

非定常経済時系列におけるトレンドの統計的問題*

田中勝人

一橋大学

いわゆる「単位根問題」と「共和分分析」は、時系列的方法に基礎を置く計量経済学の理論および実証研究のホットなトピックとして 1980 年代から存続してきた。そのインパクトは、それ以前の計量経済学の歴史において最も華々しい成果をあげた同時方程式モデルの推定問題を凌ぐほどに強く、計量経済学のテキストを書き換えねばならないほどの影響をもたらした。

この 2 つのトピックは、経済時系列の非定常性の源泉であるトレンドと深く関わっており、複数の時系列を扱う場合には不可避の対象である。まず、「単位根問題」は、トレンドが確率的なのか確定的なのかを統計的に判断する問題であり、トレンドが確率的であるという判断が下された時系列が複数個ある場合には、次に「共和分分析」に進み、複数の確率的トレンドが多重共線的な関係にあるかどうかを調べることになる。

本稿では、両者のトピックのキーワードであるトレンドに焦点をあて、単位根検定が前提としているトレンドに関する二分法、すなわち、確定的トレンドと確率的トレンドの間の識別可能性に関して議論を行う。ここでの問題意識は、「トレンドが確定的か確率的かということ、統計的に検証することは可能であろうか？」という疑問であり、結論としては、そのことが不可能となるような状況があるということを指摘したい。

* 本稿は、2000 年 9 月 16, 17 日に大阪府立大学で開かれた日本経済学会大会における招待講演の論文 (Tanaka [2001]) の内容をまとめたものである。詳細については、原論文を参照されたい。

1. はじめに

冒頭に述べたように、単位根問題や共和分分析は時系列的方法に基づく計量経済学に非常に強いインパクトを与えた。その理由としては、以下の諸点をあげることができる。

- (i) 伝統的な計量経済学の方法では、原系列が非定常と判断されるときは階差変換を施すことにより、常に定常性の枠内で処理していた。しかし、現実の経済現象との対応において、原系列そのものの変動を分析すべきである、との認識が生まれた。そうすることにより、階差変換による情報の損失や、そこから派生する階差後の系列と原系列との間に起こりうる推測の不一致性、推測における非効率性などを回避することが可能となった。
- (ii) 非定常性をランダム・ウォーク、確率的トレンド、あるいはマルチンゲールと関連させることにより、(1) 従来の景気循環論とは異なる経済変動の均衡の不安定性、(2) 確率的トレンドに支配される予測不可能性、(3) マルチンゲールがもたらす市場の効率性や合理的期待、などの諸仮説が単位根検定により検証できるようになった。これらの検定の結果がもたらすインプリケーションは経済政策にとっても重要である。しかし、検定統計量は正規分布やカイ 2 乗分布などの標準的な分布に従わず、そのための理論を展開する必要が生じた。
- (iii) 説明変数と誤差項が相関をもつ場合には LSE (最小 2 乗推定量) は一致性を失う、という事実は定常的な説明変数では真であるが、非定常な世界では適用されない。したがって、いわゆる同時方程式バイアスも消滅し、同時方程式モデルのために考案された各種推定量の優劣の正当性も根拠を失う。
- (iv) 互いに無相関な非定常変数間で回帰モデルや同時方程式モデルなどを作ると、有意な t 値や決定係数が観測されるという、見せかけの回帰 (spurious regression) が起きる。この事実は標本サイズが大きくなる漸近的な状況でも成立し、そのようなことが起きない定常的な場合と対照的である。
- (v) 定常的な諸変数の一次結合は常に定常的である。他方、非定常的な諸変数の一次結合は必ずしも非定常的ではなく、定常になる場合もある。それが共和分現象であり、確率的な多重共線関係を生み出す。原系列を分析する

場合には、共和分関係の有無を検定する手続きが必要となった。

上記のような諸事実から学ぶ教訓は、「経済時系列は原系列そのもの（レベル）を分析すべきである。ただし、その場合には、非定常時系列を扱うことになるから、非定常性がもたらす特殊性に深い注意を払うべきである。」ということになるだろう。

本稿では、経済時系列の非定常性がもたらす特殊性、あるいは非正則性に関連する事柄として、トレンドの問題を取り上げる。主要な関心事は、確率的なトレンドと確定的なトレンドの関係であり、その違いを判断することは統計的に可能であろうか、という懐疑なり疑問である。第2節では、Phillips [1998, 1999] の議論に基づいて、DGP (Data Generating Process: データ生成過程) が $I(1)$ 系列、すなわち確率的トレンドを含む場合に、誤って確定的トレンドのモデルが受容されてしまう状況を考察する。第3節では、単位根検定との関連を考察するために、Phillips の議論を拡張して、DGP が near- $I(1)$ 系列の場合を考え、その上で単位根検定の検出力を調べる。そして、標本サイズが大きくなっても、検出力が増えないという意味で単位根検定が無効となるような状況を考察する。第4節では、本稿で得られた結果から得られる示唆を述べ、結論を与える。

2. 確率的トレンドと確定的トレンド：純粋な $I(1)$ 系列の場合

本節では、DGP が純粋な $I(1)$ 系列の場合

$$(1) \quad y_t = y_{t-1} + u_t, \quad y_0 = 0, \quad (t = 1, \dots, T)$$

を考えよう。ここで、誤差項の系列 $\{u_t\}$ は、定常かつ短期記憶の線形過程に従うものとする。すなわち、

$$u_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} |\alpha_j| < \infty, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2).$$

もちろん、DGP は未知であるから、そのすべてを知ることは永遠に不可能である。それゆえに、神ならぬ人間はさまざまなモデルを想定することにより DGP に接近するわけであるが、その場合にモデルのよしあしは統計的に判断される。しかし、 $I(1)$ 系列の特殊性は、伝統的に行われている統計的判断を誤

りに導くものとなる．その典型的な例が，いわゆる「見せかけの回帰」と呼ばれる現象であり，次の事実が知られている．

・説明変数が $I(1)$ の場合の見せかけの回帰

$\{x_t\}$ を $\{y_t\}$ とは独立な $I(1)$ 系列とするととき，回帰式

$$(2) \quad y_t = \hat{\alpha} + \hat{\beta}x_t + \hat{u}_t, \quad (t = 1, \dots, T)$$

は，次の性質をもつ．

- (a) $\hat{\alpha} = O_p(\sqrt{T})$, $\hat{\beta} = O_p(1)$
- (b) 係数の t 統計量 $= O_p(\sqrt{T})$
- (c) 決定係数 $R^2 = O_p(1)$
- (d) ダービン＝ワトソン統計量 $DW = O_p(1/T)$

・説明変数が線形トレンドの場合の見せかけの回帰

$$(3) \quad y_t = \hat{\gamma} + \hat{\delta}t + \hat{v}_t, \quad (t = 1, \dots, T)$$

は，次の性質をもつ．

- (a) $\hat{\gamma} = O_p(\sqrt{T})$, $\hat{\delta} = O_p(1/\sqrt{T})$
- (b) 係数の t 統計量 $= O_p(\sqrt{T})$
- (c) $R^2 = O_p(1)$
- (d) $DW = O_p(1/T)$

・説明変数が三角関数の場合の見せかけの回帰

$$(4) \quad y_t = \hat{\mu} + \hat{\kappa} \sin \lambda t + \hat{w}_t, \quad \lambda = \frac{\pi t}{2T}, \quad (t = 1, \dots, T)$$

は，次の性質をもつ．

- (a) $\hat{\mu} = O_p(\sqrt{T})$, $\hat{\kappa} = O_p(\sqrt{T})$
- (b) 係数の t 統計量 $= O_p(\sqrt{T})$
- (c) $R^2 = O_p(1)$
- (d) $DW = O_p(1/T)$

これら 3 つの回帰は、 $\{y_t\}$ が $I(1)$ 系列である限り成立しないものである。しかし、 t -統計量や決定係数が有意性を示唆して、もっともらしい回帰をもたらす。特に、説明変数が三角関数であっても有意な回帰を与えることは驚くべき事実である。ただし、DW 統計量のオーダーが $1/T$ であることから、残差に非常に強い正の系列相関があることを示唆し、回帰が満足の行くものではないというシグナルを発している。このことは、回帰の残差系列が単位根を含む $I(1)$ 系列であることを意味しており、あてはめた回帰式が spurious であることの証左にほかならない。

