

(16) 「統計的推測理論とその応用」に関する研究報告

Etsuo Kumagai, Nobuo Inagaki (Osaka University) : An exact information loss in the multivariate gamma distribution	663
津田美幸 (中央大・COE) : 量子 Bhattacharyya 不等式	665
丸山芳人 (東京理科大・理研) : Measures of multivariate kurtosis in elliptical distributions	667
布能英一郎 (関東学院大学経済学部) : Worcester's log-linear model for four or more dimensions	669
田中研太郎 (東工大) : クロスバリデーションを用いた混合分布モデルの推定についてー統計的推測理論とその応用ー	671
藤木美江 (大阪大学大学院基礎工学研究科), 白旗慎吾 (大阪大学大学院基礎工学研究科) : 非線形モデルにおける最深回帰推定量	673
西山陽一 (統計数理研究所) : Nonparametric inference for Lévy processes by continuous observation: a martingale approach	675
藤井孝之 (大阪大学基礎工) : 連続型閾値モデルの離散観測について	677
舞原寛祐 (筑波大・数理物質科学) : 正規分布の平均に関する検定の重ね合わせによる多重決定方式	679
肖 玉山 (長春大学・理), 高田佳和 (熊本大学・工) : 順序制約のもとでの予測量の改良	681
藤澤洋徳 (統計数理研究所) : 外れ値の割合が多い場合にもバイアスが小さいロバスト推定	683
柿沢佳秀 (北海道大学大学院経済学研究科) : 多変量母集団の平均ベクトルに関する DL 法とボンフェローニ型不等式による同時信頼区間	685
岩佐 学 (熊本大学大学院自然科学研究科) : 同等性の検定と信頼区間について	687

An exact information loss in the multivariate gamma distribution

Osaka University Etsuo Kumagai
Osaka University Nobuo Inagaki

In this paper, we propose a generalized hyperbolic model as an extension of the Nile problem by Fisher [2], which showed the exact information loss with a hyperbolic curve as a parameter in the two dimensional exponential distribution, and we investigate the exact information loss with respect to MLE in the generalized hyperbolic model. In this calculation, we use a projection matrix directly, which is partially based on the projection with respect to the conditional expectation given MLE in Inagaki [3].

Let k be a positive integer and let random variables $\{X_i\}_{i=1}^k$ be independent mutually and be each distributed with the gamma distribution $\{G_A(q_i, \alpha_i^{-1})\}_{i=1}^k$, where $\{q_i\}_{i=1}^k$ and $\{\alpha_i\}_{i=1}^k$ are constants and parameters respectively which are all positive and finite. We consider the k -dimensional gamma distribution (X_1, \dots, X_k) whose probability density function is

$$\begin{aligned} p(x_1, \dots, x_k) &= \exp\{\langle \alpha, \mathbf{x} \rangle - \psi(\alpha)\} h(\mathbf{x}) \\ &= \exp\left\{-\sum_{i=1}^k \alpha_i x_i + \sum_{i=1}^k \log \alpha_i^{q_i}\right\} \prod_{i=1}^k \frac{x_i^{q_i-1}}{\Gamma(q_i)}. \end{aligned}$$

We consider the k -dimensional gamma distribution whose parameters satisfy the condition $\psi(\alpha(\theta)) = 0$, that is, parameters have the relationships as follows:

$$\alpha(\theta) = \begin{pmatrix} \alpha_1(\theta) \\ \vdots \\ \alpha_{k-1}(\theta) \\ \alpha_k(\theta) \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_{k-1} \\ \vartheta^{-1} \end{pmatrix}, \quad \theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_{k-1} \end{pmatrix},$$

where

$$\vartheta = \prod_{i=1}^{k-1} \theta_i^{r_i}, \quad 0 < \theta_1, \dots, \theta_{k-1} < \infty, \quad r_1 = \frac{q_1}{q_k}, \dots, r_{k-1} = \frac{q_{k-1}}{q_k}.$$

Note that r_1, \dots, r_{k-1} are positive constants. We call this a *generalized hyperbolic model*. The following theorem holds with respect to this generalized hyperbolic model:

