日本統計学会誌
第 41 巻, 第 2 号, 2012 年 3 月
445 頁 ～ 473 頁

マルチスケール・ブートストラップの漸近理論

下平　英寿*

Asymptotic Theories of Multiscale Bootstrap

Hidetoshi Shimodaira*

マルチスケール・ブートストラップはブートストラップ確率のスケーリング則を利用して近似的に不偏な検定の p-値を計算する方法である．ブートストラップのサンプルサイズ m をデータのサンプルサイズ n から変更して確率分布のスケールを変化させる．データを変換して多変量正規分布に従うと仮定すれば，正規化ブートストラップ確率を仮想的に負のサンプルサイズ $m = -n$ に外挿して計算したものが p-値になる．しかし指数型分布族を仮定すると正規モデルからのズレを表すバイアス修正項の計算が必要になる．もし帰無仮説の形状が錐の場合は通常の漸近理論が正当化されず，曲面のテイラー展開の代わりにフーリエ変換を使う．本稿では，このようないくつかの設定におけるマルチスケール・ブートストラップの漸近理論を紹介する．さらにブートストラップ信頼区間の標準的な漸近理論の観点からマルチスケール・ブートストラップの説明を試みる．

Multiscale bootstrap computes a p-value of an approximately unbiased test by utilizing the scaling law of bootstrap probability. The scale of the probability distribution changes when the sample size m of bootstrap is altered from the sample size n of data. Assuming that a transformed data follows the multivariate normal distribution, the p-value is computed as a normalized bootstrap probability with sample size being extrapolated to $m = -n$, meaning a fictional negative sample size. Assuming the exponential family of distributions, instead, we need to compute a bias correction term representing a deviation from the normal model. If the null hypothesis is a cone-shaped region, the ordinary asymptotic theory is not justified and we may use the Fourier transform of the surface instead of the Taylor series. In this paper, we illustrate these asymptotic theories for multiscale bootstrap in several settings. We also attempt to explain multiscale bootstrap in the light of the standard asymptotic theory of bootstrap confidence interval.

キーワード: ブートストラップ確率，バイアス修正，不偏検定，スケーリング則，平均曲率，加速定数，ブートストラップ信頼区間，漸近理論．

1.　はじめに

データにはランダムネスがあることを考慮すると，データから「計算」して得られた結果（予測値や推定値）をどれほど信頼してよいだろうか？　その「計算」を反復実行する

* 東京工業大学大学院情報理工学研究科：〒 152-8552 東京都目黒区大岡山 2-12-1.

だけで計算結果の統計的信頼度を精確に求める方法を紹介する．Efron (1979) のブートストラップは観測データのランダムネスを反映するモンテカルロ・シミュレーションである．多数のデータを生成して得られた結果のヒストグラムによってランダムネスの程度を見積もる．とくに階層型クラスタリングのように計算結果が離散値の場合，各値の観測頻度はブートストラップ確率と呼ばれる．これは Felsenstein (1985) 以来，分子進化学で系統樹を推定する問題に広く応用されている．

ブートストラップにはバイアスがあることが知られており，計算結果が連続値の場合にヒストグラムから得られる信頼区間を修正して被覆確率のバイアスを減少させる方法が研究された．サンプルサイズ $n \to \infty$ の漸近理論において被覆確率のバイアスが $O(n^{-k/2})$ のとき k 次の精度という．ブートストラップ信頼区間は 1 次の精度である．Efron (1987) の BCa (bias-corrected and accelerated) や，BCa の近似計算である DiCiccio and Efron (1992) の ABC (approximate bootstrap confidence または approximate BCa) は 2 次の精度である．もっと素朴な bootstrap-t 信頼区間でも 2 次の精度が得られる．信頼区間のバイアス修正に関する研究は Hall (1992) に良くまとめられている．

計算結果が離散値の場合にブートストラップ確率を修正するのは連続値の場合より困難である．ブートストラップ確率を仮説検定の p-値として解釈すると 1 次の精度しかない．2 次の BCa 信頼区間に相当する p-値を計算するには，ブートストラップ確率を修正するための 2 個の修正項が必要である．1 つは仮説領域の境界曲面の「平均曲率」，もう一つは分散の変化率で「加速定数」と呼ばれる．Efron *et al.* (1996), Efron and Tibshirani (1998) は離散値しか得られない状況で平均曲率と加速定数を計算する方法を提案した．しかし実用には困難があり，平均曲率を計算するためにデータではなく帰無分布からブートストラップしたり，加速定数を計算するためにヒューリスティックな近似をしていた．

データを変換して多変量正規分布に従うと仮定すれば問題が容易になり，加速定数の計算は不要になる．この正規モデルを仮定して Shimodaira (2002a) はブートストラップだけで平均曲率を計算するマルチスケール・ブートストラップを提案した．Shimodaira (2002a) のアイデアはブートストラップのサンプルサイズ m をデータのサンプルサイズ n から変更して確率分布のスケールを変化させることである．ブートストラップ確率がどのように変化するかを表すスケーリング則のプロットから曲線の傾きとして平均曲率が推定できる．しかし一般の指数型分布族を仮定すると，正規モデルからのズレを表す加速定数が必要になる．そこで Shimodaira (2004) はブートストラップだけで平均曲率と加速定数を計算する方法を提案し，さらに 3 個の修正項を追加して 3 次の精度の p-値を求めた．Shimodaira (2004) のアイデアは 2 段階ブートストラップにおける正規モデルからのズレとして加速定数を推定することである．

これまでの議論で用いた通常の漸近理論が適用できない状況でも正当化できる漸近理論

を考察した結果，Shimodaira (2008a) はマルチスケール・ブートストラップの原理をシンプルに表現することに成功した．通常の漸近理論のポイントは $n \to \infty$ の極限で曲面が平坦に近づくことである．これは仮説領域の境界曲面をテイラー展開すると k 階微分 $(k \geq 2)$ は $O(n^{-(k-1)/2})$ と表されることから導かれる．ところが重要な応用例において仮説領域は凸多面錐になり曲面が滑らかでないからテイラー展開による議論が正当化できない．さらに問題なのは，錐はスケーリングによって不変だから，そもそも曲面が平坦に近づかない．たとえば系統樹推定はこの状況であり，この分野で使われてきた Shimodaira and Hasegawa (1999) の検定は多重比較法の応用である．これは帰無分布を錐の頂点と仮定する最悪評価のため検定は保守的になる．そこで錐など曲面が滑らかでない場合でも正当化できる漸近理論を Shimodaira (2008a) は考案した．漸近理論を制御するパラメータとして n の代わりに曲面の高さ方向の変動を採用し，曲面をつぶして平坦に近づけるアイデアである．もう一つのアイデアは，曲面のテイラー展開の代わりに，曲面のフーリエ変換を用いることである．これならば，滑らかでない曲面にも適用できる．正規モデルを仮定したマルチスケール・ブートストラップをこの漸近理論で調べた結果，正規化ブートストラップ確率を $m = -n$ で計算すれば不偏検定の p-値になることが分かった．これに気づいてから Shimodaira (2002a) の結果を振り返ると，$n \to \infty$ とする通常の漸近理論においても，3 次の精度の p-値はやはり $m = -n$ の正規化ブートストラップ確率であった．

本稿の目的はマルチスケール・ブートストラップを導出するための漸近理論の現状を紹介することである．まず第 2 節で理論に立ち入らずにマルチスケール・ブートストラップの概要を説明したあと，第 3 節から第 6 節の各節ではそれぞれ異なる漸近理論の設定を紹介する．第 3 節は正規モデルを仮定して Shimodaira (2002a) の概要を紹介する．第 4 節は指数型分布族を仮定して Efron and Tibshirani (1998) を参照しながら Shimodaira (2004) の概要を紹介する．第 5 節は Hall (1992) の「滑らかな関数モデル」を仮定した信頼区間の理論を参照して，マルチスケール・ブートストラップの説明を試みた．第 4 節と比べて新しい結果はないが，ブートストラップ信頼区間の標準的な理論とマルチスケール・ブートストラップを結びつける意義がある．第 6 節は滑らかでない曲面でも正当化できる Shimodaira (2008a) の漸近理論の概要を紹介する．正規モデルを仮定するマルチスケール・ブートストラップの現状を把握するには，第 4 節と第 5 節を読まずに第 6 節へ進んでよい．

マルチスケール・ブートストラップのアルゴリズムや理論はいまのところ限定された状況で議論されているに過ぎない．これまでに述べた漸近理論は極限で仮説領域が半空間（境界は平坦）になって空間が 2 個に分割されることを仮定している．本稿では説明しないが，Shimodaira (2010) は Shimodaira (2008a) の漸近理論を拡張して，空間が 3 個の領域に分割される場合を考察した．まさにこの状況が，Efron and Tibshirani (1998) で議論されたように頻度論の p-値とベイズ事後確率の違いを明確にする．2 個に分割されるときベイズ

事後確率を頻度論の p-値に一致させる probability matching prior が定義できるが，3 個に分割されたときベイズと頻度論が異なる信頼度を与える．領域が任意の配置をするときに有効な漸近理論はまだ得られていない．

本稿の目的は理論の紹介であるから，マルチスケール・ブートストラップを実装したソフトウエアや数値例について説明しない．系統樹推定は CONSEL (Shimodaira and Hasegawa (2001))，階層型クラスタリングは pvclust (Suzuki and Shimodaira (2006)) がある．これらは第 3 節の理論を実装している．一般用は scaleboot (Shimodaira (2006)) があり，これは第 6 節の理論を実装している．日本語の文献は下平 (2002b, 2008b) がある．

2. マルチスケール・ブートストラップの使い方

2.1 ブートストラップでバラツキを調べる

サンプルサイズ n のデータ $D_n = \{x_1, x_2, \ldots, x_n\}$ から計算する関数 $f(D_n)$ の値がどれほど信頼できるかを調べたい．要素の平均値 $\bar{x}_n = (x_1 + x_2 + \cdots + x_n)/n$ を計算する $f(D_n) = \bar{x}_n$ はきわめて単純な例であるし，もっと複雑なソフトウエアとして実装されている関数でもよい．データの各要素は独立に同分布 $q(x)$ に従う確率変数の観測値とする．これを $x_1, x_2, \ldots, x_n \sim q(x)$ と表す．

真の分布 $q(x)$ からデータをサンプリングするコンピュータ・シミュレーションによって多数のデータを生成できると考えてみる．B 個のシミュレーションデータを $D_n^{(b)}, b = 1, \ldots, B$ とする (例えば $B = 10000$)．これから $f(D_n^{(b)}), b = 1, \ldots, B$ を計算して，その標準偏差などによってバラツキを調べれば，$f(D_n)$ にどれほどランダムネスがあるかの目安になる．

ところが現実のデータ解析では真の分布 $q(x)$ が未知であるから，観測した D_n を使ってシミュレーションを行う．Efron (1979) のブートストラップは n 個の要素をランダムに D_n から取り出して，D_n の複製データ $D_n^* = \{x_1^*, x_2^*, \ldots, x_n^*\}$ とする．ブートストラップは重複を許すリサンプリングであり，同じ要素を何回でも取り出して良い．この手続きを多数回繰り返し実行して，$D_n^{*(b)}, b = 1, \ldots, B$ を生成する．そして $f(D_n^{(b)}), b = 1, \ldots, B$ の代わりに $f(D_n^{*(b)}), b = 1, \ldots, B$ を計算してバラツキを測定する．

2.2 ブートストラップのサンプルサイズを変更する

Bickel *et al.* (1997) の “m out of n boostrap” は重複を許して m 個の要素をランダムに D_n から取り出して，D_n の複製データ $D_m^* = \{x_1^*, x_2^*, \ldots, x_m^*\}$ とする．つまりブートストラップしたデータのサンプルサイズを n から m へ変更する．同じ要素を何回でも取り出せるから，任意の $m > 0$ を指定できる．しかし m を変更するとバラツキの大きさが変化することを考慮する必要がある．たとえば $f(D_n) = \bar{x}_n$ の場合は $f(D_m^{*(b)}), b = 1, \ldots, B$ の標

準偏差が $1/\sqrt{m}$ に比例するから，$m=n$ と比較して分散の大きさの比は

$$\sigma^2 = \frac{n}{m} \tag{2.1}$$

である．つまりバラツキを標準偏差でみたときのスケールは σ で表される．したがって $\sigma^{-1}(f(D_m^{*(b)}) - f(D_n))$ のようにスケーリングすれば $m=n$ のときの本来のバラツキが調べられる．一般に $f(D_n)$ が平均値でない場合でも，本稿の σ^2 は (2.1) 式の意味で用いる．

m out of n bootstrap は m の値を n より十分に小さくして使う．$n \to \infty$ の漸近理論においては，たとえば $m = O(\sqrt{n})$ のように，$m \to \infty, m/n \to 0$ とする．このとき D_m^* の要素が重複することの影響は小さいので，重複を許さずに要素を取り出すサブサンプリング (Politis and Romano (1994)) でも結果はあまり変わらない．小さい m を使う目的は，そのほうが $f(D_m^*)$ の計算時間が短くなる場合があることと，$f(D_n) = \max\{x_1, \ldots, x_n\}$ など通常のブートストラップがうまく機能しない関数でよい結果が得られる場合があることである．このような利点が特に重要でない場合は m を変更せずに $m=n$ とした通常のブートストラップを使えばよい．

