

統計的推測理論の現状

久保川達也*, 江口真透**, 竹村彰通***, 小西貞則****

Recent Developments of the Theory of Statistical Inference

Tatsuya Kubokawa* Shinto Eguchi** Akimichi Takemura***
and Sadanori Konishi****

This paper consists of four independent surveys on various aspects of the theory of statistical inference. The first chapter by T. Kubokawa surveys recent developments in decision theoretic estimation theory focusing on shrinkage type estimation procedures. The second chapter by S. Eguchi treats statistical asymptotic theory from differential geometrical viewpoint. The third chapter by A. Takemura surveys various topics of testing hypotheses. The final chapter by S. Konishi surveys recent developments in Bootstrap methodology.

統計的推測理論は多方面にわたって発展しているが、ここではこの発展を、決定論的観点からの推定論、微分幾何的アプローチによる漸近理論、検定論、ブートストラップ法、の4つのトピックにわけそれぞれのトピックに章をあてて概観する。全体の内容を調整した後、第1章を久保川、第2章を江口、第3章を竹村、第4章を小西がそれぞれ執筆した。トピックごとに文献もかなり明確にわかれるため、参考文献も各章ごとに与えてある。統計的推測理論のような大きな分野の発展を概観する際には、その中で何が重要な発展であるかなどについてさまざまな観点がありえる。ここでの概観も、それぞれの執筆者の観点にある程度引き寄せた概観となっていることとお断りしておきたい。

1. 統計的推定理論の最近の展開—縮小推定を中心に—

1.1 はじめに

データから未知母数についての推定・検定を行う、いわゆる統計的推測問題において、推測方法の良さを危険関数で比較しその性質を論ずる学問を統計的決定理論という。統計的決定理論はWald (1950) に始まり、ミニマックス性や許容性などの基本的概念の一般理論が構築され1960年代にはほぼ完成された。推定理論においては、70年代以降は各論に移り、Stein identityの導入によって平均ベクトルの同時推定に関するStein問題の研究が顕著な発展を遂げた。80年代には平均のStein問題の拡張と信頼領域に関する議論や平均が未知のときの分散の区間推定に関する研究が活発に行われ、またStein, HaffによるWishart identityの導入によって、共分散行列の推定問題や行列平均の推定問題についての研究が発展してきた。そして順序制約

論文受付：1993年1月 改訂受付：1993年2月 受理：1993年2月

* 東京大学 工学部, 〒113 東京都文京区本郷 7-3-1

** 島根大学 理学部, 〒690 島根県松江市西川津町 1060

*** 東京大学 経済学部, 〒113 東京都文京区本郷 7-3-1

**** 統計数理研究所, 〒106 東京都港区南麻布 4-6-7

下での母数推定, 変量・混合線形モデルでの分散成分の推定など応用上重要でより複雑なモデルでの推定問題とその理論的解明に関心が注がれてきた。こうした応用的現場への関心の広がりには, 経済学, 医学, 工学等の分野からの新たな統計的手法の開発要請に応えながら今後ますます拡大し, その現場に対応した推測理論の構築に向けて発展していくと思う。

本章では母数モデルの統計的推定理論のうち, 平均ベクトルの Stein 問題や分散, 共分散行列の推定問題など縮小推定に関する理論的展開を中心に報告する。

興味ある母数 θ はデータ X に基づいてその関数 $\delta(X)$ によって推定されるわけであるが, その推定量を評価するために損失関数 $L(\delta(X), \theta)$ がとられる。 $\delta(X)$ と θ とのユークリッド距離で定義される二乗損失関数が用いられることが多いが, X の密度関数 $f(x, \theta)$ に対して Kullback-Leibler の距離 $\int \log\{f(t, \delta)/f(t, \theta)\}f(t, \delta)dt$ に基づいた損失関数が自然であり, エントロピー損失と呼ばれる。推定量は危険関数 $R(\theta, \delta) = E[L(\delta(X), \theta)]$ によって評価される。この危険関数に関して最悪な場合を最善にするもの, 即ち $\sup_{\theta} R(\theta, \delta^*) = \inf_{\delta} \sup_{\theta} R(\theta, \delta)$ をみたく推定量 δ^* をミニマックス推定量という。ミニマックス推定量は, 無情報事前分布に対する一般化 Bayes 推定量に関係しており, Blyth による Bayes 推定量の極限としての導出方法と, Kudo, Kiefer による変換群に関する最良共変推定量としての導出方法がある。後者は共分散行列のミニマックス推定に関して重要である。ミニマックス推定量は損失関数のとり方に強く依存しており, 例えば 2 標本の共通平均の推定問題においては, 損失関数のとり方に応じて第 1 標本の標本平均, 第 2 標本の標本平均, 全標本平均というように異なった推定量がミニマックスになってしまうことが起こる (Zacks (1970), Kubokawa (1987))。ミニマックス推定に関する最近の興味深い結果は, 正規分布の平均が閉区間に入っているという制約条件の下では平均のミニマックス推定量は標本平均ではなく, 区間の両端に 0.5 の確率をもつ事前分布に対する Bayes 推定量であるという事実で, Casella-Strawderman (1981) によって示された。

ミニマックス性と並んで重要な概念に許容性がある。ある推定量 δ が δ^* を改良するとは, 任意の θ に対して $R(\theta, \delta) \leq R(\theta, \delta^*)$ が成り立ち, ある θ_0 で真に不等号が成立することをいい, このような δ が存在しないときに, δ^* は許容的であるという。推定量の許容性の条件は Stein, Brown, Zidek 等によって与えられ, Blyth や Hodges-Lehmann などの証明方法がある (Lehmann (1983), Berger (1985))。ミニマックス推定量などの通常の推定量の多くは一般化 Bayes 推定量になっており, それらの許容性を示すには注意深い議論が必要である。例えば k 次元正規分布の平均ベクトルの推定においては, 平均二乗誤差に関して通常の推定量が $k=1, 2$ のとき許容的であるにもかかわらず $k \geq 3$ では非許容的となってしまう。これは Stein 現象と呼ばれ理論的解明がなされてきた(詳しくは 1.3 節参照)。これに関連して許容性がマルコフ連鎖の再帰性に関係しているという興味深い結果が Brown (1971), Eaton (1992) 等によって得られている。

以下の節では縮小推定を中心に理論的展開を概説するが, それ以外の推定理論の最近の発展に関してここで若干ふれておく。推定理論において骨格をなすものは最尤推定であり, それは一貫性, 漸近有効性が保証される故である(2章参照)。最尤推定量を明示的に求めることが困難な場合には尤度方程式の数値的解法が行われるが, モデルが複雑になれば尤度方程式を立てること自体が大変になってくる。不完全データの解析や混合分布からのデータの解析においては EM アルゴリズムという簡便な方法があり, その収束性や加速法などが研究されている(宮川(1987), Eguchi (1991))。また母数の順序制約下での最尤推定量の導出のための isotonic 回帰法 (Robertson et al. (1988)) や prior feedback 法 (Robert (1991)) など計算機の発達にともなって導出方法にも様々な広がりを見せている。その他ほんの一部の紹介になってしまうが, Nile 問題や変動係数一定の問題など変換群に関して不変な構造をもつ推定問題については

Kariya (1989) 等によって、多変量階層モデルにおいて長さ固定の区間推定を構成するための二段階標本抽出法については Hyakutake-Siotani (1987) 等によって様々な研究がなされてきた。

1.2 分散の推定

通常の推定量が非許容的となる代表的な例の一つに、未知の平均をもった正規分布の分散を推定する問題がある。これは分散の不偏推定量が標本平均に含まれる情報を用いて改良されるというもので、Stein (1964) によって示された興味深い結果である。その後、Brewster-Zidek (1974) は Brown (1968) の方法に基づいて滑らかな一般化 Bayes 推定量を求めた。Stein と Brewster-Zidek の二つの方法は分散の区間推定や指数分布の尺度母数の推定などに適用されてきたが、最近 Kubokawa (1991 b) は新たな統一的方法を提案し改良する推定量のクラスを構成した。本節では、分散の推定に関するこうした決定理論的な歴史的展開を概説しよう。この分野についての優れた総合報告が Maatta-Casella (1990) により書かれているので参照されたい。

1.2.1 Stein 法と Brown-Brewster-Zidek 法 実験計画や線形回帰モデルなどの標準形に現れる次なるモデルを扱う。\$S\$ をスカラー、\$X\$ を \$p\$ 次元ベクトルとし互いに独立に

$$S/\sigma^2 \sim \chi_n^2, X \sim N_p(\theta, \sigma^2 I_p) \quad (1.1)$$

なる分布に従うとする。興味ある母数は \$\sigma^2\$ で、\$X, S\$ の関数 \$\delta = \delta(X, S)\$ によって推定するわけであるが、その推定量の良さを評価するためにエントロピー損失 \$L(\delta/\sigma^2) = \delta/\sigma^2 - \log(\delta/\sigma^2) - 1\$ を採用しそれに関する危険関数 \$R(\sigma^2, \theta, \delta) = E[L(\delta/\sigma^2)]\$ を考える。この他にも二乗損失 \$(\delta/\sigma^2 - 1)^2\$ や対称な損失関数 \$\delta/\sigma^2 + \sigma^2/\delta - 2\$ などが扱われる。

\$\sigma^2\$ の最も自然な推定量は不偏推定量 \$\delta_0 = n^{-1}S\$ である。これはまた次の意味で最適となっている。\$p \times p\$ 直交行列の全体を \$O(p)\$ で表すとき、アフィン変換群 \$S \to c^2S, X \to cGX + d, \sigma^2 \to c^2\sigma^2, \theta \to c\Gamma\theta + d, c \in R, d \in R^p, \Gamma \in O(p)\$, に関して \$\sigma^2\$ の推定問題を不変にするために、\$\delta(c^2S, cGX + d) = c^2\delta(S, X)\$ なる共変推定量を考える。このとき共変推定量は \$\delta(S, X) = aS, a > 0\$, と表され、このクラスの中で危険関数を最小にするもの、即ち最良共変推定量 (BEE) が存在し \$\delta_0\$ で与えられる。

自然な推定量 \$\delta_0\$ が \$X\$ に含まれる情報を用いて改良されるという興味深い結果を最初に示したのは Stein (1964) である。彼はアフィン変換群の部分群である尺度変換群 \$S \to c^2S, X \to cGX, \sigma^2 \to c^2\sigma^2, \theta \to c\Gamma\theta\$ に関して共変な推定量のクラス \$\delta_\phi = S\phi(W), W = \|X\|^2/S\$ に注目した。ここで \$S\$ は共変量、\$W\$ は最大不変量である。\$\|X\|^2/\sigma^2\$ は未知の非心母数 \$\lambda = \|\theta\|^2/\sigma^2\$ をもった非心カイ二乗分布 \$\chi_p^2(\lambda)\$ に従うので \$\delta_\phi\$ の中に最良な推定量は存在しないが、\$\delta_0\$ を改良するものを見つけることは可能である。

Stein は \$W\$ の条件付き期待値 \$E_\lambda[L(\phi(W)S/\sigma^2) | W]\$ を最小にする関数 \$\phi_\lambda(W)\$ を求め、\$\phi_\lambda(W) \le \phi_{\lambda=0}(W) = (1+W)/(n+p)\$ をみたすことを示した。従って \$\phi^T(W) = \min\{n^{-1}, (1+W)/(n+p)\}\$ とおくと \$\phi_\lambda(W) \le \phi^T(W) \le n^{-1}\$ なる不等式が成り立ち、損失関数の凸性から \$E_\lambda[L(\phi^T(W)S/\sigma^2) | W] \le E_\lambda[L(n^{-1}S/\sigma^2) | W]\$ となり、推定量

$$\delta^{ST} = \delta_{\phi^T} = \min\{\delta_0, (S + \|X\|^2)/(n+p)\}$$

は \$\delta_0\$ を改良することがわかる。こうした改良方法を Stein 法と呼び、これによって得られた打ち切り型推定量を Stein 型と呼ぶことにする。\$\delta^{ST}\$ は仮説 \$H: \theta = 0\$ vs. \$K: \theta \neq 0\$ に関して、\$H\$ が受容される時には \$(S + \|X\|^2)/(n+p)\$, 棄却される時には \$\delta_0\$ をとるという予備検定推定量に

なっている。また Bayes の観点からは経験 Bayes 推定量として自然に導かれる (Kubokawa et al. (1992 b))。 n が小さいとき δ_0 の分散が大きくなることから、 δ^{ST} の有効性は n が小、 p が大のときに顕著に現れることがわかる。

δ_0 を改良するもう一つの流れは Brown (1968) に始まる。彼は定数 $r > 0$ で半直線 $[0, \infty)$ を二分劃し、条件付期待値 $E_\lambda[L(aS/\sigma^2)|W < r]$ を最小にする定数 $a = a_\lambda(r)$ を求め、 $a_\lambda(r) \leq a_{\lambda=0}(r) = a_0(r) < n^{-1}$ なる不等式をみたすことを示した。ここで $a_0(r)$ は

$$a_0(r) = \frac{1}{n+p} \frac{\int_0^r t^{p/2-1}(1+t)^{-(n+p)/2} dt}{\int_0^r t^{p/2-1}(1+t)^{-(n+p+2)/2} dt} \quad (1.2)$$

で与えられる。この事実と損失関数の凸性から、 $W < r$ のときには $a_0(r)S$ 、 $W \geq r$ のときには $n^{-1}S$ をとる推定量が δ_0 を改良することがわかる。 Brewster-Zidek (1974) はこうした Brown のアイデアに基づいて半直線 $[0, \infty)$ を無限に細かく分割することを考え、その極限として $\delta^{GB} = Sa_0(W)$ なる滑らかな推定量を求め、 δ_0 を改良するとともに一般化 Bayes 推定量になっていることを示した。この一連の方法を BBZ (Brown-Brewster-Zidek) 法と呼び、この方法で得られた推定量を BBZ 型と呼ぶことにする。 Brewster-Zidek は δ^{GB} が δ_ϕ のクラスで許容的であることを示し、 Proskin (1985) は推定量全体における許容性を証明した。

$p=1$ のときには δ^{GB} による改善度はほんのわずかにすぎないことが指摘されているが、 p が大きくなれば意味のある改善度が得られる。 δ^{ST} が $\lambda=0$ で最大の改善度を与えるのに対して、 δ^{GB} は $\lambda=0$ では改良されておらず λ が 0 から少し離れたところで最大の改良を与える。このことと δ^{GB} の形の複雑さを考慮すると応用上は意味のはっきりした簡便な Stein 型推定量 δ^{ST} が望ましいと思われる。

正規分布以外にも、未知の位置母数をもった指数分布の尺度母数の点推定については Stein 型、 BBZ 型がそれぞれ Arnold (1970)、 Brewster (1974) により導かれ、また一様分布、逆ガウス分布に対しても Stein 型推定量が得られてきた。最後に点推定での改良方法には Stein 法、 BBZ 法以外に Strawderman (1974) による方法があることを注意しておきたい。

1.2.2 区間推定 分散の区間推定についても点推定の場合と対応する結果が導かれる。 Tate-Klett (1959) は S のみに基づいた $[aS, bS]$ なる形の信頼区間を考えた。ここで a 、 b は信頼係数 $0 < 1 - \gamma < 1$ に対し $P[aS < \sigma^2 < bS] = 1 - \gamma$ をみたす正の定数であるが、一意に決めるのに最適規準を導入する必要がある。一つは比 b/a を最小にする最小比信頼区間、もう一つは長さ $b - a$ を最小にする最短信頼区間である。そのときの a 、 b のみたすべき等式はそれぞれ

$$a^{-1} - b^{-1} = n \log(b/a), \quad a^{-1} - b^{-1} = (n+2) \log(b/a)$$

で与えられる。最小比信頼区間は最短不偏区間にもなっている。

最短信頼区間を X を用いて改良する試みは Cohen (1970) に始まる。彼は区間の長さを変えず、真の母数を覆う確率 CP (Coverage Probability) を大きくする意味で改良された信頼区間を Brown の方法で求めた。すると点推定との類似性から分割を無限に細かくすることが考えられる。 Shorrock (1990) はこうした BBZ 型信頼区間を求めその一般化 Bayes 性を示した。さらに区間の長さを短かくし CP を大きくするという両方の意味で改良する BBZ 型信頼区間の導出とその一般化 Bayes 性が Goutis-Casella (1991) によって示された。

一方、分散の信頼区間については最短規準よりも最小比規準の方が自然であることが指摘されている。 Nagata (1989) は簡便で応用上有用な立場から Stein 型信頼区間

$$I^{ST} = [\min\{1, \phi^*(W)\}aS, \min\{1, \phi^*(W)\}bS],$$

$$\phi^*(W) = (a^{-1} - b^{-1})(1+W) / \{(n+p)\log(b/a)\}$$

を提案し最小比信頼区間を改良することを証明した。また信頼性などで重要な指数分布の場合についても簡便な Stein 型信頼区間を導出した (永田 (1991))。

1.2.3 新たな統一的方法 いままで別々に得られてきた Stein 型及び BBZ 型推定量を統一的に導く新たな方法が Kubokawa (1991 b), Takeuchi (1991) によって提案された。この基本的アイデアは、危険関数の差を積分表現することであり、IERD (Integral Expression of Risk Difference) 法と呼ぶことにする。

いま $\lim_{w \rightarrow \infty} \phi(w) = n^{-1}$ と仮定すると δ_0 と δ_ϕ との危険関数の差は定積分によって

$$\begin{aligned} E[L(\delta_0/\sigma^2)] - E[L(\delta_\phi/\sigma^2)] &= E[[L(\phi(tW)S/\sigma^2)]_{t=1}^\infty] \\ &= E\left[\int_1^\infty \frac{d}{dt} \{L(\phi(tW)S/\sigma^2)\} dt\right] \\ &= E\left[\int_1^\infty L'(\phi(tW)S/\sigma^2) \phi'(tW) WS/\sigma^2 dt\right] \end{aligned}$$

と表現でき、変数変換を行うと $\phi(w)$ について改良するための次なる条件を導くことができる。

- (a) $\phi(w)$ は単調増加で $\lim_{w \rightarrow \infty} \phi(w) = n^{-1}$,
- (b) $\phi(w) \geq a_0(w)$.

但し $a_0(w)$ は (1.2) で与えられている。こうして δ_0 を改良するクラスが得られたわけで、 $a_0(w)$, $\phi^T(w)$ が (a), (b) の条件を満たすことから δ^{GB} , δ^{ST} はこのクラスに入ることがわかり、別々の方法で導かれてきた二種類の推定量が IERD 法により統一的に得られる。

IERD 法はその簡便さ故、正規、対数正規、指数、Pareto 分布など単調尤度比をもった分布族と Bowl 型損失関数の場合への一般化を可能にする。また区間推定に対しても適用可能で

- (a) $\phi(w)$ が単調増加で $\lim_{w \rightarrow \infty} \phi(w) = 1$,
- (b) $\int_0^1 t^{\frac{p}{2}-1} \{a^{-\frac{n+p}{2}} e^{-\frac{1}{2a\phi}(1+wt)} - b^{-\frac{n+p}{2}} e^{-\frac{1}{2b\phi}(1+wt)}\} dt \geq 0$

なる条件をみたとす $\phi(w)$ に対して、最小比信頼区間は $I_\phi = [a\phi(W)S, b\phi(W)S]$ によって CP を大きくする意味で改良される。このクラスの中には Stein 型区間 I^{ST} や BBZ 型区間 I_{ϕ_0} が含まれている (Kubokawa (1991 b))。

1.2.4 多次元母数の推定への拡張 多次元母数の推定への拡張として多変量回帰モデルの共分散行列及び一般化分散の推定が取り上げられる。その標準形は互いに独立な p 次正方行列 S と $p \times r$ 行列 X を用いて

$$S \sim W_p(n, \Sigma), X \sim N_{p \times r}(\theta, \Sigma \otimes I_r) \quad (1.3)$$

と表わされる。ここで $W_p(n, \Sigma)$ は Wishart 分布、 \otimes は Kronecker 積を表わしている。

一般化分散 $|\Sigma|$ の点推定については Shorrocks-Zidek (1976) が Zonal 多項式を用いて Stein 型推定量を導出し、Zonal 多項式を用いない別証明が Sinha (1976) により与えられた。最近 Sugiura-Konno (1988) は危険関数の級数表現を与えて改善度を数値的に調べた。一方 Stein 型信頼区間などが Sarkar (1989) によって求められた。しかし $r \geq 2$ の場合 BBZ 型推定量を求めることは容易ではなく、これは最大不変量が一次元で表わせないことに起因している。Rukhin-Sinha (1991) は一般化分散の通常の推定量が X を使わなくても $p \geq 4$ なら非許容的となるという興味深い事実を証明している。

共分散行列 Σ の推定については、Sinha-Ghosh (1987) によって Stein 型推定量が導かれた。

特に $r=1$ の場合に Perron (1990) は尺度共変推定量のクラスにおいて Stein 法による特徴付けを与え, Kubokawa et al. (1992 a) は Sinha-Ghosh 推定量を改良する経験 Bayes 推定量を導出し, Kubokawa et al. (1990) は Σ の推定構造を明らかにし改良する一般化 Bayes 推定量を求めた。しかし Σ の場合と同様 $r \geq 2$ のときの BBZ 型推定量等明らかにすべき多くの問題が残されている。Xを使わなくても S 自身で改良可能なことについては 1.4 節で扱う。

1.3 平均ベクトルの同時推定

統計的決定理論において最も興味深くそして驚くべき結果は, Stein (1956) によって発見された平均ベクトルの同時推定に関する非許容性の事実であろう。これは三つ以上の推定問題を一緒にすると個々には改良できないものが改良されてしまうというもので, Stein 現象と呼ばれ Brown, Berger, Efron, Morris 等により理論的研究や応用への可能性等が論じられてきた。1970 年代から今日に至るまでこの分野が爆発的に発展しつづけてきた理由の一つには, Stein (1973) によって開発された部分積分のアプローチにより技術的取扱いが平易になった点が挙げられるだろう。この節では, Stein 問題の今日までの歩みを概説する。なお竹内 (1979), 篠崎 (1991), Brandwein-Strawderman (1990) により優れた総合報告が出されているので参照されたい。

1.3.1 Stein 現象 p -次元確率ベクトル $X=(X_1, \dots, X_p)'$ が正規分布 $N_p(\theta, I_p)$ に従うとき平均ベクトル $\theta=(\theta_1, \dots, \theta_p)'$ を X の関数 $\delta(X)$ で同時推定する問題を考えよう。ここでは推定量の良さを評価するのに二乗損失関数 $\|\delta(X)-\theta\|^2$ に関する危険関数とがとられる。

θ の自然な推定量は明らかに X 自身であり, 最小分散不偏, 最尤そしてミニマックスである。また直交行列 Γ , ベクトル d に対して, $\Gamma X+d, \Gamma\theta+d$ なる変換群に関して推定問題が不変になるためには, 推定量は共変性 $\delta(\Gamma X+d)=\Gamma\delta(X)+d$ をみたさなければならない。この共変推定量は $\delta(X)=X+d$ の形で表わされるが, X はこのクラスの中で最良なものになっている。

Stein (1956) は X の許容性に注目し, $p=1, 2$ のときには許容的であるが, $p \geq 3$ に対しては非許容的となることを証明した。事実彼は上の変換群の部分群 $\Gamma X, \Gamma\theta$ に関して共変な推定量 $\delta_\phi = \{1-\phi(\|X\|^2)\|X\|^{-2}\}X$ の中に X を改良するものが存在することを示し, James-Stein (1961) は, $\delta^{JS} = \{1-(p-2)/\|X\|^2\}X$ という形の改良型推定量の一つを明示的に与えた。この Stein 現象を説明する議論には竹内 (1979), Stigler (1990), Brandwein-Strawderman (1990) 等がある。

James-Stein 推定量はその形から何か奇異な印象をもたれるかもしれないが, Bayes 的立場からは経験 Bayes 推定量として自然に導かれる。いま母数 θ を確率変数と考え θ が事前分布 $N_p(0, \tau I_p)$ に従うとする。ここで τ は未知母数である。このとき X を与えたときの θ の事後分布は $N_p(\tau X/(1+\tau), \tau I_p/(1+\tau))$ となるから θ の Bayes 推定量は $\{1-(1+\tau)^{-1}\}X$ で与えられる。これは未知母数 τ を含むので τ を X の周辺分布 $N_p(0, (1+\tau)I_p)$ で推定することが考えられる。 $\|X\|^2$ の周辺分布が $(1+\tau)\chi_p^2$ であるから $(1+\tau)^{-1}$ の不偏推定量は $(p-2)/\|X\|^2$ となる。これを上の Bayes 推定量に代入して経験 Bayes 推定量 $\{1-(p-2)/\|X\|^2\}X$ が得られ, これが δ^{JS} に一致することがわかる。 δ^{JS} のこうした動機付けは Efron-Morris (1972) によってなされたが, 彼は同時にこの経験 Bayes の方法によって Stein 現象の証明が可能であることを示した。Stein 現象の証明方法には, この他に, 非心カイ二乗分布が中心カイ二乗分布の Poisson 混合分布で表現されることを用いた James-Stein (1961) の最初の方法と, 部分積分を利用した Stein (1973) の方法がある。後者は大変簡便で有用であるので次に紹介しておこう。

より一般的に δ_ϕ が X を改良するための ϕ についての条件を求めることから始めよう。絶対連続な関数 $h(x)$ とその微分導関数 $h'(x)$ に対して部分積分により $E[(X_i-\theta_i)h(X_i)]=$

$E[h'(X_i)]$ なる等式が成り立つ。これは Stein identity と呼ばれ、これを用いると δ_ϕ の危険関数は

$$\begin{aligned} R(\theta, \delta_\phi) &= E\left[p + \frac{\phi^2}{\|X\|^2} - 2 \sum_{i=1}^p (X_i - \theta_i) X_i \frac{\phi}{\|X\|^2}\right] \\ &= E\left[p + \frac{\phi}{\|X\|^2} \{\phi - 2(p-2)\} - 4\phi'\right] \end{aligned}$$

となる。こうして identity を使うことにより未知母数 θ_i を消すことができ、 $E[\cdot]$ の中身が危険関数の不偏推定量になるのである。 $R(\theta, X) = p$ だから、結局 X を改良するためには $\phi(t)$ が $\phi(t)\{\phi(t) - 2(p-2)\}/t - 4\phi'(t) \leq 0$ なる微分不等式をみたせばよい。例えば (1) ϕ が非減少かつ (2) $0 < \phi \leq 2(p-2)$, であればよい。こうして X を改良する推定量のクラス δ_ϕ が構成される。 $\phi(t) = p-2$ は明らかに (1), (2) の条件をみたすので δ^S はこのクラスに入り、その危険関数は $R(\theta, \delta^S) = p - (p-2)^2 E[\|X\|^{-2}]$ で与えられ、非心度が 0 のとき最大の改善が得られる。Stein identity による方法の簡便さと有能さはその後のこの分野の顕著な発展をもたらしていくことになった。

James-Stein 推定量は $\|X\|^2 < p-2$ のとき縮小し過ぎてしまい各 X_i の符号を変えてしまう。そこで $\delta^+ = \max(0, 1 - (p-2)/\|X\|^2)X$ なる positive-part Stein 推定量が考えられ、実際 δ^S を改良することがわかる。 δ^+ 自身、解析的でないため非許容的であることが一般論から導かれる。 δ^+ を改良する明示的な推定量を見つけることはここ数十年の大問題とされてきたが、最近 Shao-Strawderman (1992) はその導出に成功している。しかし彼らの発見した推定量も滑らかでなく、その意味では非許容的なままである。それでは X を改良する許容的な推定量はどのような形をしているのだろうか。Strawderman (1971) は $\phi^{GB}(t) = p-2 - 2\left[\int_0^1 z^{p/2-2} e^{(1-z)t/2} dz\right]^{-1}$ に対して $\delta_{\phi^{GB}}$ なる形の一般化 Bayes 推定量を求め、 $\phi^{GB}(t)$ が上の (1), (2) の条件をみたすこと、即ち $\delta_{\phi^{GB}}$ が X を改良することを示すとともにその許容性を証明した。 $\delta_{\phi^{GB}}$ は θ の事前分布の中の超母数に対して無情報事前分布を仮定して得られるという意味で階層 Bayes 推定量である。Stein (1973, 81) は $\nabla^2 = \sum \partial^2/\partial x_i^2$, $\nabla = (\partial/\partial x_1, \dots, \partial/\partial x_p)'$ に対して、 f が超調和条件 $\nabla^2 f(x) \leq 0$ をみたせば $\delta^S = X + \nabla \log f(X)$ なる推定量が X を改良することを示し、ポテンシャル理論との興味深い関係を暗示した。また一般化 Bayes 推定量は δ^S の形で表現され、形式的な事前測度が超調和条件をみたせば X を改良することが示された (Stein (1981), Haff (1991))。一般化 Bayes 推定量の許容性と非許容性との境界を事前分布によって特徴付けることもなされている (Brown-Hwang (1982), Berger (1985) 等)。

Kubokawa (1991 a) は分散の推定で使われた BBZ 法をこの問題に適用し、 $\delta_{\phi^{GB}}$ が δ^S を改良する BBZ 型推定量になっていることを示した。さらに Kubokawa (1991 b) は IERD 法を用いて、 δ^S を改良するクラスを明らかにし、

- (a) $\phi(t)$ は単調増加で、 $\lim_{t \rightarrow \infty} \phi(t) = p-2$,
- (b) $\phi(t) \geq \phi^{GB}(t)$,

