

徳山大学総合経済研究所モノグラフ 11

## ベイジアン単位根検定の展開

河 田 正 樹

## はじめに

2003年のノーベル経済学賞は R. Engle, C. W. J. Granger 両教授が受賞した。受賞の対象は、経済時系列分析に関するものであり、Engle が ARCH モデル、Granger が共和分 (cointegration) についての研究であった。これらの研究は現在の計量経済分析、計量ファイナンス分析に広く用いられている。

Granger の受賞対象となった共和分とはどういうものであるかを簡単に説明してみよう。階差定常な2つの時系列がその線形結合をとることによって定常な系列をつくることができるとき、2つの時系列は共和分関係にあるという。また、共和分分析とは階差定常な複数の時系列間の回帰関係を分析するものといえる。

ある時系列が階差定常な系列かどうかを検定するものが単位根検定である。共和分分析は単位根検定をおこなった結果「単位根がある」と判定された複数の系列についての分析であり、単位根検定は共和分分析の出発点となっているといえよう。

単位根検定は1970年代後半に Dickey - Fuller によって最初に考案され、数多くの実証研究に用いられている。しかし、この検定には検出力の低さなどの問題があり、現在でもさまざまな議論がなされている。なかでも Sims がベイジアン立場からおこなった Dickey - Fuller 流の単位根検定に対する痛烈な批判は大きな波紋を呼び、さまざまな研究者がベイジアン、ノンベイジアン(本書ではノンベイジアンのアプローチを伝統的アプローチとよぶことにする)それぞれの立場からこの問題についての検討をおこなった。

単位根検定には、単位根を含むモデルを帰無仮説とする検定と定常なモデルを帰無仮説とする検定の2つの種類がある。本書は1章で単位根検定とは何かについて簡単に触れた上で、2章では単位根を含むモデルを帰無仮説とする検定について、3章では定常なモデルを帰無仮説とする検定について、そ

れぞれ伝統的アプローチとベイジアンアプローチを紹介し、考察をおこなった。4章は最近考案されているこの2種類の検定を結合する試みについて触れている。これらは筆者が大学院博士後期課程在学中から執筆してきた論考に大幅な加筆修正を加え、体系だてて整理したものである。

このように本書をまとめることができたのも、さまざまな方々のご指導、ご助力のおかげである。徳山大学総合経済研究所には平成14年度1号部門研究「単位根問題におけるベイジアンアプローチに関する研究」の研究助成を受け、このような機会をいただいた。印刷を担当してくださった中西印刷株式会社の方々には、数々の難しい注文に対応していただいた。心よりお礼を申し上げたい。また学部、大学院を通じて筆者を根気強く指導をしてくださった佐竹元一郎早稲田大学名誉教授にもこの場を借りてお礼申し上げたい。

2004年3月

河田 正樹

# 目次

はじめに	i
<b>1</b> 単位根検定とは	<b>1</b>
<b>2</b> 単位根を帰無仮説とする検定	<b>5</b>
2.1 伝統的アプローチ	5
2.2 ベイジアンアプローチ	7
2.2.1 Sims の主張	7
2.2.2 Phillips の批判	11
2.2.3 事前分布に関する論争	15
2.2.4 ベイジアン単位根検定	17
<b>3</b> 定常帰無仮説の単位根検定	<b>25</b>
3.1 伝統的アプローチ	25
3.2 ベイジアンアプローチ	28
3.2.1 シミュレーションによる検証	31
<b>4</b> 結合の動き	<b>39</b>
4.1 伝統的アプローチ	39
4.2 ベイジアンアプローチ	40
おわりに	45

# 1 単位根検定とは

時間領域における時系列モデル分析は、時間順に並べた観測値の列  $\{y_t; t = 1, \dots, T\}$  について、それを生み出した確率過程を探りだし、その結果を用いて予測などをおこなうことを目的としている。確率過程を探る際に平均、分散、自己共分散といった特性値を推測する必要があるが、各期に1個の実現値しか持たない経済データの場合には、そのままでは特性値を推測することは不可能である。そのため、

$$E(y_t) = \mu \quad (1.1)$$

$$\text{Var}(y_t) = \sigma^2 \quad (1.2)$$

$$\text{Cov}(y_t, y_s) = \gamma_{t-s} \quad (t \neq s) \quad (1.3)$$

という定常性の仮定をおいて分析をおこなっている。これは、平均および分散がそれぞれ一定の値をとり、共分散は2時点間の差のみに依存するという仮定である。

しかし、経済分析に用いられるデータのほとんどは、この定常性の条件が満たされない非定常なデータである。そのため非定常系列は定常系列に変換して分析することが必要となる。Box - Jenkins 流の方法では、平均非定常な系列は階差を適当な回数とることによって定常系列へと変換し、分析が行われる。定常系列に変換されたかどうかの判断は、自己相関や偏自己相関を見ることによっておこなう。

平均非定常な系列を定常系列に変換するための方法としては、線形トレンドをあてはめることにより非定常性をとり除くという方法も考えられる。階差をとることによって非定常性をとり除くモデル(階差モデル)と、線形トレンドをあてはめることによって非定常性をとり除くモデル(トレンドモデル)

のいずれを用いるかによって、予測結果にはかなりの差が生じることが知られている。そこで研究者の恣意的判断でなく、科学的な根拠にもとづく処理方法の選択が必要となっている。処理方法の選択に際して統計的検定をおこなうために用いられるものが単位根検定である。単位根検定とは、モデルのAR部分にパラメータが1に等しくなるようなパラメータ(単位根)があるかどうかを検定するものであり、単位根を含むモデルは、階差をとる変換をほどこすことが可能となる。

単位根検定には、帰無仮説および対立仮説のとり方に、さまざまな組み合わせがある。1階の自己回帰(AR(1))モデルを例に線形トレンドの有無、単位根の有無によって分類してみると次のようになる。

表 1-1 単位根検定における仮説の分類

	単位根含む	単位根含まず
線形トレンドなし	$H_1 : y_t = y_{t-1} + u_t$	$H_3 : y_t = \rho y_{t-1} + u_t$ $H_4 : y_t = \mu' + \rho y_{t-1} + u_t$ $\mu' = (1 - \rho)\mu$
線形トレンドあり	$H_2 : y_t = \beta + y_{t-1} + u_t$	$H_5 : y_t = \mu' + \beta' t + \rho y_{t-1} + u_t$ $\mu' = (1 - \rho)\mu + \beta\rho$ $\beta' = \beta(1 - \rho)$

このモデルにおいて、 $H_2$  vs.  $H_5$ 、 $H_1$  vs.  $H_5$ 、 $H_1$  vs.  $H_4$ 、 $H_1$  vs.  $H_3$  といった検定をおこなうことが考えられる

$H_2$  vs.  $H_5$  の検定は時系列に線形トレンドが含まれていることが分かっているときの検定であり、 $H_5$  において  $\rho = 1$  かどうかを検定する。 $\rho = 1$  のとき、モデルはドリフトのあるランダムウォークとなり、 $\rho < 1$  なら線形トレンドを持つ定常な AR(1) モデルとなる。

$H_1$  vs.  $H_5$  の検定は時系列の持つトレンドが、線形トレンドなのか確率トレンド<sup>1</sup>なのかを検定するものであり、 $H_5$  において  $\rho = 1 \cap \beta = 0$  かどうか

<sup>1</sup>階差をとることによって除去されるトレンドのことを確率トレンドという。

を検定する。  $\rho = 1$  のとき、モデルはランダムウォークとなり、  $\rho < 1$  なら線形トレンドを持つ定常な AR(1) モデルとなる。

$H_1$  vs.  $H_4$  の検定は定数項は存在するが、線形トレンドは含まれないと考えられるときの検定であり、  $H_4$  において  $\rho = 1$  かどうかを検定する。  $\rho = 1$  のとき、モデルはランダムウォークとなり、  $\rho < 1$  なら定数項周辺での定常な AR(1) モデルとなる。

$H_1$  vs.  $H_3$  の検定は定数項も線形トレンドも含まれないと考えられるときの検定であり、  $H_3$  において  $\rho = 1$  かどうかを検定する。  $\rho = 1$  のとき、モデルはランダムウォークとなり、  $\rho < 1$  なら  $x$  軸周辺の定常な AR(1) モデルとなる。

さて、通常の回帰分析において

$$y_t = \alpha + \beta x_t + u_t, \quad (t = 1, \dots, T) \quad (1.4)$$

で  $H_0 : \beta = \beta_0$  vs.  $H_1 : \beta \neq \beta_0$  という検定をおこなう場合、  $\beta$  の推定値  $\hat{\beta}$  を用いて

$$t = \frac{\hat{\beta} - \beta_0}{s_{\hat{\beta}}} \quad (s_{\hat{\beta}} \text{ は } \hat{\beta} \text{ の標準誤差}) \quad (1.5)$$

という統計量を考えると、  $t$  は自由度  $T - 2$  の  $t$  分布にしたがうので  $t$  検定をおこなうことができる。しかし、単位根検定の場合には  $\rho$  の分布が  $\rho = 1$  のときのみ特殊な分布となるので、このような  $t$  検定をおこなうことができない。そのため、特別な検定方法が必要となる。それをおこなうのが単位根検定である。

## 2 単位根を帰無仮説とする検定

### 2.1 伝統的アプローチ

単位根検定は Dickey と Fuller による一連の研究 (Fuller [8] Dickey and Fuller [5] など) によって最初に提唱された。この 2 人の提唱した検定方法は Dickey - Fuller 流の検定<sup>1</sup>といわれている。

$$y_t = \rho y_{t-1} + u_t \quad (2.1)$$

という定数項なしの AR(1) モデルにおいて、 $\rho$  の最小 2 乗推定量  $\hat{\rho}$  の性質を考える。  $|\rho| < 1$  である場合には、

$$\sqrt{T}(\hat{\rho} - \rho) \quad (2.2)$$

の分布が漸近的に  $N(0, 1 - \rho^2)$  にしたがう。しかし、 $\rho = 1$  となる場合には  $\hat{\rho}$  は、定常の場合よりも早く 1 に収束し、 $T(\hat{\rho} - 1)$  が漸近的にある分布にしたがうが、その分布はかなり複雑な分布である<sup>2</sup>。

この分布は理論的に導出することが難しいために、Dickey はモンテカルロ実験によってこの経験分布を求めた。この経験分布は Fuller [8] において表の形であらわされている。最小 2 乗法で求めた  $\hat{\rho}$  から  $T(\hat{\rho} - 1)$  を計算し、この分布表の値と比較するものが DF 検定の 1 つの方法である。

DF 検定の他の方法として、通常の回帰分析における t 検定にあたるやり方が考えられる。すなわち  $\hat{\rho}$  の標準誤差を  $s_\rho$  とするとき、

$$\hat{\tau} = \frac{\hat{\rho} - 1}{s_\rho} \quad (2.3)$$

---

<sup>1</sup>以下この検定のことを DF 検定ということにする。

<sup>2</sup>Fuller [8]p. 367.