ところで、上で考えた回帰に関する漸近的な性質は、標本サイズ T が大きくなる場合のものであり、その理論を T -asymptotics と呼ぶことができる。他方、上で考えたような回帰において、標本サイズが十分大きい状況で、説明変数の個数 K を増やして行く場合の漸近的な性質を考えることも可能である。この場合の理論を以下では K -asymptotics と呼ぶことにする。

K -asymptotics を考えるために、(1) の DGP に対して、次の回帰式

$$(5) \quad y_t = \sum_{k=1}^K \tilde{b}_k \phi_k \left(\frac{t}{T} \right) + \tilde{u}_t = \tilde{b}(K)' \phi(K, t/T) + \tilde{u}_t, \quad (t = 1, \dots, T)$$

を考えよう。ここで、

$$\tilde{b}(K) = (b_1, \dots, b_K)', \quad \phi(K, t/T) = (\phi_1(t/T), \dots, \phi_K(t/T))'$$

である。 $\phi(K, t/T)$ は確定的トレンドのベクトルであり、その第 k 成分は

$$\phi_k(r) = \sqrt{2} \sin[(k - 1/2)\pi r] \quad (k = 1, \dots, K)$$

で定義される。説明変数は三角関数のトレンドであり、(4) で考えた spurious regression の拡張となっている。このモデルは、Phillips [1998] で最初に論じられたものである。

もちろん、確定的トレンドとしては、(3) のような線形トレンドや多項式トレンドを説明変数とする方が自然であるが、その場合には数学的な取扱いが複雑になるので、以下では三角関数を考える。なお、ここで考えている三角関数が特殊な形をしている理由は次の通りである。まず、(1) の DGP に対して、部分和過程

$$X_T(r) = \frac{1}{\sqrt{T}\sigma_L} y_{[Tr]} = \frac{1}{\sqrt{T}\sigma_L} \sum_{t=1}^{[Tr]} u_t$$

は、 $T \rightarrow \infty$ のとき、標準ブラウン運動 $\{W(r)\}$ に弱収束する．ここで、 σ_L は、定常な誤差項の系列 $\{u_t\}$ の長期分散

$$\sigma_L^2 = \lim_{T \rightarrow \infty} V \left(\frac{1}{T} \sum_{t=1}^T u_t \right) = \sigma^2 \left(\sum_{j=0}^{\infty} \alpha_j \right)^2$$

の正の平方根である．なお、長期分散に対応して、短期分散

$$\sigma_S^2 = V(u_t) = \sigma^2 \sum_{j=0}^{\infty} \alpha_j^2$$

を定義しておく．

上の結果は、I(1) 系列の分布がブラウン運動の分布で近似できることを述べているが、系列のパスに関しても、 $E(|\varepsilon_t|^p) < \infty$ 、ただし、 $p > 2$ ならば、

$$\sup_{0 \leq r \leq 1} T^\delta \left| \frac{1}{\sqrt{T} \sigma_L} y_{[Tr]} - W(r) \right| \rightarrow 0, \quad \left(0 < \delta = \frac{1}{2} - \frac{1}{p} < \frac{1}{2} \right)$$

が確率 1 で成り立つ (Csörgő and Horváth [1993])．そして、標準ブラウン運動 $\{W(r)\}$ は確率 1 で次の展開表現をもつことが知られている (Loève [1978], Chan and Wei [1988])．

$$W(r) = \sum_{k=1}^{\infty} \phi_k(r) \frac{\xi_k}{(k-1/2)\pi}, \quad \{\xi_k\} \sim \text{NID}(0,1).$$

すなわち、確率過程であるブラウン運動は、確定的な三角関数トレンド $\phi_k(r)$ の一次結合で表現可能（ただし、係数は確率変数）である．このことが、(5) で $\phi_k(r)$ を使うことの理由であり、数学的な演算を容易にする．なお、 $\phi_k(r)$, $(k = 1, 2, \dots)$ は区間 $[0, 1]$ 上で正規直交である．

さて、(5) のモデルを (4) と比較した場合、説明変数が任意個に拡張されているが、依然として spurious である点は変わらない．そのシグナルがどの程度発せられているかをみるために、 T -asymptotics を考えると次の結果が得られる．

・ (5) のモデルにおける T -asymptotics

(a) $\tilde{b}(K) = O_p(\sqrt{T}) \sim$ 多変量正規分布

(b) $\sum_{t=1}^T \tilde{u}_t^2 = O_p(T^2)$

(c) 係数の t 統計量 $= O_p(\sqrt{T})$

$$(d) R^2 = O_p(1)$$

$$(e) DW = O_p(1/T)$$

上の結果はより精密化することができる．例えば，(e) に関しては

$$T \times DW \implies \frac{\sigma_S^2}{\sigma_L^2} \times \frac{1}{\int_0^1 W_{\phi K}^2(r) dr}$$

が成り立つ．ここで， $\implies$ は弱収束を表す．他方，

$$W_{\phi K}(r) = W(r) - \int_0^1 \phi(K, s)' W(s) ds \phi(K, r)$$

であり，これはブラウン運動 $\{W(r)\}$ を $\phi(K, r)$ で張られる空間に射影した場合の残差過程と解釈することができる．実際， $\{W_{\phi K}(r)\}$ は (5) のモデルの残差 $\tilde{u}_t$ の次の意味での極限となっている．

$$\frac{1}{\sqrt{T} \sigma_L} \tilde{u}_{[Tr]} \implies W_{\phi K}(r).$$

結局のところ， K が固定されている限り，項数 K にかかわらず，(5) は (4) のモデルと本質的に同じ性質をもっており，残差は依然として単位根を含む系列である．

ところが， T を大きくしたあとで，さらに説明変数の項数 K を大きくして行く場合の K -asymptotics においては，次のことが成り立つ．

・ (5) のモデルにおける K -asymptotics

$$(a) \tilde{b}(K)/\sqrt{T} = O_p(1) \sim \text{多変量正規分布}$$

$$(b) \sum_{t=1}^T \tilde{u}_t^2 / T^2 = O_p(1/K)$$

$$(c) \text{係数の } t \text{ 統計量} / \sqrt{T} = O_p(\sqrt{K})$$

$$(d) R^2 \text{ は } 1 \text{ に確率収束}$$

$$(e) T \times DW = O_p(K)$$

K -asymptotics においては，係数の t -統計量は発散し，係数の有意性を示唆する．また，決定係数も 1 に近づく．このことは，残差が単位根系列を含まなくなることを意味し，回帰のもっともらしさを与える．さらに， T -asymptotics

においては回帰の無意味さのシグナルであった DW 統計量までもが回帰式の正当性を主張し始めるのである．なお，(e) の結果をより精密化すると，

$$\frac{T}{K} \times DW \longrightarrow \frac{\sigma_S^2}{\sigma_L^2} \times \pi^2$$

が確率収束の意味で得られる．

表 1 は，(1) の DGP において誤差項 $\{u_t\}$ が i.i.d. の場合に，(5) のモデルから得られる R^2 と DW の平均と分散を繰り返し数 1,000 回のシミュレーションにより求めたものである． R^2 の分布は T に依存せずに $K \rightarrow \infty$ のとき 1 に近づくこと， DW の平均と分散は T に反比例すること，また， K とともに大きくなることなどを読みとることができる．なお，今の場合， $\sigma_L^2 = \sigma_S^2$ であるから， DW の値は，上の精密化の結果から，ほぼ $\pi^2 K/T$ である．この理論値も表に示されているが，近似の程度は良好である．

表 1

以上のことから，I(1) 系列で表される確率的トレンドは，三角関数による確定的トレンドで完全に表現されるという結果が得られ，統計的にも識別不可能な状況が生まれることになる．