なる $\phi(t)$ に対して δ_ϕ は δ^S を改良することを示した。このことは、分散の推定問題が δ^S の改良に関係していることを暗示しており、 δ^+ が Stein 型、 $\delta_{\phi^{GB}}$ が BBZ 型に対応している。また分散が未知のモデル (1.1) においては、James-Stein 推定量は分散 σ^2 の最良共変推定量 σ_0^2 を用いて $\{1 - \sigma_0^2(p-2)/\|X\|^2\}$ で与えられるが、 σ_0^2 の代わりに改良型推定量を使用することが James-Stein 推定量の改善に通ずるという結果も得られている (Kubokawa et al. (1993))。

1.3.2 拡張と新たな展開 Stein 現象については様々な拡張等がなされてきたが、その主なものを正規分布の場合についてまず概説しよう。

共分散行列が未知の場合には、尺度変換群に関して不変な損失関数を扱う限り未知の共分散行列をその推定量で置き換えればよい (James-Stein (1961), Lin-Tsai (1973)). しかし不変でない損失関数に関しては一様な改良を与える推定量の導出は容易でなかったが、その明快な解答が Gleser (1986) によって与えられた。

行列平均の推定問題は共分散行列の推定と関連している点で興味深い。 $p \times r$ 確率行列 X が $N_{p \times r}(\theta, I_p \otimes I_r)$ に従うとき、Efron-Morris (1972) は通常の推定量 X が経験 Bayes 推定量 $\hat{\theta}_0^{EM} = \{I_p - (p-1)(XX')^{-1}\}X$ によって改良されることを示した。さらに Stein (1973) は $\hat{\theta}_0^{EM}$ の改良を考え、Efron-Morris (1976) は経験 Bayes の方法により $\hat{\theta}_0^{EM}$ の改良は逆共分散行列の推定問題に帰着できることを示し、 $\hat{\theta}^{EM} = \hat{\theta}_0^{EM} - (p^2 + p - 2)(\text{tr} XX')^{-1}X$ なる推定量によって改良されることを導いた。Zheng (1988) は Stein (1981) の行列平均への拡張を行った。より一般的な多変量回帰モデル (1.3) における係数行列の推定問題への展開は Bilodeau-Kariya (1989), Konno (1991), Honda (1991) 等によりなされ、成長曲線モデルでの議論が Kubokawa et al. (1992 b), Tan (1991) によってなされた。成長曲線モデルにおいて改良のための一般的な条件を求めることは今後に残された課題になっている。

θ についての事前情報に基づいて θ がある部分空間 V に入っていることが推察されるときには V の方向へ X を縮小する Stein 推定量 $\delta(V)$ が考えられ、 θ が V に近いときには大きな改善を与える。しかしそのような事前情報はもっと漠然としたものであるかもしれない。George (1986) は θ が存在すると推察される部分空間の候補が複数個 V_1, \dots, V_k 考えられるとき、それぞれに縮小する Stein 推定量 $\delta(V_i)$ の重み付きの和として表わされる適応型推定量 $\sum_{i=1}^k \rho_i(X) \delta(V_i)$ を提案した。ここで $\rho_i(X)$ は改良分の大きいと考えられる $\delta(V_i)$ に対する重みが大きくなるように作られている。

この他にも、母数 $\theta_1, \dots, \theta_p$ の間に順序制約が課せられているときの Stein 現象については Chang (1982), Sengupta-Sen (1991) により、逐次解析での Stein 現象については Takada (1984), Ghosh et al. (1987) 等により、多重回帰問題での最尤推定量の非許容性に関しては Baranchick (1973), Takada (1979), Zidek (1978) により議論された。また Stein 現象が損失関数のクラスにおいて一様に成立しているか否かの議論が Shinozaki (1980), Hwang (1985) によって、Pitman closeness なる規準の下での解明が Sen et al. (1989) によってなされた。

1.3.3 非正規分布での Stein 現象 正規分布以外の離散型・連続型分布族に対しても Stein 現象の解明がなされてきた。

連続型分布族のうち、球対称性の仮定の下では $p \geq 4$ のとき Stein 現象が生ずることが証明され (Brandwein-Strawderman (1990) 参照), またコンパクトな集合上での一様性を仮定すれば $p=3$ のときでも Stein 現象が生ずることが示された。Elliptical Contoured (EC) 分布においては、最小二乗推定量を James-Stein 推定量が改良するための条件は分布の形に依らないこと、即ち改良の頑健性が Cellier et al. (1989) により示された。

連続型指数分布族の場合、部分積分による identity が Hudson (1978) により導かれ、通常の不偏推定量を改良する縮小推定量を求めるため微分不等式の解法が Hudson (1978), Berger (1980), DasGupta (1986) により求められた。特に Berger (1980) は Gamma 分布の尺度母数の同時推定については損失関数のとり方によっては 2 次元以上で Stein 現象が生ずるという事実を発見した。その他に Hudson の identity の多次元指数分布族への拡張、エントロピー損失関数の下での Stein 現象、指数分布族での超調和条件等が議論されている。Shinozaki (1984) は一様分布、両側指数分布、 t -分布などの一次元の分布についても合算することにより Stein 現象が生ずることを示した。また分布の形が明示的にわからなくても 4 次の中心モーメントまで

の情報があれば Stein 効果が得られるという興味深い結果を与えた。正の母数の同時推定に関する非許容性の一般的結果が DasGupta (1989) により得られている。

離散型指数分布族の場合には、不偏推定量を改良するための縮小推定量を構成することは差分不等式の解を見つけることになり、Poisson 分布、負の二項分布に対して 4 次元以上のときに解が求められた (Hudson (1978))。一方 Poisson 分布の Stein 現象が 2 次元以上のとき生ずることを Clevenston-Zidek (1975) が示し、離散分布での Stein 問題の研究が活発になされてきたが、それらが Ghosh et al. (1983) によって統一的にまとめられた。

ノンパラメトリックモデルにおいても L-, M-, R-推定量の Stein 効果による改良が Sen-Saleh (1987), Shiraiishi (1991) 等によって示された。確率過程や時系列モデルでの Stein 現象についても調べられている。

1.3.4 信頼領域 1980 年代の Stein 問題の展開の一つに信頼領域の問題が挙げられる。X が $N_p(\theta, I_p)$ に従うときの通常信頼領域は $C^0(X) = \{\theta; \|\theta - X\|^2 \leq c\}$ であり、c は信頼係数 $1 - \gamma$ に対して $P(\chi_p^2 \leq c^2) = 1 - \gamma$ をみたす定数である。

ある信頼領域 $C(X)$ が $C^0(X)$ を改良するとは、(I) $P_\theta\{\theta \in C(X)\} \geq P_\theta\{\theta \in C_0(X)\}$ がすべての θ に対して成り立つことと、(II) $(C(X) \text{ の体積}) \leq (C^0(X) \text{ の体積})$ がほとんどすべての X に対して成り立つことである。Stein, Brown, Joshi らによって $p \geq 3$ のときの $C^0(X)$ の非許容性が示されてきた。しかし実際改良している信頼領域を明示的に求めたのは Hwang-Casella (1982) であり、positive-part Stein 推定量 $\delta_a^+(X) = \max\{0, 1 - a/\|X\|^2\}X$ に対して $C^a(X) = \{\theta; \|\theta - \delta_a^+(X)\|^2 \leq c\}$ なる信頼領域を考えると、 $p \geq 4$ で $0 < a \leq a_c$ なる a に対して $C^a(X)$ が $C^0(X)$ を改良することを証明した。但し a_c は方程式 $\{\sqrt{c} + \sqrt{c + a_c}\}^{p-3} = (a_c)^{(p-3)/2} e^{\sqrt{a_c c}}$ の解である。さらに Hwang-Casella (1984) では $p=3$ の場合を含むように a のみならず範囲を広げた。そして球対称な分布族への拡張や改良する信頼領域のクラスの構成がなされてきた。尺度母数が未知の場合の、区間推定での Stein 現象については Robert-Casella (1990) が多変量 t -分布を含む球対称分布族に対して証明した。しかしその議論には正規分布が含まれておらず今後の課題とされている。

以上述べてきた信頼領域の改良は、体積を等しくしたまま (I) の意味で CP を大きくする方向でなされてきた。しかし区間推定の本来の意味からは CP を一定にしたまま (II) の意味で体積のより小さい信頼領域を構成することが望まれる。Shinozaki (1989) は球 $C^0(X)$ 全体を原点に向けて縮小することによって (II) の意味で一様に改良する信頼領域を解析的に与えることに成功した。

1.4 共分散行列の推定

この節では統計的決定理論の立場から活発に研究されてきた、多変量正規分布の共分散行列の推定問題について報告する。

$p \times p$ 確率行列 S が期待値 $n\Sigma$ をもった Wishart 分布 $W_p(n, \Sigma)$ に従うとし、 Σ を $\hat{\Sigma}$ によってエントロピー損失 $\text{tr} \hat{\Sigma} \Sigma^{-1} - \log |\hat{\Sigma} \Sigma^{-1}| - p$ に関して推定する問題を考えよう。通常、不偏推定量 $\hat{\Sigma}_0 = n^{-1}S$ が用いられるが、 $\hat{\Sigma}_0$ の固有根が Σ の固有根に比べ広がってしまい、この欠点を克服するために $\hat{\Sigma}_0$ の固有根を中央に向けて縮小する必要が生ずる。この方向の仕事には、Stein (1977), Efron-Morris (1976), Haff (1980), Sugiura-Fujimoto (1982) 等がある。特に Haff (1979) は Wishart 分布での部分積分の公式 (Wishart identity) を導出し、それは共分散行列を含んだ推測問題において改良型推定量を求めるための強力な手段となっている。

一般線形群はミニマックス性についての Kiefer の条件をみたさないので最良共変推定量 $\hat{\Sigma}_0$

はミニマックスでない。James-Stein (1961) はその部分群である下三角行列による変換群 $G\ddagger$ を考え、それに関する最良共変推定量がミニマックスであり、 $\hat{\Sigma}^m = TDT'$, $S = TT'$, $T \in G\ddagger$, $D = \text{diag}(d_1, \dots, d_p)$, $d_i = (n+p+1-2i)^{-1}$, で与えられることを示した。しかし $\hat{\Sigma}^m$ は座標系のとり方に依存するので直交不変なミニマックス推定量を構成することが望まれる。

直交不変なミニマックス推定量の導出には二つの方向がある。一つは Stein (1977), Dey-Srinivasan (1985) のアプローチであり、直交行列 R , 対角行列 $L = \text{diag}(l_1, \dots, l_p)$ によって $S = RLR'$ と表わされるとき、 $\hat{\Sigma}^m$ が $\hat{\Sigma}^{ST} = R \text{diag}(l_1 d_1, \dots, l_p d_p) R'$ によって改良される。さらに Dey-Srinivasan (1985) は $p \geq 3$ のときに $\hat{\Sigma}^{ST}$ を改良する推定量を導き、Sheena-Takemura (1992) は打ち切り型推定量を考えることによって $p \geq 2$ での $\hat{\Sigma}^{ST}$ の非許容性を示した。Haff (1991) は Bayes 推定量の変分形式 VFBE (Variational Form of Bayes Estimator) を与える一般論を展開し、 Σ に対する VFBE を求め、それが $\hat{\Sigma}^{ST}$ より優れていることをシミュレーション実験によって示した。もう一つは Takemura (1984) のアプローチで、直交群 $O(p)$ 上の一様分布 μ と、 $T_r T_r' = \Gamma' S \Gamma$, $\Gamma \in O(p)$ に対して $\hat{\Sigma}^u = \int_{O(p)} \Gamma T_r D T_r' \Gamma' d\mu(\Gamma)$ なる推定量によって $\hat{\Sigma}^m$ は改良される。 $\hat{\Sigma}^u$ は $p \leq 3$ のときには明示的表現が与えられたが、 $p \geq 4$ では困難とされてきた (Takemura (1984))。その困難さはある量の比の期待値を計算するところにあるが、Perron (1992) はそれを期待値の比に置き換えて近似解を陽に求め、それが直交不変ミニマックス推定量になっていることを示した。

共分散行列に関連して二つの共分散行列の比に関する推定が DasGupta (1989), Konno (1992), Bilodeau-Srivastava (1992) 等によって議論されてきた。特に Bilodeau-Srivastava は比に関するエントロピー損失を導入し、共分散行列の場合と同様な結果が比の推定において成立することを示した。

1.5 縮小を要する推定問題

以上述べてきた、分散、共分散行列、平均ベクトルの推定は縮小推定の代表的な問題である。その他にも通常の推定量の縮小または拡大を必要とする問題は少なくないように思われる。そのときどの程度縮小または拡大すべきかの指標を与えることが重要であるが、1.2 節、1.3 節でふれた IERD 法はその一つの有用な手段であると期待される。その例のいくつかを以下に紹介しよう。

変量模型、混合模型における分散の群間成分の推定については、一般にその不偏推定量は正の確率で負値を取りえてしまい、その非合理性を排除する為に様々な試みがなされてきた。二次形式の推定量のうちでは非負な不偏推定量は存在しないことや非負な二次形式推定量は漸近的一致性を持たないことが知られている。従って、二次形式統計量の範囲を超えて正で一致性をもった改良型推定量を求めることが望まれる。IERD 法を用いると、不偏推定量を正の方に縮小することによってそのような推定量を求めることができる。多次元への拡張として、多変量混合モデルにおける共分散行列の群間成分の推定が Calvin-Dykstra (1991) によって議論され、最尤推定量を求めるアルゴリズム等が提案された。このような多次元の場合に IERD 法をいかに適用するかは今後の興味深い課題である。

非心カイ二乗分布や非心 F 分布の非心母数の推定についても不偏推定量は負値を取りえてしまい、その欠点を排除するための手段が議論されてきたが、この場合にも不偏推定量を改良する合理的な推定量の導出に対しては IERD 法が有効であることがわかる。この問題の多次元化も今後の課題である (Leung-Muirhead (1987))。

順序制約下での母数推定においては、推定量がその母数空間からはみ出しているときにはその空間に縮小または拡大する必要がある。最尤推定量はそのような手法になっていて、その導

出のための isotonic 回帰法を用いたアルゴリズム等が提案されてきた。IERD 法を用いると許容のミニマックス推定量を含んだ改良型推定量のクラスを構成することができる。また多変量線形校正問題や統計的制御問題における古典的推定量を改良する一致推定量の構成や分散比の二重縮小推定量の導出においても IERD 法の使用が有用であることがわかってきた。その他にも、決定後または検定後の推定問題 (Dahiya (1974)), 線形回帰モデルで説明変数間に多重共線性が存在するときの安定な推定量の導出, 付加的情報の使用による推定量の改善などにおいて縮小の考え方が重要であり, 一般に, 推定が過大評価または過小評価している場合や, 推定があいまいで漠然としたものだったり不安定なものだったりしたときには, より安定したより確かな方向へ縮小することが望ましく, 何らかの意味でより優れた縮小推定量を導出することが大切であると思われる。

謝辞：査読者の有益なコメントに感謝します。

参 考 文 献

- [1] Akai, T., (1986). Simultaneous estimation of location parameters of the distribution with finite support. *Ann. Inst. Statist. Math.*, **38**, 85-99.
- [2] Arnold, B. C. (1970). Inadmissibility of the usual scale estimate for a shifted exponential distribution. *J. Amer. Statist. Assoc.*, **65**, 1260-1264.
- [3] Baranchick, A. (1973). Inadmissibility of maximum likelihood estimators in some multiple regression problems with three or more independent variables. *Ann. Statist.*, **1**, 312-321.
- [4] Berger, J. O. (1980). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of gamma parameters. *Ann. Statist.*, **8**, 545-571.
- [5] Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. 2nd Ed., Springer-Verlag, New York.
- [6] Bilodeau, M. and Kariya, T. (1989). Minimax estimators in the normal MANOVA model. *J. Mult. Anal.*, **28**, 260-270.
- [7] Bilodeau, M. and Srivastava, M. S. (1992). Estimation of the eigenvalues of $\Sigma_1 \Sigma_2^{-1}$. *J. Mult. Anal.*, **41**, 1-13.
- [8] Brandwein, A. C. and Strawderman, W. E. (1990). Stein estimation: The spherically symmetric case. *Statist. Sci.*, **5**, 356-369.
- [9] Brewster, J. F. (1974). Alternative estimators for the scale parameter of the exponential distribution with unknown location. *Ann. Statist.*, **2**, 553-557.
- [10] Brewster, J. F. and Zidek, J. V. (1974). Improving on equivariant estimators. *Ann. Statist.*, **2**, 21-38.
- [11] Brown, L. D. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *Ann. Math. Statist.*, **39**, 29-48.
- [12] Brown, L. D. (1971). Admissible estimators, recurrent diffusions, and insolvable boundary value problems. *Ann. Math. Statist.*, **42**, 855-904.
- [13] Brown, L. D. and Hwang, J. T. (1982). A unified admissibility proof. In *Statistical Decision Theory and Related Topics III*, (S. S. Gupta and J. Berger, Eds.), Academic Press, New York.
- [14] Calvin, J. A. and Dykstra, R. L. (1991). Maximum likelihood estimation of a set of covariance matrices under Lowner order restrictions with applications to balanced multivariate variance components models. *Ann. Statist.*, **19**, 850-869.
- [15] Casella, G. and Strawderman, W. E. (1981). Estimating a bounded mean. *Ann. Statist.*, **9**, 870-879.
- [16] Cellier, D., Fourdrinier, D. and Robert, C. P. (1989). Robust shrinkage estimators of the location parameter for elliptically symmetric distributions. *J. Mult. Anal.*, **29**, 39-52.
- [17] Chang, Y.-T. (1982). Stein-type estimators for parameters in truncated spaces. *Keio Sci. Tech. Rep.*, **35**, 185-193.
- [18] Clevenson, M. L. and Zidek, J. V. (1975). Simultaneous estimation of the mean of independent

- Poisson laws. *J. Amer. Statist. Assoc.*, **70**, 698-705.
- [19] Cohen, A. (1972). Improved confidence intervals for the variance of a normal distribution. *J. Amer. Statist. Assoc.*, **67**, 382-387.
- [20] Dahiya, R. C. (1974). Estimation of the mean of the selected population. *J. Amer. Statist. Assoc.*, **69**, 226-230.
- [21] DasGupta, A. (1986). Simultaneous estimation in the multiparameter Gamma distribution under weighted quadratic losses. *Ann. Statist.*, **14**, 206-219.
- [22] DasGupta, A. (1989). A general theorem on decision theory for nonnegative functionals: with applications. *Ann. Statist.*, **17**, 1360-1374.
- [23] Dey, D. K. and Srinivasan, C. (1985). Estimation of covariance matrix under Stein's loss. *Ann. Statist.*, **13**, 1581-1591.
- [24] Eaton, M. L. (1992). A statistical diptych: admissible inferences—recurrence of symmetric Markov chains. *Ann. Statist.*, **20**, 1147-1179.
- [25] Efron, B. and Morris, C. (1972). Empirical Bayes on vector observations: An extension of Stein's method. *Biometrika*, **59**, 335-347.
- [26] Efron, B. and Morris, C. (1976). Multivariate empirical Bayes estimation of covariance matrices. *Ann. Statist.*, **4**, 22-32.
- [27] Eguchi, S. (1991). A conjugate look at EM algorithm and its turbo form in exponential families. Unpublished manuscript.
- [28] George, E. I. (1986). Minimax multiple shrinkage estimation. *Ann. Statist.*, **14**, 188-205.
- [29] Ghosh, M., Hwang, J. T. and Tsui, K.-W. (1983). Construction of improved estimators in multiparameter estimation for discrete exponential families. *Ann. Statist.*, **11**, 351-367.
- [30] Ghosh, M., Nickerson, D. M. and Sen, P. K. (1987). Sequential shrinkage estimation. *Ann. Statist.*, **15**, 817-829.
- [31] Gleser, L. J. (1986). Minimax estimators of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix. *Ann. Statist.*, **14**, 1625-1633.
- [32] Goutis, C. and Casella, G. (1991). Improved invariant confidence intervals for a normal variance. *Ann. Statist.*, **19**, 2015-2031.
- [33] Haff, L. R. (1979). An identity for the Wishart distribution with applications. *J. Mult. Anal.*, **9**, 531-542.
- [34] Haff, L. R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix. *Ann. Statist.*, **8**, 586-597.
- [35] Haff, L. R. (1991). The variational form of certain Bayes estimators. *Ann. Statist.*, **19**, 1163-1190.
- [36] Honda, T. (1991). Minimax estimators in the MANOVA model for arbitrary quadratic loss and unknown covariance matrix. *J. Mult. Anal.*, **36**, 113-120.
- [37] Hudson, H. M. (1978). A natural identity for exponential families with applications in multiparameter estimation. *Ann. Statist.*, **6**, 478-484.
- [38] Hwang, J. T. (1985). Universal domination and stochastic domination: Estimation simultaneously under a broad class of loss functions. *Ann. Statist.*, **13**, 295-314.
- [39] Hwang, J. T. and Casella, G. (1982). Minimax confidence sets for the mean of a multivariate normal distribution. *Ann. Statist.*, **10**, 868-881.
- [40] Hwang, J. T. and Casella, G. (1984). Improved set estimators for a multivariate normal mean. *Statist. and Decisions*, Supplement Issue **1**, 3-16.
- [41] Hyakutake, H. and Siotani, M. (1987). The multivariate heteroscedastic method: Distribution of statistics and an application. *Amer. J. Math. Manage. Sci.*, **7**, 89-111.
- [42] Isogai, E. (1987). The convergence rate of fixed-width sequential confidence intervals for a probability density function. *Sequential Anal.*, **6**, 55-69.
- [43] James, W. and Stein, C. (1961). Estimation with quadratic loss. In *Proc. 4th Berkeley Symp. Math. Statist. Probab.*, **1**, 361-379.
- [44] Kariya, T. (1989). Equivariant estimation in a model with an ancillary statistic. *Ann. Statist.* **17**, 920-928.
- [45] Konno, Y. (1991). On estimation of a matrix of normal means with unknown covariance matrix. *J. Mult. Anal.*, **36**, 44-55.

- [46] Konno, Y. (1992). On estimating eigenvalues of the scale matrix of the multivariate F distribution. *Sankhya, A*, **54**, 241-251.
- [47] Kubokawa, T. (1987). Admissible minimax estimation of a common mean of two normal populations. *Ann. Statist.*, **15**, 1245-1256.
- [48] Kubokawa, T. (1991a). An approach to improving the James-Stein estimator. *J. Mult. Anal.*, **36**, 121-126.
- [49] Kubokawa, T. (1991b). A unified approach to improving equivariant estimators. METR 91-01, Dept. Math. Engrg. Inf. Phys., Univ. Tokyo. *Ann. Statist.* (1993), to appear.
- [50] Kubokawa, T., Honda, T., Morita, K. and Saleh, A. K. Md. E. (1990). Estimating a covariance matrix of a normal distribution with unknown mean. METR 90-12, Dept. Math. Engrg. Inf. Phys., Univ. Tokyo.
- [51] Kubokawa, T., Morita, K., Makita, S. and Nagakura, K. (1992a). Estimation of the variance and its applications. *J. Statist. Plan. Inf.*, to appear.
- [52] Kubokawa, T., Robert, C. and Saleh, A. K. Md. E. (1992b). Empirical Bayes estimation of the covariance matrix of a normal distribution with unknown mean under an entropy loss. *Sankhya, Series A*, **54**, 402-410.
- [53] Kubokawa, T., Saleh, A. K. Md. E. and Morita, K. (1992c). Improving on MLE of coefficient matrix in a growth curve model. *J. Statist. Plan. Inf.*, **31**, 169-177.
- [54] Lehmann, E. L. (1983). *Theory of Point Estimation*. Wiley, Ney York.
- [55] Leung, P. L. and Muirhead, R. J. (1987). Estimation of parameter matrices and eigenvalues in MANOVA and canonical correlation analysis. *Ann. Statist.*, **15**, 1651-1666.
- [56] Lin, P.-E. and Tsai, H.-L. (1973). Generalized Bayes minimax estimators of the multivariate normal mean with unknown covariance matrix. *Ann. Statist.*, **1**, 142-145.
- [57] Maatta, J. M. and Casella, G. (1990). Developments in decision-theoretic variance estimation. *Statist. Sci.*, **5**, 90-120.
- [58] 宮川雅巳 (1987). EM アルゴリズムとその周辺. 応用統計学, **16**, 1-19.
- [59] Nagata, Y. (1989). Improvements of interval estimations for the variance and the ratio of two variance. *J. Japan Statist. Soc.*, **19**, 151-161.
- [60] 永田 靖 (1991). 指数分布の尺度母数の改良信頼区間. 品質, **21**, 5-10.
- [61] Perron, F. (1990). Equivariant estimators of the covariance matrix. *Canadian J. Statist.*, **18**, 179-182.
- [62] Perron, F. (1992). Minimax estimators of a covariance matrix. *J. Mult. Anal.*, **43**, 16-28.
- [63] Proskin, H. M. (1985). An admissibility theorem with applications to the estimation of the variance of the normal distribution. Ph. D. dissertation, Dept. Statistics, Rutgers Univ.
- [64] Robert, C. (1991). Prior feedback: A Bayesian approach to maximum likelihood estimation. Unpublished manuscript.
- [65] Robert, C. and Casella, G. (1990). Improved confidence sets for spherically symmetric distributions. *J. Mult. Anal.*, **32**, 84-94.
- [66] Robertson, T., Wright, F. T. and Dykstra, R. (1988). *Order Restricted Statistical Inference*. North-Holland, New York.
- [67] Rukhin, A. L. and Sinha, B. K. (1991). Decision-theoretic estimation of the product of Gamma scales and generalized variance. Unpublished manuscript.
- [68] Sarkar, S. K. (1989). On improving the shortest length confidence interval for the generalized variance. *J. Mult. Anal.* **31**, 136-147.
- [69] Sen, P. K., Kubokawa, T. and Saleh, A. K. Md. E. (1989). The Stein paradox in the sense of the Pitman measure of closeness. *Ann. Statist.* **17**, 1375-1386.
- [70] Sen, P. K. and Saleh, A. K. Md. E. (1987). On preliminary test and shrinkage M-estimation in linear models. *Ann. Statist.*, **15**, 1580-1592.
- [71] Sengupta, D. and Sen, P. K. (1991). Shrinkage estimation in a restricted parameter space. *Sankhya, A*, **53**, 389-411.
- [72] Shao, P. Y.-S. and Strawderman. (1992). Improving on the James-Stein positive-part estimator. Unpublished manuscript.
- [73] Sheena, Y. and Takemura, A. (1992). Inadmissibility of non-order-preserving orthogonally invar-

- iant estimators of the covariance matrix in the case of Stein's loss. *J. Mult. Anal.* **41**, 117-131.
- [74] Shinozaki, N. (1980). Estimation of a multivariate normal mean with a class of quadratic loss functions. *J. Amer. Statist. Assoc.*, **75**, 973-976.
- [75] Shinozaki, N. (1984). Simultaneous estimation of location parameters under quadratic loss. *Ann. Statist.*, **12**, 322-335.
- [76] Shinozaki, N. (1989). Improved confidence sets for the mean of a multivariate normal distribution. *Ann. Inst. Statist. Math.*, **41**, 331-346.
- [77] 篠崎信雄 (1991). Stein タイプの縮小推定量とその応用. 応用統計学 **20**, 59-76.
- [78] Shiraishi, T. (1991). On positive-part shrinkage R- and M-estimation in one-way ANOVA. *J. Japan Statist. Soc.*, **21**, 61-72.
- [79] Shorrock, G. (1990). Improved confidence intervals for a normal variance. *Ann. Statist.*, **18**, 972-980.
- [80] Shorrock, R. B. and Zidek, J. V. (1976). An improved estimator of the generalized variance. *Ann. Statist.* **4**, 629-638.
- [81] Sinha, B. K. (1976). On improved estimators of the generalized variance. *J. Mult. Anal.* **6**, 617-626.
- [82] Sinha, B. K. and Ghosh, M. (1987). Inadmissibility of the best equivariant estimators of the variance-covariance matrix, the precision matrix, and the generalized variance under entropy loss. *Statist. and Decisions*, **5**, 201-227.
- [83] Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proc. 3rd Berkeley Symp. Math. Statist. Probab.*, **1** 197-206.
- [84] Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* **16**, 155-160.
- [85] Stein, C. (1973). Estimation of the mean of a multivariate normal distribution. In *Proc. Prague Symp. Asymptotic Statist.*, 345-381.
- [86] Stein, C. (1977). Lectures on multivariate estimation theory. (In Russian.) In *Investigation on Statistical Estimation Theory I.* (Zapiski Nauchych Seminarov LOMI im. V. A. Steklova AN SSSR vol. 74). Leningrad, 4-65.
- [87] Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.*, **9**, 1135-1151.
- [88] Stigler, S. M. (1990). A Galtonian perspective on shrinkage estimators. *Statist. Sci.*, **5**, 147-155.
- [89] Strawderman, W. E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Statist.*, **42**, 385-388.
- [90] Strawderman, W. E. (1974). Minimax estimation of powers of the variance of a normal population under squared error loss. *Ann. Statist.*, **2**, 190-198.
- [91] Sugiura, N. and Fujimoto, M. (1982). Asymptotic risk comparison of improved estimators for normal covariance matrix. *Tsukuba J. Math.*, **6**, 103-126.
- [92] Sugiura, N. and Konno, Y. (1988). Entropy loss and risk of improved estimators for the generalized variance and precision. *Ann. Inst. Statist. Math.*, **40**, 329-341.
- [93] Takada, Y. (1979). A family of minimax estimators in some multiple regression problems. *Ann. Statist.*, **7**, 1144-1147.
- [94] Takada, Y. (1984). Inadmissibility of a sequential estimation rule of the mean of a multivariate normal distribution. *Sequential Anal.*, **3**, 267-271.
- [95] Takada, Y. (1992). A sequential procedure with asymptotically negative regret for estimating a normal mean. *Ann. Statist.*, **20**, 562-569.
- [96] Takemura, A. (1983). An orthogonally invariant minimax estimator of the covariance matrix of a multivariate normal population. *Tsukuba J. Math.*, **8**, 367-376.
- [97] 竹内 啓 (1979). Stein 推定量の意味とその応用. 応用統計学 **8**, 81-95.
- [98] Takeuchi, K. (1991). Personal communications.
- [99] Tan, M. (1991). Improved estimators for the GMANOVA problem with application to Monte Carlo simulation. *J. Mult. Anal.*, **38**, 262-274.
- [100] Tate, R. F. and Klett, G. W. (1959). Optimal confidence intervals for the variance of a normal distribution. *J. Amer. Statist. Assoc.*, **54**, 674-682.

