



ダブル・ブートストラップ法およびファスト・ダブル・ブートストラップ法による検定のシミュレーション比較（〈特集〉計量分析の理論と応用）

難波, 明生

(Citation)

国民経済雑誌, 209(1):67-79

(Issue Date)

2014-01

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/81008953>

(URL)

<https://hdl.handle.net/20.500.14094/81008953>



ダブル・ブートストラップ法および
ファスト・ダブル・ブートストラップ法による
検定のシミュレーション比較

難 波 明 生

国民経済雑誌 第 209 巻 第 1 号 抜刷

平成 26 年 1 月

ダブル・ブートストラップ法および ファスト・ダブル・ブートストラップ法による 検定のシミュレーション比較

難 波 明 生

現代の計量経済学・統計学において、ブートストラップ法は非常に重要な手法の一つである。ブートストラップ法を用いることによって、従来の漸近理論に基づく推定・検定の精度を改善することができる可能性があることがよく知られている。本稿では、Beran (1988) によるダブル・ブートストラップ法および Davidson and MacKinnon (2002) によるファスト・ダブル・ブートストラップ法を平均に関する検定に応用し、その特性をシミュレーションにより分析する。シミュレーションの結果から、標本がある程度大きい場合、ダブル・ブートストラップ法を行うことにより、漸近理論や通常のブートストラップ法に基づく検定よりも精度が高く、検定力のパフォーマンスも良好な検定を行える可能性があることが示される。この結果は、標本がある程度大きく、必要とされる計算時間が現実的な場合における、ダブル・ブートストラップ法の有効性を示唆するものである。

キーワード ブートストラップ法, ダブル・ブートストラップ法,
ファスト・ダブル・ブートストラップ法

1 は じ め に

現代の計量経済学・統計学において、ブートストラップ (bootstrap) 法は非常に重要な手法の一つである。計量経済学・統計学においては様々な推定法および検定法が提案されているが、その多くについては漸近的な性質しか知られていない。したがって、実際の分析では、漸近的な近似が行えるほど大きな標本ではないのにもかかわらず、漸近的な性質に基づいて、近似的な推定・検定が行われている場合が少なくない。このような状況においては、得られた推定量・検定統計量の小標本における分布が漸近分布から大きくかけ離れてしまい、望まれる精度の推定・検定が行うことができない可能性がある。このような場合に、統計量の分布を、従来の漸近理論よりも正確に近似することができる可能性がある方法として注目されているのが、Efron (1979) により提案されたブートストラップ法である。

通常の推定・検定が、統計量の分布表を用いて行われるのに対して、ブートストラップ法

では、手元にある標本から無作為標本を再度抽出（リサンプリング）し、得られた標本（ブートストラップ標本）を用いて計算した統計量の分布を用いて統計量の分布を近似する。ブートストラップ標本を用いた統計量の分布は、リサンプリングを何度も繰り返し、得られた統計量の経験分布を用いて求めることになる。したがって、ブートストラップ法においては、リサンプリングによって得られる標本を用いた統計量の反復計算が必要となる。

ブートストラップ法による分布の近似が、従来の漸近理論による近似よりも正確になるための重要な条件が、統計量が漸近的にピボット（pivot）であるということである。ピボットとは、統計量の分布が未知パラメータに依存しないということを意味する。漸近的にピボットである統計量として特に良く用いられるのが、 t 統計量である。 t 統計量の分布は、標本が大きくなるにしたがって標準正規分布に近づいていくので、 t 統計量が漸近的にピボットであることは容易に確認できる。Beran (1987, 1988) は、統計量が漸近的にピボットである場合に、ブートストラップ法を用いることにより、漸近理論に基づく信頼区間や検定の精度を改善できる可能性があることを示している¹⁾。

Beran (1988) は、通常のブートストラップ法において得られたブートストラップ標本から再度リサンプリングを行い、検定を行う方法を提案した。この方法では、 P 値（ p -value）が検定統計量として用いられることになるが、任意の連続型検定統計量の P 値は、帰無仮説の下で $(0, 1)$ 上の一様分布に従うので、 P 値はピボットであるといえる。この方法は、リサンプリングが二段階に行われるので、ダブル・ブートストラップ（double bootstrap）法と呼ばれている。ダブル・ブートストラップ法においては、第一段階のリサンプリングを B 回、第二段階のリサンプリングを C 回行うとすると、全体で $B \times C$ 回のリサンプリングを行わなければならないので、非常に多くの反復計算が必要ということになる。