によって得られる  $\hat{\tau}$  を計算するものである。  $\rho = 1$  の場合、この  $\hat{\tau}$  の値もやはり  $t$  分布にはしたがわないので、Dickey が求めた  $\hat{\tau}$  分布表と比較し、仮説検定をおこなう。

この2つのタイプの単位根検定は Dickey - Fuller の  $t$  値タイプの単位根検定といわれる検定法である。この検定は対象となるモデルによって用いる分布表が異なり、定数項を含んだモデル

$$y_t = \mu' + \rho y_{t-1} + u_t \quad (2.4)$$

の場合には  $T(\hat{\rho}_\mu - 1)$  または  $\hat{\tau}_\mu$  の分布表を用い、定数項と線形トレンドを含むモデル

$$y_t = \mu' + \beta' t + \rho y_{t-1} + u_t \quad (2.5)$$

の場合には、 $T(\hat{\rho}_\tau - 1)$  または  $\hat{\tau}_\tau$  の分布表を用いる。

DF 検定は Dickey と Fuller を中心とする研究者たちによって、いくつかの発展がなされた<sup>3</sup>。AR の次数が高次である場合や、MA 項を含むモデルなどへ拡張がなされた。これらのモデルの検定は拡張 (Augmented)DF 検定といわれる。

ADF 検定ではモデルを次のような階差の形で表現する。定数項、線形トレンドの有無によって3種類のモデルに分類できる。

$$\Delta y_t = \rho y_{t-1} + \rho_1 \Delta y_{t-1} + \cdots + \rho_p \Delta y_{t-p} + u_t \quad (2.6)$$

$$\Delta y_t = \mu'' + \rho y_{t-1} + \rho_1 \Delta y_{t-1} + \cdots + \rho_p \Delta y_{t-p} + u_t \quad (2.7)$$

$$\Delta y_t = \mu'' + \beta'' t + \rho y_{t-1} + \rho_1 \Delta y_{t-1} + \cdots + \rho_p \Delta y_{t-p} + u_t \quad (2.8)$$

これらのモデルについて  $H_0 : \rho = 1$  vs.  $H_1 : \rho < 1$  という検定をおこなうのが ADF 検定である。

---

<sup>3</sup>Dickey and Fuller [6], Said and Dickey [34] など。

この ADF 検定にノンパラメトリック修正をおこなったものが Phillips - Perron の  $Z_\rho$  検定、 $Z_t$  検定<sup>4</sup>といわれるものである。 $Z_\rho$  が  $T(\hat{\rho} - 1)$  に修正を加えたものであり、 $Z_t$  が  $\hat{\tau}$  に修正を加えたものである。この検定統計量も定数項、トレンドの有無によってその式が変わる。

伝統的アプローチの単位根検定を実証分析に適用したのものとして Nelson and Plosser [28] がある。彼らはアメリカの 14 のマクロ経済系列に ADF 検定を適用し、失業率を除く 13 の系列に単位根の存在を認めた。しかしこの結果は、ほとんどの系列において帰無仮説を棄却しないことから、単位根検定の検出力の低さが問題とされるようになった。

## 2.2 ベイジアンアプローチ

### 2.2.1 Sims の主張

Sims [36] は、伝統的アプローチによる単位根問題およびこの問題に関連した計量経済学を不毛な考え (sterile ideas) として痛烈に批判した<sup>5</sup>。

Sims の批判は、伝統的アプローチが  $\rho = 1$  の漸近分布のみを特別に取り扱うことに向けられている。 $|\rho| < 1$  のときの漸近分布と  $\rho = 1$  のときの漸近分布が異なるために、さまざまな矛盾が生じているというものである。

伝統的アプローチにおいて、仮説検定は区間推定と密接に関連しており、点推定値  $\hat{\theta}$  まわりの  $100(1 - \alpha)\%$  信頼区間をとった区間推定値を考えた場合、この区間内に  $\theta_0$  が含まれていれば、通常は  $H_0 : \theta = \theta_0$  vs.  $H_1 : \theta \neq \theta_0$  という検定を有意水準 5% でおこなった場合、帰無仮説は採択される。

しかし単位根問題の場合、 $\rho = 1$  のときのみ、その漸近分布が特殊な分布にしたがうという不連続性のため、このような性質は成り立たない。たとえ

---

<sup>4</sup>Phillips [29], Phillips and Perron [31] によって提唱されたものである。以下この検定のことを PP 検定ということにする。

<sup>5</sup>Sims [36] p.463.

ば、標本数  $T = 100$  の時系列データを定数項なし、トレンドなしの AR(1) モデルを用いて、最小 2 乗法で推定し、

$$y_t = 0.9096 \times y_{t-1} + u_t \quad (2.9)$$

(0.04177)

という結果を得たとしよう (カッコ内の数値は標準誤差をあらわす)。このとき、 $\hat{\rho}$  まわりの 95% 信頼区間は (0.8269, 0.9923) となり、区間推定値の中に  $\rho = 1$  は含まれない。しかし、 $H_0 : \rho = 1$  という仮説検定をおこなう場合、有意水準 5% の採択域は (0.8411, 1.003) となり<sup>6</sup>、帰無仮説を採択することになる。このような  $\rho = 1$  の場合のみ漸近分布が異なるという不連続性は、「本当は定常なモデルを誤って単位根を含むモデルであるとみなす」ことになると Sims は指摘している<sup>7</sup>。

Sims は、このような問題を解決するために、事前分布を一様分布としたベイジアンアプローチをおこなった。

伝統的アプローチでは、母数  $\theta$  を未知の定数と考え、そこから確率変数であるデータ  $x$  が求められる。そして、データから標本統計量を求め、それにもとづいてパラメータの推測をおこなう。一方ベイジアンアプローチは、未知の母数  $\theta$  自身を確率変数と考える。 $\theta$  に関する事前の情報を確率分布の形で表した事前分布の密度関数を  $p(\theta)$  とし、 $\theta$  を与えたときのデータ  $x$  の分布の確率密度関数を  $p(x|\theta)$  とするとき、

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} \quad (2.10)$$

というベイズの定理を用いて事後分布  $p(\theta|x)$  を導出する。この事後分布が母数  $\theta$  に関する現時点での情報である。

---

<sup>6</sup> $\rho = 1$  の場合の信頼区間の算出には、Fuller [9] p.641 の  $\hat{\tau}$  分布表を用いた。

<sup>7</sup>Sims [36] p.465.

ベイジアンアプローチでは、事前分布にどのような分布を用いるかという問題がある。この事前分布の設定に関して、母数に関する事前の知識がない場合には次のような基準が考えられる<sup>8</sup>。

1. 母数の範囲が  $-\infty$  から  $\infty$  までの間の値をとるとき、 $p(\theta)$  として、一様分布を考える。すなわち  $p(\theta) \propto \text{定数}$  である。
2. 母数の範囲が 0 から  $\infty$  までの間の値をとるとき、 $p(\theta)$  として、 $\theta$  の対数が一様分布となるような分布を考える。すなわち  $p(\log \theta) \propto \text{定数}$  である。これは、 $p(\theta) \propto \theta^{-1}$  である。

この基準は Bayes や Laplace がサンプリングの際に用いた一様な事前分布について、Jeffreys が「ごくふつうの場合であればおおむね成立する」<sup>9</sup>としたものであり、Jeffreys の基準といわれている。この基準による事前分布の密度関数は、積分して 1 になるという密度関数の性質を必ずしも満たさないため、広義事前密度関数といわれることもある。

この Jeffreys の基準によって設定された  $\rho$  の事前分布を一様分布としたベイジアンアプローチでは、伝統的アプローチのような不連続性は起きず、このようなベイジアンアプローチのほうが正しいと Sims は主張し、Sims and Uhlig [37] における次のようなシミュレーション実験によって示した<sup>10</sup>。

*Step1* 100 個の平均 0、分散 1 の正規乱数を発生させてそれを  $u_t$  とし、この  $u_t$  を用いて  $y_t$  を生成する。(ただし  $y_0 = 0$  とする。)

*Step2* 標本数  $T = 100$  の系列  $y_t$  に最小 2 乗法を適用し、推定値  $\hat{\rho}$  を求める。

*Step3* *Step1*、*Step2* の操作を 10000 回繰り返し、 $\hat{\rho}$  の条件付き分布を求める。

*Step4* *Step1*、*Step2*、*Step3* の操作を  $\rho$  について 0.80 から 1.10 の間の 0.01 刻みの値についておこない、 $\rho$  と  $\hat{\rho}$  の同時分布を求める。

<sup>8</sup>Jeffreys [11] p.117. 訳語は繁樹 [33] 54 ページを用いた。

<sup>9</sup>Jeffreys [11] p.117. 訳語は Zellner [43] 邦訳 47 ページによる。

<sup>10</sup>この実験では  $\sigma^2 = 1$  という仮定をおいている。このような仮定をおいても一般性を失わない。

このような手順を踏み、 $\hat{\rho}$  の階級幅を  $[-\infty, 0.795), [0.795, 0.805), \dots$  とし  
て度数を求め、 $(\hat{\rho} \leq 1 \cap \rho \leq 1)$  の部分についてグラフにしたものが 図 2-1 で  
ある。このグラフは Sims and Uhlig のおこなった実験と同様の方法によって  
筆者が作成した。なお、 $\hat{\rho}$  については、各階級の中央値によって代表させ、  
たとえば  $\hat{\rho} = 0.95$  とは、 $0.945 \leq \hat{\rho} < 0.955$  を意味している。

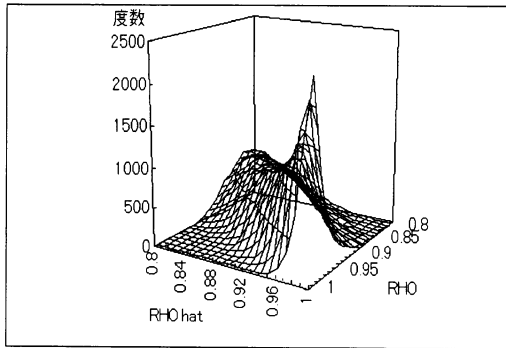


図 2-1  $\hat{\rho}$  と  $\rho$  の同時分布

伝統的アプローチは未知の母数が与えられたときの標本統計量の値を考える。  
すなわちこの同時分布をある  $\rho$  の値においてスライスした  $\hat{\rho}|\rho$  に注目  
する。

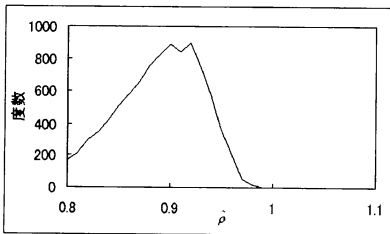


図 2-2  $\hat{\rho}|\rho = 0.9$  の分布

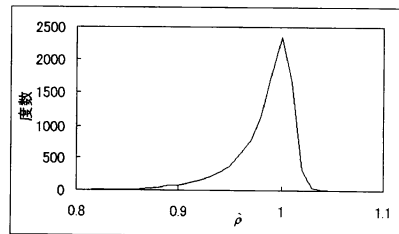


図 2-3  $\hat{\rho}|\rho = 1$  の分布

図 2-2、図 2-3 はそれぞれ  $\rho = 0.9$ 、 $\rho = 1$  のところで図 2-1 の同時分布をスライスしたものである。 $\rho < 1$  のとき、 $\hat{\rho}$  の漸近分布は正規分布となるが、 $\rho$  が 1 に近づくにつれ、正規分布への収束のスピードは遅くなる。 $\rho = 1$  のときの漸近分布は左右非対称となる。

一方、ベイジアンアプローチはデータが与えられたときの母数の分布を考える。この同時分布をある  $\hat{\rho}$  の値においてスライスした  $\rho|\hat{\rho}$  は事前分布に一樣分布を用いた場合の母数の分布とみることができる。