- [101] Wald, A. (1950). *Statistical Decision Functions*. Wiley, New York.
- [102] Woodroffe, M. and Takahashi, M. (1982). Asymptotic expansions for the error probabilities of some repeated significance tests. *Ann. Statist.*, **10**, 895-908
- [103] Zacks, S. (1970). Bayes and fiducial equivariant estimators of the common mean of two normal distributions. *Ann. Math. Statist.*, **41**, 59-69.
- [104] Zheng, Z. (1986). On estimation of matrix of normal mean. *J. Mult. Anal.*, **18**, 70-82.
- [105] Zidek, J. (1978). Deriving unbiased risk estimators of multinormal mean and regression coefficient estimators using zonal polynomials. *Ann. Statist.*, **6**, 769-782.

2. 統計的推測の漸近理論 —幾何的アプローチの展開から—

2.1 はじめに

この章では統計的推測の漸近理論を幾何的アプローチを中心にサーベイし、将来の展開への萌芽的研究も紹介したい。Kass (1987) は巻頭の紹介の中で R. A. Fisher のことを“*Our most enigmatic sage* (我々の不可思議な大いなる哲人)”と称した。Fisher が 1925 年論文の中で統計的推測の基本概念となる一致性、十分性、有効性を提出し、その三位一体の枠組の中で情報の損失とその回復が彼の推定理論の核心となっている。彼の一言々に対して多くの優秀な数理統計学者が長大な数学的道具を駆使して理解を完成して行ったのは 1980 年代である。更に驚くべきことに彼の業績の対を成すもう一方には集団遺伝学がある。自然選択の Fisher の基本定理が集団遺伝学の中心を形成している。最近の成果である進化的安定戦略 (ESS) の展開もこの枠組にある。(Fisher の数理進化論の貢献の総説は Karlin (1992) を、ESS については Lessard (1989) を参照されたい。)

2.2 漸近有効性

Fisher は標本 X から統計量 $T = T(X)$ への情報縮約による情報損失の尺度を $\Delta_n(T) \equiv I_\theta(X) - I_\theta(T)$ と定めた。ここで n は X の標本サイズ、 $I_\theta(X)$ と $I_\theta(T)$ は X と T の持つ情報行列とする。 T が十分統計量ならば $\Delta_n(T)$ は任意の n に対して消える。 X の従う分布が指数型ならば、最尤推定量は十分だから情報損失はないことより、指数型でないときが問題となった。最尤推定量は観測一回当たりの損失 $(1/n)\Delta_n$ が $n \rightarrow \infty$ の極限で消えることしか云えない。この性質 (1 次有効性) を満たす推定量のクラスの中で Fisher の主張は一般に“最尤推定量は n の極限で情報損失 Δ_n の下限を与える (2 次有効性)”と云うものである。

Fisher 予想の最尤推定量の 2 次有効性の解決の第一歩は Rao (1960, 1962, 1963) の三部作から始まった。情報損失 Δ_n の扱い易い形を与え、多項分布のワンパラメーター族の場合、最小 χ^2 推定量を含む幾つかの推定量と共に最尤推定量の情報損失の極限を求め、Fisher の計算の一部を訂正した。更にバイアス修正の操作をすれば推定量の 2 次有効性は分散の 2 次までの項を最小にすることを証明した。第二歩は Efron (1975) によって踏み出された。 X の従う分布が指数型分布の非線型ワンパラメーター族に広げられ、推定量 T の情報損失の極限は

$$\lim_{n \rightarrow \infty} \Delta_n(T) = \gamma_\theta^2 + (1/2)\beta_\theta^2(T)$$

で与えられる。ここで γ_θ は θ における指数線型性からの乖離度を表す量で Efron は統計曲率と呼んだ。 $\beta_\theta(T)$ は T に依存する量で T が最尤推定量のとき恒等的に消え、ゆえにその損失の極限は下界 γ_θ に到達することが証明された。最終版は、Amari (1982, 1985) の微分幾何による考察で成し遂げられた。確率密度関数の空間上で相異なる $p(x)$ と $q(x)$ を結ぶ測地線の一組

C_1 と C_2 を

$$C_1 = \{tp(x) + (1-t)q(x): 0 \leq t \leq 1\}, C_2 = \{c(t)p(x)^t q(x)^{1-t}: 0 \leq t \leq 1\}$$

と定める。ここで C_2 中の $c(t)$ は規格化定数である。この C_1 と C_2 によって誘導される線型接続を各々、ミクスチュア接続と指数接続と呼ぶ。モデルは指数型分布の非線型マルチパラメータ族 M に拡張され、情報損失行列は

$$\lim_{n \rightarrow \infty} \Delta_n(T) = G^2(H^{(e)}(M)) + (1/2)G^2(H^{(m)}(T))$$

と表された。ここで $H^{(e)}$ は M の指数接続に関する第2基本形式で、 $H^{(m)}$ は推定量 T の張るアンシラリー空間のミクスチュア接続に関する第2基本形式で、 G^2 は Gram 形式を表す。統計曲率 γ_0^2 の多次元版は $H^{(e)}(M)$ の Gram 形式で表現された。この二つの線型接続は Fisher 計量の下で互いに双対で、統計モデルと推定を同時に評価できる物差の役割を演じ、2次有効性理論が完成された。4節においてその後の発展の中でこの双対構造が重要な鍵を担うことになる。Eguchi (1983) は、最小コントラスト推定量の情報損失の公式を与え、最尤推定量を含む2次有効な推定量のワンパラメータ族を提出した。これより2次有効な推定量のクラスの中で更に高次漸近挙動の構造決定が問題となるがこれは未解決である。

漸近有効性の大きな貢献として竹内と赤平の業績が挙げられる。彼等による漸近有効性の定義は、分布の集中度によるものである。推定量を中央値不偏になるように操作し、真値の廻りでの確率を Neyman-Pearson 検定によって得られた上界を基準にして比較するものである。これらの結果は Akahira and Takeuchi (1981) にまとめられている。彼らの結果は曲指数族を越えるより一般の正則な分布族について成り立つが、その結果は甘利と公文による微分幾何学を用いる表現と一致している。また Pfanzagl (1985) も分布の漸近展開の方法によって漸近有効性を考察している。Edgeworth 展開の幾何公式は Amari and Kumon (1983) によって与えられている。

Le Cam (1970) は分布族の正則性について解析的に厳密に考察している。超有効性の問題の一解決は Inagaki (1970) によって与えられた。この方向をまとめたものに Lauritzen (1988), Torgerson (1991) がある。

2.3 その他の漸近的性質

統計曲率のその他の応用は次に挙げられる。検定の漸近理論において Kumon and Amari (1983) は1次、2次有効な検定の3次検出力損失関数を統計曲率を用いて具体的に示した。Le Cam (1964) の提案した推定量 T の欠損量は Skogaard (1985) によって

$$\lim_{n \rightarrow \infty} n \sup_B \{P_\theta(B) - \hat{P}_\theta(B)\}^2 \leq \gamma_0^2$$

と表された。ここで P_θ は X の確率測度、 \hat{P}_θ は T の誘導する確率測度とする。欠損量のより高度な研究は Akahira (1986) を参照されたい。Fu (1982) は統計曲率の役割を大偏差解析による推定の研究の中で次の形で明示した。

$$\lim_{\epsilon \rightarrow 0} \epsilon^{-4} \{B(\theta, \epsilon) - \beta(T, \epsilon)\} \geq (1/8)\gamma_0^2$$

ここで $B(\theta, \epsilon) = \lim_{n \rightarrow \infty} n^{-1} \log P\{\|T - \theta\|^2 > \epsilon\}$, $\beta(T, \epsilon) = \inf_{\theta^*} \{KL(\theta^*, \theta): \|T - \theta\|^2 > \epsilon\}$. 上の不等式の等号は T が最尤推定量のとき成立する。Fu and Kass (1984) も参照。Eguchi (1984) は、回帰分析における残差平方和を上モデル M へ拡張し、推定量 T に対して M の余次元が f のとき、

$$\lim_{n \rightarrow \infty} n E_{\theta} \|\bar{x} - \eta(T)\|^2 - f \geq (1/2) \gamma_{\theta}^2 - \langle H^{(e)}(M), (3/2)H^{(e)}(M) - (1/2)H^{(m)}(M) \rangle$$

が成立し、等号はやはり T が最尤推定量のとき成立する。推定された残差の最小化によって得られる重み付け最小自乗推定量が提案され極限損失の計算と集団遺伝学の問題に適用された (Eguchi (1986, 1990) 参照)。これらの3つの不等式は Rao や竹内と赤平によって工夫されたバイアス修正の操作が必要ない。これは尺度そのものがパラメータ効果がないからだと思われる。Efron (1982) に関連した考察が見られる。最尤法はパラメータ推定としてではなくパラメータによって指定された確率密度関数の決定 (サマリーと呼ばれる) に最適性があると云う主張である。統計モデルの中で最も自然と思われる正規分布の平均パラメータの推定においてさえも平均自乗誤差を損失とするととき最尤法推定量は最適でない。このような Stein 現象は前章を参照。

最尤推定値 $\hat{\theta}$ を求めるための Fisher スコアアルゴリズムはスコア関数を $S(\theta)$ とするとき、反復列 $\{\theta_k\}_k$ が初期値 θ_0 から $\theta_{k+1} = \theta_k + I^{-1}(\theta_k)S(\theta_k)$ ($k=0, 1, \dots$) と定められる。このとき、

$$\frac{\|\theta_{k+1} - \hat{\theta}\|}{\|\theta_k - \hat{\theta}\|} \propto \gamma_{\theta}$$

が成立する (Kass (1984), Eguchi (1985) 参照)。最尤推定値を求めるもう一つの Dempster, Laird and Rubin (1977) による EM アルゴリズムは収束は遅いが大域収束性を持つ方法として有力である。最尤推定値の存在性の考察は Nakamura (1991) にある。

Cox and Reid (1987) はパラメータ直交性から近似条件付き推測の構成を議論している。条件付き推測の興味ある考察は Yanagimoto (1989) にある。Barndorff-Nielsen (1986 a, b) は条件付き推測に対してもう一つの微分幾何的方法を考えた。Riemann 計量として Fisher 情報量ではなくて観測値に依存した Fisher 情報量を採用した。即ち、最尤推定量 $\hat{\theta}$ と漸近補助統計量 a が存在して観測 Fisher 情報量が $\mathfrak{F}_{\theta}(\hat{\theta}, a)$ と書けるとき計量の成分を $g(\theta) \equiv \mathfrak{F}_{\theta}(\theta, a)$ によって定義した。これにより観測値に依存した線型接続が定義され、Barndorff-Nielsen and Cox (1979) で得られた最尤推定量の条件付き密度関数の公式

$$p(\hat{\theta})|a, \theta = cL(\theta)/L(\hat{\theta}) \quad (2.1)$$

が拡張された。Efron and Hinkley (1978) 及びサドルポイント法との関連は Reid (1988) に議論されている。これはロケーションモデルにおける Fisher の公式の拡張に成功している。更に Barndorff-Nielsen, Blæsild and Eriksen (1989) は変換モデルに対してこの枠組の中で Lie 群論を展開している。

Bartlett (1937) は対数尤度比検定において帰無仮説の下での尤度比統計量の χ^2 分布の近似の為に定数倍する補正を提案した。それにより尤度比統計量の仮説の下での分布が $1/n$ のオーダーまで χ^2 分布に一致できる。この Bartlett 補正について Vos (1989) は幾何的理論を展開した。関連文献に Barndorff-Nielsen and Cox (1984) がある。

2.4 幾何的アプローチの最近の展開について

Fisher の直観を解く有力な鍵は“幾何的アプローチ”ではあったが、ある意味で既に予告されているストーリーを再解釈したに過ぎないという批判もあるだろう。しかしながら、1980年代後半からこの方法論が漸近理論の枠を超えて新しい広がりを見せるようになった。アメリカ数理統計協会が“統計科学”というキーワードと共に新雑誌を刊行した同時期である。この微分幾何的展開には二つの方向がある：一つは新しい幾何学の建設を目指す方向、もう一つは方

法論の対象を統計学に留まらず数理諸分野へ飛躍する方向である。勿論この内なる方向と外なる方向は表裏一体で数理統計と離別したものでなく、より広い翼の統計科学に含まれると思われる。

前者の方向の萌芽は Nagaoka and Amari (1982) に見られる。それに刺激を受けた Lauritzen (1987) は現代微分幾何のコーディネイトフリーの言葉によって双対線型接続を研究した。それは Riemann 空間 (M, g) 上の線型接続 ∇ と ∇^* を平均接続 $(\nabla + \nabla^*)/2$ が g に関する計量接続になる時、双対的であると呼び、更に $T(X, Y, Z) \equiv g(\nabla_X Y - \nabla_Y X, Z)$ が X, Y, Z に関して対称であるとき (M, g, T) を統計多様体と呼んだ。 ∇ と ∇^* に関する Riemann 曲率テンソルを R と R^* とするとき、

$$g(R(X, Y)Z, W) = g(Z, R^*(Y, X)W)$$

が成立することより M の R 平坦と R^* 平坦の同値性が示された。Lauritzen は $R = R^*$ の同値条件を提出した。Amari (1985) は R -平坦な区間 M 上に Legendre 変換: $\phi(\eta) \equiv \inf_{\theta} \{\theta\eta - \Psi(\theta)\}$ から作られるダイバージェンス関数 D に関してピタゴラスの定理を証明した。即ち、 M 上の三点 p, q, r に対して p から q へ結ぶ ∇ -測地線と r から q へ結ぶ ∇^* -測地線が点 q で g の意味で直角に交わる時、

$$D(p, q) + D(q, r) = D(p, r)$$

となる。この様に R -平坦な空間 M 上に双対 Euclid 的世界像が連想される。これが Fisher プログラムの“十分性”から派生した指数型分布族の数学的拡張である。実は R -平坦な空間 M を純粋に幾何的な発想から Shima (1976, 1980, 1986) は Hessian 多様体と呼んだ。実際、 M 上で g の成分は座標系 θ と η で各々、 $\phi(\theta)$ と $\phi(\eta)$ の Hessian で表される。 ∇ と g に対する Laplacian と自己共役な楕円型微分作要素の関係が明かにされている。また Kurose (1990) はアフィン微分幾何の立場から双対接続 ∇ と ∇^* を研究した。Eguchi (1985, 1992) は多様体上のコントラスト関数が、自然に計量と双対接続を生成することを示し、更に反対称部分が Riemann 曲率となるテンソルの双対対称性を研究した。野水も共同研究の中で双対接続の考察を始めた (Nomizu and Pikall (1987) と Dilen, Nomizu and Vranken (1990) 参照)。

Barndorff-Nielsen の研究グループの一連の論文活動は条件付き原理から見出された計量、共変微分、高次微分の数学的性質を抽出することに成功している。Barndorff-Nielsen and Blæsild (1987), Barndorff-Nielsen and Jupp (1988), Barndorff-Nielsen and Blæsild (1988), Barndorff-Nielsen, Blæsild and Eriksen (1989), Barndorff-Nielsen and Jupp (1989), Barndorff-Nielsen (1990), Blæsild (1991) 等、精力的な研究がなされている。彼らは微分ストリング、不変 Taylor 展開など統計的推測の応用に現れた数学的側面を更に抽象的に考え、ヨーク幾何と呼んでいる。

この様に統計的推測を対称にして生まれた“幾何的アプローチ”は Riemann 幾何学を超えて双対性が本質的な役割を果たす新しい幾何学に発展する兆しが見えつつある。これは数理物理学の世界から眺めれば自然な方向と思われる。Newton, Gauss, Einstein の例を挙げるまでもなく数学と物理学は交互に刺激し、培い合った長い歴史があり、現代も数学の一大主流をなしている。この 50 年近く経て Fisher 予想の解決、統計的推測論の理解に有効だった方法論が数学的にも興味あるものを提供すると考えておかしくないだろう。少なくとも 1980 年代まで幾何学者が全く気がつかなかった微分可能多様体の豊富な例に“統計モデル”或は“統計パラメータ”があることは云えるだろう。

もう一つの方向は幾何化の対象の拡大にある。最初に挙げられるのは上の流れとは独立に進

められた Akin (1979, 1982, 1990) の集団遺伝学の幾何化がある。彼は Shashahani (1979) が提案した計量によって生物集団の多形性、性比、適応度、エピスタシスなど世代に関する生物集団の力学系の理論を展開している。力学系を記述するベクトル場が Shashahani 計量に関するグラジェント場でないとき Hopf 分岐が起こることを証明し進化ゲーム論を展開している。ところで、この計量は多項分布の Fisher 計量そのものである。この様に Fisher が息吹を与えた統計学と集団遺伝学が独立に幾何化されたことは興味深い。

甘利は共同研究を通し、脳の神経回路網の機構を微分幾何の方法で挑戦している。Boltzmann マシンの成す多様体に双対平坦性が自然に導入されている (Amari (1990), Amari (1991), Amari, Kurata and Nagaoka (1992) Amari, Fujita and Sinomoto を参照。) この研究に先駆けてシステム制御理論と多元情報コード理論の幾何化にも成功を収めている (Amari (1987, 1989) と Amari and Han (1989), 関連する文献は Ravishanker, Melnick and Tsai (1990) を参照)。

2.5 おわりに

この 60 年の歩みの中で科学は目まぐるしい進展を遂げた。例えば、Fisher たちによって創られた集団遺伝学の中では遺伝子とは単なる数学的な記号 (A, B, \dots) に過ぎなかった。現代の生物学によると遺伝子は塩基対配列の化学物質として観測できる。その 1 次構造は各生物種について Gen Bank, EMBL などのデータバンクから手軽に利用できる時代である。アメリカではヒトの全遺伝情報を解析する計画 (HGP) が国家的事業として進められている。Neyman (1971) は早くから DNA 列の統計解析の重要性を訴えている。もし Fisher ならばどんな解析をしていたらうか? Felsenstein (1983) や Kishino and Hasegawa (1989) によって進化系統樹の最尤法が精力的に研究されている。

気象学から端を発したカオス理論、フラクタル次元論、ファジー推論など統計科学の展開に有力なアプローチが近年盛んに研究され、我々が解析できる現象が拡大されつつある (カオスの統計的予測、制御に関しては Casdaghi (1992) など *J. Roy. Statist. Soc.* の第 54 巻に特集を、フラクタル次元の推定は Taylor and Taylor (1991), ファジー推論は Toley and Manton (1992) などを参照)。

現代における統計学の目指すべき方向は科学としての統計学があると割り切ることで、不毛な思想的対立、極端なまでの自己完結性の追求、その有用性との葛藤から飛躍できるのではないかと思われる。統計学会の会報 No. 73 において竹内啓氏が環境問題について随筆で述べた様に、21 世紀を迎えつつある現代の環境問題は限りなく地球惑星そのもの、或はその物質循環のトータルシステムの構造の同定が解決を急ぐ大きな問題となっていると思われる。このテーマに対して統計学の貢献が正否の鍵を握っているのではないだろうか。

謝辞：査読者の有益なコメントに感謝します。

参 考 文 献

- [1] Akahira, M. (1986). *The structure of asymptotic deficiency of estimators*. Queen's Papers in Pure and Applied Math. 75, Queen's University Press, Kingston.
- [2] Akahira, M. and K. Takeuchi (1981). *Asymptotic efficiency of statistical estimators: Concepts and higher order asymptotic efficiency*. Lec. Note in Statist. 7, Springer-Verlag, New York.
- [3] Akin, E. (1979). *The geometry of population genetics*. Lec. Note in Biomathematics 31, Springer-Verlag, New York.
- [4] Akin, E. (1982). Exponential families and game dynamics. *Can. J. Math.* 34, 374-405.

- [5] Akin, E. (1990). The differential geometry of population genetics and evolutionary games. In "Mathematical and statistical developments of evolutionary theory", ed. S. Lessard, NATO ASI Series, 1-93.
- [6] Amari, S. (1982). Differential geometry of curved exponential families-Curvatures and information loss. *Ann. Statist.* **10**, 357-385.
- [7] Amari, S. (1982). Geometrical theory of asymptotic ancillarity and conditional inference. *Biometrika* **69**, 1-17.
- [8] Amari, S. (1985). *Differential-geometrical methods in statistics*. Lec. Note in Statist. 28, Springer-Verlag, New York.
- [9] Amari, S. (1987). Differential geometry of a parametric family of invertible linear systems-Riemannian metric, dual affine connections and divergence. *Math. System Theo.* **20**, 53-82.
- [10] Amari, S. (1987). Differential geometrical theory of statistics. In *Differential Geometry in Statistical inference*, IMS, 19-94, California,
- [11] Amari, S. (1989). Fisher information under restriction of Shannon information in multiterminal situations. *Ann. Inst. Statist. Math.* **41** 623-648.
- [12] Amari, S. (1990). Mathematical foundations of neuro computing. *Proc. IEEE* **78**, 1443-1463.
- [13] Amari, S. (1991). Dualistic geometry of the manifold of higher-order neurons. *Neural Network* **4**, 443-451.
- [14] Amari, S. and Kumon, M. (1983). Differential geometry of Edgeworth expansion in curved exponential family. *Ann. Inst. Statist. Math.* **35**, 1-24.
- [15] Amari, S. and Kumon, M. (1988). Estimation in the presence of infinitely many nuisance parameters-Geometry of estimating functions. *Ann. Statist.* **16**, 1044-1068.
- [16] Amari, S. and Han, T. S. (1989). Statistical inference under multi-terminal rate restrictions-A differential geometric approach. *IEEE Trans. Infor. Theory* **35**, 217-227.
- [17] Amari, S. Kurata, K. and Nagaoka, H. (1992). Information geometry of Boltzmann machines, *IEEE Trans. Neural Network* **3**, 260-271.
- [18] Amari, S., Fujita, N. and Sinomoto, S. Four types of learning curves. *Neural Computation* **4**, 605-618.
- [19] Barndorff-Nielsen O. E. (1986 a). Inference on full or partial parameters based on the standardized signed log likelihood ratio. *Biometrika* **73**, 307-322.
- [20] Barndorff-Nielsen O. E. (1986 b). Likelihood and observed geometris. *Ann. Statist.* **14**, 856-873.
- [21] Barndorff-Nielsen O. E. (1986 c). Strings, tensorial combinants, and Bartlett adjustments. *Proc Roy. Soc. London A* **406**, 127-137.
- [22] Barndorff-Nielsen O. E., Cox D. R. and Reid, N. (1986). The role of differential geometry in statistical theory. *Int. Statist. Rev.* **54**, 83-96.
- [23] Barndorff-Nielsen O. E. and Cox, D. R. (1979). Edgeworth and saddlepoint approximations with statistical applications (with Discussion). *J. Roy. Statist. Soc. B* **41**, 279-312.
- [24] Barndorff-Nielsen O. E. and Cox, D. R. (1984). Bartlett adjustments to likelihood ratio statistic and the distribution of the maximum likelihood estimator. *J. R. Statist. Soc. B* **46**, 483-495.
- [25] Barndorff-Nielsen O. E. and Blæsild, P. (1987). Strings: mathematical theory and statistical examples. *Proc. Roy. Soc. London A* **411**, 155-176.
- [26] Barndorff-Nielsen O. E. and Blæsild, P. (1987). Derivative strings: contravariant aspects. *Proc Roy. Soc. London A* **411**, 421-444.
- [27] Barndorff-Nielsen O. E. and Jupp, P. E. (1988). Differential geometry, profile likelihood, L-sufficiency and composite transformation models. *Ann. Statist.* **16**, 1009-1043.
- [28] Barndorff-Nielsen O. E. and Blæsild, P. (1988). Strings: mathematical theory and statistical examples. *Proc Roy. Soc. London A* **411**, 155-176.
- [29] Barndorff-Nielsen O. E., Blæsild, P. and Eriksen P. R. (1989). *Decomposition and Invariance of measures, and Statistical transformation models*. Lec. Note in Statist. 58, Springer-Verlag, New York.
- [30] Barndorff-Nielsen O. E., (1990). Approximate interval probabilities. *J. R. Statist. Soc. B* **52**, 485-496.
- [31] Barndorff-Nielsen O. E., (1991). Modified signed log likelihood ratio. *Biometrika* **78**, 557-563.

- [32] Blæsild, P. (1991). Yokes and Tensors derived from yokes. *Ann. Inst. Statist. Math.* **43**, 95-113.
- [33] Casdagli, M. (1992). Chaos and deterministic versus non-linear modeling. *J. Roy. Statist. Soc.* **54**, 303-328.
- [34] Cox, D. R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference (with Discussion). *J. R. Statist. Soc. B* **49**, 1-39.
- [35] Cook, R. D. (1986). Assessment of local influence. (with Discussion). *J. R. Statist. Soc. B* **48**, 133-169.
- [36] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977) Maximum likelihood from incomplete data via the EM algorithm (with Discussion), *J. R. Statist. Soc. B* **39**, 1-38.
- [37] Dilen, F., Nomizu, K. and Vranken, L. (1990). Conjugate connections and Radon's theorem in affine differential geometry. *Monatsefte Math.* **109**, 221-235.
- [38] Efron, B. (1975). Defining the curvature of a statistical problem. *Ann. Statist.* **3**, 1189-1242.
- [39] Efron, B. (1982). Maximum likelihood and decision theory. *Ann. Statist.* **10**, 340-356.
- [40] Efron, B. and Hinkley, D. V. (1978). Assessing the accuracy of the maximum likelihood estimator: Observed vs. expected Fisher information, (with Discussion) *Biometrika*, **65**, 457-487.
- [41] Eguchi, S (1983). Second order efficiency of minimum contrast estimators in a curved exponential family. *Ann. Statist.* **11**, 793-803.
- [42] Eguchi, S (1984). A characterization of second order efficiency in a curved exponential family. *Ann. Inst. Statist. Math.* **36 A**, 199-206.
- [43] Eguchi, S (1985). A differential geometric approach to statistical inference on the basis on contrast functionals. *Hiroshima J. Math.*, 341-391.
- [44] Eguchi, S (1986). A projection method of estimation for a subfamily of exponential families. *Ann. Inst. Statist. Math.* **38 A**, 389-398.
- [45] Eguchi, S. and M. Matsuura (1990). Testing the Hardy-Weinberg equilibrium in the HLA system. *Biometrics* **46**, 415-426.
- [46] Eguchi, S. (1992). Geometry of minimum contrast. *Hiroshima J. Math.* **22**, 631-647.
- [47] Fisher, R. A. (1925). Theory of statistical estimation. *Proc. Cambridge Phils. Soc.* **22**, 700-725.
- [48] Fisher, R. A. (1934). Two new properties of mathematical likelihood. *Proc. R. Soc. A* **144**, 285-307.
- [49] Felsenstein, J. (1983). Statistical Inference of phylogenies. *J. R. Statist. Soc. A* **146**, 246-272.
- [50] Fu, J. C. (1982). Large sample point estimation: A large deviation theory approach. *Ann. Statist.* **10**, 762-777.
- [51] Fu, J. C. and Kass, R. E. (1984). A note on the interpretation of the Bahadur bound and the rate of convergence of the maximum likelihood estimator. *Statist. Probab. Lett.* **2**, 269-273.
- [52] Inagaki, N. (1970). On the limiting distribution of a sequence of estimators with uniformity property. *Ann. Inst. Statist. Math.* **22**, 1-13
- [53] Karlin, S. (1992). R. A. Fisher and evolutionary theory. *Statist. Sci.* **7**, 13-33.
- [54] Kass, R. E. (1984). Canonical parametrization and zero parameter effect curvature. *J. Roy. Statist. Soc.* **46**, 86-92
- [55] Kass, R. E. (1987). Introduction. In *Differential Geometry in Statistical inference*, IMS, California,
- [56] Kass, R. E. (1989). The geometry of asymptotic inference. *Statist. Sci.* **4**, 188-234.
- [57] Kishino, H and Hasegawa, H (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* **29**, 170-179.
- [58] Korose T. (1990). Dual connections and affine geometry. *Math. Z.* **203**, 115-121.
- [59] Kumon, M. and Amari, S. (1983). Geometrical theory of higher-order asymptotics of test, interval estimator and conditional inference. *Proc. R. Soc. London A* **387**, 429-458.
- [60] Kumon, M. and Amari, S. (1984). Estimation of structural parameter in the presence of a large number of nuisance parameters. *Biometrika* **71**, 445-459.
- [61] Le Cam, L. (1964). Sufficiency and approximate sufficiency. *Ann. Math. Statist.* **35**, 1419-1455.
- [62] Le Cam, L. (1970). On the assumption used to prove asymptotic normality of maximum likelihood estimates. *Ann. Math. Statist.* **41**, 802-828.
- [63] Lessard, S. (1989). Resource allocation in Mendelian populations: Further in ESS theory. In

- Mathematical evolutionary theory*, 197-246. Princeton Univ. Press, Princeton.
- [64] Lauritzen, S. L. (1987). Statistical manifolds. In *Differential Geometry in Statistical inference*, IMS, California, 163-216.
- [65] Lauritzen, S. L. (1988). *Extremal families and systems of sufficient statistics*. Lec. Note in Statist. 49, Springer-Verlag, New York.
- [66] Nagaoka and Amari (1982). Differential geometry of smooth families of probability distributions, *METR* 82-7, Univ. Tokyo.
- [67] Nakamura, T. (1991). Existence of maximum likelihood estimates for Interval-censored data from some three-parameter models with shifted origins. *J. Roy. Statist. Soc.* **53**, 211-22.
- [68] Neyman, J. (1971). Molecular studies of evolution: a source of novel statistical problems. In *Statistical decision theory and related topics*, ed. S. Gupta, 1-27, Academic Press, New York.
- [69] Nomizu, K. and Pikall, O. (1987). On the geometry of affine immersions. *Math. Z.* **195**, 165-178.
- [70] Okamoto, I., Amari, S. and Takeuchi, K. (1991). Asymptotic theory of sequential estimation: Differential geometric approach. *Ann. Statist.* **19**, 961-981.
- [71] Pfanzagl, J. (1985), *Asymptotic expansion for general Statistical models*. Lec. Note in Statist. 31, Springer-Verlag, New York.
- [72] Pfanzagl, J. (1990). *Estimation in semiparametric models*. Lec. Note in statist. 63, Springer-Verlag, New York.
- [73] Rao, C. R. (1961). Asymptotic efficiency and limiting information. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1**, 531-545.
- [74] Rao, C. R. (1962). Efficient estimates and optimum inference procedures in large samples (with discussion). *J. Roy. Statist. Soc. B* **24**, 46-72.
- [75] Rao, C. R. (1963). Criteria of estimation in large samples. *Sankya* **25**, 189-206.
- [76] Ravishanker, N., Melnick, E. L. and Tsai, C-L. (1990). Differential geometry of ARMA models. *J. Time Series Anal.* **11**, 259-274.
- [77] Reid, N. (1988). Saddlepoint methods and statistical inference. *Statist. Sci.*, **3**, 213-237.
- [78] Shima, H. (1976). Symmetric spaces with invariant locally Hessian structures. *J. Math. Soc. Japan* **29**, 581-589.
- [79] Shima, H. (1980). Homogeneous Hessian manifolds. *Ann. Inst. Fourier* **30**, 91-128.
- [80] Shima, H. (1986). Vanishing theorems for compact Hessian manifolds. *Ann. Inst. Fourier* **36**, 183-205.
- [81] Skogaard, I. M. (1985). A second-order investigation of asymptotic ancillarity. *Ann. Statist.* **13**, 534-551.
- [82] Toley, D. H. and Manton K. G. (1992). Large sample properties of a discrete grade of membership model. *Ann. Inst. Statist. Math.* **44**, 85-95.
- [83] Torgerson, E. (1991). Comparison of statistical experiments. *Encyclopedia of Mathematics and its Applications*, Cambridge.
- [84] Taylor, C. C. and Taylor, S. J. (1991). Estimating the dimension of a fractal. *J. Roy. Statist. Soc.* **53**, 353-364.
- [85] Vos, P. (1989). Fundamental equations for statistical submanifolds with applications to the bartlett correction. *Ann. Inst. Statist. Math.* **41**, 429-450.
- [86] Yanagimoto, T. (1991). Estimating a model through the conditional MLE. *Ann. Inst. Statist. Math.* **43**, 735-746.

