95-J-6

縮小推定の理論と応用

久保川達也
東京大学経済学部

1995年6月

縮小推定の理論と応用

東京大学 経済学部　久保川達也

1. はじめに

統計的推測理論の研究における醍醐味は, 推測問題の数理的な'からくり'を様々な側面から解きほぐしてその推測問題の構造を明らかにすることと, 実際の使用に耐えうる簡便で有効な推測方法を提示することであろう. 推定理論において, 最尤推定量 (MLE) についての研究はそのようなものの一つである. 大標本理論での推定問題の構造が明らかになり MLE の高次漸近有効性とその微分幾何学的解釈が与えられたことは, 近年の顕著な発展の成果として高く評価される. また MLE 導出のための尤度方程式の解法は原理的には計算機によって計算させうるものであり, 従って MLE は実務家の使用に耐えうる推定方法である.

一方, 医薬生物学や工学などの分野においては, 必ずしも大きな標本が望めない場合がある. このときには事前情報の使用や通常は母数推定に組み込まない統計量を取り込むなどして情報の回復をはかり, できるだけ有効な推定手法を開発することが試みられるであろう. 小標本のこうした問題への理論的研究は統計的決定理論の枠組みの中で議論されてきた. 統計的決定理論は Wald(1950) に始まり, 不偏推定に関する Rao-Blackwell の定理, 代数的構造が埋め込まれた推定問題における最良共変推定量のミニマクス性, ミニマクス性と許容性についての一般的条件の導出など, 基本的かつ一般的な理論が 1950 年代 60 年代に構築された (竹内 (1963,65), 鍋谷 (1978),Zacks(1971),Lehmann(1983) を参照). それ以降は, おもに各推定問題での理論的解明という個別的な研究に関心が向けられ, 特に MLE, 不偏推定量など通常の推定量が許容的でないような推定問題における縮小推定の有効性に関する研究が展開され, その理論的深化と応用的広がりがはかられてきた (Berger(1985),Robert(1994),Rukhin(1995) を参照). その代表的なものが平均ベクトルの Stein 現象に関する研究であろう. p-次元正規分布の平均の同時推定において, $p \geq 3$ に対して MLE が非許容的であることを Stein(1956) が発見し, 実際改良している縮小推定量を James-Stein(1961) が導出して以来, 不可思議な現象として多くの理論家の興味を引きその理論的解明と展開がなされてきた. 特に部分積分である Stein identity の導入によって一層著しい発展を遂げた. 一方で Efron-Morris(1972) によって James-Stein 推定量の経験 Bayes 性が示されてからは, 経験 Bayes 推定の形での応用例への適用が考えられ, Efron-Morris(1975) の野球選手の打撃能力の推定やエクサルバドルにおけるトキソプラズマ症の発生率の推定への応用, Fay-Herriot(1979), Battese-Harter-Fuller(1988), Prasad-Rao(1990), Ghosh-Rao(1994) らの small area problem への応用, Wahba(1985), Li(1985), Li-Hwang(1984), Yanagimoto-Yanagimoto(1987), Ansley-Kohn-Wong(1993) らのスプライン関数によるデータの平滑化への応用, Tsutakawa-Shoop-Marienfeld(1985), 丹後 (1988) らの死亡指標の推定への応用, Copas(1983) の線形予測問題への応用など縮小の概念の適用が様々に模索されてきた. また縮小推定の実用的成功例が Morris(1983), Casella(1985) にも紹介されている. このように Stein 型縮小推定は, 理論的研究の深化が図られてきたとともに, 実用的視点からも有効な推定手法として評価されてきたものといえる.

James-Stein 型の縮小推定や経験 Bayes 推定に限らず, 広い意味での縮小もしくは修正を必要とする問題が少なからず存在する (広津 (1992)). 例えば, 5 節及び 6 節に登場する問題が

挙げられる. 変量, 混合線形モデルにおける分散の郡間成分の推定に対しては, その代表的な不偏推定量である ANOVA, MINQU 推定量は負値を取り得てしまう. この不合理な点を解決するためには不偏推定量を正の方向へ縮小 (修正) しなければならないであろう. また品質管理の分野に登場する SN 比は非心パラメータの関数であるが, その非心パラメータの推定においても同様な問題が生じ不偏推定量の負の部分を正の方向へ縮小する必要がある. 線形回帰モデルにおいて説明変数を逆予測する, いわゆる線形校正は, 化学実験, 農事試験, 計量経済の分野において登場する問題であるが, 通常の古典的推定量 (MLE) はモーメントをもたないことが知られている. この場合も古典的推定量の不安定な部分を打ち切るなどの縮小が望まれる. その他, 順序制約などなんらかの制約が母数空間に課せられている場合には, 推定量がその母数空間からはみ出すときにはその空間の方へ縮小する必要がある. また母集団選抜後や検定後の推定問題においては通常の推定量は明らかに偏りが生じてしまうのでなんらかの縮小を要する. こうした縮小という考え方は, 田口の割引推定のように実際の現場においても受け入れられてきたことである. こうしてみると, 一般に, 推定が過大評価または過小評価している場合や, 推定が曖昧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もしくは修正することが望ましく, こうした縮小・修正を要する場面はデータ解析の現場において少なからず存在すると思われる.

縮小を要する推定問題において, どの程度縮小すべきかの指標を与えることは大事な問題である. 交差検証法 (Cross Validation) は応用的に有用な一つの方法であり, 縮小の程度は予測誤差が最小になるように求められる. 階層 Bayes 推定や 経験 Bayes 推定も一つの実用的方法であり, Cramer-Rao の下限に対して Bayes 解の導出を行っている椿 (1985) の方法も興味深い. 一方 決定理論の立場からこの問題をながめると, 通常の推定量を改良しうる程度の縮小を定めることは意味のあることである. その一つの有効な方法は, 縮小推定量の危険関数の不偏推定量を導くアプローチで, Stein identity と呼ばれる部分積分のアイデアに基づいている. この方法は, 平均ベクトルの推定の理論の顕著な発展をもたらしたのであるが, 実際 MLE を改良するための縮小の程度を与えることができる. この部分積分のアイデアはカイ二乗分布や Wishart 分布の identity をもたらし, 共分散行列や平均行列の推定構造を明らかにする上で, 強力な武器として大いに役立ち, 行列の縮小推定についての理論的展開が活発になされたのである. しかしこうした identity の使用には限界があった. 例えば, 平均ベクトルの同時推定に対しては, MLE を改良する推定量のクラスを構成することができるが, さらに一歩進んで James-Stein 推定量を改良するための縮小の程度を明らかにしようとするとうまくいかないことに気がつく. 同様な問題が平均未知の正規分布の分散を推定する問題にも生ずる. 標本分散だけに基づいた通常の推定量が, 標本平均に含まれる情報を用いた Stein 型打ち切り推定量によって改良されることが知られているが, この問題に identity を適用しても改良するための縮小の程度を明確にすることはできない.

こうした中で, Kubokawa(1994a), Takeuchi(1991) は IERD(Integral Expression of Risk Difference) 法を提案した. これは簡単な定積分により危険関数の差を積分表現し変数変換を行うことによって改良する推定量のクラスを明らかにするものである. IERD 法によって, 分散の推定に対しては Stein 型打ち切り推定量や Brewster-Zidek の一般化 Bayes 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. また平均ベクトルの推定に対しても, James-Stein 推定量を改良する推定量のクラスで, positive-part Stein 推定量や Strawderman-Berger の許

容的推定量を含んだものを構成できる．こうして IERD 法を用いることにより identity の限界を超えた結論を導くことができるとともに，別々の方法で得られてきたものが統一的にしかも簡便な方法によって導くことができるようになる．さらに上で挙げた縮小・修正を要する様々な応用問題においても適用可能でそれぞれ望まれる結論が得られる．

本報告では，IERD 法の使用を中心にしながらも，より広い視野から縮小推定の理論と応用を概説することを心がけた．2 節では，分散の推定問題に関する理論的展開を概観し，縮小推定のための Stein 法，Brown-Brewster-Zidek 法，そしてそれらを統一した新たな IERD 法について述べる．また分散比の二重縮小推定についても概説する．3 節では平均ベクトルの同時推定における Stein 現象について概観し，正規及び非正規分布での様々な研究成果，IERD 法による James-Stein 推定量の改良，positive-part Stein 推定量の改良についての最近の結果，信頼集合についての一連の結果などについて述べる．4 節では共分散行列の縮小推定について概観する．その他の縮小もしくは修正を要する推定問題が 5 節でとりあげられ，変量線形モデルにおける分散成分の推定，SN 比あるいは非心母数の推定，線形校正問題や統計的制御問題などにおいては IERD 法の使用によって合理的で優れた縮小推定量を求めることができることを述べる．また制約された母数空間での推定問題，選抜後の推定問題等についてもふれる．また，多重共線性，帰無仮説の方向への縮小，判別分析，誤判別率の推定，不良率や信頼度の推定，分位点の推定，因子分析における不適解など，幅広い意味で縮小もしくは修正を必要とする推定問題について述べる．6 節では，経験 Bayes 推定の応用例として，打撃能力の推定，死亡率及び死亡指標の推定，small-area の推定問題，データの平滑化法について述べる．この報告では，混乱しない限り同じ記号が使われているので注意されたい．

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(1994a), Takeuchi(1991) は新たな統一的方法として IERD 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，IERD 法がどのようなもので今までの方法とはどのような違いがあるのかについて説明する．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．

2.1. 推定問題の定式化．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

$$S/\sigma^2 \sim \chi_n^2, \quad X \sim \mathcal{N}_p(\theta, \sigma^2 I_p) \tag{2.1}$$

なる分布に従うとする．興味ある母数は σ^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ここでは Kullback-Leibler 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. Kullback-Leibler 損失は, 一般に, 確率変数 X が密度関数 $f(x;\omega)$ に従っていて δ で ω を推定するとき, 2 つの分布 $f(x;\delta)$ と $f(x;\omega)$ の間の Kullback-Leibler の距離

$$\int \log\left\{\frac{f(x;\delta)}{f(x;\omega)}\right\} f(x;\delta)dx \tag{2.2}$$

に基づいた損失関数である.

σ^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 $O(p)$ で表わすとき, アフィン変換群 $S \to c^2 S$, $X \to c\Gamma X + d$, $\sigma^2 \to c^2\sigma^2$, $\theta \to c\Gamma\theta + d$, $c \in \mathbf{R}$, $d \in \mathbf{R}^p$, $\Gamma \in O(p)$, に関して σ^2 の推定問題を不変にするために, $\delta(c^2S, cX + d) = c^2\delta(S, X)$ 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, それが δ_0 で与えられる.

自然な推定量 δ_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 $S \to c^2S$, $X \to c\Gamma X$, $\sigma^2 \to c^2\sigma^2$, $\theta \to c\Gamma\theta$ に関して共変な推定量のクラス

$$\delta_\phi = S\phi(W), \quad W = ||X||^2/S$$

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, δ_ϕ の中に最良な推定量は存在しないが, δ_0 を改良するものを見つけることは可能である. 実際次の 2 つの方法が知られてきた.

2.2. Stein 法. Stein(1964) は $W = ||X||^2/S$ 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 $\phi_\lambda^*(W)$ を求めることを考えた. $u = ||X||^2/\sigma^2$, $v = S/\sigma^2$ とおきそれらの密度関数をそれぞれ $f_p(u;\lambda)$, $f_n(v)$ とおくと,

$$\begin{aligned}\phi_\lambda^*(w) &= \frac{1}{E[S/\sigma^2|W = w]} \\ &= \frac{\int vf_p(vw;\lambda)f_n(v)dv}{\int v^2f_p(vw;\lambda)f_n(v)dv} \\ &\le \frac{\int vf_p(vw;0)f_n(v)dv}{\int v^2f_p(vw;0)f_n(v)dv} = \phi_0^*(w)\end{aligned} \tag{2.3}$$

が成り立つ. 上の不等式は $f_p(u;\lambda)/f_p(u;0)$ が u に関して単調増加するという性質を用いて示される. さらに $\phi_0^*(w) = (1+w)/(n+p)$ となり,

$$\phi_T(W) = \min\left\{\frac{1}{n}, \frac{1+W}{n+p}\right\} \tag{2.4}$$

とおくと $\phi_\lambda^*(W) \le \phi_T(W) \le n^{-1}$ なる関係が成り立つことがわかる. 従って図 1 からわかるように, 損失関数の凸性から $E_\lambda[L(\phi_T(W)S/\sigma^2)|W] \le E_\lambda[L(n^{-1}S/\sigma^2)|W]$ が成り立ち, 推定量

$$\delta^{ST} = \delta_{\phi_T} = \min\left\{\frac{S}{n}, \frac{S + ||X||^2}{n+p}\right\} \tag{2.5}$$

は δ_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．

δ^{ST} は仮説 $H: \theta = 0$ vs. $K: \theta \neq 0$ に関して，H が受容されるときには $(S+||X||^2)/(n+p)$，棄却されるときには δ_0 をとるという予備検定推定量になっている．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，θ が 0 に近いときに顕著に現れることがわかる．δ^{ST} の危険関数の概形が図 2 に描かれている．

また Kubokawa-Robert-Saleh(1992) が示しているように，δ^{ST} は，Bayes の観点から経験 Bayes 推定量として自然に導かれる．実際 $\eta = 1/\sigma^2$ とおくとき，η を与えたときの θ の条件付き分布として未知超母数 a と予め与えられた既知のベクトル θ_0 に対して

$$\theta|\eta \sim \mathcal{N}_p(\theta_0, \frac{1}{a\eta}I_p)$$

を仮定し，η については無情報分布 $\eta^{-1}d\eta$ を仮定する．この事前分布に対する σ^2 の Bayes 推定量は $\tau = a/(1+a)$ に対して

$$\hat{\sigma}^2_B(\theta_0) = \frac{1}{E[\eta|X,S]} = \frac{\tau||X-\theta_0||^2+S}{n+p}$$

となる．ここで τ は未知だから (X,S) の周辺分布から推定しなければならない．周辺分布は

$$s^{n/2-1}\frac{\tau^{p/2}}{(\tau||x-\theta_0||^2+s)^{(n+p)/2}}$$

の定数倍で与えられるから，$0<\tau<1$ に注意すると τ の最尤推定量は

$$\hat{\tau} = \min\left\{\frac{pS}{n||X-\theta_0||^2},\ 1\right\}$$

で与えられる．この推定量 $\hat{\tau}$ を $\hat{\sigma}^2_B(\theta_0)$ の τ に代入したものが経験 Bayes 推定量 $\hat{\sigma}^2_{EB}(\theta_0)$ であり，その結果

$$\hat{\sigma}^2_{EB}(\theta_0) = \frac{\hat{\tau}||X-\theta_0||^2+S}{n+p} = \min\left\{\frac{S}{n},\ \frac{||X-\theta_0||^2+S}{n+p}\right\}$$

となり，δ^{ST} は経験 Bayes 推定量 $\hat{\sigma}^2_{EB}(0)$ として解釈されることがわかる．

θ_0 は事前情報として与えられるもので，θ が θ_0 に近いところで $\hat{\sigma}^2_{EB}(\theta_0)$ の改良度は大きく，従って事前に θ についてのなんらかの情報が得られていれば，$\hat{\sigma}^2_{EB}(\theta_0)$ は優れた推定量になる．しかし $\hat{\sigma}^2_{EB}(\theta_0)$ のよさは，その情報がたとえ間違っていてもその危険関数はつねに δ_0 の危険関数よりも小さく，主観的事前分布の誤った想定に対する実害が生じないことにある．換言すれば，事前情報に対して頑健なのである．こうした意味で経験 Bayes 推定量 $\hat{\sigma}^2_{EB}(\theta_0)$ は MLE のように事前情報を全く無視したものでもなく，また Bayes 推定量のように事前情報に完全に依存したものでもなく，事前情報を取り入れながらもその誤った想定による実害は避けたものになっている．

2.3. Brown-Brewster-Zidek(BBZ) 法. 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 $[0, \infty)$ を二分割し,

$$\hat{\sigma}^{2BR}(c) = \begin{cases} cS, & W < r\text{のとき}, \\ n^{-1}S, & W \ge r\text{のとき}, \end{cases}$$

なる推定量を考えた. 条件付期待値 $E_\lambda[L(cS/\sigma^2)|W < r]$ を最小にする定数 $c = c_\lambda^*(r)$ を求めると,

$$\begin{aligned} c_\lambda^*(r) &= \frac{\int F_p(rv;\lambda) f_n(v)dv}{\int vF_p(rv;\lambda) f_n(v)dv} \\ &\le \frac{\int F_p(rv;0) f_n(v)dv}{\int vF_p(rv;0) f_n(v)dv} = c_0^*(r) \end{aligned} \tag{2.6}$$

が成り立つことがわかる. ここで $F_p(x;\lambda) = \int_0^x f_p(t;\lambda)dt$ であり, $F_p(x;\lambda)/F_p(x;0)$ の単調性から (2.6) の不等式が導かれる. $c_0^*(r)$ は r について単調増加で $c_0^*(r) \le \lim_{r\to\infty} c_0^*(r) = n^{-1}$ が成り立ち, この事実と損失関数の凸性から Brown の推定量 $\hat{\sigma}^{2BR}(c_0^*(r))$ は 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た. $0 = r_{i,0} < r_{i,1} < \cdots < r_{i,n_i-1} < r_{i,n_i} = \infty$, $\lim_{i\to\infty} r_{i,1} = 0$, $\lim_{i\to\infty} r_{i,n_i-1} = \infty$, $\lim_{i\to\infty}\sup_j |r_{i,j} - r_{i,j-1}| = 0$ なる分割の列に対して,

$$\hat{\sigma}_i^2 = \{c_0^*(r_{i,j})S; \quad r_{i,j-1} \le W < r_{i,j}\}_{j=1,\cdots,n_i}$$

なる推定量の列を考えると, $c_0^*(r)$ の単調性から $\hat{\sigma}_i^2$ は δ_0 を改良することがわかる. 極限をとると $\lim_{i\to\infty}\hat{\sigma}_i^2 = c_0^*(W)S$ となり, この極限が Brewster-Zidek の推定量 δ^{BZ} である. Fatou の補題を使うと

$$\begin{aligned} E[L(\delta^{BZ}/\sigma^2)] &= E[L(\lim_{i\to\infty} \hat{\sigma}_i^2/\sigma^2)] \\ &\le \varliminf_{i\to\infty} E[L(\hat{\sigma}_i^2/\sigma^2)] \\ &\le E[L(\delta_0/\sigma^2)] \end{aligned}$$

が成り立ち, 極限値 δ^{BZ} が δ_0 を改良することが示される. 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 は階層的事前分布

$$\theta|\eta,\lambda \sim \mathcal{N}_p(0, \frac{1-\lambda}{\lambda}\eta^{-1}I_p), \ \frac{1}{\lambda}I_{(0,1)}(\lambda)d\lambda, \ \frac{1}{\eta}d\eta \tag{2.7}$$

に対して, δ^{BZ} が一般化 Bayes 推定量になっていることを示すとともに, δ_ϕ のクラスの中で許容的であることを示した. 後に Proskin(1985) は推定量全体の中での δ^{BZ} の許容性を証明しており, 従ってδ_0 を改良していてかつ許容的な推定量 δ^{BZ} が得られたことになる.

$p = 1$ のときには δ^{BZ} 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 (Rukhin(1987)) が, p が大きくなれば意味のある改善度が得られる. 図 2 で示されているように, δ^{ST} が $\lambda = 0$ で最大の改善度を与えるのに対して, δ^{BZ} 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^{BZ} の形の複雑さと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} が望ましいのかもしれない.

2.4. 新たな統一的方法 - IERD 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(1994a), 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 $\lim_{w\to\infty}\phi(w) = n^{-1}$ 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

$$
\begin{aligned}
&R(\theta,\sigma^2,\delta_0) - R(\theta,\sigma^2,\delta_\phi)\\
&= E[L(\phi(\infty)S/\sigma^2)] - E[L(\phi(W)S/\sigma^2)]\\
&= E[L(\phi(tW)S/\sigma^2)\Big|_{t=1}^{t=\infty}]\\
&= E[\int_1^\infty \frac{d}{dt}\{L(\phi(tW)S/\sigma^2)\}dt]\\
&= E[\int_1^\infty L'(\phi(tW)S/\sigma^2)\phi'(tW)WS/\sigma^2 dt]\\
&= \int_0^\infty\int_0^\infty\int_1^\infty L'(\phi(t\frac{u}{v})v)\phi'(t\frac{u}{v})udtf_p(u;\lambda)f_n(v)dudv
\end{aligned}
\tag{2.8}
$$

と表わされる. (2.8) の最後の等式の右辺において $w = (t/v)u$, $x = (wv)/t$ と順に変数変換すると

$$
\begin{aligned}
&R(\theta,\sigma^2,\delta_0) - R(\theta,\sigma^2,\delta_\phi)\\
&= \int_0^\infty\int_0^\infty\int_1^\infty L'(\phi(w)v)\phi'(w)\frac{wv^2}{t^2}f_p(\frac{wv}{t};\lambda)f_n(v)dtdvdw\\
&= \int_0^\infty\int_0^\infty\int_0^{wv} L'(\phi(w)v)\phi'(w)vf_p(x;\lambda)f_n(v)dxdvdw\\
&= \int_0^\infty \phi'(w)\int_0^\infty L'(\phi(w)v)vF_p(wv;\lambda)f_n(v)dvdw
\end{aligned}
\tag{2.9}
$$

という等式が成り立つ. ここで次の補題に注意しよう.

補題 2.1. 正値関数 $g(x)$, $h(x)$ に対して $h(x)/g(x)$ が単調増加すると仮定する. また $K(x)$ を実数値関数で, ある x_0 が存在して $x < x_0$ に対しては $K(x) < 0$, $x > x_0$ に対しては $K(x) \geq 0$ なるものと仮定する. このとき

$$
\int K(x)\frac{h(x)}{g(x)}dx \geq \frac{h(x_0)}{g(x_0)}\int K(x)dx
$$

なる不等式が成り立つ.

$\phi'(w) \geq 0$ を仮定すると, $F_p(x;\lambda)/F_p(x)$ の単調性と補題2.1 とから

$$\begin{aligned} &R(\theta, \sigma^2, \delta_0) - R(\theta, \sigma^2, \delta_\phi) \\ &\geq \int_0^\infty \phi'(w) \frac{F_p(wv_0;\lambda)}{F_p(wv_0)} \int_0^\infty L'(\phi(w)v)vF_p(wv)f_n(v)dvdw \end{aligned} \tag{2.10}$$

となる. 従って δ_ϕ が δ_0 を改良するためには, $\phi(w)$ について次なる条件が成り立てばよいことになる.

(a) $\phi(w)$ は単調増加で $\lim_{w\to\infty}\phi(w) = n^{-1}$,
(b) $\int_0^\infty L'(\phi(w)v)vF_p(wv)f_n(v)dv \geq 0$, 即ち$\phi(w) \geq \phi_0(w)$, 但し

$$\phi_0(w) = \frac{\int_0^\infty F_p(wv)f_n(v)dv}{\int_0^\infty vF_p(wv)f_n(v)dv}.$$

$\phi_0(w)$ 自身 (a), (b) の条件を満たすので $\phi_0(W)S$ は一つの改良型推定量となるが, (2.6) で与えられた $c_0^*(r)$ に対して $\phi_0(w) = c_0^*(w)$ であることから, これは 2.3 節で議論した Brewster-Zidek(1974) の推定量 δ^{BZ} に一致している. また (2.4) で与えられた $\phi_T(w)$ に対しては $\phi_0(w) \leq \phi_T(w)$ が成り立ちしかも $\phi_T(w)$ は (a) の条件をみたすから, $\phi_T(W)S$, 即ち Stein 型打ち切り推定量もこのクラスの中に入っている. 従って, 2.2 節と 2.3 節で別々の方法で導かれてきた 2 種類の推定量 δ^{ST}, δ^{BZ} が IERD 法により統一的に得られたことになる.

2.5. 分布の拡張と区間推定への適用. 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し, また区間推定に対しても適用可能である.

S と T を互いに独立な確率変数で

$$S/\sigma \sim g(v)I_{[v>0]}, \quad T/\sigma \sim h(u;\lambda)I_{[u>k(\lambda)]} \tag{2.11}$$

なる密度関数に従っているとする. ここで λ は実数, $k(\cdot)$ は実数値関数, σ と λ は未知母数, $I_{[\cdot]}$ は定義関数である. 例えば正規分布のときには $\lambda = ||\theta||^2/\sigma^2$, $k(\lambda) = 0$, 指数分布 $\sigma^{-1}\exp\{-(x-\mu)/\sigma\}I_{[x>\mu]}$ のときには$\lambda = \mu/\sigma$, $k(\lambda) = \lambda$ である. $L(t)$ を連続で bowl 型関数, 即ち $t < 1$ で単調減少, $t > 1$ で単調増加な関数とする. 尺度母数 σ を $\delta = \delta(S,T)$ で推定する問題を考え, 推定量 δ の良さは bowl 型損失関数 $L(\delta/\sigma)$ で測られるものとする.

S の定数倍の中で最良な推定量 $\delta_0 = c_0S$ を, $W = T/S$ に対して

$$\delta_\phi = \begin{cases} \phi(W)S, & W > 0 \text{ のとき}, \\ c_0S, & W \leq 0 \text{ のとき}, \end{cases} \tag{2.12}$$

なる推定量で改良することを考える. $H(x;\lambda) = \int_0^x h(u;\lambda)I_{[u>k(\lambda)]}du$, $H(x) = H(x;0)$, $h(x) = h(x;0)$ とし,

(A.1) $H(x;\lambda)/H(x)$ が $x>0$ に関して単調増加する

という仮定のもとで IERD 法を用いると

(a) $\phi(w)$ が単調増加で $\lim_{w\to\infty}\phi(w)=c_0$,

(b) $\int_0^\infty L'(\phi(w)v)vH(wv)g(v)dv \geq 0$,

なる条件が成り立つとき, δ_ϕ が δ_0 を改良することが示される. さらに $c_1>c_2>0$ に対して,

(A.2) $h(c_2x)/h(c_1x)$ が x に関して単調増加する

という仮定をおくと, BBZ 型及び Stein 型推定量がこのクラスに入ることがわかる. 仮定 (A.1), (A.2) は正規, 指数, 対数, Pareto 分布に対してみたされている. 指数分布のときには, Stein 型推定量が Arnold(1970), Zidek(1973) によって, BBZ 型推定量が Brewster(1974) によって提案されたものに一致する.

逆 Gauss 分布の母数推定に関する研究が最近取り上げられ, Stein 型推定量が Pal-Sinha(1989) により, BBZ 型推定量が MacGibbon-Shorrock(1994) により得られた. Kourouklis(1995) は, 逆 Gauss 分布が単調尤度比の性質をもつことを示し, それに基づいて (A.1), (A.2) がみたされ, ここでの一般的な枠組みに入ることを示した. 一様分布の尺度母数に対して改良型打ち切り推定量の導出が, Rukhin-Kuo-Dey(1990) によってなされたが, Bayes 的性質などに関する議論は今後の課題として残されている. Rukhin(1991) は, 位置尺度分布族において, 尺度母数の Stein 問題が漸近的に確率変数の 2 次式を含んだ関数の推定問題に帰着されることを示し, 漸近的に改良する Stein 型, BBZ 型推定量を導出した.