ところで， K -asymptotics は，説明変数の個数が標本サイズに近づく状況の性質である．この理由から，上の結果は I(1) 系列に限らずに成立するのではないか．例えば，(1) の DGP の代わりに定常な時系列などを想定しても同様の結果をもたらすのではないかと考えるかもしれない．この点については，シミュレーション結果をグラフ表示してみよう．図 1 は，(1) の DGP (ただし，誤差項は i.i.d.) に対して (5) のモデルから得られる R^2 の分布 (標本サイズは 400 で 1,000 回の繰り返し．以下同様) を $K = 1$ と $K = 20$ の場合に図示したものである．これに対して，図 2 は，DGP が定常な AR(1) (係数値は 0.6) の場合である．他方，図 3，図 4 は， DW の分布を，それぞれ，I(1) と定常な AR(1) の場合に図示したものである．

図 1 図 2 図 3 図 4

ここでは，定常な AR(1) との比較だけを示しておいたが，i.i.d. 系列，ARCH 系列，フラクショナルな ARIMA 系列などと比較しても同様の結果となる．

すなわち、(5) のような確定的トレンドをあてはめた場合、上で述べた K -asymptotics は $I(1)$ 系列特有のものである。

(5) のモデルは (1) の DGP をネストしないが、今度はネストするモデルとして、

$$(6) \quad y_t = \hat{\rho}y_{t-1} + \sum_{k=1}^K \hat{b}_k \phi_k\left(\frac{t}{T}\right) + \hat{u}_t, \quad (t = 1, \dots, T)$$

を考えよう。このモデルは、確定的トレンド $\phi_k(r)$ が特殊な形をしていることを除けば、単位根検定に使われるモデルと同じ形をしている。したがって、単位根検定統計量として、係数推定量に基づく Z_ρ と、 t -統計量 Z_t を考えることができる。ただし、これらは残差系列 $\{\hat{u}_t\}$ の相関をノンパラメトリックに補正したものである。

これらの単位根検定統計量に対して、

$$(7) \quad Z_\rho \implies \frac{\int_0^1 W_{\phi K}(r) dW(r)}{\int_0^1 W_{\phi K}^2(r) dr}$$

$$(8) \quad Z_t \implies \frac{\int_0^1 W_{\phi K}(r) dW(r)}{\left(\int_0^1 W_{\phi K}^2(r) dr\right)^{1/2}}$$

が成立する (Phillips [1999])。また、次の T -asymptotics が成り立つ。

・ (6) のモデルにおける T -asymptotics

- (a) $Z_\rho = O_p(1)$, $Z_t = O_p(1)$,
- (b) $\hat{b}(K) = O_p(1/\sqrt{T}) \sim$ 非正規分布
- (c) $\hat{b}(K)$ の t -統計量 $= O_p(1)$
- (d) $\sum_{t=1}^T \hat{u}_t^2 = O_p(T)$

ネストされないモデル (5) の場合には、係数推定量は $\sqrt{T}$ のオーダーで正規性に従っていたが、今の場合には非正規で、 $1/\sqrt{T}$ という微小のオーダーである。しかし、その t -統計量は定数オーダーとなり、有意性を示す可能性がある。残差は、単位根を含まない定常的な系列となる。したがって、 DW 統計量も通常の意味合いで解釈できることになる。結果的に、(6) のモデルは残差

に系列相関の可能性があるので、三角関数のトレンド項も有意となり、説明力のあるモデルとなりうる。

他方、 K -asymptotics については、次のことが成り立つ。

・ (6) のモデルにおける K -asymptotics

- (a) $Z_\rho \rightarrow N\left(-\frac{\pi^2 K}{2}, \frac{\pi^4 K}{6}\right)$
- (b) $Z_t \rightarrow N\left(-\frac{\pi\sqrt{K}}{2}, \frac{\pi^2}{24}\right)$
- (c) $\hat{b}_k$ の t -統計量 $= O_p(\sqrt{K})$

K -asymptotics においては、単位根検定統計量 Z_ρ と Z_t は $-\infty$ に発散する。ただし、発散の仕方は両者で異なる。実際、 Z_ρ は平均も分散も K の大きさに発散するが、 Z_t は分散一定で、平均だけが $\sqrt{K}$ のオーダーで発散する。他方、両者とも適当に基準化すれば正規性に従う。また、(c) の結果から、三角関数のトレンドは本来は不要であるが、統計的には有意な結果となる。なお、(a)、(b) の結果は、(7)、(8) に対する K -asymptotics であるが、これらは Phillips [1999] において得られた次の事実から導出される。

$$\int_0^1 W_{\phi K}(r) dW(r) + \frac{1}{2} = O_p\left(\frac{\sqrt{\log K}}{K}\right),$$

$$\int_0^1 W_{\phi K}^2(r) dr \longrightarrow N\left(\frac{1}{\pi^2 K}, \frac{2}{3\pi^4 K^3}\right).$$

3. 確率的トレンドと確定的トレンド：Near-I(1) 系列の場合

前節では、純粋な I(1) 系列を DGP とした場合に、確定的トレンドのみをフィットしたモデル (5) と、ラグも加えてフィットしたモデル (6) を考えた。これは Phillips [1998,1999] において議論されたものであるが、単位根検定の観点からは、この設定により帰無分布は得られるものの対立分布を求めることはできない。したがって、検出力も議論不可能である。

そこで、本節では DGP を次のような near-I(1) 系列に拡張する。

$$(9) \quad y_t = \rho y_{t-1} + u_t, \quad \rho = 1 - (c/T), \quad (t = 1, \dots, T).$$

ここで、誤差項の系列 $\{u_t\}$ は純粋な $I(1)$ の場合と同様に、定常かつ短期記憶の線形過程に従うものとする。すなわち、

$$u_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} |\alpha_j| < \infty, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2).$$

Near- $I(1)$ 系列とは、標本サイズ T が大きくなるとともに純粋な $I(1)$ に $1/T$ のオーダーで近づくような非定常時系列である。Near- $I(1)$ における漸近理論を議論するためには、 $\{y_t\}$ から作られる部分和過程

$$(10) \quad Y_T(r) = \frac{1}{\sqrt{T}\sigma_L} y_{[Tr]} = \frac{1}{\sqrt{T}\sigma_L} \sum_{t=1}^{[Tr]} \rho^{[Tr]-t} u_t$$

に対する次の FCLT (Functional Central Limit Theorem: 汎関数中心極限定理)

$$Y_T(\cdot) \Rightarrow J^c(\cdot)$$

が必要となる (Tanaka[1996])。ここで、 $\{J^c(r)\}$ は Ornstein-Uhlenbeck (O-U) 過程であり、次のように定義される。

$$J^c(r) = e^{-cr} \int_0^r e^{cs} dW(s) \Leftrightarrow dJ^c(r) = -cJ^c(r) dr + dW(r), \quad J^c(0) = 0.$$

そして、O-U 過程は次のような展開表現をもっている (Loève [1978])。

$$(11) \quad J^c(r) = \sum_{n=1}^{\infty} \frac{f_n(r)}{\sqrt{\lambda_n}} \xi_n, \quad \{\xi_n\} \sim \text{NID}(0, 1).$$

ここで、 λ_n は $J^c(r)$ の共分散関数

$$\text{Cov}(J^c(r), J^c(s)) = \frac{e^{-c|r-s|} - e^{-c(r+s)}}{2c},$$

を積分核とする積分方程式の n 番目に小さな固有値であり、 $f_n(r)$ はそれに対応する固有関数である。積分核は対称、連続かつ正値定符号であることから、固有値はすべて正、固有関数の全体は正規直交にとることができる (Courant and Hilbert[1953], Hochstadt[1973])。

Near- $I(1)$ に拡張した場合の困難さは、固有値と固有関数を明示的に求めることが不可能な点にある。実際、 λ_n は、方程式

$$\tan \sqrt{\lambda - c^2} = -\frac{\sqrt{\lambda - c^2}}{c}$$

の n 番目に小さい正値解となる (Tanaka [2001]) . この解は, 不等式

$$(n - 1/2)\pi < \sqrt{\lambda_n - c^2} < n\pi \quad \text{and} \quad (n - 1/2)^2\pi^2 + c^2 < \lambda_n < n^2\pi^2 + c^2$$

をみたすことはわかるが, 明示的に得ることはできない. ただし, コンピュータによる数値計算で求めることは可能である. また, $f_n(r)$ は,

$$f_n(r) = \frac{\sin \mu_n r}{M_n}, \quad \mu_n = \sqrt{\lambda_n - c^2}, \quad M_n = \sqrt{\frac{1}{2} - \frac{\sin 2\mu_n}{4\mu_n}}$$

のように, λ_n に依存した三角関数となり, 明示的な表現は得られないが確定的なトレンドとして使うことができる.