3. 検定論の最近の展開

3.1 はじめに

ここでの目的は検定論の最近の研究を概観することである。検定論は推定論と並んで数理統計学の中心的な課題であり、全体を詳細に把握することは困難であるので、以下では国際的に重要と思われる事項について触れるとともに、日本人の業績に重点をおいて検定論を概観する。

統計的決定理論の観点からの検定論の理論の枠組みは Lehmann (1959) の教科書に代表されるように 1950 年代に確立された。この意味では検定理論のそれ以降の発展は 50 年代に確立された枠組みの中での精緻化と言ってよい。実際、検定の一般理論に関する最近の文献はあまり多くない。しかしいくつかの分野では、検定論が活発に研究されている。例えば分布の適合度検定、ノンパラメトリック検定、多変量解析などの分野である。また、検定における漸近理論の発展も注目される。第 2 章で概観されているように漸近理論は主に推定論において発展したものであるが、漸近理論の道具となる漸近展開の手法は検定論の研究にも有用である。ただし推定論と比較して見ると、検定論における漸近理論の結論は推定論におけるほど明快なものではないと思われる。

検定の一般理論の研究があまりさかんでない一つの理由として、伝統的な検定の考え方に対する疑問があげられる。伝統的な検定論の枠組みでは、第 1 種の過誤と第 2 種の過誤が非対称的に扱われている。また帰無仮説も非常に明確に定義されたものでなければならない。このような検定の考え方は非現実的であるとして、検定の問題をより広くモデル選択の問題と考える立場が有力になりつつある。モデル選択の問題は多方面で論じられていることもあり、ここではモデル選択の問題にはこれ以上触れない。ただし例えば適合度検定は、伝統的な検定というよりはモデル選択の方法という側面が強くモデル選択の考え方に近いものと言うこともできる。

以下では、検定の一般理論について述べた後、不変検定、順序制約下での検定、尤度比検定等の漸近理論、適合度検定について述べる。多変量解析については別稿にゆずり、ノンパラメトリック検定に関してもスペース等の関係で以下では省略する。

3.2 検定の一般理論

検定の一般論に関しては現在でも Lehmann の教科書 (第 2 版: Lehmann (1986)) が最もまとまったものであると思われる。1959 年の第 1 版と比較して第 2 版は 2 倍近くの厚さとなりこの間の検定論の発展を裏づけているが、内容を比較して見ると検定論の枠組みはほとんど変化しておらず、個別的な内容の充実が主である。第 2 版で追加された部分としては、Lehmann 自身の貢献を多く含むノンパラメトリック検定の分野の諸結果、多変量線形モデルに関する 1 章、及び補助統計量 (ancillary) に基づく条件つき推測に関する最終章があげられる。ノンパラメトリック検定及び多変量線形モデルの分野では、60 年代 70 年代を通じて様々の検定の性質が一般的な検定論の枠組みの中で詳しく調べられて来ており、それらの結果が盛り込まれている。統計量 A の分布が関心のあるパラメータに依存しない時 A を補助統計量という。補助統計量が存在する場合には、統計的推測を補助統計量を与えた時の条件つき分布に基づいて行うべきだとする条件つき推測の基準はある程度の説得性を持っているが、これを支持する決定論的観点からの明快な結果は得られていないように思われる。漸近理論の枠組みでは、第 2 章に触れているように漸近的な補助統計量に基づく条件つき推測の理論が発展しつつある。

決定理論の観点からの検定論の基礎的研究として Brown, Marden, Perlman, Cohen らの継

統的な研究 (Marden (1982a, b), Cohen and Sackrowitz (1987), Brown and Marden (1989), Cohen and Marden (1989), Marden and Perlman (1990) 等) が注目される。彼らは、許容的な決定方式がベイズ決定方式の(汎弱位相に関する)極限として得られるという一般的結果を応用することによって、様々な検定問題について検関数の完全類を与えている。これにより多くの検定の許容性を示すとともに、いくつかのよく知られた検定についてはそれらが非許容的であることを示している。非許容性の結果は興味深い。また検定の不偏性についても考察している。Marden (1982a) では同一の仮説につき複数 (n 個) の独立な検定結果が得られる場合に、各検定において観測された有意水準 (p -値) $p_i, i=1, \dots, n$, を総合する検定方式の許容性を議論している。そして例えば $-2\sum \log p_i$ を自由度 $2n$ のカイ二乗分布と比較する Fisher の方式が許容的であることを示している。Cohen and Sackrowitz (1987) は複数 (k 個) の母集団において対応する母数 θ_i が等しい ($H_0: \theta_1 = \dots = \theta_k$) ことを検定する k 標本の等値性の検定の不偏性を論じ、指数型分布族で各母集団からの標本数が等しい場合について、多くの検定が不偏であることを示している。これは正規分布等に関する個別の結果を一般化したものである。Cohen and Marden (1989) では、 k 個の正規母集団の分散の等値性の k 標本問題について各母集団からの標本分散を $s_i^2, i=1, \dots, k$, とする時、 $\max s_i^2 / \sum s_i^2$ に基づく Cochran の検定及び $\max s_i^2 / \min s_i^2$ に基づく Hartley の検定が非許容的となることを示している。

3.3 不変検定

検定の不変性は検定問題が対称性を持つ場合にしばしば前提とされる基準である。数学的には対称性は群の作用という形で扱われるので、不変検定を論じる際には群及び群の作用にともなう不変測度の概念が用いられる。特に多変量解析の分野では群の作用として一般線形群や直交群などの連続群を考えるので数学的にやや高度な議論が必要となる。

多変量解析等で必要とされる不変性の群及び不変測度に関する数学的な諸結果は、R. A. Wijsman, S. A. Andersson, 刈屋, M. L. Eaton などによって整理された。Eaton (1989) 及び Wijsman (1990) は連続群及び不変測度の理論に関する教科書として有用であるとともに、不変性の観点から様々な統計的問題について整理して議論している。不変検定の観点から特に重要なのは、最大不変量の尤度比を密度関数の不変測度に関する積分の比を用いて表す Wijsman (1967) の定理である。この定理の導出と証明は Wijsman (1990), Andersson (1982), Kariya (1985), Kariya and Sinha (1988) に与えられている。

Wijsman の定理により最大不変量の尤度比が求められることから、これに Neyman-Pearson の補題を応用することにより最強力不変検定を得ることが考えられる。最強力不変検定は通常対立仮説に依存する。しかしながら刈屋は、最大不変量の尤度比の微係数に対して Neyman-Pearson の補題を適用することにより、多変量解析における GMANOVA 等のいくつかの複雑な検定問題について、局所最強力不変検定が存在することを示し明示的な形で導出した (Kariya (1978, 1981a), Eaton and Kariya (1983) 等)。局所最強力不変検定は、不変検定のうちで帰無仮説の近くの対立仮説に関して最も高い検出力を持った検定である。GMANOVA は MANOVA (多変量分散分析) の母数に制約の加わった複雑なモデルであり、局所最強力検定のような最適性を持つ検定が明示的に得られることは興味深い。これらの結果については Kariya (1985) でまとまった解説が与えられている。

また刈屋は不変検定の方法を用いて「検定のロバストネス」の問題を系統的に研究している。すなわち正規分布のもとで最適性を持つ通常の検定の標本分布や最適性が、正規分布を一般化した elliptically contoured distribution (楕円等高面分布) に関しても保存されるか、という問題である (Kariya (1981 b, c), Kariya and Sinha (1985) 等)。このうち正規分布以外でも帰

無仮説のもとでの標本分布が変わらない場合を null robustness, さらに正規分布以外でも対立仮説に対する最適性が失われない場合を nonnull robustness あるいは optimality robustness とよんでいる。刈屋は elliptically contoured distribution に関しては非常に多くの場合に null robustness が、またいくつかのケースでは optimality robustness が成立することを示した。検定のロバストネスについては Kariya and Sinha (1988) にまとまった解説が与えられている。

不変検定の手法の応用として興味深いのは、正規分布のもとでの外れ値に関する Ferguson (1960) の結果である。Ferguson は $X_i, i=1, \dots, n$ が互いに独立に正規分布 $N(\mu_i, \sigma_i^2)$ に従う時に slippage 型の検定問題、すなわち対立仮説のもとでいくつかの μ_i あるいは σ_i^2 が他のものと異なるとする検定問題、において標本歪度と尖度が局所最強力不変検定となることを示した。この Ferguson の結果は多変量正規分布の場合 (Schwager and Margolin (1982)) 及び球面対称分布の場合 (Das and Sinha (1986)) に一般化され、Mardia (1970) の多変量尖度に基づく検定が局所最強力不変検定であることが示された。ただし多変量の場合の結果は対立仮説の選び方に依存したものだと思われる。Ferguson の結果は外れ値の検定に関してであるが、これは正規分布の適合度検定の問題とも考えることができるので 3.6 節で再びふれる。

3.4 順序制約下での検定

いわゆる順序制約下での検定問題についてもかなりの文献が見られる。順序制約下の検定問題とは、例えば k 標本の平均の等値性検定 (帰無仮説 $H_0: \mu_1 = \dots = \mu_k$) で、対立仮説のもとでは平均の大きさの順序が決まっている (対立仮説 $H_1: \mu_1 \leq \dots \leq \mu_k$) 形の検定問題である。対立仮説として一方向の順序のみ ($H_1: \mu_1 \leq \dots \leq \mu_k$) を考える場合を片側検定問題、両方向の順序 ($H_1: \mu_1 \leq \dots \leq \mu_k$ or $\mu_1 \geq \dots \geq \mu_k$) を考える場合を両側検定問題という。順序制約下の検定問題は対立仮説が複数の線形不等式によって表わされるより一般の凸錐となる問題に一般化される (Kudo (1963))。すなわち対立仮説として母数ベクトルがある凸錐 C に属する問題を片側検定問題、 $C \cup (-C)$ に属する検定問題を両側検定問題という。順序制約のもとでは最尤推定量や尤度比検定が複雑となるために、通常の検定問題とは異なった理論が必要とされる。この問題に関しては日本から九州大学のグループ及び東京大学のグループの貢献が顕著なので、以下では主にこれらの貢献について紹介する。

順序制約下での統計的推測は Barlow, Bartholomew, Bremner and Brunk (1972) の教科書で分野が確立したが、その後の発展は Robertson, Wright and Dykstra (1988) の教科書にまとめられている。Robertson らの貢献についてはこの教科書から知ることができる。また広津 (1976) の第 6 章及び広津 (1992) の第 3 章にも簡単な解説がある。

片側検定問題に関しては Birnbaum (1955) 以来の議論により検定方式の最小完全類を求めることができる (Eaton (1970), Hirotsu (1982), 竹内 (1979), Marden (1982 b) 参照)。広津 (1982) に従ってこの結果を多変量正規分布 $N(\mu, \Sigma)$ の平均ベクトル μ の検定について述べれば次のようになる。帰無仮説及び対立仮説をそれぞれ

$$H_0: A'\mu = 0, \quad H_1: A'\mu \geq 0$$

とする。ただし $A'\mu \geq 0$ は $A'\mu$ の各要素が非負であることを表す。 $x \sim N(\mu, \Sigma)$ とする時、許容的な検定関数は、その受容域が凸で、かつ $(A'A)^{-1}A'\Sigma^{-1}(x - E_0(x|A^*x))$ の各要素について単調増加である。ただし E_0 は帰無仮説のもとでの期待値であり、 A^* は (A, A^*) が正則かつ $A'\Sigma A^* = 0$ となるように選ぶ。 $\Sigma = I$ の場合について述べれば、 C^- を対立仮説をなす凸錐 C の双対錐とする時、許容的な検定関数の受容域は凸でありかつ C^- の内部に向かう任意の方向に単調である。順序制約下での検定については尤度比検定をはじめいろいろな検定方式が提案さ

れているが、以上の結果によりこれらの検定の多くは許容的であることがわかる。ただし、帰無仮説についても順序制約を課した場合には尤度比検定が非許容的となる場合のあることが知られている (Warrack and Robertson (1984), Nomakuchi and Sakata (1987), Menéndez and Salvador (1991))。

順序制約下での検定方式としてよく用いられるものには尤度比検定, Abelson-Tukey 検定, 田口玄一氏による累積法 (田口玄一 (1966) 参照) およびその精密化・一般化である累積カイ二乗検定などがある。ただし累積カイ二乗検定は両側検定問題に適用されるものである。Abelson-Tukey 検定は C の中心方向に向かう対立仮説に対して高い検出力を持つ検定である。また累積カイ二乗検定は C の端辺方向に対するカイ二乗統計量を各辺について累積したものであり、さまざまな形の対立仮説についてある程度の検出力を持つように考えられた検定である。

広津, 竹内は累積カイ二乗検定が実用的でありまた検出力の観点からも良好であることに注目して, 累積カイ二乗検定の考え方を様々な形に展開している (Hirotzu (1978, 1979, 1982, 1986), Takeuchi and Hirotzu (1982) 等参照)。これにより累積カイ二乗検定の有用性が理論統計家の間でもよく知られるところとなった。Hirotzu (1986) は累積カイ二乗検定統計量の標本分布が自由度 1 のカイ二乗統計量の加重和となることを示し, 累積カイ二乗法の意味づけを明確にした。また漸近分布の形が適合度検定の一つである Anderson-Darling 検定 (3.6 節参照) の場合と同等であるという興味深い結果も示した。

工藤, 坂田, 笹渕, 野間口らは特に順序制約下の尤度比検定に関して多くの結果を出している。これらの結果の一部は野間口 (1992) でサーベイされている。順序制約下での平均ベクトルの最尤推定量は, 順序制約を満たさない推定値を順に加重平均していく PAVA (pool adjacent violator algorithm) とよばれるアルゴリズムで簡単に求められる。Sasabuchi (1980) では帰無仮説が凸錐 C の境界, 対立仮説が凸錐 C の内部となる問題について尤度比検定を導いている。Sasabuchi et al. (1983) では検定問題を複数の多変量正規分布の平均ベクトル間の順序制約に拡張した場合の PAVA による最尤推定を論じている。Nomakuchi and Shi (1988) では同じ問題について Abelson-Tukey 型の統計量を提案している。

与えられたデータに対して PAVA により最尤推定量が簡単に計算できることから, 尤度比検定統計量の値を計算することも容易である。しかしながら PAVA は場合分けを含み, 最尤推定量及び尤度比検定の標本分布は複雑である。これが実用上の尤度比検定の欠点となっている。尤度比検定統計量 $\bar{\chi}^2$ の帰無仮説のもとでの標本分布はカイ二乗分布の混合分布となることが知られており, カイ二乗分布とよばれている。すなわち G_j を自由度 j のカイ二乗分布の累積分布関数とする時, 適当な重み $\omega_i (\omega_i \geq 0, \sum \omega_i = 1)$ を用いて帰無仮説のもとで

$$P(\bar{\chi}^2 \leq c) = \sum_{i=0}^n \omega_i G_i(c)$$

と表わされる。 ω_i は最尤推定量が凸錐 C の i 次元の境界に落ちる確率にあたる。野間口 (1992) で論じられているようにこの重み ω_i の計算が満足な形で解決されていないことが尤度比検定を用いる際の障害となっている。ちなみに ω_i について $\sum_{i=0}^n (-1)^i \omega_i = 0$ という等式が成り立つ。これは一時 Shapiro (1987) の予想として話題になったが, 多面体論の分野ですでに証明されていた事実であった。

3.5 検定の漸近理論

ここでは検定論における漸近理論, 特に尤度比検定などの漸近的にカイ二乗分布に従う検定統計量の漸近理論について述べる。帰無仮説のもとで漸近的にカイ二乗分布に従う検定として

は、尤度比検定のほかにもワルド検定やスコア検定がよく用いられる。この中で尤度比検定の標本分布の帰無仮説のもとでの漸近展開が注目を集めてきた。他の検定と比較して尤度比検定の分布の漸近展開が多くの場合に非常に簡明となることが知られていた。この事実は尤度比検定統計量のバートレット補正という形で定式化され一般的な形で示された。ここではまずこのバートレット補正について述べ、その後他の検定との比較について述べる。検定の漸近理論の分野ではこの他に逐次検定に関する文献や Bahadur efficiency に関する文献も見られるがここでは省略する。

いま $p+q$ 次元の母数ベクトル θ が $\theta=(\theta_1, \theta_2)$ の形の部分ベクトルに分割されたとする。ただし θ_1 は p 次元、 θ_2 は q 次元とする。ここで θ_1 が関心のあるパラメータであり帰無仮説が $H_0: \theta_1=\theta_{10}$ の形の複合帰無仮説であるとする。 θ_2 は局外母数である。いまサンプルサイズ n の標本に基づく尤度を $L(\theta)$ とし、対数尤度を $l(\theta)=\log L(\theta)$ と表す。対数尤度比検定統計量の2倍を Λ とおくと、

$$\Lambda=2(l(\hat{\theta})-l(\bar{\theta}))$$

と表わされる。ただし $\hat{\theta}$ は対立仮説のもとでの無条件の最尤推定量であり、 $\bar{\theta}=(\theta_{10}, \bar{\theta}_2)$ は帰無仮説のもとでの最尤推定量である。帰無仮説のもとで Λ の極限分布は自由度 p のカイ二乗分布であるが、ここでは Λ の分布の漸近展開を考える。いま Λ の期待値を漸近展開すると $E(\Lambda)=p(1+b/n)+o(n^{-1})$ の形に漸近展開できる。ここで期待値を補正して

$$\tilde{\Lambda}=\frac{\Lambda}{1+b/n} \quad (3.1)$$

とおく。この補正をバートレット補正という。バートレット補正をおこなえば $E(\tilde{\Lambda})=p+o(n^{-1})$ となるから、期待値に関する限りカイ二乗分布の近似がよくなっている。ところがこの補正により（期待値のみならず）分布関数に関しても近似が改良されるのである。すなわち G_p を自由度 p のカイ二乗分布の累積分布関数とする時

$$P(\tilde{\Lambda} \leq c) = G_p(c) + o(n^{-1}) \quad (3.2)$$

となることが示される。 $o(n^{-1})$ は実際は $O(n^{-2})$ である (Barndorff-Nielsen and Hall (1988))。このことを、尤度比検定はバートレット補正可能である、という。複合帰無仮説の場合には (3.1) 式で b の値が局外母数の真の値に依存し $b(\theta_2)$ となる場合もあるが、その場合には \sqrt{n} のオーダーの一致推定量を代入して $b(\bar{\theta}_2)$ を用いればよい。バートレット補正後の誤差は n^{-2} のオーダーとなるからカイ二乗近似が良好であることが期待される。実際多くの検定については数値計算によりバートレット補正が有効であることが示されている。しかしながら、多項分布の適合度検定のような離散分布の場合にはバートレット補正が必ずしも有効ではないことを Frydenberg and Jensen (1989) は指摘している。尤度比検定以外の検定では以上のような期待値の補正をおこなってもほとんどの場合分布関数の補正にはならない。

尤度比検定のバートレット補正可能性を一般的な形で扱ったのは Lawley (1956) である。しかしながら、Lawley の計算が非常に複雑であったことと尤度比検定の漸近展開に関する結果の蓄積が少なかったことから、Lawley の結果にもかかわらず当時はバートレット補正が一般に可能であるとは認識されなかった。その後 Hayakawa (1977) が再び膨大な計算を行い、単純帰無仮説 ($q=0$) 及び指数型分布族の自然母数に関する検定についてはバートレット補正が可能であることを示したが、一般の複合仮説の場合の証明にはいたらなかった。その後 Hayakawa (1987) は Hayakawa (1977) の結果を再検討することにより複合帰無仮説の場合

にもパートレット補正が可能であることを確認した。また Cordeiro (1987) も複合帰無仮説の場合の証明を与えた (Harris (1986) も参照)。その後 Bickel and Ghosh (1990) がベイズ法によるパートレット補正可能性の証明を与えるなど他の証明も知られるようになった。分布関数を反転して、パートレット補正を Cornish-Fisher 型の確率展開の形で表せば Y を自由度 p のカイ二乗分布に従う確率変数として

$$\Lambda = Y(1 + b/n) + o(n^{-1}) \quad (3.3)$$

となる。ただし $\overset{d}{=}$ は両辺の分布が等しいことを示す。Takeuchi and Takemura (1988) は 1 母数指数型分布族の場合について、(3.3) 式の右辺を $1/n$ のべきに展開した時の一般項を明示的に評価することにより B_i を多項式とし $\Lambda = Y(1 + B_1(Y)/n + B_2(Y)/n^2 + \dots)$ と Cornish-Fisher 展開した時 B_i が $(i-1)$ 次の多項式となることを示した。 B_1 の次数が 0 すなわち定数であることがパートレット補正である。

尤度比検定のパートレット補正がなぜ可能であるかは自明なことではない。このためパートレット補正の意味についていくつかの論文で論じられている。Barndorff-Nielsen and Cox (1984) では Barndorff-Nielsen and Cox (1979) で得られた漸近補助統計量を与えた時の最尤推定量の密度関数 ((2.1) 式) の標準化定数 c とパートレット補正係数 $(1 + b/n)^{-1}$ の間に簡単な関係式が成り立つことを主張している。McCullagh and Cox (1986) では尤度比検定及びそのパートレット補正が母数の変換に関して不変であることから、パートレット補正係数を母数の変換に関して不変な形で表現し部分的に幾何的な解釈を与えている。

補助統計量に基づく条件つき推測の基準は、漸近的な枠組では有意な結果を導くために有用な基準であると期待されており、この観点から漸近的補助統計量の構成の問題及び漸近的補助統計量を与えた時の条件つき分布に関する論文が数多く見られる。これらについては Cox (1988) が包括的な説明をしている。Cox and Reid (1987) は関心のある母数が 1 次元 ($p=1$) の場合に通常の尤度比のかわりに条件つき尤度比を用いることを提案し、条件つき尤度比の(近似的な)形を導いた。Mukerjee and Chandra (1991) は $q=1$ の場合に Cox and Reid の条件つき尤度比のパートレット補正を与えている。

尤度比検定のパートレット補正に関連して興味深いトピックは標準化された符号つき対数尤度比 (standardized signed log likelihood ratio) である。関心のある母数が 1 次元である時

$$r = \text{sgn}(\hat{\theta}_1 - \theta_{10})\sqrt{\Lambda}$$

を符号つき対数尤度比という。 r は漸近的に標準正規分布に従う。 Λ はパートレット補正によって $o(n^{-1})$ のオーダーまで自由度 1 のカイ二乗分布に一致することから、 r も平均と標準偏差 (の漸近展開) で標準化すれば標準正規分布の近似が改善できると期待される。 n^{-1} のオーダーまでを考えれば $1 + b/n = E(\Lambda) = E(r^2) = \text{Var}(r) + (E(r))^2$ となるから $r' = (r - E(r))/\sqrt{1 + b/n - (E(r))^2}$ が標準化された符号つき対数尤度比である。 $E(r)$ は $n^{-1/2}$ のオーダーである。Barndorff-Nielsen (1986, 1991) は r' の形の統計量について正規近似が $o(n^{-1})$ であること、すなわち Φ を標準正規分布の分布関数として $P(r' \leq c) = \Phi(c) + o(n^{-1})$ であることを主張している。Barndorff-Nielsen の導出は漸近補助統計量を用いたわかりにくいものであり、通常の漸近展開の手法による検証が必要であると思われる。1 母数指数型分布族の場合については Nishii and Yanagimoto (1991) がこの主張を確認している。

尤度比検定の他に漸近カイ二乗検定統計量としてよく用いられるものにはワルドの統計量 W やラオのスコア検定 R がある。 I_{11} をフィッシャー情報行列 I の (1,1) ブロック、 I^{11} を I^{-1} の (1,1) ブロックとして、これらの統計量は

$$W = n(\hat{\theta}_1 - \theta_{10})[I^{11}(\bar{\theta})]^{-1}(\hat{\theta}_1 - \theta_{10}), \quad R = \frac{1}{n} \frac{\partial l}{\partial \theta_1}(\bar{\theta})[I_{11}(\bar{\theta})]^{-1} \frac{\partial l}{\partial \theta_1}(\bar{\theta})$$

で定義される。ただしこれらは若干変形されて用いられることも多い。ワルド統計量やラオのスコア検定は多変量解析の分野や計量経済学の分野の数多くの検定問題について提案され、これらの統計量の漸近展開の結果も数多く発表されている。これらの検定を比較するには、対立仮説として contiguous alternative すなわち $H_1: \theta_1 = \theta_{10} + t/\sqrt{n}$ の形の対立仮説を考え H_1 での検出力を比較する必要がある。その際に帰無仮説のもとでのサイズを必要なオーダーまでそろえておく必要がある。より具体的には帰無仮説のもとでの各検定のサイズを $\alpha + o(n^{-1})$ の形にそろえ、その上で H_1 のもとでの検出力の $1/n$ のオーダーまでの項を比較することになる。これを3次の検出力比較という。Kumon and Amari (1983) 及び Amari (1985) の6章では曲指数族の枠組みで1母数の場合について、幾何学的手法を用いて3次の検出力の比較をおこなった。その結果は以上の検定の中で一様に他よりよい検定はなく、検出力の優劣は t の値に依存するというものである。Eguchi (1991) も幾何学的観点からの検定の比較を扱っている。一般の分布族の場合の $1/\sqrt{n}$ のオーダーまでの検出力の比較は Peers (1971), Hayakawa (1975), Harris and Peers (1980) で扱われている。また Mukerjee, Chandra, Joshi らは一連の仕事 (Chandra and Joshi (1983), Mukerjee and Chandra (1987), Mukerjee (1989, 1990a, 1990b) 等) の中で検出力の3次の比較をおこなっている。そして t が小さい時には一般にスコア検定の検出力がよいことを示している。このことは Amari (1985) の枠組みでも確認されていることである。

3.6 適合度検定

すでに述べたように、分布形の適合度検定に関しては最近でも多くの文献が見られる。例えば、観測値が正規分布に従うとする正規性の検定については非常に多くの検定方法が提案されている。これは、伝統的な検定論の多くが正規分布の仮定のもとに構成されていることへの問題意識の現れであるとも言える。分布形の検定の場合には、対立仮説の次元が無次元であることもあり、対立仮説を特定することが難しい。このために1母数の場合のよう最強力検定にあたるものは存在しない。このような事情もあり正規性の検定について数多くの提案がなされている。以下では、主に1変量正規性の検定を念頭において適合度検定の研究を概観する。多変量正規性の検定についても多くの文献が見られるがスペースの関係で以下では省略する。