分散の区間推定についても点推定の場合に対応する結果が導かれる. Tate-Klett(1959) は S のみに基づいた $I_0=[aS,bS]$ なる形の信頼区間を考えた. ここで a,b は信頼係数 $0<1-\alpha<1$ に対し $P[aS<\sigma<bS]=1-\alpha$ 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. 最小比信頼区間のときには a,b のみたすべき等式は $a^{-1}g(a^{-1})=b^{-1}g(b^{-1})$ となる.

最小比信頼区間 I_0 を

$$I_\phi=\begin{cases}[a\phi(W)S,b\phi(W)S], & W>0\ \text{のとき},\\ I_0, & W\leq 0\ \text{のとき},\end{cases} \tag{2.13}$$

によって, 真の母数を覆う確率 (CP: Coverage Probability) を大きくする意味で改良するためには, 仮定 (A.1) のもとで

(a) $\phi(w)$ が単調増加で $\lim_{w\to\infty}\phi(w)=1$,

(b) $\frac{1}{a}g(\frac{1}{a\phi})H(\frac{w}{a\phi})\geq\frac{1}{b}g(\frac{1}{b\phi})H(\frac{w}{b\phi})$

なる条件が成り立てばよいことが IERD 法から導かれる. その際, 本質的な不等式の証明には補題 2.1 に対応する次の補題が使われる.

補題 2.2. 正値関数 $g(x)$, $h(x)$ に対して $h(x)/g(x)$ が単調増加すると仮定する. このとき $c_1,\ c_2>0,\ 0<x_1<x_2$ に対して

$$c_2h(x_2)-c_1h(x_1)\geq\frac{h(x_1)}{g(x_1)}\{c_2g(x_2)-c_1g(x_1)\}$$

なる不等式が成り立つ.

(A.2) を仮定すると BBZ 型及び Stein 型信頼区間がこのクラスに入ることがわかる. 正規, 指数分布に対する Stein 型信頼区間は Nagata(1989, 91) によって得られたものになっており, 例えば正規分布の場合には

$$I^{ST} = [a\phi_T(W)S,\ b\phi_T(W)S]$$
$$\phi_T(W) = \min\left\{1,\ \frac{a^{-1}-b^{-1}}{\log(b/a)}\frac{1+W}{n+p}\right\}$$

なる簡便な形の信頼区間となる.

一方, 最短信頼区間を改良する試みは Cohen(1970) に始まり, Shorrock(1990) によって BBZ 型信頼区間が求められその一般化 Bayes 性が示された. 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) によって示されている.

2.6 他の改良型推定量. 点推定での改良方法について, Stein 法と BBZ 法を中心に述べてきたが, その他に Strawderman(1974a) や Shinozaki(1995) による方法があるのでここで紹介しておく.

Strawderman(1974a) は, 非減少関数 $\phi(\cdot)$ に対して

$$\delta_\phi^{STD} = \frac{1}{n}\left\{1-\phi\left(\frac{1}{1+W}\right)\frac{1}{(1+W)^\varepsilon}\right\}S$$

なる縮小推定量を提案し, これがδ_0を改良するための, ϕ と ε に関する条件を求めた. Mathew-Sinha-Sutradhar(1992) は, この方法を分散成分の推定問題に適用した.

いままで述べてきた推定量は変換群に関して共変な推定量のクラスに属するものである. Shinozaki(1995) は, Stein 推定量 δ^{ST} が

$$\delta^{ST} = \frac{S}{n} - \frac{p}{n+p}\left(\frac{S}{n}-\frac{||X||^2}{p}\right)^+, \quad a^+ = \max(0,a)$$

と変形できることから, S/n を $\max(1,||X||^2/p)$ の方向へ縮小した非共変な推定量を 3 種類提案している. その一つは,

$$\delta_\phi^{SN} = \frac{S}{n} - \frac{p}{n+p}\phi\left(\frac{S}{S+||X||^2}\right)\left\{\frac{S}{n}-\max\left(1,\frac{||X||^2}{p}\right)\right\}^+$$

で与えられる. δ_ϕ^{SN} が δ_0 を改良するための ϕ に関する条件が得られている. またこうした非共変推定量の改良度が, Stein 推定量に比べてかなり大きいことを数値的に示している. 例えば, $n=3$, $p=1$, $\sigma^2=1$ の場合に相対的改良度を比較すると, $\theta=0$ のとき, δ^{ST} が 4.23%であるのに対してδ_ϕ^{SN} は 14.22%, また $\theta=0.5$ のときδ^{ST} が 3.98% に対して δ_ϕ^{SN} は 11.03% となり, δ_ϕ^{SN} の有効性が指摘されている. δ_ϕ^{SN}の有効性は, $\sigma^2>1>||X||^2/p$ のとき $n^{-1}S$ を $p^{-1}||X||^2$

へ縮小する Stein 推定量は明らかに縮小し過ぎであり, $n^{-1}S$ を $\max(1, p^{-1}||X||^2)$ の方向へ縮小する方が好ましいという理由によると解釈される.

2.7 分散比の二重縮小推定. 分散の推定に関連して分散比の二重縮小推定量を求める問題がある. X_1, S_1, X_2, S_2 を互いに独立な確率変数とし

$$X_i \sim \mathcal{N}_{p_i}(\theta_i, \sigma_i^2 I_{p_i}), \quad S_i \sim \sigma_i^2 \chi^2_{n_i}, \quad i = 1, 2$$

なる分布に従うとする. このとき分散比 $\rho = \sigma_2^2/\sigma_1^2$ を推定する問題を考えると, S_2/S_1 の定数倍の推定量の中で平均 2 乗誤差を最小にするものは

$$\hat{\rho}_0 = \frac{n_1 - 4}{n_2 + 2} \frac{S_2}{S_1}$$

で与えられる. Gelfand-Dey(1988) は X_1, X_2 の情報を使って $\hat{\rho}_0$ を改良する推定量として

$$\hat{\rho}_1 = \max\left\{\hat{\rho}_0, \ \frac{n_1 + p_1 - 4}{n_2 + 2} \frac{S_2}{S_1 + ||X_1||^2}\right\},$$
$$\hat{\rho}_2 = \min\left\{\hat{\rho}_0, \ \frac{n_1 - 4}{n_2 + p_2 + 2} \frac{S_2 + ||X_2||^2}{S_1}\right\},$$

を導いた. これは本質的には Stein 型と同等なもので, X_1, X_2 の両方を用いた二重縮小推定量の導出が望まれるものの容易でないとされてきた. しかし Kubokawa(1994b) は IERD 法の使用によってこの問題に一つの可能な解を与えることに成功した. 彼の導出した二重縮小推定量は

$$\begin{aligned} \hat{\rho}^{DS} &= \hat{\rho}_1 + \hat{\rho}_2 - \hat{\rho}_0 \\ &= \hat{\rho}_1 + \min\left\{0, \ \frac{n_1 - 4}{n_2 + p_2 + 2} \frac{S_2 + ||X_2||^2}{S_1} - \hat{\rho}_0\right\} \\ &= \hat{\rho}_2 + \max\left\{0, \ \frac{n_1 + p_1 - 4}{n_2 + 2} \frac{S_2}{S_1 + ||X_1||^2} - \hat{\rho}_0\right\} \end{aligned}$$

であり, これは $\hat{\rho}_1, \hat{\rho}_2$ の両方を改良することが証明された. 2 番目と 3 番目の等式の右辺の第 2 項は, それぞれ $\hat{\rho}_1, \hat{\rho}_2$ の縮小し過ぎた部分を修正している項であると解釈される.

2.8 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

$$S \sim \mathcal{W}_p(n, \Sigma), \quad X \sim \mathcal{N}_{p\times r}(\Theta, \Sigma \otimes I_r) \tag{2.14}$$

と表わされる. ここで $\mathcal{W}_p(n, \Sigma)$ 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 Shorrock-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. Sugiura-Konno(1988) はその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た．

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-Robert-Saleh(1992) は Sinha-Ghosh 推定量を改良する経験 Bayes 推定量を導出し, Kubokawa-Honda-Morita-Saleh(1993) 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 $|\Sigma|$ 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 4 節で扱う．

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 3 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 ら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), Mikhail-Vassily(1994) により優れた総合報告が出され, Judge-Bock(1978), Berger(1985), Hoffmann(1992), Robert(1994) の本の中で幅広く取り上げられているので参照されたい．

3.1. Stein 現象. p-次元確率ベクトル $X=(X_1,\ldots,X_p)'$ が正規分布 $\mathcal{N}_p(\theta, I_p)$ に従うとき平均ベクトル $\theta=(\theta_1,\ldots,\theta_p)'$ 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 2 乗損失関数 $||\delta(X)-\theta||^2$ に関する危険関数がとられる．

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 ΓX, $\Gamma\theta$ に関して共変な推定量

$$\delta_\phi = \left\{1-\frac{\phi(||X||^2)}{||X||^2}\right\}X \tag{3.1}$$

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,

$$\delta^{JS} = \left\{1-\frac{p-2}{||X||^2}\right\}X \tag{3.2}$$

という形の改良型推定量の一つを明示的に与えた．これは X を 0 の方向へ縮約しているので縮小推定量と呼ばれる．この Stein 現象を説明する議論には 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 (Efron-Morris(1972a),Robbins(1983))．いま母数 θ を確率変数と考え θ が事前分布 $\mathcal{N}_p(\theta_0, a^{-1}I_p)$ に従うとする．ここで a は未知母数で θ_0 は予め与えられた既知のベクトルとする．このとき X を与えたときの θ の事後分布は $\mathcal{N}_p(\theta_0+(1-\tau)(X-\theta_0),(1-\tau)I_p)$, $\tau=a/(1+a)$ となるから θ の Bayes 推定量は

$$\hat{\theta}_B(\theta_0)=\theta_0+(1-\tau)(X-\theta_0)$$

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 $\mathcal{N}_p(\theta_0,\tau^{-1}I_p)$ で推定することが考えられる．$||X-\theta_0||^2$ の周辺分布が $\tau^{-1}\chi_p^2$ であるからτ の不偏推定量は

$$\hat{\tau}=\frac{p-2}{||X-\theta_0||^2}$$

となる．これを上の Bayes 推定量に代入して経験 Bayes 推定量

$$\hat{\theta}_{EB}(\theta_0)=\theta_0+(1-\hat{\tau})(X-\theta_0)=\theta_0+\left\{1-\frac{p-2}{||X-\theta_0||^2}\right\}(X-\theta_0)$$

が得られ, $\hat{\theta}_{EB}(0)$ が δ^{JS} に一致することがわかる．もちろん $\hat{\theta}_{EB}(\theta_0)$ は $p\geq 3$ に対して X を改良するので, 2.2 節で分散の経験 Bayes 推定の折 述べたことと同様な意味付けが $\hat{\theta}_{EB}(\theta_0)$ についても与えられる．即ち James-Stein 推定量は事前情報を取り入れながらもその誤った想定による実害は避けたものになっている．δ^{JS} の経験 Bayes による動機付けは Efron-Morris(1972a) によってなされたが, 彼は同時にこの方法によって 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

$$E[(X_i-\theta_i)h(X_i)]=E[h'(X_i)] \tag{3.3}$$

なる等式が成り立つ (Stein(1973,81))．これは Stein identity と呼ばれ, これを用いると δ_ϕ の危険関数は

$$\begin{aligned}R(\theta,\delta_\phi)&=E\left[p+\frac{\phi^2}{||X||^2}-2\sum_{i=1}^{p}(X_i-\theta_i)X_i\frac{\phi}{||X||^2}\right]\\&=E\left[p+\frac{\phi}{||X||^2}\{\phi-2(p-2)\}-4\phi'\right]\end{aligned} \tag{3.4}$$

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 $\phi(w)$ が

$$\phi(w)\{\phi(w) - 2(p-2)\}/w - 4\phi'(w) \le 0$$

なる微分不等式をみたせばよい. 例えば

(1) $\phi(w)$ が非減少,
(2) $0 < \phi(w) \le 2(p-2)$,

であればよい (Efron-Morris(1976a)). 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. $\phi(w) = p-2$ は明らかに (1), (2) の条件をみたすので δ^{JS} はこのクラスに入り, その危険関数は $R(\theta, \delta^{JS}) = p - (p-2)^2 E[||X||^{-2}]$ で与えられ, 非心度が 0 のとき最大の改善が得られる (図 3). 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$\delta_+^{JS} = \max\{0, 1-(p-2)/||X||^2\}X$ なる positive-part Stein 推定量が考えられ, 実際 δ^{JS} を改良することが示される. δ_+^{JS} 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が, δ_+^{JS} を改良する明示的な推定量を見つけることはここ十数年の大問題とされてきた. 最近 Shao-Strawderman(1994) はその導出に成功し,

$$\delta_g^{SS}(a) = \delta_+^{JS} - \frac{a\, g(||X||^2)}{||X||^2} X I_{[p-2 \le ||X||^2 \le p]}, \tag{3.5}$$

なる改良された推定量を求めた. 但し, $g(t)$ は $t = p-1$ について対称で, $g(p-2) = g(p) = 0$,

$$g(t) = \begin{cases} t - p, & p^* \le t \le p \text{ のとき}, \\ 2p^* - p - t, & p-1 \le t < p^* \text{のとき}, \end{cases}$$

p^*, a は適当な定数とする. 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) は階層的事前分布

$$\theta|\eta, \lambda \sim \mathcal{N}_p(0, \frac{1-\lambda}{\lambda}\eta^{-1} I_p), \frac{1}{\lambda^2} I_{(0,1)}(\lambda) d\lambda, \frac{1}{\eta} d\eta$$

を想定し,

$$\phi_{GB}(w) = p - 2 - \frac{2}{\int_0^1 z^{p/2-2} e^{(1-z)w/2} dz} \tag{3.6}$$

に対して $\delta_{GB} = \delta_{\phi_{GB}} = \{1 - \phi_{GB}(||X||^2)/||X||^2\}X$ なる形の一般化 Bayes 推定量を求め, $\phi_{GB}(w)$ が上の (1),(2) の条件をみたすこと, 即ち δ_{GB} が X を改良することを示した. また Brown-Hwang(1982) の条件が成り立つことを確かめることによりその許容性が証明される. δ^{JS}, δ_{GB}, δ_+^{JS} の危険関数の挙動については図 3 で示されている. また別の形の許容的ミニマクス推定量が Alam(1073) によって与えられている.

Stein(1973,81) は $\nabla^2 = \sum \partial^2/\partial x_i^2$, $\nabla = (\partial/\partial x_1, \ldots, \partial/\partial x_p)'$ に対して, f が超調和条件 $\nabla^2 f(x) \leq 0$ をみたせば$\delta^{ST}(f) = X + \nabla \log f(X)$ 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 $\delta^{ST}(f)$ の形で表現され, しかも $f(x)$ は周辺密度で与えられ, この $f(x)$ 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)). Berger-Srinivasan(1978) は, θ の許容的推定量の, 一般化 Bayes 推定量としての特徴付けを与えた. なお, 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 (Brown(1971), Srinivasan(1981), Brown-Hwang(1982), Berger(1985) 等) や許容性とマルコフ連鎖の再帰性との関係 (Brown(1971), Johnstone(1984), Eaton(1992) 等) などが論じられている.

3.2. IERD 法による James-Stein 推定量の改良. James-Stein 推定量が非許容的であることは 3.1 節で述べたが, ここではそれを改良する推定量のクラスを構成することを考えよう. Stein identity を用いた危険関数の表現 (3.4) から, δ_ϕ が δ^{JS} を改良する条件として

$$\{\phi(w) - (p-2)\}^2 - 4w\phi'(w) \leq 0$$

が得られる. しかし実際この不等式をみたす $\phi(w)$ の一般的な条件を与えるのは困難であろう. ここに identity による特徴付けの限界があると思われる. そこで IERD 法をこの問題に適用することを試みよう. $\lim_{w\to\infty}\phi(w) = p-2$ を仮定すると, (3.4) より

$$\begin{aligned}
&R(\theta,\delta^{JS}) - R(\theta,\delta_\phi)\\
&= E\left[\frac{\phi(\infty)}{||X||^2}\{\phi(\infty)-2(p-2)\} - \frac{\phi(||X||^2)}{||X||^2}\{\phi(||X||^2)-2(p-2)\}\right] + 4E[\phi'(||X||^2)]\\
&= E\left[\int_1^\infty \frac{d}{dt}\left\{\frac{\phi(t||X||^2)}{||X||^2}\{\phi(t||X||^2)-2(p-2)\}\right\}dt\right] + 4E[\phi'(||X||^2)]\\
&= 2E[\int_1^\infty \{\phi(t||X||^2)-(p-2)\}\phi'(t||X||^2)dt] + 4E[\phi'(||X||^2)]\\
&= 2\int_0^\infty\int_1^\infty \{\phi(tx)-(p-2)\}\phi'(tx)f_p(x;\lambda)dtdx + 4E[\phi'(||X||^2)],
\end{aligned} \tag{3.7}$$

但し $\lambda = ||\theta||^2$ とする. $w = tx$, $y = w/t$ と変数変換すると (3.7) の最後の等式の左辺の第 1 項は

$$\begin{aligned}
&2\int_0^\infty\int_1^\infty \{\phi(w)-(p-2)\}\phi'(w)f_p(\frac{w}{t};\lambda)\frac{1}{t}dtdw\\
&= 2\int_0^\infty\int_0^w \{\phi(w)-(p-2)\}\phi'(w)f_p(y;\lambda)y^{-1}dydw
\end{aligned}$$

となり, 従って,

$$\begin{aligned}
&R(\theta,\delta^{JS}) - R(\theta,\delta_\phi)\\
&= 2\int_0^\infty \phi'(w)\left\{\{\phi(w)-(p-2)\}\int_0^w y^{-1}f_p(y;\lambda)dy + 2f_p(w;\lambda)\right\}dw
\end{aligned} \tag{3.8}$$

となる. (3.8) と次の不等式

$$f_p(w;\lambda)/\int_0^w y^{-1}f_p(y;\lambda)dy \geq f_p(w)/\int_0^w y^{-1}f_p(y)dy$$

とから δ_ϕ が δ^{JS} を改良するためには次の条件 (a), (b) がみたされればよいことがわかる.

(a) $\phi(w)$ は単調増加で, $\lim_{w\to\infty}\phi(w) = p-2$,

(b) $\phi(w) \geq \phi_0(w)$, 但し

$$\begin{aligned} \phi_0(w) &= p-2-2f_p(w)/\int_0^w y^{-1}f_p(y)dy \\ &= \int_0^w f_p(y)dy/\int_0^w y^{-1}f_p(y)dy. \end{aligned} \tag{3.9}$$

実は $\phi_0(w) = \phi_{GB}(w)$ であり, 従って Strawderman(1971) の一般化 Bayes 推定量 δ_{GB} は James-Stein 推定量を改良していることがわかる. Kubokawa(1991) は 分散の推定で使われた BBZ 法をこの問題に適用し δ_{GB} が分割を細かくした極限として導かれること, 即ち BBZ 型推定量であることを示している. 一方 δ_+^{JS} が Stein 型推定量に対応しており, 2 節で扱った分散の推定問題が δ^{JS} の改良に関係していることを暗示している (Rukhin(1992b)). 分散の推定との関連で述べると, 分散が未知のモデル (2.1) においては, James-Stein 推定量は分散 σ^2 の最良共変推定量 $\hat{\sigma}_0^2$ を用いて $\{1-\hat{\sigma}_0^2(p-2)/||X||^2\}X$ で与えられるが, $\hat{\sigma}_0^2$ の代わりに分散の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-Makita-Morita-Nagakura(1993)).

3.3. 拡張と新たな展開. Stein 現象については様々な拡張や新たな展開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.

James-Stein 推定量 δ^{JS} の危険関数は $p-(p-2)^2E[1/||X||^2]$ で与えられるが, このことは θ_i の中の一つでも絶対値の大きいものがあれば全体として大きな改良が望めないことを意味している. そこでこのような問題を解決するために James-Stein 推定量の修正が Efron-Morris(1972b), Stein(1981) らによって提案された. Stein(1981) の考案した打ち切り型推定量は, 成分毎に

$$\delta_i^{(\ell)} = \left(1 - \frac{(\ell-2)\min\{1, Z_{(\ell)}/|X_i|\}}{\sum_{j=1}^p X_j^2 \wedge Z_{(\ell)}^2}\right)X_i$$

で与えられる. ここで ℓ は適当な定数, $a\wedge b = \min(a,b)$ とし,

$$Z_i = |X_i|, \quad Z_{(1)} < Z_{(2)} < \cdots < Z_{(p)}$$

を $Z_1, Z_2, \ldots, Z_p$ の順序統計量とする. このとき $\delta^{(\ell)}$ の危険関数は,

$$R(\theta, \delta^{(\ell)}) = p-(\ell-2)^2E[1/\sum_{j=1}^p X_j^2 \wedge Z_{(\ell)}^2]$$

と表され, たとえ大きな X_j^2 がいくつか存在しても, $\delta^{(\ell)}$ はそれらの影響を受けないように修正されたことがわかる. またこの推定量は, 危険関数の (各θ_iについての) 成分の最大値を小さくしていることが知られている. ℓの決め方やロバスト推定に関する議論が, Dey-Berger(1983), Berger-Dey(1985) で与えられている.

不等分散をもった多変量正規モデル, $X = (X_1, \ldots, X_p)' \sim \mathcal{N}_p(\theta, D)$, $D = \mathrm{diag}(d_1, \ldots, d_p)$ での縮小推定の問題は, Efron-Morris(1975), Fay-Herriot(1979), Morris(1983) などにより様々な応用場面で議論されてきた. ミニマクスな通常の縮小推定量は,

$$\delta_i^{MS} = \left\{1 - \frac{p-2}{\Sigma_i X_i^2/d_i}\frac{1}{d_i}\right\} X_i, \quad i = 1, \ldots, p$$

で与えられるが (Berger(1976)), 他に比べて小さい d_i があると, 改良度が小さくなってしまう. そこで Shinozaki-Chang(1994) は, ミニマクス的経験 Bayes 推定量として

$$\sum_{i=1}^{p} \frac{X_i^2}{\hat{a} + d_i} = p - 2$$

の解$\hat{a}$ に対して

$$\delta_i^{SC} = \left\{1 - \frac{\Sigma d_i^2 - 2d_{max}^2}{(p-2)d_{max}^2}\frac{d_i}{\hat{a} + d_i}\right\} X_i, \quad i = 1, \ldots, p$$

を提案し, δ^{MS} との数値的比較を行った.

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 $\delta^{JS}(V)$ 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(1986a,b) は θ が存在すると推察される部分空間の候補が複数個 $V_1, \ldots, V_k$ 考えられるとき, それぞれに縮小する Stein 推定量 $\delta^{JS}(V_i)$ の重み付きの和として表わされる適応型推定量$\sum_{i=1}^k \rho_i(X)\delta^{JS}(V_i)$ を提案した. 但し $\rho_i(X)$ は改良度の大きいと考えられる $\delta^{JS}(V_i)$ に対する重みが大きくなるように作られている.

主観的事前分布の誤った想定に対してロバストなBayes推定についての研究が Berger(1980b) らによって展開された. まず $X \sim \mathcal{N}_p(\theta, \Sigma)$, Σ:既知, なるモデルを考えるとき, 事前分布 π_0: $\theta \sim \mathcal{N}_p(\mu, A)$ に対する θ の Bayes 推定量は

$$\delta^{\pi_0} = X - \Sigma(\Sigma + A)^{-1}(X - \mu)$$

となり, 超母数 μ, A は主観的に決められる. $A - \Sigma \geq 0$ とし, $0 < \lambda < 1$ に対して, $B(\lambda) = \lambda^{-1}A - \Sigma$ とおくとき, $k > 0$ に対して階層的事前分布

$$g_k(\theta) = \int_0^1 \frac{\lambda^{k-1-(p/2)}}{|B(\lambda)|^{1/2}} e^{-(\theta-\mu)'B(\lambda)^{-1}(\theta-\mu)/2} d\lambda$$

を考える. これは, Strawderman(1971) の事前分布を拡張したものであり, 一般化 Bayes 推定量は

$$\delta^{RB} = X - \frac{r_k((X-\mu)'(\Sigma + A)^{-1}(X-\mu))}{(X-\mu)'(\Sigma + A)^{-1}(X-\mu)}\Sigma(\Sigma + A)^{-1}(X - \mu)$$

$$r_k(v) = \frac{v\int_0^1 \lambda^k \exp\{-\lambda v/2\}d\lambda}{\int_0^1 \lambda^{k-1}\exp\{-\lambda v/2\}d\lambda}$$

で与えられ, $r_k(v)=\min\{p-2,v\}$ と近似される. この推定量は,

$$(X-\mu)'(\Sigma+A)^{-1}(X-\mu)\le p-2$$

のときには主観的 Bayes 推定量 δ^{π_0} になり, それ以外では Stein 型縮小推定量になっている. 従って, θ に対する μ, A の想定が正しければ Bayes 推定量δ^{π_0} が生じ, その想定が誤っていれば Stein 型縮小推定量となり, その誤りが著しければ通常の推定量 X に近づくこと, 即ちδ^{RB} は事前分布の主観性に対してロバストな性質をもっていることがわかる.

ロバストな Bayes 推定量 δ^{RB} は残念ながらどんな Σ, A に対してもミニマクス性をもっているとは限らない. そこで Berger(1982) は, Bhattacharya(1966) のアイデアを用いて, 主観的超母数 μ, A を組み込んでかつミニマクス性を合わせもつ推定量を導出した. 簡単のために, $\Sigma=\mathrm{diag}(\sigma_1^2,\ldots,\sigma_p^2)$, $A=(a_1,\ldots,a_p)$, $q_i=\sigma_i^4/(\sigma_i^2+a_i)$ とし $q_1\ge q_2\ge\cdots\ge q_p>0\equiv q_{p+1}$ とするとき, ミニマクスなロバスト Bayes 推定量は, 成分毎に

$$\begin{aligned}\delta_i^{MB}=&X_i-\frac{\sigma_i^2}{\sigma_i^2+a_i}(X_i-\mu_i)\\&\times\left[\frac{1}{q_i}\sum_{j=i}^{p}(q_j-q_{j+1})\min\left\{1,\frac{2(j-2)^+}{\sum_{k=1}^{j}(X_k-\mu_k)^2/(\sigma_k^2+a_k)}\right\}\right]\end{aligned}$$

で与えられる.