このような状況で, まず, (9) の DGP をネストしない回帰モデル

$$(12) \quad y_t = \sum_{k=1}^K \tilde{b}_k f_k\left(\frac{t}{T}\right) + \tilde{u}_t = \tilde{b}(K)' f(K, t/T) + \tilde{u}_t, \quad (t = 1, \dots, T)$$

を考えよう. 純粋な $I(1)$ 系列の場合と異なる点は, 確定的トレンドとして, $\phi_k(r)$ の代わりに $f_k(r)$ を使っていることである. 上述のように, $f_k(r)$ は明示的に表現できないが, c が与えられれば λ_k が求まるので, 計算可能である. もちろん, $c = 0$ の純粋な $I(1)$ 系列の場合には, $f_k(r)$ は $\phi_k(r)$ に帰着する.

今の場合の T -asymptotics は, DGP が純粋な $I(1)$ 系列の場合と本質的に同一であることがわかる. もちろん, 推定された係数の漸近分布などは DGP に依存するから, その点での違いはあるが. 係数が有意な値をもたらすことや, 回帰後にも残差系列が DGP の特性を受け継いでいる点は同じである. また, DW 統計量が spurious な回帰のシグナルを発していることも同様である. 他方, K -asymptotics についても, 同様のことが成り立つので, ここでは省略する. なお, 残差系列 $\{\tilde{u}_t\}$ に関しては, 次の FCLT が成り立つ.

$$\frac{1}{\sqrt{T}\sigma_L} \tilde{u}_{[Tr]} \Rightarrow J_{fK}^c(r) = J^c(r) - \int_0^1 f(K, s)' J^c(s) ds f(K, r).$$

極限の残差過程 $\{J_{fK}^c(r)\}$ は, $O-U$ 過程 $\{J^c(r)\}$ を確定的トレンドのベクトル $f(K, r)$ により張られる空間に射影したときの残差であり, この残差は $O-U$ 過程の性質を依然として受け継いでいる, というのが言わんとすることである.

次に, DGP をネストするモデル

$$(13) \quad y_t = \hat{\rho} y_{t-1} + \sum_{k=1}^K \hat{b}_k f_k\left(\frac{t}{T}\right) + \hat{u}_t = \hat{\rho} y_{t-1} + \hat{b}(K)' f(K, t/T) + \hat{u}_t,$$

を考えよう．このとき， T -asymptotics に関して次の定理が成立する．ただし，定理の中の $c(K)$ は長さが 1 の K 次元定数列ベクトルであり， $t_{c(K)'\hat{b}(K)}$ は t -統計量である．

定理 1 (9) の near-I(1) を DGP とする時系列 $\{y_t\}$ に対して，(13) のモデルを考えると，次の T -asymptotics が成り立つ．

$$\begin{aligned} Z_\rho &\Rightarrow \frac{\int_0^1 J_{fK}^c(r) dJ^c(r)}{\int_0^1 \{J_{fK}^c(r)\}^2 dr}, \\ Z_t &\Rightarrow \frac{\int_0^1 J_{fK}^c(r) dJ^c(r)}{\left(\int_0^1 \{J_{fK}^c(r)\}^2 dr\right)^{1/2}}, \\ \sqrt{T}c(K)'\hat{b}(K) &= O_p(1), \\ t_{c(K)'\hat{b}(K)} &= O_p(1). \end{aligned}$$

定理の内容は，純粋な I(1) の場合の結果と本質的に同じである．すなわち，near-I(1) においても，説明変数として，そのラグと三角関数のトレンドを使った回帰は，ラグだけでなく，本来は不要である確定的トレンドも有意性を示すことになる．

他方， K -asymptotics については，次の結果が成り立つ．

定理 2 定理 1 と同じ DGP およびモデルに対して，次の K -asymptotics が成り立つ．

$$\begin{aligned} Z_\rho &\Rightarrow N\left(-\frac{\pi^2 K}{2}, \frac{\pi^4 K}{6}\right), \\ Z_t &\Rightarrow N\left(-\frac{\pi\sqrt{K}}{2}, \frac{\pi^2}{24}\right), \\ \sqrt{T}c(K)'\hat{b}(K)/\sigma_L &\Rightarrow N\left(0, \frac{\pi^4 K^2}{4} \sum_{k=1}^K \frac{c_k^2}{\lambda_k}\right), \\ t_{c(K)'\hat{b}(K)} &= O_p(\sqrt{K}). \end{aligned}$$

K -asymptotics の結果も，純粋な I(1) 系列の場合と同様であり， Z_ρ や Z_t などの単位根分布は，I(1) と near-I(1) で不変となる．しかし，このことから，

直ちに K -asymptotics の世界では単位根検定は無効である、と主張するのは速断に過ぎる。なぜなら、ここで使われたモデル (13) は、説明変数として三角関数のトレンド $f_k(r)$ を含むが、それは c の値を既知として計算されるものである。しかし、それは $\rho = 1 - (c/T)$ の値を知っていることと同等であり、単位根検定のためのモデルとして使うことはできない。そのためのモデルは、トレンドが真の c の値に依存しないものでなければならない。そこで、適切なモデルとして、(13) の $f_k(r)$ を、純粋な $I(1)$ の場合の回帰で使ったトレンド $\phi_k(r)$ で置き換えたものを考える。

このような置き換えをすることにより、(13) を単位根検定のモデルとして使うことができる。しかし、そのときにも、本質的に定理 2 と同様の結果が成り立つ。この点に関しては、次節で簡単なシミュレーションにより確かめることにする。このことから、単位根検定の無効性が主張される。すなわち、本来存在していない確定的トレンドを考慮するあまり、確率的トレンドの存在の有無に関する判断が不可能になったと言える。このことは、確定的トレンドの項数を増やすに従って near- $I(1)$ の情報がその中に吸収されてしまい、残差の中には純粋な $I(1)$ と同じ情報しか残らなくなってしまう、ということを示唆している。

4. シミュレーション

ここでは、前節で得られた結果をシミュレーションにより確認する。簡単化のため、DGP は

$$(14) \quad y_t = \rho y_{t-1} + \varepsilon_t, \quad y_0 = 0, \quad \rho = 1 - (c/T), \quad (t = 1, \dots, T)$$

とする。ここで、 $\{\varepsilon_t\} \sim \text{NID}(0, 1)$ であり、 c は非負定数である。

他方、モデルとしては、以下の 2 つのモデル

$$(15) \quad y_t = \sum_{k=1}^K \tilde{b}_k g_k\left(\frac{t}{T}\right) + \tilde{\varepsilon}_t, \quad (t = 2, \dots, T)$$

$$(16) \quad y_t = \hat{\rho} y_{t-1} + \sum_{k=1}^K \hat{b}_k g_k\left(\frac{t}{T}\right) + \hat{\varepsilon}_t, \quad (t = 2, \dots, T)$$

を考える。ここで、 $g_k(r)$ は c に依存しない三角関数 $\phi_k(r)$ か、 c に依存する三角関数 $f_k(r)$ のいずれかを表す。なお、 $\hat{\rho}$ を求めるには (16) 式そのものを

使うよりも、まず、回帰

$$y_{t-1} = \sum_{k=1}^K \check{b}_k g_k \left(\frac{t}{T} \right) + \check{\varepsilon}_{t-1}, \quad (t = 2, \dots, T)$$

を行って残差を計算した上で、

$$\hat{\rho} = \frac{\sum_{t=2}^T \check{\varepsilon}_{t-1} \tilde{\varepsilon}_t}{\sum_{t=2}^T \check{\varepsilon}_{t-1}^2} = \frac{\sum_{t=2}^T \check{\varepsilon}_{t-1} y_t}{\sum_{t=2}^T \check{\varepsilon}_{t-1}^2}$$