正規分布に続いて適合度検定の対象となるのが、信頼性理論や生存時間解析における指数分布の仮定である。指数分布は多くの都合のよい性質をもっているため、実際の分析において指数分布を仮定できるかどうかの問題となる。生存時間解析においては、生存時間の分布がそのまま観測されず、観測値が打ち切られる(センサリング)ことが多い。センサリングの問題は生存時間解析の一つの主要な課題である。このためセンサリングの仮定のもとでの適合度検定も *Biometrika*, *Biometrics* 等の雑誌の多くの論文で議論されている。これについても以下では省略する。

適合度検定に関してまとまった書物としては D'Agostino and Stephens (1986) があげられる。この本は論文集の形態をとっているものの適合度検定に関する教科書と言ってもよく、理論応用の両面の観点から適合度検定についての結果をまとめている。また、適合度検定に用いられる経験分布関数の理論は Shorack and Wellner (1986) の大部の教科書において詳細に扱われている。

適合度検定として最初に考えられるものは、分布を区間に区切り多項分布の確率の検定に帰

着させるピアソン型のカイ二乗適合度検定であろう。これに関する漸近理論は Moore and Spruill (1975) に整理されているが、それ以降の進展は大きなものではない。特に、検出力を考慮した時に、区間や区間数をどのようにとったらよいかという簡単かつ重要な問題に対する明確な答えはあいかわらず得られていないように思われる (Quine and Robinson (1985))。

最も盛んに研究されているのは、Kolmogorov-Smirnov 検定、Cramér-Von Mises 検定、Anderson-Darling 検定等の経験分布関数 F_n に基づく検定である。経験分布関数に基づく検定は帰無仮説が単純仮説、すなわち帰無仮説のもとでの分布が特定された場合には、累積分布関数を考えることによって一様分布の場合に帰着できる。従ってこの場合の検定統計量の標本分布は帰無仮説の分布に依存しない (distribution free) なものとなり、有意点なども標準的な形で与えることができる。しかしながら、分布形の検定においては当然のことながら位置母数及び尺度母数は未知のことが多い。従って分布形の検定としては分布族の未知母数を推定するか、あるいは 3.3 節におけるように不変性により未知母数に依存しない検定を考える必要がある。経験分布関数に基づく検定においては、未知母数を推定すると検定統計量の標本分布が分布形に依存してしまい、漸近理論で考えた場合でも distribution free とならないことが問題となる。これは Durbin (1973a, b) によって示されたように、経験分布関数の漸近共分散関数が分布形及び未知母数に依存してしまうためである。従って未知母数を含む場合には経験分布関数に基づく検定の (漸近的) 有意点は、分布形及び未知母数の組み合わせごとに与える必要がある。正規分布においては位置 μ 及び尺度 σ^2 を未知母数と考えればよいし、指数分布においては尺度母数を未知と考えればよいであろう。これらの場合の Cramér-von Mises 型の検定の漸近的有意点は Stephens (1976) に与えられている。回帰分析においては残差の正規性の検定が問題となる。この問題に対して Pierce ら (Pierce and Kopecky (1979), Pierce and Gray (1982), Pierce (1985) 等) は次の有用な結果を示した。すなわち、定数項を含む回帰分析及び自己回帰過程においては、残差の経験分布関数の帰無仮説のもとでの漸近的性質は独立同一分布の場合と同等である。従って回帰分析の残差に関する正規性の検定は、位置及び尺度母数が未知の場合の独立同一分布の場合の正規性の検定と同様に行うことができ、有意点も同じものを用いることができる。

分布形の検定においては対立仮説の次元は無次元であるが、あえて対立仮説の分布族を特定化し Neyman-Pearson の補題と同様の考え方をを用いて特定の対立仮説に対して最適な検定を求めるのも一つの行き方である。すでに 3.3 節でふれたように Ferguson (1960) は標本歪度と尖度が外れ値型の対立仮説に対して局所最強力 (位置尺度) 不変であることを示した。Spiegelhalter (1977) は一様分布及び両側指数分布を対立仮説として局所最強力不変検定を導いた。また Spiegelhalter (1983) では標本歪度が漸近的には t 分布族に対する局所検定となることを示しこの意味で標本歪度の最良性を示した。Kuwana and Kariya (1991) は密度関数が $\exp(-|x|^q/2)$ の形のベキ指数分布において、正規性の検定が仮説 $H:\theta=2$ の検定となることに注目して、多変量正規性に関する局所最強力不変検定を与えた。

適合度検定で一つの興味のある問題は分布の対称性である。例えばロバスト推定では左右対称な分布の中心を推定する問題を考えることが多い。この場合分布の対称性を検定する意味がある。原点を中心とする左右対称で連続な累積分布関数は $F(x)+F(-x)=1$ という等式がなりたつ。従って経験分布関数を用いてこの関係をチェックすることが考えられる。対称の中心が 0 であることが既知の場合は分布形によらない検定が得られる。Aki (1981) は Cramér-von Mises 型の統計量について、分布の中心が未知の場合を論じ検定統計量の標本分布を評価した。この場合の標本分布は分布形に依存するものとなる。さらに Aki (1987), Nabeya (1987) はより一般化された対称性の検定を与えている。

経験分布関数と並んで適合度検定に頻繁に用いられるのが順序統計量である。順序統計量は経験分布関数 F_n の逆関数、すなわち分位点関数 (quantile function) の値である。従って順序統計量に基づく検定は経験分位点関数 F_n^{-1} に基づく検定である。経験分布関数に比して、順序統計量を用いることにより裾の重い分布に対する検出力の向上が期待できる。経験分位点関数に基づく検定としては Shapiro-Wilk 検定及びその変形がよく知られている (Shapiro and Wilk (1965), Shapiro and Francia (1972), de Wet and Venter (1972))。これは Shapiro 及び Wilk 自身のモンテカルロ研究の結果により Shapiro-Wilk 検定の検出力が良好であると理解されたことにもよる (ただし Pearson, D'Agostino and Bowman (1977) によれば Shapiro-Wilk 検定の検出力は標本歪度や尖度に基づく検定に比して高くはない)。Shapiro-Wilk 検定統計量を計算するためには正規分布の順序統計量の期待値及び共分散行列が必要となる。順序統計量の期待値はまだしも、共分散行列を評価するのは困難である。このため、Shapiro-Wilk 検定を変形したいくつかの検定が提案されることとなった。これらの変形された統計量は漸近分布の取り扱いも容易であるが、もともとの Shapiro-Wilk 検定の漸近分布の厳密な導出は Leslie, Stephens and Fotopoulos (1986) によって与えられた。経験分位点関数の漸近理論と Shapiro-Wilk 型の統計量への応用は LaRicca and Mason (1986) で扱われている。また Verrill and Johnson (1987) ではセンシングを含む場合について Shapiro-Wilk 型の統計量の漸近分布を与えている。

適合度検定におけるもう一つの興味深いアプローチは、経験分布関数の特性関数、すなわち経験特性関数を用いて検定を行うことである。正規分布は簡単な特性関数を有するので、経験特性関数を正規分布の特性関数と比較することによって正規性の検定を行うことができる。変数 t を固定すれば経験特性関数 $\phi_n(t) = (1/n) \sum_{j=1}^n \exp(itX_j)$ には中心極限定理を応用することができるので、これを用いて有意点を計算することができる。経験特性関数に基づく検定は単純帰無仮説、すなわち帰無仮説の分布が特定された場合について Feigin and Heathcote (1976) で扱われたが、Murota and Takeuchi (1981) において確率変数を標準化することにより、位置母数及び尺度母数に依存しない分布形の検定として具体的に検討され、正規性の検定として有効であることが示された。Koutrouvelis and Kellermeier (1981), Epps and Pulley (1983) も同様のアプローチである。Murota and Takeuchi の結果は Csörgö (1986) によって数学的にも精緻な形で多変量の場合に拡張された。

謝辞：査読者の有益なコメントに感謝します。

参 考 文 献

- [1] Aki, S. (1981). Asymptotic distribution of a Cramér-von Mises type statistic for testing symmetry when the center is estimated. *Ann. Inst. Stat. Math.*, **33**, 1-14.
- [2] Aki, S. (1987). On nonparametric tests for symmetry. *Ann. Inst. Stat. Math.*, **39**, 457-472.
- [3] Amari, S. (1985). 第2章参照
- [4] Andersson, S. (1982). Distribution of maximal invariants using quotient measures. *Ann. Statist.*, **10**, 955-961.
- [5] Barlow, R.E., Bartholomew, J. M., Bremner, D. J., and Brunk, H. D. (1972). *Statistical Inference under Order Restrictions*. Wiley.
- [6] Barndorff-Nielsen, O. E. and Cox, D. R. (1979). 第2章参照
- [7] Barndorff-Nielsen, O. E. and Cox, D. R. (1984). 第2章参照
- [8] Barndorff-Nielsen, O. E. (1986). 第2章参照
- [9] Barndorff-Nielsen, O. E. (1991). 第2章参照
- [10] Barndorff-Nielsen, O. E. and Hall, P. (1988). On the level-error after Bartlett adjustment of the

- likelihood ratio statistic. *Biometrika*, **75**, 374-378.
- [11] Bickel, P. J. and Ghosh, J. K. (1990). A decomposition for the likelihood ratio statistic and the Bartlett correction—A Bayesian argument. *Ann. Statist.*, **18**, 1070-1090.
- [12] Birnbaum, A. (1955). Characterization of complete classes of tests of some multiparametric hypotheses, with application to likelihood ratio tests. *Ann. Math. Statist.*, **26**, 21-36.
- [13] Brown, L. D. and Marden, J. I. (1989). Complete class results for hypothesis testing problems with simple null hypotheses. *Ann. Statist.*, **17**, 209-235.
- [14] Chandra, T. K. and Joshi, S. N. (1983). Comparison of the likelihood ratio, Rao's and Wald's tests and a conjecture of C. R. Rao. *Sankhya*, Series A, **45**, 226-246.
- [15] Cohen, A. and Marden, J. I. (1989). On the admissibility and consistency of tests for homogeneity of variances. *Ann. Statist.*, **17**, 236-251.
- [16] Cohen, A. and Sackrowith, H. B. (1987). Unbiasedness of tests for homogeneity. *Ann. Statist.*, **15**, 805-816.
- [17] Cordeiro, G. M. (1987). On the correction to the likelihood ratio statistics. *Biometrika*, **74**, 265-274.
- [18] Cox, D. R. and Reid, N. (1987). 第2章参照
- [19] Cox, D. R. (1988). Some aspects of conditional and asymptotic inference: A review. *Sankhya*, Series A, **50**, 314-337.
- [20] Csörgö, S. (1986). Testing for normality in arbitrary dimension. *Ann. Statist.*, **14**, 708-723.
- [21] D'Agostino, R. B. and Stephens, M. A. (editors) (1986). *Goodness-of-fit Techniques*. Marcel Dekker.
- [22] Das, R. and Sinha, B. K. (1986). Detection of multivariate outliers with dispersion slippage in elliptically symmetric distributions. *Ann. Statist.*, **14**, 1619-1624.
- [23] de Wet, T. and Venter, J. H. (1972). Asymptotic distribution of certain test criteria for normality. *South African Statist. J.*, **6**, 135-149.
- [24] Durbin, J. (1973a). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.*, **1**, 279-290.
- [25] Durbin, J. (1973b). *Distribution Theory for Tests Based on the Sample Distribution Function*. SIAM, Philadelphia.
- [26] Eaton, M. L. (1970). A complete class theorem for multidimensional one-sided alternatives. *Ann. Math. Statist.*, **41**, 1884-1888.
- [27] Eaton, M. L. (1989). *Group Invariance Applications in Statistics*. Regional Conference Series in Probability and Statistics, vol. 1, IMS-AMS.
- [28] Eaton, M. L. and Kariya, T. (1983). Multivariate tests with incomplete data. *Ann. Statist.*, **11**, 654-665.
- [29] Eguchi, S. (1991). A geometric look at nuisance parameter effect of local powers in testing hypothesis. *Ann. Inst. Stat. Math.*, **43**, 245-260.
- [30] Epps, T. W. and Pulley, L. B. (1983). A test for normality based on the empirical characteristic function. *Biometrika*, **70**, 723-726.
- [31] Feigin, P. D. and Heathcote, C. R. (1976). The empirical characteristic function and the Cramér-Von Mises statistic, *Sankhya*, **38**, 309-325.
- [32] Ferguson, T. S. (1960). On the rejection of outliers. In *Proc. 4th Berkeley Symp.* **1**, 253-287.
- [33] Frydenberg, M. and Jensen, J. L. (1989). Is the 'improved likelihood ratio statistic' really improved in the discrete case? *Biometrika*, **76**, 655-661.
- [34] Harris, P. (1986). A note on Bartlett adjustments to likelihood ratio tests. *Biometrika*, **73**, 735-737.
- [35] Harris, P. and Peers, H. W. (1980). The local power of the efficient scores test statistic. *Biometrika*, **67**, 525-529.
- [36] Hayakawa, T. (1975). The likelihood ratio criterion for a composite hypothesis under a local alternative. *Biometrika*, **62**, 451-460.
- [37] Hayakawa, T. (1977). The likelihood ratio criterion and the asymptotic expansion of its distribution. *Ann. Inst. Statist. Math.*, **29**, 359-378.
- [38] Hayakawa, T. (1987). Correction to "The likelihood ratio criterion and the asymptotic expansion

- of its distribution". *Ann. Inst. Statist. Math.*, **39**, 681.
- [39] 広津千尋 (1976). 分散分析. 教育出版
- [40] Hirotsu, C. (1978). Ordered alternatives for interaction effects. *Biometrika*, **65**, 561-570.
- [41] Hirotsu, C. (1979). The cumulative chi-squares method and a studentized maximal contrast method for testing an ordered alternative in a one-way analysis of variance model. *Reports of Statistical Application Research, JUSE*, **26**, 12-21.
- [42] Hirotsu, C. (1982). Use of cumulative efficient scores for testing ordered alternatives in discrete models. *Biometrika*, **69**, 567-578.
- [43] Hirotsu, C. (1986). Cumulative chi-squared statistic as a tool for testing goodness of fit. *Biometrika*, **73**, 165-173.
- [44] 広津千尋 (1992). 実験データの解析—分散分析を超えて—. 共立出版
- [45] Kariya, T. (1978). The general MANOVA problem. *Ann. Statist.*, **6**, 200-214.
- [46] Kariya, T. (1981a). Tests for the independence between two seemingly unrelated regression equations. *Ann. Statist.*, **9**, 381-390.
- [47] Kariya, T. (1981b). A robustness property of Hotelling's T^2 -test. *Ann. Statist.*, **9**, 211-214.
- [48] Kariya, T. (1981c). Robustness of multivariate tests. *Ann. Statist.*, **9**, 1267-1275.
- [49] Kariya, T. (1985). *Testing in the Multivariate General Linear Model*. Kinokuniya, Tokyo.
- [50] Kariya, T. and Sinha, B. K. (1985). Nonnull and optimality robustness of some tests. *Ann. Statist.*, **13**, 1182-1197.
- [51] Kariya, T. and Sinha, B. K. (1988). *The Robustness of Statistical Tests*. Academic Press, New York.
- [52] Koutrouvelis, I. A. and Kellermeier, J. (1981). A goodness-of-fit test based on the empirical characteristic function when parameters must be estimated. *J. R. Statist. Soc. B*, **43**, 173-176.
- [53] Kudo, A. (1963). A multivariate analogue of the one-sided test. *Biometrika*, **50**, 403-418.
- [54] Kumon, M. and Amari, S. (1983). 第2章参照.
- [55] Kuwana, Y. and Kariya, T. (1991). LIB tests for multivariate normality in exponential power distributions. *J. Multivariate Analysis*, **39**, 117-134.
- [56] LaRiccia, V. and Mason, D. M. (1986). Cramér-von Mises statistics based on the sample quantile function and estimated parameters. *J. Multivariate Analysis*, **18**, 93-106.
- [57] Lawley, D. N. (1956). A general method of approximating to the distribution of likelihood ratio criteria. *Biometrika*, **43**, 295-303.
- [58] Lehmann, E. L. (1959). *Testing Statistical Hypotheses*. 1st. ed. Wiley.
- [59] Lehmann, E. L. (1986). *Testing Statistical Hypotheses*. 2nd ed. Wiley.
- [60] Leslie, J. R., Stephens, M. A., and Fotopoulos, S. (1986). Asymptotic distribution of the Shapiro-Wilk W for testing for normality. *Ann. Statist.*, **14**, 1497-1506.
- [61] Marden, J. I. (1982a). Combining independent noncentral Chi squared or F tests. *Ann. Statist.*, **10**, 266-277.
- [62] Marden, J. I. (1982b). Minimal complete classes of tests of hypotheses with multivariate one-sided alternatives. *Ann. Statist.*, **10**, 962-970.
- [63] Marden, J. I. and Perlmen, M. D. (1990). On the inadmissibility of step-down procedures for the Hotelling T^2 problem. *Ann. Statist.*, **18**, 172-190.
- [64] Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, **57**, 519-530.
- [65] McCullagh, P. and Cox, D. R. (1986). Invariants and likelihood ratio statistic. *Ann. Statist.*, **14**, 1419-1430.
- [66] Menéndez, J. A. and Salvador, B. (1991). Anomalies of the likelihood ratio test for testing restricted hypotheses. *Ann. Statist.*, **19**, 889-898.
- [67] Moore, D. S. and Spruill, M. C. (1975). Unified large-sample theory of general chi-squared statistics for tests of fit. *Ann. Statist.*, **3**, 599-616.
- [68] Mukerjee, R. (1989). Third-order comparison of unbiased tests: a simple formula for the power difference in the one-parameter case. *Sankhya, Series A*, **51**, 212-232.
- [69] Mukerjee, R. (1990a). Comparison of tests in the multiparameter case I. Second-order power. *J. Multivariate Analysis*, **33**, 17-30.

- [70] Mukerjee, R. (1990b). Comparison of tests in the multiparameter case. II. A third-order optimality property of Rao's test. *J. Multivariate Analysis*, **33**, 31-48.
- [71] Mukerjee, R. and Chandra, T. K. (1987). Comparison between the locally most powerful unbiased and Rao's tests. *J. Multivariate Analysis*, **22**, 94-105.
- [72] Mukerjee, R. and Chandra, T. K. (1991). Bartlett-type adjustment for the conditional likelihood ratio statistic of Cox and Reid. *Biometrika*, **78**, 365-372.
- [73] Murota, K. and Takeuchi, K. (1981). The studentized empirical characteristic function and its application to test for the shape of distribution. *Biometrika*, **68**, 55-65.
- [74] Nabeya, S. (1987). On Aki's nonparametric test for symmetry. *Ann. Inst. Stat. Math.*, **39**, 473-482.
- [75] Nishii, R. and Yanagimoto, T. (1991). Asymptotic expansions on the signed log likelihood ratio. To appear in *Proceedings of the Third Pacific Area Statistical Conference*.
- [76] Nomakuchi, K. and Sakata, T. (1987). A note on testing two-dimensional normal mean. *Ann. Inst. Statist. Math.* **39**, 489-495.
- [77] Nomakuchi, K. and Shi, N. Z. (1988). A test for a multiple isotonic regression problem. *Biometrika*, **75**, 181-184.
- [78] 野間口謙太郎(1992). 線形不等式制約条件下での正規分布の平均に関する尤度比検定. 1992年春季日本数学会特別講演.
- [79] Peers, H. W. (1971). Likelihood ratio and associated test criteria. *Biometrika*, **58**, 577-587.
- [80] Pierce, D. A. (1985). Testing normality in autoregressive models. *Biometrika*, **72**, 293-298.
- [81] Pierce, D. A. and Kopecky K. J. (1979). Testing goodness of fit for the distribution of errors in regression models. *Biometrika*, **66**, 1-5.
- [82] Pierce, D. A. and Gray R. J. (1982). Testing normality of errors in regression models. *Biometrika*, **69**, 223-236.
- [83] Pearson, E. S., D'Agostino, R. B. and Bowman, K. O. (1977). Tests for departure from normality: Comparison of powers. *Biometrika*, **64**, 231-246.
- [84] Quine, M. P. and Robison, J. (1985). Efficiencies of chi-square and likelihood ratio goodness-of-fit tests. *Ann. Statist.*, **13**, 727-742.
- [85] Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. Wiley.
- [86] Sasabuchi, S. (1980). A test of a multivariate normal mean with composite hypotheses determined by linear inequalities. *Biometrika*, **67**, 429-439.
- [87] Sasabuchi, S., Inutsuka, M. and Kulatunga, D. D. S. (1983). A multivariate version of isotonic regression. *Biometrika*, **70**, 465-472.
- [88] Schwager, S. J. and Margolin, B. H. (1982). Detection of multivariate normal outliers. *Ann. Statist.*, **10**, 943-954.
- [89] Shapiro, S. S. and Francia, R. S. (1972). An approximate analysis of variance test for normality. *J. Amer. Statist. Assoc.*, **67**, 215-216.
- [90] Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, **52**, 591-611.
- [91] Shapiro, A. (1987). A conjecture related to chi-bar-squared distributions. *Amer. Math. Monthly*, **94**, 46-48.
- [92] Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Application to Statistics*. Wiley.
- [93] Spiegelhalter, D. J. (1977). A test for normality against symmetric alternatives. *Biometrika*, **64**, 415-418.
- [94] Spiegelhalter, D. J. (1983). Diagnostic tests of distributional shape. *Biometrika*, **70**, 401-409.
- [95] Stephens, M. A. (1976). Asymptotic results for goodness-of-fit statistics with unknown parameters. *Ann. Statist.*, **4**, 357-369.
- [96] 田口玄一(1966). 統計解析. 丸善.
- [97] 竹内 啓(1979). 制約条件のある対立仮説, 仮説の下での推定・検定問題. 経済学論集 **45**, 2-10.
- [98] Takeuchi, K. and Hitotsu, C. (1982). The cumulative chi-squares method against ordered alternatives in two-way contingency tables. *Reports of Statistical Application Research, JUSE*, **29**, No. 3,

1-13

- [99] Takeuchi, K and Takemura, A. (1988). On Cornish-Fisher type expansion of likelihood ratio statistic in one parameter exponential family. in *Proceedings of Fifth Japan-USSR Symposium on Probability Theory and Mathematical Statistics*, Springer.
- [100] Verrill, S. and Johnson, R. A. (1987). The asymptotic equivalence of some modified Shapiro-Wilk statistics—complete and censored sample cases. *Ann. Statist.*, **15**, 413-419.
- [101] Warrack, G. and Robertson, T. (1984). A likelihood ratio test regarding two nested but oblique order-restricted hypotheses. *J. Amer. Statist. Assoc.*, **79**, 881-886.
- [102] Wijsman, R. A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. *Proc. Fifth Berkeley Symp. on Math. Statist. and Prob.*, **1**, 389-400. Univ. of California Press.
- [103] Wijsman, R. A. (1990). *Invariant Measures on Groups and Their Use in Statistics*. IMS Lecture Notes-Monograph Series, vol. 14.

4. ブートストラップ法とその応用

4.1 はじめに

1979年にB. Efronによって提唱されたブートストラップ法は、従来、理論や数式に基づく解析的アプローチが難しかった問題に対して、有効な解を与えることができるということで注目を集めてきた。その特徴は、ブートストラップ法の実行プロセスの中で、解析的導出をコンピュータを用いた大量の反復計算で置き換えているところにある。これによって、極めて緩やかな仮定のもとで、より複雑な問題に適用できる柔軟な統計手法となった。

このコンピュータの利用を前提とした統計的数値計算法は、多くの研究者の興味を引き、ここ約15年の間に理論的・実際の両側面に渡って集中的に研究が行われ、数多くの論文が発表されてきた。これらの研究を通して、ブートストラップ法の理論構造が明らかとなり、その有用性と共に問題点も指摘されてきた。本稿は、ブートストラップ法に関して行われてきた様々な分野の研究を、その基本的な考え方と共に紹介することを目的とする。

4.2節では、Efron (1979, 1982) によるブートストラップ法の基本的な実行プロセスと、適用上のいくつかの留意点を中心に述べる。4.3節では回帰モデルへの応用研究を、4.4節では、判別分析に於ける誤判別率推定、情報量規準AICに於ける対数尤度のバイアス補正への応用を紹介する。4.5節では、これまで最も研究が進められてきた、ブートストラップ分布と信頼区間の構成に関する研究を概観する。4.6節では、ブートストラップシミュレーションの誤差に関するいくつかの研究を紹介する。

現在までに発表された膨大な量の研究論文を網羅的に紹介することはできないので、これまでの研究成果を集めた著書 (Beran and Ducharme (1991), Hall (1992 a), Mammen (1992)), 総合報告 (Hinkley (1988), DiCiccio and Romano (1988), 小西 (1990), Swanepoel (1990), Léger *et al.* (1992), 小西・本多 (1992)) 等を併せて参照していただきたい。

4.2 ブートストラップ法

4.2.1 実行のプロセス

未知の確率分布 F をもつ母集団からの大きさ n の無作為標本を、 $\mathbf{x}_n = \{X_1, \dots, X_n\}$ とする。 θ を F に関するあるパラメータとし、これを一つの推定量 $\hat{\theta}_n = \hat{\theta}(X_1, \dots, X_n)$ を用いて推定する。ただし、推定量は標本の並べ替えに対して不変であるとする。観測値 $X_1 = x_1, \dots, X_n = x_n$ が採られたとき、推定値 $\hat{\theta} = \hat{\theta}(x_1, \dots, x_n)$ をもって θ に関する推論を行なうと共に、推定の信頼度

を併せて評価することが統計的分析を行う上で重要となる。推定の誤差を捉える基本的な一次元評価尺度が、次の推定量のバイアスおよび標準誤差である。

$$(4.1) \quad b(F) = E_F[\hat{\theta}_n - \theta], \quad se(F) = \{E_F[(\hat{\theta}_n - E_F[\hat{\theta}_n])^2]\}^{1/2}$$

ここで、期待値は確率分布 F に関してとるものとする。さらに、推定量の標本分布が解れば、確率あるいは信頼度を用いて推定値とパラメータとの誤差をより明確に述べることができるし、パラメータ θ の信頼区間の構成が可能となる。

このような観点から必要となるのは、多くの場合 $T_{SD} = \sqrt{n}(\hat{\theta}_n - \theta)$ の分布およびその 100 α % 点 (パーセントイル)

$$(4.2) \quad H_n(x, F) = P_F\{\sqrt{n}(\hat{\theta}_n - \theta) \leq x\}, \quad x_\alpha^{(SD)} = \inf\{x: H_n(x, F) \geq \alpha\}$$

である。また、 $\sqrt{n}\hat{\theta}_n$ の (漸近) 分散 σ^2 に対して、何らかの推定量 $\hat{\sigma}^2$ が求めれば、スチューデント化された統計量 $T_{ST} = \sqrt{n}(\hat{\theta}_n - \theta)/\hat{\sigma}$ の分布とその 100 α % 点

$$(4.3) \quad K_n(x, F) = P_F\{\sqrt{n}(\hat{\theta}_n - \theta)/\hat{\sigma} \leq x\}, \quad x_\alpha^{(ST)} = \inf\{x: K_n(x, F) \geq \alpha\}$$

を用いることができる。

推定量の統計的誤差を計るバイアス、標準誤差、標本分布は、いずれも確率分布 F に依存する未知の量であり、データに基づいてどのように推定するかが問題となる。ブートストラップ法は、これらの量の推定を解析的に行なう代わりに、計算機の上で数値的に実行するための一つの統計手法で、基本的には次のステップを通して実行される。

(1) 未知の母集団確率分布 F を経験分布関数 \hat{F} で推定する。ここで、 \hat{F} は、 n 個の観測値 $\{x_1, \dots, x_n\}$ の各点で確率 $1/n$ をもつ分布である。

(2) 既知の確率分布 \hat{F} をもつ母集団からの大きさ n の無作為標本 (ブートストラップ標本) を $\mathbf{x}_n^* = \{X_1^*, \dots, X_n^*\}$ とし、 $\hat{\theta}_n^* = \hat{\theta}(X_1^*, \dots, X_n^*)$ とする。このとき、(4.2) 式の推定量の分布 $H_n(x, F)$ および 100 α % 点 $x_\alpha^{(SD)}$ は、標本 \mathbf{x}_n が与えられたもとで各々

$$(4.4) \quad \hat{H}_n(x, \hat{F}) = P_F\{\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}) \leq x\}, \quad \hat{x}_\alpha^{(SD)} = \inf\{x: \hat{H}_n(x, \hat{F}) \geq \alpha\}$$

と推定され、また、(4.3) 式の分布 $K_n(x, F)$ とその 100 α % 点 $x_\alpha^{(ST)}$ は

$$(4.5) \quad \hat{K}_n(x, \hat{F}) = P_F\{\sqrt{n}(\hat{\theta}_n^* - \hat{\theta})/\hat{\sigma}^* \leq x\}, \quad \hat{x}_\alpha^{(ST)} = \inf\{x: \hat{K}_n(x, \hat{F}) \geq \alpha\}$$