主観的事前分布 π_0 に対して, q だけ ε-contaminated な事前分布 $\pi=(1-\varepsilon)\pi_0+\varepsilon q$ を想定したときのロバストな Bayes 推定の議論が Berger(1985), Berger-Berliner(1986) 等で展開されている. 彼らは, 事前分布 π に対する x の周辺密度 $m(x|q)$ を最大にする q を$\hat{q}$ とかくとき, $\hat{\pi}=(1-\varepsilon)\pi_0+\varepsilon\hat{q}$ を ML-II 事前分布と呼び, これに対する Bayes 推定量が

$$\delta^{\hat{\pi}}(x)=\hat{\lambda}_\varepsilon(x)\delta^{\pi_0}(x)+(1-\hat{\lambda}_\varepsilon(x))\delta^{\hat{q}}(x)$$

なる形で表されることを示し, Stein 問題への適用と種々の議論を行った. ここで δ^{π_0} は π_0 に対する通常の Bayes 推定量, $\delta^{\hat{q}}$ は $\hat{q}$ に対する経験 Bayes 推定量を表しており, $\delta^{\hat{\pi}}$ は, δ^{π_0} と $\delta^{\hat{q}}$ の結合推定量でその比率が ε の度合いに応じて決まる形になっている. ロバストな Bayes 推定に関する一連の理論的結果が Berger-Berliner(1984) でわかりやすく解説されているので参照されたい. Berliner(1985) は, ある事前分布のクラス Γ に対して一様に Bayes risk を最小にする推定量を求める問題は, ガンマ分布の尺度母数の (同時) 推定問題に帰着でき, δ^{JS}が Γにおいて ロバストな Bayes 推定を与えるという結果を示した.

事前分布の超母数に対して, さらに事前分布を仮定する, いわゆる階層 Bayes 推定についての研究が Berger(1985), Ghosh-Sinha(1988), Berger-Robert(1990) などによってなされた. 例えば, $X\sim\mathcal{N}_p(\theta,\sigma^2I)$ (σ^2 は既知) なるモデルにおいて, $\theta\sim\mathcal{N}_p(\beta\mathbf{1},\sigma_\pi^2I)$, $\mathbf{1}=(1,\ldots,1)'$ なる第 1 ステップの事前分布と, $\beta\sim\mathcal{N}(\beta^0,A)$, $\sigma_\pi^2\sim\pi_2^2(\sigma_\pi^2)$ なる第 2 ステップの事前分布を想定すると, 一つの階層 Bayes 推定量が得られる. 簡単のために, 第 2 ステップを無情報事前分布,

即ち $A \to \infty$, $\pi_2^2(\sigma_\pi^2) = 1$ とすると, それは

$$\delta^{HB} = X - E^{\pi_2^2(\sigma_\pi^2|X)}\left[\frac{\sigma^2}{\sigma^2+\sigma_\pi^2}\right](X - \overline{x}\mathbf{1})$$

$$\pi_2^2(\sigma_\pi^2|X) \propto (\sigma^2+\sigma_\pi^2)^{-(p-1)/2}\exp\left\{-\frac{||X-\overline{x}\mathbf{1}||^2}{2(\sigma^2+\sigma_\pi^2)}\right\}, \quad \overline{x} = p^{-1}\sum_{i=1}^{p} X_i$$

と表される. 経験 Bayes 推定において問題となった縮小過多 (Over Shrinkage) による弊害が階層 Bayes 推定においては生じないことがわかる. より一般的なモデルの設定での階層 Bayes 推定量とその危険関数の不偏推定量の導出及びミニマクス性が成り立つための一般的な条件などが Berger-Robert(1990) で議論されている.

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), Baranchik(1970), Lin-Tsai(1973), Bock(1975), Alam(1975), Efron-Morris(1976a)). 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 $\mathcal{N}_{p\times r}(\Theta, I_p \otimes I_r)$ に従うとき, Efron-Morris(1972a) は通常の推定量 X が経験 Bayes 推定量

$$\hat{\Theta}_0^{EM} = \{I_p - (r-p-1)(XX')^{-1}\}X$$

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(1976b) は経験 Bayes の方法により $\hat{\Theta}_0^{EM}$ の改良は共分散行列の逆行列の推定問題に帰着できることを示し,

$$\hat{\Theta}_1^{EM} = \hat{\Theta}_0^{EM} - (p^2+p-2)(\mathrm{tr}XX')^{-1}X$$

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. 誤差共分散行列が未知の一般的な多変量回帰モデル (2.14) 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), Shieh(1993) 等によりなされ, 成長曲線モデルでの議論が Kubokawa-Saleh-Morita(1992), Tan(1991) によってなされた. また Kariya-Konno(1994), Kariya-Konno-Strawderman(1994) は二重縮小推定についての興味深い結果を導いている.

この他にも, 母数 $\theta_1, \ldots, \theta_p$ 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-Nickerson-Sen(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 Stein(1960), Baranchick(1973), Takada(1979), Zidek(1978) により議論された. Brown(1990) は, この問題を, 補助統計量を与えたときの条件付き推測及び条件無しでの推測の立場で議論し, それぞれの立場で決定理論的結果が異なること (Ancillarity Paradox) を示した. Brown の論文には多くの興味深い議論とコメントが付加されているので参照されたい. 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 Brown(1973), Shinozaki(1980), Hwang(1985) によって, Pitman closeness なる規準の下での解明が Sen-Kubokawa-Saleh(1989) によってなされた.

George(1991), Krishnamoorthy(1992), Sarkar(1994) は, 未知の不等分散をもった多変量正規分布の共通平均ベクトルを推定する問題を考え, たとえ分散を推定するための統計量が存在しなくても, 線形推定量は縮小推定量によって改良できるという興味深い結果を導出した.

Johnstone(1987) は, MLE や James-Stein 推定量に対する損失関数の推定問題を扱い, 不偏推定量の非許容性を示した. positive-part Stein 推定量や一般化 Bayes 推定量に対する損失関数の推定問題については, Lu-Berger(1989a) で議論され, 他の指数型分布族に対しては, Lele(1992) により議論されている.

3.4. 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.

[1] 楕円分布. 球面対称分布や楕円分布における Stein 問題は, Brandwein-Strawderman(1990) により概説されているので参照されたい. $G(\eta)$ を η の分布関数とするとき, 正規分布の $G(\eta)$-尺度混合分布

$$f(||x-\theta||)=\int\left(\frac{\eta}{2\pi}\right)^{p/2}e^{-(\eta/2)||x-\theta||^2}G(d\eta)$$

での議論が, Strawderman(1974b), Berger(1975), Srivastava-Bilodeau(1989), Chou -Strawderman(1990), により与えられ, $p\geq 3$ での Stein 現象が示された. また未知の分散をもった尺度混合分布での結果が Bravo-MacGibbon(1988a) によって与えられた.

尺度混合分布を仮定しない一般的な球面対象分布においては, $p\geq 4$ のときに Stein 効果が現れることが, Brandwein-Strawderman(1978, 91a,b), Bock(1985) によって証明され, 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 Ralescu-Brandwein-Strawderman(1992) により示された. Elliptical Contoured (EC) 分布においては, 最小 2 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-Fourdrinier-Robert(1989) により示された.

[2] 連続型指数分布族. 連続型指数分布族の場合 Hudson(1978) は, 部分積分による identity を導入して James-Stein 推定量に対応する改良型推定量を導出した. その後の研究においては, 特に, p 個の独立な確率変数 $X_1,\ldots,X_p$ があって, 各 X_i がガンマ分布

$$f(x_i,\theta_i)=\exp(-\theta_i x_i)x_i^{\alpha_i-1}\theta_i^{\alpha_i}/\Gamma(\alpha_i),\quad x_i>0$$

に従うとき, $\theta=(\theta_1,\ldots,\theta_p)'$ 及び $\theta^{-1}=(\theta_1^{-1},\ldots,\theta_p^{-1})'$ を同時推定する問題が活発に議論された. Berger(1980a) は, θ^{-1}を損失関数

$$L(\delta,\theta)=\sum_{i=1}^{p}\theta_i^m(1-\delta_i\theta_i)^2$$

に関して推定する問題において, 改良のための微分不等式を求め, $m=-2,-1,1$ のときには $p\geq 2$ で, また $m=0$ のときには $p\geq 3$ で解が存在することを示した. Ghosh-Parsian(1980) はその解のクラスの拡張を行い, DasGupta(1986) は一般的な損失関数

$$L(\delta,\theta)=\sum_{i=1}^{p}c_i\theta_i^{m_i}(1-\delta_i\theta_i)^2,\quad c_i>0$$

を扱い, 最良共変推定量$\delta_{0,i} = X_i/(\alpha_i+1), i=1,\ldots,p,$ が縮小推定量

$$\delta_i^{DG} = \frac{X_i}{\alpha_i+1}(1+\phi_i(X)), \quad i=1,\ldots,p,$$
$$\phi_i(X) = -c(\mathrm{sgn} m_i)x_i^{m_i/2}\prod_{j=1}^{p} x_j^{-m_j/2p}, \quad c>0$$

によって, $p \geq 2$ で改良されることを証明した. また

$$L(\delta,\theta) = \sum_{i=1}^{p} c_i\theta_i^{m_i}(\delta_i/\theta_i-1)^2, \quad c_i>0$$

の下で θ を推定する場合にも, $p \geq 2$ に対して最良共変推定量の非許容性を示した. 後にDasGupta(1989) は, 一般的な分布族での尺度母数や共分散行列の固有値など, 正値をとる母数の同時推定に対して, 縮小推定による改良についての一般論を打ち立てた.

その他に Hudson の identity の多次元指数分布族への拡張 (Chou(1988)), Kullback-Leibler 損失の下での $p \geq 3$ に対する Stein 現象 (Dey-Ghosh-Srinivasan(1987)), 指数分布族での超調和条件 (Haff-Johnson(1986)), 多重縮小推定 (Ki-Tsui(1990)), 混合分布での尺度母数推定 (Dey(1990)), $\theta^f = (\theta_1^{f_1},\ldots,\theta_p^{f_p})'$, ($f_i=1$ または $f_i=-1$) の推定 (Bilodeau(1988)) 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. 逆 Gauss 分布についての結果も Bravo-MacGibbon(1988b) によって得られている.

[3] 離散型指数分布族. ポアソン分布や負の二項分布などを含む離散型指数分布族での Stein 現象も知られている. $X_1,\ldots,X_p$ が $X_i \sim \mathcal{P}o(\theta_i)$ なる独立な確率変数とするとき, $\theta = (\theta_1,\ldots,\theta_p)'$を同時推定する際に 2 種類の損失関数 $L_0(\delta,\theta)$, $L_1(\delta,\theta)$ が主に扱われてきた. ここで,

$$L_m(\delta,\theta) = \sum_{i=1}^{p}(\delta_i-\theta_i)^2/\theta_i^m, \quad m=0,1$$

とする. L_1損失に関して, Clevenson-Zidek(1975) は,

$$\delta^{CZ} = \left[1-\frac{\beta+p-1}{\sum_{i=1}^{p}X_i+\beta+p-1}\right]X_i, \quad i=1,\ldots,p$$

なる推定量が, $p \geq 2$, $0 \leq \beta \leq p-1$ のとき, X を改良することを示すとともに, X を改良する許容的一般化 Bayes 推定量を導出した. Ghosh-Parsian(1981) は一般化 Bayes 推定量のクラスを構成し, Tsui-Press(1982) は改良する推定量のクラスを構成した. また Tsui(1984) は, 真の分布が負の二項分布であっても, Clevenson-Zidek 型推定量の優越性が成り立つことを示した.

一方 L_0 損失に関しては, Hudson(1978) が, 離散分布での identity を導入して改良するための差分不等式を導出し, $p \geq 3$ のときにその解を求めた. より一般的な損失関数$\sum(\delta_i-\theta_i)^2/\theta_i^{m_i}$ と一般的な離散指数分布族での差分不等式の解法等が, Hwang(1982), Ghosh-Hwang-Tsui(1983), Chou(1991) によって統一的にまとめられた. L_0損失, L_1損失で提案された様々な推定量の紹介とそれらの良さの比較が Jun(1993) によって与えられているので参照されたい.

その他, Kullback-Leibler 損失の下での $p \geq 3$ に対する Stein 現象 (Ghosh-Yang(1988)), 離散混合分布での結果 (Dey-Chung(1992)), 二項分布での改良結果 (Johnson(1987)), 離散分布での許容性 (Brown(1981), Johnstone(1984), Brown-Farrell(1985)) などの議論がある. また, 分割表 (Gupta-Saleh-Sen(1989), Albert(1987)), 多項分布 (Lwin-Maritz(1989)), ディリクレ分布 (Kuo(1986)) などについても議論があり, 応用と関連させながら今後の展開が期待される.

[4] 漸近理論. ノンパラメトリックモデルにおいては L-, M-, R- 推定量の Stein 効果による改良が Sen-Saleh(1985, 87), Saleh-Sen(1985), Shiraishi(1991) 等によって示された. 時系列モデルでの Stein 現象については Chaturvedi-Hoa-Shukla(1993), Nickerson-Basawa(1992), Koul-Saleh(1993) により, また Gauss 過程での決定理論については Spruill(1982), Majumdar(1994), Mandelbaum-Shepp(1987) により調べられた.

Yanagimoto(1994) は James-Stein 推定量 δ^{JS} に対して

$$E[||X-\theta||^2] = E[||X-\delta^{JS}||^2] + E[||\delta^{JS}-\theta||^2]$$

なる関係が成り立つことに注目し, これを平均ピタゴラス関係と呼んだ. 最近, こうした関係が漸近的に成りたつことを Eguchi-Yanagimoto(1994) が示している. 密度関数 $p(x;\theta)$, $\theta \in \mathbf{R}^p$, の Fisfer 情報量の逆数を $(g^{ij}(\theta))_{ij}$ とし,

$$(\mathrm{grad} f)^i = \sum_{j=1}^{p} g^{ij}(\theta)\frac{\partial}{\partial\theta_j} f(\theta)$$

と定義する. MLE $\hat{\theta}$ に対して,

$$\hat{\theta}^* = \hat{\theta} + \frac{1}{n}\mathrm{grad} u(\hat{\theta})$$

とするとき, 漸近的に平均ピタゴラス関係

$$R_n(\hat{\theta},\theta) = R_n(\hat{\theta},\hat{\theta}^*) + R_n(\hat{\theta}^*,\theta) + O(n^{-1})$$

が成立するための必要十分条件は $\exp(u)$ が調和条件をみたすことを示した. ここで $r_n(\hat{\theta},\theta)$ は Kullback-Leibler 損失に関する危険関数を表している. さらに, $\exp(u)$ が超調和条件をみたせば, $\hat{\theta}$ は $\hat{\theta}^*$ によって漸近的に改良されることがいえる. また Komaki(1994) は, 縮小推定量が漸近的に改良するための微分幾何学的解釈を与えた.

3.5. 信頼集合. 1980 年代の Stein 問題の展開の一つとして信頼集合の構成問題が挙げられる. Robert-Saleh(1989) がこの問題の解説論文をかいているので参照されたい. X が $\mathcal{N}_p(\theta, I_p)$ に従うときの通常の信頼集合は $C_0(X)=\{\theta; ||\theta-X||^2 \leq c\}$ であり, c は信頼係数 $1-\alpha=\gamma$ に対して $P(\chi_p^2 \leq c^2)=\gamma$ をみたす定数である.

ある信頼集合 $C(X)$ 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,

(II) ($C(X)$ の体積)$\leq$($C_0(X)$ 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.

Brown(1966), Joshi(1967) によって $p \geq 3$ のときの $C_0(X)$ の非許容性が示され, その後, 改良を与える様々な信頼集合が提案されてきた. Berger(1980b) は, 一般化 Bayes 推定に基づいた信頼集合を導出した. これは計算が困難であることや一様な改良結果が得られていないという欠点はあるものの, θを含む確率においても体積においても著しい改善を与えることを示した. 厳密に $C_0(X)$ を改良している信頼集合を明示的に求めたのは Hwang-Casella(1982) が最初であり, positive-part Stein 推定量$\delta^+(a, X) = \max\{0, 1 - a/||X||^2\}X$ に対して

$$C(a, X) = \{\theta; ||\theta - \delta^+(a, X)||^2 \leq c\}$$

なる信頼集合を考えると, $p \geq 4$ で $0 < a \leq a_c$ なる a に対して $C(a, X)$ が $C_0(X)$ を (I) の意味で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 のみたす範囲を広げた. こうして, 一様な改良結果を保証するとともに簡便で実用的な信頼集合が得られたことになる. 一様分布, 両側指数分布, 多変量 t-分布などの球面対称な分布族への拡張や改良する信頼集合のクラスの構成が, Hwang-Chen(1986), Ki-Tsui(1985) によって, また部分空間へ縮小する問題での結果等が Casella-Hwang(1987) によって得られた.

尺度母数 (分散) が未知の場合, F-統計量に基づいた通常の信頼集合を縮小型信頼集合によって漸近的もしくは数値的に改良することが, Chen-Hwang(1988), Hwang-Ullah(1994) によって示された. 多変量 t-分布を含む球面対称分布族に対する改良結果の厳密な証明が Robert-Casella(1990) によってなされたが, その議論には正規分布が含まれておらず今後の興味深い課題として残されている.

以上述べてきた信頼集合の改良は, 体積を等しくしたまま (I) の意味で真の母数を覆う確率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むしろ (II) の意味で体積のより小さい信頼集合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同じ信頼係数を保ちながら (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.

一方, 経験 Bayes 信頼集合についての研究は Morris(1983), Casella-Hwang(1983) らによって展開されてきた. $I_C(\theta) = 1; \theta \in C, = 0; \theta \notin C$ とし, $k_0 = \exp(-c^2/2)/(2\pi)^{p/2}$ とおくと, $C_0(X)$ は損失関数

$$L(\theta, C) = k_0(C\text{の体積}) - I_C(\theta)$$

に関してミニマクスとなる. Casella-Hwang(1983) は, この損失関数に関して経験 Bayes 信頼集合を構成した. $\theta \sim \mathcal{N}_p(0, \tau^2 I)$ なる事前分布に対する Bayes 信頼集合は, $B = \tau^2/(\tau^2 + 1)$ とおくとき,

$$C_B(X) = \left\{\theta; ||\theta - BX||^2 \leq B[c^2 - p\log B]\right\}$$

で与えられ, τ, 即ち B を周辺分布から推定することによって,

$$C_{EB}(X) = \left\{\theta; ||\theta - \delta^+(p-2, X)||^2 \leq v_E(||X||)\right\}$$

なる経験 Bayes 信頼集合が得られる．ここで,

$$v_E^2(||X||) = \left(1-\frac{p-2}{c^2}\right)\left[c^2 - p\log\left(1-\frac{p-2}{c^2}\right)\right], \quad ||X|| \le c \text{ のとき},$$
$$= \left(1-\frac{p-2}{||X||^2}\right)\left[c^2 - p\log\left(1-\frac{p-2}{||X||^2}\right)\right], \quad ||X|| \le c \text{ のとき},$$

である．$C_{EB}(X)$ は, (II) の意味で $C_0(X)$ を一様に改良しているものの (I) についての解析的な結果は得られていないが, p が小さくなければ $(p \ge 5)$, 要求された信頼係数をみたしていることが数値的に示されている．経験 Bayes risk による評価が最近 He(1992) によって与えられた.

点推定において損失関数 $||X-\theta||^2$ の推定問題について述べたが, 同様な議論が信頼集合においてもなされ, $I_{C_0(X)}(\theta) = I(\theta \in C_0(X))$ の不偏推定量 γ の非許容性や許容性の結果が, Lu-Berger(1989b), Hwang-Brown(1991), Robert-Casella(1994), George-Casella(1994), Casella-Hwang-Robert(1994) らによって得られた．仮説検定においても同様な推定問題が議論され, 決定理論的結果が Hwang et al.(1992) において議論されている.

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 $\mathcal{W}_p(n,\Sigma)$ に従うとし, Σ を $\hat{\Sigma}$ によって Kullback-Leibler 損失 $\mathrm{tr}\hat{\Sigma}\Sigma^{-1} - \log|\hat{\Sigma}\Sigma^{-1}| - p$ 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_T^+ を考え, それに関する最良共変推定量がミニマクスであり, $S = TT'$, $T \in G_T^+$ なる T に対して,

$$\hat{\Sigma}^{JS} = TDT', \quad D = \mathrm{diag}(d_1, \ldots, d_p), \quad d_i = (n+p+1-2i)^{-1},$$

で与えられることを示した．しかし $\hat{\Sigma}^{JS}$ は座標系のとり方に依存するので直交不変なミニマ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 $L = \mathrm{diag}(\ell_1, \ldots, \ell_p)$ によって $S = RLR'$ と表わされるとき, $\hat{\Sigma}^{JS}$ が

$$\hat{\Sigma}^{ST} = R\mathrm{diag}(\ell_1 d_1, \ldots, \ell_p d_p)R'$$

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, $T_\Gamma T_\Gamma' = \Gamma' S \Gamma$, $\Gamma \in O(p)$, $T_\Gamma \in G_T^+$ に対して

$$\hat{\Sigma}^{TK} = \int_{O(p)} \Gamma T_\Gamma D T_\Gamma' \Gamma' d\mu(\Gamma)$$

なる推定量によって $\hat{\Sigma}^{JS}$ は改良される. $\hat{\Sigma}^{TK}$ 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クス推定量になっていることを示した. 最近 Yang-Berger(1994) は Σ の Bayes 推定に関する議論を展開している.

その他に, 共分散行列の逆行列 Σ^{-1} の推定 (Krishnamoorthy-Gupta(1989), Dey-Ghosh-Srinivasan(1990) など), 成長曲線モデルでの共分散行列の推定 (Konno(1995)), Σ の固有値の同時推定 (Dey(1988), Dey-Gelfand(1989), Jin(1993), DasGupta(1989)), Σ のコレスキー分解された行列の推定 (Eaton-Olkin(1987)), 2 つの共分散行列 Σ_1, Σ_2 の同時推定 (Loh(1991a,b)) などの議論がある.

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 Kullback-Leibler 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.

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り, おもに理論的な側面からの活発な研究がなされてきたものである. その他にも通常の推定量の縮小・拡大など修正を必要とする問題は少なくないように思われる. そのときどの程度修正すべきかの指標を与えることが重要であるが, 2 節, 3.2 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するとともに, 修正を要する推定問題をいくつか列挙してみよう.

5.1. 分散成分の推定. 繰り返し数が等しい一元配置の変量モデル

$$y_{ij} = \mu + a_i + e_{ij}, \quad i = 1, \ldots, k, \; j = 1, \ldots, m$$
$$a_i \sim \mathcal{N}(0, \sigma_A^2), \quad e_{ij} \sim \mathcal{N}(0, \sigma^2)$$

を考える. ここで $\{a_i\}$ と $\{e_{ij}\}$ は独立とし, μ, σ_A^2, σ^2 を未知母数とする. $\overline{Y}_i = \Sigma_{j=1}^m Y_{ij}/m$, $\overline{Y} = \Sigma_{i=1}^k \Sigma_{j=1}^m Y_{ij}/(mk)$, $S_1 = \Sigma_{i=1}^k \Sigma_{j=1}^m (Y_{ij} - \overline{Y}_i)^2$, $S_2 = m\Sigma_{j=1}^k (\overline{Y}_i - \overline{Y})^2$ とおくと, $\overline{Y}$, S_1,

S_2 は最小十分統計量でそれぞれ独立に

$$\begin{aligned}
&\overline{Y} \sim \mathcal{N}(\mu, (\sigma^2 + m\sigma_A^2)/(mk)),\\
&S_1 \sim \sigma^2 \chi_n^2, \quad n = k(m-1),\\
&S_2 \sim (\sigma^2 + m\sigma_A^2)\chi_{k-1}^2
\end{aligned} \tag{5.1}$$

なる分布に従う.

まず σ_A^2 の推定問題を考えよう. σ_A^2 の最小分散不偏推定量 (UMVUE) は

$$\delta_U = \frac{1}{m}\left(\frac{S_2}{k-1} - \frac{S_1}{n}\right)$$

で与えられるが, 負値を取り得るという不合理な性質をもっている. この性質を取り除く必要があるわけであるが, 2 次形式推定量の中だけで考察するのには問題がある. というのは, LaMotte(1973) が '非負の 2 次形式推定量で不偏なものは存在しない' ことを示し, Kleffe-Rao(1986) が '不偏でなくても非負の 2 次形式推定量は m を固定し $k \to \infty$ としたとき一致性をもたない' ことを示した. 従って 2 次形式推定量の枠を超えて非負な一致推定量を求めることが望まれる. 平均 2 乗誤差で推定量の良さを評価すると δ_U は

$$\delta_0 = \frac{1}{m}\left(\frac{S_2}{k+1} - \frac{S_1}{n}\right)$$

によって改良されることがわかる. この δ_0 をさらに改良するために

$$\delta_\phi = \frac{1}{m}\left\{S_2\phi\left(\frac{S_1}{S_2}\right) - \frac{S_1}{n}\right\}$$

なる縮小推定量を考えると, δ_ϕ が δ_0 を改良するための次なる条件が IERD 法から得られる.

(a) $\phi(w)$ が単調増加で, $\phi(0) = 1/(k+1)$,

(b) $\phi(w) \le \phi_0(w)$, 但し

$$\phi_0(w) = \frac{1}{n}\frac{\int_w^\infty z^{n/2}(1+z)^{-(n+k+3)/2}dz}{\int_w^\infty z^{n/2-1}(1+z)^{-(n+k+3)/2}dz}.$$

この結果から次なる推定量が得られる (Kubokawa-Saleh-Makita(1993)).

$$\begin{aligned}
\delta^{PT} &= \max\left\{\delta_0, \frac{2}{mn(n+k+3)}(S_1+S_2)\right\}\\
\delta_0^+ &= \max\{\delta_0, 0\}.
\end{aligned} \tag{5.2}$$

こうして正の改良型推定量 δ^{PT} が得られ, 我々が提案したいものである. また δ_0^+ は Klotz-Milton-Zacks(1969) により導かれたものである. $\sigma^2 = 1$, $\sigma_A^2 = 0.1$, $k = m = 3$ のときの推定量 δ_U, δ^{PT}, δ_0^+ の分布の様子が, 図 4 に描かれている. δ_U が負の範囲にも小さくない確率で分布するのに対して, δ^{PT} は $\sigma_A^2 = 0.1$ の周りに分布することがわかる. 推定量 δ_U, δ_0, δ^{PT}, δ_0^+

に対する危険関数 $E[(\delta-\sigma_A^2)^2/(\sigma^2+m\sigma_A^2)^2]$ の 50,000 回の繰り返しによるシミュレーション結果が表 1 に与えられている. δ^{PT} がかなり良い推定量であることがわかる. 繰り返しが異なる場合には最小十分統計量が完備でないから UMVUE を求めることができず, ANOVA 推定量や MINQU 推定量など様々な不偏推定量が提案されてきた. Kubokawa(1995) は IERD 法によって 通常の ANOVA 推定量を改良する正の推定量を導出している.

誤差分散 σ^2 の推定については明快な決定理論的結果が得られる. σ^2 の推定量 $\hat{\sigma}_0^2=(n+2)^{-1}S_1$ が

$$\hat{\sigma}_\psi^2=S_1\psi\left(\frac{S_2}{S_1}\right)$$

なる縮小推定量によって平均 2 乗誤差の意味で改良されるための条件として, 次なるものが IERD 法によって得られる.