Davidson and MacKinnon (2002) は、ダブル・ブートストラップ法における反復計算を軽減する方法として、ファスト・ダブル・ブートストラップ（fast double bootstrap）法を提案した。ダブル・ブートストラップ法においては、第二段階のリサンプリングが C 回行われるのに対して（ C は十分に大きな自然数）、ファスト・ダブル・ブートストラップ法では、第二段階のリサンプリングは 1 回しか行われないので（つまり $C=1$ ）、全体では $2B$ 回の反復計算を行えば良いということになる。

通常のブートストラップ法については、多くの理論的研究が行われ、実証分析においても多数の応用が行われている。ダブル・ブートストラップ法およびファスト・ダブル・ブートストラップ法についても、通常のブートストラップ法についてではないにせよ、様々な研究が行われているが（Caers and Dyck (1998), DiCiccio et al. (1992), Hinkley and Shi (1989), McKnight et al. (2000), Letson and McCullough (1998), McCullough and Vinod (1998), Nankervis (2005), Vinod (1995), Vinod and McCullough (1995) 等を参照）、未知の部分も多

い。そこで本稿では、いくつかの分布の平均に関する検定にダブル・ブートストラップ法およびファスト・ダブル・ブートストラップ法を応用し、その小標本特性をコンピューターを用いたシミュレーションにより分析する。本稿の構成は以下の通りである。第2節では、通常のブートストラップ法、ダブル・ブートストラップ法およびファスト・ダブル・ブートストラップ法を用いた検定について簡単に説明する。第3節ではそれぞれの検定の検定力をシミュレーションにより比較する。第4節では、得られた結果および今後の展望を簡単に述べる。

2 検 定 方 法

2.1 漸近理論に基づく検定

$X_1, X_2, \dots, X_n$ を平均 μ 、分散 σ^2 の母集団から得られた互いに独立な標本とする。本稿では、帰無仮説 $H_0: \mu=0$ を対立仮説 $H_1: \mu>0$ に対して検定することを考えよう。帰無仮説 $H_0: \mu=\mu_0 (\neq 0)$ を検定したい場合には、 $Y_i=X_i-\mu_0$ として、 Y_i の平均が 0 であるという帰無仮説を検定すれば良いので、平均が 0 であるという帰無仮説のみを考えることにより一般性は失われない。

H_0 を検定するために t 統計量

$$T = \frac{\bar{X}}{s/\sqrt{n}} \quad (1)$$

を考えよう。ただし $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, s は s^2 の正の平方根である。 H_0 の下で $n \rightarrow \infty$ のとき $T \xrightarrow{D} N(0, 1)$ となり、 T の漸近分布は未知パラメータに依存しないので、 t 統計量は漸近的にピボットであるといえる。

z_α を正規分布の上側 $100 \times \alpha\%$ 点とすると、 T の実現値が z_α より大きいとき、帰無仮説 H_0 は有意水準 $100 \times \alpha\%$ で棄却されることになる。この検定は次のように行うこともできる。 Z を標準正規分布に従う確率変数とし、 P 値を $p = P(Z > T)$ で定義する。²⁾ $p < \alpha$ であれば H_0 は有意水準 $100 \times \alpha\%$ で棄却される。上記の P 値は、 $F_T(\cdot)$ を帰無仮説の下での t 統計量の分布関数とすると、 $p = 1 - F_T(T)$ と定義される。したがって、帰無仮説が正しいとき、 p は $(0, 1)$ 上の一様分布に従い、 p の分布は未知パラメータに依存しないので、 p はピボットである。

以上の検定は T が漸近的に標準正規分布に従うことを利用している。しかしながら、標本の大きさが有限の場合は、 T の分布は厳密には正規分布ではないので、この検定は近似的なものである。

2.2 ブートストラップ法

上記の漸近理論に基づく検定が正規分布表を用いて行われるのに対し、ブートストラップ法では、統計分布表を用いず、手元にある標本からのリサンプリングにより検定を行う。具体的な手順は以下の通りである。

A1. $X_1, X_2, \dots, X_n$ から大きさ n の標本を復元抽出する。この過程をリサンプリングと呼び、得られた標本 $X_1^*, X_2^*, \dots, X_n^*$ はブートストラップ標本と呼ばれる。つまり、リサンプリングは $P(X_i^* = X_j) = 1/n$, $(i, j = 1, 2, \dots, n)$ という分布から無作為標本を抽出することと同じことである。