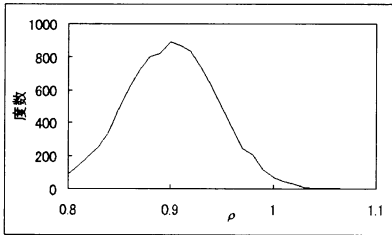


図 2-4  $\rho|\hat{\rho} = 0.9$  の分布

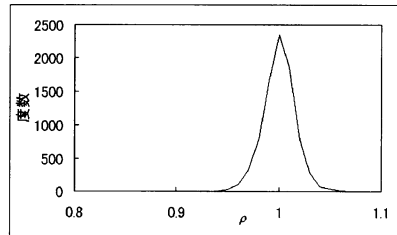


図 2-5  $\rho|\hat{\rho} = 1$  の分布

図 2-4、図 2-5 はそれぞれ  $\hat{\rho} = 0.9$ 、 $\hat{\rho} = 1$  のところで図 2-1 の同時分布をスライスしたものである。散らばりの度合いは異なるものの、 $\hat{\rho} = \rho$  に関して左右対称な分布となっている。このように、 $\rho|\hat{\rho}$  の分布を考えると、分布が左右対称になり、たとえば  $\hat{\rho} = 0.95$  が得られた場合にモデルが  $\rho = 0.9$  となる確率と  $\rho = 1$  となる確率はほぼ等しくなるということがいえる。このようなことから、Sims and Uhlig は伝統的アプローチではなく、ベイジアンのアプローチで単位根の問題を取り扱うほうが良いと主張している。

### 2.2.2 Phillips の批判

このような Sims の提唱したベイジアンアプローチに対して、Phillips [30] は同じベイジアン立場をとりながら、次のような批判をおこなった。

Phillips は Sims のおこなった事前分布として一様分布を採用したベイジアン手法について、説明変数が変動を持たない通常の回帰分析であるならば、事前分布として一様分布を用いる手法でもかまわないが、時系列分析の場合は事前分布の設定の際に時系列データの成り立ちを考慮に入れなくてはならないと考えた。すなわち、通常の回帰分析において係数はある説明変数とそれに対応する平均応答に影響しているにすぎないが、時系列分析においては、係数はデータの自己相関構造全体に影響を与えるものであり、より多くの情報を含んだものとなる。

Phillips はこのような時系列データの成り立ちを考慮に入れた、客観的に無知の状態を示す事前分布として、次のようなものを採用した。

$$p(\theta) \propto |\mathbf{I}_{\theta\theta}|^{1/2} \quad (2.11)$$

ここで  $\theta$  は母数を表している。また、 $\mathbf{I}_{\theta\theta}$  はフィッシャーの情報行列といわれ、

$$\mathbf{I}_{\theta\theta} = E\left\{-\frac{d^2}{d\theta d\theta'} \log p(\mathbf{y}|\theta)\right\} \quad (2.12)$$

と定義されるものである。 $\mathbf{y}$  は得られたデータ系列を表している。

この事前分布は前節において示した  $p(\theta) \propto \theta^{-1}$  のような広義事前密度関数が  $\theta$  のべき乗に関して不変であるという性質に注目し、それを Jeffreys が一般化したものである<sup>11</sup>。この事前分布は Jeffreys 型の事前分布といわれるものである。

(2.1) 式で表されるようなモデルにおいて、 $y_0$  が与えられているという条件のもとで、 $\mathbf{y}$  のデータ分布は

$$p(\mathbf{y}|\rho, \sigma, y_0) = (2\pi)^{-T/2} \sigma^{-T} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^T (y_t - \rho y_{t-1})^2\right\} \quad (2.13)$$

---

<sup>11</sup>Zellner [43] 邦訳 52 ページ。

となる。これを用いて事前分布を求める。 $\theta = (\rho, \sigma)'$  とすると、

$$I_{\theta\theta} = \begin{pmatrix} I_{\rho\rho} & 0 \\ 0 & I_{\sigma\sigma} \end{pmatrix} \quad (2.14)$$

となり、(2.13) 式から、

$$I_{\rho\rho} = \begin{cases} \frac{T}{1-\rho^2} - \frac{1}{1-\rho^2} \frac{1-\rho^{2T}}{1-\rho^2} + \left(\frac{y_0}{\sigma}\right)^2 \frac{1-\rho^{2T}}{1-\rho^2} & (\rho \neq 1) \\ \frac{T(T-1)}{2} + T\left(\frac{y_0}{\sigma}\right)^2 & (\rho = 1) \end{cases} \quad (2.15)$$

$$I_{\sigma\sigma} = \frac{2T}{\sigma^2} \quad (2.16)$$

となり、事前分布は

$$p(\rho, \sigma) \propto \frac{1}{\sigma} I_{\rho\rho}^{1/2} \quad (2.17)$$

で与えられる。この事前分布は図 2-6 のような形状となる。ただし、 $y_0 = 0$  とおき、標本数  $T = 100$  および  $T = 200$  の場合について示している。

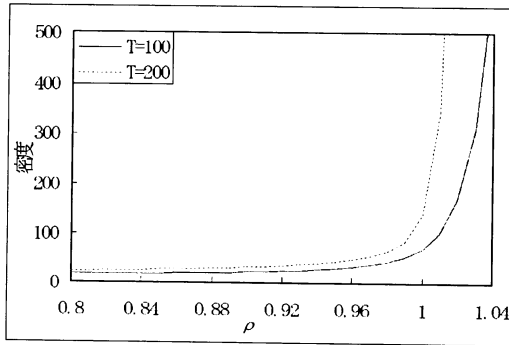


図 2-6 Jeffreys 型の事前分布



この事前分布を用いて  $\rho$  と  $\sigma$  の同時事後分布を計算すると、

$$p(\rho, \sigma | y, y_0) \propto \sigma^{-T-1} \exp\left\{-\frac{1}{2\sigma^2} [m(\hat{u}) + (\rho - \hat{\rho})^2 m(y)]\right\} I_{\rho\rho}^{1/2} \quad (2.18)$$

となる。ここで  $m(y) = \sum_{t=1}^T y_{t-1}^2$ ,  $m(\hat{u}) = \sum_{t=1}^T (y_t - \rho y_{t-1})^2$  である。

この同時事後分布から  $y_0 = 0$  のときの  $\rho$  の周辺事後分布を導出すると、

$$p(\rho | y) \propto I_{\rho\rho}^{1/2} [m(\hat{u}) + (\rho - \hat{\rho})^2 m(y)] \quad (2.19)$$

となる。この  $\rho$  の周辺事後分布と、事前分布を一様分布に設定した場合の  $\rho$  の周辺事後分布を、シミュレーションで作成した  $T = 100$  の2つの系列について表したものが図 2-7 および図 2-8 である。

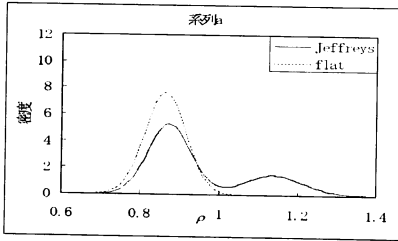


図 2-7  $\rho$  の周辺事後分布 (系列 a)

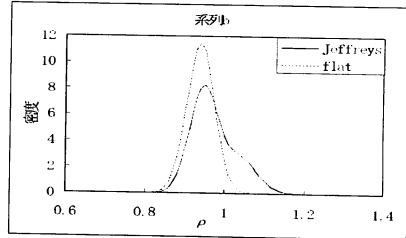


図 2-8  $\rho$  の周辺事後分布 (系列 b)

図 2-7 および図 2-8 で用いた系列の特性値は表 2-1 のようになっている。

	$T$	$\hat{\rho}$	$m(\hat{u})$	$m(y)$
系列 a	100	0.8636	107.364	402.633
系列 b	100	0.9390	1209.667	143.996

この  $\rho$  の事後分布は、事前分布として一様分布を用いた場合と異なり、 $\hat{\rho}$  に関して対称ではなく、また系列 a のようにしばしば  $|\rho| > 1$  のところにも

う1つの山ができる双峰型分布をすることがある。Phillipsはこのもう1つの山について、データの相互依存関係の結果として生じるものであるとみなしている。

また Phillips は事前分布に一樣分布を用いた場合と Jeffreys 型の事前分布を用いた場合についてシミュレーションをおこなっている。その結果から事前分布に一樣分布を用いた場合には事後分布に下方バイアスがあり、それゆえ  $\rho = 1$  となる確率を低く見積もってしまう<sup>12</sup>と主張している。単位根検定の問題に関しては、Jeffreys 型の事前分布を用いた場合には、伝統的アプローチを用いた場合とほぼ同様の結果が得られる<sup>13</sup>と主張している。

### 2.2.3 事前分布に関する論争

Phillips の結果と Sims の結果はともに無知をあらわす事前分布を用いたにもかかわらず大きく異なる。このことから単位根問題におけるベイジアンアプローチでは、無知をあらわす事前分布として、どのような分布を用いるべきかについての論争<sup>14</sup>がおこった。

Koop and Steel [17], Leamer [21], Kim and Maddala [16] などの Phillips [30] に対するコメントでは、Phillips の指摘に一定の評価をしつつ、非定常の部分に多くのウエイトをおく事前分布を合理的なものとはいえないとしている。

Phillips が Jeffreys 型の事前分布を用いて導出した事後分布は、尤度関数自体が時系列データの自己相関構造に関する情報を含んでいるため、この情報を過大視したものである<sup>15</sup>。また、Jeffreys 型の事前分布は非定常の部分に多くのウエイトをおきすぎるため、定常の場合に双峰型の分布を生み出し

---

<sup>12</sup>Phillips [30] p.348.

<sup>13</sup>Phillips [30] p.348.

<sup>14</sup>その論争は *Journal of Applied Econometrics*, 6 や *Econometric Theory*, 10, などに見ることができる。

<sup>15</sup>Kim and Maddala [16] p.376.

てしまう<sup>16</sup>と述べている。

Kim and Maddala [16] は Sims の手法と Phillips の手法をシミュレーションで比較した。その結果から、事前分布に一樣分布を用いた場合には事後分布は下方バイアスを持つが、一方で Jeffreys 型の事前分布を用いた場合には上方バイアスをもち、そのバイアスの大きさは事前分布として一樣分布を用いた場合よりも大きい<sup>17</sup>と主張している。

また、Sims [38] は Phillips [30] の批判に対するコメントにおいて、Phillips の用いた Jeffreys 型の事前分布について、標本数に強く依存し、標本数が大きくなるにつれて非定常なモデルに対するウエイトが非常に高まることを問題としている。このような標本数に依存した事前分布は、主観主義的なベイジアン立場にとっても、データ解析的ベイジアン立場にとっても受け入れられることのできないものであると反論した。

無知をあらわす事前分布としてどのような分布を選ぶべきか、Lubrano [24] や Berger and Yang [2] などにおいて事前分布が考えられた。

Lubrano [24] は  $\rho$  の範囲について、 $|\rho| > 1$  の極端に大きな値をとることはないと考えた。そして小さな正の値  $v$  を用いて  $[-\sqrt{1+v}, \sqrt{1+v}]$  が  $\rho$  のとりうる範囲と考え、有限範囲内の分布としてベータ分布

$$p(\rho) \propto (\sqrt{1+v} + \rho)^{p-1} (\sqrt{1+v} - \rho)^{q-1} \quad (2.20)$$

を考えた。この分布で  $p = q = 1$  のときには一樣分布となる。Lubrano [24] が考えたのは  $p = 1$  で  $q = 0$  の分布であり、

$$p(\rho) \propto (\sqrt{1+v} - \rho)^{-1} \quad (2.21)$$

となる。この分布の形状は Jeffreys 型の事前分布に近い形となり、 $\rho > 1$  のところで密度が無限大に近づくが、 $v$  を適当に選ぶことによってそのスピードと事前分布の範囲をおさえることができる。また標本数  $T$  の関数になって

---

<sup>16</sup>Kim and Maddala [16] p.378.

<sup>17</sup>Kim and Maddala [16] p.379.