Theorem 1 In the k -dimensional generalized hyperbolic model, let $P^*(\theta)$ be the projection matrix into the subspace spanned by $\Sigma^{\frac{1}{2}}(\theta)\dot{\alpha}(\theta)$, that is,

$$P^*(\theta) = \Sigma^{\frac{1}{2}}(\theta)\dot{\alpha}(\theta) \left({}^t\dot{\alpha}(\theta)\Sigma(\theta)\dot{\alpha}(\theta) \right)^{-1} {}^t\dot{\alpha}(\theta)\Sigma^{\frac{1}{2}}(\theta),$$

so that the generalized statistical curvature $\Gamma(\theta)^2$ is represented as follows:

$$\begin{aligned} \Gamma(\theta)^2 &= \left({}^t\dot{\alpha}(\theta)\Sigma(\theta)\dot{\alpha}(\theta) \right)^{-2} \left\{ {}^t\ddot{\alpha}(\theta) \left[\mathbf{I}_{k-1} \otimes \Sigma^{\frac{1}{2}}(\theta)(\mathbf{I}_k - P^*(\theta))\Sigma^{\frac{1}{2}}(\theta) \right] \ddot{\alpha}(\theta) \right\} \\ &= \frac{1}{q_1 + \cdots + q_k} \mathbf{I}_{k-1}, \end{aligned}$$

where \mathbf{I}_{k-1} and \mathbf{I}_k are the $k-1$ and k dimensional identity matrices, respectively, and the notation \otimes means the Kronecker product. \square

Theorem 2 The exact Fisher information loss of k -dimensional hyperbolic model is represented by

$$I_n(\theta) - I_T(\theta) = \frac{n}{n(q_1 + \cdots + q_k) + 1} I(\theta).$$

\square

The exact information loss converges to the following:

$$\lim_{n \rightarrow \infty} \{I_n(\theta) - I_T(\theta)\} = \frac{1}{q_1 + \cdots + q_k} I(\theta),$$

which is equivalent to a natural expansion of the definition of statistical curvature by Efron[1].

References

- [1] B.Efron, Defining the curvature of a statistical problem (with applications to second order efficiency), Ann. Statist. 6 (1975) 1189–1242.
- [2] R.A.Fisher, *Statistical Methods and Scientific Inference*, Hafner Press, New York, 1973.
- [3] N.Inagaki, The asymptotic distribution of a conditional likelihood ratio function given a maximum likelihood estimator in *Statistical Theory and Data Analysis (K.Matusita(Editor))*, North-Holland, Amsterdam, 1985.
- [4] N.Inagaki and E.Kumagai, Exact information loss in Fisher's circle model, *Mathematica Japonica*, 44 (1996) 455–467.
- [5] K.Inoue, N.Inagaki and E.Kumagai, The exact information loss of the maximum likelihood estimator in the k -dimensional sphere model, *Mathematica Japonica*, 52 (2000) 289–310.
- [6] E.Kumagai and N.Inagaki, Exact information loss in the multivariate gamma distribution, *Scientiae Mathematicae Japonicae*, 62 (2005), 91–101.

量子 Bhattacharyya 不等式

中央大・COE 津田美幸

先日の統計関連学会連合大会での報告に引き続き、不偏推定に関する Bhattacharyya 不等式の量子力学系への拡張を説明する。古典との違いは以下のポイントにある。

- (1) 密度 (=状態) とその (高階) 微分との比の定義の仕方に任意性があること。
- (2) 下限を達成する測定の POVM による実現可能性条件があること。

(1) は作用素の積の非可換性に由来する。良い下限を得るためには、微分の情報を最大限に引き出すような乗算規則に基づいて比を定める必要がある。典型的には S 型/R 型/L 型の 3 種類があり、母数空間 Θ の次元や被推定量の形などにより使い分ける。

(2) は物理系の観測に関わる上で最低限必要な条件である。(他の問題は更に、局所性など、別の条件が追加される。) 古典と同様、不等式の下限の達成条件は Schwartz の不等式の等号成立条件に基づいて表され、その等式で得られる作用素が θ に依存せず、しかもその作用素が表す測定が POVM で実現できるならば、一様に最適な不偏推定量が与えられる。しかし一般には、測定の実現可能性条件は扱いにくい形をしている。特に問題となるのは複数の実数値物理量の同時測定可能性である。(例えば $\Theta = \mathbb{R}^2$ の場合、不偏推定量の分散行列 V の同時最小化問題において、一般にはトレードオフを引き起こすため、リスクとして $\text{Tr} V$ などが用いられる。) ところで、被推定量が実数か複素数ならば、任意の推定量は、半正規 (subnormal) な一般化オブザーバブルの正規拡大、というシンプルな議論に帰着され (Halmos (1950), Yuen&Lax (1973), Hayashi&Sakaguchi (2000)), しかも、量子ガウス状態モデルなど、重要な応用例をもつ。