A2. $\chi = \{X_1, X_2, \dots, X_n\}$ を所与としたときのブートストラップ標本の期待値は

$$E[X_i^* | \chi] = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$$

であり、平均 0 という帰無仮説を満たさないので、 $X_{bi} = X_i^* - \bar{X}$, $(i = 1, 2, \dots, n)$ と変換する。この変換はリセンタリング (recentering) と呼ばれる。明らかに $E[X_{bi} | \chi] = 0$ であるから、 X_{bi} は帰無仮説を満たす。

A3. X_{bi} を用いて t 統計量

$$T_b = \frac{\bar{X}_b}{s_b / \sqrt{n}}$$

を計算する。ただし、 $\bar{X}_b = \frac{1}{n} \sum_{i=1}^n X_{bi}$, $s_b^2 = \frac{1}{n-1} \sum_{i=1}^n (X_{bi} - \bar{X}_b)^2$, s_b は s_b^2 の正の平方根である。

A4. A1.-A3. を B 回繰り返す。 b は何回目の反復計算かを表す添字であるとする (つまり $b = 1, 2, \dots, B$)。得られた B 個の T_b の経験分布を用いて t 統計量の帰無仮説の下での分布を近似し、検定を行う。そのために、ブートストラップ法による P 値として

$$p_B = \frac{\sum_{b=1}^B 1(T_b > T)}{B} \quad (2)$$

を計算する。ただし $1(A)$ は事象 A が起こったとき 1, それ以外では 0 をとる指示関数 (indicator function) であり T は通常の t 統計量の実現値である。通常のブートストラップ法による検定では、 $p_B < \alpha$ であれば、有意水準 $100 \times \alpha \%$ で帰無仮説を棄却する。

実際の手順においては、A1. のリサンプリングと A2. のセンタリングを同時に行うために、 $X_i - \bar{X}$, $(i = 1, 2, \dots, n)$ からリサンプリングを行えば良い。 $n \rightarrow \infty$ のときに $T_b | \chi \xrightarrow{D}$

$N(0, 1)$ となるので、 T の漸近分布を χ を所与とした T_b の漸近分布で近似できることから、ブートストラップ法が漸近的に有効であることが分かる。 B 回の繰り返し計算は、 T_b の分布をシミュレーションするために行っていると考えることができる。このようなブートストラップ法による検定は、漸近的にピボットである T の分布を近似しているために、漸近理論に基づく通常の検定よりも高い精度を持つことが Beran (1987) により示されている。

2.3 ダブル・ブートストラップ法

ダブル・ブートストラップ法においては、通常のブートストラップ法における B 回の繰り返しのそれぞれのステップにおいて得られたブートストラップ標本に対して、さらに C 回のリサンプリングを行うことになる。具体的には、ダブル・ブートストラップ法による検定は以下のように行われる。

B1. 通常のブートストラップ法の A1.-A2. と同様に $X_1, X_2, \dots, X_n$ をリセンタリングした上でリサンプリングを行い、 X_{bi} , ($i=1, 2, \dots, n$) を得る。得られた X_{bi} を用いて T_b を計算する。

B2. X_{bi} , ($i=1, 2, \dots, n$) にリセンタリングを行った上でリサンプリングを行う。つまり、 $X_{bi} - \bar{X}_b$, ($i=1, 2, \dots, n$) から大きさ n の無作為標本を復元抽出する。得られた標本を X_{bci} , ($i=1, 2, \dots, n$) とする。

B3. X_{bci} を用いて t 統計量

$$T_{bc} = \frac{\bar{X}_{bc}}{s_{bc}/\sqrt{n}} \quad (3)$$

を計算する。ただし、 $\bar{X}_{bc} = \frac{1}{n} \sum_{i=1}^n X_{bci}$, $s_{bc}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_{bci} - \bar{X}_{bc})^2$, s_{bc} は s_{bc}^2 の正の平方根である。