2.3 負のサンプルサイズへバラツキを外挿する

Shimodaira (2002a, 2004, 2008a) のマルチスケール・ブートストラップ (multiscale bootstrap) は m out of n bootstrap と全く同様に，ブートストラップしたデータのサンプルサイズを n から m へ変更する．ただし m out of n bootstrap では m を n より十分に小さくしていたが，マルチスケール・ブートストラップでは m の値を n から大きく変えることはせず，n の数分の 1 から数倍程度の範囲で変化させる．$n \to \infty$ の漸近理論においては $m = O(n)$ とする．

マルチスケール・ブートストラップでは D_m^* を生成して $f(D_m^*)$ のバラツキを測定し，その結果を外挿して仮想的に負のサンプルサイズ $m = -n$ のときのバラツキを調べる．外挿するために $m=n$ の周辺でいくつかの m の値を指定してバラツキを測定する．複数のスケール σ においてブートストラップを実行するので，マルチスケール・ブートストラップと名付けた．

m out of n bootstrap がスケーリングによって $m=n$ のときのバラツキを調べるのは常識的といえるが，マルチスケール・ブートストラップがサンプルサイズの符号を反転して $m=-n$ とするのは非常識かもしれない．その目的はなにか？ (2.4 節)，どうやって外挿するのか？ (2.5 節)，なぜサンプルサイズの符号を反転するのか？ (3.7 節)，これらの疑問に答えていく．

2.4　ブートストラップ確率の計算

マルチスケール・ブートストラップでは $f(D_n)$ は離散値をとる関数とする．ここでは特に 0 または 1 の二値関数とする．観測値は $f(D_n)=0$ か $f(D_n)=1$ のどちらかであるが，その信頼度を調べたい．

Felsenstein (1985) のブートストラップ確率 (bootstrap probability) は，

$$\widehat{\mathrm{BP}}(\sigma^2)=\frac{\#\{f(D_m^{*(b)})=1\}}{B}=\frac{1}{B}\sum_{b=1}^{B}f(D_m^{*(b)})$$

である．十分に B が大きい場合を考えて，ここでは頻度ではなく確率

$$\mathrm{BP}(\sigma^2)=P\Big\{f(D_m^*)=1\mid D_n\Big\}=E\Big\{f(D_m^*)\mid D_n\Big\}$$

をブートストラップ確率の定義としておく．ブートストラップ確率は m の関数になるが，m の代わりに σ^2 をパラメータとして用いる．通常のブートストラップでは $\sigma^2=1$ すなわち $m=n$ とおいて

$$\mathrm{BP}=\mathrm{BP}(1)$$

が $f(D_n)=1$ の信頼度として用いられる．

十分に B が大きければ誤差を無視して $\widehat{\mathrm{BP}}(\sigma^2)\approx\mathrm{BP}(\sigma^2)$ と考えてよい．$\hat p=\widehat{\mathrm{BP}}(\sigma^2)$ の分散は $V(\hat p|D_n)\approx\hat p(1-\hat p)/B$ である．たとえば $\hat p=0.05$ のとき，$B=10000$ なら標準誤差は $(0.05*0.95/10000)^{1/2}=0.002$ となり十分小さいが，$B=100$ では $(0.05*0.95/100)^{1/2}=0.02$ となり不十分だろう．

ここで $f(D_n)$ の信頼度の意味を考えておく．真の分布 $q(x)$ からサンプリングを考えると $f(D_n)$ は確率変数である．しかし $n\to\infty$ の極限を考えるとバラツキが減少して $f(D_\infty)=\lim_{n\to\infty}f(D_n)=0$ または 1 が定まるとする．この $f(D_\infty)$ が $f(D_n)$ の真値である．観測値が $f(D_n)=1$ だったとしても真値が $f(D_\infty)=1$ とは限らない．マルチスケール・ブートストラップの目的は，$f(D_\infty)=1$ であるという仮説検定の頻度論的な p-値を求めることである．BP はそのような p-値として利用できるがバイアスが大きく，典型的な問題では false positive を生じやすいことが知られている．

2.5　近似的に不偏な検定の p-値を計算

Shimodaira (2008a) は次式を正規化ブートストラップ確率 (normalized bootstrap probability) と定義した．

$$\mathrm{AU}(\sigma^2)=\Phi\Big\{\sigma\Phi^{-1}(\mathrm{BP}(\sigma^2))\Big\}\tag{2.2}$$

ただし $\phi(x)=(2\pi)^{-1/2}e^{-x^2/2}$, $\Phi(x)=\int_{-\infty}^{x}\phi(t)\,dt$, $x\in\mathbb{R}$ は標準正規分布 $N(0,1)$ の密度関数と累積分布関数，$\Phi^{-1}(\cdot)$ は $\Phi(\cdot)$ の逆関数である．定義 (2.2) より

$$\mathrm{BP}=\mathrm{AU}(1)$$

である．後ほどの都合で上側確率の記号

$$\bar{\Phi}(z) = 1 - \Phi(z) = \Phi(-z), \quad \bar{\Phi}^{-1}(p) = \Phi^{-1}(1-p) = -\Phi^{-1}(p)$$

を定義すると，$\mathrm{AU}(\sigma^2) = \bar{\Phi}(\sigma\bar{\Phi}^{-1}(\mathrm{BP}(\sigma^2)))$ と書いても良い．そして近似的に不偏 (approximately unbiased) な検定の p-値を

$$\mathrm{AU} = \mathrm{AU}(-1)$$

と定義する．もちろんブートストラップは $m > 0$ しか実行できないので, $\mathrm{BP}(\sigma^2)$ と $\mathrm{AU}(\sigma^2)$ が定義されるのは本来 $\sigma^2 > 0$ である．そのため $\mathrm{AU}(\sigma^2)$ の $\sigma^2 > 0$ における測定値を $\sigma^2 \le 0$ へ外挿して $\mathrm{AU}(-1)$ を計算する．

ブートストラップ確率を σ^2 の関数として表すパラメトリックモデルを用意する．

$$\mathrm{BP}(\sigma^2) = \bar{\Phi}(\sigma^{-1}\psi(\sigma^2; \beta_0, \ldots, \beta_{k-1}))$$

ただし $(\beta_0, \ldots, \beta_{k-1})$ は k 個のパラメータで，たとえば多項式

$$\psi_{\mathrm{poly}.k}(\sigma^2; \beta_0, \ldots, \beta_{k-1}) = \sum_{j=0}^{k-1} \beta_j \sigma^{2j} \tag{2.3}$$

が考えられる．このように与えたモデルはブートストラップ確率のスケーリング則を表す．

モデルのパラメータを推定するために k 個以上の $m > 0$ の値を定めてブートストラップを実行し，$\mathrm{BP}(\sigma^2)$ を測定する．この測定結果にモデル式を当てはめて推定したパラメータを $(\hat{\beta}_0, \ldots, \hat{\beta}_{k-1})$ とする．そして

$$\mathrm{AU}(\sigma^2) = \bar{\Phi}(\psi(\sigma^2; \hat{\beta}_0, \ldots, \hat{\beta}_{k-1}))$$

によって $\sigma^2 \le 0$ に $\mathrm{AU}(\sigma^2)$ を外挿する．多項式モデル (2.3) 式は問題無く $\sigma^2 = -1$ へ外挿できるが，ψ の定義によっては $\sigma^2 > 0$ のみで意味をもつ場合がある．そのような場合の外挿法は 6.8 節の (6.8) 式で与えている．

2.6 仮説の領域

次節以降の幾何的な議論の参考のため，二値関数 $f(D_n)$ を確率分布の空間の領域として説明しておく．まずデータ D_n を経験分布

$$\hat{q}_n(x) = \frac{1}{n}\sum_{t=1}^{n} \delta_{x_t}(x)$$

で表す．δ_{x_t} は $x = x_t$ の point mass である．ブートストラップは経験分布からのサンプリングであり，$x_1^*, x_2^*, \ldots, x_m^* \sim \hat{q}_n(x)$ と表される．分布関数に関して $\hat{q}_n \to q$ は一様収束するので，ブートストラップによって得られる分布は真の分布の近似といえる．そして

$$\hat{q}_m^* = \frac{1}{m}\sum_{t=1}^{m} \delta_{x_t^*}$$

は D_m^* を表し，$m \to \infty$ で $\hat{q}_m^* \to \hat{q}_n$ である．

確率分布から実数への連続な汎関数 $\theta(\cdot)$ とある定数 θ_1 に関して，$\theta(\hat{q}_n) \leq \theta_1$ のとき $f(D_n) = 1$，$\theta(\hat{q}_n) > \theta_1$ のとき $f(D_n) = 0$ と仮定する．確率分布の空間における領域

$$\mathcal{H} = \{q \mid \theta(q) \leq \theta_1\}$$

を定義すると，$\hat{q}_n \in \mathcal{H}$ のとき $f(D_n) = 1$，$\hat{q}_n \notin \mathcal{H}$ のとき $f(D_n) = 0$ である．$\hat{q}_m^{*(b)}, b = 1, \ldots, B$ は $\hat{q}_n$ の周りに分布していて，ブートストラップ確率は $\widehat{\mathrm{BP}}(\sigma^2) = \#\{\hat{q}_m^{*(b)} \in \mathcal{H}\}/B$ と表される．

興味のある帰無仮説は $q \in \mathcal{H}$ と表されるので，$\mathcal{H}$ を仮説の領域という．θ の真の値を $\theta_0 = \theta(q)$ と書けば，帰無仮説を $\theta_0 \leq \theta_1$ と書いてもよい．この検定の p-値を $p(D_n)$，有意水準を α とすれば，仮説 $\mathcal{H}$ を棄却する確率は $\beta(q) = P(p(D_n) < \alpha)$ である．不偏検定とは，$q \in \mathcal{H}$ のとき Type-I error が $\beta(q) \leq \alpha$, $q \notin \mathcal{H}$ のとき Type-II error が $1 - \beta(q) \leq 1 - \alpha$ を満たすことであるから，結果として真の分布が仮説領域の境界上 $\partial\mathcal{H} = \{q \mid \theta(q) = \theta_1\}$ のとき，言い換えると $\theta_0 = \theta_1$ のとき，$\beta(q) = \alpha$ である．したがって，任意の $0 < \alpha < 1$ で不偏検定になっていれば,

$$P(p(D_n) < \alpha) = \alpha, \quad 0 < \alpha < 1, \quad q \in \partial\mathcal{H}$$

つまり $p(D_n)$ は区間 (0,1) の一様分布に従う．

左辺は有意水準 α における棄却確率であるが, これと α との差をバイアスと呼ぶ．$n \to \infty$ の漸近理論においてバイアスが $O(n^{-k/2})$ のとき，k 次の精度の近似的に不偏な検定という．その検定の p-値を k 次の精度の近似的に不偏な p-値と呼ぶことにする．$q \notin \partial H$ の場合のブートストラップ確率を議論するため，q から ∂H への隔たりが $O(1/\sqrt{n})$ で近づく local alternatives の設定を用いる．

3. なぜサンプルサイズの符号を反転するとベイズが頻度論に変わるのか？

3.1 多変量正規モデルを仮定した漸近理論

Shimodaira (2002a) は多変量正規モデルを仮定してマルチスケール・ブートストラップを導出した．Efron *et al.* (1996) や Efron and Tibshirani (1998) で与えられた漸近理論を利用して幾何的な議論を行う．漸近理論は $O(n^{-1})$ の項まで計算して，$O(n^{-3/2})$ の項は誤差として無視する．

第 3 節を通して K 次元多変量正規モデル

$$Y \mid \mu \sim N(\mu, I_K) \tag{3.1}$$

を仮定する．ただし確率分布から $\mathbb{R}^K$ への汎関数 $g_n(\cdot)$ を用いて $Y = g_n(\hat{q}_n)$, $\mu = g_n(q)$ と変換したものが (3.1) 式を満たすと考える．確率変数 Y の実現値を y と書いて区別する．