本研究では、 $\Theta = \mathbb{R}$, $\Theta = \mathbb{C}$ として量子版の Bhattacharyya 不等式を導出し、達成可能性条件を明らかにした。また、量子ガウス状態モデルの振幅 θ の (正則とは限らない) 多項式 $g(\theta)$ の推定に応用し、以下のことを示した。

新しい結果						従来の結果	
Θ	\mathbb{R}		\mathbb{C}			\mathbb{R}	\mathbb{C}
被推定量	θ^2	θ^3	正則	反正則	実数値	θ	θ
一様最適不偏推定量	存在	非存在	存在	存在	存在	存在	存在
達成する下限	S 型	—	R 型	L 型	R 型 & L 型	S 型	R 型
正規拡大	不要	—	必要	必要	不要	不要	必要
測定の古典物理的解釈	不可能	—	ヘテロ	ヘテロ	不可能	ホモ	ヘテロ

(正則 $\Leftrightarrow dg/d\bar{\theta} = 0$, 反正則 $\Leftrightarrow dg/d\theta = 0$, 実数値 $\Leftrightarrow g(\theta) = g(\bar{\theta})$,
ホモ=ホモダイナ検波, ヘテロ=ヘテロダイナ検波.)

量子ガウス状態は古典的正規分布の自然な拡張である。古典ならば、Hermite 多項式を用いて任意の多項式の UMVUE が構成できる。この性質は、 $\Theta = \mathbb{C}$ の量子ガウスに受け継がれている。ただし、最適推定量の記述には Laguerre 陪多項式が用いられる。

意外なことに、 $\Theta = \mathbb{R}$ だと θ^3 のように最適推定量が存在しない例が存在する。これは、量子版の推定量のクラスが古典に比べて大きいため、不偏性の制約が相対的に緩くなり、リスクを一様に最小にする推定量が存在し得ない、ことを示している。ただし、不偏性の下で許容的な θ^3 の推定量は存在する。

基本設定 量子力学系は、ヒルベルト空間 \mathcal{H} を用いて記述される。 \mathcal{H} 上の非負定値自己共役線形作用素の全体を \mathcal{L}_+ とする。系の状態は、 $\text{Tr} \rho = 1$ を満たす $\rho \in \mathcal{L}_+$ で表される。集合 Ω の元を観測値とする測定 (Ω 上の測定) は、 Ω の部分集合の族 \mathcal{B} から \mathcal{L}_+ への写像 M で表される。ただし、次が成り立つものとする: $M(\emptyset) = 0$, $M(\Omega) = \text{恒等作用素}$, 互いに素な \mathcal{B} の元の列 $\{A_n\}$ に対して $M(\bigcup A_n) = \sum M(A_n)$. 事象 $A \in \mathcal{B}$ が観測される確率は $\text{Tr}(\rho M(A))$ で与えられる。これにより、 Ω 上に確率分布が定まる。

量子 Bhattacharyya 不等式 $\Theta = \mathbb{R}$ または $\Theta = \mathbb{C}$ で, 系の状態が未知の母数 $\theta \in \Theta$ によって ρ_θ の形で与えられているとする. ρ_θ のサポートは θ によらず一定とする. 多項式 $g(\theta)$ の推定量 T は $g(\theta)$ 上の測定によって定まる確率変数で, 無数に存在する. T に関する不偏性条件 $E_\theta(T) \equiv g(\theta)$ の下で, $\text{MSE}_\theta(T) = E_\theta(|T - g(\theta)|^2)$ を最小化する問題において, 次が成り立つ.