B4. B2.-B4. を C 回繰り返し (つまり $c=1, 2, \dots, C$)、得られた C 個の T_{bc} の値を用いて、 T_b の P 値の推定値

$$\tilde{p}_b = \frac{\sum_{c=1}^C 1(T_{bc} \geq T_b)}{C} \quad (4)$$

を求める。

B5. B1.-B4. を B 回繰り返し、ダブル・ブートストラップ法による P 値

$$p_D = \frac{\sum_{b=1}^B 1(\tilde{p}_b < p_B)}{B} \quad (5)$$

を求める。 $p_D < \alpha$ であれば帰無仮説 H_0 を有意水準 $100 \times \alpha\%$ で棄却する。

ダブル・ブートストラップ法においては、ピボットである P 値の値を近似していることになるため、通常の正規分布による近似より精度の高い検定を行うことができることが Beran (1988) により示されている。

2.4 ファスト・ダブル・ブートストラップ法

前述のステップからも分かるように、ダブル・ブートストラップ法においては、 $B \times C$ 回のリサンプリングが必要になる。一般に第二段階のリサンプリングの回数 C は第一段階のリサンプリングの回数 B より小さくて良いといわれているが (Booth and Hall (1994) を参照)、全体としてかなりの計算時間が必要になる。ダブル・ブートストラップ法における計算時間の問題を解決するために提案されたのが、Davidson and MacKinnon (2002) によるファスト・ダブル・ブートストラップ法である。ファスト・ダブル・ブートストラップ法は、第二段階のリサンプリングを 1 回しか行わないという点でダブル・ブートストラップ法と異なる。ファスト・ダブル・ブートストラップ法による検定は以下のように行われる。

C1. 通常のブートストラップ法の A1.-A2. と同様に $X_1, X_2, \dots, X_n$ をリセンタリング上でリサンプリングを行い、 $X_{bi}, (i=1, 2, \dots, n)$ を得る。得られた X_{bi} を用いて T_b を計算する。

C2. $X_{bi}, (i=1, 2, \dots, n)$ にリセンタリングを行った上で 1 回のリサンプリングを行う。つまり、 $X_{bi} - \bar{X}_b, (i=1, 2, \dots, n)$ から大きさ n の無作為標本を復元抽出する。得られた標本を $X_{bli}, (i=1, 2, \dots, n)$ とする。

C3. X_{bli} を用いて t 統計量

$$T_{b1} = \frac{\bar{X}_{b1}}{s_{b1}/\sqrt{n}} \quad (6)$$

を計算する。

C4. C1.-C3. を B 回繰り返し、 B 個の T_{b1} の p_B 分位点 Q_F を求める。言い換えれば、 Q_F は

$$p_B = \frac{\sum_{b=1}^B 1(T_{b1} > Q_F)}{B} \quad (7)$$

を満たす Q_B の値である。

C5. ファスト・ダブル・ブートストラップ法による P 値

$$p_F = \frac{\sum_{b=1}^B 1(T_b > Q_F)}{B} \quad (8)$$

を求める。 $p_F < \alpha$ であれば帰無仮説 H_0 を有意水準 $100 \times \alpha\%$ で棄却する。

C4. における Q_F の値は、 $B=100$, $p_B=0.05$ であれば、100個の T_{b1} を降順に並べたときの6番目の値となるので、 $B \times p_B$ が整数となる場合には、 B 個の T_{b1} を降順に並べたときの $B \times p_B + 1$ 番目の値を用いれば良い。

以上の手順から分かるように、ダブル・ブートストラップ法においては $B \times C$ 回の繰り返し計算が必要であるのに対し、ファスト・ダブル・ブートストラップ法では、第二段階のリサンプリングを1回しか行わないために、全体で $2B$ 回の繰り返し計算を行うだけで良い。ダブル・ブートストラップ法では、第二段階のリサンプリングを利用して T_b のP値を推定しているのに対し、ファスト・ダブル・ブートストラップ法では T_b の分位点を推定してP値の分布を近似していることになる。