仮説の領域 $\mathcal{H}$ は $\mathbb{R}^K$ の領域 $H = g_n(\mathcal{H})$ に変換され，$\mathcal{H} = g_n^{-1}(H)$ を仮定すれば，

$$\mu \in H$$

が帰無仮説である．第 2 節のブートストラップは $y^{*(b)} = g_n(\hat{q}_m^{*(b)})$, $b = 1, \ldots, B$ より $\widehat{\mathrm{BP}}(\sigma^2) = \#\{y^{*(b)} \in H\}/B$ と書ける．これをモデル化するためブートストラップ $Y^* = g_n(\hat{q}_m^*)$ は

$$Y^* \mid y \sim N(y, \sigma^2 I_K) \tag{3.2}$$

に従う確率変数と仮定して

$$\mathrm{BP}(\sigma^2) = P(Y^* \in H|y)$$

とする．(3.2) 式はパラメトリック・ブートストラップならば正当化されるが，D_n からリサンプリングした D_m^* の場合は近似であり，十分に n が大きいとき正当化される．

簡単な例を挙げる．q が多変量正規分布 $x_t \sim N(\eta, I_K)$ で，興味のある仮説が $\|\eta\| \leq 1$ とする．$\bar{x}_n$ の分散は I_K/n だから，$g_n(q) = \sqrt{n}\int xq(x)\,dx$ とすれば，$Y = \sqrt{n}\bar{x}_n$, $\mu = \sqrt{n}\eta$ が (3.1) を満たす．仮説の領域は $H = \{\mu \mid \|\mu\| \leq \sqrt{n}\}$ であり n に依存して変化する．

第 3 節では多変量正規モデルを仮定してマルチスケール・ブートストラップが近似的に不偏な検定の p-値を計算することを示す．しかし実際の BP や AU の計算において正規モデル (3.1) 式や (3.2) 式を直接利用するわけではない．BP や AU は第 2 節に述べた手順で計算するが，その性質を調べるために正規モデルを利用する．このような正規モデルへ変換する $g_n(\cdot)$ や適切な次元 K を実際に知る必要はなく $g_n(\cdot)$ や K の存在が仮定できればよい．この変換は必ずしも厳密である必要はなく，少なくとも近似的に正規モデルに変換できればマルチスケール・ブートストラップは有効に機能する．興味のあるパラメータの漸近正規性が十分によい近似を与える場合はこれに該当するだろう．

3.2 ブートストラップはベイズ

ベイズで考えると BP は H の事後確率と解釈できる．μ の事前分布は $\mathbb{R}^K$ の一様分布とする．(3.1) 式より $Y|\mu \sim const. \times \exp(-\|y-\mu\|^2/2)$ であるから，$\mu|y \sim const. \times \exp(-\|\mu - y\|^2/2)$ となり，μ の事後分布は

$$\mu|y \sim N(y, I_K) \tag{3.3}$$

である．(3.3) 式の μ の分布は，(3.2) 式の Y^* の分布で $\sigma^2 = 1$ としたものに等しい．したがって，

$$\mathrm{BP} = P(\mu \in H|y)$$

は $\mu \in H$ となる事後確率である．

3.3 仮説の境界を曲面で表す

仮説領域の境界 ∂H が滑らかな曲面と仮定する．∂H 上のある1点を原点とする座標系を考える．∂H の接平面の成分を $u = (u_1, u_2, \ldots, u_{K-1}) \in \mathbb{R}^{K-1}$，その法線方向の成分を $v \in \mathbb{R}$ として，$\mathbb{R}^K$ の点を (u, v) と表す．原点の近傍で仮説領域を

$$H = \{(u, v) \mid v \leq -h(u), \quad u \in \mathbb{R}^{K-1}\} \tag{3.4}$$

と書く．この領域の境界は曲面

$$\partial H = \{(u, v) \mid v = -h(u), \quad u \in \mathbb{R}^{K-1}\}$$

である．曲面の関数をテイラー展開すると（$\sum_{i,j}, \sum_{i,j,k}$ など省略して）

$$h(u) = d_{ij} u_i u_j + e_{ijk} u_i u_j u_k + O(n^{-3/2})$$

$$d_{ij} = \frac{1}{2} \frac{\partial^2 h(u)}{\partial u_i \partial u_j}\Big|_0, \quad e_{ijk} = \frac{1}{6} \frac{\partial^3 h(u)}{\partial u_i \partial u_j u_k}\Big|_0$$

と書ける．ただしひとつの項に同じ添え字が2回あるときは和を取る記法を採用する．h の2階微分は $d_{ij} = O(n^{-1/2})$，3階微分は $e_{ijk} = O(n^{-1})$ であり，一般に k 階微分は $O(n^{-(k-1)/2})$, $k \geq 2$ である．(3.1) 式を満たすように $g_n(\cdot)$ が $\sqrt{n}$ 倍のリスケーリングをする結果，k 階微分の分子の h が $\sqrt{n}$ 倍，分母の各 u_i 成分も $\sqrt{n}$ 倍されて，全体として $n^{1/2}/n^{k/2}$ 倍される．

3.4 距離と曲率が事後確率を決める

仮説領域の境界上に μ を制限した最尤推定は

$$\hat{\mu}(y) = \arg\max_{\mu \in \partial H} \ell(\mu; y)$$

である．(3.1) 式の対数尤度関数は $\ell(\mu; y) = -\|\mu - y\|^2/2$ であるから，$\hat{\mu}(y)$ は y に最も近い ∂H 上の点である．つまり $\hat{\mu}(y)$ は y の ∂H 上への射影である．ここで (u, v) 座標系の原点を $\hat{\mu}(y)$ に一致させた局所座標系を考える．とくに $y_1 = u_1, \ldots, y_{K-1} = u_{K-1}, y_K = v$ と座標軸を定めれば，この座標系では $\hat{\mu}(y) = (0, 0)$, $y = (0, v)$ である．$v = \pm\|y - \hat{\mu}(y)\|$ は符号付き距離と呼ばれ，符号は $y \in H$ のとき $v \leq 0$，$y \notin H$ のとき $v > 0$ である．

Efron and Tibshirani (1998) の (2.19) 式より次式が得られる（$\sum_i$, $\sum_{i,j}$ を省略している）．

$$\mathrm{BP} = \bar{\Phi}(v + d_{ii} - v d_{ij} d_{ij}) + O_p(n^{-3/2}) = \bar{\Phi}(v + d_{ii}) + O_p(n^{-1}) \tag{3.5}$$

y が H から遠ざかるほど符号付き距離 v が大きくなり BP は小さくなる．d_{ii} は ∂H の $\hat{\mu}(y)$ における曲率を反映していて，∂H が H 自身の方向に曲がって H の体積が減少すると d_{ii} は大きくなり BP が小さくなる．たとえば H が凸なら $d_{ii} > 0$ である．対称行列 $(d_{ij}) = D$

とおき，その固有値を $\lambda_1,\dots,\lambda_{K-1}$ とすると，$d_{ii} = \mathrm{tr}(D) = \sum_{i=1}^{K-1}\lambda_i = O_p(n^{-1/2})$, $d_{ij}d_{ij} = \mathrm{tr}(D^2) = \sum_{i=1}^{K-1}\lambda_i^2 = O_p(n^{-1})$ である．d_{ii} は $K-1$ 個の固有ベクトル方向の曲率の和になるので平均曲率と呼ばれる．

もし ∂H が平坦なら $(Y_1^*,\dots,Y_{K-1}^*)$ は BP に無関係になり $Y_K^* \sim N(v,1)$ の 1 次元を考えればよいので，$\mathrm{BP} = P(Y_K^* \le 0|y) = \bar\Phi(v)$ である．たしかに (3.5) 式に $D=0$ を代入すれば $\mathrm{BP} = \bar\Phi(v) + O_p(n^{-3/2})$ が得られる．

もし $y \in \partial H$ の場合，$y = \hat\mu(y)$ だから $v=0$ である．このときの BP を BP_0 と書く．(3.5) 式に $v=0$ を代入すると

$$\mathrm{BP}_0 = \bar\Phi(d_{ii}) + O_p(n^{-3/2}) \tag{3.6}$$

であるから，平均曲率は $d_{ii} = \bar\Phi^{-1}(\mathrm{BP}_0) + O_p(n^{-3/2})$ と書けて，BP_0 から d_{ii} が計算できる．一般に y が与えられたとき，d_{ii} を計算するには (3.2) 式ではなく

$$Y^* \mid y \sim N(\hat\mu(y), I_K) \tag{3.7}$$

から Y^* を生成してブートストラップ確率を計算すれば BP_0 が得られる．

3.5　頻度論の p-値も距離と曲率で決まる

符号付き距離を $v(y)$ と表す．この値が大きくなるほど y が H から遠ざかるから，仮説 $\mu \in H$ が正しい可能性は小さくなるだろう．そこで頻度論的な信頼度を

$$\mathrm{PV} = P\left(v(Y^*) \ge v(y) \mid \hat\mu(y)\right) \tag{3.8}$$

と定める．ただし Y^* は (3.2) 式ではなく (3.7) 式から生成して $v(Y^*)$ の帰無分布における確率を計算している．この PV は 3 次の精度の近似的に不偏な p-値になる．つまり

$$P\left(\mathrm{PV} < \alpha \mid \mu\right) = \alpha + O(n^{-3/2}), \quad 0<\alpha<1, \quad \mu \in \partial H$$

を満たす．そして Efron (1985) から導かれる Efron and Tibshirani (1998) の (2.17) 式と PV は $O_p(n^{-3/2})$ の誤差で等価であり，漸近展開は次式で与えられる．

$$\mathrm{PV} = \bar\Phi(v - d_{ii} + v d_{ij} d_{ij}) + O_p(n^{-3/2}) = \bar\Phi(v - d_{ii}) + O_p(n^{-1}) \tag{3.9}$$

検定の Type-I error を厳密にコントロールするためには，帰無仮説が棄却しにくくなる最悪ケースの帰無分布を考えて

$$p(y) = \sup_{\mu \in \partial H} P(v(Y) \ge v(y) \mid \mu) \tag{3.10}$$

と p-値を定義すべきである．例えば多重比較法は H が凸錐の場合に相当して，sup は μ が H の頂点で得られる (ただし統計量 $v(y)$ の定義を変更する)．しかし最悪ケース以外の一般

的な状況を考えると，(3.10) 式は Type-II error が大きくなり保守的な結果 (false negative) が得られやすい．これに対して，(3.8) 式では y に最も近い $\hat{\mu}(y)$ を使って帰無分布を与えることにより Type-I error と Type-II error の両者を近似的にコントロールする．

3.6　曲率の反転でベイズが頻度論に

ところで係数 β_0, β_1 を

$$\beta_0 = v = O_p(1), \quad \beta_1 = d_{ii} - v d_{ij} d_{ij} = O_p(n^{-1/2})$$

と定義すれば，(3.5) 式と (3.9) 式は

$$\mathrm{BP} = \bar{\Phi}(\beta_0 + \beta_1) + O_p(n^{-3/2}), \quad \mathrm{PV} = \bar{\Phi}(\beta_0 - \beta_1) + O_p(n^{-3/2}) \tag{3.11}$$

である．もし $\beta_1 > 0$ なら BP $<$ PV だから $P(\mathrm{BP} < \alpha) \geq P(\mathrm{PV} < \alpha)$ となる．つまり BP は Type-I error をコントロールせず，仮説は棄却されやすくなる (false positive).