により求める方がよい。

以下のシミュレーション結果では、標本サイズは $T = 400$ 、繰り返し数は 1,000 に固定する。まず、図 5 は DGP がランダム・ウォークの場合の $\hat{\rho}$ の分布のヒストグラム（上段は $K = 1$ 、下段は $K = 20$ ）である。また、曲線は、 K -asymptotics

$$T(\hat{\rho} - 1) \Rightarrow N \left(-\frac{\pi^2 K}{2}, \frac{\pi^4 K}{6} \right)$$

に基づく正規分布の密度関数である。 $K = 1$ の場合は、通常の単位根分布であり、左にゆがんだ形状であるが、 $K = 20$ と大きくなると正規分布に近づく様子が見てとれる。

図 5

次に、near-I(1) の DGP として $\rho = 0.95$ の場合を考える。 $T = 400$ であるから、 $c = T(1 - \rho) = 20$ である。この場合の $\hat{\rho}$ の分布が図 6 に示されている。どちらも $K = 20$ であるが、確定的トレンドとして、上段では $f_k(r)$ 、下段では $\phi_k(r)$ が使われている。両者のヒストグラムはほとんど同じである。さらに、図 5 の下段に示された $\rho = 1$ の場合のヒストグラムとも形状がよく似ている。このことは、 $\hat{\rho}$ の分布が、曲線で示された $\rho = 1$ の場合の理論的分布でよく近似されていることを意味する。

図 6

本論では、確定的トレンドとしては不自然な三角関数を使ってきたが、最後に通常の単位根モデル

$$(17) \quad y_t = \hat{\rho} y_{t-1} + \hat{b}_0 + \sum_{k=1}^K \hat{b}_k t^k + \hat{u}_t, \quad (t = 2, \dots, T)$$

を取り上げよう．DGP は (14) の near-I(1) である．この場合も， $\hat{\rho}$ を求めるには，三角関数のトレンドの場合と同様に二段階に分けて計算する方がよい．ただし，上の多項式回帰は説明変数から作られる行列（ヒルベルト行列）が非常に不安定なので， K が 10 程度になると行列計算が困難となる．そこで，ルジャンドル多項式などの直交多項式を使った回帰を考えることにする．このようにしても，結果に影響はない．

通常の単位根検定では，(17) において $K = 1$ の場合がよく使われる．そのときの $\hat{\rho}$ の分布は図 7 のようになる．ここでは， $\rho = 1, 0.975, 0.95$ の 3 つの場合を示している．左側の網掛けの部分は，単位根検定における有意水準 5% に対応する棄却域あるいは検出力である．単位根検定は，この他にもさまざまなものが考えられ，局所対立仮説のもとでの検出力の観点から相互比較が可能である（Tanaka [1996]）．この点に関する議論はしないが， $K = 1$ である今の場合の単位根検定は一致性をもっている．

図 7

しかし， K が大きくなるときの検定は無効になる．図 8 は， $K = 20$ の場合の $\hat{\rho}$ のヒストグラムを図示したものである．上段は $\rho = 1$ ，下段は $\rho = 0.95$ の場合であるが，両者は同じような形状を示しており，このモデルに基づく単位根検定は明らかに無効となる．なお，曲線は， $\rho = 1$ の場合に Nabeya [1999] により得られた結果

$$T(\hat{\rho} - 1) \Rightarrow N(-4K, 16K)$$

に基づく正規分布の密度関数である．この結果は上段の場合にはよく当てはまり，下段についても乖離は見られないが，near-I(1) の場合にも理論的に成り立つかどうかは未解決である．

図 8

5. 示唆と結論

本稿では，まず，Phillips [1998] の議論に沿って，確定的トレンド（多項式あるいは三角関数など）を含まない純粋に確率的な I(1) 過程（確率的トレンド）を DGP とする時系列に対して，確定的トレンドを使った回帰は，大標本

のもとでトレンドの項数 K を増やすとき、対応する係数が有意となり、回帰の決定係数が 1 に近づくという意味で spurious でないことを理論的に解明した。さらに、Phillips [1999] に基づいて、単位根検定のモデル、すなわち、1 期前のラグと確定的トレンドを説明変数とする回帰モデルを使って、大標本のもとで確定的トレンドの項数 K が増えるときの単位根検定統計量の帰無分布に関する漸近的な挙動を調べた。

さらに、単位根検定の検出力を調べるために、Phillips [1998, 1999] の議論を拡張して、near- $I(1)$ を DGP とする場合の K -asymptotics を調べた。

これらの議論から得られる結果と示唆は次の通りである。

- (i) 確率的トレンドである $I(1)$ 過程は、無数の確定的なトレンドの一次結合として表現することが可能である（ただし、係数は確率変数）。したがって、非定常時系列が「確定的トレンド」をもつのか、あるいは「確率的トレンド」をもつのか、という二者択一の言い方は必ずしも適切ではない。ここで興味深いことは、トレンドから連想されるイメージとは異なって、 $I(1)$ 過程を近似する確定的トレンドは必ずしも多項式とは限らずに、本稿で使った三角関数のように有界な周期関数でも構わない、という点である。
- (ii) ラグ変数に加えて確定的トレンドを説明変数とするモデルにおいては、DGP が $I(1)$ であるにもかかわらず、確定的トレンドもラグ変数とともに有意となる。すなわち、確率的トレンドの動きが本来は不要な確定的トレンドも使うことで説明される。したがって、確率的トレンドと確定的トレンドの関係は、必ずしもお互いの否定、あるいは代替とはならない。
- (iii) K -asymptotics では、near- $I(1)$ における単位根分布が $I(1)$ におけるそれと変わらない、という意味で単位根検定は無効になる。この原因も、本来は不要である確定的トレンドをモデルに含めたことに帰せられる。

トレンドが実際に存在するとしても、それが確定的か確率的かを判断することのむずかしさは、トレンドそのものが観測不可能であることに由来する。観測不可能ゆえ、我々はトレンドをいかようにもモデル化することができる。しかし、上述のごとく、 K -asymptotics の世界においてはトレンドが確定的か確率的かを判断するための統計的検定も無力となってしまう。

もちろん、ここで考えたモデルは、DGP をネストしない、あるいはネストしているにしても、DGP からかなり離れたモデルである。もし、DGP の構造に関する何らかの情報があれば、このような極端なモデルは選択しないであろう。しかし、事前情報が何もない場合には、今まで考察してきた例のようにトレンドの特定化は困難になる、ということが言える。トレンドは実体があるわけではなく、観測されるデータの実現プロセスをモデル化する際に分析者が自由に定式化できるものである。いずれのモデル化が優れているかの判断は、データに基づいて統計的に行われるが、モデルによってはそれが不可能な状況があるということを指摘したい。非定常時系列におけるトレンドのモデル化とよさの評価は、古くて新しいチャレンジングな統計的問題である。

- Chan, N. H. and C. Z. Wei (1988) “Limiting distributions of least squares estimates of unstable autoregressive processes,” *Annals of Statistics*, Vol. 16, pp. 367-401.
- Courant, R. and D. Hilbert (1953) *Methods of Mathematical Physics*, Vol. I, New York: Wiley.
- Csörgő, K. and L. Horváth (1993) *Weighted Approximations in Probability and Statistics*, New York: Wiley.
- Durlauf, S. N. and P. C. B. Phillips (1988) “Trends versus Random Walks in Time Series Analysis,” *Econometrica*, Vol. 56, pp. 1333-1354.
- Hochstadt, H. (1973) *Integral Equations*, New York: Wiley.
- Loève, M. (1978)] *Probability Theory II*, Fourth Edition, New York: Springer-Verlag.
- Nabeya, S. (1999) “Unit Root Seasonal Autoregressive Models with a Polynomial Trend of Higher Degree,” To appear in *Econometric Theory*.
- Phillips, P. C. B. (1987) “Time Series Regression with a Unit Root,” *Econometrica*, Vol. 55, pp. 277-301.
- (1998) “New Tools for Understanding Spurious Regressions,” *Econometrica*, Vol. 66, pp. 1299-1325.
- (1999) “New Unit Root Asymptotics in the Presence of Deterministic Trends,” Discussion Paper, Cowles Foundation for Research in Economics, Yale University.
- Tanaka, K. (1990) “The Fredholm Approach to Asymptotic Inference on Nonstationary and Noninvertible Time Series Models,” *Econometric Theory*, Vol. 6, pp. 411-432.