いないので標本数が増えるにしたがって、発散のスピードが速くなるということはない。

Berger and Yang [2] は reference prior として

$$p(\rho) = \begin{cases} \frac{1}{2\pi\sqrt{1-\rho^2}} & (|\rho| < 1 \text{ のとき}) \\ \frac{1}{2\pi|\rho|\sqrt{\rho^2-1}} & (|\rho| \geq 1 \text{ のとき}) \end{cases} \quad (2.22)$$

という分布を考えた。彼らの考えた reference prior は AR(1) モデルでは  $|\rho| < 1$  の部分しか求められず、この事前分布は  $|\rho| = 1$  のところで対称となるような分布である。また  $|\rho| = 1$  のときのみ特別に扱っている。

これらの事前分布を用いたときの比較が Bauwens et al. [1] pp. 174 – 193. においてなされている。

#### 2.2.4 ベイジアン単位根検定

ベイジアンアプローチでは母数を確率変数とみなすため、事後分布から  $H_0$  となる確率 ( $p_0$ ) や  $H_1$  となる確率 ( $p_1$ ) を直接求めることができる。この事後確率の比 (事後オッズ比)  $p_0/p_1$  を事前分布において  $H_0$  になる確率 ( $\pi_0$ ) と  $H_1$  になる確率 ( $\pi_1$ ) の比 (事前オッズ比)  $\pi_0/\pi_1$  で割ったベイズファクター

$$B = \frac{p_0/p_1}{\pi_0/\pi_1} = \frac{p_0\pi_1}{p_1\pi_0} \quad (2.23)$$

を用いる方法がベイジアンアプローチの仮説検定の1つである。ベイズファクターは、データによって  $H_0$  となる確率がどの程度増したかどうかを測るものであり、 $B > 1$  かどうかは1つの判断基準となる。また  $\pi_0 = \pi_1$  とすればベイズファクターは事後オッズ比に等しくなる。

さて、(2.1) 式のモデルにおける単位根検定は

$$H_0 : \rho = 1 \quad \text{vs.} \quad H_1 : \rho < 1$$

という検定であり、単純帰無仮説 vs. 複合対立仮説の検定である。このような検定では  $\rho$  の事前分布を

$$p(\rho) = \begin{cases} \pi_0 & (\rho = 1 \text{ のとき}) \\ (1 - \pi_0)g(\rho) & (\rho < 1 \text{ のとき}) \end{cases} \quad (2.24)$$

とすることが考えられる。 $\rho$  の分布が連続型である場合、 $P(\rho = 1) = 0$  である。そのため微小な  $\varepsilon$  を用いて、 $P(\rho \in (1 - \varepsilon, 1 + \varepsilon))$  を考える。この事前分布は、 $P(\rho \in (1 - \varepsilon, 1 + \varepsilon))$  に  $\pi_0$  という確率をおき、残りの確率  $1 - \pi_0$  をある分布  $g(\theta)$  にしたがって割りふるものである。この事前分布を用いたベイズファクターは

$$B = \frac{p(x|\rho = 1)}{\int_{\rho < 1} g(\rho)p(x|\rho)} \quad (2.25)$$

となる。

この仮説検定で問題となるのが、対立仮説のもとでの事前分布  $g(\rho)$  である。Sims [36] が考えた  $-\infty$  から  $\infty$  までの一様分布は、積分しても 1 に等しくならない変則 (improper) 事前分布であり、ベイズファクターを計算することができない。そのため解決策として、ある範囲内での一様分布をとることが考えられる。

Sims [36] は、(2.24) 式のような事前分布で  $g(\rho)$  が尤度の大部分をカバーする一様分布を考えた。そこで問題となるのが一様分布の下限であるが、下限を 0.5 とし、 $\pi_0$  の値を 0.5 とすると、Dickey - Fuller の t 値タイプの検定統計量  $\tau = (1 - \hat{\rho})/\sigma_\rho = 2$ 、(ただし  $\sigma_\rho = \sqrt{\sigma^2/\sum y_{t-1}^2}$ ) と、事後オッズ比が 1 をともに満たす、すなわち伝統的アプローチとベイジアン事後オッズ比によるアプローチが同一の結果となるような  $\hat{\rho}$  の値は  $\hat{\rho} = 0.92$  のときであることを導いた。

Schotman and van Dijk [35] もやはり、事前分布として  $[a, 1)$  の区間に一様分布を用いるもの考えた。ここで特徴的なのは、 $a$  の値はあらかじめ決め

ておくのではなく、事前分布の範囲内に尤度関数の大部分を含むようにデータから分析のたびに求めるということである。

$a$  の値は

$$a^* = \hat{\rho} + s_{\hat{\rho}} F^{-1}(\alpha F(-\hat{\tau})) \quad (2.26)$$

となる  $a^*$  を用いる。ここで  $F(\cdot)$  は t 分布の累積密度関数を表し、自由度は標本数  $T$  のとき  $T - 1$  となる。 $\hat{\tau}$  は DF 検定の統計量  $\hat{\tau} = (\hat{\rho} - 1)/s_{\hat{\rho}}$  である。0.001 から 0.1 までの範囲の  $\alpha$  を入れることにより、尤度関数の  $\rho = 1$  以下の部分のうち、 $(1 - \alpha)$  だけを事前分布がカバーすることになる。

このとき、事後オッズ比は

$$\frac{p_0}{p_1} = \frac{p(\rho = 1|X)}{\frac{1}{1-a} \int_a^1 p(\rho|X) d\rho} \quad (2.27)$$

となる。また、 $p(\rho = 1|X)$  としては、 $\lim_{\rho \rightarrow 1} p(\rho|X)$  を用いている。

単純仮説 vs. 複合仮説の仮説検定では、伝統的アプローチとベイジアンアプローチでその結果が大きく異なることが知られている。伝統的アプローチにおける有意水準 5% の検定で、ちょうど臨界値となるようなデータが得られたとする。ベイジアンアプローチにおいてこのデータを用いた場合、帰無仮説を採択する事後確率は多くの場合 0.05 より大きい。そして標本数が大きくなるほどこの事後確率は大きくなり、伝統的アプローチでは帰無仮説を棄却するようなデータが、ベイジアンアプローチでは帰無仮説を採択するという結果を招くことがある。この問題は、Jeffreys [11] が考え、それを Lindley [23] がパラドックスとして指摘したもので、Lindley のパラドックス、または Jeffreys のパラドックスといわれる。

Schotman and van Dijk [35] はデータから事前分布の下限を求めることによって、Lindley のパラドックスを避けられると考えた。彼らが考えた下限を用いることによって、標本数  $T$  が大きくなるにつれて、事前分布の範囲が小さくなるからである。

まず Lindley のパラドックスが、単位根検定の場合にも起きているかどうかを調べてみる。DF 検定における有意水準  $\alpha$  のときの臨界値となる  $\hat{\rho}$  を求め、対応する事後確率  $p_0$  を計算するのである。図 2-1 で求めた  $\rho$  と  $\hat{\rho}$  の同時分布を尤度とし、(2.24) 式のような事前分布で、 $g(\rho)$  として 0.8 から 0.99 まで 0.1 刻みの離散型の一様分布を考える。そして  $\pi_0 = 0.5$  としたときの事後確率  $p_0$  を計算したものが表 2-2 である。ここでは Fuller [9] p.641 の表から臨界値に対応する  $\hat{\rho}$  の値を求め、最も近い  $\hat{\rho}$  について  $H_0$  が採択される事後確率  $p_0$  を求めた。たとえば  $T = 100, \alpha = 0.05$  のとき、DF 検定の臨界値  $T(\hat{\rho} - 1) = -7.9$  なので、対応する  $\hat{\rho} = -7.9/100 + 1 = 0.921$  なので、最も近い値として  $\hat{\rho} = 0.92$  のときの事後確率  $p_0$  を求めた。

表 2-2 単位根検定における事後確率  $p_0$   
(事前分布  $UD[0.80, 0.99], \pi_0 = 0.5$ )

$\alpha$	$T$			
	50	100	250	500
0.05	0.118	0.121	0.333	0.475
0.1	0.174	0.239	0.483	0.848

この事後確率は、導出の途中で近似的な計算を何度も用いており、精度はかなり低いものと考えられる。しかし、 $T$  が大きくなるにつれて  $p_0$  の値が大きくなるという、Lindley のパラドックスがおこっているということは、この表から見てとることができよう。

次に、Schotman and van Dijk [35] の提唱した事前分布について検証してみる。(2.1) 式における  $\rho$  を 0.9 から 1 まで 0.1 ずつ変えて標本数  $T = 50, 100, 250, 500$  の系列を 100 個ずつ発生させる。これらの系列について推定値  $\hat{\rho}$  と、その標準誤差  $s_{\hat{\rho}}$  を求め、一様分布の下限に  $a^*$  を用いた場合と下限を 0.8 とした場合とした場合のそれぞれの事後オッズ比を計算した。次の表 2-3 は、発生させた 100 個の系列の中で事後オッズ比  $\frac{p_0}{p_1}$  が 1 より大きくなる、すなわち単位根があるという帰無仮説を採択する系列の数である。こ

ここで事前確率  $\pi_0 = 0.5$  としたので、事後オッズ比はベイズファクターに等しい。

表 2-3 一様な事前分布の下限についての比較

$\rho$		$T$			
		50	100	250	500
0.9	下限 0.8	8	4	0	0
	下限 $a^*$	13	0	0	0
0.91	下限 0.8	14	5	0	0
	下限 $a^*$	17	1	0	0
0.92	下限 0.8	20	13	0	0
	下限 $a^*$	28	9	0	0
0.93	下限 0.8	34	18	0	0
	下限 $a^*$	42	4	0	0
0.94	下限 0.8	39	22	0	0
	下限 $a^*$	40	10	0	0
0.95	下限 0.8	42	32	2	0
	下限 $a^*$	44	16	0	0
0.96	下限 0.8	47	42	16	0
	下限 $a^*$	51	26	0	0
0.97	下限 0.8	56	49	25	6
	下限 $a^*$	58	29	9	4
0.98	下限 0.8	60	57	53	28
	下限 $a^*$	64	43	33	4
0.99	下限 0.8	71	77	78	67
	下限 $a^*$	72	56	53	15
1	下限 0.8	71	76	96	93
	下限 $a^*$	73	75	81	72



この結果から標本数  $T$  が小さい場合には、Schotman and van Dijk [35] の事前分布のほうが帰無仮説を採択しやすくなっており、 $T$  が大きくなるにつれて帰無仮説を棄却しやすくなっていることがわかる。すなわち、Schotman and van Dijk [35] の事前分布は標本数  $T$  が大きい場合には有効であるが、 $T = 50$  などの標本数の場合には注意が必要であるといえる。

次に Lindley のパラドックスとの関連を見るために、この結果を  $\hat{\rho}$  の値によってまとめなおしてみる。

表 2-4 採択系列数 ( $T = 50$ )

$\hat{\rho}$	総数	下限 0.8	下限 $a^*$
-0.9	381	0	0
0.91	66	0	0
0.92	54	0	0
0.93	69	0	5
0.94	68	0	35
0.95	74	74	74
0.96	77	77	77
0.97	72	72	72
0.98	68	68	68
0.99	56	56	56
1	47	47	47
1.01	43	43	43
1.02	13	13	13
1.03	7	7	7
1.04	4	4	4
1.05	1	1	1

表 2-5 採択系列数 ( $T = 100$ )

$\hat{\rho}$	総数	下限 0.8	下限 $a^*$
-0.9	291	0	0
0.91	61	0	0
0.92	69	0	0
0.93	84	0	0
0.94	100	0	0
0.95	93	0	0
0.96	104	104	2
0.97	89	89	58
0.98	70	70	70
0.99	68	68	68
1	46	46	46
1.01	21	21	21
1.02	2	2	2
1.03	2	2	2
1.04	0	0	0
1.05	0	0	0