と推定される。ただし、 $\hat{\sigma}^*$ は、ブートストラップ標本 \mathbf{x}_n^* に基づく推定量とする。

(3) (4.4), (4.5) 式の値は、経験分布関数 \hat{F} が既知の確率分布であることを利用して、モンテカルロ法によって数値的に近似する。すなわち、 \hat{F} から大きさ n のブートストラップ標本を B 回反復抽出し、各標本に対する推定値 $\hat{\theta}_{(i)}^*$ を計算し、観測データに基づく推定値 $\hat{\theta}$ に対して $\varphi_{SD}(i) = \sqrt{n}(\hat{\theta}_{(i)}^* - \hat{\theta})$ とおく。このとき、(4.4) 式は

$$(4.6) \quad P_F\{\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}) \leq x\} \approx \{B \text{ 個の } \varphi_{SD}(i) \text{ の中で } x \text{ 以下の個数}\} / B, \\ \hat{x}_\alpha^{(SD)} \approx \{\varphi_{SD}(i); i=1, \dots, B\} \text{ の } 100 \alpha \text{ \% 点}$$

と近似される。(4.5) 式のブートストラップ分布とそのパーセントイルに対しても同様である。すなわち、 i 番目のブートストラップ標本に対する推定値と分散の値 $\{\hat{\theta}_{(i)}^*, \hat{\sigma}_{(i)}^{*2}\}$ に対して、 $\varphi_{ST}(i) = \sqrt{n}(\hat{\theta}_{(i)}^* - \hat{\theta})/\hat{\sigma}_{(i)}^*$ と置き換えればよい。

同様にして、(4.1) 式のバイアス、標準誤差のブートストラップ推定値はそれぞれ

$$(4.7) \quad \hat{b}(\hat{F}) = E_{\hat{F}}[\hat{\theta}_n^*] - \hat{\theta} \approx \sum_{i=1}^B \hat{\theta}_{(i)}^* / B - \hat{\theta},$$

$$(4.8) \quad s\hat{e}(\hat{F}) = \{E_{\hat{F}}[(\hat{\theta}_n^* - E_{\hat{F}}[\hat{\theta}_n^*])^2]\}^{1/2} \approx \left\{ \sum_{i=1}^B (\hat{\theta}_{(i)}^* - \hat{\theta}^*)^2 / (B-1) \right\}^{1/2}$$

で与えられる。ただし、 $\hat{\theta}^* = \sum_{i=1}^B \hat{\theta}_{(i)}^* / B$ とする。

以上が、ブートストラップ法の基本的な実行プロセスである。ブートストラップ法の手法紹介と、様々な分野の問題への応用を総合的に述べたのが、Efron (1982), Efron and Gong (1983), Efron and Tibshirani (1986) である。その他、手法を極めて平易に紹介した Diaconis and Efron (1983), 小西 (1988, 8章) がある。また、Efron (1992 a) は、ブートストラップ法の研究を通して生じたいくつかの問題を上げ、Efron 自身のこれまでの研究の流れの中でこれらの問題を検討している。

(モンテカルロアルゴリズム) ブートストラップ法が、複雑な推測問題に対して幅広く適用できるのは、上述のステップ (3) のモンテカルロ法に基づく数値計算法を実行できる点にある。これは、バイアス、分散、確率分布のブートストラップ推定値は、結局、ブートストラップ標本の同時分布 ($\Pi_{a=1}^n \hat{F}$) に関する期待値の計算であり、これを \hat{F} が既知の分布であることを利用してモンテカルロ法による数値近似を行っている。すなわち、経験分布関数からの大きさ n の標本の反復抽出とは、観測データ $\{x_1, \dots, x_n\}$ からの大きさ n の標本の復元抽出と同値であることを利用している。もし、独立、同一分布に従う標本でなければ、このような標本の反復抽出は基本的には実行できないことが分かる。

(反復抽出の回数) モンテカルロ法による数値近似の誤差は、ブートストラップ反復抽出の回数 B を無限大とすると無視できるものである。実際には反復抽出の回数は、バイアスおよび標準誤差の推定に対しては、 $B=50 \sim 200$ 、これに対してパーセンタイルの推定では、 $B=1000 \sim 2000$ は必要であることが Efron (1987, 9節) で報告されている。その他、Hall (1986 a) の区間推定に於ける反復回数の理論研究がある。4.6節では、モンテカルロシミュレーションによる近似誤差を減少させるための各種手法を紹介する。

(スチューデント化された統計量) もし推定量の分散が有効に推定できれば、スチューデント化された統計量の利用は、近似精度を改善するという意味で有用である (4.5節を参照)。ノンパラメトリックモデルでの分散の推定法としては、デルタ法 (Siotani *et al.* (1985, p. 88), Hall (1992 a, p. 76)), ジャックナイフ法 (Quenouille (1949, 1956), Tukey (1958), Miller (1974)), ブートストラップ法等の利用が考えられる。ただし、ブートストラップ分散推定法を用いる場合、二段階ブートストラップ法 (例えば、小西 (1990, p. 149) を参照) を実行する必要がある。

スチューデント化された統計量に基づく方法が有効に働くのは、分散の安定した推定値が得られるときである。分散の安定した推定が難しい標本相関係数のような場合、4.5.3節で検討する有効な変換を行うか、あるいはブートストラップ反復法の適用が考えられる。

(他の手法との関係) 推定量のバイアス、標準誤差のブートストラップ推定と、統計的リサンプリング手法として古くから用いられているジャックナイフ法との関係、さらには影響関数 (例えば、Hampel *et al.* (1986, p. 84)), デルタ法にもとづく推定法、無限小ジャックナイフ法 (Jaeckel (1972)) との関係が、Efron (1982), Efron and Gong (1983), Parr (1983), 高橋 (1985) など明らかにされた。また、ブートストラップリサンプリング法と Hartigan (1969, 1971, 1975) の *Subsample* 法との関係については、Efron (1979, p. 24) で論じられており、そ

の他 Babu (1992) の研究がある。

4.2.2 確率分布の推定

ブートストラップ法の基本的な考え方は、未知の確率分布 F からの標本に基づく推測過程を、既知の経験分布関数 \hat{F} からのブートストラップ標本に基づく設定へと置き換えているところにある。したがって、確率分布 F としてパラメトリックモデルを想定した場合、あるいは他の何らかの方法で未知の確率分布を推定した場合にも、ブートストラップ法は実行できる。

(パラメトリックブートストラップ法) 母集団の変動を表す確率分布として、パラメトリックモデル $f(x|\eta)$ を想定できるものとする。このようなモデル設定では、未知の母集団確率分布は、例えば、分布を規定するパラメータベクトル η をその最尤推定量で置き換えた $f(x|\hat{\eta})$ で推定することができる。このとき、ブートストラップ標本は $f(x|\hat{\eta})$ に従って取り出され、4.2.1 節と同様のプロセスを適用できる。

想定したモデルあるいは推定量の複雑さにもよるが、推定量のブートストラップ分布を、解析的に陽に現すことも可能である (小西 (1990, p. 142))。解析的アプローチが難しい場合、既知の分布 $f(x|\eta)$ に従う乱数を反復発生させることによって、モンテカルロ近似を行うこともできる。これが、パラメトリックブートストラップ法と呼ばれる手法である。

(平滑化ブートストラップ法) 基本的なブートストラップ法は、経験分布関数 \hat{F} によって未知の確率分布 F を推定した。これに対して、適当な平滑化を行った分布関数 \hat{F}_h で推定したのが、平滑化ブートストラップ法である。Efron (1982, p. 30) は、標本相関係数に対する Fisher の z -変換の標準誤差推定の問題を取り上げ、数値的に \hat{F} より \hat{F}_h に基づくブートストラップ推定の方が良い場合があることを示した。Silverman and Young (1987) は、核関数による平滑化法 (Silverman (1986)) に基づいて、推定量の平均二乗誤差の推定問題を考察した。Young (1988), De Angelis and Young (1992) は、標本相関係数に対する z -変換の標準誤差推定について、また Hall *et al.* (1989) は、推定パーセンタイルの分散の推定について、ブートストラップ法と平滑化ブートストラップ法の比較検討を行った。

もし、適当に平滑化パラメータを選ぶことによって、 \hat{F} より \hat{F}_h に基づくブートストラップ法の方がある意味で良いということが言えれば、平滑化ブートストラップ法の有用性は増す。しかし、推定量の誤差評価の問題に対しては、一般に平滑化パラメータは汎関数と F に依存し、その決め方については種々の問題があり今後の研究課題である。

(その他の手法) Rubin (1981) は、ブートストラップ標本に含まれる各観測データの割合を、ディリクレ分布で制御したベイズ的ブートストラップ法を提唱した。関連研究に、Lo (1987, 1988), Banks (1988) 等がある。有限母集団、層別抽出法に対するブートストラップ法の研究は、Bickel and Freedman (1984), Chao and Lo (1985), Rao and Wu (1988), Sitter (1992) 等があり、また方向性データに基づく統計的推測への応用は、Ducharme *et al.* (1985), Fisher and Hall (1989 a, b) によって研究された。

4.3 回帰モデル

目的変数 y と p 個の説明変数 $\mathbf{x}=(x_1, \dots, x_p)'$ に関して、 n 個の観測値 $\{(y_i, \mathbf{x}_i); i=1, \dots, n\}$ が得られたとする。このとき、 y と \mathbf{x} の関係を表すモデルとして $y_i=f(\mathbf{x}_i;\boldsymbol{\beta})+\varepsilon_i (i=1, \dots, n)$ を仮定する。ただし、 \mathbf{x}_i は事前に設定された既知の定数ベクトル、 $\boldsymbol{\beta}$ は未知のパラメータベクトル、 f の関数形は既知とする。また、誤差項 $\boldsymbol{\varepsilon}=(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ は、互いに独立に同一の未知の確率分布 F に従うものとし、 $E[\varepsilon_i]=0, E[\varepsilon_i^2]=\sigma^2$ と仮定する。

未知のパラメータベクトル $\boldsymbol{\beta}$ を、例えば最小二乗法を用いて推定し、これを $\hat{\boldsymbol{\beta}}$ と置く。推定量 $\hat{\boldsymbol{\beta}}$ の統計的誤差をブートストラップ法によって推定するとき、モデルの構造を反映するとい

う立場に立てば、以下のプロセスを通して実行される。

(1) 各点の残差 $\varepsilon_i = y_i - f(x_i; \hat{\beta})$ ($i=1, \dots, n$) を求め、その平均を $\hat{\varepsilon}_{(\cdot)} = \sum_{i=1}^n \varepsilon_i / n$ とおく。残差の平均は必ずしも 0 ではないから、平均を補正した残差を $e_i = \varepsilon_i - \hat{\varepsilon}_{(\cdot)}$ とおき、 $\{e_1, \dots, e_n\}$ に基づいて経験分布関数 \hat{F} を構成する。

(2) \hat{F} からの大きさ n の標本 e_1^*, \dots, e_n^* に対して、 $y_i^* = f(x_i; \hat{\beta}) + e_i^*$ ($i=1, \dots, n$) とおき、ブートストラップ標本 $\{(y_i^*, x_i); i=1, \dots, n\}$ をつくる。

(3) $\min_{\beta} \sum_{i=1}^n \{y_i^* - f(x_i; \beta)\}^2$ の解を $\hat{\beta}_{(n)}^*$ とおく。(2), (3) のプロセスを例えば B 回繰り返すことによって得られる $\hat{\beta}_{(n)}^*, \hat{\beta}_{(n)}^*, \dots, \hat{\beta}_{(n)}^*$ に基づいて、推定量 $\hat{\beta}$ に関する誤差評価を行うことができる。

回帰モデルにおいて、大きさ n の標本 $z_i = (y_i, x_i)$ ($i=1, \dots, n$) は、互いに独立で同一の $(p+1)$ 次元確率分布 F から観測されたと考えられることもできる。この場合、経験分布関数は $\{z_1, \dots, z_n\}$ によって構成され、ブートストラップ誤差推定のプロセスは 4.2.1 節の通りである。

(線形重回帰モデル) 線形重回帰モデル $y_i = x_i' \beta + \varepsilon_i$ ($i=1, \dots, n$) の回帰係数 $\beta = (\beta_1, \dots, \beta_p)'$ の最小二乗推定量を、ステップ (2) のブートストラップ標本から構成すると $\hat{\beta}^* = (X'X)^{-1} X' y^*$ となる。ただし、 $y^* = (y_1^*, \dots, y_n^*)'$, $X = (x_1', \dots, x_n')$ とおき、 $n \times p$ 行列 X の階数は p ($p \leq n$) とする。定数項を含む線形重回帰モデルで、説明変数行列 X の第一列の要素をすべて 1 とした場合、ステップ (1) の残差の平均は 0 となり残差の補正は必要ないことに注意する。なお、Bickel and Freedman (1983) は残差の $\{n/(n-p)\}^{1/2}$ 補正を、Weber (1984), Stine (1985) はスチューデント化された残差の利用を提案している。

線形重回帰モデルへのブートストラップ法の応用、特に、 $(X'X)^{1/2}(\hat{\beta} - \beta)/\sigma$ に対するブートストラップ分布の有効性、漸近理論の研究は、Freedman (1981), Bickel and Freedman (1983) によるところが大きい。さらに、Freedman (1984), Freedman and Peters (1984 a, b) では、一般化最小二乗推定量、ダイナミック線形モデルにおける二段階最小二乗推定量などに対するブートストラップ法の適用研究が行なわれた。Peters and Freedman (1984) には、これらの研究を通して得られた適用上の留意点が簡潔にまとめられている。なお、回帰係数の推定に M -推定量を用いた場合 (Shorack (1982), Lahiri (1992), より一般には Arcones and Giné (1992)), L_1 -norm 推定量を用いたときの種々の数値比較 (Stangenhuis (1987), Dielman and Pfaffenberger (1988)) の研究がある。

回帰係数に対する信頼区間の構成への応用は、Robinson (1987), Hall (1989 a) の研究が、リッジ回帰におけるリッジパラメータの推定への応用は、Delaney and Chatterjee (1986) にみられる。Wu (1986) は、誤差項に対する仮定の一つである等分散性が満たされない場合、 $\hat{\beta}$ のブートストラップ分散推定の漸近的一致性は成立しないことを示し、種々の対処法を提唱した。関連研究として Shao (1988) などがある。

将来観測されるデータ $y_0 = x_0' \beta + \varepsilon_0$ に対する予測域を構成するには、 y_0 の予測値 $\hat{y}_0 = x_0' \hat{\beta}$ との差 $R(y, F) = y_0 - \hat{y}_0$ の分布を推定する必要がある。ブートストラップ法では、 \hat{F} からのブートストラップ標本 $\{e_1^*, \dots, e_n^*\}$ および e_0^* を抽出して、 $\hat{\beta}^*$ と $y_0^* = x_0' \hat{\beta} + e_0^*$ を求める。このとき、基本的には $R(y^*, \hat{F}) = y_0^* - x_0' \hat{\beta}^*$ の分布の 100 α % 点 \hat{x}_α を用いて、 $[x_0' \hat{\beta} + \hat{x}_\alpha, x_0' \hat{\beta} + \hat{x}_{1-\alpha}]$ と構成される。予測域の構成については Stine (1985) を、また同時信頼領域の構成に関しては Hall and Pittelkow (1990) を参照されたい。予測誤差のブートストラップ推定に関しては、Efron (1983, 1986), Bunke and Droge (1984), Kipnis (1992) などの研究がある。

(ノンパラメトリック回帰, 密度関数の推定) 核関数を用いたノンパラメトリック回帰および密度関数の推定に於ける共通の問題点は、平滑化パラメータの決め方にある。確率密度関数 f からの大きさ n の無作為標本 x_n に基づく、核関数を利用した密度関数の推定量は $\hat{f}_n(x; h) =$

$(nh)^{-1} \sum_{i=1}^n K\{(x-X_i)/h\}$ で与えられる。ただし、 h は平滑化パラメータとする。平滑化パラメータは、 $E \int \{\hat{f}_n(x;h) - f(x)\}^2 dx$ を最小にするように選択される場合が多い。平均二乗誤差は、推定量のバイアスと分散に分解され、その推定にブートストラップ法の適用が考えられる。

経験分布関数 \hat{F} に基づく通常のブートストラップ法を適用するとき、ブートストラップ標本 x_n^* から $\hat{f}_n^*(x;h) = (nh)^{-1} \sum_{i=1}^n K\{(x-X_i^*)/h\}$ をもとめ、 $\{\hat{f}_n(x;h), f(x)\}$ を各々 $\{\hat{f}_n^*(x;h), \hat{f}_n(x;h)\}$ で置き換えて実行する。ところが、一般に核関数に基づく推定量に対しては、このような方法ではブートストラップバイアス推定は有効に働かない (例えば、Hall (1992 a, p. 205))。これは、推定量が n 個の標本の線形関数となっているため、期待値をとると $E_F[\hat{f}_n^*(x;h)] = \hat{f}_n(x;h)$ となり、理論上バイアスを 0 と推定してしまうことによる。

ブートストラップ法を適用するに当たって、この点を克服するためのいくつかの方法が提唱されてきた。Taylor (1989), Faraway and Jhun (1990) は、経験分布関数に代えて平滑化ブートストラップ法を適用し、Hall (1990 a) は、大きさが n より小さいブートストラップ標本に基づく手法を提案した。これらの手法の概説は、Hall (1992 a, 4.4 節), Marron (1992) にみられる。

核関数を利用したノンパラメトリック回帰へのブートストラップ法の応用は、Härdle and Bowman (1988), Härdle *et al.* (1988), Dikta (1990), Faraway (1990), Cao-Abad (1991), Härdle and Marron (1991) 等があり、この分野の研究は Hall (1992 a, 4.5 節), Mammen (1992) にまとめられている。その他、Romano (1988 a, 1988 b) のモード推定への応用などがある。

4.4 予測誤差推定

ブートストラップ法を予測誤差の推定に応用することによって、従来解析的アプローチが難しかった複雑な問題に対して、有効な解を与えることが可能となりつつある。このような例として、判別分析における誤判別率の推定、情報量規準 AIC (Akaike (1973)) に於ける対数尤度のバイアス補正の問題を取り上げ、その基本的な考え方と関連研究を紹介する。

4.4.1 判別分析における誤判別率推定

二つの母集団 (群) Π_1, Π_2 があり、各々未知の p 次元確率分布 $F_1(x), F_2(x)$ をもつとする。群 $\Pi_1(\Pi_2)$ からの標本が $n_1(n_2)$ 個観測されたとし、これらを $x_n = \{x_i^{(j)}; j=1, \dots, n_i; i=1, 2\}$ ($n = n_1 + n_2$) とおく。ここでは、各群から観測されたデータを通して得られる情報を基に確率分布モデルを想定し、これらを $g_i(x|\theta_i)$ ($i=1, 2$) とする。また、標本 x_n に基づいて、何らかの方法で推定された確率分布モデルを $\hat{g}_i(x|x_n)$ とおく。このとき、例えば、判別関数を $h(x|x_n) = \hat{g}_1(x|x_n)/\hat{g}_2(x|x_n)$ と置くと、標本空間を二分する判別領域 $\hat{R}_1 = \{x: h(x|x_n) > c\}$, $\hat{R}_2 = \{x: h(x|x_n) \leq c\}$ が構成され、新たに観測されたデータ x_0 に対して、 $x_0 \in \hat{R}_1$ のとき x_0 は群 Π_1 へ属すると判別を実行することができる。

判別分析に於ける予測誤差とは、初期標本 x_n とは独立に群 Π_i から採られたデータ、すなわち確率分布 F_i に従って取り出されたデータを、誤って群 Π_j ($j \neq i$) からのものと予測してしまう確率

$$\eta_i(F_i; x_n) = \int_{\hat{R}_j} dF_i(x) \quad (i, j=1, 2; i \neq j)$$

であり、一般に実際の誤判別率または条件付き誤判別率と呼ばれている。

実際の誤判別率の一つの推定法は、上式に含まれる未知の確率分布 F_i を群 Π_i からの初期標本に基づく経験分布関数 \hat{F}_i で置き換えた推定量

$$\hat{\eta}_i(\hat{F}_i; \mathbf{x}_n) = \int_{\hat{R}_i} d\hat{F}_i(\mathbf{x}) = \int I(\mathbf{x}|\hat{R}_i) d\hat{F}_i(\mathbf{x}) = \frac{1}{n_i} \sum_{a=1}^{n_i} I(\mathbf{x}_a^{(i)}|\hat{R}_i)$$

を用いることである。ここで、 I は、 $I(\mathbf{x}|\hat{R}_i) = 1(\mathbf{x} \in \hat{R}_i); = 0(\mathbf{x} \notin \hat{R}_i)$ で与えられる定義関数とする。これは、群 Π_i からの初期標本のうち、判別領域 \hat{R}_i へ入った結果誤判別された標本の割合であり、よく知られた見かけ上の誤判別率に他ならない。

ブートストラップ法は、実際の誤判別率を見かけ上の誤判別率という一つの推定量で推定したときのバイアス $b_i(F_1, F_2) = E_{F_1, F_2}[\hat{\eta}_i(\hat{F}_i; \mathbf{x}_n) - \eta_i(F_i; \mathbf{x}_n)]$ の推定に用いられた。ブートストラップ法を適用すると、このバイアスは

$$\hat{b}_i(\hat{F}_1, \hat{F}_2) = E_{\hat{F}_1, \hat{F}_2} \left[\frac{1}{n_i} \sum_{a=1}^{n_i} I(\mathbf{x}_a^{(i)*}|\hat{R}_i^*) - \frac{1}{n_i} \sum_{a=1}^{n_i} I(\mathbf{x}_a^{(i)}|\hat{R}_i^*) \right]$$

と推定され、経験分布関数の同時分布に関する期待値は、モンテカルロ法によって数値的に近似される。このとき、見かけ上の誤判別率のバイアスを補正した $\hat{\eta}_i(\hat{F}_i; \mathbf{x}_n) - \hat{b}_i(\hat{F}_1, \hat{F}_2)$ を、実際の誤判別率の推定量とする（詳細な実行プロセスは、McLachlan (1992, 10 節)、小西・本多 (1992, 3 節) を参照されたい）。

判別分析に於ける誤判別率推定へのブートストラップ法の応用は、Efron (1979, 1983, 1986) の研究によるところが大きい。Efron は、バイアス補正を施した見かけ上の誤判別率は、交差検証法 (Lachenbruck and Mickey (1968), Stone (1974), Geisser (1975)) と比較して、変動の小さい推定量としてその有効性を示した。しかし、母集団の設定によっては、ブートストラップバイアス補正は十分に働かないことも指摘された (Efron (1983), Chatterjee and Chatterjee (1983))。Davison and Hall (1992) は両母集団の平均間の差が $n^{-1/2}$ のオーダーで 0 に収束するとき、ブートストラップバイアス補正を交差検証法との比較に於いて理論的に検討した。

Efron (1983) は、実際の誤判別率に対する推定量の平均二乗誤差に基づく評価から出発して、様々な改良法を提唱した。これらは、確率化ブートストラップ法 (Randomized bootstrap)、二段階ブートストラップ法 (Double bootstrap)、0.632 推定量と呼ばれ、中でも 0.632 推定量の有効性を数値的に示した。

ブートストラップ法に基づく推定法を含めた種々の誤判別率推定法の有効性、問題点は、主として数値実験を通して検証されつつある。正規母集団からの標本に基づいて線形判別分析を実行したときの手法の様相は、McLachlan (1980), Chernick *et al.* (1985, 1986), Wernecke and Kalb (1987), Sánchez and Cepeda (1989), Fitzmaurice *et al.* (1991) 等で研究された。Snapinn and Knoke (1988), Ganeshanandam and Krzanowski (1990), Konishi and Honda (1990) 等では、非正規モデルのもとでも合わせて検討され、Wang (1986) は多項分布モデルのもとで検討している。Jain *et al.* (1987) は二次判別分析における様相も調べている。

変数選択への応用は、Honda and Konishi (1988), Snapinn and Knoke (1989) が線形判別に基づいて、Efron and Gong (1983), Gong (1986) がロジスティック判別 (回帰) への適用を通して検証している。また、各種推定法を総合的に報告した論文、著書としては、McLachlan (1986, 1987, 1992; 10 節), Hand (1986), 小西・本多 (1992) がある。

4.4.2 情報量規準

この節では、予測誤差を確率分布に基づきグローバルな観点から捉えた情報量規準 AIC (Akaike (1973)) を取り上げ、ブートストラップ法の枠組みの中で、推定量のバイアス補正について検討する。はじめに、Akaike (1973, 1974), 竹内 (1976), 坂元 他 (1983), Shibata (1989) などを参考に、情報量規準の基本的な考え方を簡単に整理する。

未知の確率分布関数 $F(x)$ からの大きさ n の無作為標本を \mathbf{x}_n とする。確率分布関数 $F(x)$ の密度関数を $f(x)$ とし、これに対して想定したモデルの密度関数を $g(x|\theta)$ とする。モデルに含まれる未知のパラメータ $\theta (\in \Theta)$ は、 p 次元パラメータベクトルとする。このような設定のもとで、将来観測されるデータ z に対する(予測)確率分布 $\bar{g}(z|\mathbf{x}_n)$ を構成したいとする。一つの方法は、想定したモデルの確率分布 g に含まれるパラメータ θ を何らかの方法で推定し、求めた推定値 $\hat{\theta}$ で置き換えた $\bar{g}(z|\mathbf{x}_n) = g(z|\hat{\theta})$ を用いる方法である。その他、ベイズ的方法による予測分布、ABIC 最小化事後モード法 (Akaike (1980)) など様々な方法が提案されている。

大きさ n の標本に基づいて推定された一つの予測確率分布 $\bar{g}(z|\mathbf{x}_n)$ と、この標本を生成した真の確率分布 $f(z)$ との距離を Kullback-Leibler 情報量で測るとする。このとき、標本 \mathbf{x}_n によって推定される種々の予測確率分布の違いは、平均対数尤度と呼ばれる項

$$(4.9) \quad \eta(F; \mathbf{x}_n) = \int f(z) \log \bar{g}(z|\mathbf{x}_n) dz = \int \log \bar{g}(z|\mathbf{x}_n) dF(z)$$

が関係する。平均対数尤度は、真の確率分布 F と予測確率分布の推定を通して標本 \mathbf{x}_n に依存する未知の量である。そこで、平均対数尤度の一つの推定量として、(4.9) 式に含まれる未知の確率分布 F を、標本 \mathbf{x}_n に基づく経験分布関数 \hat{F} で置き換えた

$$(4.10) \quad \hat{\eta}(\hat{F}; \mathbf{x}_n) = \int \log \bar{g}(z|\mathbf{x}_n) d\hat{F}(z) = \frac{1}{n} \sum_{a=1}^n \log \bar{g}(X_a|\mathbf{x}_n)$$

を用いることができる。

予測確率分布を構成するために用いたデータを再び利用して、未知の確率分布 $F(z)$ を推定していることから、推定量 (4.10) は見かけ上の推定量であるといえる。判別分析に於ける見かけ上の誤判別率のバイアスを補正したと同様に、対数尤度と呼ばれる一つの推定量で平均対数尤度を推定したときのバイアス $b(F) = E_F[\hat{\eta}(\hat{F}; \mathbf{x}_n) - \eta(F; \mathbf{x}_n)]$ の補正が必要となる。したがって、このバイアスを何らかの方法で推定できれば、対数尤度のバイアスを補正した一つの情報量規準

$$IC(\hat{F}; \mathbf{x}_n) = \frac{1}{n} \sum_{a=1}^n \log \bar{g}(X_a|\mathbf{x}_n) - b(F) \text{ の推定量}$$

がもたらす。情報量規準 AIC は、基本的には、このように対数尤度で平均対数尤度を推定したときの漸近的なバイアスを補正した推定量として与えられた。

ブートストラップ法を適用するとバイアスは

$$\hat{b}(\hat{F}) = E_{\hat{F}}[\hat{\eta}(\hat{F}^*; \mathbf{x}_n^*) - \hat{\eta}(\hat{F}; \mathbf{x}_n^*)] = E_{\hat{F}}\left[\frac{1}{n} \sum_{a=1}^n \log \bar{g}(X_a^*|\mathbf{x}_n^*) - \frac{1}{n} \sum_{a=1}^n \log \bar{g}(X_a|\mathbf{x}_n^*)\right]$$

と推定される。ただし、 \mathbf{x}_n^* は大きさ n のブートストラップ標本、 \hat{F}^* はブートストラップ標本の各点に確率 $1/n$ をもつ経験分布関数とする。期待値は、モンテカルロ法によって数値的に計算される。

ブートストラップ法は極めて緩やかな仮定のもとで、解析的に煩雑な手続きなしで適用できる。しかし、ブートストラップバイアス推定には、標本変動とシミュレーションによる変動(4.6節を参照)があり、その大きさには十分注意を払う必要がある。これに対して、最尤推定量を用いた情報量規準 AIC は、漸近的バイアス評価に関して解析的導出の困難さも推定による変動も取り除かれ、実際問題への適用上極めて有用な手法といえる。

ブートストラップ法を適用して対数尤度のバイアスを補正する方法は, Wong (1983), Efron (1986) によって暗に示され, 後に石黒, 北川, 坂元がその重要性を指摘した. この方法は EIC と名づけられ, Ishiguro and Sakamoto (1991), 石黒 他 (1992), 北川 (1991), 北川 他 (1992), 坂元 他 (1992) によって, その有用性が検討されつつある. 小西 (1992 a, 1992 b) は, 想定した確率分布モデルを規定するパラメータを, 統計的汎関数で定義される推定量で推定したとき, 対数尤度のバイアスは, 推定量の影響関数とモデルのスコア関数で表されることを示した. 漸近理論に基づく解析的アプローチとブートストラップ法による数値的アプローチとの関係については, 小西 (1992 b) を参照されたい.

ここでは, ブートストラップ法の応用という立場から, 標本は独立同分布に従うと仮定した. この仮定を取り除いた場合, 回帰モデル, 時系列モデルにおけるブートストラップ法の応用で研究されつつあるように, モデルの構造を反映するように手法の修正が必要である. 時系列モデルにおけるブートストラップ法の応用に関しては, Freedman (1984), Findley (1986), Swanepoel and van Wyk (1986), Bose (1988), Künsch (1989), Basawa *et al.* (1989), Stoffer and Wall (1991), Franke and Härdle (1992) 等の研究がある. これらの論文のいくつかに関しては, Léger *et al.* (1992, 3 節) の中で, 実際のデータの分析例を含めて簡潔にまとめられているので参照されたい.

