

(4) 「事前情報を持つ統計モデルの解析のための基礎理論とその応用」
に関する研究報告

- 笹渕祥一 (九州芸術工科大学・芸術工学部・共通講座) : 順序制約の下での 2
項分布の母数の均一性の検定に関する注意と検定統計量の比較 171
- 布能英一郎 (関東学院大学経済学部) : 母数制約のある推定問題について 173
- Toshio Sakata (Department of Computer Science Kumamoto University)・Ryuichi
Sawae (Department of Applied mathematics Okayama University of Science) :
Gröbner basis and walks on a set of contingency tables 175
- 工藤昭夫 (兵庫大学経済情報学部)・山本義郎 (北海道大学大学院工学研究科) :
(p, q) 配列上の順序制約の検定 177
- Lukman Thalib (Griffith University, Australia & NIAES, Japan) : Principal Compo-
nent Analysis for Grouped Bio-diversity Data 179
- 安楽和夫 (西南学院大学・文学部) : 順序制約の下での情報量基準の提案 181
- 栗木 哲 (統計数理研究所)・竹村彰通 (東京大学経済学部) : Anderson-Stephens
統計量の漸近帰無分布の裾確率 183
- 森 隆一 (京都産業大学・理学部)・三澤哲也 (名古屋市立大学・経済学部) :
Simulated Annealing による時系列データの Wavelet 近似 185
- 越水 孝 (大阪電気通信大学・大学院工学研究科)・辻谷将明 (大阪電気通信
大学・情報工学部) : ニューロ判別分析におけるモデル診断について
Model Diagnostics on Neural Discriminant Analysis 187
- 種村正美 (統計数理研究所)・Aarhus Univ. (E.B.Vedel Jensen)・Aalborg Univ.
(J.Møller) : MCMC 法によるランダムネットワークのベイズ的推論 189
- Yosihiko Ogata (The Institute of Statistical Mathematics) : Statistical analysis of un-
certain paleoearthquake occurrences times to forecast the hazard of rupture on a
fault segment 191

TAKA-AKI SHIRAISHI (Department of Mathematical Sciences Yokohama City University) : STUDENTIZED ROBUST STATISTICS FOR MAIN EFFECTS IN A TWO-FACTOR MANOVA	193
大津起夫 (北海道大学・文) : 等質性分析とペナルティ制約付き対応分析	195
A. J. Hayter (School of Industrial and Systems Engineering Georgia Institute of Technology) · Tetsuhisa Miwa (Laboratory of Statistics National Institute of Agro-Environmental Sciences) · Wei Liu (Department of Mathematics University of Southampton) : Combining the Advantages of One-sided and Two-sided Test Procedures for Comparing Several Treatments with a Control	197

順序制約の下での2項分布の母数の均一性の検定に関する注意と検定統計量の比較

九州芸術工科大学 芸術工学部 共通講座 笹渕祥一

X_1, \dots, X_k は互いに独立な確率変数で、 X_i は2項分布 $B(n_i, p_i)$ に従う ($i = 1, \dots, k$) とする。

次の検定問題を考える。

帰無仮説 $H_1: "p_1 = p_2 = \dots = p_k"$ 、対立仮説 $K_1: "p_1 \leq p_2 \leq \dots \leq p_k$ かつ H_1 ではない"。

1. 「用量・反応関係の解析の問題」との相違についての注意

まず、本講演で扱う問題は、「用量・反応関係の解析の問題」とは若干異なる事を注意しておく。

例えば、粟の量を d_1, d_2, \dots, d_k ($d_1 < d_2 < \dots < d_k$) とした時の反応確率 $P(d_1), P(d_2), \dots, P(d_k)$ に対し、帰無仮説を " $P(d_1) = P(d_2) = \dots = P(d_k)$ " とし対立仮説を " $P(d_1) \leq P(d_2) \leq \dots \leq P(d_k)$ " とする検定を行う問題、あるいは、関数 P の推定を行う問題は、「用量・反応関係の解析の問題」と呼ばれる。この場合には、 d_1, d_2, \dots, d_k の値を組み入れた統計量を用いるべきであり、Cochran(1954), Armitage(1955) を始めとする多くの研究がなされてきている。よく知られている Cochran・Armitage 検定等の統計量は d_1, d_2, \dots, d_k に依存している。(廣津(1982)第2章、柳川(1986)第4章、等参照。)

本講演では、 d_1, d_2, \dots, d_k に対応するデータはなく、単に k 個の2項分布確率変数のみが観測される場合 を考える。

2. 4つの検定方法の紹介

定義 1 (Barlow, Bartholomew, Bremner, and Brunk(1972, 1.2節))

x_1, \dots, x_k を与えられた k 個の実数、 w_1, \dots, w_k を与えられた k 個の正数とする。" $\theta_1 \leq \dots \leq \theta_k$ " なる条件の下で $\sum_{i=1}^k (x_i - \theta_i)^2 w_i$ を最小にする $(\theta_1, \dots, \theta_k) = (\theta_1^*, \dots, \theta_k^*)$ を、" w_1, \dots, w_k を重みとする x_1, \dots, x_k の Isotonic Regression (単調回帰)" と呼ぶ。簡単のために、" θ_i^* は w_i を重みとする x_i の IR" と略記する。

$\hat{p}_i = X_i/n_i$ ($i = 1, \dots, k$) とおく。この講演では、次の4つの検定について考察する。

検定 I : 逆正弦変換に基づく Isotonic Test (Barlow et al.(1972, 4.3節))

$$T_I = \sum_{i=1}^k 4n_i(U_i^* - \bar{U})^2 \geq c_1 \text{ の時、} H_1 \text{ を棄却。}$$

ここで、 $\bar{U} = \sum_{i=1}^k n_i U_i / \sum_{i=1}^k n_i$ 、 $U_i = \text{Sin}^{-1}\{(\hat{p}_i)^{1/2}\}$ 、 U_i^* は n_i を重みとする U_i の IR、 c_1 は正の定数。

検定 II : Isotonic Test (Bartholomew(1959), Shorack(1967), Barlow et al.(1972, 4.3節))

$$T_{II} = \sum_{i=1}^k \frac{n_i(\hat{p}_i^* - \hat{p})^2}{\hat{p}(1 - \hat{p})} \geq c_2 \text{ の時、} H_1 \text{ を棄却。}$$

ここで、 $\hat{p} = \sum_{i=1}^k X_i / \sum_{i=1}^k n_i$ 、 \hat{p}_i^* は n_i を重みとする \hat{p}_i の IR、 c_2 は正の定数。

検定 III : 尤度比検定 (Robertson and Wegman(1978), Robertson, Wright, and Dykstra(1988, 4.1節))

$$T_{III} = 2 \sum_{i=1}^k [n_i \hat{p}_i \log(\hat{p}_i^* / \hat{p}) + n_i(1 - \hat{p}_i) \log\{(1 - \hat{p}_i^*) / (1 - \hat{p})\}] \geq c_3 \text{ の時、} H_1 \text{ を棄却。}$$

ここで、 c_3 は正の定数。

検定 IV : 逆正弦変換に基づく Contrast Test (漸近最近迫局所最強検定)

(Abelson and Tukey (1963)、Schaafsma and Smid (1966)、Poon (1980))

$$T_{IV} = \frac{\sum_{i=1}^k 4n_i w_i U_i}{(\sum_{i=1}^k 4n_i w_i^2)^{1/2}} \geq c_4 \text{ の時、} H_1 \text{ を棄却。}$$

ここで、 c_4 は正の定数。スコア： w_1, \dots, w_k については、Robertson et al.(1988, 4.2節)等を参照。

3. 4つの検定方法に関する考察

H_1 の下で、 $n_i \rightarrow \infty (i = 1, \dots, k)$ の時、 T_I, T_{II}, T_{III} の漸近分布は、同一の“ χ^2 分布” (χ^2 分布の重み付き和)となる。(Collings, Margolin, and Oehlert(1981)、Robertson et al.(1988, 4.1節)、Kulatunga, Asai, and Sasabuchi(1996)) また、 T_{IV} の漸近分布は標準正規分布となる。従って、いずれの検定も、漸近分布により棄却点を定めて検定を行う事が考えられる。しかし、これが小標本の場合にも妥当であるかどうか問題となる。

Poon(1980)は、 T_{II} を用いて、 H_1 の下での条件付分布に基づく(漸近分布を用いない) exactな検定方法を提案している。そして、検定IV(及び別の2つの検定)と比較して、小標本の場合には、提案した検定が検出力等の点で優れている、という事を示している。が、大標本の場合には、この方法は計算が面倒、という事も述べている。Collings et al.(1981)は、漸近分布による棄却点を用いた場合には、検定IIは適切なサイズを持つが、検定Iはサイズが大きくなりすぎるため用いるのは不適當である、という事を示している。また、Poon(1980)とCollings et al.(1981)はいずれも、検定IIは、対立仮説 K_1 に対してだけでなく、unimodal(umbrella)対立仮説に対しても比較的高い検出力を持つ、という事も示している。一方、検定IIIについては、Robertson et al.(1988, 4.1節)が、“検定IやIIと似た挙動を示すであろう”と示唆している。

Kulatunga et al.(1996)は、 $n_1 = \dots = n_k = n$ として、 $n = 10, 25, 50$ の場合に、検定I~IVの各々について、漸近分布により棄却点を定めて検定を行う場合の実際の検定のサイズと検出力を、Monte Carlo simulation によって推定して、おおむね次のような結果を得た。(なお、この論文は、検定IIIを他の検定と比較するのが主な目的で、検定IとIIに対してはPoon(1980)やCollings et al.(1981)と同様の結論を得ている。)

(1) 検定IとIVは、漸近分布による棄却点を用いた場合、実際の検定のサイズが大きくなりすぎるため、用いるのは不適當である。(逆正弦変換の影響か?) (2) 検定IIとIIIは適切なサイズを持つ。(3) 検定IIIは、サイズや検出力に関して、検定IIと似た挙動を示すが、検定Iとはかなり異なる。(4) n が小さい時は、検定IIとIIIとでは、IIの方が検出力が高い場合が多い。 n が大きい時は、2つの検定間の検出力の差はあまりない。

4. 「非心度」を用いた検出力の比較

笹淵・寺西(1997)は、理論的に検出力を調べるために「非心度」を用いる事を考えた。ただし、前節の(1)により、検定IとIVは考察の対象からはずし、検定IIとIIIのみを考察した。

簡単のために、 $n_1 = \dots = n_k = n$ として、一般的な「非心度」の定義を述べる。

定義 2 θ を未知の k 次元母数ベクトルとし、 θ に関する検定問題: H vs. K を考察する。 Y_n は、 k 次元確率ベクトルで、 θ の不偏推定量かつ一致推定量であるとする。 Y_n の関数 $T_n(Y_n)$ に対し、「 $T_n(Y_n) \geq C \Rightarrow H$ を棄却」なる検定を考える。ただし、 C は $\lim_{n \rightarrow \infty} \sup_H P_\theta(T_n(Y_n) \geq C) = \alpha$ を満たす正の定数である。このとき、 $T_n(\theta)$ を「 $T_n(Y_n)$ の非心度」と呼ぶ。

「非心度」によって検出力を比較する理論的根拠についての考察は、笹淵・寺西(1997)参照。

しかし、非心度からは、検出力の大小の比較の目安は得られても、検出力の値そのものは得られない。そこで、笹淵・寺西(1997)は、 K_1 の下での T_{II} と T_{III} の漸近正規性を用いて、検出力の近似を行った。(実は、漸近正規性が成り立つ場合には、「非心度」は漸近平均に他ならない。)ただし、 T_{II} と T_{III} の式には、順序制約: “ $p_1 \leq p_2 \leq \dots \leq p_k$ ”の下での最尤推定量 (\hat{p}_i^*)が含まれているため、母数の真値が K_1 の境界上にある場合には、漸近正規性は自明とは言えず、これを示すには若干の工夫が必要である。(阪大の岩佐学氏による。)

笹淵・寺西(1997)は、まず、 T_{II} と T_{III} に対して、各々の非心度(漸近平均)と漸近分散を計算した。次に、 $n = 10, 25, 50$ の場合に、 K_1 の様々な点において、Monte Carlo simulationを用いて検出力を推定し、非心度と検出力の関係、正規近似による検出力の近似値と実際の検出力との差、等を調べた。その主な結果は次の通り。

(1) 検定統計量を1つ固定して、対立仮説内の相異なる点での検出力を比較する場合。 n の大きさによって若干の違いはあるが、検出力の比較は、非心度や正規近似によってかなりできる。(2) 対立仮説内の点を固定して、2つの検定統計量の検出力を比較する場合。 n が小さい時には、非心度や正規近似によって検出力を比較する事は困難。 n が大きい時には、検出力の比較は、非心度や正規近似によってかなりできる。(3) 検出力の大小を、非心度あるいは正規近似によって判断する場合、正規近似による判断の方が若干正確なようである。(4) 正規近似による検出力の近似値は、実際の検出力の値より小さい場合が多い。