(a) $\psi(w)$ が単調増加で, $\lim_{w\to\infty}\psi(w)=1/(n+2)$,

(b) $\psi(w)\geq\psi_0(w)$, 但し

$$\psi_0(w)=\frac{1}{n+k+1}\frac{\int_0^w z^{(k-1)/2-1}(1+z)^{-(n+k+1)/2}dz}{\int_0^w z^{(k-1)/2-1}(1+z)^{-(n+k+3)/2}dz}.$$

このクラスに入るものとして $\hat{\sigma}_{\psi_0}^2$, $\min\{\hat{\sigma}_0^2,(S_1+S_2)/(n+k+1)\}$ などが考えられるが, 実は $\hat{\sigma}_{\psi_0}^2$ は Portnoy(1971) によって提案された一般化 Bayes 推定量に一致しており, ミニマクス性に関する彼の予想が解決されたことになる.

Kubokawa-Saleh-Konno-Wagatsuma(1995) は, (σ^2,σ_A^2) を同時推定する問題において, 次のような Kullback-Leibler 損失関数

$$L(\hat{\sigma}^2,\hat{\sigma}_A^2,\sigma^2,\sigma_A^2)=n\left\{\frac{\hat{\sigma}^2}{\sigma^2}-\log\frac{\hat{\sigma}^2}{\sigma^2}-1\right\}+(k-1)\left\{\frac{\hat{\sigma}^2+m\hat{\sigma}_A^2}{\sigma^2+m\sigma_A^2}-\log\frac{\hat{\sigma}^2+m\hat{\sigma}_A^2}{\sigma^2+m\sigma_A^2}-1\right\}$$

を考え, IERD 法を用いることにより明解な決定理論的結果が得られることを示した. 即ち, 通常の不偏推定量 $(S_1/n,\delta_U)$ がミニマクスであり, それらの打ち切り推定量 $(\hat{\sigma}_{EB}^2,\delta_{EB})$, 但し

$$\hat{\sigma}_{EB}^2=\min\left\{\frac{S_1}{n},\frac{S_1+S_2}{n+k-1}\right\},\quad \delta_{EB}=\max(\delta_U,0),$$

が経験 Bayes 推定量として解釈されしかも不偏推定量を改良していること, また不偏推定量を改良する一般化 Bayes 推定量も IERD 法による通常のアプローチで導出できることが示された. 特に常に正値をとる改良型推定量として $(\hat{\sigma}_{EB}^2,\delta^{PT})$, 但し

$$\delta^{PT}=\frac{1}{m}\left[\max\left\{\frac{S_2}{k-1},\frac{S_1+S_2}{n+k-3}\right\}-\min\left\{\frac{S_1}{n},\frac{S_1+S_2}{n+k-1}\right\}\right],$$

を提案し, その良さを調べている. 回帰係数ベクトルの一般化最小 2 乗推定量 (GLSE) や GLS F-検定, small-area 推定問題 (6.3 節) においては, 分散成分をそれらの推定量で置き換える, いわゆる 2 段階推測手法がとられるが, その際ここで議論された分散成分の推定量が使われる.

多変量混合モデルの共分散成分の推定についても同様な問題が生ずるが, 共分散成分の半正定値推定量の導出が Amemiya(1985), Mathew-Niyogi-Sinha(1994), Remadi-Amemiya(1994) によりなされ, Calvin-Dykstra(1991) により restricted MLE を求めるアルゴリズムの構成など活発な議論が展開されている.

5.2. SN 比と非心母数の推定. 品質管理の分野において測定器や測定方法の比較に用いられる規準として SN(Signal-Noise) 比がある. 測定誤差の確率分布に正規性や等分散性を仮定できる場合には, SN 比の推定が本質的に非心 F-分布の非心母数の推定問題と同等であることが三輪 (1979) によって示された. モデル (2.1) を仮定すると非心母数は $\lambda = ||\theta||^2/\sigma^2$ と表わされ, その不偏推定量は

$$\delta_0 = (n-2)||X||^2/S - p$$

で与えられる. これは分散成分の推定と同様負値を取り得るという不合理な性質をもち, これを修正する工夫が様々な形でなされてきた (椿 (1982), Neff-Strawderman(1976)). ここでは

$$\delta_\phi = (n-2)\frac{||X||^2}{S} - \phi\left(\frac{||X||^2}{S}\right)$$

なる縮小推定量を考えよう. すると IERD 法を用いると, δ_ϕ が δ_0 を改良する条件として

(a) $\phi(w)$ が単調増加で, $\lim_{w\to\infty}\phi(w) = p$,

(b) $\phi(w) \geq \phi_0(w)$, 但し

$$\phi_0(w) = (n-2)\frac{\int_0^w \int vtf_p(vt)f_n(v)dvdt}{\int_0^w \int vf_p(vt)f_n(v)dvdt}$$

が得られる (Kubokawa-Robert-Saleh(1993)). 例えばこのクラスに入りかつ正値をとるものとして

$$\delta^{PT} = \max\left\{\delta_0, \frac{2(n-2)}{p+2}\frac{||X||^2}{S}\right\} \tag{5.3}$$

なる推定量が提案できる.

分散 σ^2 が既知のときには非心χ^2-分布の非心母数の推定になるが, これについては経験 Bayes 的アプローチが Perlman-Rasmussen(1975) により, また完全類の構成が Chow(1987) によって, MLE の非許容性が Saxena-Alam(1982) によって示された. IERD 法による条件は上の推定量において $n \to \infty$, $S/n \to \sigma^2$ とすることによって導出され, 不偏推定量 $||X||^2/\sigma^2 - p$ が $\max\{\delta_0, \frac{2}{p+2}\frac{||X||^2}{\sigma^2}\}$ によって改良されることがわかる. 最近 Shao-Strawderman(1995) は, positive-part Stein 推定量の改良において用いた論法を使って, $\max(\delta_0, 0)$ を改良する推定量の導出に成功した. 非心度の推定は重相関係数の推定にも関係しており, Muirhead(1985), Leung-Muirhead(1987) により議論されている.

5.3. 線形校正問題と統計的制御問題. 線形校正は回帰の逆推定の問題であり広く使われている統計手法の一つである. 線形校正においては独立変数 x と従属変数 y の期待値との間に線形回帰モデル $E[y|x] = \alpha + \beta x$ が仮定される. まず値が既知の標準試料 $x_1, \ldots, x_n$ を用いて,

$$y_i = \alpha + \beta x_i + e_i, \quad i = 1, \ldots, n$$

を観測し, 次に値 x_0 が未知の k 個の試料について観測を行ない

$$y_{0j} = \alpha + \beta x_0 + e_{0j}, \quad j = 1, \ldots, k$$

の値を得たとしよう. ここで e_i, e_{0j} は互いに独立に $\mathcal{N}_p(0, \sigma^2 I_p)$ に従うとし, α, β は未知の p-次元ベクトル, σ^2 は未知の誤差分散, x_i, x_0 はスカラー, y_i, y_{0j} は p-次元ベクトルとする. このときデータ y_i, y_{0j} に基づいて x_0 を逆推定する問題が線形校正 (Linear Calibration) である. x_0 の推測に必要な統計量を整理すると, x_0 の推測問題は次のモデルの x の推測問題に帰着される. y, z, T を互いに独立な確率変数とし,

$$y \sim \mathcal{N}_p(\theta, \sigma^2 I_p), \quad z \sim \mathcal{N}_p(\theta x, \sigma^2 I_p), \quad T \sim \mathcal{W}_p(n+k-3, \sigma^2 I_p) \tag{5.4}$$

なる分布に従い, θ, x, σ^2 が未知母数である. もちろん T はさらに $s = \mathrm{tr}T$ に縮められ, $s \sim \sigma^2 \chi^2_q$, $q = p(n+k-3)$, に従うことが注意されよう.

x の逆推定には 2 つの代表的な方法が知られている. 一つは古典的推定量と呼ばれるもので,

$$\delta_C = \frac{y'z}{||y||^2}$$

で与えられる. これは, $||z - yx||^2$ を最小にする x を求めることにより導かれ, 一致性をもった推定量である. $p \geq 3$ のときには δ_C の平均 2 乗誤差が有界になるものの, $p = 1, 2$ のときには 2 次のモーメントが有界でないという不都合な性質をもっている. この性質を補うため, Krutchkoff(1967) は$\Sigma_{i=1}^n \{\gamma'(y_i - \mathbf{\bar{y}}) - (x_i - \mathbf{\bar{x}})\}^2$ を最小にする γ を求めて逆回帰させることを考え,

$$\delta_I = \frac{y'T^{-1}z}{1 + y'T^{-1}y}$$

なる逆回帰推定量を導出した. δ_I のモーメントは有限であるが, 一致性をもたないことが知られている. Hoadley(1970) は $p = 1$ のとき δ_I が一般化 Bayes 推定量になっていることを示し, Kubokawa-Robert(1994) は一般に

$$\delta_B = \frac{y'z}{s + ||y||^2}$$

が proper Bayes であり, 故に許容的であることを示している. $p = 1$ のときは $\delta_I = \delta_B$ である. δ_B の非一致性を改善するために 三輪 (1985) は

$$\delta^{MW}(d) = \frac{y'z}{ds/q + ||y||^2}$$

なる一致性をもった一般化逆回帰推定量を提案し, SN 比 $||\theta||^2/\sigma^2$ が大きいところで最良な d の近似解を与えている.

線形校正問題と統計的制御問題との関係を指摘するのは興味深いことである. 上で与えた推定量は $\delta_g = g(||y||^2/s)y'z/s$ なる形に表わされる. このとき δ_g の平均 2 乗誤差は

$$E[(\delta_g - x)^2] = E\left[g^2\left(\frac{||y||^2}{s}\right)\frac{||y||^2}{s^2}\sigma^2\right] + x_0^2 E\left[\left\{g\left(\frac{||y||^2}{s}\right)\frac{y'\theta}{s} - 1\right\}^2\right]$$

と分解され, より小さい平均2乗誤差をもった推定量を求めるためには

(I) $g(t)$ を小さくすること,

(II) $E\left[\left\{g\left(||y||^2/s\right)y'\theta/s-1\right\}^2\right]$ を小さくすること,

が成り立てばよいことになる. 実は問題 (II) は統計的制御問題と呼ばれており, Takeuchi(1968), Zaman(1981), Berliner(1983), Berger-Berliner-Zaman(1982) 等により一連の結果が得られる.

古典的推定量は y が 0 に近いところで不安定になるので, その部分を縮小する必要があるであろう. そこで Kubokawa-Robert(1994) は

$$\delta_\phi=\left\{1-\phi\left(\frac{||y||^2}{s}\right)\right\}\frac{y'z}{||y||^2}$$

なる形の縮小推定量を考え, δ_ϕ が δ_Cを改良するための ϕ の条件を IERD 法を用いて導出している. そして δ_C を改良する一つの縮小推定量として

$$\delta^{KR}=\min\left\{\frac{1}{||y||^2},\ \frac{q+p-2}{s+||y||^2}\right\}y'z \tag{5.5}$$

なる打ち切り推定量を提案した. これはすべての $p\geq 1$ に対して有界なモーメントと一致性をもった望ましい推定量である.

多変量線形校正が Brown(1982), Nishii-Krishnaiah(1988), Fujikoshi-Nishii(1986) などによって議論されているが, 小標本での議論では解決すべき多くの問題が残されているようである. 線形校正についての優れた総合報告が Osborne(1991) により与えられているので参照されたい.

5.4. 制約された母数空間での推定. 母数空間に制約が入っているときの推定問題として最も単純なものは, 正規分布の正の平均の推定であろう. X が $\mathcal{N}(\mu,1)$ に従う確率変数であるとき, $\mu>0$ なる制約が課せられておりこの正の平均 μ を推定する問題である. 通常の推定量 X はミニマクスであるが, 明らかに不都合を生ずるので, $\delta_\phi=X-\phi(X)$ なる縮小推定量を考える必要がある. IERD 法を用いると δ_ϕ が X を改良するための条件として

(a) $\phi(w)$ が単調増加で $\lim_{w\to\infty}\phi(w)=0$,

(b) $\phi(w)\geq\phi_0(w)$, 但し

$$\phi_0(w)=\frac{\int_{-\infty}^{w}ze^{-z^2/2}dz}{\int_{-\infty}^{w}e^{-z^2/2}dz}$$

が得られる. $\phi_T(w)=\min(0,w)$, $\phi_0(w)$ は (a), (b) の条件をみたし, Stein 型推定量 $\delta^{ST}=\max(X,0)$, BBZ 型推定量 $\delta^{BZ}=X-\phi_0(X)$ が得られる. δ^{BZ} は事前分布 $d\mu I_{[\mu>0]}$ に対する一般化 Bayes 推定量である.

3.2 節で James-Stein 推定量の改良が 2 節の分散の推定問題に関連していることを述べたが, Rukhin(1992b) は $n\to\infty$, $||\theta||^2/(p\sigma^2)\to\mu$, $p\to\infty$ なる漸近理論において, いずれの問題もここで扱われている正規分布の正の平均の推定問題と同等になることを証明した. 即ち 2 つの問題は漸近的には同じものであることが示されたのである.

母数空間が両側から制約されている場合, 即ち $\mu \in [-m, m]$, $m > 0$, のときには, 推定量 X はミニマクスでさえなくなってしまう. Casella-Strawderman(1981) は, 両点 $-m$, m に確率をもった事前分布に対する Bayes 推定量が

$$\delta_B(m) = m \ \tanh(mX) = m\frac{e^{mX} - e^{-mX}}{e^{mX} + e^{-mX}}$$

で与えられることを示し, $m < 1.05$ のときには $\delta_B(m, X)$ がミニマクスで, MLE

$$\delta^{ML} = XI_{[|X|<m]} - mI_{[X \leq -m]} + mI_{[X \geq m]}$$

や不偏推定量 X を改良するという興味深い結果を証明した. Bickel(1981) は, m についての 2 次漸近ミニマクス性を, また Johnstone-MacGibbon(1992) はポアソン分布の場合を議論している.

Moors(1981) は, 制約された母数の推定問題において, 対称性を考慮に入れると更なる改善が得られることを示した. 例えば, 確率 p で表 (H) のでるコイン投げ実験において $0.2 < p < 0.8$ なる情報が事前に仮定されているとしよう. 非現実的ではあるが簡単のため, 1 回の試行だけで p を推定するとし, 表がでたら $\hat{p}(H) = z$ で推定し, 裏 (T) がでたら対称的に $\hat{p}(T) = 1 - \hat{p}(H) = 1 - z$ で p を推定するものとする. このとき 2 乗損失 $(\hat{p} - p)^2$ による危険関数は,

$$\begin{aligned} MSE(p, \hat{p}) &= p(z - p)^2 + (1 - p)((1 - z) - p)^2 \\ &= z^2 + 2(2p(1 - p) - 1)z + 1 - 3p(1 - p) \end{aligned}$$

となり,

$$z > \max_{0.2<p<0.8}\{1 - 2p(1 - p)\} = 0.68$$

のときには MSE はすべての p に対して, z に関して単調増加することがわかる. 通常は $p < 0.8$ より, 表がでれば $\hat{p}(H) = 0.8$ で推定されるであろうが, 対称性を考慮にいれることにより, $\min(\hat{p}(H), 0.68) = 0.68$ で推定した方が MSE をより小さくできることになる. Moors(1981) は, こうした議論を回帰モデル $y_i = \alpha + \beta x_i + e_i$, $0 < \beta < 1$, の母数推定等にも適用した.

医薬生物実験, 農事試験などの分野において, 平均や分散の間に順序制約が想定できる場合がある. $Y_1, \ldots, Y_k$ を互いに独立な確率変数で $Y_i \sim \mathcal{N}(\mu_i, \sigma^2/n_i)$ なる分布に従い, 平均の間に $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$ なる順序制約が仮定できるとしよう. このとき μ_i の非制約 (不偏) 推定量 Y_i は明らかに非許容的で, Lee(1981) は μ_i の MLE が Y_i を平均 2 乗誤差の意味で改良することを示した. $(\mu_1, \ldots, \mu_k)$ の MLE は $\Sigma_{i=1}^k (Y_i - \theta_i)^2 n_i$ を$\theta_1 \leq \theta_2 \leq \ldots \leq \theta_k$ なる制約のもとで最小化する $(\theta_1, \ldots, \theta_k)$ として与えられることが知られている. これは isotonic 回帰推定量と呼ばれ, μ_i の MLE は

$$\hat{\mu}_i = \min_{s \geq i} \max_{r \leq i} \frac{\Sigma_{j=r}^s Y_j n_j}{\Sigma_{j=r}^s n_j}$$

によって与えられる. これも制約された母数空間への一種の縮小推定量ととらえることができるが, 一般化 Bayes 推定量など明らかにすべきことが残されている. 最近, Hwang-Peddada(1994) は, 順序制約下での信頼区間について決定論的結果を与えたので参照されたい.

もう少し曖昧な制約 $\mu_1 \leq \mu_i$, $i = 2, \ldots, k$ しか仮定できないときには, 最小の平均 μ_1 の推定問題は簡単になり, IERD 法によって, Y_1 を改良していてかつ一般化 Bayes 推定量を含んだもののクラスを構成することができる (Kubokawa-Saleh(1994)). その他, 制約された母数空間での推測については Barlow-Bartholomew-Bremner-Brunk(1972), Robertson-Wright-Dykstra(1988) などに詳しく書かれている. また順序制約下での Stein 型縮小推定量に関する議論が Chang(1982), Sengupta-Sen(1991) により与えられている. $k = 2$ の場合の μ_1 または μ_2 の推定に関する決定論的研究については Blumenthal-Cohen(1968a,b), Cohen-Sackrowitz(1970) によってなされた.

母数に等号制約が入った場合曲がったモデルが生ずるが, その推測理論の微分幾何学的構造が Amari(1982) などにより研究されてきた. 変動係数一定の問題やナイル問題はその代表的なものであり, いずれも適当な群構造を当てはめることにより, 補助統計量が最大不変量として解釈されて最良共変推定量が求まり, それが MLE を改良することがわかる (Kariya(1989)). 前者は (2.1) のモデルにおいては, $\theta'\theta/\sigma^2 = const.$ なる制約の下で μ を推定する問題として表されるが, 最良共変推定量の具体的な形の導出や拡張したモデルでの議論が Kariya-Giri-Perron(1988), Marchand(1994) らによって与えられた.

分散の異なるいくつかの正規母集団の共通平均の推定も, 平均がすべて等しいという等号制約が埋め込まれた問題としてとらえることができる. 例えば 2 標本の場合,

$$\overline{X} \sim \mathcal{N}(\mu, \sigma_1^2/n_1), \quad S_1 \sim \sigma_1^2 \chi^2_{n_1-1}$$
$$\overline{Y} \sim \mathcal{N}(\mu, \sigma_2^2/n_2), \quad S_2 \sim \sigma_2^2 \chi^2_{n_2-1}$$

なるモデルに縮約される. 共通平均 μ の MLE を陽に求めることができないため, $\overline{X}$ と $\overline{Y}$ をそれらの精度の比で内分した結合推定量

$$\hat{\mu}^{GD} = \frac{cS_1^{-1}}{cS_1^{-1} + S_2^{-1}}\overline{X} + \frac{S_2^{-1}}{cS_1^{-1} + S_2^{-1}}\overline{Y}$$

が考えられる. この 2 段階の一般化最小 2 乗推定量 (GLSE) は, 特に Graybill-Deal(1959) 推定量と呼ばれ, $\hat{\mu}^{GD}$ が $\overline{X}$ もしくは $\overline{Y}$ を改良するための条件やその拡張等が Brown-Cohen(1974), Cohen-Sackrowitz(1974), Bhattacharya(1984) らによって議論された. この問題はまた, 古くから研究されてきた, 釣合型不完備ブロック計画におけるブロック間情報の回復問題とも関連している (Yates(1940), Seshadri(1963), Shah(1964), Stein(1966)). Kubokawa(1987) は, 一つの許容的なミニマクス推定量を導出したが, $\hat{\mu}^{GD}$ の許容性は未解決の問題として残されたままである.

5.5. 選抜後の推定問題. スポーツの試合では, 予選での成績が 1 位の選手が決勝ではかんばしくないことがしばしば起こる. これは, 予選で 1 位の選手に対してはその成績を本人の実力と判断してはいけないことを暗示している. これが選抜後の推定の問題である. 簡単のために, $X = (X_1, \ldots, X_p)'$, $\theta = (\theta_1, \ldots, \theta_p)'$ に対して, $X \sim \mathcal{N}_p(\theta, I_p)$ とし, 平均 θ_i の最も大きいものを選ぶ問題を考えよう. このとき

$$X_{i^*} = \max_{1 \leq i \leq p} X_i$$

となる i^* を選べばよい. しかし選ばれた i^* の平均 θ_{i^*} の推定量を X_{i^*} としてよいであろうか. θ_{i^*} は

$$\theta_{i^*} = \sum_{i=1}^{p} \theta_i I(X_i = \max_j X_j)$$

と表現でき, $\theta_1 = \cdots = \theta_p = \theta_0$ のときには$\theta_{i^*} = \theta_0$で, $c_p > 0$ に対して

$$E[X_{i^*}] = \theta_0 + c_p$$

となり, X_{i^*} は上方に偏った推定量となっている. 従って X_{i^*} は下方へ縮小する必要があり, 竹内 (1980), Dahiya(1974), Cohen-Sackrowitz(1982) らによって様々な方法が提案された. 特に Venter(1988) は, 標準正規分布の分布関数 Φ と密度関数 ϕ に対して, X_{i^*} のバイアスが

$$E[X_{i^*} - \theta_{i^*}] = \sum_{i=1}^{p} \int z\phi(z) \prod_{j \neq i} \Phi(z + \theta_i - \theta_j) dz$$

で与えられることを示し, このバイアスの推定量を, 定数λに対して

$$b(X, \lambda) = \lambda \sum_{i=1}^{p} \int z\phi(z) \prod_{j \neq i} \Phi(z + \lambda(X_i - X_j)dz$$

とし, $b(X, \lambda)$ だけ縮小した推定量

$$\delta_{i^*}^{VT} = X_{i^*} - b(X, \lambda)$$

を提案してバイアスとリスクの挙動を調べている. しかし, θ_{i^*} の不偏推定量が未だ求められておらず今後の課題になっている.

Hwang(1993) は, この問題を, $\theta_i = \mu + a_i$, $a_i \sim \mathcal{N}(0, \sigma_A^2)$ なる変量モデルの枠組の中で考察し, 経験 Bayes 推定量

$$\delta_{i^*}^{EB} = \overline{X} + \left[1 - \frac{p-3}{\Sigma(X_j - \overline{X})^2}\right]^+ (X_{i^*} - \overline{X}), \quad \overline{X} = p^{-1} \sum_{i=1}^{p} X_i$$

が X_{i^*} を改良することを証明した.

指数分布に対しても同様な問題が考えられるが, この場合には UMVUE やミニマクス推定量などが明示的に導出される (Sackrowitz-SamuelCahn(1984)). Vellaisamy(1992) はガンマ分布のときに決定論的考察を行った.

5.6. その他の縮小・拡大など修正を要する推定問題. 上で述べてきた問題の他に, 何らかの縮小・拡大などの修正を要する推定問題についてここで簡単に紹介する.

[1] **多重共線性.** 線形回帰モデル $y = X\beta + \varepsilon$ において, 説明変数間に多重共線性が存在するときには, 通常の最小 2 乗推定量$\hat{\beta} = (X'X)^{-1}X'y$ は不安定なものになってしまう. そこで, Hoerl-Kennard(1970) はリッジ回帰推定量 $\hat{\beta}^R(k) = (X'X + kI)^{-1}X'y$, $k > 0$ を提案した.

しかし, $\hat{\beta}^R(k)$ は $\hat{\beta}$ を一様には改良しないので, k を適当に推定することにより $\hat{\beta}^R(\hat{k})$ が $\hat{\beta}$を一様に改良するように $\hat{k}$ を定めることができる. しかし, Casella(1980) により指摘されているように, $\hat{\beta}^R(\hat{k})$ は再び多重共線性の影響を受けて不安定性が生じてしまう. こうして, 両方の性質を合わせもつ推定量を得ることができない. 両方の性質を適当に調整する基準とその基準に沿った新たな縮小推定量の提案が望まれる. Hill-Judge(1990) は, 縮小推定量の危険関数の不変推定量に基づいて取り除く変数を決定する方法を提案している.

Casella(1985) は, 安定性とミニマクス性を同時にみたすための, 計画行列 X の条件を導出した. P を $p \times p$ 直交行列とし,

$$P'X'XP = D_\lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p), \quad \lambda_1 \geq \cdots \geq \lambda_p$$

とするとき,

$$\hat{\beta}^{GR}(K) = P(D_\lambda + K)^{-1}P'X'Y, \quad K = \mathrm{diag}(k_1, \ldots, k_p)$$

なる一般化リッジ回帰推定量を考え, 安定性を測る基準として, 条件数 (Condition Number)

$$\kappa[\hat{\beta}(K)] = \max_{1\leq i\leq p}(\lambda_i + k_i) / \min_{1\leq i\leq p}(\lambda_i + k_i)$$

を定義し, $\kappa[\hat{\beta}(K)] < \lambda_1/\lambda_p$ のとき条件数を改良するということにした. 誤差分散 σ^2 の不偏推定量 $\hat{\sigma}^2$ と定数 a_i に対して$\hat{k}_i = a_i\hat{\sigma}^2/\hat{\beta}'X'X\hat{\beta}$ とするとき, $\hat{\beta}^{GR}(\hat{K})$ がミニマクスでかつ $\kappa[\hat{\beta}(\hat{K})] \leq \lambda_1/\lambda_p$ をみたすための必要十分条件は

$$\sum_{i=1}^{p-1} \lambda_i^{-2} > (\lambda_1\lambda_p)^{-1}$$

であることが証明された. 従って計画行列 X がこの不等式をみたすとき, 上記の意味において安定性とミニマクス性を合わせもつリッジ回帰推定量が得られる.

[2] 帰無仮説の方向への縮小. 確率変数 X が, σ^2 が既知の正規分布 $\mathcal{N}(\mu, \sigma^2)$ に従い, μ が帰無仮説 $H_0 : \mu = \mu_0$ に近いことが予想される場合, X を μ_0 の方向へ縮小した推定量

$$\hat{\mu}(k) = k(T)X + \{1 - k(T)\}\mu_0, \quad T = (X - \mu_0)/\sigma$$

が考えられる. ここで縮小関数 $k(T)$ の決め方にはいくつかの方法が提案されている. Hirano(1970) は, 予備検定推定量ととらえて, $k_1(t) = I(|t| \geq z_{\alpha/2})$ とし, 赤池情報量基準 (AIC) を最小にするα を決めることを提案した. 但し $z_{\alpha/2}$ は $\mathcal{N}(0,1)$ の $100(1-\alpha/2)$%点を表している. Thompson(1968) は $k_2(t) = t^2/(1+t^2)$ を提案した. これは, 推定量 $cX + (1-c)\mu_0$ の危険関数を最小化する c が $c = \{(\mu-\mu_0)^2/\sigma^2\}/\{1 + (\mu-\mu_0)^2/\sigma^2\}$ で与えられるので, μ をその推定量で置き換えたものである. Inada(1984) は両方のアイデアを取り入れて

$$k_3(t) = d^*I(|t| < z_{\alpha/2}) + I(|t| \geq z_{\alpha/2})$$

とし, (α, d^*) をミニマクス・リグレット基準によって決定することを考えた. また別の基準が Hawkins-Han(1989) により提案され, 尺度母数の場合の議論が Kambo-Handa-AlHemyani(1990)

により与えられた. 理論上は X 自身が許容的であり X を一様に改良する $\hat{\mu}(k)$ を求めることはできないが, 応用上は, 帰無仮説に近いと予想される場合があり, そうした状況で有用な推定方式であると考えられる.

