイズが安定しており，かつ，単調な検出力を持つ構造変化の検定に関する研究が展開されていくだろう．その際，定常な誤差項ばかりではなく，系列相関が非常に強い，長期記憶過程や単位根過程に対しても頑健な検定の構築が期待される．構造変化の検定の歴史は長いが，以上のようにまだ未解決な問題が多数残っており，今後も研究が進められていくだろう．

参考文献

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica*, **59**, 817–858.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point, *Econometrica*, **61**, 821–856 (Erratum, **71**, 395–397).

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica*, **62**, 1383–1414.

Andrews, D. W. K., Lee, I. and Ploberger, W. (1996). Optimal changepoint tests for normal linear regression, *J. Econom.*, **70**, 9–38.

Aue, A. and Horváth, L. (2013). Structural breaks in time series, *J. Time Ser. Anal.*, **34**, 1–16.

Bai, J. (1997). Estimating multiple breaks one at a time, *Econ. Theory*, **13**, 315–352.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes, *Econometrica*, **66**, 47–78.

Brown, B., Durbin, J. and Evans, J. (1975). Techniques for testing the constancy of regression relationships over time, *J. R. Stat. Soc. Ser. B*, **37**, 149–172.

Crainiceanu, C. M. and Vogelsang, T. J. (2007). Nonmonotonic power for tests of a mean shift in a time series, *J. Stat. Comput. Simul.*, **77**, 457–476.

Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*, Wiley, New York.

Deng, A. and Perron, P. (2008). A non-local perspective on the power properties of the CUSUM and CUSUM of squares tests for structural change, *J. Econom.*, **142**, 212–240.

Elliott, G. and Müller, U. K. (2006). Efficient tests for general persistent time variation in regression coefficients, *Rev. Econ. Stud.*, **73**, 907–940.

Hansen, B. E. (1997). Approximate asymptotic P values for structural-change tests, *J. Bus. Econ. Stat.*, **15**, 60–67.

Juhl, T. and Xiao, Z. (2009). Tests for changing mean with monotonic power, *J. Econom.*, **148**, 14–24.

Kejriwal, M. (2009). Tests for a mean shift with good size and monotonic power, *Econ. Lett.*, **102**, 78–82.

Kiefer, N. M. and Vogelsang, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests, *Econ. Theory*, **21**, 1130–1164.

Kurozumi, E. and Tuvaandorj, P. (2011). Model selection criteria in multivariate models with multiple structural changes, *J. Econom.*, **164**, 218–238.

Park, J. Y. (1990). Testing for unit roots and cointegration by variable addition, in *Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots*, (eds. T. Fomby and R. Rhodes), pp. 107–134, Jai Press.

Park, J. Y. and Choi, B. (1988). A new approach to testing for a unit root, CAE Working paper 88-23, Cornell University.

Perron, P. (2006). Dealing with structural breaks, in *Palgrave Handbook of Econometrics, Vol. 1: Econ. Theory*, (eds. T. C. Mills and K. Patterson), Palgrave Macmillan, New York.

Perron, P. and Yamamoto, Y. (2014). On the usefulness or lack thereof of optimality criteria for structural change tests, *Econom. Rev.*, forthcoming.

Qu, Z. and Perron, P. (2007). Estimating and testing structural changes in multivariate regressions, *Econometrica*, **75**, 459–502.

Sayginsoy, Ö. and Vogelsang, T. J. (2008). Testing for a shift in trend at an unknown date: A fixed-B analysis of heteroskedasticity autocorrelation robust OLS based tests, Working paper, Department of Economics, Michigan State University.

Sayginsoy, Ö. and Vogelsang, T. J. (2011). Testing for a shift in trend at unknown date: A fixed-B analysis of heteroskedasticity autocorrelation robust OLS-based tests, *Econ. Theory*, **27**, 992–1025.

Shao, X. and Zhang, X. (2010). Testing for change points in time series, *J. Am. Stat. Assoc.*, **105**, 1228–1240.

Sowell, B. (1996). Optimal tests for parameter instability in the generalized method of moments framework, *Econometrica*, **64**, 1085–1107.

Vogelsang, T. J. (1999). Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series, *J. Econom.*, **88**, 283–299.

Yang, J. and Vogelsang, T. J. (2011). Fixed-B analysis of LM-type tests for a shift in mean, *Econom. J.*, **14**, 438–456.