BP と PV の違いは β_1 の前の符号だけである．H の形を変更して $\hat{\mu}(y)$ を中心に ∂H を裏返した領域を $\bar{H} = \{(u,v) \mid v \leq h(u),\, u \in \mathbb{R}^{K-1}\}$ と定義すると $\partial \bar{H}$ の曲率は $\bar{D} = -D$ となるので，$\bar{H}$ の BP は

$$\overline{\mathrm{BP}} = \bar{\Phi}(v - d_{ii} - v d_{ij} d_{ij}) + O_p(n^{-3/2}) = \mathrm{PV} + O_p(n^{-1})$$

である．つまり曲率を反転すると $O_p(n^{-1})$ の誤差でベイズの BP が頻度論の PV になる．

Efron and Tibshirani (1998) は BP_0 を利用して BP のバイアス修正を試みた．

$$\mathrm{BC} = \bar{\Phi}\left\{\bar{\Phi}^{-1}(\mathrm{BP}) - 2\bar{\Phi}^{-1}(\mathrm{BP}_0)\right\} \tag{3.12}$$

Efron (1987) では，このアイデアを連続値パラメータのブートストラップ信頼区間に適用したものを bias-corrected bootstrap と呼んでいる．(3.5) 式と (3.6) 式より

$$\mathrm{BC} = \bar{\Phi}(v + d_{ii} - v d_{ij} d_{ij} - 2 d_{ii}) + O_p(n^{-3/2}) = \overline{\mathrm{BP}} + O_p(n^{-3/2})$$

と表される．BC は $O_p(n^{-3/2})$ の誤差で $\overline{\mathrm{BP}}$ に等価になり，$\mathrm{BC} = \mathrm{PV} + O_p(n^{-1})$ は 2 次の精度の近似的に不偏な p-値になる．BC の実装ではいかに (3.7) 式に相当するリサンプリングを実現するかが問題になる．

3.7　負のサンプルサイズで曲率が反転

Shimodaira (2002a) は $\sigma^2 = 1$ の BP である (3.5) 式を一般化して任意の $\sigma^2 > 0$ において

$$\mathrm{BP}(\sigma^2) = \bar{\Phi}(\beta_0 \sigma^{-1} + \beta_1 \sigma) + O_p(n^{-3/2}) \tag{3.13}$$

を与えた．これはスケーリング則，つまりスケール σ を変化させたときのブートストラップ確率の変化を表している．幾何的に考えると (3.13) 式を容易に導出できる．Y^* と仮説

H を含めた $\mathbb{R}^K$ 全体を σ^{-1} 倍しても事象は変化せず $\mathrm{BP}(\sigma^2)$ も変化しない．ところが Y^* の分布 (3.2) 式は

$$Y^*/\sigma \sim N(\mu/\sigma, I_K)$$

に変化する．分散をみると $\sigma^2 = 1$ だからブートストラップ確率の式は (3.5) 式が利用できる．次に述べるように，この σ^{-1} 倍するリスケーリングによって β_0 は $\beta_0\sigma^{-1}$，β_1 は $\beta_1\sigma$ に置き換わるので，(3.11) 式にこれらを代入すると (3.13) 式が得られる．

リスケーリングは長さに相当する量をすべて σ^{-1} 倍する．v は σ^{-1} 倍されるので，β_0 は $\beta_0\sigma^{-1}$ で置き換わる．d_{ij} は h の 2 階微分であったから，分子の h が σ^{-1} 倍，分母の各 u_i 成分も σ^{-1} 倍されて，全体として $\sigma^{-1}/(\sigma^{-1})^2 = \sigma$ 倍される．したがって，d_{ii} は $d_{ii}\sigma$，$vd_{ij}d_{ij}$ は $(v/\sigma)(d_{ij}\sigma)(d_{ij}\sigma) = vd_{ij}d_{ij}\sigma$ に置き換わり，結果として β_1 は σ 倍される．

正規化ブートストラップ確率 $\mathrm{AU}(\sigma^2)$ の定義 (2.2) 式に (3.13) 式を代入すると正規化ブートストラップ確率のスケーリング則が得られる．

$$\mathrm{AU}(\sigma^2) = \bar{\Phi}(\beta_0 + \beta_1\sigma^2) + O_p(n^{-3/2}) \tag{3.14}$$

これを BP として解釈すると，符号付き距離は β_0 に保ったままで，曲率 β_1 を σ^2 倍したときのブートストラップ確率である．したがって曲率を反転するには $\sigma^2 = -1$ とおけばよい．つまり $\mathrm{AU} = \mathrm{AU}(-1)$ は

$$\mathrm{AU} = \bar{\Phi}(\beta_0 - \beta_1) + O_p(n^{-3/2}) = \mathrm{PV} + O_p(n^{-3/2})$$

となって，AU が 3 次の精度の近似的に不偏な p-値であることが分かった．

4. 指数型分布族を仮定した漸近理論

4.1 正規モデルからのズレを表す加速定数

Shimodaira (2004) は指数型分布族を仮定してマルチスケール・ブートストラップをさらに修正する方法を議論した．K 次元の確率ベクトル $Y = g_n(\hat{q}_n)$ の確率分布は多変量正規モデル (3.1) 式を一般化してポテンシャル関数 $\varphi(\mu)$ の指数型分布族を仮定する．ただし $\mu = E(Y)$ は期待値パラメータである．たとえば多変量正規モデルのポテンシャル関数は $\varphi(\mu) = \|\mu\|^2/2$ である．原点においてこの正規モデルと同じ平均と分散を持つように座標変換して $\partial\varphi/\partial\mu_i = 0$, $(\partial^2\varphi/\partial\mu_i\partial\mu_j) = I_K$ と仮定しても一般性を失わない．ブートストラップ $Y^* = g_n(\hat{q}^*_m)$ は (3.2) 式の代わりにポテンシャル関数 $\varphi(y)/\sigma^2$ の指数型分布族を仮定する．これらの分布を便宜的に

$$Y|\mu \sim \varphi(\mu, 1), \quad Y^*|y \sim \varphi(y, \sigma^2)$$

と書くことにする．なお，指数型分布族の密度関数は $Y \sim c(y)\exp(\nu_i y_i - \kappa(\nu))$ の形に書ける．$\nu = \partial\varphi(\mu)/\partial\mu$ は自然パラメータという．キュムラント関数 $\kappa(\nu)$ はルジャンドル変換 $\kappa(\nu) = \max_\mu\{\mu_i\nu_i - \varphi(\mu)\}$ によって与えられる．

Shimodaira (2004) の漸近理論は誤差が $O(n^{-3/2})$ だが, ここでは簡単のため誤差が $O(n^{-1})$ の議論をする．まずスケーリング則 (3.14) 式は次のように修正される．

$$\mathrm{AU}(\sigma^2) = \bar\Phi(\beta_0 - 2a\beta_0^2 + (\beta_1 - a)\sigma^2) + O_p(n^{-1}) \tag{4.1}$$

ここで $a = O_p(n^{-1/2})$ は正規モデルからのズレを表す量で,

$$a = -\frac{1}{6}\frac{\partial^3\varphi(\mu)}{\partial\mu_K^3}\Big|_{\hat\mu}$$

と定義され, とくに正規モデルでは $a = 0$ である．Efron (1987) は a を加速定数 (acceleration constant) と呼んだ．Shimodaira (2004) の (7.15) 式は a の他に $(\partial^3\varphi/\partial\mu_i\partial\mu_K^2)^2$, $\partial^4\varphi/\partial\mu_K^4$ を使って (4.1) 式を $O_p(n^{-3/2})$ の誤差で求めた結果であり，これら正規分布からのズレを表す量をゼロにすれば (3.14) 式に一致する．第 2 節のマルチスケール・ブートストラップは $\mathrm{AU} = \mathrm{AU}(-1)$ を計算する．したがって，(4.1) 式に $\sigma^2 = -1$ を代入すると

$$\mathrm{AU} = \bar\Phi(\beta_0 - \beta_1 + a(1 - 2\beta_0^2)) + O_p(n^{-1})$$

が p-値として計算される．

4.2 マルチスケール・ブートストラップは 1 次の精度しかない

正規モデルの時と同様に (3.8) 式で PV を定義すると，指数型分布属でも 3 次の精度の近似的に不偏な p-値になる．Shimodaira (2004) の (7.13) 式は PV の漸近展開を $O_p(n^{-3/2})$ の誤差で求めた結果であり，3 次の精度の近似的に不偏な p-値は必ずその形に書けることが示されている．これを $O_p(n^{-1})$ の誤差で書くと

$$\mathrm{PV} = \bar\Phi(\beta_0 - \beta_1 + a(1 - \beta_0^2)) + O_p(n^{-1})$$

となり，2 次の精度の近似的に不偏な p-値はこの形に書ける．

マルチスケール・ブートストラップが計算する $\mathrm{AU} = \mathrm{AU}(-1)$ との差を求めると

$$\bar\Phi^{-1}(\mathrm{PV}) - \bar\Phi^{-1}(\mathrm{AU}) = a\beta_0^2 + O_p(n^{-1}) = O_p(n^{-1/2}) \tag{4.2}$$

である．$a = 0$ でなければ一般にバイアスが $O(n^{-1/2})$ であるから，AU は 1 次の精度の近似的に不偏な p-値でしかない．通常のブートストラップ確率 $\mathrm{BP} = \mathrm{AU}(1)$ は

$$\mathrm{BP} = \bar\Phi(\beta_0 + \beta_1 - a(1 + 2\beta_0^2)) + O_p(n^{-1}) \tag{4.3}$$

であるから PV との差は

$$\bar{\Phi}^{-1}(\mathrm{PV}) - \bar{\Phi}^{-1}(\mathrm{BP}) = -2\beta_1 + a(2+\beta_0^2) + O_p(n^{-1}) = O_p(n^{-1/2})$$

であり，BP も 1 次の精度である．AU と BP のどちらも 1 次の精度であるが，PV からのズレの内容をみると，AU は曲率 β_1 の影響を完全に修正している点で BP より望ましい．

4.3　ブートストラップの反復は計算量が大きい

Hall (1992) はブートストラップを反復して高次項を修正していく原理を述べている．Efron and Tibshirani (1998) は Remark A で BP にこの原理を適用した．ここでは BP が y の関数であることを明示するため $p(y)$ と書き，これを修正して $p'(y)$ を計算する．まずブートストラップは $E(Y^*|y) = y$ ではなく $E(Y^*|y) = \hat{\mu}(y)$ となるように

$$Y^*|y \sim \varphi(\hat{\mu}(y), 1) \tag{4.4}$$

から Y^* を生成する．正規モデルで言えば (3.2) 式ではなく (3.7) 式から生成することに相当する．(4.4) 式から B 個の $y^{*(b)}$, $b = 1, \ldots, B$ を生成して

$$p'(y) = P(p(Y^*) > p(y) \mid \hat{\mu}(y)) \approx \#\{p(y^{*(b)}) > p(y)\}/B$$

を計算する．各 $p(y^{*(b)})$ の計算には B 個の $Y^{**} \sim \varphi(y^{*(b)}, 1)$ を生成するから，全体では B^2 個の Y^{**} が生成され $Y^{**} \in H$ の判定を行う．このようにブートストラップによって BP を修正する方法をダブルブートストラップとも言う．正規モデルのとき $p'(y)$ は 3 次の精度の近似的に不偏な検定を与えることを Efron and Tibshirani (1998) が述べている．じつは指数型分布族でも 3 次の精度であることを Shimodaira (2004) は示した．

ダブルブートストラップの p' に再び上記の修正を適用すれば，さらに精度の向上が期待できる．k 回反復ブートストラップでは，$k = 1$ が最初の BP, それから得られた p'（つまりダブルブートストラップ）が $k = 2$ に相当する．計算量は B^k に比例するので，通常 $k = 2$ でもかなり厳しい．

4.4　加速定数がわかれば BP を補正できる

Efron (1987) は実数パラメータのブートストラップ信頼区間の理論を示した．これを利用して Efron and Tibshirani (1998) は BP の修正を次式で与えた．

$$\mathrm{ABC} = \bar{\Phi}\left\{\bar{\Phi}^{-1}(\mathrm{BP}) - 2\bar{\Phi}^{-1}(\mathrm{BP}_0) + a(\bar{\Phi}^{-1}(\mathrm{BP}))^2\right\} \tag{4.5}$$

これは正規モデルの (3.12) 式を指数型分布族へ一般化したものである．(3.7) 式から Y^* を生成して BP_0 を計算した時と同様に，(4.4) 式から Y^* を生成して計算した BP を BP_0 と書く．(4.3) 式に $\beta_0 = 0$ を代入すれば

$$\mathrm{BP}_0 = \bar{\Phi}(\beta_1 - a) + O_p(n^{-1})$$

が得られる．これと (4.3) 式を (4.5) 式に代入すると

$$\mathrm{ABC} = \bar{\Phi}(\beta_0 + \beta_1 - a(1+2\beta_0^2) - 2(\beta_1 - a) + a\beta_0^2) + O_p(n^{-1}) = \mathrm{PV} + O_p(n^{-1})$$

つまり ABC は 2 次の精度の近似的に不偏な p-値である．ABC において加速定数 a を用いた補正項の重要性を確かめるため，正規モデルのための (3.12) 式を計算してみると

$$\mathrm{BC} = \bar{\Phi}(\beta_0 + \beta_1 - a(1+2\beta_0^2) - 2(\beta_1 - a)) + O_p(n^{-1}) = \mathrm{AU} + O_p(n^{-1})$$

になってしまう．

問題は加速定数 a の計算方法である．Efron and Tibshirani (1998) の (6.9) または (7.13) にあるように ∂H の法線方向に関する情報が必要になり，何らかの形でモデルを特定する必要がある．計算が比較的容易な Efron *et al.* (1996) の方法はヒューリスティックで近似が良いとはいえない．