以上の事から、非心度は、検定手法を1つ定めた場合に、その検定法の検出力の対立仮説の中での挙動を知るのには有効であると思われる。

母数制約のある推定問題について

関東学院大学経済学部 布能英一郎

1. Introduction

二項分布または多項分布で母数に制限がある場合、自乗損失下で MLE が非許容的となることがある。たとえば、 $X \sim \text{Binomial}(n, \theta)$ で $\theta \in \Theta^* = [\zeta, 1 - \zeta] \subset [0, 1]$ (但し、 $0 < \zeta < 1/n$ を仮定) の場合、MLE よりも、より優れた推定量を構成することができる。他方、二項分布で母数 θ が $0 < \theta < \alpha$ 但し $0 < \alpha < 1$ に制限されている場合、 θ_{MLE} を優越する推定量を構成することが大変困難である。しかしながら、Brown による完備性定理を用いると、サンプルサイズ n がある程度大きければ、自乗損失下で MLE の非許容性が示せる。

2. Complete class theorem を用いた非許容性の証明

定義、仮定条件 \mathcal{X} を標本空間とする。そして \mathcal{X} は有限と仮定する。 \mathcal{D} : 決定空間, L : 損失関数, Θ : 母数空間とし、統計的決定問題 $(\mathcal{X}, \Theta, \mathcal{D}, L)$ を (Θ, \mathcal{X}) で略記する。

\mathcal{X} の空でない部分集合 \mathcal{X}' に対し、 $\Theta(\mathcal{X}') \equiv \{\theta \in \Theta : g(\theta) = \sum_{y \in \mathcal{X}'} P(y|\theta) > 0\}$ と定める。そうすると、標本空間を \mathcal{X}' 母数空間を $\Theta(\mathcal{X}')$ とする restricted probability distribution $P_{\mathcal{X}'}(x|\theta) = P(x|\theta)/g(\theta)$ が well-defined である。

Theorem 2.1 (Brown 1981, Ann.Stat.) δ が (Θ, \mathcal{X}) で許容的とする。 \mathcal{X} の部分集合 \mathcal{X}' が $\Theta(\mathcal{X}') \neq \emptyset$ で、更に

- (1) 各 $x \in \mathcal{X}'$ に対し $P_{\mathcal{X}'}(x|\theta)$ が $\overline{\Theta(\mathcal{X}')}$ 上で well defined
- (2) 各 $d \in \mathcal{D}$ に対して、損失関数 $L(\theta, d)$ は $\Theta(\mathcal{X}')$ 上 θ の連続関数
- (3) $\overline{\Theta(\mathcal{X}')}$ は compact

の条件が満たされるならば、 δ がベイズとなるような $\overline{\Theta(\mathcal{X}')}$ 上の確率測度 $d\tau(\theta)$ が存在する

Theorem 2.2 $X \sim \text{Binomial}(n, \theta)$ 但し $\Theta = [0, \alpha]$, $2/n < \alpha < 1$. このとき θ_{MLE} は自乗損失下で非許容的

証明の概要: θ_{MLE} が許容的と仮定する。 $\mathcal{X}' = \{1, 2, \dots, n\}$ に選ぶ。 $\Theta(\mathcal{X}') = (0, \alpha]$, $\overline{\Theta(\mathcal{X}')} = [0, \alpha] = \Theta$ であり、 $\lim_{\theta \rightarrow 0} P_{\mathcal{X}'}(x|\theta)$ は計算により、 $x=1$ に対しては 1、 $x=2, 3, \dots, n$ に対しては 0。

Theorem 2.1 により $\exists d\tau(\theta)$ on $[0, \alpha]$ such that θ_{MLE} は $d\tau(\theta)$ を prior とするベイズ。

$d\tau(\theta)$ を $\theta = 0$ に集中する部分、 $\theta = \alpha$ に集中する部分、 $0 < \theta < \alpha$ の部分に分解して考える。 $X=1$ にてベイズ解と MLE が等しいことから $d\tau(\theta)$ が $\theta = 0$ に 100 パーセント集中することはない。次に $X = n$ にてベイズ解と MLE が等しいことから $d\tau(\theta)$ は $0 < \theta < \alpha$ の部分に確率がなく。そうすると $X = 2$ を観測した時のベイズ解は、計算により α となるが MLE は $2/n$ で一致せず、矛盾。

3. センサリングメカニズムを持つような多項分布からの観測への応用

Meeden, Ghosh, Srinivasan, Vardeman(1989, Ann.Stat.) は Kaplan-Meier 推定量の自乗損失下での許容性を証明したが、同時に、多項分布からの観測が樹木構造のセンサリングメカニズムを持つような状況を考え、この場合の自乗損失下での MLE の許容性を証明した。他方、センサリングメカニズムが樹木構造でなく、オーバーラップが生じる時の MLE は単純計算では求められず、EM アルゴリズム等の繰返し計

算で求めることになる。そして MLE が許容的であるか否かの考察は、おこなわれていなかった。

さて、上記のセンサリングメカニズムが樹木構造でない場合「見かけ上、母数に制限がなくとも、何らかの制約によって、標本空間の端点以外を観測しても MLE が母数空間の境界に値を取る」ことがわかり、このような時に、完備性定理を用いる方法を適用してみたところ、自乗損失下で MLE の非許容性が示せた。

Example 3.1 $(X_1, X_2, X_3) \sim \text{Multinomial}(n, (\theta_1, \theta_2, \theta_3))$ であるがセンサリングにより

outcome	probability	number of observations
$\{X_1, X_2\}$	$\xi_{12} = (\theta_1 + \theta_2)/2$	x_{12}
$\{X_2, X_3\}$	$\xi_{23} = (\theta_2 + \theta_3)/2$	x_{23}
$\{X_1, X_3\}$	$\xi_{13} = (\theta_1 + \theta_3)/2$	x_{13}

で観測されている状況を考える。 θ には $0 \leq \theta_1, \theta_2, \theta_3 \leq 1, \theta_1 + \theta_2 + \theta_3 = 1$ 以外の制限はないが、各 ξ_{ij} には $0 \leq \xi_{ij} \leq 1/2$ なる母数制約がある。このため

(1) $x_{12} > x_{23} + x_{13}$ の場合

$$\xi_{12\text{MLE}} = 1/2, \xi_{23\text{MLE}} = x_{23}/2(x_{23} + x_{13}), \xi_{13\text{MLE}} = x_{13}/2(x_{23} + x_{13})$$

なぜなら、(1) の場合、 $\xi_{12\text{MLE}} = x_{12}/(x_{12} + x_{23} + x_{13}) > 1/2$ であるから $\xi_{12\text{MLE}} = 1/2$ である。更に $Likelihood \propto \xi_{23}^{x_{23}} \xi_{13}^{x_{13}} = \xi_{23}^{x_{23}} (1/2 - \xi_{23})^{x_{13}}$ により $\xi_{23\text{MLE}} = x_{23}/2(x_{23} + x_{13}), \xi_{13\text{MLE}} = x_{13}/2(x_{23} + x_{13})$. これを linear parameter transformation することで

$$\theta_{1\text{MLE}} = x_{13}/(x_{23} + x_{13}), \theta_{2\text{MLE}} = x_{23}/(x_{23} + x_{13}), \theta_{3\text{MLE}} = 0$$

を得る。同様に、

(2) $x_{23} > x_{12} + x_{13}$ の場合

$$\theta_{1\text{MLE}} = 0, \theta_{2\text{MLE}} = x_{12}/(x_{12} + x_{13}), \theta_{3\text{MLE}} = x_{13}/(x_{12} + x_{13})$$

(3) $x_{13} > x_{12} + x_{23}$ の場合

$$\theta_{1\text{MLE}} = x_{12}/(x_{12} + x_{23}), \theta_{2\text{MLE}} = 0, \theta_{3\text{MLE}} = x_{23}/(x_{12} + x_{23})$$

このとき $\mathcal{X}' = \{(x_{12}, x_{23}, x_{13}) \in \mathcal{X} | x_{12} \geq 1, x_{23} \geq 1, x_{13} \geq 1\}$ に選び、Theorem 2.2 と同様な考え方で議論を行うと $n \geq 6$ にて θ_{MLE} は自乗損失下で非許容的。

Example 3.2 $(X_1, X_2, X_3) \sim \text{Multinomial}(n, (\theta_1, \theta_2, \theta_3))$ であるがセンサリングにより

outcome	probability	number of observations
X_1	$\xi_1 = \theta_1 \zeta$	x_1
X_2	$\xi_2 = \theta_2 \zeta$	x_2
X_3	$\xi_3 = \theta_3 \zeta$	x_3
$\{X_1, X_2\}$	$\xi_{12} = (\theta_1 + \theta_2)(1 - \zeta)/2$	x_{12}
$\{X_2, X_3\}$	$\xi_{23} = (\theta_2 + \theta_3)/(1 - \zeta)2$	x_{23}
$\{X_1, X_3\}$	$\xi_{13} = (\theta_1 + \theta_3)/(1 - \zeta)2$	x_{13}

で観測されている状況を考える。 (θ, ζ) には ζ が $0 < \zeta \leq 1$ の範囲を自由に動くのなら、Example 3.1 のような restriction はないが、 ζ を $0 \leq \zeta < 1$ に固定すると、Example 3.1 同様の restriction がある。そこで、特に $\zeta = 0$ の場合を念頭に入れると、Example 3.1 の考察を拡張した考えで θ_{MLE} の自乗損失下での非許容性が (n がある程度大きいときに) 示せる。

Example 3.2 の一般化は容易に行える。更に censoring mechanism が informative であっても、 θ_{MLE} の自乗損失下で非許容的となる場合が見つかっている。

Gröbner basis and walks on a set of contingency tables

Toshio Sakata and Ryuichi Sawae

Department of Computer Science Department of Applied mathematics
Kumamoto University Okayama University of Science

[1]. **Introduction.** Metropolis walks on the set of contingency tables were used in Sakata-Yanagawa[2] and Sakata-Nomakuchi[3]. A new theory of the relation between the generation of Metropolis walks and Gröbner basis was found in Diaconis-Strumfels(1998)[1]. According to their theory we made several programs of generating Metropolis walks on the set of contingency tables and reported several simulation results. Our programs will be useful for our future studies.

[2] **Markov Basis and Gröbner Basis.** Let \mathcal{X} be a finite set and T be a statistic over \mathcal{X} and $g_1, g_2, \dots, g_L : \mathcal{X} \rightarrow \mathcal{Z}$ where $\mathcal{Z} = \{0, \pm 1, \pm 2, \dots\}$ and \mathcal{G}_t be the set $\{g : \mathcal{X} \rightarrow \mathcal{N} : \sum_x g(x)T(x) = t\}$. If the conditions

- (a) $\sum_x g_i(x)T(x) = 0$ for $1 \leq i \leq L$
- (b) for any t and any pair g and $g' \in \mathcal{G}_t$, there are $(\epsilon_1, g_{i_1}), (\epsilon_1, g_{i_2}), \dots, (\epsilon_A, g_{i_A})$ such that

$$g' = g + \sum_{j=1}^A \epsilon_j g_{i_j}, \text{ and } g + \sum_{j=1}^a \epsilon_j g_{i_j} \geq 0 \text{ for } 1 \leq a \leq A, \epsilon_j \in \{-1, 1\},$$

are satisfied, then g_1, \dots, g_L are said to be a Markov basis. From a Markov basis a Metropolis walk is derived as follows.

Let assume that we are now at the position $g \in \mathcal{G}_t$. Choose I from $\{1, 2, \dots, L\}$ randomly and ϵ from $\{-1, 1\}$ randomly. If $g + \epsilon g_I$ is nonnegative, the chain moves $g + \epsilon g_I$, and otherwise stay at g . Then in the limit the Markov chain has the uniform distribution on \mathcal{G}_t . Note that the condition (a) implies that $g + \epsilon g_I \in \mathcal{G}_t$ and that the condition (b) implies that the chain is connected. Next proposition gives the method that realizes a hypergeometric distribution on \mathcal{G}_t .

Proposition Let $\sigma(g)$ be a positive function on \mathcal{G}_t . Even if $g + \epsilon g_I \geq 0$ we move only with probability

$$\min \left\{ \frac{\sigma(g + \epsilon g_I)}{\sigma(g)}, 1 \right\}$$

then the limit distribution is proportional to $\sigma(g)$.

To calculate Gröbner bases they used MACAULAYII due to Grayson and Stillman. However we used Risa /Asir, which has been developed at Fujitsu Laboratory Limited.