表 2-6 採択系列数 ( $T = 250$ )

$\hat{\rho}$	総数	下限 0.8	下限 $a^*$
-0.96	720	0	0
0.97	110	0	0
0.98	93	93	0
0.99	106	106	105
1	67	67	67
1.01	4	4	4

表 2-7 採択系列数 ( $T = 500$ )

$\hat{\rho}$	総数	下限 0.8	下限 $a^*$
-0.96	712	0	0
0.97	100	0	0
0.98	94	0	0
0.99	113	113	14
1	81	81	81
1.01	0	0	0

$T = 100, \alpha = 0.05$  のときの DF 検定の臨界値は  $\hat{\rho} = 0.92$  であるが、表 2-5 から 2 つの検定ともにすべての系列について帰無仮説を棄却することがわかる。 $T = 500$  のときの有意水準 0.1 の臨界値に対応する  $\hat{\rho}$  の値は 0.99 であるが、この場合は表 2-7 が示すように、Schotman and van Dijk [35] の事前分布のほうが帰無仮説を棄却しやすくなっているということがいえよう。下限 0.8 の一様分布もその範囲が狭いため、Lindley のパラドックスの問題はあまり起きないが、Schotman and van Dijk [35] の事前分布は Lindley のパラドックスを避けることができているといえよう。

この事前分布は Lindley のパラドックスを避けることができ、実用的な分布である。しかしこの事前分布を用いた場合、「同一の尤度から導かれる統計的結論は同じでなくてはならない」という尤度原理を満たさなくなる<sup>18</sup>。それゆえベイジアンアプローチにおける事前分布としては、論理的な整合性を欠くという問題を含んだものといえよう。

最後に、一様な事前分布の上限について考える。Sims [36] や Schotman and van Dijk [35] が用いた事前分布はともに、 $\rho < 1$  の範囲である。これは、与えられた系列が先験的に、定常である ( $\rho < 1$ ) か、階差定常である ( $\rho = 1$ ) かのいずれかであることがわかっているというものである。もし、このような検定を非定常な系列 ( $\rho > 1$ ) であるか、定常または階差定常な系列 ( $\rho \leq 1$ )

<sup>18</sup>Koop and Steel [18] p. 97.

であるか先験的にわからない系列に対してこの検定を用いた場合には、誤った結論を導きかねないであろう。実際に、特に標本数  $T$  が 50 や 100 の場合には、 $\rho > 1$  を仮定しているのに、 $\hat{\rho} < 1$  となる系列が少なからずあり、 $T = 50$  の場合では、 $\rho = 1.01$  とした場合に、 $\hat{\rho} < 1$  となる系列の割合は 50.5%、 $\rho = 1.02$  とした場合で 39.3% となる。よって、このような事前分布が有効となるのは、あらかじめ、 $\rho > 1$  であるか  $\rho \leq 1$  であるかを検証した後のみといえよう。

### 3 定常帰無仮説の単位根検定

#### 3.1 伝統的アプローチ

DF 検定の検出力の低さへの対応は、ノンベジアンの間でもいくつかの検討がなされた。

伝統的アプローチにおける仮説検定では有意水準が5%とか1%の小さな値が用いられ、帰無仮説はそれが成り立たないという強い証拠があるときのみ棄却される。一方、ベイズファクターなどを用いるベジアンアプローチでは、帰無仮説と対立仮説は同等に扱われる。そのような観点から、帰無仮説を単位根モデルとする検定を補うものとして、帰無仮説を定常モデルとする検定が考えられた。その1つに Kwiatkowski et al. [20] の検定<sup>1</sup>がある。KPSS 検定は、1種の状態空間モデルを仮定し、状態方程式の誤差項についてラグランジュ乗数 (LM) 検定をおこなったものであるといえる。

Kwiatkowski et al. [20] では、原系列  $y_t$  を線形トレンド、ランダムウォーク、定常な誤差項の3つに分解することを仮定している。すなわち、

$$y_t = \xi t + \tau_t + e_t \quad (3.1)$$

となり、ここで、 $\tau_t$  は

$$\tau_t = \tau_{t-1} + u_t \quad (3.2)$$

を満たすランダムウォークであり、 $u_t \sim (0, \sigma_u^2), i.i.d.$  である。

このモデルにおいて、 $u_t$  の分散  $\sigma_u^2$  に着目し、 $H_0 : \sigma_u^2 = 0$  vs.  $H_1 : \sigma_u^2 \neq 0$  という検定をおこなうのが、KPSS 検定である。帰無仮説のもとで、 $\tau_t$  は時点  $t$  に依存せず一定となるので、 $y_t$  は線形トレンドまわりの定常な系列とな

<sup>1</sup>以下この検定のことを KPSS 検定ということにする。

る。一方、対立仮説のもとでは、 $\tau_t$  はランダムウォークとなり、 $y_t$  は単位根を含む系列であるとみなされる。

このモデルは状態空間モデルの 1 種とみることができる。状態空間モデルとは、時系列  $\mathbf{Y}_t$  を、状態変数  $\mathbf{X}_t$  の線形関数にノイズを加えた観測方程式と、 $\mathbf{X}_t$  の前時点の状態からの推移にノイズを加えた状態方程式の 2 つの方程式によって表現するモデルであり、ARIMA モデルを含むさまざまな時系列を表現できる。

一般的には

$$\mathbf{Y}_t = \mathbf{G}_t \mathbf{X}_t + \mathbf{W}_t \quad (3.3)$$

$$\mathbf{X}_{t+1} = \mathbf{F}_t \mathbf{X}_t + \mathbf{V}_t \quad (3.4)$$

とあらわされ、(3.3) 式が観測方程式、(3.4) 式が状態方程式である。

Kwiatkowski et al. [20] のモデルにおいて、(3.1) 式を観測方程式、(3.2) 式を状態方程式とみれば状態空間モデルである。

KPSS 検定では Nabeya and Tanaka [27] で考えられた統計量が用いられる。 $y_t$  を定数項と線形トレンドに対して回帰した残差を  $\hat{e}_t$  とする<sup>2</sup>。残差の部分 and

$$S_t = \sum_{i=1}^t \hat{e}_i, \quad t = 1, 2, \dots, T \quad (3.5)$$

とし、 $\hat{\sigma}_e^2$  を誤差項の分散の推定値とすると、LM 統計量は  $e_t \sim N(0, \sigma_e^2), i.i.d.$  という仮定のもとで、

$$LM = \sum_{t=1}^T \frac{S_t^2}{\hat{\sigma}_e^2} \quad (3.6)$$

となる。

---

<sup>2</sup>(3.1) 式で  $\xi = 0$  のとき、 $\hat{e}_t = y_t - \bar{y}$  となる。

しかし、(3.6) 式の  $e_t \sim N(0, \sigma_e^2), i.i.d.$  という仮定は強すぎて、非現実的である。そのため、Kwiatkowski et al. [20] は  $i.i.d.$  が成り立たない場合の誤差項分散の推定値として、

$$s^2(l) = T^{-1} \sum_{t=1}^T e_t^2 + 2T^{-1} \sum_{s=1}^l w(s, l) \sum_{t=s+1}^T e_t e_{t-s} \quad (3.7)$$

を考えた。ここで、 $w(s, l) = 1 - s/(l+1)$  であり、誤差項の異時点間の積にウエイトを与えている。 $l$  はラグを何期まで考えるかについてのパラメータである。

また検定統計量は分子も基準化し、

$$\hat{\eta} = \frac{T^{-2} \sum S_t^2}{s^2(l)} \quad (3.8)$$

というものが用いられている。Kwiatkowski et al. [20] ではこの統計量の臨界値をシミュレーションによって求めた。

ところで、この KPSS 検定は MA 単位根の検定と密接な関係がある。MA 単位根とは差分過剰の状態をあらわすものである。

$$y_t = u_t - \psi u_{t-1}, \quad u_t \sim (0, \sigma_u^2), i.i.d. \quad (3.9)$$

という定数項なしの MA(1) モデルを考えたとき、 $|\psi| = 1$  のとき、MA 単位根となる。 $y_t = \Delta z_t$  であるような系列  $z_t$  を考えると、 $\psi = 1$  のとき、

$$z_t = z_0 + \sum_{i=1}^t y_0 = u_t + (z_0 - u_0) \quad (3.10)$$

となる。 $z_0 = u_0$  を仮定すると、 $z_t$  は定常な誤差項となるので、 $y_t$  はこれをさらに差分をとったものとなり、差分過剰の状態である。

次のような MA(1) モデルを考える。

$$y_t = dt + v_t \quad (3.11)$$

$$\Delta v_t = (1 - \theta L)\epsilon_t \quad (3.12)$$

ここで、 $v_0 = \epsilon_0$  とおき、 $\Delta v_t$  を変形し累積すると、

$$v_t = (1 - \theta) \sum_{s=0}^t \epsilon_s + \theta \epsilon_t \quad (3.13)$$

となる。

一方、(3.2) 式で、 $u_t = (1 - \theta)\epsilon_t$  とおくと、

$$\tau_t = \tau_0 + (1 - \theta) \sum_{s=1}^t \epsilon_s \quad (3.14)$$

となるので、初期値を適当に選ぶことによって (3.1), (3.2) 式は (3.11), (3.12) 式の MA(1) モデルと等しくなる。

(3.13) 式で  $|\theta| < 1$  のとき、 $(1 - \theta) \sum_{s=0}^t \epsilon_s$  は  $I(1)$ <sup>3</sup>、 $\theta \epsilon_t$  は  $I(0)$  なので、 $y_t$  は  $I(1)$  となる。一方  $\theta = 1$  のときは  $v_t = \epsilon_t$  となるので  $y_t$  は  $I(0)$  となる。すなわち MA モデルにおける単位根検定からは、定常モデルを帰無仮説とした仮説検定が自然と導かれる。

## 3.2 ベイジアンアプローチ

KPSS 検定のモデルにベイジアンアプローチを適用したのが Koop and van Dijk [19] である。彼らは、最初に線形トレンドのない次のようなモデルを考えた。

$$y_t = \tau_t + e_t \quad (3.15)$$

$$\tau_t = \tau_{t-1} + u_t \quad (3.16)$$

ここで誤差項には  $e_t \sim N(0, \sigma_e^2), i.i.d.$ 、 $u_t \sim N(0, \sigma_u^2), i.i.d.$  という強い仮定をおいている。

<sup>3</sup> $I(d)$  はその系列の  $d$  回差分をとったものが定常系列となることをあらわす。

系列が単位根を含むかどうかの検定は、 $H_0 : \sigma_u^2 = 0$  vs.  $H_1 : \sigma_u^2 \neq 0$  ということであるが、この検定は

$$\theta = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \quad (3.17)$$

とした場合、 $H_0 : \theta = 0$  vs.  $H_1 : \theta \neq 0$  という検定をおこなうことと同値である。このような変換をほどこすことによって、 $0 \leq \sigma_u^2 < \infty$  のパラメータ区間が、 $0 \leq \theta < 1$  という有限の区間に変換される。

$H_0 : \theta = 0$  の検定は、単純仮説 vs. 複合仮説の検定なので、事前分布を

$$p(\theta) = \begin{cases} \pi_0 & (\theta = 0 \text{ のとき}) \\ (1 - \pi_0)g(\theta) & (\theta \neq 0 \text{ のとき}) \end{cases} \quad (3.18)$$