4.5 ブートストラップのステップで加速定数を計算

ABC はモデルの特定や $\hat{\mu}(y)$ の計算が応用上の制限になり好ましくない．これらの必要なく，ブートストラップだけで BP を修正する方法を Shimodaira (2004) が与えた．まず D_n から m_1 個の要素をランダムに取り出して $D^*_{m_1}$ とする．次に $D^*_{m_1}$ から m_2 個の要素をランダムに取り出して $D^{**}_{m_2}$ とする．このブートストラップの各段階をステップと呼び，とくに 2 段階の方法を 2 ステップ・マルチスケール・ブートストラップと呼ぶ．これを指数型分布族で次のようにモデル化する．

$$Y|\mu \sim \varphi(\mu, 1), \quad Y^*|y \sim \varphi(y, \sigma_1^2), \quad Y^{**}|y^* \sim \varphi(y^*, \sigma_2^2) \tag{4.6}$$

$$\sigma_1^2 = \frac{n}{m_1}, \quad \sigma_2^2 = \frac{n}{m_2}$$

そして，2 ステップ・ブートストラップ確率を

$$\mathrm{BP}(\sigma_1^2, \sigma_2^2) = P(Y^{**} \in H \mid y)$$

と定義する．この計算はダブルブートストラップに似ているが計算量は $O(B^2)$ ではなく $O(B)$ しかない．ダブルブートストラップでは各 y^* から生成する Y^{**} は B 個であったが，2 ステップ法では 1 個だけしか生成しない．

もし多変量正規モデルが仮定できるなら (4.6) 式は

$$Y|\mu \sim N(\mu, I_K), \quad Y^*|y \sim N(y, \sigma_1^2 I_K), \quad Y^{**}|y^* \sim N(y^*, \sigma_2^2 I_K)$$

となる．2 段階のブートストラップをまとめて考えると，分散共分散行列を足しあわせて，$Y^{**}|y \sim N(y, (\sigma_1^2 + \sigma_2^2) I_K)$ となる．これは (3.2) 式で $\sigma^2 = \sigma_1^2 + \sigma_2^2$ としたものだから，

$$\mathrm{BP}(\sigma_1^2, \sigma_2^2) = \mathrm{BP}(\sigma_1^2 + \sigma_2^2)$$

である．しかし一般の指数型分布族ならば，正規モデルからのズレが次のように検出できる

$$\bar{\Phi}^{-1}(\mathrm{BP}(\sigma_1^2,\sigma_2^2)) - \bar{\Phi}^{-1}(\mathrm{BP}(\sigma_1^2+\sigma_2^2)) = \frac{a\sigma_1^2\sigma_2^2(\beta_0^2-(\sigma_1^2+\sigma_2^2))}{(\sigma_1^2+\sigma_2^2)^{5/2}} \tag{4.7}$$

したがって 3 個以上の (σ_1^2,σ_2^2) の組み合わせで $BP(\sigma_1^2,\sigma_2^2)$ を測定し，(4.1) 式と (4.7) 式を当てはめることにより β_0, β_1, a が推定できる．

計算の都合を考えて $s_1=(\sigma_1^2+\sigma_2^2)^{-1/2}$, $s_2=\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2)^{-2}$, $\gamma_1=\beta_0-2a\beta_0^2$, $\gamma_2=\beta_0(a-\beta_1)$, $\gamma_3=\beta_0 a$ と定義すると

$$\bar{\Phi}^{-1}(\mathrm{BP}(\sigma_1^2,\sigma_2^2)) = s_1\gamma_1(1+s_2\gamma_3) - \frac{\gamma_2+s_2\gamma_3}{s_1\gamma_1}$$

と書ける．これを $\mathrm{BP}(\sigma_1^2,\sigma_2^2)$ の測定値に当てはめて，γ_1, γ_2, γ_3 を推定する．そして

$$\mathrm{PV} = \bar{\Phi}\left\{\gamma_1(1+\gamma_3)+\frac{\gamma_2}{\gamma_1}\right\} + O_p(n^{-1})$$

によって 2 次の精度の近似的に不偏な p-値を計算する．Shimodaira (2004) はさらに 1 段階増やした 3 ステップ・マルチスケール・ブートストラップも与えた．$\gamma_1,\gamma_2,\gamma_3$ に加えてさらに $O(n^{-1})$ の係数を 3 個推定することにより，3 次の精度の p-値が計算できる．

5. 滑らかな関数モデルを仮定した漸近理論

5.1 パラメータ推定量の分布関数を漸近展開する

実数パラメータ $\theta \in \mathbb{R}$ の推定量 $\hat{\theta}_n$ の漸近正規性を仮定する．$n\to\infty$ の極限で，平均と分散が $E(\hat{\theta}_n)\to\theta_0$, $E(n(\hat{\theta}_n-\theta_0)^2)\to\tau^2$ とすれば，$\hat{\theta}_n$ を標準化した

$$S_n = \sqrt{n}\,\frac{\hat{\theta}_n-\theta_0}{\tau}$$

の漸近分布は $N(0,1)$ である．モーメントが $E(S_n) = n^{-1/2}k_1 + O(n^{-3/2})$, $E((S_n - E(S_n))^2) = 1 + n^{-1}k_2 + O(n^{-2})$, $E((S_n - E(S_n))^3) = n^{-1/2}k_3 + O(n^{-3/2})$ と書けるとき，$p_1(s) = -(k_1+k_3(s^2-1)/6)$ を定義する．S_n の分布関数のエッジワース展開は

$$P(S_n\le s) = \Phi(s) + n^{-1/2}p_1(s)\phi(s) + O(n^{-1}) = \Phi(s+n^{-1/2}p_1(s)+O(n^{-1})) \tag{5.1}$$

である．(5.1) 式の二つ目の等号はコーニッシュ・フィッシャー展開に相当するが，s の周りで右辺をテイラー展開するだけで等号が確認できる．この右辺を以下で用いる．

この漸近展開が正当化される状況として，Hall (1992) は滑らかな関数モデル (smooth function model) を考えた．D_n の要素 x_t, $t=1,\ldots,n$ の次元は d とする．x_t の平均と期待値を $\bar{x}_n = n^{-1}\sum_{t=1}^n x_t$, $\eta = E(x_t)$ と書く．滑らかな関数 $A:\mathbb{R}^d\to\mathbb{R}$ で $A(\eta)=0$ を満たすものを考える．一般に，$\lim_{n\to\infty} E(nA(\bar{x}_n)^2) = \tau^2$ とおいて，

$$S_n = \sqrt{n}\,A(\bar{x}_n)\tau^{-1}$$

を定義する．とくに $\hat{\theta}_n = \theta(\bar{x}_n)$, $\theta_0 = \theta(\eta)$ とおけば先ほどの S_n は $A(\bar{x}_n) = \theta(\bar{x}_n) - \theta(\eta)$ で与えられる．ただし 2.6 節の $\theta(\cdot)$ の定義は q の汎関数だったが，ここでは期待値パラメータ $\eta = \int xq(x)\,dx$ だけを通して値が定まるので $\theta(\eta)$ と書いた．

Hall (1992) の Theorem 2.2 によれば (5.1) 式の $p_1(s)$ は次式で与えられる．

$$p_1(s) = -(A_1\tau^{-1} + \tfrac{1}{6}A_2\tau^{-3}(s^2-1)) = p_1(0) - \tfrac{1}{6}A_2\tau^{-3}s^2 \tag{5.2}$$

ただし，定数 A_1, A_2 は（各項に同じ添え字が２回出てきたら Σ を省略する記法で）

$$A_1 = \tfrac{1}{2}a_{ij}\eta_{ij}, \quad A_2 = a_i a_j a_k \eta_{ijk} + 3a_i a_j a_{kl}\eta_{ik}\eta_{jl}$$

添え字は $1 \le i,j,k,l \le d$ の範囲を動く．a_{ij}, η_{ij} 等の項は次のように与えられる．まず $x, \eta \in \mathbb{R}^d$ の第 i 成分を $x^{(i)}$, $\eta^{(i)}$ と書く．そして $j \ge 1$, $i_1,\ldots,i_j \in \{1,2,\ldots,d\}$ に対して

$$a_{i_1,\ldots,i_j} = \left.\frac{\partial^j A(x)}{\partial x^{(i_1)}\cdots x^{(i_j)}}\right|_{x=\eta}, \quad \eta_{i_1,\ldots,i_j} = E[(x^{(i_1)} - \eta^{(i_1)})\cdots(x^{(i_j)} - \eta^{(i_j)})]$$

5.2 n を m で置き換えて，ブートストラップ分布の漸近展開を求める

ブートストラップ D_m^* の要素 x_t^*, $t = 1,\ldots,m$ の平均を $\bar{x}_m^* = m^{-1}\sum_{t=1}^m x_t^*$ と書く．パラメータ推定値 $\hat{\theta}_m^* = \theta(\bar{x}_m^*)$ を標準化した

$$S_m^* = \sqrt{m}\,\frac{\hat{\theta}_m^* - \hat{\theta}_n}{\hat{\tau}_n}$$

の漸近分布を考える．ただし $\hat{\tau}_n = \tau + O_p(n^{-1/2})$ は τ の一致推定量で，$\tau : \mathbb{R}^d \to \mathbb{R}$ を用いて $\hat{\tau}_n = \tau(\bar{x}_n)$, $\tau = \tau(\eta)$ と書けると仮定する．Hall (1992) の (3.17) 式より，とくに $m = n$ の時の漸近分布は (5.1) 式の $p_1(s)$ を推定量 $\hat{p}_1(s) = p_1(s) + O_p(n^{-1/2})$ で置き換えた

$$P(S_n^* \le s|D_n) = \Phi(s + n^{-1/2}\hat{p}_1(s) + O_p(n^{-1}))$$

である．一般の $m = O(n)$ では，$n^{-1/2}$ を $m^{-1/2}$ で置き換えて

$$P(S_m^* \le s|D_n) = \Phi(s + m^{-1/2}p_1(s) + O_p(n^{-1})) \tag{5.3}$$

とすればよい．ただし $m^{-1/2}\hat{p}_1(s)$ の代わりに $m^{-1/2}p_1(s)$ を使っても，その差は $O_p(n^{-1})$ の誤差として無視できる．この結果より $\hat{\theta}_m^*$ のブートストラップ分布

$$\hat{F}_m(\theta) = P(\hat{\theta}_m^* \le \theta \mid D_n)$$

が直ちに得られる．$s = \sqrt{m}\,(\theta - \hat{\theta}_n)/\hat{\tau}_n$ とおくと $\hat{F}_m(\theta) = P(S_m^* \le s|D_n)$ である．ただし $s = O_p(1)$ となるように $\theta = \theta_0 + O(n^{-1/2})$ の local alternatives を仮定する．(5.2) 式を (5.3) 式に代入すると，ブートストラップ分布の漸近展開が得られる．

$$\hat{F}_m(\theta) = \bar{\Phi}\left[\sqrt{m}\,\frac{\hat{\theta}_n - \theta}{\hat{\tau}_n} + \frac{1}{\sqrt{m}}\left(\frac{A_1}{\tau} - \frac{A_2}{6\tau^3}\right) + \frac{1}{\sqrt{m}}\frac{A_2}{6\tau^3}\left(\sqrt{m}\,\frac{\hat{\theta}_n - \theta}{\hat{\tau}_n}\right)^2\right] + O_p(n^{-1})$$

5.3 ブートストラップ確率のスケーリング則を導出

ある定数 θ_1 を指定して実数の集合 $H = \{\theta \mid \theta \le \theta_1\}$ を考える．未知パラメータ θ_0 に関する帰無仮説は，$\theta_0 \in H$ である．ここでは $\theta_1 = \theta_0 + O(n^{-1/2})$ を仮定して，この仮説のブートストラップ確率の漸近展開を与える．

$$t_n = \sqrt{n}\,\frac{\hat{\theta}_n - \theta_1}{\hat{\tau}_n}$$

とおいて $\mathrm{BP}(\sigma^2) = \hat{F}_m(\theta_1)$ を整理すると

$$\mathrm{BP}(\sigma^2) = \bar{\Phi}\left[\sigma^{-1}\left\{t_n - n^{-1/2}\sigma^2 p_1(0) + n^{-1/2}(A_2\tau^{-3}/6)t_n^2\right\}\right] + O_p(n^{-1})$$

また $\mathrm{AU}(\sigma^2) = \bar{\Phi}(\sigma\bar{\Phi}^{-1}(\mathrm{BP}(\sigma^2)))$ より

$$\mathrm{AU}(\sigma^2) = \bar{\Phi}\left[t_n - n^{-1/2}\sigma^2 p_1(0) + n^{-1/2}(A_2\tau^{-3}/6)t_n^2\right] + O_p(n^{-1}) \tag{5.4}$$

5.4 信頼区間の被覆確率に関する一般的な結果を確認しておく

Hall (1992) の Proposition 3.1 は片側信頼区間の被覆確率を与える．$z_\alpha = \Phi^{-1}(\alpha)$ とおいて，θ_0 の片側信頼区間

$$\hat{\mathcal{I}}_n = \left[\hat{\theta}_n + n^{-1/2}\hat{\tau}_n z_\alpha + n^{-1}\hat{\tau}_n\hat{s}_1(z_\alpha) + O_p(n^{-3/2}), \infty\right) \tag{5.5}$$