[3] **Example.** Here we consider $2 \times 2 \times 2$ contingency tables and conditional inference with fixed marginals $x_{.11}, x_{.12}, x_{.21}, x_{.22}$ which correspond to the variables $y_{11}, y_{12}, y_{21}, y_{22}$. Another fixed marginals are $x_{1.1}, x_{1.2}, x_{2.1}, x_{2.2}$, which correspond to the variable $z_{11}, z_{12}, z_{21}, z_{22}$. And the remaining fixed marginals are $x_{11.}, x_{12.}, x_{21.}, x_{22.}$ which correspond to the variables $u_{11}, u_{12}, u_{21}, u_{22}$. If we set

$$B = [\begin{array}{l} x_{111} - y_{11} * z_{11} * u_{11}, x_{112} - y_{12} * z_{12} * u_{11}, x_{121} - y_{21} * z_{11} * u_{12}, x_{122} - y_{22} * z_{12} * u_{12}, \\ x_{211} - y_{11} * z_{21} * u_{21}, x_{212} - y_{12} * z_{22} * u_{21}, x_{221} - y_{21} * z_{21} * u_{22}, x_{222} - y_{22} * z_{22} * u_{22} \end{array}]$$

and

$$V = [\begin{array}{l} y_{11}, y_{12}, y_{21}, y_{22}, z_{11}, z_{12}, z_{21}, z_{22}, u_{11}, u_{12}, u_{21}, u_{22}, x_{111}, x_{112}, x_{121}, x_{122}, \\ x_{211}, x_{212}, x_{221}, x_{222} \end{array}]$$

then Asir put out the following basis. $gr(B, V, 2)$;

$$\begin{aligned}
& [x_{221} * x_{212} * x_{122} * x_{111} - x_{222} * x_{211} * x_{121} * x_{112}, u_{22} * x_{212} * x_{122} * u_{11} - x_{222} * u_{21} * u_{12} * x_{112}, \\
& -x_{221} * u_{21} * u_{12} * x_{111} + u_{22} * x_{211} * x_{121} * u_{11}, x_{222} * u_{12} * z_{12} - u_{22} * z_{22} * x_{122}, \\
& -x_{221} * x_{212} * u_{12} * z_{12} * x_{111} + u_{22} * z_{22} * x_{211} * x_{121} * x_{112}, x_{212} * z_{12} * u_{11} - z_{22} * u_{21} * x_{112}, \\
& x_{221} * z_{22} * u_{21} * x_{122} * x_{111} - x_{222} * x_{211} * x_{121} * z_{12} * u_{11}, x_{221} * u_{12} * z_{11} - u_{22} * z_{21} * x_{121}, \\
& u_{22} * x_{212} * z_{21} * x_{122} * x_{111} - x_{222} * x_{211} * u_{12} * x_{112} * z_{11}, -u_{21} * z_{21} * x_{111} + x_{211} * u_{11} * z_{11}, \\
& x_{221} * x_{212} * x_{122} * u_{11} * z_{11} - x_{222} * u_{21} * z_{21} * x_{121} * x_{112}, x_{221} * z_{22} * x_{122} * z_{11} \\
& -x_{222} * z_{21} * x_{121} * z_{12}, -x_{212} * z_{21} * z_{12} * x_{111} + z_{22} * x_{211} * x_{112} * z_{11}, \\
& -u_{22} * z_{22} * y_{22} + x_{222}, x_{212} * x_{122} * u_{11} - z_{22} * u_{21} * y_{22} * u_{12} * x_{112}, \\
& x_{221} * z_{22} * u_{21} * y_{22} * u_{12} * x_{111} - x_{222} * x_{211} * x_{121} * u_{11}, -y_{22} * u_{12} * z_{12} + x_{122}, \\
& x_{221} * x_{122} * z_{11} - u_{22} * z_{21} * y_{22} * x_{121} * z_{12}, u_{22} * x_{212} * z_{21} * y_{22} * z_{12} * x_{111} \\
& -x_{222} * x_{211} * x_{112} * z_{11}, x_{222} * z_{21} * y_{21} - x_{221} * z_{22} * y_{22}, -x_{212} * z_{21} * x_{122} * y_{21} * x_{111} \\
& + z_{22} * x_{211} * y_{22} * x_{121} * x_{112}, -u_{22} * z_{21} * y_{21} + x_{221}, u_{21} * z_{21} * x_{122} * y_{21} * x_{111} \\
& -x_{211} * y_{22} * x_{121} * z_{12} * u_{11}, -u_{21} * z_{21} * u_{12} * y_{21} * x_{111} + x_{211} * x_{121} * u_{11}, \\
& -x_{212} * z_{21} * u_{12} * y_{21} * z_{12} * x_{111} + z_{22} * x_{211} * x_{121} * x_{112}, x_{122} * y_{21} * z_{11} - y_{22} * x_{121} * z_{12}, \\
& x_{221} * x_{212} * y_{22} * z_{12} * x_{111} - x_{222} * x_{211} * y_{21} * x_{112} * z_{11}, \\
& -u_{12} * y_{21} * z_{11} + x_{121}, x_{222} * u_{21} * y_{12} - u_{22} * x_{212} * y_{22}, \\
& -x_{221} * u_{21} * x_{122} * y_{12} * x_{111} + u_{22} * x_{211} * y_{22} * x_{121} * x_{112}, x_{122} * y_{12} * u_{11} - y_{22} * u_{12} * x_{112}, \\
& x_{221} * x_{212} * y_{22} * u_{12} * x_{111} - x_{222} * x_{211} * x_{121} * y_{12} * u_{11}, -z_{22} * u_{21} * y_{12} + x_{212}, \\
& -x_{212} * z_{21} * u_{12} * y_{21} * x_{111} + z_{22} * x_{211} * x_{121} * y_{12} * u_{11}, u_{21} * z_{21} * x_{122} * y_{12} * x_{111} \\
& -x_{211} * y_{22} * u_{12} * x_{112} * z_{11}, x_{221} * u_{21} * z_{12} * y_{12} * x_{111} - u_{22} * x_{211} * y_{21} * x_{112} * z_{11}, \\
& -z_{12} * y_{12} * u_{11} + x_{112}, -u_{21} * z_{21} * z_{12} * y_{12} * x_{111} + x_{211} * x_{112} * z_{11}, x_{221} * x_{212} * y_{11} \\
& -u_{22} * z_{22} * x_{211} * y_{21} * y_{12}, -u_{12} * y_{21} * z_{12} * y_{12} * x_{111} + x_{121} * x_{112} * y_{11}, x_{221} * u_{21} * y_{11} \\
& -u_{22} * x_{211} * y_{21}, -u_{12} * y_{21} * x_{111} + x_{121} * u_{11} * y_{11}, x_{212} * z_{21} * y_{11} - z_{22} * x_{211} * y_{12} \\
& , -u_{21} * z_{21} * y_{11} + x_{211}, -z_{12} * y_{12} * x_{111} + x_{112} * z_{11} * y_{11}, x_{111} - u_{11} * z_1 * y_{11}]
\end{aligned}$$

Among these only one element is the Gröbner basis for the ideal \mathcal{J}_T (for the definition see [1]), $x_{221} * x_{212} * x_{122} * x_{111} - x_{222} * x_{211} * x_{121} * x_{112}$. Thus the ideal \mathcal{J}_T is generated by the unique polynomial which corresponds to the simple movement,

$$\begin{aligned}
& + - + - \\
& - + - + .
\end{aligned}$$

[IV] Conclusion. For more complex contingency tables we will have non-trivial movements automatically. That is the aim of this theory. Our plan is to construct c-programs which (1) automatically visualize all admissible movements corresponding to bases put out by Asir for a given type of tables, (2) automatically generate Metropolis walks and calculate various p-values by using the walk. They are still under construction. Finally we note that both of calculating times of Gröbner bases and generating time of the walks are increase rapidly if the size of tables becomes bigger.

[V] Reference.

1. P.Diaconis and B.Strumfels(1998).Algebraic algorithms for sampling from conditional distributions. A.S. Vol.26,No.1,363-397.
2. Toshio Sakata and Takasi Yanagawa(1998). Determination of the No-Observed-Adverse-Effect-Level by controlling over-estimation probability. Statistical research report No.SRR 007-98,Australian National University.
3. Generalized shift test method and Metropolis walk(1997).Bull. of Info. Cybernetics, **29**,1-13

(p, q) 配列上の順序制約の検定

兵庫大学経済情報学部 工藤昭夫
北海道大学大学院工学研究科 山本義郎

1 An optimum contrast

x_{ij} , $i = 1, 2, \dots, p$, $j = 1, 2, \dots, q$ を $E(x_{ij}) = \theta_{ij}$, $\text{Var}(x_{ij}) = 1$ である独立正規変量とする. x_{ij} とその母数 θ_{ij} の $p \times q$ 配列はそれぞれ行列 X, Θ と書ける, つまり

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1q} \\ \vdots & & \vdots \\ x_{p1} & \cdots & x_{pq} \end{pmatrix}, \quad \Theta = \begin{pmatrix} \theta_{11} & \cdots & \theta_{1q} \\ \vdots & & \vdots \\ \theta_{p1} & \cdots & \theta_{pq} \end{pmatrix}. \quad (1)$$

一般性を失うこと無く, 以降 $p < q$ であるものとする. このように格子点上に配置された母数に対して, Loop order や Tree order がある場合の検定問題を考える. まず, 各行と列にそれぞれ simple order を導入した場合を考えると順序制約は以下のように与えられる.

$$\left\{ \begin{array}{l} \theta_{11} \geq \cdots \geq \theta_{1q} \\ \vee \qquad \qquad \qquad \vee \\ \vdots \qquad \qquad \qquad \vdots \\ \vee \qquad \qquad \qquad \vee \\ \theta_{p1} \geq \cdots \geq \theta_{pq} \end{array} \right. \quad (2)$$

この semi-order を, (p, q) 格子上の Loop order と呼ぶ. $p = q = 2$ のときは simple loop order である. 平均値の同一性を帰無仮説 H とし, (2) のような順序を対立仮説 K とする検定問題を考える. この検定問題に対する尤度比検定の理論は存在する (cf. Kudo & Choi(1975)) が, 統計量の計算及び帰無仮説の下での確率分布の計算は不等式数 $p(q-1) + (p-1)q$ が変量数 $p \times q$ より多いため不可能である.

本報告では, 次の形式をした optimum contrast が定理 1 で与えられることを報告した.

$$L(X, C) = \sum_{i=1}^p \sum_{j=1}^q c_{ij} x_{ij} \quad \text{with} \quad \sum_{i=1}^p \sum_{j=1}^q c_{ij} = 0 \quad (3)$$

ここで係数 c_{ij} の集合を C で表わす. 有為水準 α の optimum contrast test は

$$L(X, C) / \sqrt{\sum_{i=1}^p \sum_{j=1}^q c_{ij}^2} \geq u_\alpha \quad (4)$$

であるとき仮説を棄却する. ただし u_α は標準正規分布の上側 α 点である.

定理 1 $C = (c_{ij})$ が以下の条件を満たすとき $L(X, C)$ は帰無仮説 H に対する (2) の順序を対立仮説 K_1 とした検定に対する optimum contrast である.

$$c_{ij} = \frac{1}{2}(p+q) + 1 - i - j \quad (5)$$

2 最適性について

S を直交補空間上の原点を中心とする単位円の表面とし, \bar{C} を $\bar{C} = C \cap S$ とする. x を (p, q) 配列 X のベクトル形とし a を C の点に対応するベクトルとする. (a, x) が大きすぎるときに母数の同一性の仮説

を棄却する検定手法； ϕ_a が帰無仮説を棄却するのは次の場合である。

$$(\bar{a}, x) \geq u_a \text{ 但し } \bar{a} = a/\|a\| \in \bar{C} = C \cap P \quad (6)$$

この検定は、正規性と等分散性が仮定される場合は、 $\lambda\theta$ ($\lambda > 0$) で与えられる半直線上で一様最強力という意味で適切である。 ψ を線形統計量、例えば (b, x) , $(b \in \bar{C})$, にもとづく検定とし、検定 ϕ_a と ψ の点 a における検出力をそれぞれ $P(a, \phi_a)$, $P(a, \psi)$ で表わす。 $P(a, \phi_a) \geq P(a, \psi)$ であり、 ψ の点 a における損失をその差で表すこととするとそれは正である。すなわち $P(a, \phi_a) - P(a, \psi) \geq 0$ である。また $Q(\psi)$ を全ての $a \in \bar{C}$ において最大の損失であるとする、つまり

$$Q(\psi) = \sup_{a \in \bar{C}} \{P(a, \phi_a) - P(a, \psi)\}. \quad (7)$$

optimum linear test は全ての線形検定の中で $Q(\psi)$ を最小にするものである。Schaafsma, W. and Smid, L. J. (1966) はこの optimum linear test が存在しただ一つであることを緩やかな条件のもとで示している。

3 基本となる定理

$m(= p \times q)$ 本のベクトルを $\theta = (\theta'_1, \dots, \theta'_q)'$ (ただし $\theta_j = (\theta_{1j}, \theta_{2j}, \dots, \theta_{pj})'$ ($j = 1, \dots, q$)) とする。また (m, n) 行列を $A = (a_1, \dots, a_n)$ とし N を 1 から $n(= p \times (q-1) + q \times (p-1))$ までの自然数とする、不等式制約は次のように表わせる

$$A'\theta \geq 0. \quad (8)$$

Shi により N の部分集合 $N_1 \in N$ で次の Proposition を満たすものの存在が示されている。これより $\mathbf{1}_i = (1, \dots, 1)$ は全て 1 からなる i 次ベクトルを表わすものとする。二つのベクトル間の不等号 $>$ と \geq は全ての要素に関して同じ不等号が成立することを意味するものである。