とおく。

ここで、対立仮説のもとでの事前分布について考えてみる。 $\theta$  と  $\sigma_e^2$  に、無知をあらわす事前分布を設定するならば、Jeffreys の基準により、 $\theta$  に関しては一様分布、 $\sigma_e^2$  に関してはその対数が一様分布となる分布を設定することができる。しかし、これらの分布は積分が 1 とならない improper な事前分布である。そこで、proper な事前分布として  $\sigma_e^2$  の逆数  $h_e$  についてガンマ分布を考える<sup>4</sup>。

$$g(\theta, h_e) \sim f_G(s_e^{-2}, \nu_e) \quad (3.19)$$

ここで、 $s_e^{-2}$  はガンマ分布の平均を、 $\nu_e$  は自由度を表している<sup>5</sup>。これは自然共役事前分布であるが、 $\nu_e$  としてごく小さな値<sup>6</sup>をとると、無知をあらわす事前分布に近くなる。

<sup>4</sup> $\theta$  には一様分布を設定している。

<sup>5</sup>ガンマ分布のこのような表現は Poirier [32] p.100 などに見ることができる。

<sup>6</sup>Koop and van Dijk [19] では  $10^{-300}$  を考えた。



一方、尤度関数について考える。(3.15)(3.16) 式をベクトルで表現すると、

$$\mathbf{y} = \boldsymbol{\tau} + \mathbf{e}, \quad \boldsymbol{\tau} = \mathbf{C}\mathbf{u} \quad (3.20)$$

となる。ただし、

$$\mathbf{y} = (y_1, \dots, y_t)', \quad \boldsymbol{\tau} = (\tau_1, \dots, \tau_t)', \quad \mathbf{e} = (e_1, \dots, e_t)', \quad \mathbf{u} = (u_1, \dots, u_t)'$$

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

である。したがって、

$$\mathbf{y} \sim N(0, \sigma_e^2 \mathbf{V})$$

となる。ただし、 $\mathbf{V} = \mathbf{I}_T + \frac{\theta}{1-\theta} \mathbf{C}\mathbf{C}'$  である。

事前分布と尤度をかけあわせ、局外母数である  $h_e$  を積分することによって、 $\theta$  の事後分布は

$$p(\theta|Data) \propto |\mathbf{V}|^{-1/2} (\mathbf{y}'\mathbf{V}^{-1}\mathbf{y} + \nu_e s_e^2)^{-(T+\nu_e)/2} \quad (3.21)$$

となる。

$\nu_e$  としてごく小さい値をとり、 $s_e^2 = 1$  とすると、この項は無視してもよく、事前分布において  $\pi_0 = 0.5$  としたときの事後オッズ比は

$$\frac{p_0}{p_1} = \frac{(\mathbf{y}'\mathbf{y})^{-T/2}}{|\mathbf{V}|^{-1/2} (\mathbf{y}'\mathbf{V}^{-1}\mathbf{y})^{-T/2}} \quad (3.22)$$

となる。これはベイズファクターに等しくなる。

### 3.2.1 シミュレーションによる検証

Koop and van Dijk [19] の考えたベイジアンアプローチを検証するために、シミュレーション実験をおこなった。

(3.15)(3.16) 式において  $e_t$  と  $u_t$  について正規乱数を発生し、 $\tau_0 = 0$  として、 $T = 100$  個の  $y_t$  の系列をつくった。ここで、 $e_t$  の分散  $\sigma_e^2$  は 1 とし、 $u_t$  の分散  $\sigma_u^2$  は与えた  $\theta$  から  $\sigma_u^2 = \frac{\theta}{1-\theta}$  として導出した。

まず系列  $y_t$  を、データ発生プロセス (DGP) の  $\theta$  の値を変えて 3 つ発生し、それぞれの系列の周辺事後分布の形状を調べてみる。

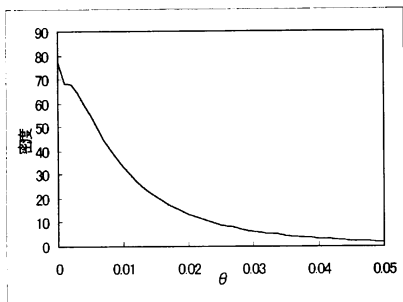


図 3-1  $\theta = 0$  の周辺事後分布

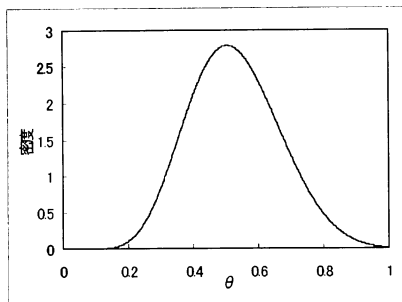


図 3-2  $\theta = 0.5$  の周辺事後分布

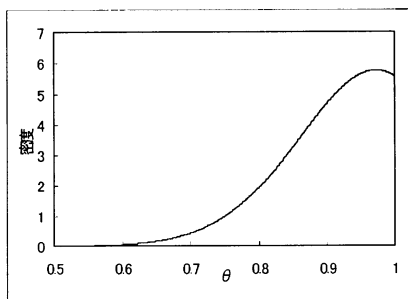


図 3-3  $\theta = 1$  の周辺事後分布

図 3-1 は  $\theta = 0$  のときの周辺事後分布であり、0 のごく近くの範囲の密度が非常に高くなり、他の部分の密度は非常に低い。このデータでは  $B = 83.56$  となり、帰無仮説が採択され、 $I(0)$  となる。

図 3-2 は  $\theta = 0.5$  すなわち  $\sigma_u^2 = \sigma_e^2$  のときの周辺事後分布である。0.5 の近くを山とするほぼ左右対称の形状をしている。またこのデータでは  $B = 4.2 \times 10^{-27}$  となり、対立仮説が採択され、 $I(1)$  となる。

図 3-3 は  $\theta = 1$  すなわち  $\sigma_e^2 = 0$  のときの周辺事後分布である。ここで  $\sigma_u^2 = 1$  とした。この場合、1 の近くにかかなり大きな山ができる。またこのデータでは  $B = 2.4 \times 10^{-111}$  となり、対立仮説が採択され、 $I(1)$  となる。

次に、DGP の  $\theta$  を 0 から 0.1 まで 0.01 刻みで変化させ、個々の  $\theta$  について  $T = 100$  の系列を 1000 個つくり、そのうち  $B > 1$  となる系列がいくつあるかを数えてみた。

表 3-1  $H_0$  を採択する系列の数 (1000 系列中)

$\theta$	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
系列数	999	332	168	125	62	43	31	23	19	11	12

$\theta$  が少し変化しただけで、採択される系列が急激に減っていくことがわかる。この結果からは DF 検定でおこる、帰無仮説が棄却されにくいという問題は、ある程度回避できているということが想像できる。さらに KPSS 検定や DF 検定などの、伝統的アプローチをおこなった場合との結果の比較を試みる。

まず、DGP の  $\theta = 0, 0.01, 0.05, 0.1, 0.5, 0.99$  について、それぞれ  $T = 100$  の系列を 100 個つくり、個々の系列にベイズファクターによる検定、KPSS 検定、DF 検定をおこなった。ここで、

1.  $B > 1$  なら  $I(0)$
2.  $\hat{\eta} < 0.463$  なら  $I(0)$
3.  $\hat{\rho} > 0.921$  なら  $I(1)$

とした。KPSS 検定は Kwiatkowski et al. [20] p.166 から、DF 検定は Fuller [8] p.641 からそれぞれ有意水準 5% の片側検定の臨界値を用いた。

表 3-2 において、左側は DF 検定で  $I(0)$  とみなされたものであり、右側は  $I(1)$  とみなされたものである。それぞれについて、ベイズファクターによる検定と KPSS 検定の結果をクロス集計している。

表 3-2 検定どうしの関係 (1)

		I(0) (DF 検定)			I(1) (DF 検定)					
$\theta = 0$			<i>B</i>					<i>B</i>		
			I(0)	I(1)	計			I(0)	I(1)	計
	$\hat{\eta}$	I(0)	97	0	97	$\hat{\eta}$	I(0)	0	0	0
		I(1)	2	1	3		I(1)	0	0	0
計		99	1	100	計		0	0	0	
$\theta = 0.01$			<i>B</i>					<i>B</i>		
			I(0)	I(1)	計			I(0)	I(1)	計
	$\hat{\eta}$	I(0)	22	19	41	$\hat{\eta}$	I(0)	0	0	0
		I(1)	9	50	59		I(1)	0	0	0
計		31	69	100	計		0	0	0	
$\theta = 0.05$			<i>B</i>					<i>B</i>		
			I(0)	I(1)	計			I(0)	I(1)	計
	$\hat{\eta}$	I(0)	5	7	12	$\hat{\eta}$	I(0)	0	0	0
		I(1)	1	84	85		I(1)	0	3	3
計		6	91	97	計		0	3	3	
$\theta = 0.1$			<i>B</i>					<i>B</i>		
			I(0)	I(1)	計			I(0)	I(1)	計
	$\hat{\eta}$	I(0)	1	0	1	$\hat{\eta}$	I(0)	0	0	0
		I(1)	0	85	85		I(1)	0	14	14
計		1	85	86	計		0	14	14	
$\theta = 0.5$			<i>B</i>					<i>B</i>		
			I(0)	I(1)	計			I(0)	I(1)	計
	$\hat{\eta}$	I(0)	0	0	0	$\hat{\eta}$	I(0)	0	0	0
		I(1)	0	24	24		I(1)	0	76	76
計		0	24	24	計		0	76	76	
$\theta = 0.99$			<i>B</i>					<i>B</i>		
			I(0)	I(1)	計			I(0)	I(1)	計
	$\hat{\eta}$	I(0)	0	0	0	$\hat{\eta}$	I(0)	0	0	0
		I(1)	0	5	5		I(1)	0	95	95
計		0	5	5	計		0	95	95	

この結果から、まず DF 検定について、データに新たなノイズが加わった場合には、結果の信頼性が低いということがいえよう。 $\theta = 0.5$  のとき、 $\sigma_e^2 = \sigma_u^2$  であるので、 $y_t$  はランダムウォークに、その誤差項と等分散のホワイトノイズを加えたものとみなせる。よって、その程度のノイズで検定のサイズが変わってしまうことが確認できる。

単位根を含むモデルを2つの検定では、ベイズファクターを用いたもののほうが検出力が高いことがわかる。これは、伝統的アプローチでは帰無仮説を否定する強い証拠がないと対立仮説が採択されないのに対して、ベイジアンアプローチでは、帰無仮説と対立仮説を同等に扱うということにもよる。しかし、ベイズファクターでは棄却されるのに、伝統的アプローチで棄却されない系列もある。そこで、 $\theta = 0.01$  の4系列について、 $y_t$  の折れ線グラフとそれぞれの系列の周辺事後分布を次に示す。