3 シミュレーション分析

本節では、前節で紹介した検定方法の検定力をシミュレーションにより分析する。シミュレーションの設定は以下の通りである。前節において説明したように、帰無仮説 $H_0: \mu=0$ を対立仮説 $H_1: \mu=\mu_1$ に対して検定することを考える。分析対象となる分布としては、正規分布 (N)、一様分布 (U)、自由度1のカイ2乗分布 (X)、自由度3の t 分布 (T3) および自由度4の t 分布 (T4) の乱数を用い³⁾、すべての分布について、平均が μ_1 、分散が1になるように調整を行った。すべての分布について右側検定を行うことを考え、 μ_1 については、0, 0.1, 0.2, 0.4, 0.8を用いた。また、カイ2乗分布については、左右対称の分布ではないので、 μ_1 として0, -0.1, -0.2, -0.4, -0.8を用い、左側検定も行った (X')。標本の大きさ n は15, 20, 30, 50, 100, 200とした。ブートストラップ法における繰り返し回数は、第一段階の繰り返し回数 B については $B=500$ 、第二段階の繰り返し回数 C については $C=300$ とした。検定の有意水準は10% (0.10)、5% (0.05)、1% (0.01) を用い、10,000回の繰り返し計算を行い、帰無仮説 H_0 が棄却される割合を計算することにより、検定力を求めた。なお、第2節で説明した漸近理論に基づく検定では、正規分布を用いて棄却点を求めているが、標本抽出が行われた分布が正規分布の場合には t 統計量は厳密に t 分布に従うので、シミュレーションにおいてはすべての分布に対して t 分布の棄却点を用いた (表中では Normal Approximation と表記する)。標本の大きさが大きくなれば、 t 分布と正規分布のどちらの棄却点を用いてもほとんど差はない。対立仮説として $H_1: \mu=\mu_1=0$ を用いた場合には、検定のサイズが検定力として求められていることになるため、この場合の検定力の値が有意水準に近い検定が、サイズ・ディストーションが少ないという意味で、正確な検定であるといえる。そこで、シミュレーションによって得られた検定のサイズが有意水準に等しいという帰無仮説を、2項分布の正規近似を用いて検定を行った。表中の *, †, ‡ はこの帰無仮説がそれぞれ10%, 5%, 1%で棄却されたことを表している (つまり、これらの記号が全くついていない