Proposition 1 次の二つの命題が成り立つ。

1. 以下の 3 条件を満たす、大きさ n_1 の部分集合 $N_1 \subset N$ が少なくとも 1 つ以上存在する。
 $N = \{1, 2, \dots, n-1, n\}$ は置換により、 N_1 は N のうちの最初の n_1 個の要素とすることが出来るので、行列の分割が $A = (A_1, A_2)$ であるとして 3 つの条件を記述する。
 (a) $(k \times n_1)$ 行列 A_1 の階数は n_1 であり、それゆえ $A'_1 A_1$ は正則である。
 (b) $(A'_1 A_1)^{-1} \mathbf{1}_{n_1} > \mathbf{0}$
 (c) $(A'_2 A_1)(A'_1 A_1)^{-1} \mathbf{1}_{n_1} \geq \mathbf{1}_{n_2}$
2. 上の条件 (b) と (c) は N_1 が optimum contrast を与える必要十分条件であり、その m 次の係数ベクトルは $A_1(A'_1 A_1)^{-1} \mathbf{1}_{n_1}$ により与えられる。

optimum contrast は尺度母数を除いて唯一であるが、上の条件を満たす A_1 または N_1 は一意ではない。本報告では、 N_1 をの構成することにより、その N_1 が条件 (a),(b),(c) を満たすことを示す。この方法により定理 1 を証明した。さらにその拡張である Double loop order と Umbrella Order に対する Optimum contrast の予想を与えた。

参考文献

- [1] Kudo, A. and Choi, J-R. (1975) A generalized multivariate Analogue of the one Sided test. *Memo. Fac. Sci. Kyushu Univ. A.* 29, pp.303-328.
- [2] Schaafsma, W. and Smid, L. I. (1966) Most stringent somewhere most powerful tests against the alternative restricted by a number of linear inequalities. *Ann. Math. Stat.*, Vol.37., pp. 1161-72.
- [3] Shi, N-Z. and Kudo, A. (1987) The most powerful one-sided test of the multivariate normal mean. *Memo. Fac. Sci. Kyushu Univ. A.* 41, pp.37-44.

Principal Component Analysis for Grouped Bio-diversity Data

Griffith University, Australia & NIAES, Japan

Lukman Thalib

1. Introduction:

Two of the most popular descriptive multivariate methods are principal component analysis (PCA) and canonical variate analysis (CVA). It is widely accepted that whenever the multivariate data are grouped and discriminating the group differences using multiple variables are of the main concern, CVA is the most appropriate technique to use. PCA on the other hand, is used to develop linear combinations that successively maximise the total variance of a sample where there is no known group structure.

However, when there are more variables than within-group degrees of freedom, due to singularity of the within-group covariance matrix, canonical variates (CV) cannot be derived. In such situations, it is shown here, using a real life example from a bio-diversity study and a further computer simulation, that PCA can be a viable substitute, if the between-group differences are prominent.

Krzanowski (1992) has reported that similar problems arise in chemometrics and analysis of sensory data. He has shown in such situations, the best approach is to discriminate the groups using ranked ordered PCA. This is done by ranking the PC using the canonical variate criterion and choosing the components according to this rank order rather than the ones yielded by the ordinary PCA. He has also shown that the first few components based on this ranking are the best orthogonal components that would discriminate the groups and the order obtained would generally be different to the order obtained by ordinary PCA. The ranking of PC can be obtained by conducting separate one-way ANOVAs on the PC scores for each of the components, and using the corresponding F-statistics to rank them. These F-tests correspond to the tests of components of CVA.

In this paper, we attempt to show that there are cases where PCA can be substituted totally for CVA without any modifications. Real life data that we have analysed here, with a further computer simulation study, show that PCA extracts all the information that CVA would produce, provided there are significant between-group differences for all or most of the variables measured. In such situations the rank order of the PC provided by ordinary PCA will also be the same as that based on a canonical variate criterion. We extend the discussion further to describe how the biodiversity index that we have developed using PCA has the same characteristics as the one yielded by the CVA in terms of classifying forest types.

2. Background to the Study

The rapid assessment of the biodiversity of any ecosystem is a key, current issue amongst ecologists and conservation biologists. The main objective of this study was to discriminate the forest quality based on the set of subfamily assemblages of the moths found in the forest using an appropriate multivariate method. The idea was that if an appropriate linear combination of the assemblages of subfamilies, could be developed then they can be used as a rapid biodiversity measure to classify the forest quality of sites that have not been examined in generating the original data set.

The data matrix consists of (three forest types with three replicates) 9 sites x 22 variables (moth counts of each subfamily) and the data was standardised before carrying out a PCA as CVA cannot be used to due to lack of sites.

Further computer simulation was carried out to test the differences among linear combinations derived from PCA with those of CVA.

3. Results

The first PC explained about 34% of the total variability in the data and the second explained about 18%. The F statistics from one-way ANOVA on the PC scores indicated that first two PC were significant

in discriminating the forest types at 5% probability. The first component has an F- value of 125.67 (df = 2,8) while the second component has a F- value of 5.47 (df = 2,8).

The analysis indicated that the most important linear component in terms of discriminating the forest type is also the first principal component followed by the second. The rest of the components were not important. The linear combinations that were found to explain most of the total variability in the data are also the same that discriminate the forest types. Hence, the loadings of PC1 and PC2 can be used to develop the linear indices to discriminate and classify the forest types.

Results from the analyses of the simulated data were the same as the PCA on the original data. The simulated data also extracted only two significant components to discriminate the forest types. The first PC (F= 410, df = 2, 26) was the most important component followed by the second (F= 65.16, df = 2, 26). The variability explained by these two components were about 57% and for the real data it was 52%.

Having demonstrated that the simulated data give rise to the same PCA results as the real data, the simulated data matrix was subjected to CVA. The linear discrimination obtained using the CVA was then compared with that of PCA. The CVA also yielded two significant components with similar linear transformations as those of PCA, in discriminating the forest types.

The Pearson's linear correlation between corresponding PC and the CV thus obtained from the simulated data was very high. The first PC was highly correlated with first CV ($r = -0.99$) and second PC with the second CV ($r = 0.92$). These results suggest that the linear combinations which discriminate among the forest types are the same regardless of the analysis used.

Linear combinations derived from the PCA are the eigenvectors of the total sample variance/covariance matrix, represented by S , while the combinations derived by CVA are the eigenvectors of the $W^{-1}B$ where W is the within-group variance covariance matrix while the B is the between-group variance/covariance matrix. It should be noted here that $S = W + B$. The linear combinations of PCA attempt to maximise the total sample variation while the linear combinations of the CVA attempt to maximise the between-group variability. Our argument is that if the between-group variances dominate the total variability in the sample, for all or most of the variables in the original data, then PCA would essentially discriminate between the groups while reducing the dimensions in the total sample variability. We found in our biodiversity data, the between-group sum of squares formed the substantial proportion of the total sum of squares for most of the variables. If the between-group variances are not similar to the total sample variances one cannot expect to get similar results from both PCA and CVA.

4. Biodiversity Indices

The biodiversity indices that classify the forest types are the linear combinations of the first and the second PCA. We use the PCA on the original data to obtain the PCA loadings for these indices.

5. Conclusions

In this paper we have illustrated the use of PCA as a substitute for CVA, in developing linear combinations to discriminate the groups, whenever CV cannot be derived due to matrix singularity arising from lack of within-group degrees of freedom. We identified that when total sample covariance matrix is not significantly different from $W^{-1}B$ then PCA and CVA would give rise to linear combinations that are the same in terms of their significance, importance and biological meanings. An appropriate test and other comparative techniques that can be used to identify the situations where PCA could be substituted for CVA are also provided.

Reference:

Krzanowski, W.J. (1992) Ranking Principal Components to Reflect Group Structure, *Journal of Chemometrics*, Vol 6, 97-102

1 序

k 個の母集団から得られたデータがあり、それぞれの母集団分布の母数 $\theta_1, \theta_2, \dots, \theta_k$ の間に単順序 $\theta_1 \leq \theta_2 \leq \dots \leq \theta_k$ が仮定できるとするとき、これらの θ の中で真に異なるものを見つけるため、適当な情報量規準を用いたい。AIC を初め、多くの情報量規準では、このような母数への制約を考慮してないので、より有効な情報量規準が望まれる。本報告では、AIC の考えに沿って、順序制約下での最尤推定量を用いた場合の情報量規準を新たに提案した。またこの方法は最初の変化点あるいは閾値を見つける方法として応用できる。

2 順序制約下での情報量規準

確率分布が p -次元の母数 θ で表されているとき、対数尤度を $l(\theta)$ で表すと、赤池情報量規準 AIC は $AIC(\hat{\theta}) = l(\hat{\theta}) - p$ により表される (Akaike, 1973)。ここに $\hat{\theta}$ は制約を置かない通常最尤推定量である。ペナルティ項と呼ばれる p は最尤推定量の場合には妥当であると考えられるが、母数に制約があるような推定量の場合には、適当なものではない。なお、順序制約下での θ の最尤推定量を $\hat{\theta}$ とする。

推定された確率分布 $\hat{g}(x)$ の真の確率分布 $g(x)$ からの乖離度を測る尺度として Kullback-Leibler の情報量 (Kullback & Leibler, 1951)

$$I(g, \hat{g}) = \int g(x) \log g(x) d\nu(x) - \int g(x) \log \hat{g}(x) d\nu(x).$$

を用いる。よく知られているように、上式右辺の 2 つの項はともに未知であるが、2 番目の項 (期待対数尤度と呼ばれる) は対数尤度で推定できる。

今、観測値 x_{ij} ($i = 1, \dots, k; j = 1, \dots, n_i$) がそれぞれ正規分布 $N(\theta_i, \tau_i)$ (τ_i : 既知) から得られたものであるとする。また $\theta = (\theta_1, \dots, \theta_k)$ での対数尤度を $l(\theta)$ とし、正規分布 $N(\theta_i, \tau_i)$ の密度関数を $f(\cdot, \theta_i, \tau_i)$ と表す。ここで、 $\hat{\theta}$ での期待対数尤度を対数尤度 $l(\hat{\theta})$ で推定すると次のようなバイアスを生じる。

$$\begin{aligned} B_1(\theta) &:= E \left\{ l(\hat{\theta}) - \sum_{i=1}^k n_i \int f(t_i, \theta_i, \tau_i) \log f(t_i, \hat{\theta}_i, \tau_i) dt_i \middle| \theta \right\} \\ &= \frac{k}{2} + \frac{1}{2} E \left\{ \sum_{i=1}^k \frac{n_i (\hat{\theta}_i - \theta_i)^2}{\tau_i} - \sum_{i=1}^k \frac{n_i (\hat{\theta}_i - \tilde{\theta}_i)^2}{\tau_i} \middle| \theta \right\}, \end{aligned}$$

このバイアスに関して、 $\theta \in M = \{y = (y_1, \dots, y_k) \mid y_1 \leq \dots \leq y_k\}$ のとき、 $\sum_{i=1}^k P(i, k, w) i \leq B_1(\theta) < k$ であることがわかる。ここで、下限は $\theta \in H = \{y = (y_1, \dots, y_k) \mid y_1 = \dots = y_k\}$ のとき得られる。ただし、 $w_i = n_i / \tau_i$ であり、 $P(i, k, w)$ は $\theta \in H$ のとき $\tilde{\theta}_1, \dots, \tilde{\theta}_k$ の中に異なるものが i 個生じる確率である。

この不等式からわかるように、 $\hat{\theta}$ を用いるとき、AIC のようなペナルティを一意に決めることはできない。そこで、上に示した $B_1(\theta)$ の範囲の中で、単順序のモデル M にもっとも有利なペナルティ項として $C_p = \inf_{\theta \in M} B_1(\theta)$ を用いることを提案する。すなわち、順序制約のある場合に対する情報量規準として次を提案する。

$$ORIC(\tilde{\theta}) := l(\tilde{\theta}) - \inf_{\theta \in M} B_1(\theta) = l(\tilde{\theta}) - \sum_{i=1}^k P(i, k, w) i. \quad (1)$$

なお、特に $w_1 = \dots = w_k$ あるいは $k = 2$ のときは $\sum P(i, k, w) i = \sum 1/i$ である。

x_{ij} が正規分布 $N(\theta_i, \tau_i \sigma^2)$ (τ_i : 既知, σ^2 : 未知) からの観測値である場合も同様に、

$$ORIC(\bar{\theta}, \bar{\sigma}^2) := l(\bar{\theta}, \bar{\sigma}^2) - \sum_{i=1}^k P(i, k, w) i - 1.$$

を得る。ここで、 $\bar{\theta}_i, \bar{\sigma}^2$ はそれぞれ $\theta \in M$ の制約下での θ_i, σ^2 の最尤推定量である。

また、 x_{ij} がある指数型分布族 (Robertson et al., 1988) に属するような分布から得られたとする場合、上と同様な考えに依るならば、漸近的な評価により (1) と同じ形の情報量規準が導かれる。