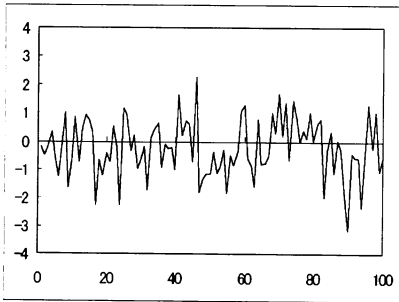


図 3-4  $B : I(0), \hat{\eta} : I(0)$  の系列

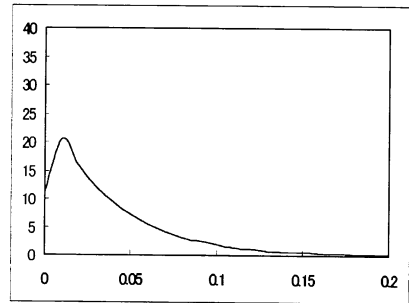


図 3-5 左の系列の周辺事後分布

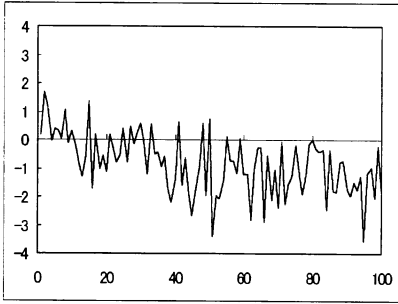


図 3-6  $B : I(1), \hat{\eta} : I(1)$  の系列

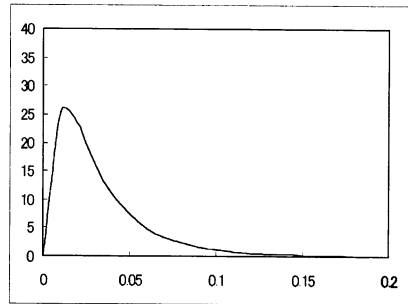


図 3-7 左の系列の周辺事後分布

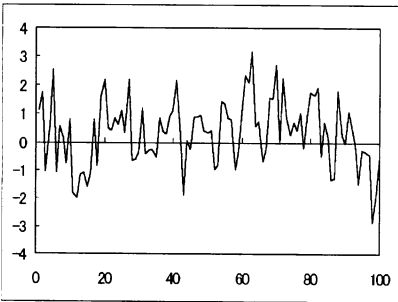


図 3-8  $B : I(1), \hat{\eta} : I(0)$  の系列

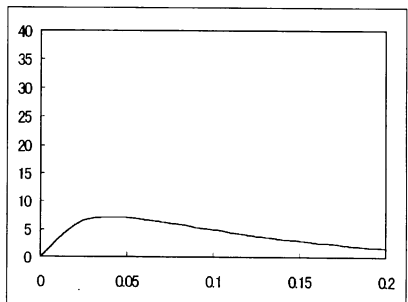


図 3-9 左の系列の周辺事後分布

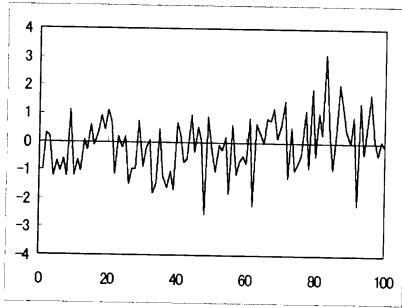


図 3-10  $B : I(0), \hat{\eta} : I(1)$  の系列

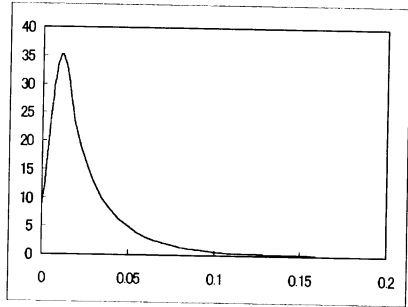


図 3-11 左の系列の周辺事後分布

この4つの系列をみると、 $y_t$  のグラフからは図 3-6 が非定常であることは明らかにわかる。しかし、他の系列については  $y_t$  のグラフからは定常か非定常かの判断はできない。

周辺事後分布をみると、ベイズファクターによって  $I(0)$  となる系列は  $\theta = 0$  に高い密度を持つ。一方、KPSS 検定では 0 の周辺の密度が高くなる場合には  $I(1)$  となる。図 3-9 と図 3-11 が対照的であり、図 3-9 は  $\theta = 0$  およびその近くの密度が低いのに対し、図 3-11 は  $\theta = 0$  およびその近くで密度が高くなる。前者はベイズファクターで  $I(1)$ 、KPSS 検定で  $I(0)$  となるのに対し、後者はベイズファクターで  $I(0)$ 、KPSS 検定で  $I(1)$  となる。

次に DGP として  $y_t = \rho y_{t-1} + e_t$  という AR(1) モデルを考える。そこで  $\rho = 1, 0.85, 0.5$  について  $T = 100$  の系列を 100 個つくり、3 種類の検定をおこなった。

このモデルは、(3.15)(3.16) 式のモデルに一致するものではないが、モデルを誤って定式化した場合として考察してみる。



表 3-3 検定どうしの関係 (2)

		I(0) (DF 検定)			I(1) (DF 検定)				
		B			B				
		I(0)	I(1)	計	I(0)	I(1)	計		
$\rho = 1$	$\hat{\eta}$	I(0)	0	0	0	I(0)	0	0	0
		I(1)	0	2	2	I(1)	0	98	98
		計	0	2	2	計	0	98	98
$\rho = 0.85$	$\hat{\eta}$	I(0)	0	15	15	I(0)	0	0	0
		I(1)	0	83	83	I(1)	0	2	2
		計	0	98	98	計	0	2	2
$\rho = 0.5$	$\hat{\eta}$	I(0)	18	43	61	I(0)	0	0	0
		I(1)	5	34	39	I(1)	0	0	0
		計	23	77	100	計	0	0	0

この結果をみると、 $\rho = 1$  のときには、ベイズファクターによる検定と KPSS 検定は、ともにすべての系列を I(1) と結論づけており、DF 検定より正確な結論を導いている。

一方、 $\rho = 0.85, 0.5$  の定常な AR(1) モデルの場合は、ベイズファクターによる検定と KPSS 検定はともに、誤って I(1) という結論を出すことが多い。特にベイズファクターではかなりの問題点がある。ベイズファクターによる検定では、モデルの定式化を誤った場合には、誤った結論を導く可能性がかなり高いということが示唆されよう。

## 4 結合の動き

単位根を含むモデルを帰無仮説とする仮説検定と、定常モデルを帰無仮説とする仮説検定ではその結果が大きく異なってしまう。

検定の利用者の立場では、結果が大きく異なることは非常に不便なことであるといえよう。そのため最近、このような2種類の検定を結合する動きがある。

### 4.1 伝統的アプローチ

伝統的アプローチにおいて2種類の検定を結合する試みとして、Charemza and Syczewska [4], Carrion et al. [3] などがある。

Charemza and Syczewska [4] は定常な AR(1) モデル

$$y_t = \beta + \rho y_{t-1} + \varepsilon_t, \quad |\rho| < 1 \quad (4.1)$$

について、DF 検定の統計量と KPSS 検定の統計量の同時分布を考えた。この同時分布から DF 検定と KPSS 検定でともに  $I(0)$  となる確率 (Probability of Joint Confirmation: PJC) を求めることができる。ある PJC (たとえば 0.95) を満たすような検定統計量の組み合わせ  $\{z_D^{PJC}, z_K^{PJC}\}$  は数多く存在するが、その中でそれぞれの統計量の周辺分布において  $I(0)$  となる確率 (Marginal PProbabilities: MBR) が等しくなる組み合わせをシミュレーションによって求めた。

Carrion et al. [3] は、Charemza and Syczewska [4] のシミュレーションを  $\rho = 1$  の場合についておこなった。 $T = 50, 100, 150, 300$  の系列についての臨界値を求めたのである。これはトレンド、定数項の有無によって3種類に分

けられる。また DF 検定と KPSS 検定の組み合わせだけでなく、PP 検定と KPSS 検定の組み合わせについても臨界値を求めた<sup>1</sup>。

Carrion et al. [3] はさらに

$$y_t = f(t) + \rho y_{t-1} + u_t \quad (4.2)$$

$$u_t = \varepsilon_t + \theta \varepsilon_{t-1} \quad (4.3)$$

という ARMA(1, 1) モデルに  $\rho = \{0.5, 0.85, 0.9, 1\}$ ,  $\theta = \{-0.8, -0.5, 0\}$  という値の組み合わせを考え、臨界値のパフォーマンスをチェックした。 $\rho \neq 1$  のときには  $f(t) = \{0, 1, 1 + 0.2t\}$  とし、 $\rho = 1$  のときには  $f(t) = \{0, 0, 0.2\}$  という値をとることによって、定数項なし、トレンドなし、トレンドありの 3 種類のモデルをつくった。

シミュレーションの結果では  $\rho$  が定常のときには、DF 検定と KPSS 検定をそれぞれおこなった場合の結果が最も良いが、 $\rho = 1$  のときは Carrion et al. [3] が考えた、単位根を帰無仮説として同時に検定する方法が良いということが示された。

## 4.2 ベイジアンアプローチ

ベイジアンアプローチでは、Koop and van Dijk [19] が考えた (3.15)(3.16) に AR 項を加えた次のようなモデルが 2 種類の検定を結合するものといえよう。

$$y_t = \rho y_{t-1} + \tau_t + e_t \quad (4.4)$$

$$\tau_t = \tau_{t-1} + u_t \quad (4.5)$$

---

<sup>1</sup>これらの結果は Carrion et al. [3] pp. 4 - 5. の表にまとめられている。

このモデルを変形すると

$$y_t = \rho y_{t-1} + \sum_{i=1}^t u_i + e_t \quad (4.6)$$

となる。

このモデルは  $\rho$  と  $\theta$  の値によって次のように分類される。

1. ( $\rho = 1 \cap \theta = 0$ ) のとき I(2)
2. ( $\rho = 1 \cap \theta \neq 0$ ) または ( $\rho \neq 1 \cap \theta = 0$ ) のとき I(1)
3. ( $\rho \neq 1 \cap \theta \neq 0$ ) のとき I(0)

このモデルについて  $\rho$  と  $\theta$  としていくつかの値を選んでデータを発生し、KPSS 検定とベイズファクターによる検定をおこなってみた。その結果をまとめたのが表 4-1 である。

ここでのベイズファクターは  $H_0 : \theta = 0$  vs.  $H_1 : \theta \neq 0$  を検定するもので、

$$B = \frac{(\mathbf{y}'_{-1} \mathbf{y}_{-1})^{-1/2} (\mathbf{y}' \mathbf{M} \mathbf{y})^{-(T-1)/2}}{\int_{\theta \neq 0} |\mathbf{V}|^{-1/2} (\mathbf{y}'_{-1} \mathbf{V}^{-1} \mathbf{y}_{-1})^{-1/2} (s^2)^{-(T-1)/2} d\theta} \quad (4.7)$$

である。ここで、

$$\begin{aligned} \mathbf{M} &= \mathbf{I} - \mathbf{y}'_{-1} (\mathbf{y}'_{-1} \mathbf{y}_{-1})^{-1} \mathbf{y}_{-1} \\ s^2 &= (\mathbf{y} - \hat{\rho} \mathbf{y}_{-1})' \mathbf{V}^{-1} (\mathbf{y} - \hat{\rho} \mathbf{y}_{-1}) \\ \hat{\rho} &= (\mathbf{y}'_{-1} \mathbf{V}^{-1} \mathbf{y}_{-1})^{-1} \mathbf{y}'_{-1} \mathbf{V}^{-1} \mathbf{y} \end{aligned}$$