最近, Kotz and Johnson (eds.) (1992) に, ここ約 100 年の間に発表された統計学に関する論文の中で, その発展に著しく寄与した論文 39 編が選ばれた. Akaike (1973) と Efron (1979) の論文は共にその中に選ばれ, 情報量規準 AIC, 統計的数値計算法としてのブートストラップ法が, 極めて適用範囲の広い柔軟な手法であることをうかがわせる. これらの方法は, より複雑な問題に対して有効な解を与える統計手法としての潜在性も十分に残している.

4.5 近似信頼区間の構成

ブートストラップ法を適用して推定量の分布およびパーセンタイルを推定したとき, その近似精度はどの様に評価したらよいであろうか. また, 従来の推定量の漸近正規性に基づく手法と比べて, その有効性は保証されるのであろうか. 本節では, このような問題から出発して, 種々の近似信頼区間の提唱へと展開して行ったブートストラップ法の理論研究を概観する. さらに, 近似信頼域の構成, 検定問題への応用について簡単にふれる.

4.5.1 ブートストラップ分布の近似精度

ブートストラップ分布の近似精度を評価するためには, 真の分布とそのブートストラップ分布との差, あるいは真のパーセンタイルとそのブートストラップ推定値との差を, 何らかの方法で陽に表す必要がある. そのために用いられた基本的な道具が, エッジワース展開と呼ばれる漸近展開式と, パーセンタイルの展開式であるコーニッシュ・フィッシャー (逆) 展開であった. これらの展開式については, 竹内 (1975), Petrov (1975), 清水 (1976), Bhattacharya and Rao (1976), Pfanzagl (1985), Barndorff-Nielsen and Cox (1989), Hall (1992 a) を参照されたい.

例えば, スチューデント化された統計量 $T_{ST} = \sqrt{n}(\hat{\theta}_n - \theta)/\hat{\sigma}$ に基づくブートストラップ分布の精度は, 有効なエッジワース展開 (Bhattacharya and Ghosh (1978), Hall (1992 a, 2 章)) が可能であるという仮定のもとで, 次のように評価される.

統計量 T_{ST} の分布のエッジワース展開は, 極限分布である標準正規分布関数 $\Phi(x)$ を第一項として, $1/\sqrt{n}$ の巾のオーダーで展開された式で与えられる. 各項は, 標準正規密度関数とその係数が母集団確率分布 F のモーメントの関数 (キュムラント) として与えられる多項式との積となっている. この各係数のモーメントを標本モーメントで置き換えた式が, ブートストラップ

分布 $P_F(T_{ST}^* \leq x)$ のエッジワース展開であり、経験エッジワース展開と呼ばれている。ここで、 $T_{ST}^* = \sqrt{n}(\hat{\theta}_n^* - \theta) / \hat{\sigma}^*$ とする。

これによって、スチューデント化された統計量の分布を、ブートストラップ分布で近似したときの誤差のオーダーは、 $P_F(T_{ST}^* \leq x) - P_F(T_{SD} \leq x) = O_p(n^{-1})$ と評価される。これは、 T_{ST}^* のブートストラップ分布が、推定量の漸近的なバイアス、歪みを、自動的にある程度捉えていることを意味する。一方、 $T_{SD} = \sqrt{n}(\hat{\theta}_n - \theta)$ に対するブートストラップ分布に関しては、 $P_F(T_{SD}^* \leq x) - P_F(T_{SD} \leq x) = O_p(n^{-1/2})$ となり、近似誤差のオーダーを通して近似精度の違いが明らかとなる。推定量の漸近正規性に基づく近似誤差のオーダーも $n^{-1/2}$ であるから、スチューデント化された統計量に対するブートストラップ分布の有効性をみることができる。ただし、確率分布 F が格子点分布の場合、エッジワース展開の $n^{-1/2}$ の項に含まれる連続補正項のブートストラップ近似は有効に働かず、このような議論は成り立たないので注意が必要である (Singh (1981), Hall (1987a, 1992a: p. 90)). なお、記号 O, O_p については、例えば、小西 (1990, p. 144), Hall (1992 a, p. 88) を参照されたい。

ブートストラップ分布の近似誤差のオーダーを、エッジワース展開を道具として評価する方法は、Bickel and Freedman (1981), Singh (1981) によって行われ、その後のブートストラップ法の理論研究に大きな影響を与えた。スチューデント化された統計量を含めた研究は、Babu and Singh (1983, 1984, 1985), Beran (1982, 1984 a, 1984 b), Hall (1986 b, 1988 a, 1990 b), Bhattacharya and Qumsiyeh (1989), Bickel (1992) 等によって幅広く研究された。このような研究に関しては、レビュー論文 DiCiccio and Romano (1988, 3節) および Hall (1992 a, 3章) を参照されたい。

ブートストラップ法による分布推定が有効に作用しない場合も検討された。Bickel and Freedman (1981) は、 U -統計量、極値統計量に基づいて構成された反例を上げ、Mammen (1992, 2章) は、ノンパラメトリック回帰に於ける例を述べた。その他、DiCiccio and Romano (1988, 3.3節) にいくつかの反例がまとめられている。

4.5.2 近似信頼区間の精度

4.2節(4.3)式のスチューデント化された統計量 T_{ST} の真の $100\alpha\%$ 点 $x_\alpha^{(ST)}$ が、仮に求めたとする。このとき、信頼係数 $1-\alpha$ の片側信頼区間は、 $I_{ST}^{(Y)} = (-\infty, \hat{\theta}_n - n^{-1/2} \hat{\sigma} x_\alpha^{(ST)})$ あるいは $[\hat{\theta}_n - n^{-1/2} \hat{\sigma} x_\alpha^{(ST)}, \infty)$ で与えられる。また、パラメータ θ がこの片側信頼区間に含まれる確率は、点 $x_\alpha^{(ST)}$ が真の $100\alpha\%$ 点であることから正確に $1-\alpha$ になる。

実際には、 $x_\alpha^{(ST)}$ の値は未知であるから、ブートストラップ法で推定された(4.5)式の推定値 $\hat{x}_\alpha^{(ST)}$ を用いて、近似片側信頼区間 $\hat{I}_{ST}^{(Y)} = (-\infty, \hat{\theta}_n - n^{-1/2} \hat{\sigma} \hat{x}_\alpha^{(ST)})$ あるいは $[\hat{\theta}_n - n^{-1/2} \hat{\sigma} \hat{x}_\alpha^{(ST)}, \infty)$ を構成する。このとき、 θ がこの区間に含まれる確率は、 $0 < \alpha < 1$ に対して

$$(4.11) \quad P_F\{\theta \leq \hat{\theta}_n - n^{-1/2} \hat{\sigma} \hat{x}_\alpha^{(ST)}\} = 1 - P_F\{\sqrt{n}(\hat{\theta}_n - \theta) / \hat{\sigma} < \hat{x}_\alpha^{(ST)}\} = 1 - \alpha + (\text{誤差項})$$

と表わせる。誤差項は被覆誤差 (Coverage error) と呼ばれ、近似誤差を評価するための一つの目安となる。また、近似信頼限界とその真の信頼限界を直接比較して

$$(4.12) \quad \hat{\theta}_n - n^{-1/2} \hat{\sigma} \hat{x}_\alpha^{(ST)} - \{\hat{\theta}_n - n^{-1/2} \hat{\sigma} x_\alpha^{(ST)}\} = (\text{確率的誤差項})$$

を評価することもできる。ブートストラップ法に基づく近似信頼区間の精度は、 $n \rightarrow \infty$ のとき、誤差の項が0に収束するオーダーを評価することによって捉えることができる。この分野の研究は、特に、Hall (1988 a, 1992 a; 3章) によるところが大きい。実際、(4.11), (4.12) 両式の誤差項のオーダーはそれぞれ

$$P_F\{\theta \in \hat{I}_{ST}^{(F)}\} = 1 - \alpha + O(n^{-1}), \hat{\theta}_n - n^{-1/2} \hat{\sigma} \hat{x}_a^{(ST)} - \{\hat{\theta}_n - n^{-1/2} \hat{\sigma} x_a^{(ST)}\} = O_p(n^{-3/2})$$

となることが示される (Hall (1988 a)). 特に後半の式は、スチューデント化された統計量の分布に基づく理論信頼限界とブートストラップ信頼限界が、 $n^{-1} = (n^{-1/2})^2$ の項まで一致していることを示し、このとき、近似信頼限界は二次の精度をもつと定義する。

これに対して、統計量 $T_{SD} = \sqrt{n}(\hat{\theta}_n - \theta)$ に基づく 4.2 節 (4.6) 式の 100 α % 点のブートストラップ推定値 $\hat{x}_a^{(SD)}$ を用いた $1 - \alpha$ 片側信頼区間に対しては、その被覆誤差は $O(n^{-1/2})$ となることが示される。また、ブートストラップ信頼限界と理論信頼限界の差は $O_p(n^{-1})$ となり、これは一次の精度しかもたないことを意味する。両側信頼区間に対しては、基準量 T_{SD}, T_{ST} どちらに基づいても共にその被覆誤差は $O(n^{-1})$ であり、片側信頼区間の場合と収束のオーダーが変わることに注意する。

ブートストラップ法に基づく区間推定の近似精度の評価は、このように標本数に関して漸近的な様相を調べ、一つの目安を与えることで行われてきた。スチューデント化された統計量を基準量として用いることによって近似精度の改善が計られるということは、Efron (1982), Babu and Singh (1983), Hinkley and Wei (1984) 等によって示唆され、区間推定に関しては、Hall (1988 a) によって包括的な研究が行われた。そこでは、例えば $P_F(|T_{ST}^*| \leq \hat{x}_a)$ のパーセンタイルを用いた対称信頼区間 (Hall (1988 b)), 区間の長さを考慮した信頼区間に対する近似精度も研究された。

有効なエッジワース展開に基づく上述の議論を行うためには、パラメータ θ の推定量に制約を置く必要がある。ブートストラップ法の理論研究でしばしば用いられたのが、多変量ベクトル平均の十分滑らかな関数として表される推定量である (Bhattacharya and Ghosh (1978), その他, 小西 (1990, p. 157), Hall (1992 a, p. 52)). また、十分滑らかな統計的汎関数で定義される推定量も用いられた (Sen (1988), Akahira and Takeuchi (1991), Konishi (1991) 他). 統計的汎関数のエッジワース展開については、von Mises (1947), Reeds (1976), Withers (1983), Beran (1984 a), Pfanzagl (1985), Takahashi (1988) 等を参照されたい。

4.5.3 近似精度の改善

ブートストラップ法の近似精度は、エッジワース展開を通して捉えることができ、有効に基準量を設定すれば精度の改善が計られることが分かる。本節では、近似精度改善のために提案されてきたいくつかの手法を紹介する。

(変換に基づく方法) Abramovitch and Singh (1985) は、スチューデント化された統計量に対してコーニッシュ・フィッシャー展開を適用することにより、二次の多項式で与えられる基準量 (T_{AS}) を提案した。この基準量の分布を正規近似した場合、誤差項のオーダーは n^{-1} であるのに対して、 T_{AS} に基づくブートストラップ分布を用いれば、より高次の精度が得られることを示した。しかし、この方法は多項式変換に基づくことから、歪みの大きな推定量に対しては、 T_{AS} の定義域内で単調性がくずれ適用できない場合があることから利用上注意を要する。これに対して、Konishi (1991) は、正規化変換の理論 (Konishi (1981, 1987)) に基づいて、この欠点を克服した単調性を有する基準量を提唱し、近似精度に関しても T_{AS} と同様の精度をもつことを示した。Hall (1992 b) は、三次の多項式変換を用いて単調性を有する基準量を提唱した。

では、例えば標本相関係数のように、有効な分散推定が難しい場合にはどうすればよいか。一つの方法は、推定量を変換して、その分散ができるだけ未知の確率分布 F に依存しないようにする方法が考えられる。Tibshirani (1988) は、二段階ブートストラップ法と平滑化法を併用して、分散安定化変換を数値的に求めるアルゴリズムを提案した。関連研究として DiCiccio

and Romano (1990) がある。しかし、ノンパラメトリックモデルのもとで、変換形をどのように推定するかは今後の研究課題である。なお、Davison *et al.* (1992) は、Tibshirani (1988) の提唱したものと基本的には同じアルゴリズムによって、部分尤度関数の推定を試みている。

(Efron の方法) Efron (1981, 1982) によって提唱された近似信頼区間は、当初ブートストラップ分布 $\hat{G}(x) = P_F(\hat{\theta}_n^* \leq x)$ の 100 α % 点を用いるものであった。その近似精度は、推定量のバイアス、歪みの大きさに影響され、精度の点で問題があることが指摘された。これに対して Efron (1981, 1985, 1987) は、その一連の論文の中で、基本的には変換に基づくが、他とは異なる観点から近似信頼区間の精度の改良を試みてきた。なお、 $\hat{\theta}_n^*$ のブートストラップ分布を用いた場合と $\sqrt{n}(\hat{\theta}_n^* - \hat{\theta})$ の分布を用いた場合とでは、異なる信頼限界を与えることに注意する (Hall (1988 a, p. 933; 1992 a, p. 95)).

Efron (1981, 1985) は、まずバイアスの修正を数値的に取り入れる方法を提唱し、次に Efron (1987) で、バイアスと解析的に求めた歪みの修正を同時に取り入れた近似信頼区間を提唱した。Efron の提唱したこの近似信頼区間は、BC α 区間 (Accelerated bias-corrected percentile interval) と呼ばれ、その理論的基礎を推定量の変換理論においている。その特徴は、変換のための関数を具体的に導出する過程を、ブートストラップ反復抽出で置き換えている点にある。Konishi (1991) は、統計的汎関数の枠組みの中で、この変換が分散安定化変換と正規化変換の合成関数で表されることを示した。その他 Efron の方法に基づく区間推定法としては、DiCiccio and Tibshirani (1987), DiCiccio and Romano (1989 a) 等がある。また、Efron (1992 b) では、パーセンタイル、バイアス等のブートストラップ推定値が、個々の観測データの影響をどの程度受けているかを診断する方法が提案されている。

(ブートストラップ反復法) パーセンタイルの推定値の精度を逐次改善する観点から捉えた手法が、Loh (1987, 1991) の Calibrated method である。いま、大きさ n の標本 \mathbf{x}_n に基づいて構成された信頼係数 $1 - \alpha$ の近似信頼区間を $(-\infty, \hat{\theta}(\alpha, \mathbf{x}_n)]$ とする。この区間がパラメータ θ を含む確率 $\pi(\alpha) = P_F\{\theta \leq \hat{\theta}(\alpha, \mathbf{x}_n)\}$ は、近似的に $1 - \alpha$ である。そこで、信頼限界 $\hat{\theta}(\alpha, \mathbf{x}_n)$ を構成したときの α あるいはパーセンタイルをうまく修正すれば、 $\pi(\beta_\alpha) = P_F\{\theta \leq \hat{\theta}(\beta_\alpha, \mathbf{x}_n)\} = 1 - \alpha$ とできると考えられる。

関数 $\pi(\cdot)$ は未知であるから、これをブートストラップ法で推定し $\hat{\pi}(\hat{\beta}_\alpha) = P_F\{\hat{\theta} \leq \hat{\theta}(\hat{\beta}_\alpha, \mathbf{x}_n^*)\} = 1 - \alpha$ を満たす $\hat{\beta}_\alpha$ を求め、初めの信頼限界を修正した $\hat{\theta}(\hat{\beta}_\alpha, \mathbf{x}_n)$ を新たな信頼限界とする。このプロセスを繰り返すことによって、理論上は θ に対する被覆確率を限りなく $1 - \alpha$ に近づけることができる。

これを分布関数の観点からみると、ブートストラップ分布の近似精度は、その分布ができるだけ未知の確率分布 F に依存しないような基準量 (漸近的枢軸変量) を構成することによって改善が可能となる。言い換えると、何らかの基準量 $R(\mathbf{x}_n, T(F))$ に対して、 $P_F\{R(\mathbf{x}_n, T(F)) \leq x\} = \Phi(x) + O(n^{-j/2})$ が、できるだけ大きな j に対して成り立てばよい。スチューデント化された統計量に対しては、 $j=1$ である。Beran (1987) は、このような基準量を分布関数による変換を逐次行って求める方法を提唱した。同様な考え方に基づく方法として、Hall (1986 b) の Additive correction, DiCiccio and Romano (1989 a) の Automatic percentile method があり、ブートストラップ反復法の理論研究としては、Hall and Martin (1988), Martin (1990) がある。

DiCiccio *et al.* (1992) は、ブートストラップ反復法の実行プロセスの中に鞍点近似法 (Saddlepoint approximation; Daniels (1954), Reid (1988)) を組み入れることによって、計算量を減少させる方法を提案した。ブートストラップ分布を解析的に近似するための鞍点法の利用は、Davison and Hinkley (1988), Daniels and Young (1991) 等によって研究された。

ブートストラップ反復法によって理論的に精度の改善が計られたとしても、実行上、多段階ブートストラップ標本、すなわち枝分かれ的にブートストラップ反復抽出を行う必要があり、膨大な計算時間を必要とする。また、反復のプロセスをどの段階まで実行すればよいかという問題も残る。数値的にこれらの方法が有効であるかどうかは、今後十分に検討する余地がある。

ブートストラップ反復法は、解析的アプローチに基づく Hall (1983), Withers (1983, 1984), Peers and Iqbal (1985), Abramovitch and Singh (1985) 等の研究に対応する。すなわち、エッジワース、コーニッシュ・フィッシャー展開に基づいて、推定量の高次のキュムラントを推定し、逐次取り入れていくことで達成できる。実際には、高次のキュムラントを陽に表すことは難しい場合がほとんどで、これを数値的に実行するアルゴリズムがブートストラップ反復法であるといえる。

4.5.4 近似信頼域の構成と検定

p 次元パラメータベクトル θ のある推定量を $\hat{\theta}$ とする。もし、推定量の分散共分散行列の有効な推定量が存在すれば、これを $\hat{\Sigma}$ とおく。基準量として $t_{SD} = \sqrt{n}(\hat{\theta} - \theta)$ あるいは $t_{ST} = \sqrt{n}\hat{\Sigma}^{-1/2}(\hat{\theta} - \theta)$ を用いて、 $P_F(t_{\alpha}) \in R_{\alpha} = 1 - \alpha$ となる最小の領域を何らかの方法で推定し、これを $\hat{R}_{\alpha}^{(S)}$ とおく。このとき有意水準 $1 - \alpha$ の信頼域は、 $\{\hat{\theta} - n^{-1/2}x; x \in \hat{R}_{\alpha}^{(SD)}\}$ または $\{\hat{\theta} - n^{-1/2}\hat{\Sigma}^{1/2}x; x \in \hat{R}_{\alpha}^{(ST)}\}$ で与えられる。

Hall (1987 b) は、反復抽出されたブートストラップ標本に基づく t_{ST} の値に、密度関数の推定法を適用し $\hat{R}_{\alpha}^{(ST)}$ を推定した。しかし、同時に種々の問題点があることも指摘した。同時信頼区間を含めた研究としては、標本分散共分散行列の関数として与えられる推定量に対する Beran and Srivastava (1985) の研究、因子分析における因子負荷量に対する Ichikawa and Konishi (1992) の研究、その他 Beran and Millar (1986), Beran (1988 a) の理論研究などがある。

信頼区間、信頼域を構成するための手法としては、Owen (1988) の提唱した経験尤度法がある。この手法については、Owen の論文と共に、その基本的な考え方を手短かに述べた小西 (1990, p. 155) および Hall and Scala (1990) を参照されたい。経験尤度法の漸近的な性質を中心とした研究は、Owen (1990), DiCiccio *et al.* (1991) で行われ、また、DiCiccio *et al.* (1989) は、あるパラメトリックモデルの尤度関数と経験尤度関数の比較を行っている。さらに、符号付き経験尤度比統計量の被覆誤差の改良法 (DiCiccio and Romano (1989 b)), 経験尤度に基づく信頼域の精度の改善法 (Hall (1990 c)), 経験尤度法の回帰モデルへの応用 (Owen (1991)) がそれぞれ研究されている。

ブートストラップ法の検定問題への応用研究も行なわれた。一つの方法は、構成した近似信頼区間あるいは信頼域へ、設定したパラメータの値が含まれるか否かによって検定できる。あるいは、ブートストラップ法の考え方を応用して、仮説のもとで検定統計量の分布を近似したり、 p 値の推定を行うことも基本的には可能である。しかし、複雑な問題設定に対しては種々の工夫が必要となる。各種検定問題への応用については、Beran and Srivastava (1985), Beran (1986), Beran (1988 b), Romano (1988 c), Boos and Brownie (1989), Hall and Hart (1990), Hall and Wilson (1991), Nagao and Srivastava (1992), Zhang and Boos (1992) 等を参照されたい。

4.6 有効なブートストラップシミュレーション

例えば、大きさ n の標本 x_n に基づく推定量 $\hat{\theta}_n$ のパラメータ θ に対するブートストラップバヤス推定は、次のように実行される (4.2 節 (4.1), (4.7) 式)。

$$(4.13) \quad b(F) = E_F[\hat{\theta}_n - \theta] \leftarrow \text{推定} - \hat{b}(\hat{F}) = E_F[\hat{\theta}_n^* - \hat{\theta}] \approx \hat{b}_B = \frac{1}{B} \sum_{i=1}^B \hat{\theta}_{(i)}^* - \hat{\theta}$$

すなわち、バイアス $b(F)$ は、モンテカルロ法によって数値的に近似された \hat{b}_B でもって推定される。近似値 \hat{b}_B は、標本 \mathbf{x}_n が与えられたという条件のもとで、ブートストラップ反復回数を無限大にすると、確率1で $\hat{b}(\hat{F})$ に収束する。従って、有限な B に対してはシミュレーション誤差が生じ、この誤差を制御するための種々の方法が研究されてきた。有効なブートストラップシミュレーション法 (Efficient bootstrap simulation) とは、標本 \mathbf{x}_n が与えられたもとで、 \hat{b}_B の分散を可能な限り小さくするための手法と考えることができる。分散、確率分布、パーセンタイルの推定に対しても同様である。

有効なブートストラップシミュレーション法を適用すれば、通常の実験分布関数からリサンプリングを行う方法 (一様リサンプリング) と比較して、相対的にブートストラップ標本の反復抽出の回数を減らすことも可能となる。これは、4.5.3節で述べた多段階のブートストラップ標本の抽出を必要とするような、ブートストラップ反復法を適用する際に特に有効となる。

これまでに提唱されたいくつかの手法の中で、最も実用性が高いと思われるのが、Davison *et al.* (1986) による釣合い型ブートストラップ (Balanced bootstrap) である。これは、 n 個の観測データの各々が、 B 回のリサンプリングの中で同じ回数だけ現れるようにしたもので、簡単に述べると次のようにしてブートストラップ標本を抽出する。例えば、大きさ3の標本 $\{x_1, x_2, x_3\}$ に対して、仮に3回の反復抽出を行うとする。まず、データの添え字のコピーを3組作り $\{1, 2, 3, 1, 2, 3, 1, 2, 3\}$ と並べる。次に、これをランダムに並べ替え、その結果 $\{2, 1, 1, 3, 2, 1, 3, 2, 3\}$ となったとする。釣合い型ブートストラップでは、 $\{x_2, x_1, x_1\}$, $\{x_3, x_2, x_1\}$, $\{x_3, x_2, x_3\}$ をブートストラップ反復標本とする。

この釣合型ブートストラップ法に対して、計算実行上の異なる3通りのアルゴリズムが Gleason (1988) で与えられた。Graham *et al.* (1990) には、高次の釣合い型ブロック計画との関係が述べられている。Hall (1989 b, 1990 d) は、推定量が多変量平均ベクトルの滑らかな関数として与えられるとき、釣合型ブートストラップの漸近的性質を明らかにした。

Davison *et al.* (1986) は、統計的汎関数で定義される推定量 $\hat{\theta}_n$ に対して、 $\hat{\theta}_n^* - \hat{\theta}$ から線形項、すなわち (経験) 影響関数の項を引き去った統計量に、通常のブートストラップ法の適用を提案した。これは、線形近似法 (Linear approximation method) と呼ばれている。Efron (1990) は、線形項を他の方法で置き換える方法を提案し、これは、中心化法 (Centring method) と呼ばれている (Hall (1992 c, 4節), Do and Hall (1992) を参照)。Hall (1989 b) は、推定量が多変量平均ベクトルの滑らかな関数として与えられるとき、釣合型ブートストラップ、線形近似法、中心化法の三つの手法は、漸近的には同値であることを示した。

一般に、平均、分散が共に等しく、共分散が負の二つの推定量を等ウェイトで合わせると、新たな推定量の平均はもとの推定量に等しいが、分散は小さくなる。Hall (1989 c) は、この性質を利用してリサンプリングのアルゴリズムを提案し、Antithetic resampling と名づけた。Do (1992) は、釣合型および Antithetic リサンプリング法を数値的に比較検討し、釣合型ブートストラップ法の有効性を指摘した。

Johns (1988) は、Hammersley and Handscomb (1964) の Importance sampling を応用したリサンプリング法を提唱し、パーセンタイルの推定に於いて、一様リサンプリングを基準として、反復抽出の回数を一けた減らすことができることを示した。基本的なブートストラップ法は、 n 個の観測値の各点に一樣に確率 $1/n$ を付与した経験分布関数を用いた。これに対して、各観測値にあるウェイトをおいた確率分布を構成し、そこから標本抽出を実行する。直感

的には、推定量の分布の裾の部分のパーセンタイルの推定に有効な標本の抽出確率を高くするよう、観測値に基づいて新たな分布関数を再構成したといえる。基本的な考え方の説明は、例えば小西 (1990, p. 156), Hall (1992 a, p. 298) を参照されたい。

Hinkley and Shi (1989) は、Importance sampling の考え方を、Beran (1987) の信頼区間の構成法へ応用し、二段階ブートストラップ法の反復回数を減少させる方法を提示した。Do and Hall (1991) は、一様リサンプリングを用いて Importance sampling の基本となるウエイトの一つの決め方を提示した。

謝辞：査読者の有益なコメントに感謝します。

参 考 文 献

- [1] Abramovitch, L. and Singh, K. (1985). Edgeworth corrected pivotal statistics and the bootstrap. *Ann. Statist.* **13**, 116-132.
- [2] Akahira, M. and Takeuchi, K. (1991). Bootstrap method and empirical process. *Ann. Inst. Statist. Math.* **43**, 297 - 310.
- [3] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Proceedings of 2nd International Symposium on Information Theory* (Petrov, B. N. and Csaki, F. eds.), Akademiai Kiado, Budapest, pp. 267-281.
- [4] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Contr.*, **AC-19**, 716-723.
- [5] Akaike, H. (1980). Likelihood and the Bayes procedure (with discussion). In *Bayesian Statistics* (Bernardo, J. M., De Groot, M. H., Lindley, D. U. and Smith, A. F. M. eds.), University Press, Valencia, Spain, pp. 143-166, pp. 185-203.
- [6] Arcones, M.A. and Giné, E. (1992). On the bootstrap of M -estimators and other statistical functionals. In *Exploring the Limits of Bootstrap*, (R. LePage and L. Billard eds.), Wiley, pp. 13-47.
- [7] Babu, G. J. (1992). Subsample and half-sample methods. *Ann. Inst. Statist. Math.* **44**, 703-720.
- [8] Babu, G. J. and Singh, K. (1983). Inference on means using the bootstrap. *Ann. Statist.* **11**, 999-1003.
- [9] Babu, G. J. and Singh, K. (1984). On one term Edgeworth correction by Efron's bootstrap. *Sankhyā Ser. A* **46**, 219-232.
- [10] Babu, G. J. and Singh, K. (1985). Edgeworth expansions for sampling without replacement from finite populations. *J. Multivariate Anal.* **17**, 261-278.
- [11] Banks, D. L. (1988). Histospline smoothing the Bayesian bootstrap. *Biometrika* **75**, 673-684.
- [12] Barndorff-Nielsen, O.E. and Cox, D. R. (1989). *Asymptotic Techniques for Use in Statistics*. Chapman and Hall, London.
- [13] Basawa, I. V., Mallik, A. K., McCormick, W. P. and Taylor, R. L. (1989). Bootstrapping explosive autoregressive processes. *Ann. Statist.* **17**, 1479-1486.
- [14] Beran, R. (1982). Estimated sampling distributions: The bootstrap and competitors. *Ann. Statist.* **10**, 212-225.
- [15] Beran, R. (1984a). Bootstrap methods in statistics. *Jber. Deutsch. Math.-Verein* **86**, 14-30.
- [16] Beran, R. (1984b). Jackknife approximations to bootstrap estimates. *Ann. Statist.* **12**, 101-118.
- [17] Beran, R. (1986). Simulated power functions. *Ann. Statist.* **14**, 151-173.
- [18] Beran, R. (1987). Prepivoting to reduce level error of confidence sets. *Biometrika* **74**, 457-468.
- [19] Beran, R. (1988a). Balanced simultaneous confidence sets. *J. Amer. Statist. Assoc.* **83**, 679-686.
- [20] Beran, R. (1988b). Prepivoting test statistics: A bootstrap view of asymptotic refinements. *J. Amer. Statist. Assoc.* **83**, 687-697.
- [21] Beran, R. and Ducharme, G. R. (1991). *Asymptotic Theory for Bootstrap Methods in Statistics*. Centre de Recherches Mathématiques, Université de Montréal, Canada.
- [22] Beran, R. and Millar, P. W. (1986). Confidence sets for a multivariate distribution. *Ann. Statist.*

- 14, 431-443.
- [23] Beran, R. and Srivastava, M. S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. *Ann. Statist.* **13**, 95-115.
- [24] Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6**, 434-451.
- [25] Bhattacharya, R. N. and Qumsiyeh, M. (1989). Second order and L^p -comparisons between the bootstrap and empirical Edgeworth expansion methodologies. *Ann. Statist.* **17**, 160-169.
- [26] Bhattacharya, R. N. and Rao, R. R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.
- [27] Bickel, P. J. (1992). Theoretical comparison of different bootstrap t confidence bounds. In *Exploring the Limits of Bootstrap*, (R. LePage and L. Billard eds.), Wiley, pp. 65-76.
- [28] Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9**, 1196-1217.
- [29] Bickel, P. J. and Freedman, D. A. (1983). Bootstrapping regression models with many parameters. In *A Festschrift for Erich L. Lehmann* (P. J. Bickel, K. A. Doksum, and J. C. Hodges, Jr., eds.), Wadsworth, Belmont, pp. 28-48.
- [30] Bickel, P. J. and Freedman, D. A. (1984). Asymptotic normality and the bootstrap in stratified sampling. *Ann. Statist.* **12**, 470-482.
- [31] Boos, D. D. and Brownie, C. (1989). Bootstrap methods for testing homogeneity of variances. *Technometrics* **31**, 69-82.
- [32] Bose, A. (1988). Edgeworth correction by bootstrap in autoregressions. *Ann. Statist.* **16**, 1709-1722.
- [33] Bunke, O. and Droge, B. (1984). Bootstrap and cross-validation estimators of the prediction error for linear regression models. *Ann. Statist.* **12**, 1400-1424.
- [34] Cao-Abad, R. (1991). Rate of convergence for the wild bootstrap in nonparametric regression. *Ann. Statist.* **19**, 2226-2231.
- [35] Chao, M. T. and Lo, S.-H. (1985). A bootstrap method for finite population. *Sankhyā Ser. A* **47**, 399-405.
- [36] Chatterjee, S. and Chatterjee, S. (1983). Estimation of misclassification probabilities by bootstrap methods. *Commun. Statist. B* **12**, 645-656.
- [37] Chernick, M. R., Murthy, V. K. and Nealy, C. D. (1985). Application of bootstrap and other resampling techniques: evaluation of classifier performance. *Pattern Recognition Letters* **3**, 167-178.
- [38] Chernick, M. R., Murthy, V. K. and Nealy, C. D. (1986). Correction note to 'Application of bootstrap and other resampling techniques: evaluation of classifier performance'. *Pattern Recognition Letters* **4**, 133-142.
- [39] Daniels, H. E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Statist.* **25**, 631-650.
- [40] Daniels, H. E. and Young, G. A. (1991). Saddlepoint approximation for the studentized mean, with an application to the bootstrap. *Biometrika* **78**, 169-179.
- [41] Davison, A. C. and Hall, P. (1992). On the bias and variability of bootstrap and cross-validation estimates of error rate in discriminant problems. *Biometrika* **79**, 279-284.
- [42] Davison, A. C. and Hinkley, D. V. (1988). Saddlepoint approximations in resampling methods. *Biometrika* **75**, 417-431.
- [43] Davison, A. C., Hinkley, D. V. and Schechtman, E. (1986). Efficient bootstrap simulation. *Biometrika* **73**, 555-566.
- [44] Davison, A. C., Hinkley, D. V. and Worton, B. J. (1992). Bootstrap likelihoods. *Biometrika* **79**, 113-130.
- [45] De Angelis, D. and Young, G. A. (1992). Bootstrapping the correlation coefficient: A comparison of smoothing strategies. *J. Statist. Comput. Simul.* **40**, 167-176.
- [46] Delaney, N. J. and Chatterjee, S. (1986). Use of the bootstrap and cross-validation in ridge regression. *J. Bus. Econ. Statist.* **4**, 255-262.
- [47] Diaconis, P. and Efron, B. (1983). Computer-intensive methods in statistics. *Sci. Amer.* **248**, 116-130 (松原 望訳 (1983): コンピューターがひらく新しい統計学, サイエンス, **13**, 頁 58-75).