Ghosh-Saleh-Sen(1989) は, 回帰モデル $y = X_1\beta_1 + X_2\beta_2 + \varepsilon$ において, $H_0 : \beta_2 = 0$ が予想されるときのβ_1 の経験 Bayes 推定量

$$\hat{\beta}_1^{EB} = \hat{\beta}_1 + \left(1 - \frac{c}{F}\right)(\tilde{\beta}_1 - \hat{\beta}_1)$$

を導出し, これがミニマクス推定量 $\tilde{\beta}_1$ を改良するための c に関する条件を求めた. ここで, F は, $H_0 : \beta_2 = 0$ vs. $H_A : \beta_2 \neq 0$ に対する尤度比検定統計量に対応するもので,

$$F = \frac{\tilde{\beta}_2' X_2'(I - X_1(X_1'X_1)^{-1}X_1')X_2\tilde{\beta}_2}{||y - X_1\tilde{\beta}_1 - X_2\tilde{\beta}_2||^2}$$

で与えられる. $\hat{\beta}_1$ は H_0 の下での制限付き最小 2 乗推定量, $(\tilde{\beta}_1, \tilde{\beta}_2)$ は (β_1, β_2) の制限なしでの最小 2 乗推定量を表している. 経験 Bayes 推定量は, $\tilde{\beta}_1$ と $\hat{\beta}_1$ とを折衷した形をしており, H_0 が正しそうなときには $\hat{\beta}_1$ の方へ傾き, 反対に H_0 が疑わしいときには $\tilde{\beta}_1$ の方に傾くことがわかる.

[3] 判別分析. 多変量判別分析の問題において, 次元が高くなると共分散行列の逆行列の推定が不安定になるため, 線形及び 2 次判別関数に基づく分析があまり好ましくないものになってしまうことが, Peck-Van Ness(1982), Friedman(1989), 松田-吉村-藤本 (1990) らにより指摘された. 例えば, 共通な共分散行列をもった 2 つの正規母集団からのデータが得られている場合を考えよう.

$$\Pi_i : X_{ij} \sim \mathcal{N}_p(\theta_i, \Sigma), \quad j = 1, \ldots, n_i, \ i = 1, 2.$$

それぞれの標本平均を $\overline{X}_1$, $\overline{X}_2$ とし, Σ の不偏推定量を S とすると, 線形判別関数は,

$$h(x; \overline{X}_1, \overline{X}_2, S) = \{x - 2^{-1}(\overline{X}_1 + \overline{X}_2)\}' S^{-1}(\overline{X}_1 - \overline{X}_2)$$

で与えられ, $h(x; \overline{X}_1, \overline{X}_2, S) > 0$ なら Π_1 , $h(x; \overline{X}_1, \overline{X}_2, S) \leq 0$ なら Π_2 と判別する. この関数において, p が大きくなれば S^{-1} は不安定になり, もし $p > n_1 + n_2 - 2$ なら S^{-1} は存在しなくなる. Haff(1986), Dey-Srinivasan(1991) は, この線形判別関数の係数ベクトルの推定問題に注目し, $(\overline{X}_1, \overline{X}_2, S)$ に基づいて $\eta = \Sigma^{-1}(\theta_1 - \theta_2)$ を推定する問題を考察し, Σ^{-1} の縮小推定についての結果を適用して, $c(S + u\phi(u)I)^{-1}(\overline{X}_1 - \overline{X}_2)$, $u = 1/\mathrm{tr}S^{-1}$ なる推定量が $cS^{-1}(\overline{X}_1 - \overline{X}_2)$ なる形の通常の推定量を改良するための条件を与えた. また Rukhin(1992a) は通常の推定量が許容的となるための c に関する条件を求めている.

[4] 誤判別率の推定. 判別分析を行ったときの誤判別率を推定する際, 経験分布関数に基づいた見かけ上の推定量や MLE 型推定量は真の誤判別率を過小に推定してしまう傾向があることが知られている. 例えば, 上記 (5.6 節 [3]) の判別問題において, x が母集団 Π_1 からとられているのに Π_2 に属すると誤って判断してしまう確率は

$$\begin{aligned}\lambda(\mu_1, \Sigma, \overline{X}_1, \overline{X}_2, S) &= P[h(x; \overline{X}_1, \overline{X}_2, S) \leq 0 | x \sim \mathcal{N}_p(\mu_1, \Sigma)] \\ &= \Phi\left\{-\frac{\{\mu_1 - 2^{-1}(\overline{X}_1 + \overline{X}_2)\}' S^{-1}(\overline{X}_1 - \overline{X}_2)}{\sqrt{(\overline{X}_1 - \overline{X}_2)' S^{-1}\Sigma S^{-1}(\overline{X}_1 - \overline{X}_2)}}\right\}\end{aligned}$$

で与えられる．但し Φ は標準正規分布の分布関数を表している．この実際の誤判別率を $\overline{X}_1$, $\overline{X}_2$, S に基づいて推定するわけで, $\hat{\mu}_1 = \overline{X}_1$, $\hat{\Sigma} = S$ を代入した MLE 型推定量

$$\hat{\lambda}^{ML} = \Phi(-D/2), \quad D^2 = (\overline{X}_1 - \overline{X}_2)' S^{-1} (\overline{X}_1 - \overline{X}_2)$$

が通常考えられるが, 実際の誤判別率を過小推定してしまう．McLachlan(1974) は, Okamoto(1963) による$\lambda(\mu_1, \Sigma, \overline{X}_1, \overline{X}_2, S)$ の期待値の漸近展開に基づいて, バイアス修正を施した推定量を提案した．Efron(1979) は, ブートストラップ法を導入し, ノンパラメトリックなモデルの下で, 経験分布に基づいた見かけ上の推定量をバイアス修正した推定方式を提案した．ジャックナイフ法や交差検証法を含めていままで提案されてきたバイアス修正推定量の数値的及び理論的比較等の研究成果が, Konishi-Honda(1990), 小西-本田 (1992) で与えられいる．

[5] 不良率や信頼度の推定． 正規分布 $\mathcal{N}(\mu, \sigma^2)$ に従う独立な確率変数 $X_1, \ldots, X_n$ に基づいて, $P[X_i > c]$ なる片側確率 (裾の確率) を推定する問題は, 例えば品質管理において製品の寸法等の品質特性が上側規格限界 c をみたすことを要求されているときの不良率の推定に現れる．Brown-Rutemiller(1973) はその UMVUE と MLE の数値的比較を行い, 藤野 (1987) は線形回帰モデルへの拡張と MLE 型推定量の漸近展開によるバイアス修正推定量の有効性を示し, Peszek-Rukhin(1993) は一般化ベイズ推定量の導出とその許容性を議論した．2 つの独立な確率変数 X と Y に対する $P[Y < X]$ の推定問題が, Eris-Geisser(1971), Reiser-Guttman(1987) により議論されている．

指数分布 $\sigma^{-1}\exp(-x/\sigma)$ においては $P[X > t] = \exp(-t/\sigma)$ は信頼度と呼ばれ, この推定問題についての議論も活発になされてきた (Zacks-Even(1966), Chiou(1993) など)．2 母数指数分布に対する信頼度の Bayes 推定と許容性については, Pierce(1973), Varde(1969), Rukhin-Ananda(1989) で議論されている．

[6] 分布関数の分位点の推定． 正規分布 $\mathcal{N}(\mu, \sigma^2)$ の分位点は $\theta = \mu + a\sigma$ (a は定数) の形に表されるが, Zidek(1969, 71) は θ の通常のミニマクス推定量が打ち切り推定量により改良されることを証明した．2 母数指数分布における分位点の推定も活発に研究され, 決定理論的結果が Rukhin-Strawderman(1982), Rukhin-Zidek(1985), Rukhin(1986), Sharma-Kumar(1994) により与えられている．

[7] 因子分析における不適解． 次のようなランダムな共通因子ベクトル $\mathbf{f}$ をもった因子分析モデルを考えるとき, 独自分散行列 $\Psi = \mathrm{diag}(\psi_1, \ldots, \psi_p)$ の推定に対して, 反復計算の結果, 不適解がしばしば生ずることが知られている (丘本 (1986))．

$$\mathbf{x}_i = \mu + \Lambda \mathbf{f}_i + \mathbf{e}_i, \quad i = 1, \ldots, n$$

ここで, $\mathbf{x}_i \in \mathbf{R}^p$ は観測ベクトル, $\mu \in \mathbf{R}^p$ とΛ ($p \times k$ 行列) は未知母数, $\mathbf{f}_i \in \mathbf{R}^k$, $\mathbf{e}_i \in \mathbf{R}^p$ は独立な確率変数で, $\mathbf{f}_i \sim \mathcal{N}_k(0, I)$, $\mathbf{e}_i \sim \mathcal{N}_p(0, \Psi)$ と仮定する．このとき,

$$S = n^{-1} \sum_{i=1}^{n} (\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})',$$
$$\Sigma = Cov(\mathbf{x}_i) = \Lambda\Lambda' + \Psi$$

とおくと, Ψ と Λ の MLE は

$$q(\Psi,\Lambda)=\mathrm{tr}(S\Sigma^{-1})-\log|S\Sigma^{-1}|$$

を最小にする解として与えられる. この近似解を求めるためのアルゴリズムがいくつか提案されたが, 反復計算の結果, Ψ の推定量として非正の解 (不適解) が得られることがしばしば起こり, この問題を解決するためペナルティー関数の利用 (Lee(1980)) やラグランジェ未定係数法 (Lee(1981)) などが考えられてきた. そもそも Ψ の MLE は $\hat{\Psi}^{ML}\geq 0$ であり, 竹内 (1986) は, 最尤解の存在性とその構造を明らかにし, $\hat{\psi}_i^{ML}=0$ となる状況等について詳しく調べた. $\Psi>0$, $\mathrm{rank}\Lambda=k$ であっても $\hat{\psi}_i^{ML}=0$ となることがありうるので, 何らかの方法で $\hat{\Psi}>0$ となる推定量を導出することが望まれる. Akaike(1987) は, 上記の因子分析モデルが Bayes モデルとしてとらえられることを指摘した. また Martin-McDonald(1975) と同様な事前分布を想定し,

$$q^*(\Psi,\Lambda)=\mathrm{tr}(S(\Lambda\Lambda'+\Psi)^{-1})-\log|S(\Lambda\Lambda'+\Psi)^{-1}|+\delta\ \mathrm{tr}\Lambda'\Psi^{-1}\Lambda$$

を最小化する解を求めることを提案した. 右辺の第 2 項は, ψ_i のどれかが 0 に近づくにつれて大きくなり, $\hat{\psi}_i=0$ となる解が生じないようなペナルティーとしてはたらいていることがわかる. このように, ここでの考え方は, 推定量をどこかに縮小するというよりは, 0 から遠ざけるように修正を施している点に注意されたい. δ は, そのペナルティーの程度を調整している超母数で, 経験 Bayes 的にデータから推定される必要があるだろう. また Bayes 解もしくは経験 Bayes 解が Ψ の推定量としてどのような性質をもっているのかについて, 一致性や有効性などの点から議論することも残されている.

6. 応用例

ここでは, James-Stein 推定量やそれに関連した経験 Bayes 推定量の応用例, また分散成分モデルの適用例を取り上げ, 実際のデータ解析においてどのような場面で有効に使われているのかを考えてみよう.

6.1. 打撃能力の推定. Efron-Morris(1975) は, 野球選手の打撃能力の推定に対して James-Stein 推定量の適用を行った. 大リーグの選手 18 人について, 1970 年のシーズン中の最初の 45 打数の打率 $Y_1,\ldots,Y_{18}$ に基づいてシーズン終了時の打率 $p_1,\ldots,p_{18}$ を求める問題を考えよう. 但し, 最終打率 p_i をその選手の打撃能力とみなしている. 各 i に対して nY_i は 2 項分布 $\mathcal{B}in(n,p_i)$ に従うので, p_i の MLE は $\hat{p}_i^{ML}=Y_i$ である. しかし, MLE で p_i を推定することには疑問が残る. 例えば, クレメンテは $Y_1=0.400$ と高打率であるが, シーズン終了時までこの高打率を維持することは常識的には考えにくく, 少しは割り引いて推定した方がよさそうである. 即ち, 最初の 35 試合の成績はたまたま調子が良かった, あるいは調子が悪かったのかもしれないため, そうした標本誤差を考慮に入れた推定方法が望まれるのである.

Efron-Morris は Y_i に 2 項分布の分散安定化変換

$$X_i=\sqrt{n}\ \arcsin(2Y_i-1),\quad n=45$$
$$\theta_i=\sqrt{n}\ \arcsin(2p_i-1),$$

を行うことによって, $k=18$ に対して,

$$X_i \sim \mathcal{N}(\theta_i, 1), \quad i=1,\ldots,k$$

とし, $\theta_1,\ldots,\theta_k$ の同時推定の問題に縮小推定量を適用した. それは, 各 X_i を全体の平均 $\overline{X}=k^{-1}\Sigma_{i=1}^k X_i$ の方向へ縮小した Lindley の推定量

$$\hat{\theta}_i^L = \overline{X} + \left\{1 - \frac{k-3}{\Sigma_{i=1}^k (X_i - \overline{X})^2}\right\}(X_i - \overline{X}),$$

で与えられるが, これを逆変換することによって, 割り引いた推定量$\hat{p}_i^L$ を求めている. 各選手の MLE $\hat{p}_i^{ML}$, Lindley 推定量 $\hat{p}_i^L$, 真の打率 p_i の値が図 5 に示され, 全体の平均の方向へ縮小されている様子がみてとれる. 18 人中 16 人については, MLE より Lindley 縮小推定量の方が最終打率に近い値を与えている.

6.2. 死亡率及び死亡指標の推定. Tsutakawa-Shoop-Marienfeld(1985), 丹後 (1988), Clayton-Kaldor(1987), Manton et al.(1989) らは, 胃ガンの都市別死亡率及び死亡指標の推定に経験 Bayes 推定量が有効であることを示した. ミズリー州の 84 都市のうち i 番目の都市の人口を n_i, 胃ガンの観測死亡数を d_i, 観測死亡率を $\hat{p}_i = d_i/n_i$ で表す. これらの値が上記の文献で与えられているので, それに基づいて $(n_i, \hat{p}_i)$ をプロットしたのが図 6 で, 人口の小さい都市の死亡率のバラツキがかなり大きくなっていることがわかる (但し, 図 6 は $\log n_i$, $100,000\hat{p}_i$のスケールで描かれている). そこで, こうした観測誤差によるバラツキを考慮した安定な推定手法の導出が望まれる. i 番目の都市の真の死亡率を p_i とするとき, $k=84$ に対して, 死亡数 d_i がポアソン分布

$$d_i \sim \mathcal{P}o(n_i p_i), \quad i=1,\ldots,k$$

に従うことが仮定される. 丹後 (1985) は p_i の事前分布にガンマ分布 $\mathcal{G}(\alpha,\beta)$ を仮定し, p_i の経験 Bayes 推定量を求めている. 実際, p_i の Bayes 推定量は,

$$\hat{p}_i^B(\alpha,\beta) = E[p_i|d_i,\alpha,\beta] = \frac{d_i+\beta}{n_i+\alpha}$$

であり, 超母数 α, β を d_i の周辺密度

$$f(d_i|\alpha,\beta) = \frac{\Gamma(g_i+\beta)}{\Gamma(\beta)d_i!}\left(\frac{\alpha}{\alpha+n_i}\right)^{\beta}\left(\frac{n_i}{\alpha+n_i}\right)^{d_i}$$

から最尤法によって推定し, その推定値 $\hat{\alpha}$, $\hat{\beta}$ を Bayes 推定量に代入して

$$\hat{p}_i^{EB} = \hat{p}_i^B(\hat{\alpha},\hat{\beta})$$

なる経験 Bayes 推定量が導かれる. ミズリー州の都市別胃ガン死亡率の経験 Bayes 推定値を求めて, $(n_i, \hat{p}_i^{EB})$ をプロットしたのが図 7 である. 図 6 で示された小標本での観測死亡率のバラツキが経験 Bayes 推定を行うことによって取り除かれていることがわかる.

死亡指標の大小に応じて色分けして疾病地図を描く際, 観測死亡率のバラツキの影響が問題となる. 代表的な死亡指標として, 年齢調整死亡率 (DAR, Directly Age-Adjusted Mortality Rate)

$$DAR_i = \sum_{j=1}^{a} \frac{d_{ij}}{n_{ij}} \frac{N_j}{N}, \quad i = 1, \ldots, k, \; j = 1, \ldots, a$$

が知られている. 但し, n_{ij}, d_{ij} は i 地域, j年齢階級の人口及び観測死亡数, N_j は標準人口集団の j年齢階級人口, $N = N_1 + \cdots + N_a$ である. i 地域, j年齢階級の真の死亡率を p_{ij} とすると, DAR_i は

$$\theta_i = \sum_{j=1}^{a} p_{ij} \frac{N_j}{N}$$

の推定量となっているわけであるが, n_{ij} が小さいときの d_{ij} のバラツキが大きいため, わずかな死亡数の変動が死亡指標 DAR_i を大きく変化させてしまい, ときには異常な高値を示す欠点があることが Tsutakawa, 丹後らによって指摘された. 丹後 (1985) は, DAR_i にとって代わるものとして, 各 p_{ij}に上述の $\hat{p}_{ij}^{EB}$ を代入した経験 Bayes 年齢調整死亡率

$$EBDAR_i = \sum_{j=1}^{a} \frac{d_{ij} + \hat{\beta}_j}{n_{ij} + \hat{\alpha}_j} \frac{N_j}{N}$$

を提案し, 1975 年日本における職業別年齢調整死亡率の推定, 1980 年東京都における胃ガンの都市別年齢調整死亡率の推定に適用している.

小標本でのバラツキが大きいデータの解析に, 縮小推定の考え方が有効であることを示したのは, Efron-Morris(1975) が最初であろう. 彼らは, エルサルバドルの 36 都市のトキソプラズマ症の発生率の推定に縮小推定量の適用を試みた. 図 8 において上側にプロットされているのが, 各都市の発生率を全国平均と比較して表した値 x_i とその標準偏差である. 下側にプロットされているのが, 各都市の発生率の James-Stein 推定量の値である. James-Stein 推定量は, 標準誤差の大きい x_i に対しては大幅に縮小され, 標準誤差の小さい x_i に対してはあまり縮小されていないことがわかる.

6.3. small-area の推定問題. small-area 推定という問題が, 合衆国を中心に近年 活発に議論され研究されてきた. small-area (小さい地域) に対する通常の推定量は, その地域における標本数が少ないためしばしば非常に大きな推定誤差を生じてしまう. そこで small-area に対する推定の精度を高めるため, その関連した地域からの情報を使用する様々な統計的方法が提案されてきた. この問題の詳しい内容と最近の発展については, Fay-Herriot(1979), Prasad-Rao(1990), Ghosh-Rao(1994) を参照されたい.

ここでは Battese-Harter-Fuller(1988) が扱った small-area に対する穀物 (とうもろこしと大豆) の作付面積の推定問題について簡単に紹介する. アイオワ州の北部中央地域の 12 の郡 (county) について各郡が約 250h の segment に細分されている. i 番目の郡における segments の個数を N_i とし, そのうち n_i 個の segments が標本抽出される. n_i は 1~5 程度で, 12 個の郡全体で 36 個の segments が標本抽出された. 標本抽出された各 segment に対して, 連邦農務

省 (USDA) は直接農家にインタヴューすることによってその segment におけるとうもろこし (または大豆) の作付面積についてのデータをとった．一方, LANDSAT からの観測により, 約 0.45h の pixel(picture element, 識別する単位) に対していずれの穀物が作付けされているのかが識別され, 12 の郡すべてにわたってこうした衛星データがとられた．連邦農務省の調査データと衛星データを用いて, 各 segment のとうもろこし (または大豆) 作付面積の郡平均を推定するのが目的となる．

各 segment に対して, 連邦農務省の調査データと衛星データとの間には線形関係が認められるため,

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_{ij}, \quad i = 1, \ldots, k, \ j = 1, \ldots, n_i \tag{6.1}$$

なるモデルが想定される．i は郡に対する指数, j は各郡の中の segment に対する指数, y_{ij} は i 番目の郡における j 番目の segment のとうもろこし (または大豆) の作付面積についての調査データ, x_{1ij}, x_{2ij} はそれぞれ, 各 segment において衛星データにより, とうもろこし及び大豆と識別された pixel の個数を表している．誤差項 u_{ij} は, 郡効果 (地域別差異) を考慮して

$$u_{ij} = v_i + e_{ij} \tag{6.2}$$

なる構造をもつと考える．v_i, e_{ij} はすべて互いに独立で

$$v_i \sim \mathcal{N}(0, \sigma_v^2), \quad e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$$

を仮定する．このとき, i 番目の群に対する各 segment のとうもろこし作付面積の群平均 μ_i を推定したいとする．有限母集団の枠組みで考えると, 推定したい母数は$\mu_i = N_i^{-1} \sum_{j=1}^{N_i} y_{ij} = \beta_0 + \beta_1 \overline{x}_{1i(p)} + \beta_2 \overline{x}_{2i(p)} + v_i + \overline{e}_{i(p)}$ となるが, ここでは簡単のために

$$\mu_i = \beta_0 + \beta_1 \overline{x}_{1i(p)} + \beta_2 \overline{x}_{2i(p)} + v_i \tag{6.3}$$

を推定する問題を考えよう．但し, $\overline{x}_{\ell i(p)} = N_i^{-1} \sum_{j=1}^{N_i} x_{\ell ij}$, $\ell = 1, 2$, $\overline{e}_{i(p)} = N_i^{-1} \sum_{j=1}^{N_i} e_{ij}$ とする．

ここで記述を簡便にするために行列表現を与える．$\mathbf{y}_i = (y_{i1}, \ldots, y_{in_i})'$, $\mathbf{y} = (\mathbf{y}_1', \ldots, \mathbf{y}_k')'$, $\mathbf{x}_{ij} = (1, x_{1ij}, x_{2ij})$, $\mathbf{x}_i = (\mathbf{x}_{i1}', \ldots, \mathbf{x}_{in_i}')'$, $\mathbf{X} = (\mathbf{x}_1', \ldots, \mathbf{x}_k')'$, $\beta = (\beta_0, \beta_1, \beta_2)'$ とし, $\mathbf{u}$ も $\mathbf{y}$ と同様に定義すると, モデル (6.1), (6.2) は

$$\begin{aligned}
\mathbf{y} &= \mathbf{X}\beta + \mathbf{u}, \\
\mathbf{V}(\omega) &= E[\mathbf{u}\mathbf{u}'] = \text{block diag}(\mathbf{V}_1, \ldots, \mathbf{V}_k), \ \omega = (\sigma_v^2, \sigma_e^2) \\
\mathbf{V}_i &= \mathbf{J}_i \sigma_v^2 + \mathbf{I}_i \sigma_e^2
\end{aligned}$$

と表現される．但し, $\mathbf{J}_i$ はすべての要素が 1 の $n_i \times n_i$ 行列, $\mathbf{I}_i$ は単位行列である．また (6.3) で与えられた, 推定したい μ_i は

$$\mu_i = \overline{\mathbf{x}}_{i(p)} \beta + v_i$$

と表わされる．但し $\overline{\mathbf{x}}_{i(p)} = N_i^{-1} \sum_{j=1}^{N_i} \mathbf{x}_{ij}$ である．

$\omega = (\sigma_v^2, \sigma_e^2)$ が既知のときには, β の一般化最小 2 乗推定量 (GLSE) は

$$\tilde{\beta}(\omega) = (\mathbf{X}'\mathbf{V}(\omega)^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}(\omega)^{-1}\mathbf{y}$$

で与えられる. また, v_i と $\overline{u}_{i\cdot} = n_i^{-1}\sum_{j=1}^{n_i} u_{ij}$ の共分散行列は

$$\begin{pmatrix} \sigma_v^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 + n_i^{-1}\sigma_e^2 \end{pmatrix}$$

で与えられるから, $\overline{u}_{i\cdot}$ を与えたときの v_i の条件付期待値は

$$E[v_i|\overline{u}_{i\cdot}] = \overline{u}_{i\cdot}\frac{\sigma_v^2}{\sigma_v^2 + n_i^{-1}\sigma_e^2}$$

となる. ここで, $\overline{\mathbf{x}}_{i\cdot} = n_i^{-1}\sum_{j=1}^{n_i}\mathbf{x}_{ij}$ とおくとき, $\overline{u}_{i\cdot}$ は $\overline{y}_{i\cdot} - \overline{\mathbf{x}}_{i\cdot}\tilde{\beta}(\omega)$ で推定されるので, 結局, μ_i の最良線形不偏推定量 (BLUE) は

$$\tilde{\mu}_i(\omega) = \overline{\mathbf{x}}_{i(p)}\tilde{\beta}(\omega) + (\overline{y}_{i\cdot} - \overline{\mathbf{x}}_{i\cdot}\tilde{\beta}(\omega))\frac{\sigma_v^2}{\sigma_v^2 + n_i^{-1}\sigma_e^2} \tag{6.4}$$

となる.

$\overline{\mathbf{x}}_{i(p)}\tilde{\beta}(\omega)$ は回帰統合推定量 (Regression Synthetic Estimator), $\overline{\mathbf{x}}_{i(p)}\tilde{\beta}(\omega) + (\overline{y}_{i\cdot} - \overline{\mathbf{x}}_{i\cdot}\tilde{\beta}(\omega))$ は $\overline{y}_{i\cdot}$に依存するので調査回帰推定量 (Survey Regression Estimator) と呼ばれ, $\tilde{\mu}_i(\omega)$ はこれら 2 つの推定量をσ_v^2 と σ_e^2/n_i の分散比で内分した形をしている. 即ち, v_i の e_{ij} に対する相対的バラツキが小さければ, $\overline{y}_{i\cdot}$ を $\overline{\mathbf{x}}_{i(p)}\tilde{\beta}(\omega)$ の方向へ縮小することによって点推定の安定化を図っているのである.