表1 $n=30$ の場合の検定力

Dist	μ_0	Normal Approximation			Bootstrap			Double Bootstrap			Fast Double Bootstrap		
		10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
N	0.0	0.0958	0.0503	0.0098	0.0960	0.0500	0.0109	0.0991	0.0527	0.0134 [‡]	0.0958	0.0512	0.0125 [‡]
	0.1	0.2285	0.1311	0.0360	0.2277	0.1288	0.0370	0.2316	0.1315	0.0388	0.2274	0.1293	0.0378
	0.2	0.4207	0.2848	0.0989	0.4165	0.2770	0.0965	0.4151	0.2818	0.1007	0.4088	0.2773	0.0979
	0.4	0.8077	0.6877	0.4020	0.8048	0.6800	0.3837	0.8030	0.6765	0.3736	0.7974	0.6688	0.3672
	0.8	0.9990	0.9965	0.9675	0.9991	0.9959	0.9584	0.9992	0.9952	0.9326	0.9988	0.9936	0.9105
U	0.0	0.1005	0.0514	0.0109	0.0944 [*]	0.0446 [†]	0.0065 [‡]	0.1028	0.0539 [*]	0.0113	0.0960	0.0469	0.0079 [‡]
	0.1	0.2318	0.1360	0.0373	0.2251	0.1232	0.0238	0.2369	0.1417	0.0379	0.2253	0.1238	0.0282
	0.2	0.4142	0.2763	0.0977	0.4103	0.2643	0.0714	0.4216	0.2874	0.0995	0.4084	0.2593	0.0735
	0.4	0.8071	0.6797	0.3958	0.8146	0.6850	0.3521	0.8199	0.7071	0.4068	0.8073	0.6784	0.3234
	0.8	0.9994	0.9974	0.9765	0.9995	0.9982	0.9794	0.9994	0.9984	0.9805	0.9993	0.9976	0.9278
X	0.0	0.0525 [‡]	0.0140 [‡]	0.0008 [‡]	0.0895 [‡]	0.0363 [‡]	0.0040 [‡]	0.0990	0.0468	0.0092	0.0941 [†]	0.0414 [‡]	0.0057 [‡]
	0.1	0.1648	0.0583	0.0040	0.2440	0.1316	0.0225	0.2589	0.1542	0.0389	0.2454	0.1372	0.0295
	0.2	0.3821	0.1992	0.0270	0.4918	0.3358	0.1036	0.5094	0.3710	0.1487	0.4936	0.3413	0.1171
	0.4	0.9224	0.8103	0.3696	0.9601	0.9099	0.6912	0.9629	0.9231	0.7477	0.9565	0.9071	0.6640
	0.8	1.0000	1.0000	0.9996	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9842
X'	0.0	0.1829 [‡]	0.1280 [‡]	0.0630 [‡]	0.1221 [‡]	0.0696 [‡]	0.0202 [‡]	0.1241 [‡]	0.0717 [‡]	0.0213 [‡]	0.1169 [‡]	0.0657 [‡]	0.0204 [‡]
	-0.1	0.3150	0.2346	0.1293	0.2245	0.1396	0.0426	0.2276	0.1389	0.0461	0.2182	0.1336	0.0446
	-0.2	0.4972	0.3922	0.2391	0.3717	0.2457	0.0923	0.3695	0.2464	0.0910	0.3576	0.2369	0.0903
	-0.4	0.7750	0.6839	0.5036	0.6421	0.4890	0.2310	0.6223	0.4638	0.2104	0.6199	0.4701	0.2213
	-0.8	0.9767	0.9542	0.8803	0.9206	0.8419	0.5948	0.8827	0.7865	0.5064	0.9061	0.8222	0.5606
T3	0.0	0.1022	0.0492	0.0086	0.1203 [‡]	0.0665 [‡]	0.0172 [‡]	0.1215 [‡]	0.0697 [‡]	0.0204 [‡]	0.1265 [‡]	0.0734 [‡]	0.0224 [‡]
	0.1	0.2646	0.1575	0.0439	0.2828	0.1847	0.0687	0.2759	0.1850	0.0723	0.2886	0.1941	0.0769
	0.2	0.4847	0.3444	0.1338	0.4863	0.3638	0.1734	0.4713	0.3510	0.1700	0.4849	0.3665	0.1854
	0.4	0.8481	0.7606	0.5245	0.8201	0.7329	0.5353	0.7944	0.6999	0.4869	0.8096	0.7268	0.5324
	0.8	0.9894	0.9821	0.9510	0.9744	0.9617	0.9114	0.9598	0.9350	0.8588	0.9708	0.9576	0.8951
T4	0.0	0.0993	0.0496	0.0093	0.1097 [‡]	0.0601 [‡]	0.0149 [‡]	0.1086 [‡]	0.0610 [‡]	0.0172 [‡]	0.1121 [‡]	0.0643 [‡]	0.0189 [‡]
	0.1	0.3177	0.1968	0.0609	0.3256	0.2172	0.0818	0.3194	0.2130	0.0840	0.3282	0.2184	0.0892
	0.2	0.6204	0.4853	0.2285	0.6107	0.4867	0.2575	0.5958	0.4665	0.2410	0.6053	0.4818	0.2579
	0.4	0.9565	0.9167	0.7657	0.9387	0.8896	0.7306	0.9218	0.8589	0.6685	0.9325	0.8814	0.7048
	0.8	0.9989	0.9985	0.9940	0.9959	0.9923	0.9813	0.9902	0.9839	0.9575	0.9941	0.9911	0.9603

検定方法が理想的な方法である)。ここでは $n=30$ と $n=200$ の場合の結果のみを示すが、他の場合についても同様の傾向が得られる。

まず、正規分布において $n=30$ の場合には、正規近似および通常のブートストラップ法により、正確なサイズの検定が行われていることが分かる。これに対し、ダブル・ブートストラップ法およびファスト・ダブル・ブートストラップ法による検定は、有意水準が1%のときにあまり正確でない。しかしながら、標本が大きくなり、 $n=200$ の場合には、このような不正確さはなくなり、ダブル・ブートストラップ法もファスト・ダブル・ブートストラップ法も正確なサイズの検定が行われるようになることが分かる。検定力については、どの方法でも大きな差はないものの、ファスト・ダブル・ブートストラップ法による検定力は、他の方法よりも若干低いように見える。

次に、一様分布において $n=30$ の場合は、正規近似による検定のサイズは正確であるが、通常のブートストラップ法による検定はあまり正確でないという結果になっている。しかしながら、ダブル・ブートストラップ法およびファスト・ダブル・ブートストラップ法を用いることにより、検定の精度は改善している。 $n=200$ の場合には、すべての方法において検定精度の改善が見られ、正確なサイズの検定が行えることが分かる。この場合も、ファスト・ダブル・ブートストラップ法による検定力は、他の方法のものよりも若干低いようである。