- [48] DiCiccio, T. J. and Romano, J. P. (1988). A review of bootstrap confidence intervals. *J. R. Statist. Soc. Ser. B* **50**, 338-354.
- [49] DiCiccio, T. J. and Romano, J. P. (1989a). The automatic percentile method : accurate confidence limits in parametric models. *Canadian J. Statist.* **17**, 155-169.
- [50] DiCiccio, T. J. and Romano, J. P. (1989b). On adjustments based on the signed root of the empirical likelihood ratio statistic. *Biometrika* **76**, 447-456.
- [51] DiCiccio, T. J. and Romano, J. P. (1990). Nonparametric confidence limits by resampling methods and least favorable families. *Int. Stat. Rev.* **58**, 59-76.
- [52] DiCiccio, T. J. and Tibshirani, R. (1987). Bootstrap confidence intervals and bootstrap approximations. *J. Amer. Statist. Assoc.* **82**, 163-170.
- [53] DiCiccio, T. J. , Hall, P. and Romano, J. P. (1989). Comparison of parametric and empirical likelihood functions. *Biometrika* **76**, 465-476.
- [54] DiCiccio, T. J. , Hall, P. and Romano, J. P. (1991). Empirical likelihood is Bartlett-correctable. *Ann. Statist.* **19**, 1053-1061.
- [55] DiCiccio, T. J. , Martin, M. A. and Young, G. A. (1992). Fast and accurate approximate double bootstrap confidence intervals. *Biometrika* **79**, 285-295.
- [56] Dielman, T. E. and Pfaffenberger, R. C. (1988). Bootstrapping in least absolute value regression : An application to hypothesis testing. *Commun. Statist.- Simul. Comput.* **17**, 843-856.
- [57] Dikta, G. (1990). Bootstrap approximation of nearest neighbor regression function estimates. *J. Mult. Anal.* **32**, 213-229.
- [58] Do, K.-A. (1992). A simulation study of balanced and antithetic bootstrap resampling methods. *J. Statist. Comput. Simul.* **40**, 153-166.
- [59] Do, K.-A. and Hall, P. (1991). On importance resampling for the bootstrap. *Biometrika* **78**, 161-167.
- [60] Do, K.-A. and Hall, P. (1992). Distribution estimation using concomitants of order statistics, with application to Monte Carlo simulation for the bootstrap. *J. R. Statist. Soc. B* **54**, 595-607.
- [61] Ducharme, G. R., Jhun, M., Romano, J.P., and Truong, K. N. (1985). Bootstrap confidence cones for directional data. *Biometrika* **72**, 637-645.
- [62] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7**, 1-26.
- [63] Efron, B. (1981). Nonparametric standard errors and confidence intervals (with discussion). *Canad. J. Statist.* **9**, 139-172.
- [64] Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM, Philadelphia.
- [65] Efron, B. (1983). Estimating the error rate of a prediction rule : Improvement on cross-validation. *J. Amer. Statist. Assoc.* **78**, 316-331.
- [66] Efron, B. (1985). Bootstrap confidence intervals for a class of parametric problems. *Biometrika* **72**, 45-58.
- [67] Efron, B. (1986). How biased is the apparent error rate of a prediction rule?. *J. Amer. Statist. Assoc.* **81**, 461-470.
- [68] Efron, B. (1987). Better bootstrap confidence intervals (with discussion). *J. Amer. Statist. Assoc.* **82**, 171-200.
- [69] Efron, B. (1990). More efficient bootstrap computations. *J. Amer. Statist. Assoc.* **85**, 79-89.
- [70] Efron, B. (1992a). Six questions raised by the bootstrap. In *Exploring the Limits of Bootstrap*, (R. LePage and L. Billard eds.), Wiley, pp. 99-126.
- [71] Efron, B. (1992b). Jackknife-after-bootstrap standard errors and influence functions (with discussion). *J. R. Statist. Soc. B* **54**, 83-127.
- [72] Efron, B. and Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *Amer. Statist.* **37**, 36-48.
- [73] Efron, B. and Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* **1**, 54-77.
- [74] Faraway, J. J. (1990). Bootstrap selection of bandwidth and confidence bands for nonparametric regression. *J. Statist. Comput. Simul.* **37**, 37-44.
- [75] Faraway, J. J. and Jhun, M. (1990). Bootstrap choice of bandwidth for density estimation. *J. Amer. Statist. Assoc.* **85**, 1119-1122.

- [76] Findley, D. F. (1986). On bootstrap estimates of forecast mean square errors for autoregressive processes. In *Computer Science and Statistics: The Interface* (D. M. Allen ed.), Elsevier Science Pub., pp. 11-17.
- [77] Fisher, N. I. and Hall, P. (1989a). Bootstrap confidence regions for directional data. *J. Amer. Statist. Assoc.* **84**, 996-1002.
- [78] Fisher, N. I. and Hall, P. (1989b). Bootstrap methods for directional data. In *Proceedings on Symposium on the Analysis of Statistical Information*, Institute of Statistical Mathematics, Tokyo, pp. 275-299.
- [79] Fitzmaurice, G. M., Krzanowski, W. J. and Hand, D. J. (1991). A Monte Carlo study of the 632 bootstrap estimator. *J. of Classification* **8**, 239-250.
- [80] Franke, J. and Härdle, W. (1992). On bootstrapping kernel spectral estimates. *Ann. Statist.* **20**, 121-145.
- [81] Freedman, D. A. (1981). Bootstrapping regression models. *Ann. Statist.* **9**, 1218-1228.
- [82] Freedman, D. A. (1984). On bootstrapping two-stage least-squares estimates in stationary linear models. *Ann. Statist.* **12**, 827-842.
- [83] Freedman, D. A. and Peters, S. C. (1984a). Bootstrapping a regression equation: Some empirical results. *J. Amer. Statist. Assoc.* **79**, 97-106.
- [84] Freedman, D. A. and Peters, S. C. (1984b). Bootstrapping an econometric model: Some empirical results. *J. Bus. Econ. Statist.* **2**, 150-158.
- [85] Ganeshanandam, S. and Krzanowski, W. J. (1990). Error-rate estimation in two-group discriminant analysis using the linear discriminant function. *J. Statist. Comput. Simul.* **36**, 157-175.
- [86] Geisser, S. (1975). The predictive sample reuse method with applications. *J. Amer. Statist. Assoc.* **70**, 320-328.
- [87] Gleason, J. R. (1988). Algorithms for balanced bootstrap simulations. *Amer. Statist.* **42**, 263-266.
- [88] Gong, G. (1986). Cross-validation, the jackknife, and the bootstrap: Excess error estimation in forward logistic regression. *J. Amer. Statist. Assoc.* **81**, 108-113.
- [89] Graham, R. L., Hinkley, D. V., John, P. W. M., and Shi, S. (1990). Balanced design of bootstrap simulations. *J. R. Statist. Soc. Ser. B* **52**, 185-202.
- [90] Hall, P. (1983). Inverting an Edgeworth expansion. *Ann. Statist.* **11**, 569-576.
- [91] Hall, P. (1986a). On the number of bootstrap simulations required to construct a confidence interval. *Ann. Statist.* **14**, 1453-1462.
- [92] Hall, P. (1986b). On the bootstrap and confidence intervals. *Ann. Statist.* **14**, 1431-1452.
- [93] Hall, P. (1987a). On the bootstrap and continuity correction. *J. R. Statist. Soc. Ser. B* **49**, 82-89.
- [94] Hall, P. (1987b). On the bootstrap and likelihood-based confidence regions. *Biometrika* **74**, 481-493.
- [95] Hall, P. (1988a). Theoretical comparison of bootstrap confidence intervals (with discussion). *Ann. Statist.* **16**, 927-985.
- [96] Hall, P. (1988b). On symmetric bootstrap confidence intervals. *J. R. Statist. Soc. Ser. B* **50**, 35-45.
- [97] Hall, P. (1989a). Unusual properties of bootstrap confidence intervals in regression problems. *Probab. Th. Fields* **81**, 247-273.
- [98] Hall, P. (1989b). On efficient bootstrap simulation. *Biometrika* **76**, 613-617.
- [99] Hall, P. (1989c). Antithetic resampling for the bootstrap. *Biometrika* **76**, 713-724.
- [100] Hall, P. (1990a). Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *J. Multivariate. Anal.* **32**, 177-203.
- [101] Hall, P. (1990b). On the relative performance of bootstrap and Edgeworth approximations of a distribution function. *J. Multi. Anal.* **35**, 108-129.
- [102] Hall, P. (1990c). Pseudo-likelihood theory for empirical likelihood. *Ann. Statist.* **18**, 121-140.
- [103] Hall, P. (1990d). Performance of balanced bootstrap resampling in distribution function and quantile problems. *Probab. Theory Rel. Fields.* **85**, 239-260.
- [104] Hall, P. (1992a). *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New York.
- [105] Hall, P. (1992b). On the removal of skewness by transformation. *J. R. Statist. Soc. Ser. B*, **54**, 221-228.
- [106] Hall, P. (1992c). Efficient bootstrap simulation. In *Exploring the Limits of Bootstrap*, (R. LePage

- and L. Billard eds.), Wiley, pp. 127-143.
- [107] Hall, P. and Hart, J. D. (1990). Bootstrap test for difference between means in nonparametric regression. *J. Amer. Statist. Assoc.* **85**, 1039-1049.
- [108] Hall, P. and Martin, M. A. (1988). On bootstrap resampling and iteration. *Biometrika* **75**, 661-671.
- [109] Hall, P. and Pittelkow, Y. E. (1990). Simultaneous bootstrap confidence bands in regression. *J. Statist. Comput. Simul.* **37**, 99-113.
- [110] Hall, P. and Scala, B. L. (1990). Methodology and algorithms of empirical likelihood. *Int. Stat. Rev.* **58**, 109-127.
- [111] Hall, P. and Wilson, S. R. (1991). Two guidelines for bootstrap hypothesis testing. *Biometrics* **47**, 757-762.
- [112] Hall, P., DiCiccio, T. J., and Romano, J. P. (1989). On smoothing and the bootstrap. *Ann. Statist.* **17**, 692-704.
- [113] Hammersley, I. M. and Handscomb, D. C. (1964). *Monte Carlo Methods*. Wiley, New York.
- [114] Hampel, F. R., Rousseeuw, P. J., Ronchetti, E. M. and Stahel, W. A., (1986). *Robust Statistics (The approach based on influence functions)*, Wiley, New York.
- [115] Hand, D. J. (1986). Recent advances in error rate estimation. *Pattern Recognition Letters* **4**, 335-346.
- [116] Härdle, W. and Bowman, A. W. (1988). Bootstrapping in nonparametric regression: Local adaptive smoothing and confidence bands. *J. Amer. Statist. Assoc.* **83**, 102-110.
- [117] Härdle, W. and Marron, J. S. (1991). Bootstrap simultaneous error bars for nonparametric regression. *Ann. Statist.* **19**, 778-796.
- [118] Härdle, W., Hall, P. and Marron, J. S. (1988). How far are automatically chosen regression smoothing parameters from their optimum? *J. Amer. Statist. Assoc.* **83**, 86-101.
- [119] Hartigan, J. A. (1969). Using subsample values as typical values. *J. Amer. Statist. Assoc.* **64**, 1303-1317.
- [120] Hartigan, J. A. (1971). Error analysis by replaced samples. *J. R. Statist. Soc. Ser. B* **33**, 98-110.
- [121] Hartigan, J. A. (1975). Necessary and sufficient conditions for asymptotic joint normality of a statistic and its subsample values. *Ann. Statist.* **3**, 573-580.
- [122] Hinkley, D. V. (1988). Bootstrap methods. *J. R. Statist. Soc. Ser. B* **50**, 321-337.
- [123] Hinkley, D. V. and Shi, S. (1989). Importance sampling and the nested bootstrap. *Biometrika* **76**, 435-446.
- [124] Hinkley, D. V. and Wei, B.-C. (1984). Improvements of jackknife confidence limit methods. *Biometrika* **71**, 331-339.
- [125] Honda, M. and Konishi, S. (1988). Comparison of variable selection procedures in discriminant analysis under nonnormal populations, In *Statistical Theory and Data Analysis II* (K. Matusita, ed.), North-Holland, Amsterdam, 337-347.
- [126] Ichikawa, M. and Konishi, S. (1992). Application of the bootstrap methods in factor analysis. Research Memorandum No. 457, Institute of Statistical Mathematics, Tokyo.
- [127] Ishiguro, M. and Sakamoto, Y. (1991). WIC: An Estimator-free information criterion. Research Memorandum No. 410, Institute of Statistical Mathematics, Tokyo.
- [128] 石黒真木夫, 坂元慶行, 北川源四郎(1992). ベイズモデルとEIC. 第60回日本統計学会講演予稿集, 頁264-266.
- [129] Jaeckel, L. (1972). The infinitesimal Jackknife. Bell Lab Memorandum MM 72-1215-11.
- [130] Jain, A. K., Dubes, R. C. and Chen, C.-C. (1987). Bootstrap techniques for error estimation. *IEEE Trans. Patt. Anal. Math. Intel.* **PAMI-9**, 628-633.
- [131] Johns, V. (1988). Importance sampling for bootstrap confidence intervals. *J. Amer. Statist. Assoc.* **83**, 709-714.
- [132] 北川源四郎(1991). 対数尤度のブートストラップについて. 統数研共同研究レポート **31**, 時系列に関する推測の理論と応用, 175-179.
- [133] 北川源四郎, 石黒真木夫, 坂元慶行 (1992). EICによるモデルの予測評価. 第60回日本統計学会講演予稿集, 頁258-260.
- [134] Kipnis, V. (1992). Bootstrap assessment of prediction in exploratory regression analysis. In *Exploring the Limits of Bootstrap*, (R. LePage and L. Billard eds.), Wiley, pp. 363-387.

- [135] Konishi, S. (1981). Normalizing transformations of some statistics in multivariate analysis. *Biometrika* **68**, 647-651.
- [136] Konishi, S. (1987). Transformations of statistics in multivariate analysis. In *Advances in Multivariate Statistical Analysis* (Ed. A. K. Gupta). D. Reidel, Dordrecht, 213-231.
- [137] 小西貞則(1988). ブートストラップ法による推定量の誤差評価, 赤池弘次監修: パソコンによるデータ解析, 朝倉書店, 頁123-142.
- [138] 小西貞則(1990). ブートストラップ法と信頼区間の構成, 応用統計学 **19**, 137-162.
- [139] Konishi, S. (1991). Normalizing transformations and bootstrap confidence intervals. *Ann. Statist.* **19**, 2209-2225.
- [140] 小西貞則(1992a). 対数尤度のバイアス補正と情報量規準, 第60回日本統計学会講演予稿集, 頁39-40.
- [141] 小西貞則(1992b). ブートストラップ法と予測誤差推定, 東京大学統計学輪講資料.
- [142] Konishi, S. and Honda, M. (1990). Comparison of procedures for estimation of error rates in discriminant analysis under nonnormal populations. *J. Statist. Comput. Simul.* **36**, 105-115.
- [143] 小西貞則, 本多正幸(1992). 判別分析における誤判別率推定とブートストラップ法, 応用統計学 **21**, 67-100.
- [144] Kotz, S. and Johnson, N. L. (eds.) (1992). *Breakthroughs in Statistics*, Vol. I: Foundations and Basic Theory, Vol. II: Methodology and Distribution, Springer-Verlag, New York.
- [145] Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17**, 1217-1241.
- [146] Lachenbruch, P. A. and Mickey, M. R. (1968). Estimation of error rates in discriminant analysis. *Technometrics* **10**, 1-11.
- [147] Lahiri, S.N. (1992). Bootstrapping M -estimators of a multiple linear regression parameter. *Ann. Statist.* **20**, 1548-1570.
- [148] Léger, C., Politis, D. N. and Romano, J. P. (1992). Bootstrap technology and applications. *Technometrics* **34**, 378-398.
- [149] Lo, A. Y. (1987). A large sample study of the Bayesian bootstrap. *Ann. Statist.* **15**, 360-375.
- [150] Lo, A. Y. (1988). A Bayesian bootstrap for a finite population. *Ann. Statist.* **16**, 1684-1695.
- [151] Loh, W.-Y. (1987). Calibrating confidence coefficients. *J. Amer. Statist. Assoc.*, **82**, 155-162.
- [152] Loh, W.-Y. (1991). Bootstrap calibration for confidence interval construction and selection. *Statistica Sinica* **1**, 477-491.
- [153] Mammen, E. (1992). *When Does Bootstrap Work? Asymptotic Results and Simulations*. Springer-Verlag, New York.
- [154] Martin, M.A. (1990). On bootstrap iteration for coverage correction in confidence intervals. *J. Amer. Statist. Assoc.* **85**, 1105-1118.
- [155] Marron, J.S. (1992). Bootstrap bandwidth selection. In *Exploring the Limits of Bootstrap*, (R. LePage and L. Billard eds.), Wiley, pp. 249-262.
- [156] McLachlan, G. J. (1980). The efficiency of Efron's "bootstrap" approach applied to error rate estimation in discriminant analysis. *J. Statist. Comput. Simul.* **11**, 273-279.
- [157] McLachlan, G. J. (1986). Assessing the performance of an allocation rule. *Comp. & Maths. with Appls.* **12A**, 261-272.
- [158] McLachlan, G. J. (1987). Error rate estimation in discriminant analysis: recent advances. In *Advances in Multivariate Statistical Analysis* (A. K. Gupta, ed.), D. Reidel, 233-252.
- [159] McLachlan, G. J. (1992). *Discriminant Analysis and Statistical Pattern Recognition*. John Wiley, New York.
- [160] Miller, R. (1974). The jackknife-a review. *Biometrika* **61**, 1-17.
- [161] Nagao, H. and Srivastava, M. S. (1992). On the distributions of some test criteria for a covariance matrix under local alternatives and bootstrap approximations. *J. Multivariate Anal.* **43**, 331-350.
- [162] Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237-249.
- [163] Owen, A. B. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18**, 90-120.
- [164] Owen, A. B. (1991). Empirical likelihood for linear models. *Ann. Statist.* **19**, 1725-1747.
- [165] Parr, W.C. (1983). A note on the jackknife, the bootstrap and the delta method estimators of bias and variance. *Biometrika* **70**, 719-722.

- [166] Peers, H. W. and Iqbal, M. (1985). Asymptotic expansions for confidence limits in the presence of nuisance parameters, with applications. *J. R. Statist. Soc. Ser. B* **47**, 547-554.
- [167] Peters, S. C. and Freedman, D.A. (1984). Some notes on the bootstrap in regression problems. *J. Bus. Econ. Statist.* **2**, 406-409.
- [168] Petrov, V. V. (1975). *Sums of Independent Random Variables*. Springer-Verlag, Berlin.
- [169] Pfanzagl, G. (1985). *Asymptotic Expansions for General Statistical Models*. Lecture Notes in Statistics **31**, Springer, Berlin.
- [170] Quenouille, M. H. (1949). Approximate tests of correlation in time series. *J. R. Statist. Soc. Ser. B* **11**, 18-84.
- [171] Quenouille, M. H. (1956). Notes on bias in estimation. *Biometrika* **43**, 353-360.
- [172] Rao, J. N. K. and Wu, C. F. J. (1988). Resampling inference with complex survey data. *J. Amer. Statist. Assoc.* **83**, 231-241.
- [173] Reeds, J. A. (1976). On the Definition of von Mises Functionals. Ph. D. thesis, Harvard University.
- [174] Reid, N. (1988). Saddlepoint methods and statistical inference (with discussion). *Statist. Sci.* **3**, 213-238.
- [175] Robinson, J. (1987). Nonparametric confidence intervals in regression: The bootstrap and randomization methods. In *New Perspectives in Theoretical and Applied Statistics* (M. L. Puri, J. P. Vilaplana, and W. Wertz, eds.), Wiley, New York, pp. 243-255.
- [176] Romano, J. P. (1988a). On weak convergence and optimality of kernel density estimates of the mode. *Ann. Statist.* **16**, 629-647.
- [177] Romano, J. P. (1988b). Bootstrapping the mode. *Ann. Inst. Statist. Math.* **40**, 565-586.
- [178] Romano, J. P. (1988c). A bootstrap revival of some nonparametric distance tests. *J. Amer. Statist. Assoc.* **83**, 698-708.
- [179] Rubin, D. B. (1981). The Bayesian bootstrap. *Ann. Statist.* **9**, 130-134.
- [180] 坂元慶行, 石黒真木夫, 北川源四郎 (1983). 情報量統計学. 共立出版.
- [181] 坂元慶行, 石黒真木夫, 北川源四郎 (1992). ABIC 最小化法と EIC. 第 60 回日本統計学会講演予稿集, 頁 261-263.
- [182] Sánchez, J. M. P. and Cepeda, X. L. O. (1989). The use of smooth bootstrap techniques for estimating the error rate of a prediction rule. *Commun. Statist.-Simula. Comput.* **18**, 1169-1186.
- [183] Sen, P. K. (1988). Functional jackknifing: Rationality and general asymptotics. *Ann. Statist.* **16**, 450-469.
- [184] Shao, J. (1988). On resampling methods for variance and bias estimation in linear models. *Ann. Statist.* **16**, 986-1008.
- [185] Shibata, R. (1989). Statistical aspects of model selection. In *From Data to Model*, (J. C. Willems ed.), Springer-Verlag, pp. 215-240.
- [186] 清水良一. (1976). 中心極限定理. 教育出版.
- [187] Shorack, G.R. (1982). Bootstrapping robust regression. *Commun. Statist.—Theor. Meth.* **11**, 961-972.
- [188] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.
- [189] Silverman, B. W. and Young, G. A. (1987). The bootstrap: To smooth or not to smooth?. *Biometrika* **74**, 469-479.
- [190] Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9**, 1187-1195.
- [191] Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook*. American Sciences Press, Columbus.
- [192] Sitter, R.R. (1992). A resampling procedure for complex survey data. *J. Amer. Statist. Assoc.* **87**, 755-765.
- [193] Snapinn, S. M. and Knoke, J. D. (1988). Bootstrapped and smoothed classification error rate estimators. *Commun. Statist.-Simula. Comput.* **17**, 1135-1153.
- [194] Snapinn, S. M. and Knoke, J. D. (1989). Estimation of error rates in discriminant analysis with selection of variables. *Biometrics* **45**, 289-299.
- [195] Stangenhuis, G. (1987). Bootstrap and inference procedures for L_1 regression. In *Statistical Data Analysis Based on the L_1 -Norm and Related Methods* (Y. Dodge ed.); Elsevier Science Publishers

- B. V. (North-Holland), pp. 323-332.
- [196] Stine, R. A. (1985). Bootstrap prediction intervals for regression. *J. Amer. Statist. Assoc.* **80**, 1026-1031.
- [197] Stoffer, D. S. and Wall, K.D. (1991). Bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter. *J. Amer. Statist. Assoc.* **86**, 1024-1033.
- [198] Stone, M. (1974). Cross-validation choice and assessment of statistical predictions. *J. R. Statist. Soc. B* **36**, 111-147.
- [199] Swanepoel, J. W. H. (1990). A review of bootstrap methods. *South African Statist. J.* **24**, 1-34.
- [200] Swanepoel, J. W. H. and van Wyk, J. W. J. (1986). The bootstrap applied to power spectral density function estimation. *Biometrika* **73**, 135-141.
- [201] 高橋 一 (1985). Bootstrap 推定量の多項式近似-ノンパラメトリック統計学の話題から-。一橋論叢, **94**, 44-65.
- [202] Takahashi, H. (1988). A note on Edgeworth expansions for the von Mises functionals. *J. Multivariate Anal.* **24**, 56-65.
- [203] 竹内 啓 (1975). 確率分布の近似。教育出版。
- [204] 竹内 啓 (1976). 情報統計量の分布とモデルの適切さの規準。数理科学 **153**, 12-18.
- [205] Taylor, C.C. (1989). Bootstrap choice of the smoothing parameter in kernel density estimation. *Biometrika* **76**, 705-712.
- [206] Tibshirani, R. (1988). Variance stabilization and the bootstrap. *Biometrika* **75**, 433-444.
- [207] Tukey, J. (1958). Bias and confidence in not quite large samples. Abstract, *Ann. Math. Statist.* **29**, 614.
- [208] von Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* **18**, 309-348.
- [209] Wang, M-C. (1986). Re-sampling procedures for reducing bias of error rate estimation in multinomial classification. *Comput. Statist. and Data Analysis* **4**, 15-39.
- [210] Weber, N.C. (1984). On resampling techniques for regression models. *Statist. Probab. Lett.* **2**, 275-278.
- [211] Wernecke, K.-D. and Kalb, G. (1987). Estimation of error rates by means of simulated bootstrap distributions. *Biom. J.* **29**, 287-297.
- [212] Withers, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals. *Ann. Statist.* **11**, 577-587.
- [213] Withers, C. S. (1984). Asymptotic expansions for distributions and quantiles with power series cumulants. *J. R. Statist. Soc. Ser. B* **46**, 389-396.
- [214] Wong, W.-H. (1983). A note on the modified likelihood for density estimation. *J. Amer. Statist. Assoc.* **78**, 461-463.
- [215] Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis (with discussion). *Ann. Statist.* **14**, 1261-1350.
- [216] Young, G. A. (1988). A note on bootstrapping the correlation coefficient. *Biometrika* **75**, 370-373.
- [217] Zhang, J. and Boos, D. D. (1992). Bootstrap critical values for testing homogeneity of covariance matrices. *J. Amer. Statist. Assoc.* **87**, 425-429.