2つの異なる (階層的な) 確率モデルに対する AIC の場合と同様に、ORIC と AIC の間には次のような関係が成り立つ。

$$\begin{aligned} 2\{ORIC(\bar{\theta}_M) - AIC(\hat{\theta}_H)\} &= 2\{l(\bar{\theta}_M) - l(\hat{\theta}_H)\} - 2E(\bar{\chi}_{01}^2 | \theta \in H), \\ 2\{AIC(\hat{\theta}_K) - ORIC(\bar{\theta}_M)\} &= 2\{l(\hat{\theta}_K) - l(\bar{\theta}_M)\} - 2E(\bar{\chi}_{12}^2 | \theta \in H), \end{aligned}$$

ここで $2\{l(\bar{\theta}_M) - l(\hat{\theta}_H)\} = \bar{\chi}_{01}^2$ は、同等性の帰無仮説を単順序の対立仮説に対して検定する尤度比検定統計量であり、また $2\{l(\hat{\theta}_K) - l(\bar{\theta}_M)\} = \bar{\chi}_{12}^2$ は単順序の帰無仮説をそれ以外とする対立仮説に対する尤度比検定統計量である。

3 変化点の検出

真に異なる母数の組を見つけるために次の $2^{k-1} - 1$ 通りの場合について ORIC を計算する。

$$\theta_1 = \dots = \theta_{k-1} \leq \theta_k; \quad \theta_1 = \dots \leq \theta_{k-1} = \theta_k; \quad \dots; \quad \theta_1 \leq \dots \leq \theta_k,$$

これと $\theta_1 = \dots = \theta_k$ の場合に対する AIC を計算し、これら 2^n 個の中で最大値を与えるものを見つればよいであろう。この方法は同時に変化点を検出していることにもなる。これを用いて、葉の投与量の最大無影響量の検出に適用できる。

また $\theta^* (\in M)$ を真の母数とし、 $\theta^* \in M_\alpha, \theta^* \notin M_\beta$ とする。さらに、 $\bar{\theta}^{(\alpha)}, \bar{\theta}^{(\beta)}$ をそれぞれモデル M_α, M_β の下での最尤推定量とすると、ある定数 $c (> 0)$ に対して、 $\frac{1}{N}\{ORIC(\bar{\theta}^{(\alpha)}) - ORIC(\bar{\theta}^{(\beta)})\} \xrightarrow{P} c (N \rightarrow \infty)$ が成り立つ。

4 シミュレーション

モンテカルロ・シミュレーションにより、単順序が仮定できる場合に真の母数の組を見つけるために、AIC と ORIC との比各評価を行い、ORIC の有効性を確かめた。また、閾値あるいは、最大無影響量のような最初の変化点を見つけるために、Williams の多重比較法、Kikuchi et al. (1991) による AIC に基づく方法と ORIC による方法を比較し、ORIC に基づく方法の有効性を確認した。

REFERENCES

- [1] Akaike, H. : *2nd International Symposium on Information Theory*, 267-81, (1973).
- [2] Kikuchi, Y., Yanagawa, T. and Nishiyama, H.: *Statistical Sciences and Data Analysis*. 345-56.(1993).
- [3] Kullback, S. and Leibler, R.A. : *Ann. Math. Statist.* **22**, 79-86.(1951).
- [4] Robertson, T., Wright, F. T. and Dykstra, R. L. : *Order Restricted Statistical Inference*. (1988).
- [5] Takeuchi, K. : *Mathematical Sciences* (数理科学) No. 153, pp. 12-8 (1976).

Anderson-Stephens 統計量の漸近帰無分布の裾確率

統計数理研究所 栗木 哲
東京大学経済学部 竹村 彰通

p 次元ベクトル $x_1, \dots, x_n \in R^p$ が i.i.d. サンプルとして得られているとする. x_i の方向成分を $z_i = x_i / \|x_i\| \in S^{p-1}$ (R^p の単位球面) とおく. ここで方向の一様性に関する仮説

$$H : z_i \sim \text{Uniform}(S^{p-1})$$

の検定を考える. この問題に関し Anderson and Stephens (1972) は

$$S(u) = (1/n) \sum_{i=1}^n (u' z_i)^2 \quad (u \in S^{p-1})$$

の u に関する最大値, 最小値を

$$S_{\max} = \max_{u \in S^{p-1}} S(u), \quad S_{\min} = \min_{u \in S^{p-1}} S(u)$$

とおき, S_{\max} が大きいとき, あるいは S_{\min} が小さいときに仮説を棄却する検定を提案した. この他にも自然な検定統計量として $S_{\max} - S_{\min}$ が考えられる. 本発表では S_{\max} , $-S_{\min}$ および $S_{\max} - S_{\min}$ の $n \rightarrow \infty$ での極限帰無分布の上側裾確率について報告する.

最初に次の事実に注意する.

補題 1 $A = (a_{ij})$ を対称正規分布に従う $p \times p$ 対称行列とする. すなわち A の対角成分 a_{ii} は $N(0, 1)$, 非対角成分 $a_{ij} = a_{ji}$ は $N(0, 1/2)$ に全て独立に従うとする. また

$$B = \sqrt{\frac{p}{p-1}} \left(A - \frac{\text{tr}(A)}{p} I_p \right)$$

とおく. このとき $n \rightarrow \infty$ のとき分布収束

$$\begin{aligned} \sqrt{n}(S_{\max} - 1/p), -\sqrt{n}(S_{\min} - 1/p) &\xrightarrow{\mathcal{D}} \sqrt{\frac{2(p-1)}{p^2(p+2)}} \lambda_1(B), \\ \sqrt{n}(S_{\max} - S_{\min}) &\xrightarrow{\mathcal{D}} \sqrt{\frac{2}{p(p+2)}} (\lambda_1(A) - \lambda_p(A)) \end{aligned}$$

が成り立つ. ただし $\lambda_1(\cdot)$, $\lambda_p(\cdot)$ で最大, 最小固有値を表すこととする.

$p \times p$ 対称行列全体のなす計量ベクトル空間を $\text{Sym}(p)$ とおく. ここで $\text{Sym}(p)$ の内積は $\text{tr}(XY)$, $X, Y \in \text{Sym}(p)$ とする. $\text{Sym}(p)$ の単位球面

$$\{X \in \text{Sym}(p) \mid \text{tr}(X^2) = 1\}$$

の2つの部分多様体を考える:

$$M_1 = \left\{ \sqrt{\frac{p}{p-1}}(uu' - (1/p)I_p) \mid u \in S^{p-1} \right\},$$

$$M_2 = \left\{ \frac{1}{\sqrt{2}}(uu' - vv') \mid u, v \in S^{p-1}, u'v = 0 \right\}.$$

このとき

$$\lambda_1(B) = \max_{U \in M_1} \text{tr}(UA),$$

$$\frac{1}{\sqrt{2}}(\lambda_1(A) - \lambda_p(A)) = \max_{U \in M_2} \text{tr}(UA)$$

と表すことができる. ここで A は $p \times p$ 対称正規分布行列である. このことから $\lambda_1(B)$ および $\frac{1}{\sqrt{2}}(\lambda_1(A) - \lambda_p(A))$ の上側裾確率は漸近的に χ^2 分布の上側確率の線形和で表せること, また線形和の係数は M_1, M_2 の各点における主曲率の基本対称式を体積要素により積分することにより求められることが分かる (Kuriki and Takemura (1998)).

以下の定理において, $\bar{G}_\nu(\cdot)$ で自由度 ν の χ^2 分布の上側確率を表すものとする.

定理 1

$$P(\lambda_1(B) \geq x) = \sum_{\substack{e=0 \\ e:\text{even}}}^{p-1} w_{p-e} \bar{G}_{p-e}(x^2) + O\left(\bar{G}_p\left(\frac{2p-2}{p-2}x^2\right)\right)$$

$$w_{p-e} = \frac{1}{2} \left(\frac{2p}{p-1}\right)^{\frac{1}{2}(p-1)} \left(-\frac{p+1}{2p}\right)^{e/2} \frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p-e+1}{2}\right)\left(\frac{e}{2}\right)!}$$

定理 2

$$P\left(\frac{1}{\sqrt{2}}(\lambda_1(A) - \lambda_p(A)) \geq x\right) = \sum_{\substack{e=0 \\ e:\text{even}}}^{2p-3} w_{2p-2-e} \bar{G}_{2p-2-e}(x^2) + O\left(\bar{G}_{2p-2}\left(\frac{4}{3}x^2\right)\right)$$

$$w_{2p-2-e} = 2^{p-2} \left(-\frac{1}{2}\right)^{e/2} \binom{p}{e/2}$$

定理2における係数 w の導出の過程で, 次の関係式も示される.

補題 2

$$E[\det(A)^2] = \begin{cases} 2^{-p} \frac{2}{3} (p+2)!! p!! & p: \text{ odd} \\ 2^{-p} \frac{2p+3}{3} (p+1)!! (p-1)!! & p: \text{ even} \end{cases}$$

参考文献

Anderson, T. W. and Stephens, M. A. (1972). Tests for randomness of directions against equatorial and bimodal alternatives. *Biometrika*, **59**, 613–621.

Kuriki, S. and Takemura, A. (1998). Tail probabilities of the maxima of multilinear forms and their applications. Discussion Paper CIRJE-F-4, Faculty of Economics, Univ. of Tokyo.

Simulated Annealing による時系列データの Wavelet 近似

京都産業大学・理学部 森 隆一
 名古屋市立大学・経済学部 三澤 哲也

等間隔に並んだ有限個の時系列データを解析する方法として、山本らは ([5]) Wavelet 関数系を用いた滑らかな近似を行い、得られた近似関数の導関数を調べることにより、データの変化の特徴をとらえる方法を提唱した。Wavelet 関数系を用いる利点として、

- (i) 近似曲線の導関数を求めることができるので、それを調べることによりデータの変化の特徴をとらえ易くなる。
- (ii) 上の関数系は局所的な関数をもとにスケールの変換と平行移動により得られる関数系であるため、スケールが異なる同じような要因の積み重ねにより変化しているとみなせるデータの解析に有効に働くことが期待できる。

の2点が挙げられる。しかしながら、山本らの方法では、上記の近似関数の係数を求めるさいに連立一次方程式を用いるため、Wavelet 関数の個数と時系列データの個数を等しくとることが本質的に必要となり、それが近似精度の向上を考える上で本質的に大きな障害となっている。

このことを考慮して、我々は時系列データの折れ線グラフを近似するという立場から、時系列データと近似関数との近似誤差の他に時系列データの1階差分と導関数との近似誤差を考え、両者の和として与えられる近似評価関数を導入し、それを最小にする最適近似関数を Simulated Annealing により求める方法 Wavelet Interpolation Method with Simulated Annealing (WISAM) を得た。

近似関数系として [5] で用いられた Meyer Wavelet $\{\psi_{j,k}\}$ を用いる。この関数系は、

$$\psi(t) = \frac{1}{\pi} \int_{\frac{2\pi}{3}}^{\frac{4\pi}{3}} \cos(tu \sqrt{1-A(u)}) du + \frac{1}{\pi} \int_{\frac{4\pi}{3}}^{\frac{8\pi}{3}} \cos(tu \sqrt{A(2u)}) du$$

$$A(u) = 3(-3u/(4\pi) + 2)^2 - 2(-3u/(4\pi) + 2)^3$$

で与えられる ψ を用いて、 $\psi_{j,k}(t) = C\psi(2^j(t-k))$ として構成される。ここで、 j は -5 から -1 までとし、 k は 0 よりデータ数 -1 とする。各時刻においてデータの変化への寄与が $\psi_{-1,k}, \dots, \psi_{-5,k}$ となる要因があり、これらを $c_{j,k}$ 倍したものの和としてデータが得られていると見ることになる。(C および以下の C_0 は定数とする。)

時間間隔を 1 とする T 個の時系列データ $\{f(t)\}_{t=0,1,\dots,T-1}$ を考える。これに対する近似関数 $\sum_{k=0}^{T-1} \sum_{j=-5}^{-1} c_{j,k} \psi_{j,k}(t)$ の評価関数 E として

$$E(\Phi) = C_1 \sum_{t=0}^{T-1} (\Phi(t) - f(t))^2 + C_2 \sum_{t=0}^{T-2} \sum_{\ell=0}^{d-1} \left(\Phi' \left(t + \frac{\ell}{d} \right) - \{f(t+1) - f(t)\} \right)^2 \quad (1)$$

を用いる。我々の方法はこの評価関数を最小にする係数 $\{c_{j,k}\}$ を求めることである。このとき、時系列データの個数は T 個であり、決定すべき係数の個数は $5T$ 個であることと、 $j < 0$ の場合 $\{\psi_{j,k}; k = 0, \pm 1, \pm 2, \dots\}$ は直交性を持たないことにより、通常の解析的手法により係数を求めることはきわめて困難と考えられる。

これに替わって、我々は係数を有限区間の整数に制限することにより、有限集合上の (正値) 関数の最小値を求める確率論的手法である Simuated Annealing を用いて係数

を求める方法を採用する。この計算は、以下のような、係数 $\{c_{j,k}\}$ のとる値（それを $S = [-N, \dots, N]^{5T}$ とする）を状態とする $5T$ 次元の simple random walk を修正した時間的に一様でない Markov 連鎖 $\{X_n\}$ の simulation を行うことになる。