- (1993) “An Alternative Approach to the Asymptotic Theory of Spurious Regression, Cointegration, and Near Cointegration,” *Econometric Theory*, Vol. 9, pp. 36-61.
- (1996) *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*, New York: Wiley.
- (2001) “ K -Asymptotics Associated with Deterministic Trends in the Integrated and Near-Integrated Processes,” *Japanese Economic Review*, Vol. 52, pp. 35-63.
- Whittaker, E. T. and G. N. Watson (1958) *A Course of Modern Analysis*, Fourth Edition, London: Cambridge University Press.

表 1 R^2 と DW : I(1) 系列の DGP に (5) のモデルをあてはめた場合

	$K = 1$	2	5	10	20	50
R^2						
$T = 400$						
Mean	0.594	0.752	0.887	0.943	0.971	0.989
SD	0.325	0.236	0.121	0.060	0.032	0.013
$T = 800$						
Mean	0.591	0.752	0.888	0.942	0.971	0.988
SD	0.314	0.235	0.117	0.063	0.032	0.012
DW						
$T = 400$						
Mean	0.039	0.064	0.136	0.256	0.484	1.100
理論値	0.025	0.049	0.123	0.247	0.493	1.234
SD	0.024	0.031	0.047	0.063	0.081	0.104
$T = 800$						
Mean	0.020	0.033	0.069	0.129	0.249	0.586
理論値	0.012	0.025	0.062	0.123	0.247	0.617
SD	0.012	0.016	0.023	0.032	0.043	0.062

図 1 R^2 の分布：I(1) 系列に (5) のモデルをあてはめた場合
 (上段： $K = 1$ 下段： $K = 20$)

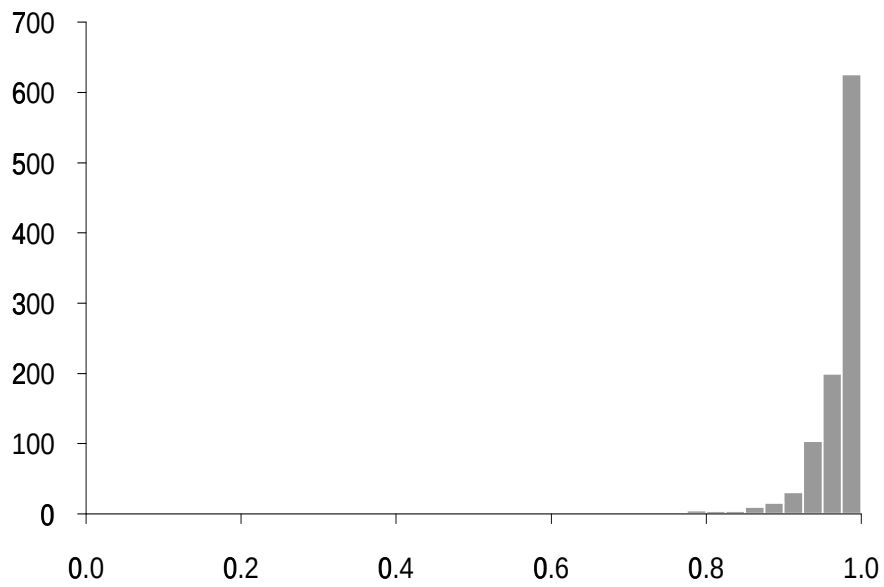
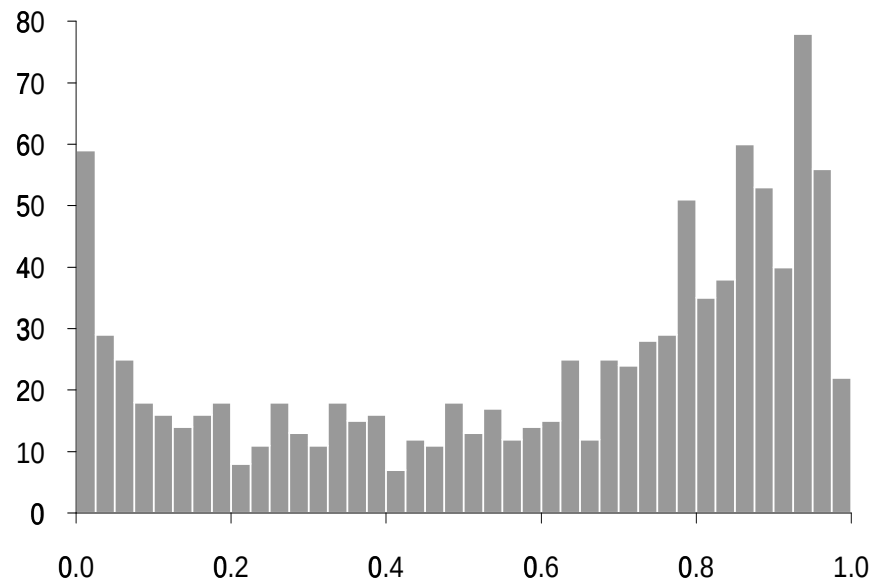


図2 R^2 の分布：定常な AR(1) に (5) のモデルをあてはめた場合
(上段： $K = 1$ 下段： $K = 20$)

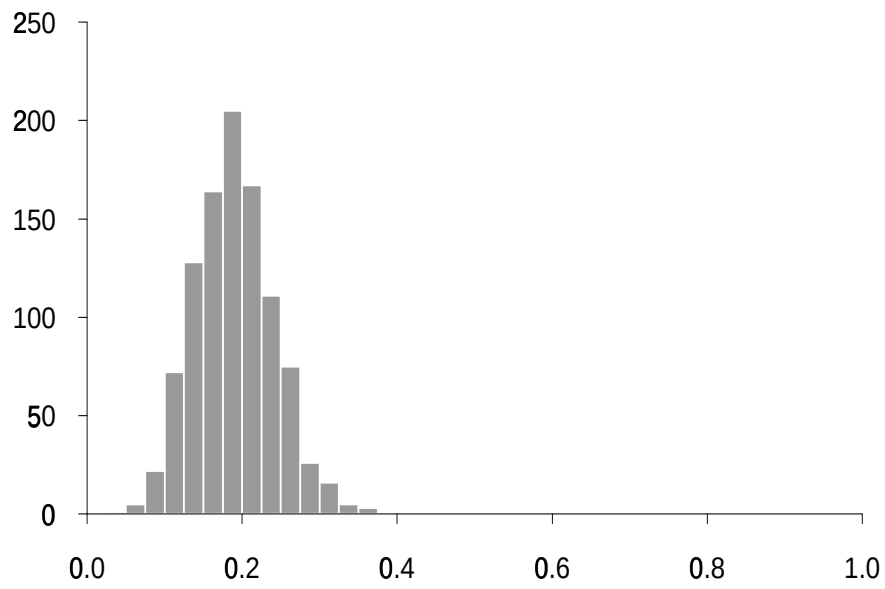
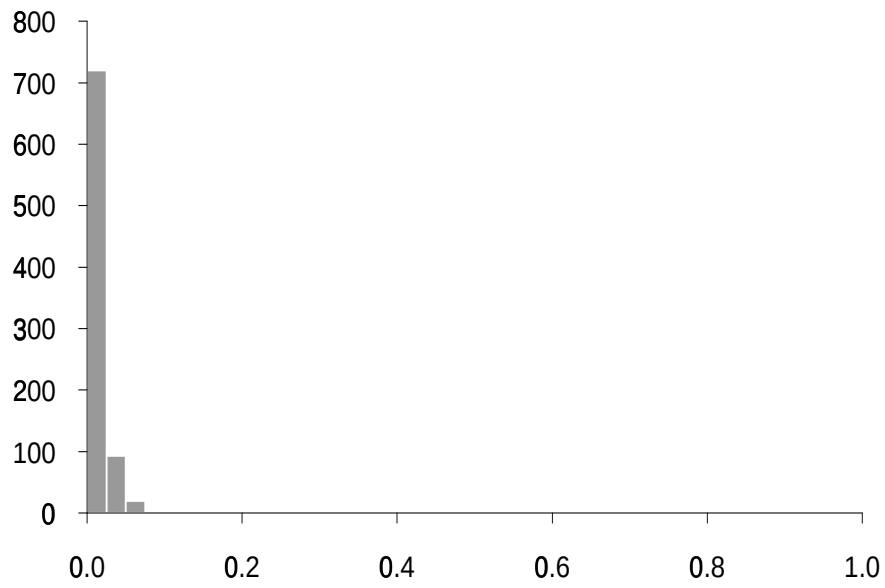