自由度1のカイ2乗分布の右側検定では、 $n=30$ の場合、正規近似による検定のサイズは全く正確ではない。通常のブートストラップ法を用いることによりこの不正確さはある程度解消されているものの、正確なサイズが得られているとはいえない。以上の結果は $n=200$ の場合でも同様である。これに対し、ダブル・ブートストラップ法による検定のサイズは $n=30$ の場合でも正確である。また、ファスト・ダブル・ブートストラップ法による検定は、 $n=30$ の場合にはあまり正確でないものの、正規近似および通常のブートストラップ法の検定よりも正確であり、 $n=200$ の場合には正確なサイズの検定が行えることになる。検定力に関しては、ダブル・ブートストラップ法の検定力は非常に高く、ファスト・ダブル・ブートストラップ法による検定力はこの場合にも若干低い傾向にあるようである。

自由度1のカイ2乗分布において左側検定を行ったとき、 $n=30$ の場合は、正規近似による検定のサイズは正確でない。右側検定の場合と同様に、通常のブートストラップ法を利用することで、かなりサイズの改善は見られるものの、正確なサイズを得ることはできない。また、この場合に関しては、ダブル・ブートストラップ法およびファスト・ダブル・ブートストラップ法に関しても、検定の精度は通常のブートストラップ法と同等である。しかし、 $n=200$ の場合は、正規近似は依然として正確でないものの、他の三つの方法については、かなりの改善が見られる。特に、ダブル・ブートストラップ法による検定の精度はかなり高いようである。検定力については、正規近似に基づく検定以外のものについてはほぼ同等で

ある。

自由度3の t 分布の場合については、正規近似以外の検定のサイズはあまり正確ではない。標本の大きさが $n=200$ の場合でも、検定のサイズが有意水準に等しいとは言いがたい。これは、ブートストラップ法による近似が正規近似よりも正確であるためには、高次のモーメントが存在することが必要であるのにもかかわらず、自由度3の t 分布では3次以上のモーメントが存在しないという理論的な側面から生じていると考えられる。しかしながら、標本が $n=30$ から $n=200$ へと大きくになるにつれて、検定のサイズはある程度改善しているようである。また、 $n=200$ の場合について見ると、ファスト・ダブル・ブートストラップ法による検定は、帰無仮説が正しい場合に過大に棄却する傾向があるにもかかわらず、検定力はあまり高くないようである。

自由度4の t 分布の場合には、 $n=30$ の場合には正規近似以外の検定のサイズは正確ではないが、 $n=200$ に標本数が増加した場合には、正規近似以外の検定の精度も改善していることが分かる。特に、ダブル・ブートストラップ法による検定は、 $n=200$ の場合には正確なサイズが得られる。また、ファスト・ダブル・ブートストラップ法に関しては、この場合についても、帰無仮説を過大に棄却する傾向があるにもかかわらず、検定力は若干低いようである。

以上の結果を要約すると以下の通りである。

1. ダブル・ブートストラップ法およびファスト・ダブル・ブートストラップ法による検定は、標本の大きさが大きくなるにつれて精度の高いものとなる。
2. 正規近似や通常のブートストラップ法による検定が正確でない場合でも、ダブル・ブートストラップ法およびファスト・ダブル・ブートストラップ法を用いることにより、検定の精度がかなり改善する場合がある。特に、標本が大きいとき、ダブル・ブートストラップ法による検定は、他の方法と比べて高い精度を持つケースが多い。
3. ファスト・ダブル・ブートストラップ法による検定は、サイズの精度が高い場合でも、検定力は他の方法に比べて低いことが多いようである。

以上のことを踏まえると、標本がある程度大きく、繰り返し計算に要する時間が現実的なものである場合には、ダブル・ブートストラップ法を用いる利点があるように思われる。

4 結 論

本稿では、ダブル・ブートストラップ法およびファスト・ダブル・ブートストラップ法を用いた検定の検定力を、漸近理論に基づいた正規近似による検定および通常のブートストラッ

ブ法に基づく検定の検定力とシミュレーションにより比較した。シミュレーションの結果から、標本がある程度大きい場合、ダブル・ブートストラップ法による検定は良好なパフォーマンスを持つことが示された。特に、通常のブートストラップ法による検定や正規近似による検定があまり正確でない場合でも、ダブル・ブートストラップ法を用いることにより検定サイズの改善が期待できることが示された。このことから、標本がある程度大きく、反復計算に要する時間が現実的である場合のダブル・ブートストラップ法の有効性が示唆される。また、ダブル・ブートストラップ法の反復計算に要する時間が現実的でない場合でも、ダブル・ブートストラップ法ほどには正確な検定サイズや高い検定力は得られないかもしれないが、ファスト・ダブル・ブートストラップ法を用いることにより、既存の方法より正確な検定が行える可能性がある。