を考える．ここで $\hat{s}_1(z_\alpha) = s_1(z_\alpha) + O_p(n^{-1/2})$ は信頼区間を特徴付ける多項式である．このとき，信頼区間が真のパラメータ値を含む被覆確率は

$$P(\theta_0 \in \hat{\mathcal{I}}_n|\theta_0) = 1 - \alpha + n^{-1/2}(q_1(z_\alpha) - s_1(z_\alpha))\phi(z_\alpha) + O(n^{-1}) \tag{5.6}$$

である．多項式 $q_1(z_\alpha)$ は後ほど述べる．被覆確率のバイアス $= (1-\alpha)-(5.6)$ 式だから

$$-n^{-1/2}(q_1(z_\alpha) - s_1(z_\alpha))\phi(z_\alpha) + O(n^{-1})$$

の大きさが小さいほど良いと考える．

多項式 $q_1(z_\alpha)$ は次のように定める．$T_n = \sqrt{n}\,(\hat{\theta}_n - \theta_0)/\hat{\tau}_n$ の分布関数を

$$P(T_n \le s) = \Phi(s) + n^{-1/2}q_1(s)\phi(s) + O(n^{-1}) = \Phi(s + n^{-1/2}q_1(s) + O(n^{-1}))$$

と書いたときに, (5.1) 式の $p_1(s)$ に相当する多項式が $q_1(s)$ である．滑らかな関数 $B : \mathbb{R}^d \to \mathbb{R}$ を $B(x) = A(x)/\tau(x)$ と定めれば $T_n = \sqrt{n}\,B(\bar{x}_n)$ であるから，$A(\bar{x}_n)\tau^{-1}$ の代わりに $B(\bar{x}_n)$ へ (5.2) 式を適用すれば $q_1(s)$ の具体的な表現が得られる．

$$q_1(s) = -(B_1 + \tfrac{1}{6}B_2(s^2-1)) = q_1(0) - \tfrac{1}{6}B_2 s^2$$

各成分は $B_1 = b_{ij}\eta_{ij}/2$, $B_2 = b_i b_j b_k \eta_{ijk} + 3 b_i b_j b_{kl} \eta_{ik} \eta_{jl}$ である．ただし

$$b_{i_1,\dots,i_j} = \frac{\partial^j B(x)}{\partial x^{(i_1)} \cdots x^{(i_j)}}\Big|_{x=\eta}, \quad c_i = \frac{\partial(\tau^2)}{\partial x^{(i)}}\Big|_{x=\eta}$$

と定義しておく．$A(x)$ と $B(x)$ の関係を整理すると $a_i = b_i \tau$, $a_{ij} = b_{ij}\tau + (a_i c_j + a_j c_i)\tau^{-2}/2$, $B_1 = A_1 \tau^{-1} - a_i c_j \eta_{ij} \tau^{-3}/2$, $B_2 = (A_2 - 3 a_i c_j \eta_{ij})\tau^{-3}$, $B_1 - \frac{1}{6}B_2 = A_1\tau^{-1} - \frac{1}{6}A_2\tau^{-3} = -p_1(0) = -q_1(0)$ である．

5.5 bootstrap-t 信頼区間に等価な p-値を求める

T_n のブートストラップ $T_n^* = \sqrt{n}\,(\hat{\theta}_n^* - \hat{\theta}_n)/\hat{\tau}_n^* = \sqrt{n}\,B(\bar{x}_n^*)$ により

$$P(T_n^* > \hat{c}_\alpha | D_n) = \alpha$$

を満たすように $\hat{c}_\alpha$ を定義する．これより bootstrap-t 信頼区間

$$\hat{\mathcal{I}}_n = \left[\hat{\theta}_n - n^{-1/2}\hat{\tau}_n \hat{c}_\alpha, \infty\right)$$

が定義される．この信頼区間が θ_1 を含まないとき仮説 H は棄却される．仮説 H を検定するための p-値として

$$\mathrm{PV} = P(T_n^* > t_n | D_n)$$

を定義すると，$\theta_1 \notin \hat{\mathcal{I}}_n \Leftrightarrow \theta_1 < \hat{\theta}_n - n^{-1/2}\hat{\tau}_n\hat{c}_\alpha \Leftrightarrow t_n > \hat{c}_\alpha \Leftrightarrow \mathrm{PV} < \alpha$ だから，$\hat{\mathcal{I}}_n$ と PV は等価になる．(5.1) 式から (5.3) 式を導出したときと同様に $P(T_n^* \le t_n | D_n) = P(T_n \le t_n) + O_p(n^{-1})$ だから PV の漸近展開は

$$\begin{aligned}\mathrm{PV} &= \bar{\Phi}(t_n + n^{-1/2} q_1(t_n) + O_p(n^{-1})) \\ &= \bar{\Phi}\left[t_n + n^{-1/2} p_1(0) - n^{-1/2}(B_2/6) t_n^2\right] + O_p(n^{-1}) \qquad (5.7)\end{aligned}$$

5.6 bootstrap-t 信頼区間は 2 次の精度の近似的に不偏な p-値を与える

先ほどの PV の漸近展開より

$$\mathrm{PV} < \alpha \Leftrightarrow t_n + n^{-1/2} q_1(t_n) + O_p(n^{-1}) > z_{1-\alpha}$$

これが等号になるのは $t_n = z_{1-\alpha} + O_p(n^{-1/2})$ つまり $q_1(t_n) = q_1(z_{1-\alpha}) + O_p(n^{-1/2})$ のときなので，

$$\begin{aligned}\mathrm{PV} < \alpha &\Leftrightarrow t_n + n^{-1/2}\hat{q}_1(z_{1-\alpha}) + O_p(n^{-1}) > z_{1-\alpha} \\ &\Leftrightarrow \theta_1 < \hat{\theta}_n + n^{-1/2}\hat{\tau}_n z_\alpha + n^{-1}\hat{\tau}_n \hat{q}_1(z_\alpha) + O_p(n^{-3/2}) \qquad (5.8)\end{aligned}$$

なお，ここでは PV を経由して上式を導出したが，コーニッシュ・フィッシャー展開を $\hat{c}_\alpha$ に適用すれば，$\hat{c}_\alpha = z_{1-\alpha} - n^{-1/2}\hat{q}_1(z_{1-\alpha}) + O_p(n^{-1}) = -z_\alpha - n^{-1/2}\hat{q}_1(z_\alpha) + O_p(n^{-1})$ か

ら直ちに同じ結果が得られる．(5.5) 式と (5.8) 式を比較すると，bootstrap-t の信頼区間は

$$s_1(z_\alpha) = q_1(z_\alpha) = p_1(0) - \tfrac{1}{6}B_2 z_\alpha^2$$

とおけばよい．そして被覆確率は

$$P(\theta_0 \in \hat{\mathcal{I}}_n|\theta_0) = 1 - \alpha + O(n^{-1})$$

となる．PV に関して言えば $\theta_0 = \theta_1$ のとき

$$P(\mathrm{PV} < \alpha) = \alpha + O(n^{-1}), \quad 0 < \alpha < 1$$

すなわち PV は 2 次の精度の近似的に不偏な p-値である．被覆確率の (5.6) 式を見ると，誤差が $O(n^{-1})$ となるのは $s_1(z_\alpha) = q_1(z_\alpha) + O(n^{-1/2})$ のときのみなので，2 次の精度の近似的に不偏な p-値はすべて $\mathrm{PV} + O_p(n^{-1})$ と書けることも分かる．

5.7 AU のバイアスは加速定数に比例する

(5.4) 式と (5.7) 式を比較すると

$$\bar{\Phi}^{-1}(\mathrm{PV}) - \bar{\Phi}^{-1}(\mathrm{AU}(\sigma^2)) = n^{-1/2}(1+\sigma^2)p_1(0) + at_n^2 + O_p(n^{-1})$$

ただし滑らかな関数モデルにおいて加速定数 a は次式で表される．

$$a = -\frac{1}{6}n^{-1/2}\left(A_2\tau^{-3} + B_2\right) = O(n^{-1/2})$$

ここで $\sigma^2 = -1$ とおくと $\bar{\Phi}^{-1}(\mathrm{PV}) - \bar{\Phi}^{-1}(\mathrm{AU}) = at_n^2 + O_p(n^{-1}) = O(n^{-1/2})$ となって，(4.2) 式と同じ結果が得られる．ただし $\beta_0 = t_n + O_p(n^{-1/2})$ と考えて良い．もし $a = 0$ ならば $\mathrm{AU} = \mathrm{PV} + O_p(n^{-1})$ は少なくとも 2 次の精度の近似的に不偏な p-値である．一般に $a \neq 0$ ならば AU は 1 次の精度しかない．これは 4.2 節の結果を確認するものである．

AU に相当する信頼区間の被覆確率を確かめておく．(5.4) 式より $\mathrm{AU}(\sigma^2) < \alpha \Leftrightarrow \theta_1 < \hat{\theta}_n + n^{-1/2}\hat{\tau}_n z_\alpha + n^{-1}\hat{\tau}_n \hat{s}_1(z_\alpha) + O_p(n^{-3/2})$ となるためには

$$s_1(z_\alpha) = -\sigma^2 p_1(0) + \tfrac{1}{6}A_2\tau^{-3}z_\alpha^2$$

とおけばよい．この $s_1(z_\alpha)$ を使えば (5.5) 式は $\mathrm{AU}(\sigma^2)$ と等価な信頼区間を与える．

$$n^{-1/2}(q_1(z_\alpha) - s_1(z_\alpha)) = n^{-1/2}(1+\sigma^2)p_1(0) + az_\alpha^2$$

を (5.6) 式に代入すれば被覆確率が得られる．$\sigma^2 = -1$ とおくと $\theta_0 = \theta_1$ のとき

$$P(\mathrm{AU} < \alpha) = 1 - P(\theta_0 \in \hat{\mathcal{I}}_n|\theta_0) = \alpha - az_\alpha^2\phi(z_\alpha) + O(n^{-1})$$

つまりバイアスは加速定数に比例する．

5.8 加速定数は歪度

推定量 $\tau(\bar{x}_n)$ の形式が，モデルを仮定せずに推定したものか，もしくは $x_1,\dots,x_n$ の分布としてとくに指数型分布族を仮定して最尤推定する場合を考える．すると Hall (1992) の (3.88) 式, (3.89) 式に示されるように，

$$2a_i a_j a_{kl}\eta_{ik}\eta_{jl} = a_i c_j \eta_{ij} - a_i a_j a_k \eta_{ijk}$$

がいえて，加速定数は次式のように $\hat{\theta}_n$ の歪度 (skewness) と解釈できる．

$$a = \frac{1}{6} n^{-1/2}\, a_i a_j a_k \eta_{ijk} \tau^{-3}$$

5.9 Efron の BC は AU に等価，ABC は PV に等価

BP のバイアスを修正するための Efron の二つの方法について述べておく．次に示すように，$O_p(n^{-1})$ の誤差を無視すると，BC は AU に等価，ABC は PV に等価である．これは 4.4 節の結果を確認するものである．

$$\mathrm{BP} = \hat{F}_n(\theta_1) = \bar{\Phi}\left[t_n - n^{-1/2}p_1(0) + n^{-1/2}(A_2\tau^{-3}/6)t_n^2\right] + O_p(n^{-1})$$
$$\mathrm{BP}_0 = \hat{F}_n(\hat{\theta}_n) = \bar{\Phi}\left[-n^{-1/2}p_1(0)\right] + O_p(n^{-1})$$

を (3.12) 式の BC に代入すると

$$\mathrm{BC} = \bar{\Phi}\left[t_n + n^{-1/2}p_1(0) + n^{-1/2}(A_2\tau^{-3}/6)t_n^2\right] + O_p(n^{-1})$$

これは (5.4) 式で $\sigma^2 = -1$ としたものに等しいので，$\mathrm{BC} = \mathrm{AU} + O_p(n^{-1})$.

一方，$\bar{\Phi}^{-1}(\mathrm{BP}) = t_n + O_p(n^{-1/2})$ に注意すると，(4.5) 式は

$$\mathrm{ABC} = \bar{\Phi}\left[t_n + n^{-1/2}p_1(0) + n^{-1/2}(A_2\tau^{-3}/6)t_n^2 + a t_n^2\right] + O_p(n^{-1})$$

これに $a = -n^{-1/2}\,(A_2\tau^{-3} + B_2)/6$ を代入すると $\mathrm{ABC} = \mathrm{PV} + O_p(n^{-1})$.