- (i) 初期値として、全ての係数を 0 とする。
- (ii) 各時刻 n で、 l, m と $h = \pm 1$ を等確率で選び、 $c_{l,m}$ を $c_{l,m} + h$ としたことによる近似誤差の増分 ΔE を求める。
- (iii) $\Delta E > 0$ の場合、確率 $1 - e^{-\beta(n)\Delta E}$ で係数を変化させない。これ以外は、 $c_{l,m}$ を $c_{l,m} + h$ に変更する。ただし $\beta(n) = C_b \log n$ とする。また、成分の絶対値が N を越えるときは変化しない。

Gemann & Gemann, Hajek, Catoni らの定理によると ([1]), C_b が小さいとき、十分な反復計算を行えば、(1) 式の評価関数の最小値を実現する係数が得ることが保証される。実際の simulation では一定の計算ステップ数で計算を打ち切り、そこまでに得られた最小値を与える係数を求める。

WISAM による近似の特徴をつかむため、データの個数を 8 とする幾つかの仮想テスト・データに対する近似を行った。この結果は、定数を含む線形的データに対しては近似評価関数値は小さい値を得、折れ方が大きくなるほど大きな値を得た。これより、近似評価関数値は直線的变化との差を反映している、或いは、直線的な変化を妨げる力の大きさを反映していることが予想される。

WISAM の応用として、(I) 運動選手の身長時系列データ (森・卯野・大西・辻井 [3]) と (II) 株価時系列データ (森・三澤 [4]) の分析が試みられている。

前者では、小学校より高等学校までの体格検査で得られる 12 年間のデータに対して得られた近似関数の導関数のグラフを成長速度曲線とし、この特徴により運動選手の分類を行い、運動選手としての成績等との関連を調べることにより選手指導の参考にすることを目標としている。これは、Wavelet 補間の利点 (i) を主としたものである。後者では利点 (ii) と上記予測に基づいて、次のようなデータ分析の方法が提案され、任天堂 1990 年第 9 週よりの株価週足データに対する試みが行われている。

各時刻 t において次のものを求める。

($t-7$ より t までの 8 個のデータに対する近似評価関数値)

– ($t-7$ より $t-1$ までの 7 個のデータに対する近似評価関数値)

ここで得られた数値は、時刻 $t-1$ より t の変化と線形的な変化との差を妨げる力の大きさを表していると考えられ、この値の大きいことは異常な変化に起きていると予測される。

[1] Aarts E. and Korst J. , Simulated Annealing and Boltzmann Machines, John Wiley & Sons, 1989.

[2] 森 隆一：シミュレーテッド・アニーリングを用いた時系列データのウェーブレット近似, 準備中.

[3] 森 隆一, 卯野 優, 大西 健, 辻井 芳樹：WISAM による陸上競技部投擲選手の発育速度曲線の分析, 準備中.

[4] 森 隆一, 三澤 哲也：Simulated Annealing による株価時系列データの Wavelet 近似, 日本統計学会, 1998.

[5] 中村 正治, 長谷川 泰洋, 山本 浩：ウェーブレット補間, 第 21 回 OR 学会中部支部研究発表会, 1994.

ニューロ判別分析におけるモデル診断について Model Diagnostics on Neural Discriminant Analysis

大阪電気通信大学 大学院工学研究科 越水 孝

大阪電気通信大学 情報工学部 辻谷 将明

1. はじめに

本報告では、階層型ニューラルネットワークを判別問題(Discriminant Problem)に適用し、出力値の確率的解釈を行うことによってネットワーク尤度を構成して尤度原理に基づく統計的推測を行う。そして、ニューラルネットワークモデルの適合度評価、情報量基準に基づくモデル選択とその診断法、およびブートストラップ法に基づく誤判別率の推定を試みる。

2. モデル構築と統計的推測

階層型ニューラルネットにおいて判別問題を取り扱う場合、ある入力パターンを提示したときの出力値は、出力層のあるユニットに入るベイズの事後確率と考えられる。出力値 $o^{<d>} = o^{<d>}(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\beta})$ は、教師値 $t^{<d>}$ が 1 である条件付き確率 $Pr\{t^{<d>} = 1 | \mathbf{x}^{<d>}\} \cong o^{<d>}$ となり、対数尤度は、

$$\ln L(\boldsymbol{\alpha}, \boldsymbol{\beta}; \mathbf{x}, \mathbf{t}) = \sum_{d=1}^D \left\{ t^{<d>} \ln o^{<d>} + (1 - t^{<d>}) \ln(1 - o^{<d>}) \right\} \quad (1)$$

で与えられる。(1)式を最大にする $\hat{o}^{<d>}$ (すなわち、 $\hat{\alpha}_j, \hat{\beta}_k$) が最尤推定量(MLE)となる。

データへのモデル適合度を総括的に評価し得る指標として、幾つかの統計量が提案されている。ニューラルネットワークの場合、逸脱度(deviance)は、

$$\begin{aligned} Dev &= 2[\ln L(\max) - \ln L(\boldsymbol{\alpha}, \boldsymbol{\beta}; \mathbf{x}, \mathbf{t})] \\ &= 2 \left[\sum_{d=1}^D \left\{ t^{<d>} \ln \left(\frac{t^{<d>}}{o^{<d>}} \right) + (1 - t^{<d>}) \ln \left(\frac{1 - t^{<d>}}{1 - o^{<d>}} \right) \right\} \right] \sim \chi^2_{D-p} \end{aligned} \quad (2)$$

と定義でき、また、Pearson 適合度統計量として

$$\chi^2 = \sum_{d=1}^D \left\{ \frac{(t^{<d>} - o^{<d>})^2}{o^{<d>}(1 - o^{<d>})} \right\} \sim \chi^2_{D-p} \quad (3)$$

を用いることもできる。

3. ブートストラップ法によるバイアス推定

3.1 情報量規準

競合するモデルが複数個ある場合には、AIC

$$AIC = -2 \ln L(\boldsymbol{\alpha}, \boldsymbol{\beta}; \mathbf{x}, \mathbf{t}) + 2p \quad (4)$$

を比較することによって最適なモデルを選択することができる。AIC は、モデルの平均対数尤度を対数尤度 $\ln L(\boldsymbol{\alpha}, \boldsymbol{\beta}; \mathbf{x}, \mathbf{t})$ で推定した時のバイアスを未知パラメータ数で近似補正している。しかし、AIC の導出については、比較的少数のパラメータを持つモデルを想定しているのに対し、ニューラルネットワークでは、各層のユニット数が増えるにつれて未知パラメータ(リンク荷重およびしきい値)数は飛躍的に増えていく。さらに、ニューラルネットワークではより少ない隠れユニット数で出力値を実

現できる場合にリンク荷重の非一意性も指摘されている。

EIC は、対数尤度のバイアスをブートストラップ法を用いて直接推定している。モデルの密度関数を $f(\mathbf{x}|\boldsymbol{\theta})$, $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$ とし、ブートストラップ法を用いると

$$C^* = E_x \{ \ln f(\mathbf{x}^* | \hat{\boldsymbol{\theta}}(\mathbf{x}^*)) - \ln f(\mathbf{x} | \hat{\boldsymbol{\theta}}(\mathbf{x}^*)) \} \quad (5)$$

によってバイアスの近似値が得られる。ここで、 \mathbf{x} は初期標本、 \mathbf{x}^* はブートストラップ標本 ($b=1, \dots, B$) で、 $\hat{\boldsymbol{\theta}}(\mathbf{x})$, $\hat{\boldsymbol{\theta}}(\mathbf{x}^*)$ はそれぞれ初期標本、ブートストラップ標本に基づいた $\boldsymbol{\theta}$ の推定値とする。このとき、

$$EIC = -2 \ln f(\mathbf{x} | \hat{\boldsymbol{\theta}}(\mathbf{x})) + 2C^* \quad (6)$$

によってブートストラップ法に基づく情報量規準が得られる。 EIC が最小のモデルを最適モデルとして選択することができる。

3.2 誤判別率の推定

判別分析では、初期標本に基づいて何らかの判別ルールを構築するが、この初期標本に対する誤判別率は、見かけ上の誤判別率(Apparent error rate)と呼ばれる。しかし、実際の誤判別率(Actual error rate)とは、初期標本に基づいて判別ルールが得られたという条件のもとで、将来観測されるデータに対して予測を誤ってしまう確率である。これを見かけ上の誤判別率で置き換えて推定することは、実際の誤判別率を過小に推定する傾向がある。また、ニューラルネットワークモデルにおいても、過大なネットワークを用いると汎化能力が低下することも指摘されている。

本稿では、ブートストラップ法を用いて実際の誤判別率と見かけ上の推定量との差、すなわちバイアス

$$\hat{b}(\hat{F}_i)_{bs} = E \{ \hat{e}_i(\hat{F}_i^*; \mathbf{X}_n^*) - \hat{e}_i(F_i; \mathbf{X}_n^*) \} \quad (7)$$

を推定する。ここに、ブートストラップ標本 \mathbf{X}_n^* に基づいて、構築された判別ルールでの見かけ上の誤判別率を $\hat{e}_i(\hat{F}_i^*; \mathbf{X}_n^*)$ 、初期標本に対する誤判別率を $\hat{e}_i(F_i; \mathbf{X}_n^*)$ とする。そして、バイアス補正を施した

$$\hat{e}_i(\hat{F}_i; \mathbf{X}_n) - \hat{b}(\hat{F}_i)_{bs} \quad (8)$$

を実際の誤判別率の推定量として用い、ネットワークの性能を評価する。

4. 適用例

二群判別問題の実際例として、前立腺癌のリンパ節転移データを取上げる。このデータは、すでに外科的処置によって癌のリンパ節転移の有無が確認されている前立腺癌患者 53 例に対して、5つの術前の所見が観察されている。これら 53 例の前立腺癌患者のデータを解析することにより、前立腺癌と診断された患者のリンパ節転移を5つの予後因子に基づいて正確に予測する。

MCMC 法によるランダムネットワークのベイズ的推論

統計数理研究所

Aarhus Univ.

Aalborg Univ.

種村 正美

E.B.Vedel Jensen

J.Møller

研究報告要旨

空間統計学において、点配置データが与えられたとき、データを生成したプロセス、そこに含まれるパラメータ等を推論することは重要な課題である。この問題は所与の点配置データが規則型と見なされる場合には有力な手法がすでに開発されている。

しかし、点配置データが集中型に属する場合は、どのデータにも同程度に有効な一般的手法を開発することは非常に困難であろうと考えられる。

われわれは、本講演において、集中型に属する点配置データの中でネットワーク構造を持つようなデータの解析法を考察した。

Karlsson & Liljeborg (1994) は半透明アルミナにおいて空孔（ポア：格子欠陥の一種）の3次元配置を観測し、アルミナの内部で空孔が定常ポアソン点過程に比べて強い集中型の傾向があることを K 統計量などの2次モーメント量を用いて示した。このデータの場合、空孔が結晶粒界の近傍に場所を占めることは非常に確からしいといえる。事実、空孔の3次元配置をステレオ図で眺めると、そのことが確認された。

上述のアルミナの空孔に関する3次元配置の観測によって、結晶粒界のようなランダムなネットワーク構造を発生させる基準とする空間点過程の統計学という新しい問題群が生じる。そこで、一般に、所与の点配置データがランダムなネットワーク構造から派生した場合、元のネットワーク構造をデータからベイズ的に推論する問題を考える。このような問題は現実に、上述の材料科学における例の他に、医学、地球科学などで現れる重要な統計的問題であり、画像解析や不均質点過程における事象の空間構造の探索など幅広い問題と密接な関連がある。

講演では、この問題に関してベイズ的アプローチの枠組みを示した。われわれはまず、観測データの背景となるランダムネットワークを Voronoi 分割ネットワーク（以下、Voronoi ネットワーク）で表す。これは、比較的少数個のパラメータでネットワーク構造を表現できるモデルであり、必要なら重み付き Voronoi 分割などを考慮することによって柔軟性に富んだモデル化が可能であるためである。

今、観測データを x とし、Voronoi ネットワークの母点を y で表すことにする。そして、 x の各々は Voronoi ネットワークのある辺上の点から確率過程によって発生したものとし、具体的には観測点 x と辺上の点 z との誤差が分散 σ^2 の正規分布に従うと仮定する。以上の設定によって、 y の事前分布 $p(y)$ および z の条件つき事前分布

$p(z|y)$ 、そして x の尤度関数 $p(x|z, y)$ を用いると、事後分布 $p(y, z|x, \sigma)$ が構成される。 y, z の事前分布として Poisson 分布を仮定すると、 z に関する積分が実行できて周辺事後分布 $p(y|x, \sigma)$ が求められる。

以上のベイズ型枠組みに基づいて、MCMC (Markov Chain Monte Carlo) 法によって、観測データ x に対応して (y, σ) の系列を計算機実験で実現し、Voronoi ネットワークの構造と正規分布のパラメータ σ を推定することが可能になる。

われわれは、2次元の人工データに対して、このベイズ型アプローチが実際に作動することを確認した。まず、直径が一定の円によるランダム逐次充填 (充填密度 0.5) によって得られた中心点を母点とする 50 個の点の配置 y を一辺の長さ 10 の正方形領域の中に生成し、その Voronoi 分割を構成して Voronoi ネットワークを準備した。次に、このネットワークにおいて辺上にサンプルした点 z を中心として、一定の分散の正規分布に従って得られた点の配置 ($N = 1000$) を観測データとする。

講演では、人工データに対する観測データにもとづいて、正規分布の分散および背景のネットワークのベイズ推定の一例を与え、真の分散の値が一定の範囲内にある場合には、推定値が良好であること、そして推定されたネットワーク構造が元の構造をよく再現することを示した。今後は、実データの解析を 2次元、3次元の場合を含めて行うことが課題である。

文 献

- [1] Karlsson, L.M. and Liljeborg, A. (1994). Second-order stereology for pores in translucent alumina studied by confocal scanning laser microscopy. *Journal of Microscopy*, **175**, pp.186-194.
- [2] Tanemura, M., Vedel Jensen, E.B. and Møller, J. (1998). Random tessellation network and some inference problems (Abstracts of the 9-th International Workshop on Stereology, Stochastic Geometry and Image Analysis). *Advances in Applied Probability*, **30**, No.2, 291 - 292.