図3 DW の分布 : I(1) 系列に (5) のモデルをあてはめた場合
(上段 : $K = 1$ 下段 : $K = 20$)

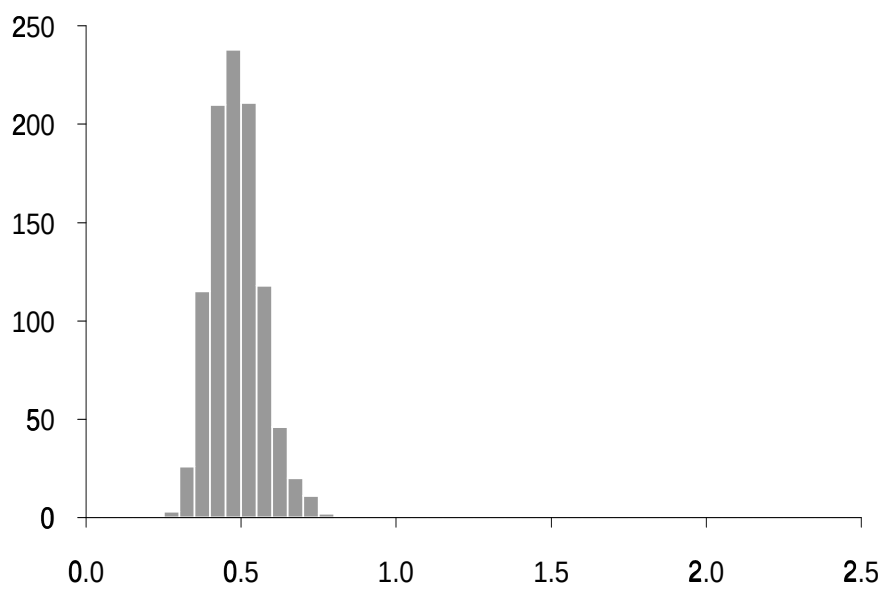
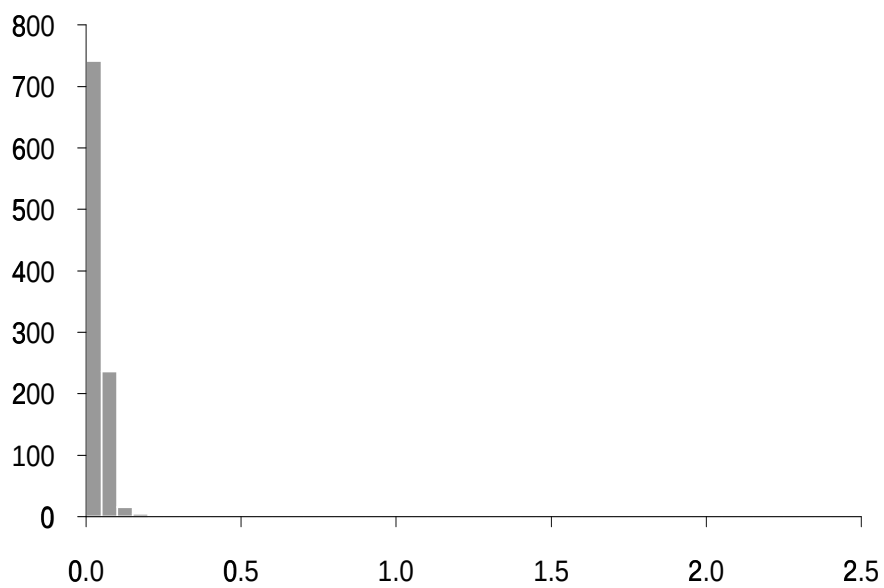


図4 DW の分布：定常な $AR(1)$ に (5) のモデルをあてはめた場合
(上段： $K = 1$ 下段： $K = 20$)

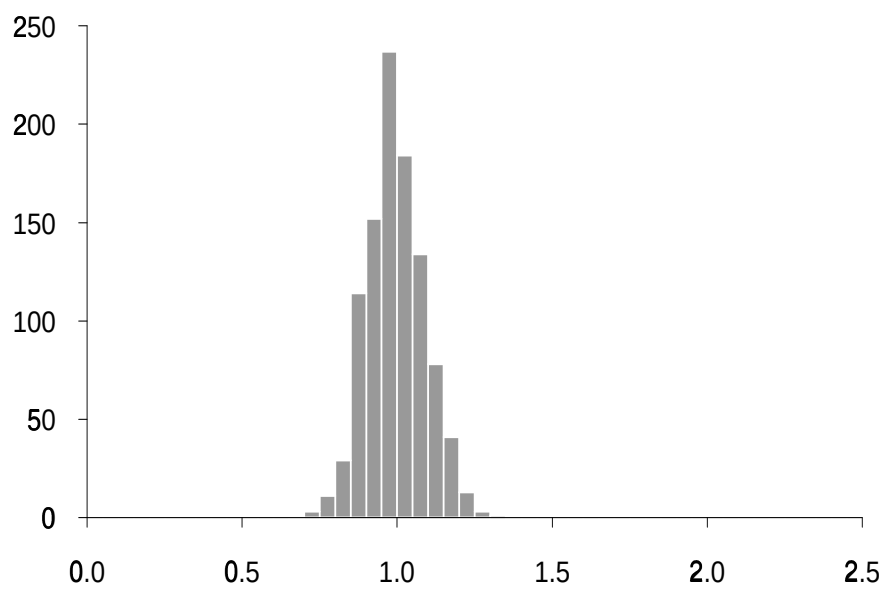
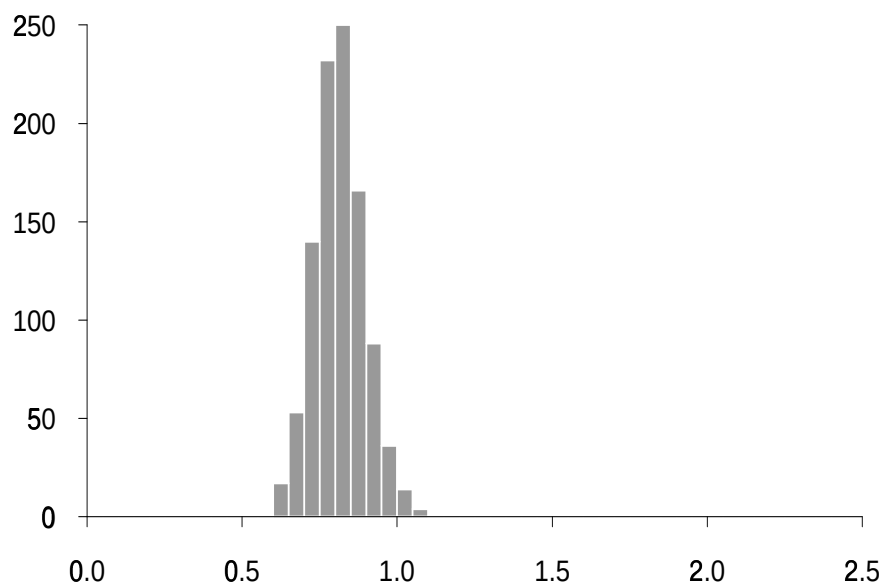


図5 $T(\hat{\rho} - 1)$ の分布

(上段: $\rho = 1, K = 1$ 下段: $\rho = 1, K = 20$)