6. ほぼ平坦な曲面を仮定する漸近理論

6.1 錐はスケール変換で変化しない

これまで第 3 節から第 5 節で導出した正規化ブートストラップ確率のスケーリング則は，(3.14) 式, (4.1) 式, (5.4) 式 を見ると，σ^2 の一次式になっている．係数 β_0, β_1 の定義はそれぞれ異なるが

$$\mathrm{BP}(\sigma^2) \approx \bar{\Phi}(\beta_0\sigma^{-1} + \beta_1\sigma), \quad \mathrm{AU}(\sigma^2) \approx \bar{\Phi}(\beta_0 + \beta_1\sigma^2)$$

の形式だった．β_0 は主に ∂H までの距離，β_1 は主に ∂H の曲率に相当する．ところが，仮説領域 H が錐の場合を考えると，全く別のスケーリング則が得られる．3.7 節の議論では

σ^{-1} 倍するリスケーリングによって β_1 が σ 倍されたが，H が錐の場合はリスケーリングしても H の形状は不変だから β_1 は変化しない．ただし錐の β_1 は曲率ではなく頂点の角度に相当する．したがって，H が錐の場合のスケーリング則は

$$\mathrm{BP}(\sigma^2) \approx \bar{\Phi}(\beta_0\sigma^{-1} + \beta_1), \quad \mathrm{AU}(\sigma^2) \approx \bar{\Phi}(\beta_0 + \beta_1\sigma)$$

である．形式的に $\sigma^2 = -1$ を $\mathrm{AU}(\sigma^2)$ に代入すると $\sigma = \sqrt{-1}$ の項があるので，これでは AU を計算することすらできない．

6.2 曲面の変動量を小さくしていく

これまでの議論では ∂H が滑らかな曲面であることを仮定していた．Shimodaira (2008a) は ∂H が必ずしも滑らかではなく，錐や多面体の頂点や辺のように特異性がある場合も含めて考えた．第 3 節と同様に K 次元多変量正規モデル (3.1) 式, (3.2) 式を仮定して，仮説領域 H を (3.4) 式で表す．ただし 3.4 節と異なり，(u, v) 座標系は y に依存せず固定して考え，原点で ∂H に接することも仮定しない．

曲面の関数 h が滑らかでない場合を扱うので h のテイラー展開は利用できない．その代わり，h のフーリエ変換とその逆変換を利用する．

$$\tilde{h}(\omega) = \mathcal{F}h(\omega) = \int_{\mathbb{R}^{K-1}} e^{-i\omega\cdot u} h(u)\, du$$
$$h(u) = \mathcal{F}^{-1}\tilde{h}(u) = \frac{1}{(2\pi)^{K-1}} \int_{\mathbb{R}^{K-1}} e^{i\omega\cdot u} \tilde{h}(\omega)\, d\omega$$

漸近理論も通常の $n \to \infty$ とする方法は機能しない．h が滑らかな場合，h の k 階微分は $O(n^{-(k-1)/2})$ で小さくなり，∂H は平坦に近づいた．ところが H が錐の場合はリスケーリングで H の形状は変化しないから，∂H は不変である．通常の漸近理論では ∂H が平坦に近づくことが計算を進める上でポイントであった．そこで，$n \to \infty$ とする漸近理論をやめて，「∂H が平坦に近づく」という漸近理論を考えることにする．通常の漸近理論で $n \to \infty$ とするとき，$(u, v) \in \mathbb{R}^K$ のすべての方向に関して H を $n^{1/2}$ 倍するが，ここで考える漸近理論では v 方向だけ H を逆に小さくしていくので，錐であっても ∂H は平坦に近づく．

なお $n \to \infty$ とする漸近理論でも真の分布を固定した fixed alternatives を考えれば錐の頂点などは一般に遠くに離れていくので影響を無視できる．ここではむしろ頂点などの影響を評価するために真の分布を仮説境界に近づける local alternatives を考えている．

Shimodaira (2008a) は連続関数 h が次の条件を満たすとき，nearly flat surface と定義した．ただし λ は漸近理論を制御するパラメータで $\lambda \to 0$ の極限を考える．

$$\|h\|_1 = \int_{\mathbb{R}^{K-1}} |h(u)|\, du < \infty, \quad \|\mathcal{F}h\|_1 < \infty, \quad \|h\|_\infty = \sup_{u\in\mathbb{R}^{K-1}} |h(u)| = O(\lambda)$$

はじめの二つはフーリエ変換のための条件である．三番目の条件より $\lambda = 0$ のとき $h(u) \equiv 0$ となり ∂H は平坦である．すべての u で $h(u) = O(\lambda)$ であり，k 階微分もすべて $O(\lambda)$ に

なる．以下の議論では $O(\lambda)$ の代わりに $O(h)$ と書くことにする．曲面 h が多項式の場合や H が錐の場合を考えると $\|h\|_\infty = \infty$ となってしまうので，漸近展開による近似をするときは形式的に $\|u\|$ を有界に制限した H を考える．

このように $h \to 0$ とするのは，あくまで近似の都合である．そもそも $n \to \infty$ とする通常の漸近理論も近似の都合であり，n が十分に大きいとき近似誤差が小さくなる．$h \to 0$ とする漸近理論は，∂H が十分に平坦のとき近似誤差が小さくなる．

6.3 期待値オペレータでスケーリング則を表す

データ $y = (u, v)$ を観測したときのブートストラップ $Y^* = (U^*, V^*)$ の分布は

$$U^*|u \sim N(u, \sigma^2 I_{K-1}), \quad V^*|v \sim N(v, \sigma^2)$$

である．この U^* の分布に関して $h(U^*)$ の期待値をとるオペレータ $\mathcal{E}_{\sigma^2}$ を定義する．

$$\mathcal{E}_{\sigma^2} h(u) = E(h(U^*)|u; \sigma^2)$$

これを用いて次式のスケーリング則が容易に得られる．導出は Shimodaira (2008a) の (5.3) を参照．

$$\mathrm{AU}(\sigma^2) = \bar{\Phi}(v + \mathcal{E}_{\sigma^2} h(u)) + O(h^2) \tag{6.1}$$

6.4 滑らかな曲面ではスケーリング則のモデル式が σ^2 の多項式になる

曲面 ∂H が滑らかな場合はスケーリング則が σ^2 のベキ級数で表される．

$$\mathcal{E}_{\sigma^2} h(u) = \sum_{j=0}^{\infty} \sigma^{2j} \gamma_j(u)$$

係数 $\gamma_j(u)$ は u に依存して次式で与えられる．

$$\gamma_j(u) = \frac{1}{2^j j!} \sum_{j_1+\cdots+j_{K-1}=j} \frac{j!}{j_1! \cdots j_{K-1}!} \frac{\partial^{2j} h(u)}{\partial u_1^{2j_1} \cdots \partial u_{K-1}^{2j_{K-1}}}$$

つまり $\gamma_j(u)$ は $h(u)$ の $2j$ 階微分を平均化したものである．とくに $j = 0, 1, 2$ では

$$\gamma_0(u) = h(u), \quad \gamma_1(u) = \frac{1}{2} \sum_{i=1}^{K-1} \frac{\partial^2 h(u)}{\partial u_i^2}, \quad \gamma_2(u) = \frac{1}{8} \sum_{i=1}^{K-1} \sum_{j=1}^{K-1} \frac{\partial^4 h(u)}{\partial u_i^2 \partial u_j^2}$$

マルチスケール・ブートストラップでスケーリング則を推定する場合は (2.3) の多項式モデル $\psi_{\mathrm{poly}.k}(\sigma^2; \beta_0, \ldots, \beta_{k-1})$ が適切である．推定値は $\beta_0 \approx v + \gamma_0(u)$, $\beta_1 \approx \gamma_1(u)$, $\beta_2 \approx \gamma_2(u)$, $\beta_3 \approx \gamma_3(u), \ldots$ などとなり，(6.1) に代入するとスケーリング則は

$$\mathrm{AU}(\sigma^2) = \bar{\Phi}(\beta_0 + \beta_1 \sigma^2 + \beta_2 \sigma^4 + \beta_3 \sigma^6 + \cdots) + O(h^2)$$

である．BP と AU は次式で表される．

$$\mathrm{BP} = \mathrm{AU}(+1) = \bar{\Phi}(\beta_0 + \beta_1 + \beta_2 + \beta_3 + \cdots) + O(h^2)$$

$$\mathrm{AU} = \mathrm{AU}(-1) = \bar{\Phi}(\beta_0 - \beta_1 + \beta_2 - \beta_3 + \cdots) + O(h^2)$$

6.5 期待値オペレータは σ^2 に関して加法的

まず $U \sim N(0, \sigma^2 I_{K-1})$ の密度関数を $\phi(u; \sigma^2) = (2\pi\sigma^2)^{-(K-1)/2} e^{-\|u\|^2/2\sigma^2}$ と書くと

$$\mathcal{E}_{\sigma^2} h(u) = \int_{\mathbb{R}^{K-1}} \phi(u - u^*; \sigma^2)\, h(u^*)\, du^*$$

は $\phi(u; \sigma^2)$ と $h(u)$ の畳み込み (convolution) である．$\mathcal{F}\phi(\cdot; \sigma^2)(\omega) = e^{-\sigma^2\|\omega\|^2/2}$ より

$$\mathcal{E}_{\sigma^2} h(u) = \mathcal{F}^{-1}\left[\tilde{h}(\omega) e^{-\sigma^2\|\omega\|^2/2}\right](u) \tag{6.2}$$

したがって $\sigma_1^2 > 0$, $\sigma_2^2 > 0$ に対して $\mathcal{E}_{\sigma_1^2}\mathcal{E}_{\sigma_2^2} h(u) = \mathcal{E}_{\sigma_1^2+\sigma_2^2} h(u)$ である．形式的に $\sigma_1^2 = \sigma^2$, $\sigma_2^2 = -\sigma^2$ とおけば $\mathcal{E}_{\sigma^2}\mathcal{E}_{-\sigma^2} h(u) = \mathcal{E}_0 h(u) = h(u)$，つまり

$$\mathcal{E}_{\sigma^2}^{-1} h(u) = \mathcal{E}_{-\sigma^2} h(u) = \mathcal{F}^{-1}\left[\tilde{h}(\omega) e^{\sigma^2\|\omega\|^2/2}\right](u) \tag{6.3}$$

が $\mathcal{E}_{\sigma^2} h(u)$ の逆演算になる．ただし形式的に負の $\sigma_2^2 = -\sigma^2$ を代入したので逆演算が必ずしも存在するとは限らない．周波数空間で $e^{\sigma^2\|\omega\|^2/2}$ の乗算は高周波成分（大きな $\|\omega\|$ における成分）を指数的に増幅してしまうので，$\tilde{h}(\omega)$ の高周波成分が十分に早くゼロに収束しなければ (6.3) 式の $\mathcal{F}^{-1}$ が発散してしまう．多項式は滑らかで高周波成分は少なく，超関数のフーリエ変換を使うと (6.3) 式が定義できる．

6.6 期待値オペレータで棄却領域を表す

有意水準 α の近似的に不偏な検定の棄却領域を

$$R = \{(u, v) \mid v > z_{1-\alpha} - r(u), \quad u \in \mathbb{R}^{K-1}\} \tag{6.4}$$

と書く．$r(u) = O(h)$ は nearly flat surface と仮定する．$\mu = (u, -h(u)) \in \partial H$ のとき，

$$P(Y \in R \mid \mu) = \bar{\Phi}(z_{1-\alpha} + h(u) - \mathcal{E}_1 r(u)) + O(h^2) \tag{6.5}$$

である．これは $\mathrm{BP} = \bar{\Phi}(v + \mathcal{E}_1 h(u)) + O(h^2)$ の v を $z_{1-\alpha} + h(u)$ でおきかえ，$\mathcal{E}_1 h(u)$ を $-\mathcal{E}_1 r(u)$ でおきかえれば得られる．不偏検定ならば (6.5) 式 $= \alpha$ になるはずだから，$O(h^2)$ の誤差を無視して $h(u) - \mathcal{E}_1 r(u) = O(h^2)$ である．これを r について解けば

$$r(u) = \mathcal{E}_{-1} h(u) + O(h^2)$$

が棄却領域 (6.4) を定める．ところで $\alpha = \bar{\Phi}(v + r(u))$ を (6.4) 式に代入すると $y = (u, v) \in \partial R$ であるから，この α が検定の p-値である．したがって，

$$\mathrm{PV} = \bar{\Phi}(v + \mathcal{E}_{-1} h(u)) \tag{6.6}$$

はバイアス $O(h^2)$ の近似的に不偏な p-値である．実はもう少し良い結果が得られており，Shimodaira (2008a) の Theorem 1 によれば (6.6) 式のバイアスは $O(h^3)$ である．ところで (6.1) 式より $\mathrm{AU} = \bar{\Phi}(v + \mathcal{E}_{-1} h(u)) + O(h^2)$ であるから，$\mathrm{PV} = \mathrm{AU} + O(h^2)$ である．つまり AU はバイアス $O(h^2)$ の近似的に不偏な p-値である．