Statistical analysis of uncertain paleoearthquake occurrences times to forecast the hazard of rupture on a fault segment

The Institute of Statistical Mathematics, Yoshihiko Ogata

Introduction

This paper eventually aims at the provision of methods for practical risk assessment of a large earthquakes on each segment of active faults. Shallow earthquakes on the land causes serious disasters in a range of region near the ruptured fault. The growth of major cities in hazard prone areas, and the public anxiety associated with risks to critical facilities such as nuclear reactors, has focused attention on the problems of insurance against such hazard, disaster mitigation, and disaster prevention. These are assuming ever greater economic importance.

Usually, recurrence times of fault ruptures in the land is very long, order of 1000 years or more. However, active faults are densely distributed in Japan. Therefore it is important to reveal as many paleoearthquake occurrence data as possible in order to make assessment of the risk of a forthcoming earthquake associated with each fault segment. In order to provide such occurrence data, trenching study has been carried out since 1980s, and even more intensively after the 1995 Hyogo-Ken Nanbu Earthquake (the Kobe Earthquake) in Japan. The past history of earthquake on a specific fault is inferred from the stratigraphic record. The occurrence times are estimated by deciphering preserved histories of displacement on the fault and also estimated from ages of strata using radiocarbon dating and related geological informations besides topography. Therefore, in many cases, an occurrence time is vaguely estimated such as in an interval.

Conclusions

The renewal processes with four different distributions and a stationary Poisson process have been applied to paleoearthquake occurrence time data which are deciphered from a number of trenching sites in central Japan. The conclusions and consequences of the proposed methods are listed as follows:

- For the risk assessment it is useful to show the change of the conditional intensity, or hazard rate, relative to the constant rate (the average) of the stationary Poisson process.
- It is more informative, especially when the sample size is small, to consider the exact likelihood using the data including the two interval ends before the first event and after the last event in the observed period.
- An objective Bayesian method is proposed to estimate the parameters of the renewal models in case where the occurrence times of the events are not given explicitly, for example, where they are given in estimated time intervals owing to the limited geological information .
- The normalizing transformation of parameters are useful not only to stabilize the maximization of the log likelihood functions but also to set the standard parameter values for each renewal models. This enables to find a simpler but

most useful predictive model by restricting one or two parameters to the standard values for each renewal process, especially in case where the sample size is very small and the uncertain time spans are wide.

- Goodness-of-fit of the models to uncertain occurrence data are compared from the predictive viewpoint by using AIC and ABIC, not only among the above cases of nested models but also among the renewal processes of different distributions, as well as among the Bayesian models of different priors including the Dirac's delta distribution for the centered data (namely, comparison between AIC and ABIC values).
- The uncertainty and estimation errors of the occurrence times owing to the limited information can be reduced by looking at estimated marginal posterior distribution of an occurrence time based on the assumption that the events on a fault segment follow renewal process.
- An assessment procedure of the current and future hazard of the forthcoming event is proposed in case where the occurrence time of the last event is uncertain.

The conclusions and consequences of the present analysis of data sets are listed as follows:

- All the considered renewal process models fit significantly better than the stationary Poisson process in spite of small number of events and in spite of wide uncertainty of occurrence times. This and estimated hazard functions indicate that all data sets of the paleoearthquake occurrences deciphered from trenching of a number of active fault segments exhibit the strong inhibitory feature of occurrences for substantially long time span after the last event.
- The maximum (mode) of each marginal posterior function tends to locate equidistantly to the neighboring modes, and the mode itself is not always located in the center of the uncertain occurrence interval.
- The current hazard rates at the first three fault segments (the West Nagano Basin fault, Atera fault and Neodani fault) are very small relative to the corresponding stationary Poisson hazard rates (reciprocal of the mean recurrence times). Also, the current hazard rate at the Kamishiro fault does not reach the level of the stationary Poisson hazard rate. In contrast, the current hazard rate at the Gofukuji fault is about 3 to 5 times as much as the Poisson hazard rate (0.1% per year), and increasing rapidly thereafter. Namely, for example, probability of the forthcoming rupture within 30 years on Gofukuji fault is about 10%.

The computer programs used in this papers are now being prepared for release.

Reference

- Ogata, Y., Estimating the hazard of rupture using uncertain occurrence times of paleoearthquakes, *Research Memorandum*, No. 702, The Institute of Statistical Mathematics, Tokyo, 1998.

STUDENTIZED ROBUST STATISTICS FOR MAIN EFFECTS IN A TWO-FACTOR MANOVA

TAKA-AKI SHIRAISHI

Department of Mathematical Sciences

Yokohama City University

22-2 Seto Kanazawa-Ku Yokohama 236, Japan

Abstract

We construct classes of robust tests and estimators in a two-factor manova with no interaction. For the two-factor factorial design with A (rows) and B (columns), the k -th response $\mathbf{X}_{ijk} = (X_{ijk}^{(1)}, \dots, X_{ijk}^{(p)})'$ is expressed as

$$\mathbf{X}_{ijk} = \boldsymbol{\mu} + \boldsymbol{\beta}_i + \boldsymbol{\tau}_j + \mathbf{e}_{ijk}, \quad (k = 1, \dots, n_{ij}, i = 1, \dots, I, j = 1, \dots, J) \quad (0.1)$$

where $\sum_{i=1}^I \boldsymbol{\beta}_i = \sum_{j=1}^J \boldsymbol{\tau}_j = \mathbf{0}$. In (0.1), \mathbf{e}_{ijk} is the error term with $E(\mathbf{e}_{ijk}) = \mathbf{0}$. It is assumed that \mathbf{e}_{ijk} 's are independent and identically distributed with continuous distribution function $F(x^{(1)}/\sigma^{(1)}, \dots, x^{(p)}/\sigma^{(p)})$, where $V(e_{ijk}^{(\ell)}) = \{\sigma^{(\ell)}\}^2$ and $e_{ijk}^{(\ell)}$ is the ℓ -th element of \mathbf{e}_{ijk} . We consider statistical inference relative to $\boldsymbol{\beta}_i$'s and $\boldsymbol{\tau}_j$'s, but we will concentrate our effects on the statistical inference relative to $\boldsymbol{\tau}_j$'s as we recognize that the statistical inference relative to $\boldsymbol{\beta}_i$'s can be obtained simply by reversing the rows and columns.

For the respective parameters, the null hypothesis of interest and the alternative are respectively

$$H; \boldsymbol{\tau}_j = \mathbf{0} \text{ for } j = 1, \dots, J \text{ v.s. } A; \boldsymbol{\tau}_j \neq \mathbf{0} \text{ for some } j.$$

We propose test procedures based on studentized robust statistics for these hypotheses in the model (0.1). We derive their asymptotic properties as the number of blocks n tends to infinity. We also propose robust estimators studentized by scale-estimators for the treatment effects and derive their asymptotic normality.

Multivariate M -tests and M -estimators based on studentized robust statistics were discussed by Singer and Sen (1985) for full rank linear models. Further discussions of studentized M -procedures were done by Koenker and Portnoy (1990) and Jurečková and Sen (1995). The linear models do not include our model (0.1), which is not of full rank. In all of those theoretical discussions, Fisher's consistency: $\int_{-\infty}^{\infty} \psi(x) dF(x) = 0$ was needed. Shiraishi (1990) discussed multivariate M -tests without assuming Fisher's consistency. However Shiraishi's test statistics were not studentized. For many applied models, the scale-parameter of the underlying distribution is unknown and Fisher's consistency does not hold. We need to construct flexible statistical procedures. For the model (0.1), we propose studentized robust tests and computable robust estimators. The asymptotic noncentral

χ^2 -distributions for the test statistics and asymptotic normality for robust estimators are derived, assuming only the finiteness of Fisher's informations. Fisher's consistency : $\int_{-\infty}^{\infty} \psi_{\ell}(x^{(\ell)}) dF_{\ell}(x^{(\ell)}) = 0$ is not assumed. In our discussion, it is verified that the *ARE* of the proposed tests (proposed estimators) relative to the *F*-test (least squares estimator (*LSE*)) goes to the *ARE* of the Huber's (1964) *M*-estimators relative to the one-sample sample mean as the block size goes to infinity. By simulation study, even for the small sample sizes, it can be seen that the proposed estimators are more efficient than least squares estimators except for the case where the underlying distribution is normal. Especially the proposed estimators are remarkably efficient for the asymmetric underlying distributions.

In Section 2, studentized robust procedures are proposed. In Section 3, asymptotic linearity is studied for robust statistics. In Section 4, the asymptotic χ^2 -distributions of the test statistics for respective parameters are derived. Furthermore, asymptotic multivariate normality for the robust estimators is studied. In Section 5, the asymptotic efficiencies of the proposed procedures relative to parametric procedures are studied. In Section 6, simulation study for small sample sizes is done.

References

- Dunn, O. J. and Clark, V. A. (1987) *Applied Statistics: Analysis of Variance and Regression*, Second Edition, New York: John Wiley.
- Good, P. (1993) *Permutation tests*, Springer.
- Hieritier, S. and Ronchetti, E. (1994) Robust bounded-influence tests in general parametric models. *Journal of the American Statistical Association*, **89**, 897-904.
- Huber, P. J. (1964) Robust estimation of a location parameter. *Annals of Mathematical Statistics*, **35**, 73-101.
- Jurečková, J. and P. K. Sen (1996) *Robust Statistical Procedures: Asymptotics and Interrelations*, New York: John Wiley.
- Koenker, R. and Portnoy, S. (1990) M-estimation of multivariate regressions. *Journal of the American Statistical Association*, **85**, 1060-1068.
- Muirhead, R. J. (1982) *Aspects of Multivariate Statistical Theory*, New York: John Wiley.
- Puri, M. L. and Sen, P. K. (1985) *Nonparametric Methods in General Linear Models*, New York: John Wiley.
- Shiraishi, T. (1989) R-estimators and confidence regions for treatment effects in multi-response experiments. *Journal of Multivariate Analysis*, **31**, 30-39.
- Shiraishi, T. (1990) M-tests in multivariate models. *Metrika*, **37**, 189-197.
- Shiraishi, T. (1996) On scale-invariant M-statistics in multivariate k samples. *Journal of the Japan Statistical Society*, **26**, 241-253.
- Singer, J. M. and Sen, P. K. (1985) M-methods in multivariate linear models. *Journal of Multivariate Analysis*, **17**, 168-184.