分散成分 σ_v^2, σ_e^2 は未知だからデータから推定する必要があり, そのために普通 Henderson(1953) の第 3 の方法が使われる. $\overline{x}_{\ell i\cdot} = \sum_j x_{\ell ij}/n_i$ とし, $y_{ij} - \overline{y}_{i\cdot}$ を $\{x_{1ij} - \overline{x}_{1i\cdot}, x_{2ij} - \overline{x}_{2i\cdot}\}$ へ OLS 回帰したときの残差を $\hat{e}_{ij}$ とする. また y_{ij} を $\{1, x_{1ij}, x_{2ij}\}$ へ OLS 回帰したときの残差を r_{ij} とする. $S_1 = \sum\sum\hat{e}_{ij}^2$, $S_2 = \sum\sum(r_{ij}^2 + \hat{e}_{ij}^2)$, $\nu = \sum n_i - k - 2$, $M = \{\sum n_i - \mathrm{tr}[(\mathbf{X}'\mathbf{X})^{-1}\sum_i n_i^2\overline{\mathbf{x}}_{i\cdot}\overline{\mathbf{x}}_{i\cdot}']\}/(k-1)$ とおくとき, σ_v^2, σ_e^2 の不偏推定量は,

$$\hat{\sigma}_v^{2U} = \frac{1}{\nu}S_1, \quad \hat{\sigma}_e^{2U} = \frac{1}{M}\left\{\frac{S_2}{k-1} - \frac{S_1}{\nu}\right\}$$

で与えられ, 5.1 節で議論された様々な推定量が考えられる. 例えば,

$$\hat{\sigma}_v^{2PT} = \frac{1}{M}\left[\max\left\{\frac{S_2}{k-1}, \frac{S_1+S_2}{\nu+k-3}\right\} - \min\left\{\frac{S_1}{\nu}, \frac{S_1+S_2}{\nu+k-1}\right\}\right],$$
$$\hat{\sigma}_e^{2EB} = \min\left\{\frac{S_1}{\nu}, \frac{S_1+S_2}{\nu+k-1}\right\}$$

は, 1 つの望ましい推定量であり, $\hat{\omega} = (\hat{\sigma}_v^{2PT}, \hat{\sigma}_e^{2EB})$ を (6.4) に代入したもの$\hat{\mu}_i = \tilde{\mu}_i(\hat{\omega})$ が small-area の平均値 μ_i の 2 段階推定量となる.

計量経済学の分野においては, 各 small-area i でのデータが時系列的に得られている場合が多い. Rao-Yu(1994) は, こうした状況を,

$$\begin{gathered} y_{it} = \mathbf{x}_{it}\beta + v_i + u_{it} + e_{it}, \\ u_{it} = \rho u_{i,t-1} + \varepsilon_{it}, \quad |\rho| < 1, \\ v_i \sim \mathcal{N}(0, \sigma_v^2),\ e_{it} \sim \mathcal{N}(0, \sigma_e^2),\ \varepsilon_{it} \sim \mathcal{N}(0, \sigma^2) \\ \{v_i\}, \{e_{it}\}, \{\varepsilon_{it}\} \text{ は独立}, \end{gathered}$$

なるモデルとして表現し, small-area の平均値の 2 段階推定量とその MSE の推定量の導出等を行った.

6.4. データ平滑化法としての縮小推定. 点 $t_1, \ldots, t_n$ $(0 \le t_1 < \cdots < t_n \le 1)$ において, 観測値 $y_1, \ldots, y_n$ が得られているとき, これらのデータに対して回帰モデル

$$\begin{aligned} y_i &= f(t_i) + e_i, \quad i = 1, \ldots, n, \\ e &= (e_1, \ldots, e_n)' \sim \mathcal{N}_n(0, \sigma^2 I_n) \end{aligned} \tag{6.5}$$

を当てはめることを考える. 但し, f は 1 次導関数 $f^{(1)}$ が絶対連続で, 2 次導関数 $f^{(2)} \in L^2[0,1]$ なる関数とする. 通常は, 平滑化パラメータ $\lambda > 0$ に対して,

$$\sum_{i=1}^{n} (y_i - f(t_i))^2 + \lambda \int_0^1 \{f^{(2)}(t)\}^2 dt \tag{6.6}$$

を最小にする関数 f を求める. この解は 3 次の自然スプライン関数であり, それは, 3 次 B-スプライン関数 $B_1(t), \ldots, B_n(t)$ の一次結合

$$f(t) = \sum_{i=1}^{n} \gamma_j B_j(t) \tag{6.7}$$

で一意的に表わすことができる. $\mu_i = f(t_i)$, $B_{ij} = B_j(t_i)$ に対して, $\mu = (\mu_1, \ldots, \mu_n)'$, $B = (B_{ij})$ とおき, また$\gamma = (\gamma_1, \ldots, \gamma_n)'$ とおくと,

$$\mu = B\gamma$$

と表わされるので, μ を推定すれば, γ が求められ (6.7) で与えられるスプライン関数が得られることになる. $\Omega_{ij} = \int_0^1 B_i^{(2)}(t) B_j^{(2)}(t) dt$, $\Omega = (\Omega_{ij})$ に対して $D = B'^{-1} \Omega B^{-1}$ とおき, $y = (y_1, \ldots, y_n)'$ とすると, (6.6) は

$$||y - \mu||^2 + \lambda \mu' D \mu$$

と書き直すことができる. $\lambda > 0$ を与えたときの解は

$$\mu(\lambda) = M(\lambda) y, \quad M(\lambda) = (I_n + \lambda D)^{-1}$$

で与えられることが知られており, 従って, 一つの平滑化スプライン関数 $f_\lambda(t)$ が得られる.

パラメータ λ は平滑化の程度を表わしており, これをどのように決定するかが重要な問題である. これには, 2 つの方法が知られてきた. 一つは, Wahba(1983) らの提唱する一般化交差検証法 (Generalized Cross Validation, GCV) で,

$$\frac{||(I - M(\lambda))y||^2}{\{\mathrm{tr}(I - M(\lambda))\}^2}$$

を最小にする $\lambda = \hat{\lambda}^{GCV}$ を選ぶというものである. もう一つは, Wahba(1985) らによって提案された経験 Bayes の方法である. Yanagimoto-Yanagimoto(1987) は, データが等間隔に得られているとき $(t_i = i)$, $\mu \sim \mathcal{N}_{n-2}(0, (\sigma^2/\lambda)D^-)$ なる事前分布を想定した. 但し, D は $\mu' D\mu = \Sigma(\mu_i - 2\mu_{i+1} + \mu_{i+2})^2$ なる $n \times n$ 対称行列である. このとき, Bayes 推定量は $\hat{\mu}^B(\lambda) = M(\lambda)y$ で与えられ, λ の推定量 $\hat{\lambda}^{EB}$ は y の周辺尤度に σ^2 の推定量を代入したもの

$$U(\lambda) = (n-2)\log\{y'(I - M(\lambda))y\} - (n-2)\log\lambda + \log|I + \lambda D|$$

を最小にする λ を求めることで与えられる.

3.1 節で述べたように, 経験 Bayes 推定量と James-Stein 推定量とは強い関係があると思われる. そこで上で導かれた経験 Bayes 推定量が Stein 型縮小推定量に対応するものなのかを調べてみよう. 実は $D = I_n$を $U(\lambda)$ に代入すると

$$U(\lambda) = n\log||y||^2 + 2\log(1 + \lambda)$$

となり $\hat{\lambda} = 0$ となってしまうため, 上の経験 Bayes 推定量は Stein 型縮小推定量になっていないことがわかる. これは, 周辺尤度に σ^2 の推定量を代入したところに起因していると思われる. そこでモデル (6.5) の代わりに, $\eta = 1/\sigma^2$ に適当な事前分布 $G(\eta)$ を仮定した次のモデルを考えて, Stein 型縮小推定量に対応する経験 Bayes 推定量を導出することを考えよう.

$$y_i = f(i) + e_i, \quad i = 1, \ldots, n \tag{6.8}$$

として $e = (e_1, \ldots, e_n)'$ が

$$\int (\sqrt{2\pi})^{-n} \eta^{n/2} e^{-(\eta/2)||e||^2} G(d\eta)$$

なる球面対称分布に従うものとする. これは

$$y|\eta \sim \mathcal{N}_n(\mu, \eta I)$$
$$\eta \sim G(\eta)$$

なる正規混合モデルとしてみることができる. ここで, η を与えたときの μ の条件付き事前分布として

$$\mu|\eta \sim \mathcal{N}_{n-2}(0, \frac{1}{\eta\lambda} D^-)$$

を想定しよう．このとき Bayes 推定量は $\hat{\mu}^B(\lambda) = M(\lambda)y$ となり，λ の推定量 $\hat{\lambda}^{SEB}$ は周辺尤度

$$\int \sqrt{|I - M(\lambda)|^*}\eta^{(n-2)/2}e^{-(\eta/2)y'(I-M(\lambda))y}G(d\eta) \tag{6.9}$$

を最大にするものとして与えられる．但し，$|A|^*$ は A の 0 でない固有値の積である．

シミュレーション実験を行うと，$\hat{\mu}^B(\hat{\lambda}^{SEB})$ は，$\hat{\mu}^B(\hat{\lambda}^{GCV})$，$\hat{\mu}^B(\hat{\lambda}^{EB})$ とほぼ同等な性質を示していることがわかる．$\hat{\lambda}^{SEB}$ の導出は数値計算に頼らざるをえない．そのため (6.9) の近似解を明示的に求めることが試みられる．その一つの近似的な解は

$$\hat{\lambda} = \frac{a}{\int \eta G(d\eta)y'Dy}$$

で与えられ，従って経験 Bayes 推定量

$$\hat{\mu}^{EB} = \hat{\mu}^B(\hat{\lambda}) = \left(I + \frac{a}{\int \eta G(d\eta)y'Dy}D\right)^{-1} y$$

が得られる．ここで，a は適当な定数で $0 < a \leq 2(\mathrm{tr}D - 2ch_{max}(D))/ch_{max}(D)$ をみたすならば，$E[||\hat{\mu}^{EB} - \mu||^2] \leq E[||y - \mu||^2]$ が一様に成り立つことが示される (Berger(1975))．$D = I$ のときには，$\hat{\mu}^{EB} = [1 - a/\{\int \eta G(d\eta)||y||^2 + a\}]y$ となり Stein 型の縮小推定量になっていることがわかる．しかし，真の関数 $f(t)$ が直線からずれしかも分散が大きくなると，$\hat{\mu}^B(\hat{\lambda}^{SEB})$ に比べ，大きな改善が得られないことが数値的にわかり，別の有効な近似解の導出が望まれている．

参考文献

Akai, T. (1986). Simultaneous estimation of location parameters of the distribution with finite support. *Ann. Inst. Statist. Math.*, **38**, 85-99.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, **52**, 317-332.

Alam, K. (1973). A family of admissible minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.* , **1**, 517-525.

Alam, K. (1975). Minimax and admissible minimax estimators of the mean of a multivariate normal distribution for unknown covariance matrix. *Ann. Statist.* , **5**, 83-95.

Albert, J.H. (1987). Empirical Bayes estimation in contingency tables. *Comm. Statist.-Theory Methods*, **16**, 2459-2485.

Amari, S. (1982). Differential geometry of curved exponential families - Curvature and information loss. *Ann. Statist.* , **10**, 357-385.

Ansley, C.F., Kohn, R. and Wong, C.(1993). Nonparametric spline regression with prior information. *Biometrika*, **80**, 75-88.

Arnold, B.C. (1970). Inadmissibility of the usual scale estimate for a shifted exponential distribution. *J. Amer. Statist. Assoc.*, **65**, 1260-1264.

Baranchik, A.J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution. *Ann. Math. Statist.* , **41**, 642-645.

Baranchick, A. (1973). Inadmissibility of maximum likelihood estimators in some multiple regression problems with three or more independent variables. *Ann. Statist.*, **1**, 312-321.

Barlow, R.F., Bartholomew, D.D., Bremner, J.M. and Brunk, H.D. (1972). *Statistical Inference under Order Restrictions.* Wiley, New York.

Battese, G.E., Harter, R.M. and Fuller, W.A. (1988). An error-components model for prediction of county crop areas using survey and satellite data. *J. Amer. Statist. Assoc.*, **83**, 28-36.

Berger, J. (1975). Minimax estimation of location vectors for a wide class of densities. *Ann. Statist.*, **3**, 1318-1328.

Berger, J. (1976). Admissible minimax estimation of a multivariate normal mean with arbitrary quadratic loss. *Ann. Statist.*, **4**, 223-226.

Berger, J.O. (1980a). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of gamma parameters. *Ann. Statist.*, **8**, 545-571.

Berger, J.O. (1980b). A robust generalized Bayes estimator and confidence region for a multivariate normal mean. *Ann. Statist.*, **8**, 716-761.

Berger, J.O. (1982). Selecting a minimax estimator of a multivariate normal mean. *Ann. Statist.*, **10**, 81-92.

Berger, J.O.(1985). *Statistical Decision Theory and Bayesian Analysis.* 2nd Ed., Springer-Verlag, New York.

Berger, J.O. and Berliner, L.M. (1984). Bayesian input in Stein estimation and a new minimax empirical Bayes estimator. *J. Econometrics*, **25**, 87-108.

Berger, J.O. and Berliner, L.M. (1986). Robust Bayes and empirical Bayes analysis with ε-contaminated priors. *Ann. Statist.*, **14**, 461-486.

Berger, J.O., Berliner, L.M. and Zaman, A. (1982). General admissibility results for estimation in a control problem. *Ann. Statist.*, **10**, 838-856.

Berger, J.O. and Dey, D.K. (1985). Truncation of shrinkage estimators of normal means in the nonsymmetric case. In *Multivariate Analysis* VI (P.R. Krishnaiah, ed.), 43-56. North-Holland, Amsterdam.

Berger, J.O. and Robert, C. (1990). Subjective hierarchical Bayes estimation of a multivariate normal mean: On the frequentist interface. *Ann. Statist.*, **18**, 617-651.

Berger, J.O. and Srinivasan, C. (1978). Generalized Bayes estimators in multivariate problems. *Ann. Statist.*, **6**, 783-801.

Berger, J. and Wolpert, R. (1983). Estimating the mean function of a Gaussian process and the Stein effect. *J. Multivariate Anal.*, **13**, 401-424.

Berliner, L.M. (1983). Improving on inadmissible estimators in control problems. *Ann. Statist.*, **11**, 814-826.

Berliner, L.M. (1985). A decision theoretic structure for robust Bayesian analysis with applications to the estimation of a multivariate normal mean. In *Bayesian Statistics* 2 (J.M. Bernardo, et al., eds.), 619-628. North-Holland, Amsterdam.

Bhattacharya, C.G. (1984). Two inequalities with an application. *Ann. Inst. Statist. Math.*, **36**, 129-134.

Bhattacharya, P.K. (1966). Estimating the mean of a multivariate normal population with general quadratic loss function. *Ann. Math. Statist.*, **37**, 1819-1824.

Bickel, P.J. (1981). Minimax estimation of the mean of a normal distribution when the parameter space is restricted. *Ann. Statist.*, **9**, 1301-1309.

Bilodeau, M. (1988). On the simultaneous estimation of scale parameters. *Canad. J. Statist.*, **16**, 169-174.

Bilodeau, M. and Kariya, T. (1989). Minimax estimators in the normal MANOVA model. *J. Multivariate Anal.*, **28**, 260-270.

Bilodeau, M. and Srivastava, M.S. (1992). Estimation of the eigenvalues of $\Sigma_1 \Sigma_2^{-1}$. *J. Multivariate Anal.*, **41**, 1-13.

Blumenthal, S. and Cohen, A. (1968a). Estimation of the larger translation parameter. *Ann. Math. Statist.*, **39**, 502-516.

Blumenthal, S. and Cohen, A. (1968b). Estimation of two ordered translation parameters. *Ann. Math. Statist.*, **39**, 517-530.

Bock, M.E. (1975). Minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.*, **3**, 209-218.

Bock, M.E. (1985). Minimax estimators that shift towards a hypersphere for location vectors of spherically symmetric distributions. *J. Multivariate Anal.*, **17**, 127-147.

Brandwein, A.C. and Strawderman, W.E. (1978). Minimax estimation of location parameters for spherical symmetric unimodal distributions. *Ann. Statist.*, **6**, 377-416.

Brandwein, A.C. and Strawderman, W.E. (1990). Stein estimation: The spherically symmetric case. *Statist. Science*, **5**, 356-369.

Brandwein, A.C. and Strawderman, W.E. (1991a). Generalizations of James-Stein estimators under spherical symmetry. *Ann. Statist.*, **19**, 1639-1650.

Brandwein, A.C. and Strawderman, W.E. (1991b). Improved estimates of location in the presense of an unknown scale. *J. Multivariate Anal.*, **39**, 305-314.

Bravo, G. and MacGibbon, B. (1988a). Improved shrinkage estimators for the mean vector of a scale mixture of normals with unknown variance. *Canad. J. Statist.*, **16**, 237-245.

Bravo, G. and MacGibbon, B. (1988b). Improved estimation for the parameters of an inverse Gaussian distribution. *Comm. Statist.-Theory Methods*, **17**, 4285-4299.

Brewster, J.F. (1974). Alternative estimators for the scale parameter of the exponential distribution with unknown location. *Ann. Statist.*, **2**, 553-557.

Brewster, J.F. and Zidek, J.V. (1974). Improving on equivariant estimators. *Ann. Statist.*, **2**, 21-38.

Brown, G.G. and Rutemiller, H.C. (1973). The efficiencies of maximum likelihood and minimum variance unbiased estimators of fraction defective in the normal case. *Technometrics*, **15**, 849-855.

Brown, L.D. (1966). On the admissibility of invariant estimators of one or more location parameters. *Ann. Math. Statist.*, **37**, 1087-1136.

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.

Brown, L.D. (1971). Admissible estimators, recurrent diffusions, and insolvable boundary value problems. *Ann. Math. Statist.*, **42**, 855-904.

Brown, L.D. (1973). Estimation with incompletely specified loss functions. *J. Amer. Statist. Assoc.*, **70**, 471-427.

Brown, L.D. (1981). A complete class theorem for statistical problems with finite sample spaces. *Ann. Statist.*, **9**, 1289-1300.

Brown, L.D. (1990). An ancillarity paradox which appears in multiple linear regression. *Ann. Statist.*, **18**, 471-538.

Brown, L.D. and Cohen, A. (1974). Point and confidence estimation of a common mean and recovery of interblock information. *Ann. Statist.*, **2**, 963-976.

Brown, L.D. and Farrell, R.H. (1985). Complete class theorems for estimation of multivariate Poisson means and related problems. *Ann. Statist.*, **13**, 706-726.

Brown, L.D. and Hwang, J.T. (1982). A unified admissibility proof. In *Statistical Decision Theory and Related Topics* III (S.S. Gupta, J. Berger, eds.), 205-230. Academic Press, New York.

Brown, P.J. (1982). Multivariate calibration. *J. Roy. Statist. Soc.* (Ser. B), **44**, 287-321.

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 multivaraite variance components models. *Ann. Statist.*, **19**, 850-869.

Casella, G. (1981). Minimax ridge regression estimation. *Ann. Statist.*, **8**, 1036-1056.

Casella, G. (1985a). Condition numbers and minimax ridge regression estimators. *J. Amer. Statist. Assoc.*, **80**, 753-758.

Casella, G. (1985b). An introduction to empirical Bayes data analysis. *Amer. Statist.*, **39**, 83-87.

Casella, G. and Hwang, J.T. (1983). Empirical Bayes confidence sets for the mean of a multivariate normal distribution. *J. Amer. Statist. Assoc.*, **78**, 688-698.

Casella, G. and Hwang, J.T. (1987). Employing vague prior information in the construction of confidence sets. *J. Multivariate Anal.*, **21**, 79-104.

Casella, G., Hwang, J.T.G. and Robert, C.P. (1994). Loss functions for set estimation. In *Statistical Decision Theory and Related Topics* V (S.S. Gupta, J.O. Berger, eds.), 237-251. Springer-Verlag, New York.

Casella, G. and Strawderman, W.E. (1981). Estimating a bounded mean. *Ann. Statist.*, **9**, 870-879.

Cellier, D., Fourdrinier, D. and Robert, C. (1989). Robust shrinkage estimators of the location parameter for elliptically symmetric distributions. *J. Multivariate Anal.*, **29**, 39-52.

Chang, Y.-T. (1982). Stein-type estimators for parameters in truncated spaces. *Keio Sci. Tech. Rep.*, **35**, 185-193.

Chaturvedi, A., Hoa, T.V. and Shukla, G. (1993). Performance of the Stein-rule estimators when the disturbances are misspecified as spherical. *Economic Studies Quarterly*, **44**, 97-107.

Chen, J. and Hwang, J.T. (1988). Improved set estimators for the coefficients of a linear model when the error distribution is spherically symmetric with unknown variances. *Canad. J. Statist.*, **16**, 293-299.

Chiou, P. (1993). Empirical Bayes shrinkage estimation of reliability in the exponential distribution. *Comm. Statist.-Theory Methods*, **22**, 1483-1494.

Chou, J.-P. (1988). An identity for multidimensional continuous exponential families and its applications. *J. Multivariate Anal.*, **24**, 129-142.

Chou, J.-P. (1991). Simultaneous estimation in discrete multivariate exponential families. *Ann. Statist.*, **19**, 314-328.

Chou, J.-P. and Strawderman, W.E.(1990). Minimax estimation of means of multivariate normal mixtures. *J. Multivariate Anal.*, **35**, 141-150.

Chow, M.S. (1987). A complete class theorem for estimating a noncentrality parameter. *Ann. Statist.*, **15**, 800-804.

Clayton, D. and Kaldor, J.(1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, **43**, 671-681.

Clevenson, M.L. and Zidek, J.V. (1975). Simultaneous estimation of the mean of independent Poisson laws. *J. Amer. Statist. Assoc.*, **70**, 698-705.

Cohen, A. (1972). Improved confidence intervals for the variance of a normal distribution. *J. Amer. Statist. Assoc.*, **67**, 382-387.

Cohen, A. and Sackrowitz, H.B. (1970). Estimation of the last mean of a monotone sequence. *Ann. Math. Statist.*, **41**, 2021-2034.

Cohen, A. and Sackrowitz, H.B. (1974). On estimating the common mean of two normal distributions. *Ann. Statist.*, **2**, 1274-1282.

Cohen, A. and Sackrowitz, H.B. (1982). Estimating the mean of the selected population. In *Statistical Decision Theory and Related Topics* III (S.S. Gupta, J.O. Berger, eds.), **1**, 243-270. Academic Press, New York.

Copas, J.B. (1983). Regression, prediction, shrinkage. *J. Roy. Statist. Soc.* (Ser. B), **45**, 311-354.

Dahiya, R.C. (1974). Estimation of the mean of the selected population. *J. Amer. Statist. Assoc.*, **69**, 226-230.

DasGupta, A. (1986). Simultaneous estimation in the multiparameter gamma distribution under weighted quadratic losses. *Ann. Statist.*, **14**, 206-219.

DasGupta, A. (1989). A general theorem on decision theory for nonnegative functionals: with applications. *Ann. Statist.*, **17**, 1360-1374.

Dey, D. (1988). Simultaneous estimation of eigenvalues. *Ann. Inst. Statist. Math.*, **40**, 137-147.

Dey, D. (1990). Estimation of scale parameters in mixture distributions. *Canad. J. Statist.*, **18**, 171-178.

Dey, D. and Berger, J.O. (1983). On truncation of shrinkage estimators in simultaneous estimation of normal means. *J. Amer. Statist. Assoc.*, **78**, 865-869.

Dey, D. and Chung, Y. (1992). Compound Poisson distributions: Properties and estimation. *Comm. Statist.-Theory Methods*, **21**, 3097-3121.

Dey, D. and Gelfand, A.E. (1989). Improved estimation of a patterned covariance matrix. *J. Multivariate Anal.*, **31**, 107-116.

Dey, D., Ghosh, M. and Srinivasan, C. (1987). Simultaneous estimation of parameters under entropy loss. *J. Statist. Plann. Inference*, **15**, 347-363.

Dey, D., Ghosh, M. and Srinivasan, C. (1990). A new class of improved estimators of a multinormal precision matrix. *Statist. Decisions*, **8**, 141-151.

Dey, D.K. and Srinivasan, C. (1985). Estimation of covariance matrix under Stein's loss. *Ann. Statist.*, **13**, 1581-1591.

Dey, D.K. and Srinivasan, C. (1991). On estimation of discriminant coefficients. *Statist. Probab. Lett.*, **11**, 189-193.

Eaton, M.L. (1992). A statistical diptych: admissible inferences - recurrence of symmetric Markov chains. *Ann. Statist.*, **20**, 1147-1179.

Eaton, M.L. and Olkin, I. (1987). Best equivariant estimators of a Cholesky decomposition. *Ann. Statist.*, **15**, 1639-1650.

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

Efron, B. and Morris, C. (1972a). Empirical Bayes on vector observations: An extension of Stein's method. *Biometrika*, **59**, 335-347.

Efron, B. and Morris, C. (1972b). Limiting the risk of Bayes and empirical Bayes estimators - Part II: The empirical Bayes case. *J. Amer. Statist. Assoc.*, **67**, 130-139.

Efron, B. and Morris, C. (1975). Data analysis using Stein's estimator and its generalizations. *J. Amer. Statist. Assoc.*, **70**, 311-319.

Efron, B. and Morris, C. (1976a). Families of minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.*, **4**, 11-21.

Efron, B. and Morris, C. (1976b). Multivariate empirical Bayes estimation of covariance matrices. *Ann. Statist.*, **4**, 22-32.

Eguchi, S. and Yanagimoto, T. (1984). Second-order structure of relative entropy risk by Laplace-Beltrami operator. Unpublished manuscript.

Enis, P. and Geisser, S. (1971). Estimation of the probability that $Y < X$. *J. Amer. Statist. Assoc.*, **66**, 162-168.

Fay, R.E. and Herriot, R. (1979). Estimates of income for small places: An application of James-Stein procedures to census data. *J. Amer. Statist. Assoc.*, **74**, 269-277.

Friedman, J.H. (1989). Regularized discriminant analysis. *J. Amer. Statist. Assoc.*, **84**, 165-175.

Fujikoshi, Y. and Nishii, R. (1986). Selection of variables in a multivariate inverse regression problem. *Hiroshima Math. J.*, **16**, 267-277.

藤野和建 (1987). 正規分布の片側確率の推定. 応用統計学, **16**, 119-130.

Gelfand, A.E. and Dey, D.K. (1988). On the estimation of a variance ratio. *J. Statist. Plann. Inference*, **19**, 121-131.

George, E.I. (1986a). Minimax multiple shrinkage estimation. *Ann. Statist.*, **14**, 188-205.

George, E.I. (1986b). Combining minimax shrinkage estimators. *J. Amer. Statist. Assoc.*, **81**, 437-445.

George, E.I. (1991). Shrinkage domination in a multivariate common mean problem. *Ann. Statist.*, **19**, 952-960.

George, E.I. and Casella, G. (1994). An empirical Bayes confidence report. *Statistica Sinica*, **4**, 617-638.

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.

Ghosh, M., Nickerson, D.M. and Sen, P.K. (1987). Sequential shrinkage estimation. *Ann. Statist.*, **15**, 817-829.

Ghosh, M. and Parsian, A. (1980). Admissible and minimax multiparameter estimation in exponential families. *J. Multivariate Anal.*, **10**, 551-564.

Ghosh, M. and Parsian, A. (1981). Bayes minimax estimation of multiple Poisson parameters. *J. Multivariate Anal.*, **11**, 280-288.

Ghosh, M. and Rao, J.N.K. (1994). Small area estimation: An appraisal. *Statist. Science*, **9**, 55-93.