である。

表 4.1 検定どうしの関係

	$\rho = 0$		$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.85$		$\rho = 1$											
	$B$	$I(0) I(1)$ 計	$B$	$I(0) I(1)$ 計	$B$	$I(0) I(1)$ 計	$B$	$I(0) I(1)$ 計	$B$	$I(0) I(1)$ 計										
$\theta = 0$	$I(0)$	97	0	97	$I(0)$	79	0	79	$I(0)$	63	0	63	$I(0)$	10	0	10	$I(0)$	0	0	0
	$\hat{\eta} I(1)$	2	1	3	$\hat{\eta} I(1)$	21	0	21	$\hat{\eta} I(1)$	37	0	37	$\hat{\eta} I(1)$	82	8	90	$\hat{\eta} I(1)$	11	89	100
	計	99	1	100	計	100	0	100	計	100	0	100	計	92	8	100	計	11	89	100
$\theta = 0.01$	$I(0)$	22	19	41	$I(0)$	15	8	23	$I(0)$	15	3	18	$I(0)$	2	1	3	$I(0)$	0	0	0
	$\hat{\eta} I(1)$	9	50	59	$\hat{\eta} I(1)$	20	57	77	$\hat{\eta} I(1)$	35	47	82	$\hat{\eta} I(1)$	25	72	97	$\hat{\eta} I(1)$	18	82	100
	計	31	69	100	計	35	65	100	計	50	50	100	計	27	73	100	計	18	82	100
$\theta = 0.05$	$I(0)$	5	7	12	$I(0)$	2	4	6	$I(0)$	3	2	5	$I(0)$	0	0	0	$I(0)$	0	0	0
	$\hat{\eta} I(1)$	1	87	88	$\hat{\eta} I(1)$	7	87	94	$\hat{\eta} I(1)$	21	74	95	$\hat{\eta} I(1)$	10	90	100	$\hat{\eta} I(1)$	4	96	100
	計	6	94	100	計	9	91	100	計	24	76	100	計	10	90	100	計	4	96	100
$\theta = 0.5$	$I(0)$	0	0	0	$I(0)$	0	0	0	$I(0)$	0	0	0	$I(0)$	0	0	0	$I(0)$	0	0	0
	$\hat{\eta} I(1)$	0	100	100	$\hat{\eta} I(1)$	4	96	100	$\hat{\eta} I(1)$	5	95	100	$\hat{\eta} I(1)$	0	100	100	$\hat{\eta} I(1)$	0	100	100
	計	0	100	100	計	4	96	100	計	5	95	100	計	0	100	100	計	0	100	100
$\theta = 0.99$	$I(0)$	0	0	0	$I(0)$	0	0	0	$I(0)$	0	0	0	$I(0)$	0	0	0	$I(0)$	0	0	0
	$\hat{\eta} I(1)$	0	100	100	$\hat{\eta} I(1)$	2	98	100	$\hat{\eta} I(1)$	0	100	100	$\hat{\eta} I(1)$	0	100	100	$\hat{\eta} I(1)$	0	100	100
	計	0	100	100	計	2	98	100	計	0	100	100	計	0	100	100	計	0	100	100

表 4-1 を見ると、KPSS 検定では  $\rho$  の値、 $\theta$  の値が増加するにつれ、I(1) と判断される系列が単調に増えていくのがわかる。一方、ベイズファクターは  $\theta = 0.01$  のとき、 $\rho$  が 0 から増加するにつれて、いったん I(0) と判断される系列が増加し、 $\rho = 0.5$  を境に I(1) と判断される系列が増加する。この傾向は  $\theta = 0.05$  のときにもみられ、非常に特徴的な変動をしているといえよう。これは定常な AR 項がベイズファクターによる検定に何らかの影響を与えているものと考えられる。

一方、 $\theta = 0$  のときは KPSS 検定では  $\rho$  の値が大きくなるにつれ I(1) と判断される系列が増えるのに対し、ベイズファクターの検定では正しい判断を下せていることがわかる。

この  $\theta$  と  $\rho$  の相互関係については今後もう少し検討していく必要があるだろう。

## おわりに

本書では単位根検定の展開について、ベイジアンアプローチを中心にまとめてみた。

このような議論の出発点となったのは、DF 検定の検出力の低さであり、そのためベイジアンアプローチや定常帰無仮説の検定などの解決するための試みがなされた。

しかし定常帰無仮説の検定は、DF 検定をはじめとする単位根を帰無仮説とする検定に取って代わるものではない。この検定はあくまで補助的に使われるものであって、系列が単位根を含むかどうかの問題について確証を与えてくれる<sup>2</sup>ものである。そのため、この2種類の検定を結合する動きは非常に注目すべきものであるといえよう。

ベイジアンアプローチについても同様のことがいえる。ベイジアンアプローチの特徴として、帰無仮説と対立仮説を同等に扱えるということがある。そうであれば、単位根を帰無仮説とする検定と定常帰無仮説の検定では等しい結果を導かなくてはならない。しかし、そのためにはどのようなモデルを設定するかが問題となってくる。その意味では Koop and van Dijk [19] が考えた (4.4)(4.5) のモデルはより深い検討が必要となるであろう。この問題については今後の研究課題としていきたい。

---

<sup>2</sup>Maddala and Kim [25] p. 126.

## 参考文献

- [1] Bauwens, L., M. Lubrano and J. F. Richard (1999), *Bayesian Inference in Dynamic Economic Models*, Oxford University Press.
- [2] Berger, J. O. and R. Y. Yang (1994), “Noninformative priors and the AR(1) model,” *Econometric Theory*, **10**, 461 – 482.
- [3] Carrion, J. L. I., A. I. Sansó and M. A. Ortuño, (2001), “Unit root and stationary tests’ wedding,” *Economics Letters*, **70**, 1 – 8.
- [4] Charemza, W. W., E. M. Syczewska, (1998), “Joint application of Dickey-Fuller and KPSS tests,” *Economics Letters*, **61**, 17 – 21.
- [5] Dickey, D. A. and W. A. Fuller, (1976), “Distribution of first order autoregressive estimators,” *1976 Proceedings of the Business and Economic Statistics Section, American Statistical Association*, 278 – 281.
- [6] Dickey, D. A. and W. A. Fuller, (1979), “Distribution of estimators for autoregressive time series with a unit root,” *Journal of the American Statistical Association*, **74**, 427 – 431.
- [7] Dickey, D. A. and S. E. Said, (1981), “Testing ARIMA(p, 1, q) versus ARMA(p + 1, q),” *1981 Proceedings of the Business and Economic Statistics Section, American Statistical Association*, 318 – 322.
- [8] Fuller, W. A., (1976), *Introduction to Statistical Time Series*, John Wiley & Sons.



- [9] Fuller, W. A., (1996), *Introduction to Statistical Time Series*, 2nd edition, John Wiley & Sons.
- [10] Gabriel, V. J., (2003), “Cointegration and joint confirmation hypothesis,” *Economics Letters*, **78**, 17 – 25.
- [11] Jeffreys, H., (1961), *Theory of Probability*, 3rd edition, Oxford University Press.
- [12] 河田 正樹 (1996) 「単位根問題におけるベイジアンアプローチ」 『早稲田経済学研究』 **44**, 135 – 143.
- [13] 河田 正樹 (1999) 「ベイジアンアプローチによる単位根検定」 『徳山大学論叢』 **52**, 1 – 18.
- [14] 河田 正樹 (2001) 「状態空間モデルにおける単位根検定 – ベイジアン視点から –」 『徳山大学論叢』 **55・56**, 1 – 15.
- [15] 河田 正樹 (2003) 「AR 単位根と MA 単位根」 『徳山大学論叢』 **59**, 71 – 85.
- [16] Kim I. M. and G. S. Maddala (1991), “Flat priors vs. ignorance priors in the analysis of the AR(1) model,” *Journal of Applied Econometrics*, **6**, 375 – 380.
- [17] Koop, G. and M. F. Steel, (1991), “A comment on: ‘To criticize the critics: an objective Bayesian analysis of stochastic trends’, by Peter C. B. Phillips,” *Journal of Applied Econometrics*, **6**, 365 – 370.
- [18] Koop, G. and M. F. J. Steel (1994) “A decision-theoretic analysis of unit-root hypothesis using mixtures of elliptical models,” *Journal of Business & Economic Statistics*, **12**, 95 – 107.

- [19] Koop, G. and H. K. van Dijk, (2000), “Testing for integration using evolving trend and seasonal models: A Bayesian approach,” *Journal of Econometrics*, **97**, 261 – 291.
- [20] Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin, (1992), “Testing the null hypothesis of stationarity against the alternative of a unit root,” *Journal of Econometrics*, **54**, 159 – 178.
- [21] Leamer, E. E., (1991), “Comment on ‘To criticize the critics’,” *Journal of Applied Econometrics*, **6**, 371 – 373.
- [22] Lee, P. M. (1997) *Bayesian Statistics: An Introduction*, 2nd edition, Arnold.
- [23] Lindley, D. V. (1957) “A statistical paradox,” *Biometrika*, **44**, 187 – 192.
- [24] Lubrano, M. (1995), “Testing for unit roots in a Bayesian Framework,” *Journal of Econometrics*, **69**, 81 – 109.
- [25] Maddala, G. S. and I. M. Kim (1998), *Unit Roots, Cointegration and Structural Change*, Cambridge University Press.
- [26] 森棟 公夫 (1999) 「単位根時系列分析」 『日本統計学会誌』 **29**, 307 – 326.
- [27] Nabeya, S. and K. Tanaka, (1988), “Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative,” *The Annals of Statistics*, **16**, 218 – 235.
- [28] Nelson, C. R. and C. I. Plosser, (1982), “Trends and random walks in macroeconomic time series,” *Journal of Monetary Economics*, **10**, 139 – 162.

- [29] Phillips, P. C. B. (1987), “Time series regression with a unit root,” *Econometrica*, **55**, 277 – 301.
- [30] Phillips, P. C. B. (1991), “To criticize the critics: an objective Bayesian analysis of stochastic trends,” *Journal of Applied Econometrics*, **6**, 333 – 364.
- [31] Phillips, P. C. B. and P. Perron, (1988), “Testing for unit root in Time series regression,” *Biometrika*, **75**, 335–346.
- [32] Poirier, D. J., (1995), *Intermediate Statistics and Econometrics*, The MIT Press.
- [33] 繁桝 算男 (1985) 『ベイズ統計入門』 東京大学出版会.
- [34] Said, S. E. and D. A. Dickey (1984), “Testing for unit roots in autoregressive - moving average models of unknown order,” *Biometrika*, **71**, 599 – 607.
- [35] Schotman, P. C. and H. K. van Dijk (1991), “A unit root in real exchange rates,” *Journal of Econometrics*, **49**, 195 – 238.
- [36] Sims, C. A., (1988), “Bayesian skepticism on unit root econometrics,” *Journal of Economic Dynamics and Control*, **12**, 463 – 474.
- [37] Sims, C. A. and H. Uhlig, (1991), “Understanding unit rooter: a helicopter tour,” *Econometrica*, **59**, 1591 – 1599.
- [38] Sims, C. A. (1991), “Comment by Christopher A. Sims on ‘To criticize the critics’, by Peter C. B. Phillips,” *Journal of Applied Econometrics*, **6**, 423 – 434.

- [39] Stock, J. H., (1994), “Unit roots, structural breaks and trends,” in Engle, R. F. and D. L. McFadden, (eds.) *Handbook of Econometrics*, vol. 4, North - Holland, 2739 – 2841.
- [40] Tanaka, K., (1996), *Time Series Analysis*, John Wiley & Sons.
- [41] 和合 肇 (1998) 「ベイズ計量経済分析における最近の発展」 『日本統計学会誌』 **28**, 253 – 305.
- [42] 山本 拓 (1988) 『経済の時系列分析』 創文社.
- [43] Zellner, A., (1971), *An Introduction to Bayesian Inference in Econometrics*, John Wiley & Sons. (福場 庸・大澤 豊 共訳 『ベイジアン計量経済学入門』 培風館, 1986)

## 著者略歴

かわだ まさき  
河田 正 樹

- 1971 年 東京都世田谷区生まれ  
1994 年 早稲田大学政治経済学部経済学科卒業  
1999 年 早稲田大学大学院経済学研究科博士後期課程単位取得  
同 年 徳山大学経済学部専任講師  
現 在 徳山大学経済学部助教授

## ベイジアン単位根検定の展開

徳山大学総合経済研究所モノグラフ 11

---

---

2004 年 3 月 17 日印刷

2004 年 3 月 25 日発行

著 者 河田 正樹

©Kawada Masaki 2004

発 行 徳山大学総合経済研究所

〒 745-8566

山口県周南市久米栗ヶ追 843-4-2

印刷所 中西印刷株式会社

〒 602-8048

京都府京都市上京区下立売小川東入る

---

---