等質性分析とペナルティ制約付き対応分析

北海道大学・文 大津起夫

1 はじめに

対応分析および多重対応分析はともに計算が簡単であり広い範囲で応用されているが、いくつかの欠点を持っている。そのひとつは、カテゴリーを無構造なものとして取り扱っているために、外的に利用可能な情報を分析に反映できず、また少頻度のカテゴリーに対応するスコアがしばしば不安定になることである。

これらを解消するためのひとつの方法は、カテゴリースコアの推定に、事前の多様な情報を表現するための制約を導入することである。名義尺度のカテゴリカルデータを対象とする非線形主成分分析(等質性分析, homogeneity analysis)は、アイテムとカテゴリーが構成する階層的な構造を利用することにより多重対応分析の欠点を回避しうる場合がある(Bekker & de Leeuw, 1988; 大津, 1993)。また、高根ら(Takane & Shibayama, 1991; 高根, 1995)は、主成分分析と対応分析の解に線形の制約を導入することにより、事前情報を利用した分析を行なうことを試みている。本報告では、解スコアに事前の仮定を緩やかな制約として課すことにより、解釈が容易でありかつ安定した結果の得られることを示す。

2 対応分析への制約の導入

2重分類データを対象とし対応分析の拡張を考える。2重分類表をあらわす頻度データ行列を $\mathbf{D}_{M \times N} = (d_{ij})$ とする。また、 $\mathbf{D}\mathbf{1} = \mathbf{f}$, $\mathbf{D}'\mathbf{1} = \mathbf{g}$, $\mathbf{F} = \text{Diag}\mathbf{f}$, $\mathbf{G} = \text{Diag}\mathbf{g}$ とする。ここで、 $\mathbf{1}$ は要素が全て1である縦ベクトルをあらわす。ベクトル \mathbf{f} の第 i 要素は、 \mathbf{D} の第 i 行の頻度の和をあらわし、また、ベクトル \mathbf{g} の第 j 要素は、第 j 列の頻度の和をあらわす。記号 $'$ は転置行列をあらわし、 Diag は引数として指定されたベクトルを対角成分とする対角行列をあらわす。さらに、 $\tilde{\mathbf{D}} = \mathbf{F}^{-1/2}\mathbf{D}\mathbf{G}^{-1/2}$ と置く。

対応分析の解は、特異値分解 $\tilde{\mathbf{D}} = \mathbf{U}_{all}\mathbf{A}\mathbf{V}'_{all}$ の解を得て、さらに $\mathbf{P}_{all} = \mathbf{F}^{-1/2}\mathbf{U}_{all}$, $\mathbf{Q}_{all} = \mathbf{G}^{-1/2}\mathbf{V}_{all}$ を計算することによって得られる。 \mathbf{P}_{all} は $M \times \min(M, N)$ の行列であり行スコアをあらわす。また、 \mathbf{Q}_{all} は $N \times \min(M, N)$ の行列であり列スコアをあらわす。ただし、 $\mathbf{P}_{all}, \mathbf{Q}_{all}$ の第1列は、トリビアルな解(特異値1に対応する解)である。

スコアを与えることによって生成される $2K$ 次元の分布の共分散行列はつぎのようになる。

$$\mathbf{S} = \begin{pmatrix} \mathbf{P}'\mathbf{F}\mathbf{P} & \mathbf{P}'\mathbf{D}\mathbf{Q} \\ \mathbf{Q}'\mathbf{D}'\mathbf{P} & \mathbf{Q}'\mathbf{G}\mathbf{Q} \end{pmatrix} \quad (1)$$

ただし、 \mathbf{P} は \mathbf{P}_{all} のうち、ノントリビアルな K 個の上位の解に対応する列ベクトルであり、 \mathbf{Q} も同様である。対応分析が定める制約条件より、 $\mathbf{P}'\mathbf{F}\mathbf{P}$ と $\mathbf{Q}'\mathbf{G}\mathbf{Q}$ とは単位行列であり、また $\mathbf{P}'\mathbf{D}\mathbf{Q}$ は対角行列である。

ここで、スコアがつくり出す空間配置の望ましさをつぎの値を用いて表すことにする。ここで n はデータの総件数である。

$$l(\hat{\Sigma}|\mathbf{S}) = -\frac{n}{2} \log(\det \mathbf{S}) \quad (2)$$

対応分析はこの指標を最大にする \mathbf{S} を \mathbf{P}, \mathbf{Q} についての標準的な制約下で求める方法であるとみなせる。これを尤度とみなすことは難しいが、対応分析の解の望ましさを指標としては妥当なものと考えられる。

行スコアと列スコアの線形関数に事前正規分布を仮定して導かれる対数尤度は、

$$l(\mathbf{P}|\mathbf{R}_A) = \text{constant} - \frac{1}{2} \text{trace} \mathbf{P}'\mathbf{A}'\mathbf{R}_A^{-1}\mathbf{A}\mathbf{P}, \quad l(\mathbf{Q}|\mathbf{R}_B) = \text{constant} - \frac{1}{2} \text{trace} \mathbf{Q}'\mathbf{B}'\mathbf{R}_B^{-1}\mathbf{B}\mathbf{Q} \quad (3)$$

となる。係数行列 \mathbf{A}, \mathbf{B} はペナルティの構造を定義し、 \mathbf{R}_A と \mathbf{R}_B は、ペナルティの強度を定義する正定値対称行列である。下に示す例では $\mathbf{R}_A, \mathbf{R}_B$ は単位行列の定数倍とし、 \mathbf{A} は行の2元配置の構造を反映するもの

としている。目的関数を (2) および (3) の和とし、これを最大にする P および Q を得れば、指定された制約を考慮した対応分析の解を得ることが期待できる。

3 計算例

高根 (1995) に取り上げられている Greenacre(1993) のデータについて分析を行なった。これは 1988 年第 4 四半期にアメリカで購入された乗用車の台数を、購入者の年齢・収入と車種別に集計したものである。年齢層は 7 階層、収入は 9 階層である。また車種は 14 種である。集計表は表側 (行) が年齢×収入、表頭 (列) が車種の 63×14 の大きさをもつ。データの総件数は、約 58 万件であるが、これから 5,000 件分をランダムにサンプリングし分析対象とした。以下の性質が確認された。

1. A に 2 階差分または、高根 (1995) と類似の制約 (部分空間制約) を指定すると、全てのデータを用いた場合に近い解が得られた。特に、このデータについては後者の制約が妥当である。
2. リサンプリングによって制約強度の決定を試みたが、データが表の大きさにくらべてスパースなため、制約の効果について妥当な評価を得ることができない。評価用データの生成法を工夫する必要がある。

図 1(a) が制約を加えない通常に対応分析による行スコアであり、図 1(b) が行スコアに部分空間制約を課した場合の解である。後者の方が解釈の容易な配置になっており、また、シミュレーションの結果、適合度 (2) の期待値がより大きくなることが確認された。

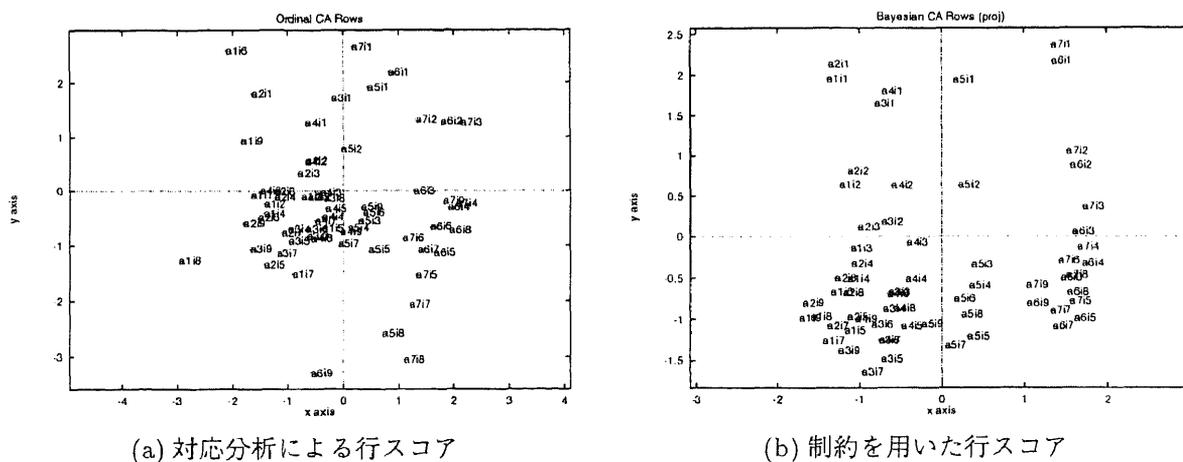


図 1: 行スコアの比較

参考文献

- [1] Bekker, P. & de Leeuw, J. (1988) Relation between variants of non-linear principal component analysis, in van Rijckevorsel & de Leeuw (1988), 1-31.
- [2] 大津起夫 (1993), OSMOD とその拡張: 人工データによる特性の検討, 行動計量学, 20, No.2, 9-23. ;
- [3] Takene, Y. & Shibayama, T. (1991). Principal component analysis with external information on both subjects and variables. *Psychometrika*, 56, 97-120.;
- [4] 高根芳雄, (1995), 制約付き主成分分析法, 東京:朝倉書店.

Combining the Advantages of One-sided and Two-sided Test Procedures for Comparing Several Treatments with a Control

A. J. Hayter

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0205, U.S.A.

Tetsuhisa Miwa

Laboratory of Statistics
National Institute of Agro-Environmental Sciences
3-1-1 Kannondai, Tsukuba 305-8604, Japan

Wei Liu

Department of Mathematics
University of Southampton
Southampton, S09 5NH, England

Abstract

Consider the standard problem of investigating whether any of k treatments are better than a control. A two-sided test procedure provides both upper and lower bounds on the differences between each treatment and the control, whereas a one-sided test procedure only provides lower bounds on these differences. However, the one-sided test procedure allows sharper inferences regarding which treatments can be declared to be better than the control. In this paper we develop a new test procedure which combines the good aspects of both the one-sided and the two-sided test procedures. This new test procedure maintains the inferential sensitivity of the one-sided test procedure while also providing both upper and lower bounds on the differences between each treatment and the control. A new set of critical points is needed which are tabulated for the balanced case where the treatment sample sizes are all equal. More generally, the new test procedure is applicable to any set of correlated or uncorrelated parameter estimates.

Consider the one-sided and two-sided critical points $T_{k,\alpha,\nu,\rho}^{(1)}$ and $T_{k,\alpha,\nu,\rho}^{(2)}$ which are the solutions to the equations

$$\int_{s=0}^{\infty} \int_{z=-\infty}^{\infty} \phi(z) g(s) \left[\Phi \left(z \sqrt{\frac{\rho}{1-\rho}} + s T_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{1-\rho}} \right) \right]^k dz ds = 1 - \alpha$$

and

$$\int_{s=0}^{\infty} \int_{z=-\infty}^{\infty} \phi(z) g(s) \left[\Phi \left(z \sqrt{\frac{\rho}{1-\rho}} + sT_{k,\alpha,\nu,\rho}^{(2)} \sqrt{\frac{1}{1-\rho}} \right) - \Phi \left(z \sqrt{\frac{\rho}{1-\rho}} - sT_{k,\alpha,\nu,\rho}^{(2)} \sqrt{\frac{1}{1-\rho}} \right) \right]^k dz ds$$

And define the critical point $C_{k,\alpha,\nu,\rho}^{(1)}$ by

$$\int_{s=0}^{\infty} \int_{z=-\infty}^{\infty} \phi(z) g(s) \left[\Phi \left(z \sqrt{\frac{\rho}{1-\rho}} + sT_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{1-\rho}} \right) \right]^{k-1} \left[\Phi \left(z \sqrt{\frac{\rho}{1-\rho}} + sC_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{1-\rho}} \right) - \Phi \left(z \sqrt{\frac{\rho}{1-\rho}} - sC_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{1-\rho}} \right) \right] dz ds = 1 - \alpha.$$

The following are $1 - \alpha$ level simultaneous confidence intervals for comparisons with a control in the balanced case with $\rho = n_T/(n_T + n_0)$, where n_T is the common treatment sample size and n_0 is the control sample size.

- If $\bar{X}_i - \bar{X}_0 > ST_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{n_T} + \frac{1}{n_0}}$ for each i , $1 \leq i \leq k$, then

$$\mu_i - \mu_0 \in \left(\max \left\{ 0, \bar{X}_i - \bar{X}_0 - ST_{k,\alpha,\nu,\rho}^{(2)} \sqrt{\frac{1}{n_T} + \frac{1}{n_0}} \right\}, \bar{X}_i - \bar{X}_0 + ST_{k,\alpha,\nu,\rho}^{(2)} \sqrt{\frac{1}{n_T} + \frac{1}{n_0}} \right), 1 \leq$$

- Otherwise,

$$\mu_i - \mu_0 \in \left(\max \left\{ 0, \bar{X}_i - \bar{X}_0 - SC_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{n_T} + \frac{1}{n_0}} \right\}, \bar{X}_i - \bar{X}_0 + SC_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{n_T} + \frac{1}{n_0}} \right)$$

for i such that $\bar{X}_i - \bar{X}_0 > ST_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{n_T} + \frac{1}{n_0}}$, and

$$\theta_i \in \left(\bar{X}_i - \bar{X}_0 - ST_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{n_T} + \frac{1}{n_0}}, \max \left\{ 0, \bar{X}_i - \bar{X}_0 + SC_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{n_T} + \frac{1}{n_0}} \right\} \right)$$

for i such that $\bar{X}_i - \bar{X}_0 \leq ST_{k,\alpha,\nu,\rho}^{(1)} \sqrt{\frac{1}{n_T} + \frac{1}{n_0}}$.

In this paper a new test procedure has been developed for making inferences on a set of treatment effects with possibly correlated estimates, where a one-sided direction of interest has been specified. In particular, the new test procedure can be applied to the common problem of comparing a set of treatments with a control. The new procedure combines the advantages of traditional one-sided and two-sided test procedures by providing at the same specified confidence level the same directional inferences as the one-sided test procedure as well as providing both upper and lower confidence bounds.

The new procedure is developed by inverting a set of suitably defined acceptance sets specified throughout the parameter space. As well as the one-sided and two-sided critical points, additional critical points are required to implement the new procedure. Tables of these new critical points are provided for the case where the treatment effect estimates are equally correlated, which includes the problem of comparisons with a control when the treatment sample sizes are all equal.