Ghosh, M., Saleh, A.K.Md.E. and Sen, P.K. (1989). Empirical Bayes subset estimation in regression models. *Statist. Decisions*, **7**, 15-35.

Ghosh, M. and Sinha, B.K. (1988). Empirical and hierarchical Bayes competitors of preliminary test estimators in two sample problems. *J. Multivariate Anal.*, **27**, 206-227.

Ghosh, M. and Yang, M.-C. (1988) Simultaneous estimation of Poisson means under entropy loss. *Ann. Statist.*, **16**, 278-291.

Gleser, L.J. (1986). Minimax estimators of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix. *Ann. Statist.*, **14**, 1625-1633.

Goutis, C. and Casella, G. (1991). Improved invariant confidence intervals for a normal variance. *Ann. Statist.*, **19**, 2015-2031.

Graybill, F.A. and Deal, R.B. (1959). Combining unbiased estimators. *Biometrics*, **15**, 543-550.

Gupta, A.K., Saleh, A.K.Md.E. and Sen, P.K. (1989). Improved estimation in a contingency tables: Independence structure. *J. Amer. Statist. Assoc.*, **84**, 525-532.

Haff, L.R. (1979). An identity for the Wishart distribution with applications. *J. Multivariate Anal.*, **9**, 531-542.

Haff, L.R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix. *Ann. Statist.*, **8**, 586-597.

Haff, L.R. (1986). On linear log-odds and estimation of discriminant coefficients. *Comm. Statist.-Theory Methods*, **15**, 2131-2144.

Haff, L.R. (1991). The variational form of certain Bayes estimators. *Ann. Statist.*, **19**, 1163-1190.

Haff, L.R. and Johnson, R.W. (1986). The superharmonic condition for simultameous estimation of means in exponential families. *Canad. J. Statist.*, **14**, 43-54.

Hawkins, D.L. and Han, C.-P. (1989). A minimum average risk approach to shrinkage estimators of the normal mean. *Ann. Inst. Statist. Math.*, **41**, 347-363.

He, K. (1992). Parametric empirical Bayes confidence intervals based on James-Stein estimator. *Statist. Decisions*, **10**, 121-132.

Henderson, C.R. (1953). Estimation of variance and covariance components. *Biometrics*, **9**, 226-252.

Hill, R.C. and Judge, G.G. (1990). Improved estimation under collinearity and squared error loss. *J. Multivariate Anal.*, **32**, 296-312.

Hirano, K. (1977). Estimation procedures based on preliminary test, shrinkage technique and information criterion. *Ann. Inst. Statist. Math.*, **29**, 21-34.

広津千尋 (1992). QC テクノロジー. 品質, **22**, 238-258.

Hoadley, B. (1970). A Bayesian look at inverse linear regression. *J. Amer. Statist. Assoc.*, **65**, 356-369.

Hoerl, A.E. and Kennard, R.W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, **12**, 55-68.

Hoffmann, K. (1992). *Improved Estimation of Distribution Parameters: Stein-Type Estimators.* B.G. Teubner Verlagsgesellschaft, Stuttgart.

Honda, T. (1991). Minimax estimators in the MANOVA model for arbitrary quadratic loss and unknown covariance matrix. *J. Multivariate Anal.*, **36**, 113-120.

Hudson, H.M. (1978). A natural identity for exponential families with applications in multiparameter estimation. *Ann. Statist.*, **6**, 478-484.

Hwang, J.T. (1982). Improving upon standard estimators in discrete exponential families with applications to Poisson and negative binomial cases. *Ann. Statist.*, **10**, 857-867.

Hwang, J.T. (1985). Universal domination and stochastic domination: Estimation simultaneously under a broad class of loss functions. *Ann. Statist.*, **13**, 295-314.

Hwang, J.T. (1993). Empirical Bayes estimation for the means of the selected populations. *Sankhya* (Ser. A), **55**, 285-311.

Hwang, J.T. and Brown, L.D. (1991). Estimated confidence under the validity constraint. *Ann. Statist.*, **19**, 1964-1977.

Hwang, J.T. and Casella, G. (1982). Minimax confidence sets for the mean of a multivariate normal distribution. *Ann. Statist.*, **10**, 868-881.

Hwang, J.T. and Casella, G. (1984). Improved set estimators for a multivariate normal mean. *Statist. Decisions*, Supplement Issue **1**, 3-16.

Hwang, J.T., Casella, G., Robert, C., Wells, M.T. and Farrell, R.H. (1992). Estimation of accuracy in testing. *Ann. Statist.*, **20**, 490-509.

Hwang, J.T. and Chen, J. (1986). Improved confidence sets for the coefficients of a linear model with spherically symmetric errors. *Ann. Statist.*, **14**, 444-460.

Hwang, J.T. and Peddada, S.D. (1994). Confidence interval estimation subject to order restrictions. *Ann. Statist.*, **22**, 67-93.

Hwang, J.T. and Ullah, A. (1994). Confidence sets centered at James-Stein estimators. *J. Econometrics*, **60**, 145-156.

Inada, K. (1984). A minimax regret estimator of a normal mean after preliminary test. *Ann. Inst. Statist. Math.*, **36**, 207-215.

James, W. and Stein, C. (1961). Estimation with quadratic loss. In *Proc. Fourth Berkeley Symp. Math. Statist. Probab.*, **1**, 361-379. University of California Press, Berkeley.

Jin, C. (1993). A note on simultaneous estimation of eigenvalues of a multivariate normal covariance matrix. *Statist. Probab. Lett.*, **16**, 197-203.

Johnson, R.W. (1987). Simultaneous estimation of binomial N's. *Sankhya* (Ser. A), **49**, 264-267.

Johnstone, I. (1984). Admissibility, difference equations and recurrence in estimating a Poisson mean. *Ann. Statist.*, **12**, 1173-1198.

Johnstone, I. (1987). On inadmissibility of some unbiased estimates of loss. In *Statistical Decision Theory and Related Topics* IV (S.S. Gupta, J.O. Berger, eds.). Springer-Verlag, New York.

Johnstone, I.M. and MacGibbon, K.B. (1992). Minimax estimation of a constrained Poisson vector. *Ann. Statist.*, **20**, 807-831.

Joshi, V.M. (1967). Admissibility of the usual confidence sets for the mean of a univariate or bivariate normal population. *Ann. Math. Statist.*, **40**, 1042-1067.

Judge, G. and Bock, M.E. (1978). *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics.* North-Holland, Amsterdam.

Jun, C.-H. (1993). Heuristic shrinkage estimators of Poisson means and a Monte Carlo comparison. *J. Statist. Comput. Simul.*, **47**, 181-193.

Kambo, N.S., Handa, B.R. and Al-Hemyari, Z.A. (1990). On shrunken estimators for exponential scale parameter. *J. Statist. Plann. Inference*, **24**, 87-94.

Kariya, T. (1989). Equivariant estimation in a model with an ancillary statistic. *Ann. Statist.*, **17**, 920-928.

Kariya, T., Giri, N.C. and Perron, F. (1988). Equivariant estimation of a mean vector μ of $\mathcal{N}(\mu, \Sigma)$ with $\mu'\Sigma^{-1}\mu = 1$ or $\Sigma^{-1/2}\mu = c$ or $\Sigma = \sigma^2\mu'\mu I$. *J. Multivariate Anal.*, **27**, 270-283.

Kariya, T. and Konno, Y. (1994). Double shrinkage estimators in the GMANOVA model. Discussion Paper Series A, Inst. Econ. Res., Hitotsubashi University.

Kariya, T., Konno, Y. and Strawderman, W.E. (1994). Construction of improved estimators for the regression coefficient matrix in GMANOVA model. Discussion Paper Series A, Inst. Econ. Res., Hitotsubashi University.

Ki, F. and Tsui, K.-W. (1985). Improved confidence set estimators of a multivariate normal mean and generalizations. *Ann. Inst. Statist. Math.*, **37**, 487-498.

Ki, F. and Tsui, K.-W. (1990). Multiple-shrinkage estimators of means in exponential families. *Canad. J. Statist.*, **18**, 31-46.

Kleffe, J. and Rao, J.N.K. (1986). The existence of asymptotically unbiased non-negative quadratic estimates of variance components in ANOVA models. *J. Amer. Statist. Assoc.*, **81**, 692-698.

Klotz, J.H., Milton, R.C. and Zacks, S. (1969). Mean square efficiency of estimators of variance components. *J. Amer. Statist. Assoc.*, **64**, 1383-1402.

Komaki, F. (1994). On asymptotic properties of predictive distributions. Technical Report of Dept. Math. Eng. Inf. Phy. (METR 94-21), University of Tokyo.

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.

小西貞則, 本田正幸 (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.

Konno, Y. (1991). On estimation of a matrix of normal means with unknown covariance matrix. *J. Multivariate Anal.*, **36**, 44-55.

Konno, Y. (1992). On estimating eigenvalues of the scale matrix of the multivariate F distribution. *Sankhya* (Ser. A), **54**, 241-251.

Konno, Y. (1995). Estimation of a normal covariacne matrix with incomplete data under Stein's loss. *J. Multivariate Anal.*, **52**, 308-324.

Koul, H.L. and Saleh, A.K.Md.E. (1993). R-estimation of the parameters of autoregressive $[AR(p)]$ models. *Ann. Statist.*, **21**, 534-551.

Kourouklis, S. (1995). A new property of the inverse Gaussian distribution with applications. Unpublished manuscript, University of Patras, Greece.

Krishnamoorthy, K. (1992). On a shrinkage estimator of a normal common mean vector. *J. Multivariate Anal.*, **40**, 109-114.

Krishnamoorthy, K. and Gupta, A.K. (1989). Improved minimax estimation of a normal precision matrix *Canad. J. Statist.*, **17**, 91-102.

Krutchkoff, R.G. (1967). Classical and inverse regression methods of calibration. *Technometrics*, **9**, 429-439.

Kubokawa, T. (1987). Admissible minimax estimation of a common mean of two normal populations. *Ann. Statist.*, **15**, 1245-1256.

Kubokawa, T. (1991). An approach to improving the James-Stein estimator. *J. Multivariate Anal.*, **36**, 121-126.

Kubokawa, T. (1994a). A unified approach to improving equivariant estimators. *Ann. Statist.*, **22**, 290-299.

Kubokawa, T. (1994b). Double shrinkage estimation of ratio of scale parameters. *Ann. Inst. Statist. Math.*, **46**, 95-116.

Kubokawa, T. (1995). Estimation of variance components in mixed linear models. *J. Multivariate Anal.*, **53**, 210-236.

Kubokawa, T., Honda, T., Morita, K. and Saleh, A.K.Md.E. (1993). Estimating a covariance matrix of a normal distribution with unknown mean. *J. Japan Statist. Soc.*, **23**, 131-144.

Kubokawa, T., Morita, K., Makita, S. and Nagakura, K. (1993). Estimation of the variance and its applications. *J. Statist. Plann. Inference*, **35**, 319-333.

Kubokawa, T. and Robert, C.P. (1994). New perspectives on linear calibration. *J. Multivariate Anal.*, **51**, 178-200.

Kubokawa, T., Robert, C. and Saleh, A.K.Md.E. (1992). Empirical Bayes estimation of the variance parameter of a normal distribution with unknown mean under an entropy loss. *Sankhya* (Ser. A), **54**, 402-410.

Kubokawa, T., Robert, C.P. and Saleh, A.K.Md.E. (1993). Estimation of noncentrality parameters. *Canad. J. Statist.*, **21**, 45-57.

Kubokawa, T. and Saleh, A.K.Md.E. (1994). Estimation of location and scale parameters under order restrictions. *J. Statist. Research*, **28**, 41-51.

Kubokawa, T., Saleh, A.K.Md.E., Konno, Y. and Wagatsuma, K. (1995). Estimation of variance components under Kullback-Leibler loss with applications. Unpublished manuscript.

Kubokawa, T., Saleh, A.K.Md.E. and Makita, S. (1993). On improved positive estimators of variance components. *Statist. Decisions*, Supplement Issue **3**, 1-16.

Kubokawa, T., Saleh, A.K.Md.E. and Morita, K. (1992). Improving on MLE of coefficient matrix in a growth curve model. *J. Statist. Plann. Inference*, **31**, 169-177.

Kuo, L. (1986). A note on Bayes empirical Bayes estimation by means of Dirichlet processes. *Statist. Probab. Lett.*, **4**, 145-150.

LaMotte, L.R. (1973). On non-negative quadratic unbiased estimation of variance components. *J. Amer. Statist. Assoc.*, **68**, 728-730.

Lee, C.C. (1981). The quadratic loss of isotonic regression under normality. *Ann. Statist.*, **9**, 686-688.

Lee, S.-Y. (1980). Estimation of covariance structure models with parameters subject to functional restraints. *Psychometrika*, **45**, 309-324.

Lee, S.-Y. (1981). The multiplier method in constrained estimation of covariance structure models. *J. Statist. Comput. Simul.*, **12**, 247-257.

Lehmann, E.L. (1983). *Theory of Point Estimation.*, Wiley, New York.

Lele, C. (1992). Inadmissibility of loss estimators. *Statist. Decisions*, **10**, 309-322.

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.

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.

Loh, W.-L. (1991). Estimating covariance matrices. *Ann. Statist.*, **19**, 283-296.

Loh, W.-L. (1991). Estimating covariance matrices II. *J. Multivariate Anal.*, **36**, 163-174.

Louis, T.A. (1991). Using empirical Bayes methods in biopharmaceutical research. *J. Multivariate Anal.*, **36**, 163-174.

Lu, K.L. and Berger, J.O. (1989a). Estimation of normal means: Frequentist estimation of loss. *Ann. Statist.*, **17**, 890-906.

Lu, K.L. and Berger, J.O. (1989b). Estimated confidence procedures for multivariate normal means. *J. Statist. Plann. Inference*, **23**, 1-19.

Lwin, T. and Maritz, J.S. (1989). Empirical Bayes approach to multiparameter estimation: With special reference to multinomial distribution. *Ann. Inst. Statist. Math.*, **41**, 81-99.

Maatta, J.M. and Casella, G. (1990). Developments in decision-theoretic variance estimation. *Statist. Science*, **5**, 90-120.

MacGibbon, B. and Shorrock, G. (1994). Estimation of the lamda parameter of an inverse Gaussian distribution. *Statist. Probab. Lett.*, to appear.

Majumdar, S. (1994). Bayes compound and empirical Bayes estimation of the mean of a Gaussian distribution on a Hilbert space. *J. Multivariate Anal.*, **48**, 87-106.

Mandelbaum, A. and Shepp, L.A. (1987). Admissibility as a touchstone. *Ann. Statist.*, **15**, 252-268.

Manton, K.G., Woodbury, M.A., Stallard, E., Riggan, W.B., Creason, J.P. and Pellom, A.C. (1989). Empirical Bayes procedures for stabilizing maps of U.S. cancer mortality rates. *J. Amer. Statist. Assoc.*, **84**, 637-650.

Marchand, E. (1994). On the estimation of the mean of a $\mathcal{N}_p(\mu, \Sigma)$ population with $\mu'\Sigma^{-1}\mu$ known. *Statist. Probab. Lett.*, **21**, 69-75.

Martin, J.K. and McDonald, R.P. (1975). Bayesian estimation in unrestricted factor analysis: A treatment for Heywood cases. *Psychometrika*, **40**, 505-516.

Mathew, T., Sinha, B.K. and Sutradhar, B.C. (1992). Nonnegative estimation of variance components in unbalanced mixed models with two variance components. *J. Multivariate Anal.*, **42**, 77-101.

Mathew, T., Niyogi, A. and Sinha, B.K. (1994). Improved nonnegative estimation of variance components in balanced multivariate mixed models. *J. Multivariate Anal.*, **51**, 83-101.

松田眞一, 吉村　功, 藤本　隆 (1990). 分散行列の縮小推定量を用いたロバストな 2 次判別関数. 応用統計学, **19**, 33-51.

McLachlan, G.J. (1974). An asymptotic unbiased technique for estimating the error rates in discriminant analysis. *Biometrics*, **30**, 239-249.

Mikhail, N. and Vassily, V. (1994). A review of the results on the Stein approach for estimators improvement. Unpublished manuscript.

三輪哲久 (1979). SN 比に関する推測について. 日本品質管理学会第 9 回年次大会研究発表要旨集, 37-40.

三輪哲久 (1985). 線形校正における点推定量の平均二乗誤差にもとづく比較. 応用統計学, **14**, 83-93.

Moors, J.J.A. (1981). Inadmissibility of linearly invariant estimators in truncated parameter spaces. *J. Amer. Statist. Assoc.*, **76**, 910-915.

Morris, C.N. (1983). Parametric empirical Bayes inference: Theory and application. *J. Amer. Statist. Assoc.*, **78**, 47-65.

Muirhead, R.J. (1985). Estimating a particular function of the multiple-correlation coefficient. *J. Amer. Statist. Assoc.*, **80**, 923-925.

鍋谷清治. (1978). 数理統計学. 共立出版, 東京.

Nagata, Y. (1989). Improvements of interval estimations for the variance and the ratio of two variances. *J. Japan Statist. Soc.*, **19**, 151-161.

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.

Neff, N. and Strawderman, W.E. (1976). Further remarks on estimating the parameter of a noncentral chi-squared distribution. *Comm. Statist.* , **A5**, 65-76.

Nickerson, D.M. and Basawa, I.V. (1992). Shrinkage estimation for linear models with correlated errors. *Sankhya* (Ser. A), **54**, 411-424.

Nishii, R. and Krishnaiah, P.R. (1988). On the moments of classical estimates of explanatory variables under a multivariate calibration model. *Sankhya* (Ser. A), **50**, 137-148.

Okamoto, M. (1963). An asymptotic expansion for the distribution of the linear discriminant function. *Ann. Math. Statist.*, **34**, 1286-1301.

丘本 正 (1986). 因子分析の基礎. 日科技連出版社, 東京.

Osborne, C. (1991). Statistical calibration: A review. *Inter. Statist. Review*, **59**, 309-336.

Pal, N. and Sinha, B.K. (1989). Improved estimators of dispersion of an inverse Gaussian distribution. In *Statistical Data Analysis and Inference* (Y. Dodge, ed.), 215-222. Elsevier Science Publishers.

Peck, R. and Van Ness, J. (1982). The use of shrinkage estimators in linear discriminant analysis. *IEEE Trans. Pattern Anal. Mach. Intelligence*, **4**, 530-537.

Perlman, M.D. and Rasmussen, V.A. (1975). Some remarks on estimating a noncentrality parameter. *Comm. Statist.*, **4**, 455-468.

Perron, F. (1990). Equivariant estimators of the covariance matrix. *Canad. J. Statist.*, **18**, 179-182.

Perron, F. (1992). Minimax estimators of a covariance matrix. *J. Multivariate Anal.*, **43**, 16-28.

Peszek, I. and Rukhin, A.L. (1993). Estimating normal distribution function and normal density. *Statist. Decisions*, **11**, 391-406.

Pierce, D.A. (1973). Fiducial, frequency, and Bayesian inference on reliability for the two parameter negative exponential distribution. *Technometrics*, **15**, 249-253.

Portnoy, S. (1971). Formal Bayes estimation with application to a random effect model. *Ann. Math. Statist.*, **42**, 1379-1402

Prasad, N.G.N. and Rao, J.N.K. (1990). The estimation of the mean squared error of small-area estimators. *J. Amer. Statist. Assoc.*, **85**, 163-171.

Proskin, H.M. (1985). An admissibility theorem with applications to the estimation of the variance of the normal distribution. Ph.D. dissertation, Dept. Statist., Rutgers University.

Ralescu, S. Brandwein, A.C. and Strawderman, W.E. (1992). Stein estimation for non-normal spherically symmetric location families in three dimensions. *J. Multivariate Anal.*, **42**, 35-50.

Rao, J.N.K. and Yu, M. (1994). Small-area estimation by combining time-series and cross-sectional data. *Canad. J. Statist.*, **22**, 511-528.

Reiser, B. and Guttman, I. (1987). A comparison of three point estimators for $P(Y < X)$ in the normal case. *Compt. Statist. Data Anal.*, **5**, 59-66.

Remadi, S. and Amemiya, Y. (1994). Asymptotic properties of the estimators for multivariate components of variance. *J. Multivariate Anal.*, **49**, 110-131.

Robbins, H. (1983). Some thoughts on empirical Bayes estimation. *Ann. Statist.*, **11**, 713-723.

Robert, C.P. (1994). *The Bayesian Choice : A Decision-Theoretic Motivation.* Springer-Verlag, New York.

Robert, C. and Casella, G. (1990). Improved confidence sets for spherically symmetric distributions. *J. Multivariate Anal.*, **32**, 84-94.

Robert, C.P. and Casella, G. (1994). Improved confidence statements for the usual multivariate normal confidence set. In *Statistical Decision Theory and Related Topics* V (S.S. Gupta, J. Berger, eds.), 351-368. Springer-Verlag, New York.

Robert, C. and Saleh, A.K.Md.E. (1989). Recentered confidence sets: A review. MSI Technical Report, Cornell University.

Robertson, T., Wright, F.T. and Dykstra, R. (1988). *Order Restricted Statistical Inference.* North-Holland, New York.

Rukhin, A.L. (1986). Admissibility and minimaxity results in the estimation problem of exponential quantiles. *Ann. Statist.*, **14**, 220-237.

Rukhin, A.L. (1987). How much better are better estimators of a normal variance ? *J. Amer. Statist. Assoc.*, **82**, 925-928.

Rukhin, A.L. (1991). Asymptotic variance estimation in multivariate distributions. *J. Multivariate Anal.*, **38**, 366-384.

Rukhin, A.L. (1992a). Generalized Bayes estimators of a normal discriminant function. *J. Multivariate Anal.*, **41**, 154-162.

Rukhin, A.L. (1992b). Asymptotic risk behavior of mean vector and variance estimators and the problem of positive normal mean. *Ann. Inst. Statist. Math.*, **44**, 299-311.

Rukhin, A.L. (1995). Admissibility: Survey of a concept in progress. *Inter. Statist. Review*, **63**, 95-115.

Rukhin, A.L. and Ananda, M.M.A. (1989). Estimating exponential reliability function and exponential density. *Statist. Decisions*, **7**, 277-296.

Rukhin, A.L., Kuo, L. and Dey, D.K. (1990). A class of minimax estimators of the scale parameter of the uniform distribution. *Statist. Probab. Lett.*, **9**, 317-321.

Rukhin, A.L. and Sinha, B.K. (1991). Decision-theoretic estimation of the product of Gamma scales and generalized variance. *Calcutta Statist. Assoc. Bull.*, **40**, 257-265.

Rukhin, A.L. and Strawderman, W.E. (1982). Estimating a quantile of an exponential distribution. *J. Amer. Statist. Assoc.*, **77**, 159-162.

Rukhin, A.L. and Zidek, J.V. (1985). Estimation of linear parametric functions for several exponential samples. *Statist. Decisions*, **3**, 225-238.

Sackrowitz, H. and Samuel-Cahn, E. (1984). Estimation of the mean of a selected negative exponential population. *J. Roy. Statist. Soc.*, **46**, 242-249.

Saleh, A.K.Md.E. and Sen, P.K. (1985). Nonparametric shrinkage estimation in a parallelism problem. *Sankhya* (Ser. A), **47**, 156-165.

Sarkar, S.K. (1989). On improving the shortest length confidence interval for the generalized variance. *J. Multivariate Anal.*, **31**, 136-147.

Sarkar, S.K. (1994). Shrinkage domination of some usual estimators of the common mean of several multivariate normal populations. *J. Statist. Plann. Inference*, **39**, 43-55.

Saxena, K.M.L. and Alam, K. (1982). Estimation of the noncentrality parameter of a chi-squared distribution. *Ann. Statist.*, **10**, 1012-1016.

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.

Sen, P.K. and Saleh, A.K.Md.E. (1985). On some shrinkage estimators of multivariate location. *Ann. Statist.*, **13**, 272-281.

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.

Sengupta, D. and Sen, P.K. (1991). Shrinkage estimation in a restricted parameter space. *Sankhya* (Ser. A), **53**, 389-411.

Seshadri, V. (1963). Combining unbiased estimators. *Biometrics*, **19**, 163-170.

Shah, K.R. (1964). Use of inter-block information to obtain uniformly better estimators. *Ann. Math. Statist.*, **35**, 1064-1078.

Shao, P.Y.-S. and Strawderman. (1994). Improving on the James-Stein positive-part estimator. *Ann. Statist.*, **22**, 1517-1538.

Shao, P.Y.-S. and Strawderman. (1995). Improving on the positive part of the UMVUE of a noncentrality parameter of a noncentral chi-square distribution. *J. Multivariate Anal.*, **53**, 52-66.

Sharma, D. and Kumar, S. (1994). Estimating quantiles of exponential populations. *Statist. Decisions*, **12**, 343-352.

Sheena, Y. and Takemura, A. (1992). Inadmissibility of non-order-preserving orthogonally invariant estimators of the covariance matrix in the case of Stein's loss. *J. Multivariate Anal.*, **41**, 117-131.

Shieh, G. (1993). Empirical Bayes minimax estimators of matrix normal means for arbitrary quadratic loss and unknown covariance matrix. *Statist. Decisions*, **11**, 317-341.

Shinozaki, N. (1980). Estimation of a multivariate normal mean with a class of quadratic loss functions. *J. Amer. Statist. Assoc.*, **75**, 973-976.

Shinozaki, N. (1984). Simultaneous estimation of location parameters under quadratic loss. *Ann. Statist.*, **12**, 322-335.

Shinozaki, N. (1989). Improved confidence sets for the mean of a multivariate normal distribution. *Ann. Inst. Statist. Math.*, **41**, 331-346.

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.

Shinozaki, N. and Chang, Y.-T. (1994). Minimaxity of empirical Bayes estimators of the means of independent normal variables with unequal variances. *Comm. Statist. - Theory Methods*, **22**, 2147-2169.

Shinozaki, N. (1995). Some modifications of improving estimators of a normal variance. *Ann. Inst. Statist. Math.*, to appear.

Shiraishi, T. (1991). On positive-part shrinkage R- and M-estimation in one-way ANOVA. *J. Japan Statist. Soc.*, **21**, 61-72.

Shorrock, G. (1990). Improved confidence intervals for a normal variance. *Ann. Statist.*, **18**, 972-980.

Shorrock, R.B. and Zidek, J.V. (1976). An improved estimator of the generalized variance. *Ann. Statist.*, **4**, 629-638.

Sinha, B.K. (1976). On improved estimators of the generalized variance. *J. Multivariate Anal.*, **6**, 617-626.

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. Decisions*, **5**, 201-227.

Spruill, C. (1982). Admissibility of the natural estimator of the mean of a Gaussian process. *J. Multivariate Anal.*, **12**, 568-574.

Srinivasan, C. (1981). Admissible generalized Bayes estimators and exterior boundary value problems. *Sankhya* (Ser. A), **43**, 1-25.

Srivastava, M.S. and Bilodeau, M. (1989). Stein estimation under elliptical distributions. *J. Multivariate Anal.*, **28**, 247-259.

Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proc. Third Berkeley Symp. Math. Statist. Probab.*, **1**, 197-206. University of California University, Berkeley.