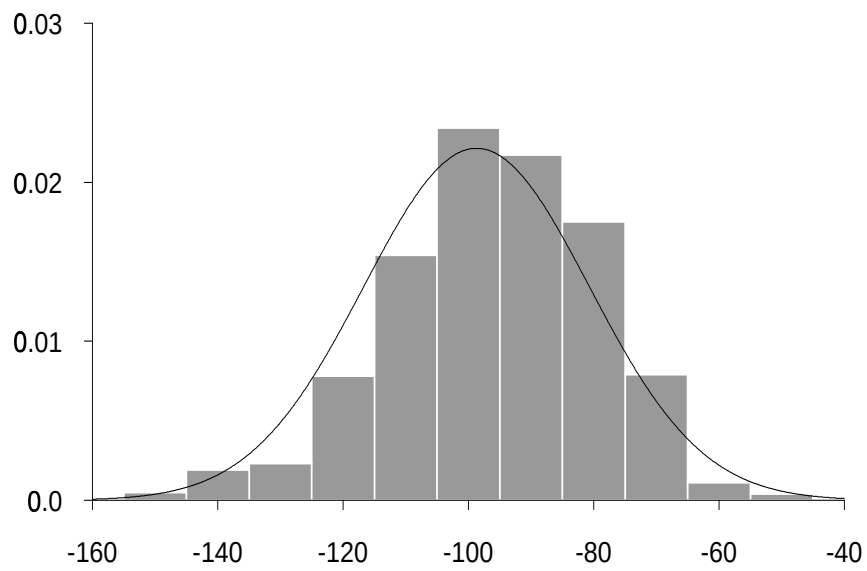
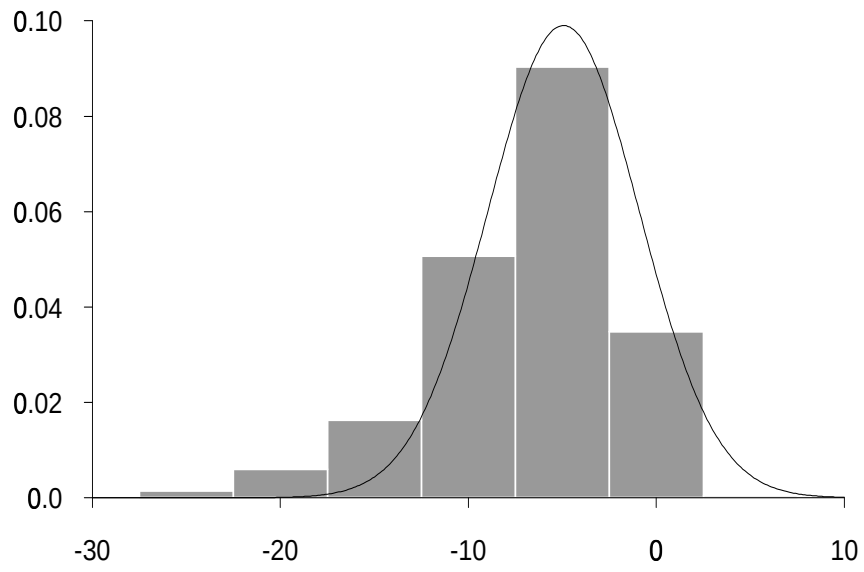


図6 $T(\hat{\rho} - 1)$ の分布 ($\rho = 0.95, K = 20$)

(上段：トレンド = $f_k(r)$ 下段：トレンド = $\phi_k(r)$)

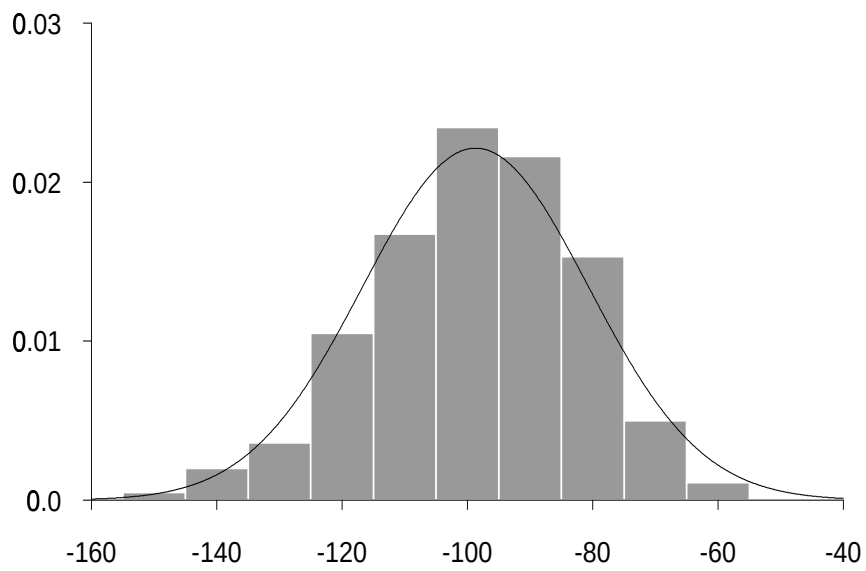
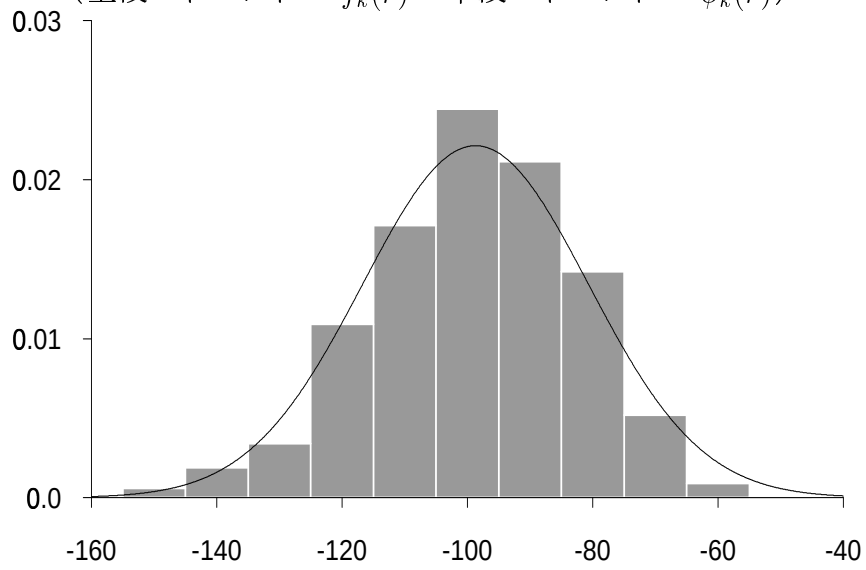


図7 通常の単位根分布

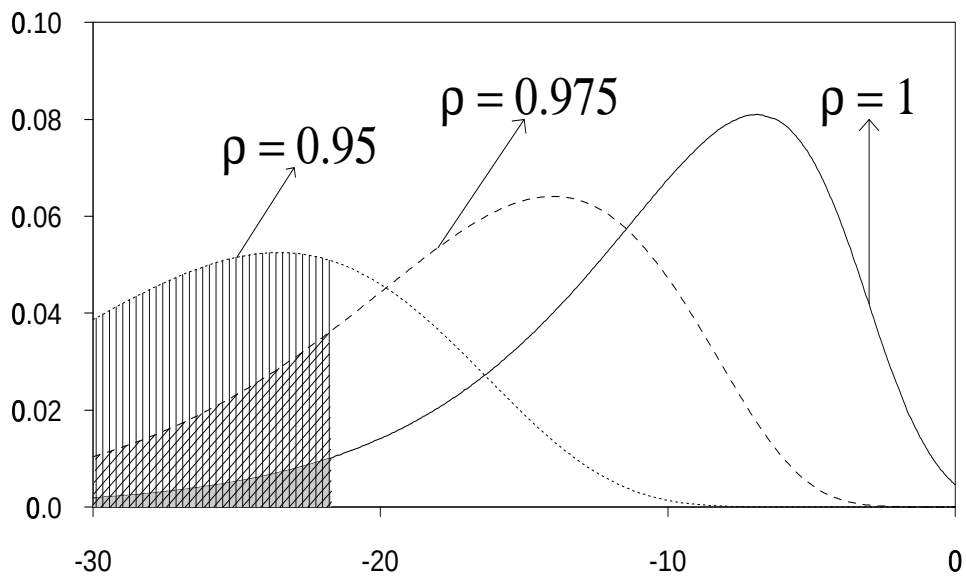


図8 $T(\hat{\rho} - 1)$ の分布 (多項式トレンド $K = 20$ の場合)
 (上段: $\rho = 1$ 下段: $\rho = 0.95$)