6.7 錐のスケーリング則は $\sigma^2 = -1$ へ外挿できない

仮説領域 H が原点を頂点とする錐の場合を考える．境界曲面 ∂H を定義する関数 h は，単位球面 S^{K-2} から実数への関数 c を使って $h(u) = c(u/\|u\|)\|u\|$ と表せる．$y = (u, v)$ は原点の近傍とする．T^* の分布が S^{K-2} 上の一様分布とすれば

$$\mathcal{E}_{\sigma^2} h(u) = \sigma e^{-\|u\|^2/2\sigma^2} \sum_{j=0}^{\infty} (j!\sigma^j)^{-1} E(\chi_m^{j+1}) E(c(T^*)(T^* \cdot u)^j)$$

となることが示せて，σ のベキで項をまとめると

$$\mathcal{E}_{\sigma^2} h(u) = \sum_{j=0}^{\infty} \sigma^{1-j} \gamma_j(u)$$

の形式に展開できる．ただし $\|u\| \to 0$ とすると係数は $\gamma_j(u) = O(\|u\|^j)$ である．最初の2項をみると $\gamma_0(u)\sigma + \gamma_1(u)$ であるので，$\psi_{\mathrm{poly}.k}$ の当てはまりは良くない．そこで

$$\psi(\sigma^2; \beta_0, \beta_1) = \beta_0 + \beta_1 \sigma \tag{6.7}$$

を当てはめると，$\beta_0 \approx v + \gamma_1(u)$, $\beta_1 \approx \gamma_0(u)$ が推定される．(6.7) 式と $\psi_{\mathrm{poly}.k-1}$ の両方を含むモデルとして，ヒューリスティックであるが

$$\psi_{\mathrm{sing}.k}(\sigma^2; \beta_0, \ldots, \beta_{k-1}) = \beta_0 + \sum_{j=1}^{k-2} \frac{\beta_j \sigma^{2j}}{1 + \beta_{k-1}(\sigma - 1)}, \quad 0 \le \beta_{k-1} \le 1$$

が考えられる（$k \ge 3$）．

錐におけるスケーリング則 $\mathrm{AU}(\sigma^2) = \bar{\Phi}(\beta_0 + \beta_1\sigma + \cdots)$ では $\sigma = \sqrt{\sigma^2}$ が含まれるため，形式的に $\sigma^2 = -1$ が代入できない．実は (6.6) 式において $\mathcal{E}_{-1} h(u)$ が発散する．Lehmann (1952) の議論から，そもそも錐に不偏検定は存在しないのである．不偏検定が存在しないのに $\mathrm{AU}(-1)$ が計算できたら，むしろ矛盾といえるだろう．

6.8 スケーリング則のモデルを有限項で打ち切る

十分に当てはまりの良いスケーリング則のモデル ψ を使って $\mathrm{AU}(\sigma^2) = \bar{\Phi}(\psi(\sigma^2))$, $\sigma^2 > 0$ が得られたと仮定する．H が滑らかな曲面ならば $\psi(\sigma^2)$ は σ^2 のべき級数で表され $\sigma^2 \le 0$ へ外挿できる．しかし H が錐の場合は $\psi(\sigma^2)$ を $\sigma^2 \le 0$ へ外挿できず，$\mathrm{AU}(-1)$ を計算できない．そこで適当に $\sigma_0^2 > 0$ を定め（例えば $\sigma_0^2 = 1$），その周りで $\psi(\sigma^2)$ をテイラー展開する．最初の k 項で打ち切って次式を定義する．

$$\mathrm{AU}_k(\sigma^2) = \bar{\Phi}\left[\sum_{j=0}^{k-1} \frac{(\sigma^2 - \sigma_0^2)^j}{j!} \frac{\partial^j \psi(\sigma^2)}{\partial(\sigma^2)^j}\Big|_{\sigma_0^2}\right] \tag{6.8}$$

錐のように $\mathrm{AU}(-1)$ が計算できない場合でも $\mathrm{AU}_k(-1)$ は計算できる．

$$\mathrm{AU}_k = \mathrm{AU}_k(-1)$$

は近似的に不偏な検定の p-値である．用いる k の値によって AU_k は異なる値になる．

ここで新たに定義した AU_k のフーリエ変換による表現を求めておく．$\psi(\sigma^2) = v + \mathcal{E}_{\sigma^2} h(u) + O(h^2)$ と (6.2) 式を (6.8) 式に代入する．フーリエ変換と微分の順序を交換すると，$\partial^j \psi / \partial(\sigma^2)^j$ は $(-\|\omega\|^2/2)^j$ をかけることに相当する．そして $\sigma^2 = -1$ を代入すると

$$\mathrm{AU}_k = \bar{\Phi}\left[v + \mathcal{F}^{-1}\left\{\tilde{h}(\omega)e^{\|\omega\|^2/2}(1 - J_k(\omega))\right\}(u)\right] + O(h^2) \tag{6.9}$$

と表現できる．ただしガンマ関数 $\Gamma(k)$，不完全ガンマ関数 $\gamma(k, x)$ を使って

$$J_k(\omega) = \gamma(k, (1+\sigma_0^2)\|\omega\|^2/2)/\Gamma(k) \tag{6.10}$$

によって関数 $J_k(\omega)$ を定義する．一方，(6.1) 式に (6.2) 式を代入すると

$$\mathrm{AU} = \bar{\Phi}\left[v + \mathcal{F}^{-1}\left\{\tilde{h}(\omega)e^{\|\omega\|^2/2}\right\}(u)\right] + O(h^2)$$

である．(6.9) 式の $1 - J_k(\omega)$ は $\|\omega\|$ の単調減少関数でありローパスフィルターと解釈できるので，AU の高周波成分をカットしたものが AU_k である．k を大きくするとローパスフィルターのカットオフ周波数が大きくなり，もし AU が存在すれば，AU_k は AU に近づく．

6.9　k を大きくすると AU_k のバイアスは小さくなる

近似的に不偏な検定の棄却域とバイアスを求めておく．$\mathrm{AU}_k < \alpha$ となる領域を (6.4) 式で表し，その $r(u)$ を $r_k(u)$ と書く．(6.6) 式を導いたときの議論より $\mathrm{AU}_k = \bar{\Phi}(v + r_k(u)) + O(h^2)$ であるから，(6.9) 式と比較すると $r_k(u) = \mathcal{F}^{-1}\{\tilde{h}(\omega)e^{\|\omega\|^2/2}(1 - J_k(\omega))\}(u)$ とおける．一方，検定のバイアスは (6.5) 式 $-\alpha = -\phi(z_\alpha)(h(u) - \mathcal{E}_1 r_k(u)) + O(h^2)$ である．

$$h(u) - \mathcal{E}_1 r_k(u) = h(u) - \mathcal{F}^{-1}\left\{\tilde{h}(\omega)(1 - J_k(\omega))\right\}(u) = \mathcal{F}^{-1}\left\{\tilde{h}(\omega)J_k(\omega)\right\}(u)$$

であるから，$|J_k(\omega)|$ が小さいほどバイアスが小さくなる．実際 $k \to \infty$ で $J_k(\omega) \to 0$ がいえて，バイアス $\to 0$ となる．

一般に (6.10) で定義したマルチスケール・ブートストラップの $J_k(\omega)$ に限らず，任意に与えた $J_k(\omega)$ から (6.9) 式で AU_k を定義しても同様の結果が得られる．例えば，4.3 節の k 回反復ブートストラップに相当する $J_k(\omega)$ を求めると

$$J_k(\omega) = (1 + e^{-\|\omega\|^2/2})(1 - e^{-\|\omega\|^2/2})^k \tag{6.11}$$

であるが，これから定義される AU_k もバイアス $\to 0$ である．

Shimodaira (2008a) の Theorem 2 によると，ω の各点で $\lim_{k\to\infty} J_k(\omega) = 0$ などの条件を満たせば，$\mu \in \partial H$ の各点で

$$\lim_{k\to\infty} P(\mathrm{AU}_k < \alpha) = \alpha + O(h^2)$$

である．さらにもう一つの条件として，$J_k(\omega)=\sum_{j=k}^{\infty}a_{k,j}\|\omega\|^{2j}$ の形に書けることを仮定する．このとき，$h(u)$ が次数 $2k-1$ 以下の多項式ならば，$\mu\in\partial H$ のとき

$$P(\mathrm{AU}_k<\alpha)=\alpha+O(h^2)$$

である．AU_k は $h(u)$ の $2k-1$ 次までの多項式成分に由来するバイアスを修正する．これらの条件を (6.10) 式と (6.11) 式は満たし定理が適用できる．つまり，マルチスケール・ブートストラップと反復ブートストラップは異なる AU_k を与えるが，仮説境界の曲面 $h(u)$ が $2k-1$ 次の多項式ならば $O(h^2)$ の誤差を無視して同等になる．

6.7 節で述べたように錐の場合は不偏検定が存在せず AU が発散する．ところが $k\to\infty$ で AU_k のバイアス $\to 0$ である．これは矛盾に見えるが，次のように説明できる．k を大きくするとバイアスを小さくするために棄却領域の境界曲面 $r_k(u)$ は次第にはげしく振動するようになる．$k\to\infty$ では $h(u)-\mathcal{E}_1 r_k(u)\to 0$ が成り立つが AU_k の値は発散してしまう．このことは円錐など簡単な例で解析的又は数値的に確認できる．実際に AU_k を応用した経験では $k=3$ 程度が適切と考える．

謝辞

マルチスケール・ブートストラップの着想に至る過程で統計数理研究所の栗木哲教授から様々なアドバイスを受けたことに感謝します．青嶋誠編集委員長および編集委員の方から本稿を改訂するための適切なコメントをいただきました．ありがとうございます．

参考文献

Bickel, P., Götze, F. and van Zwet, W. (1997). Resampling fewer than n observations: Gains, losses and remedies for losses, *Statistica Sinica*, **7**, 1–31.

DiCiccio, T. and Efron, B. (1992). More accurate confidence intervals in exponential families, *Biometrika*, **79**(2), 231–245.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife, *Ann. Stat.*, **7**, 1–26.

Efron, B. (1985). Bootstrap confidence intervals for a class of parametric problems, *Biometrika*, **72**, 45–58.

Efron, B. (1987). Better bootstrap confidence intervals, *J. Am. Stat. Assoc.*, **82**, 171–185.

Efron, B. and Tibshirani, R. (1998). The problem of regions, *Ann. Stat.*, **26**, 1687–1718.

Efron, B., Halloran, E. and Holmes, S. (1996). Bootstrap confidence levels for phylogenetic trees, *Proc. Natl. Acad. Sci. USA*, **93**, 13429–13434.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap, *Evolution*, **39**, 783–791.

Hall, P. (1992). *The Vootstrap and Edgeworth Expansion*, New York, Springer-Verlag.

Lehmann, E. L. (1952). Testing multiparameter hypotheses, *Ann. Math. Stat.*, **23**, 541–552.

Politis, D. and Romano, J. (1994). Large sample confidence regions on subsamples under minimal assumptions, *Ann. Stat.*, **22**, 2031–2050.

Shimodaira, H. (2002a). An approximately unbiased test of phylogenetic tree selection, *Syst. Biol.*, **51**, 492–508.

下平英寿 (2002b). 「ブートストラップ法によるクラスタ分析のバラツキ評価」『統計数理』**50**，33–44.

Shimodaira, H. (2004). Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling, *Ann. Stati.*, **32**, 2616–2641.

Shimodaira, H. (2006). *Scaleboot: Approximately Unbiased P-values via Multiscale Bootstrap*, R package is available from `http://CRAN.R-project.org/package=scaleboot`.

Shimodaira, H. (2008a). Testing regions with nonsmooth boundaries via multiscale bootstrap, *J. Stat. Plann. Inference*, **138**, 1227–1241.

下平英寿 (2008b). 「滑らかでない境界をもつ領域のためのマルチスケール・ブートストラップ法の理論と実装」『統計数理』**56**, 3–18.

Shimodaira, H. (2010). Frequentist and Bayesian measures of confidence via multiscale bootstrap for testing three regions, *Ann. Inst. Stat. Math.*, **62**, 189–208.

Shimodaira, H. and Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference, *Mol. Biol. Evol.*, **16**, 1114–1116.

Shimodaira, H. and Hasegawa, M. (2001). CONSEL: For assessing the confidence of phylogenetic tree selection, *Bioinformatics*, **17**, 1246–1247.

Suzuki, R. and Shimodaira, H. (2006). Pvclust: An R package for assessing the uncertainty in hierarchical clustering, *Bioinformatics*, **22**, 1540–1542. R package is available from `http://CRAN.R-project.org/package=pvclust`.