以上の結果から、既存の方法による検定があまり正確でない場合や、検定統計量の漸近分布が標準的でない場合、あるいは局外母数 (nuisance parameter) に依存する場合等に対してダブル・ブートストラップ法やファスト・ダブル・ブートストラップ法を応用するのは非常に興味深いことであるといえよう。このような課題については、今後更なる研究が待たれるところである。

注

本研究は JSPS 科研費23243038の助成を受けている。

- 1) Hall (1992) にも詳細な説明がある。
- 2) 統計学に関する書物の多くでは、アルファベットの大きい文字が確率変数、小さい文字がその実現値として用いられているが、表記が煩雑になるのを防ぐため、本稿ではこのような区別は行わない。文中で用いられる記号が確率変数とその実現値のどちらを表すかは、文脈から明らかであろう。
- 3) ブートストラップ法の理論的分析においては、エッジワース展開を用いるために、高次のモーメントの存在が必要とされる。本稿で扱う片側検定においては、ブートストラップ法が正規近似よりも正確であるという結論を解析的に得るには、3次のモーメントの存在が必要である。自由度3および自由度4の t 分布を用いたのは、 t 分布は自由度より小さい次数のモーメントしか存在しないという性質を利用して、3次のモーメントの存在の影響を見るためである。

参考文献

- Beran, R., 1987, "Prepivoting to Reduce Level Error of Confidence Sets," *Biometrika*, 74, 457-468.
- Beran, R., 1988, "Prepivoting Test Statistics: A Bootstrap View of Asymptotic Refinements," *Journal of the American Statistical Association*, 83, 687-697.
- Booth, J. G., and P. Hall, 1994, "Monte Carlo Approximation and the Iterated Bootstrap," *Biometrika*, 81, 331-340.
- Caers, J., and J. V. Dyck, 1998, "Nonparametric Tail Estimation Using a Double Bootstrap Method," *Computational Statistics & Data Analysis*, 29, 191-211.

- Davidson, R., and J. G. MacKinnon, 2002, "Fast Double Bootstrap Tests of Nonnested Linear Regression Models," *Econometric Reviews*, 21, 419-429.
- Davidson, R., and J. G. MacKinnon, 2007, "Improving the Reliability of Bootstrap Tests with the Fast Double Bootstrap," *Computational Statistics & Data Analysis*, 51, 3259-3281.
- Davison, A. C., and D. V. Hinkley, 1997, *Bootstrap Methods and Their Application*, vol. 1, Cambridge University Press.
- DiCiccio, T. J., M. A. Martin, and G. A. Young, 1992, "Fast and Accurate Approximate Double Bootstrap Confidence Intervals," *Biometrika*, 79, 285-295.
- Efron, B., 1979, "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, 7, 1-26.
- Hall, P., 1992, *The Bootstrap and Edgeworth Expansion*, Springer-Verlag Inc.
- Hinkley, D., and S. Shi, 1989, "Importance Sampling and the Nested Bootstrap," *Biometrika*, 76, 435-446.
- Letson, D., and B. McCullough, 1998, "Better Confidence Intervals: the Double Bootstrap with No Pivot," *American Journal of Agricultural Economics*, 80, 552-559.
- McCullough, B., and H. D. Vinod, 1998, "Implementing the Double Bootstrap," *Computational Economics*, 12, 79-95.
- McKnight, S. D., J. W. McKean, and B. E. Huitema, 2000, "A Double Bootstrap Method to Analyze Linear Models with Autoregressive Error Terms," *Psychological Methods*, 5, 87-101.
- Nankervis, J. C., 2005, "Computational Algorithms for Double Bootstrap Confidence Intervals," *Computational Statistics & Data Analysis*, 49, 461-475.
- Vinod, H. D., 1995, "Double Bootstrap for Shrinkage Estimators," *Journal of Econometrics*, 68, 287-302.
- Vinod, H. D., and B. McCullough, 1995, "Estimating Cointegration Parameters: an Application of the Double Bootstrap," *Journal of Statistical Planning and Inference*, 43, 147-156.