Stein, C. (1960). Multiple regression. In *Contributions to Probability and Statistics Essays in Honor of Harold Hotelling* (I. Olkin, et al., eds.), 424-443. Stanford University Press.

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.

Stein, C. (1966). An approach to the recovery of inter-block information in balanced incomplete block designs. In *Research Papers in Statistics* (Neyman Festschrift, F. N. David, ed.), 351-366. Wiley, New York.

Stein, C. (1973). Estimation of the mean of a multivariate normal distribution. In *Proc. Prague Symp. Asymptotic Statist.*, 345-381.

Stein, C. (1977). Lectures on multivariate estimation theory. (In Russian.) In *Investigation on Statistical Estimation Theory* I, 4-65. Zapiski Nauchych Seminarov LOMI im. V.A. Steklova AN SSSR vol. 74, Leningrad.(In Russian)

Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.*, **9**, 1135-1151.

Stigler, S.M. (1990). A Galtonian perspective on shrinkage estimators. *Statist. Science*, **5**, 147-155.

Strawderman. W.E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Statist.*, **42**, 385-388.

Strawderman. W.E. (1974a). Minimax estimation of powers of the variance of a normal population under squared error loss. *Ann. Statist.*, **2**, 190-198.

Strawderman. W.E. (1974b). Minimax estimation of location parameters for certain spherically symmetric distributions. *J. Multivariate Anal.*, **4**, 255-264.

Strawderman. W.E. (1978). Minimax adaptive generalized ridge regression estimators. *J. Amer. Statist. Assoc.*, **73**, 623-626.

Sugiura, N. and Fujimoto, M. (1982). Asymptotic risk comparison of improved estimators for normal covariance matrix. *Tsukuba J. Math.*, **6**, 103-126.

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.

Takada, Y. (1979). A family of minimax estimators in some multiple regression problems. *Ann. Statist.*, **7**, 1144-1147.

Takada, Y. (1984). Inadmissibility of a sequential estimation rule of the mean of a multivariate normal distribution. *Sequential Anal.*, **3**, 267-271.

Takemura, A. (1984). An orthogonally invariant minimax estimator of the covariance matrix of a multivariate normal population. *Tsukuba J. Math.*, **8**, 367-376.

竹内 啓 (1963). 統計的推定論 I. 数学 **14**, 193-209.

竹内 啓 (1965). 統計的推定論 II. 数学 **16**, 139-149.

Takeuchi, K. (1968). On the problem of fixing the level of independent variables in a linear regression function. IMM 367, Courant Institute of Math. Sciences, New York University.

竹内 啓 (1979). Stein 推定量の意味とその応用. 応用統計学 **8**, 81-95.

竹内 啓 (1980). 現象と行動のなかの統計数理. 新曜社.

竹内 啓 (1986). 因子分析モデルにおける最尤推定量の構造について. 応用統計学 **15**, 29-45.

Takeuchi, K. (1991). Personal communications.

Tan, M. (1991). Improved estimators for the GMANOVA problem with application to Monte Carlo simulation. *J. Multivariate Anal.*, **38**, 262-274.

丹後俊郎 (1988). 死亡指標の経験的ベイズ推定量について-疾病地図への適用-. 応用統計学 **17**, 81-96.

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.

Thompson, J.R. (1968). Some shrinkage techniques for estimating the mean. *J. Amer. Statist. Assoc.*, **63**, 113-122.

椿広計 (1982). 非心度・SN 比・対数 SN 比の点推定について. 21th JSQC 研究発表会, 56-59.

椿広計 (1985). 点推定における偏りの設計. 統計輪講資料.

Tsui, K.-W. (1984). Robustness of Clevenson-Zidek-type estimators. *J. Amer. Statist. Assoc.*, **79**, 152-157.

Tsui, K.-W. and Press, S.J. (1982). Simultaneous estimation of several Poisson parameters under K-normalized squared error loss. *Ann. Statist.*, **10**, 932-100.

Tsutakawa, R.K., Shoop, G.L. and Marienfeld, C.J. (1985). Empirical Bayes estimation of cancer mortality rates. *Statist. Medicine*, **4**, 201-212.

Varde, S.D. (1969). Life testing and reliability estimation for the two parameter exponential distribution. *J. Amer. Statist. Assoc.*, **64**, 621-631.

Vellaisamy, P. (1992). Inadmissibilty results for the selected scale parameters. *Ann. Statist.*, **20**, 2183-2191.

Venter, J.H. (1988). Estimation of the mean of the selected population. *Comm. Statist.-Theory Methods*, **17**, 791-805.

Wahba, G. (1983). Bayesian "confidence intervals" for the cross validated smoothing spline. *J. Royal Statist. Soc.* (Ser. B), **45**, 133-150.

Wahba, G. (1985). Comparison of GCV and GML for choosing the smoothing parameter in the generalized smoothing problem. *Ann. Statist.* , **13**, 1378-1402.

Wald, A. (1950). *Statistical Decision Functions.* Wiley, New York.

Yanagimoto, T. (1994). The Kullback-Leibler risk of the Stein estimator and the conditional MLE. *Ann. Inst. Statist. Math.*, **46**, 29-41.

Yang, R. and Berger, J.O. (1994). Estimation of a covariance matrix using the reference prior. *Ann. Statist.* , **22**, 1195-1211.

Yates, F. (1940). The recovery of inter-block information in balanced incomplete block designs. *Ann. Eugenics*, **10**, 317-325.

Zacks, S. and Even, M. (1966). The efficiency in small samples of the maximum likelihood and best unbiased estimators of reliability functions. *J. Amer. Statist. Assoc.*, **61**, 1033-1051.

Zacks, S. (1970). Bayes and fiducial equivariant estimators of the common mean of two normal distributions. *Ann. Math. Statist.*, **41**, 59-69.

Zacks, S. (1971). *The Theory of Statistical Inference*, Wiley, New York.

Zaman, A. (1981). A complete class theorem for the control problem, and further results on admissibility and inadmissibility. *Ann. Statist.*, **9**, 812-821.

Zheng, Z. (1986). On estimation of matrix of normal mean. *J. Multivariate Anal.*, **18**, 70-82.

Zidek, J. (1969). Inadmissibility of the best invariant estimator of extreme quantiles of the normal law under squared error loss. *Ann. Math. Statist.*, **40**, 1801-1808.

Zidek, J. (1971). Inadmissibility of a class of estimators of a normal quantile. *Ann. Math. Statist.*, **42**, 1444-1447.

Zidek, J. (1973). Estimating the scale parameter of the exponential distribution with unknown location. *Ann. Statist.*, **1**, 264-278.

Zidek, J. (1978). Deriving unbiased risk estimators of multinormal mean and regression coefficient estimators using zonal polynomials. *Ann. Statist.*, **6**, 769-782.

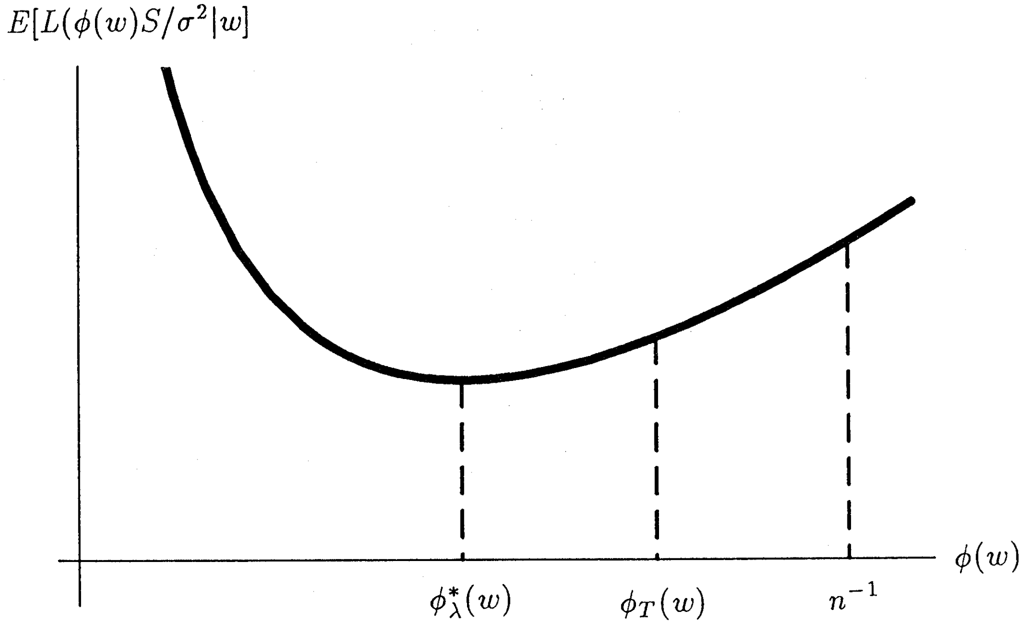

図 1.　条件付期待値 $E[L(\phi(w)S/\sigma^2|w]$ のグラフ.

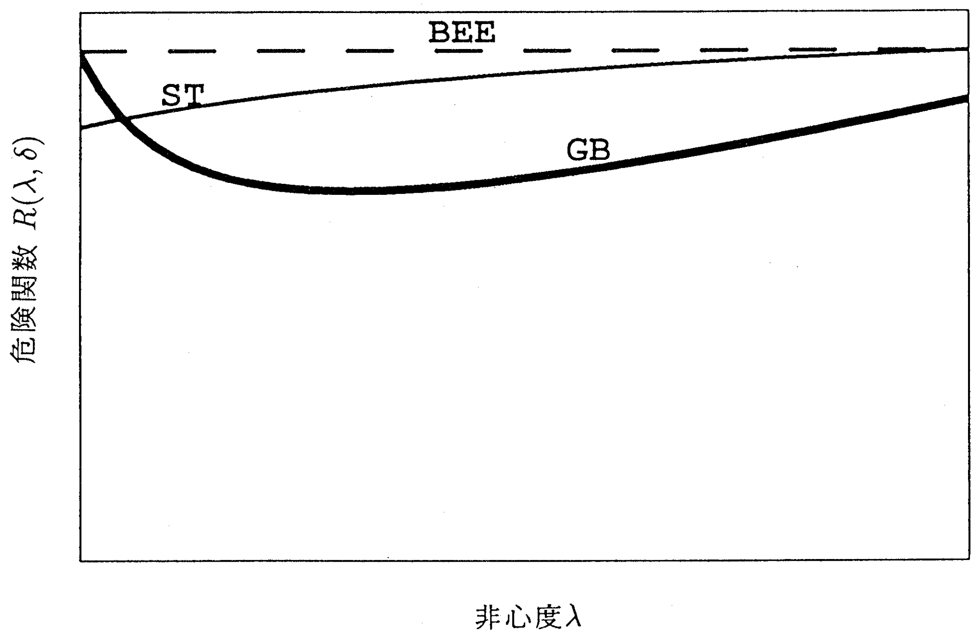

図 2. 分散の最良共変推定量δ_0(BEE), Stein 推定量δ^{ST}(ST), Brewster-Zidek 推定量δ^{BZ}(GB) の危険関数の挙動.

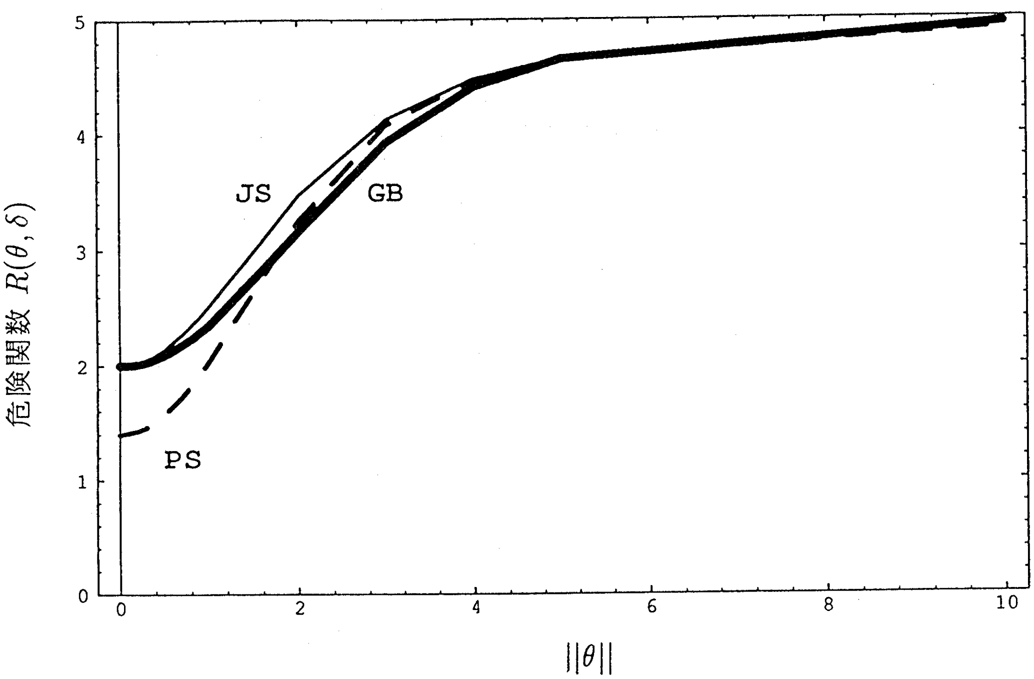

図 3. 平均ベクトルθの James-Stein 推定量δ^{JS}(JS), positive-part Stein 推定量δ_{+}^{JS}(PS), 一般化 Bayes 推定量δ_{GB}(GB) の危険関数の挙動.

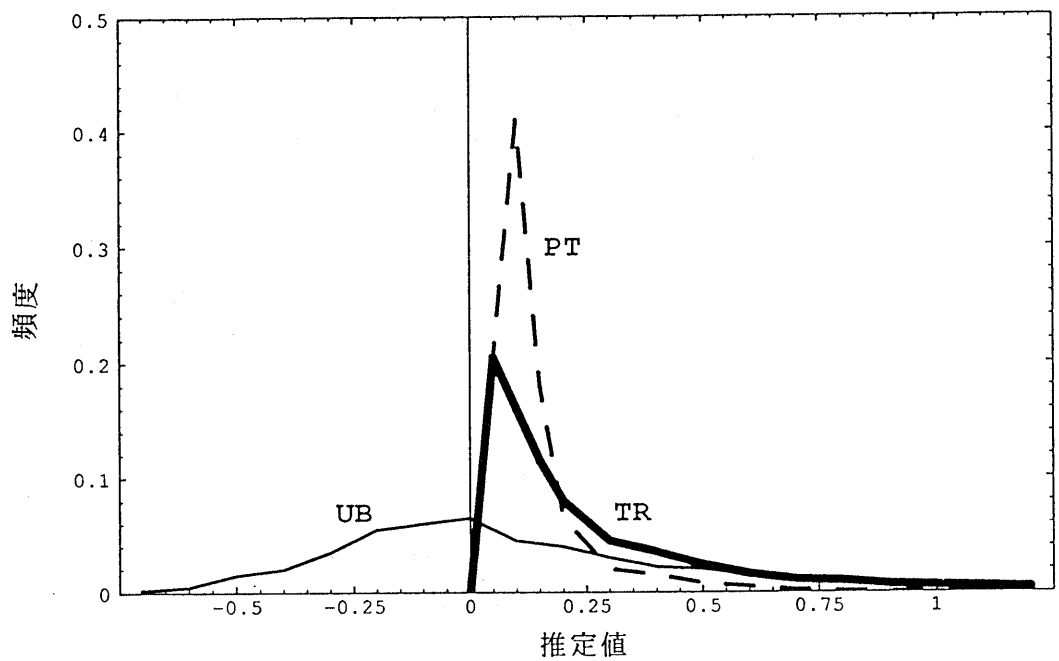

図 4. 不偏推定量δ_U(UB), 正値打ち切り推定量δ^{PT}(PT), 打ち切り推定量δ_0^+(TR) の分布の概形.

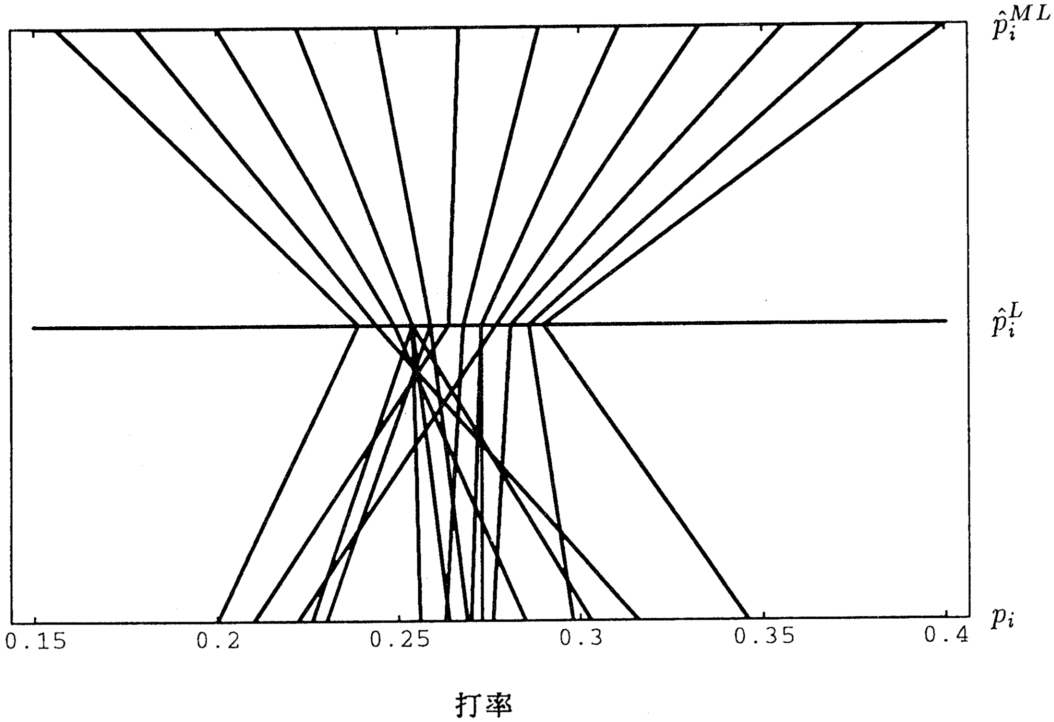

図 5. 野球選手の最終平均打率 (p_i) に対する最尤推定値 (データ平均)$(\hat{p}_i^{ML})$, 縮小推定値 $(\hat{p}_i^{L})$ とその縮小の様子.

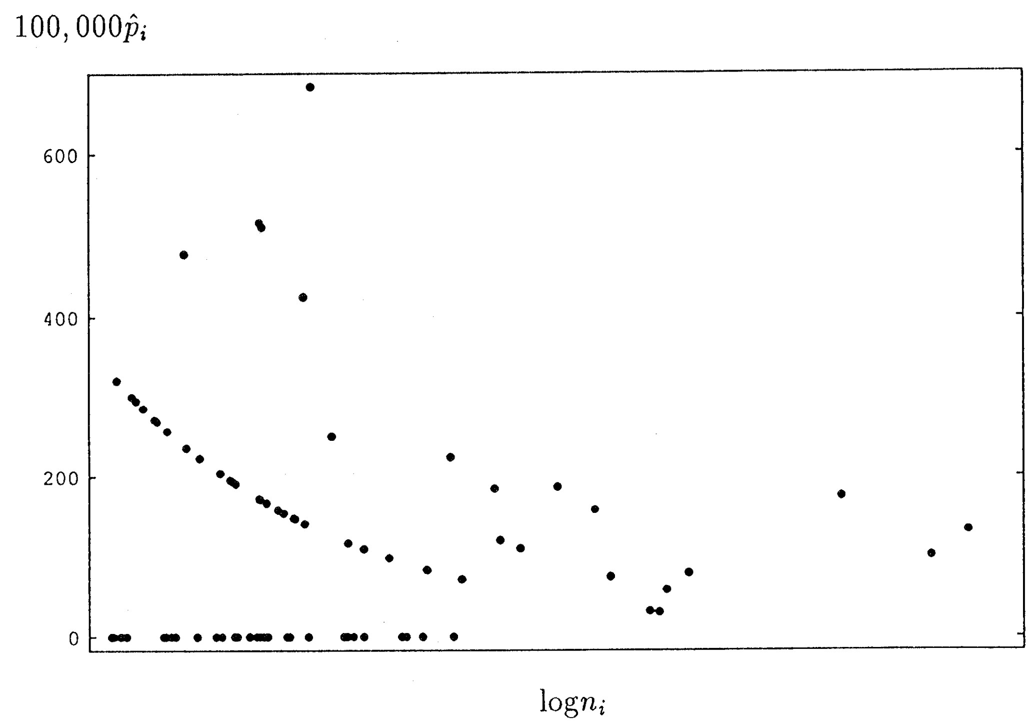

図 6.　胃ガンの都市別死亡率 (最尤推定値) の散布図. 横軸は logn_i, 縦軸は 100,000$\hat{p}_i$ のスケールで描いている.

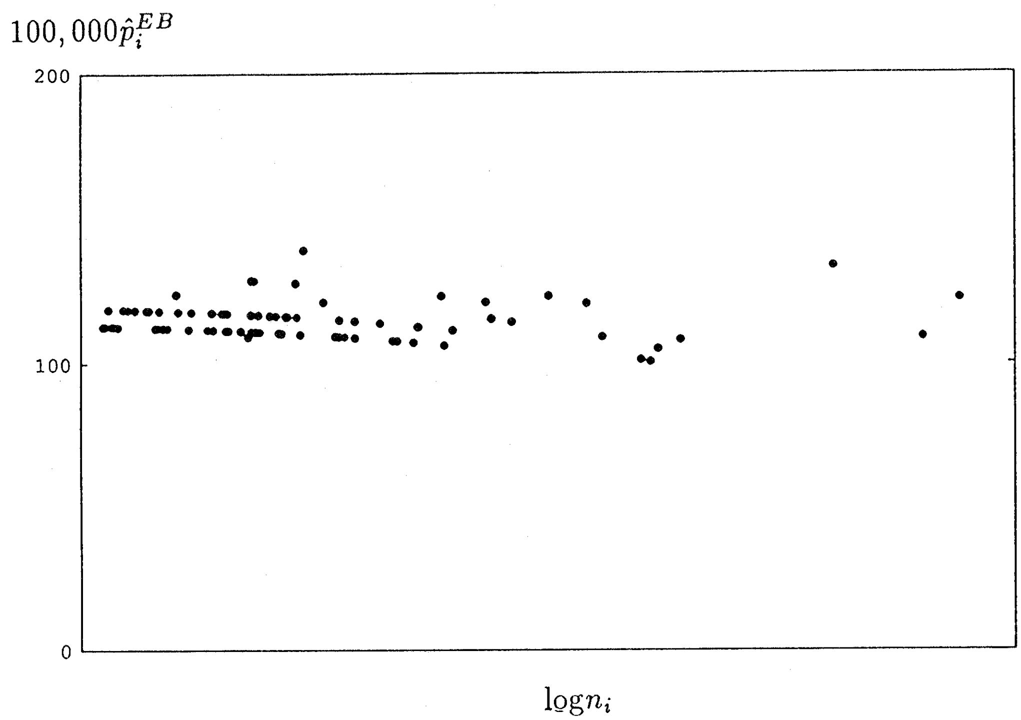

図 7. 胃ガンの都市別死亡率の経験 Bayes 推定値の散布図. 横軸は $\log n_i$, 縦軸は $100,000\hat{p}_i^{EB}$ のスケールで描いている.

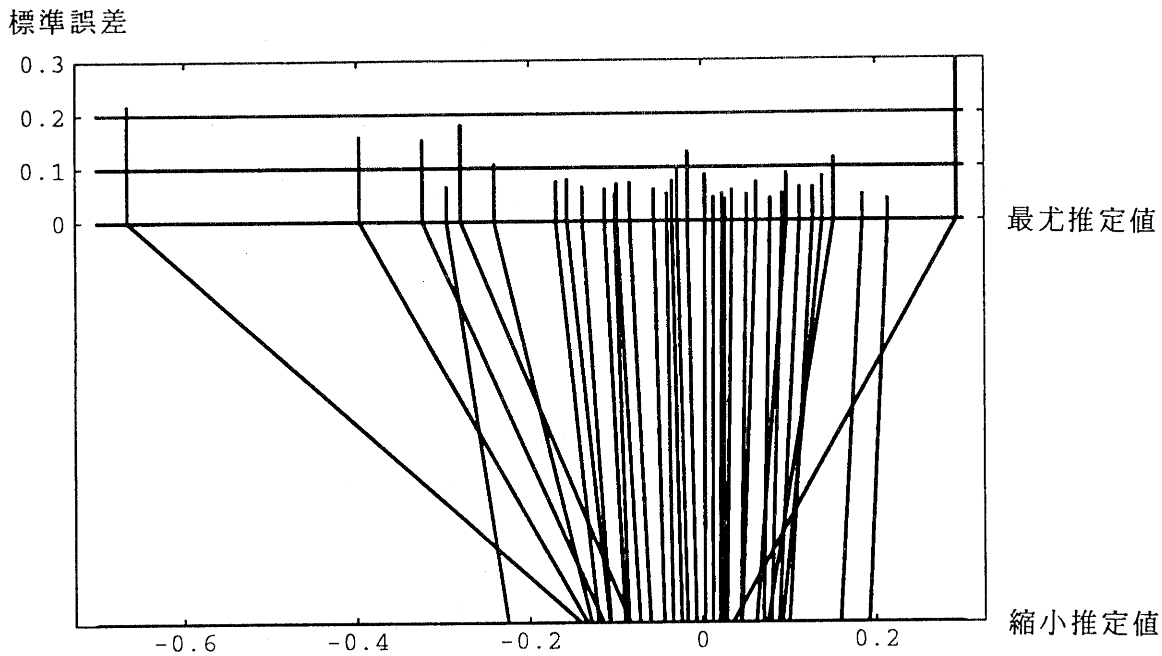

図 8. **トキソプラズマ症の都市別発生率に対する最尤推定値，その標準誤差及び経験 Bayes 型縮小推定値と縮小の様子．** 都市別発生率は全国平均からの差をとっている．バラツキ (標準誤差) の大きい推定値ほどより大きく縮小されていることがわかる．

表 1. 分散成分の推定量の危険関数の値

($\sigma^2 = 1$, $\sigma_A^2 = 0.1, 0.5, 1.0, 5.0, 10.0.$)

	σ_A^2	δ_U	δ_0	δ^{PT}	δ_0^+
	0.1	.130	.076	.010	.014
$m = 3$	0.5	.114	.060	.027	.033
	1.0	.111	.057	.040	.044
$k = 3$	5.0	.109	.055	.053	.054
	10.0	.109	.055	.054	.054
	0.1	.055	.042	.010	.012
$m = 3$	0.5	.047	.034	.023	.025
	1.0	.045	.032	.028	.029
$k = 6$	5.0	.044	.031	.031	.031
	10.0	.044	.031	.031	.031
	0.1	.029	.015	.004	.004
$m = 6$	0.5	.028	.014	.010	.011
	1.0	.027	.013	.012	.012
$k = 3$	5.0	.027	.013	.013	.013
	10.0	.027	.013	.013	.013