定理 (量子 Bhattacharyya 不等式) $\Theta = \mathbb{R}$ で T が $g(\theta)$ の不偏推定量ならば,

$$(S \text{ 型}) \quad \text{MSE}_\theta(T) \geq {}^t(\bar{D}_k^R[g(\theta)])(J_k^S)^{-1}\bar{D}_k^R[g(\theta)].$$

$\Theta = \mathbb{C}$ で T が $g(\theta)$ の不偏推定量ならば,

$$(R \text{ 型}) \text{MSE}_\theta(T) \geq (\bar{D}_k^C[g(\bar{\theta})])^\dagger (J_k^R)^{-1} \bar{D}_k^C[g(\bar{\theta})], \quad (L \text{ 型}) \text{MSE}_\theta(T) \geq (\bar{D}_k^C[g(\bar{\theta})])^\dagger (J_k^L)^{-1} \bar{D}_k^C[g(\bar{\theta})].$$

ただし, $\bar{D}_k^R = {}^t(\frac{d}{d\theta}, \frac{d^2}{d\theta^2}, \dots, \frac{d^k}{d\theta^k})$, $\bar{D}_k^C = {}^t(\frac{d}{d\theta}, \frac{d}{d\bar{\theta}}, \frac{d^2}{d\theta^2}, \frac{d^2}{d\theta d\bar{\theta}}, \frac{d^2}{d\bar{\theta}^2}, \dots, \frac{d^k}{d\theta^k}, \dots, \frac{d^k}{d\theta^l d\bar{\theta}^{k-l}}, \dots, \frac{d^k}{d\bar{\theta}^k})$ で, J_k は \bar{L}_k に基づいて次のように定められるものである ($\text{Tr}[\cdot]$ は成分毎のトレース):

$$J_k^S = \text{Tr}[\rho_\theta \bar{L}_k^S (\bar{L}_k^S)^\dagger], \quad J_k^R = \text{Tr}[\rho_\theta \bar{L}_k^R (\bar{L}_k^R)^\dagger], \quad J_k^L = \text{Tr}[\bar{L}_k^L \rho_\theta (\bar{L}_k^L)^\dagger], \\ \bar{D}_k^R[\rho_\theta] = \frac{\rho_\theta \bar{L}_k^S + \bar{L}_k^S \rho_\theta}{2}, \quad \bar{D}_k^C[\rho_\theta] = \rho_\theta \bar{L}_k^R, \quad \bar{D}_k^C[\rho_\theta] = \bar{L}_k^L \rho_\theta.$$

等号達成条件 上記定理の3つの不等式の等号部分は Schwartz の不等式の等号成立条件が満たされる場合に限り成り立つ. 従って, 作用素 $\bar{T} = \bar{D}_k[g(\bar{\theta})]^\dagger (J_k)^{-1} \bar{L}_k$ が θ に依存せず, しかも \bar{T} が半正規であることが必要十分である. それが成り立つ場合, 最適推定量は \bar{T} の正規拡大 \hat{T} で与えられる.

量子ガウスの振幅の多項式の推定 複素振幅が θ の量子ガウス状態は次で定義される:

$$\rho_\theta = \frac{1}{\pi N} \int_{\mathbb{R}^2} |\alpha\rangle \langle \alpha| \exp\left(-\frac{|\alpha - \theta|^2}{N}\right) dx dy \quad (\alpha = x + \sqrt{-1}y).$$

ただし, $|\alpha\rangle$ はコヒレントベクトルで, \mathcal{H} の正規直交基底 $\{e_n\}_{n=0}^\infty$ に対して $|\alpha\rangle = e^{-|\alpha|^2/2} \sum_n \alpha^n / \sqrt{n!} e_n$ で定義される.

このモデルでは R 型と L 型の \bar{L}_k , J_k の一般形が明示的に書ける. $k = 2, 3$ の場合の \bar{L}_k^S , J_k^S は, R 型を用いて具体的に書ける. 次が成り立つ.

定理 $\Theta = \mathbb{R}$ の場合, $g(\theta) = \theta^2$ ならば次の推定量が S 型下限を一樣に達成する:

$$\hat{T} = \frac{N(N+1)}{(2N+1)^2} (a^2 + a^\dagger^2) + \frac{N^2 + (N+1)^2}{(2N+1)^2} a^\dagger a - N \frac{N^2 + (N+1)^2}{(2N+1)^2}.$$

ただし, a は \mathcal{H} 上の消滅作用素で, $a = \sum_n \sqrt{n+1} e_n e_{n+1}^\dagger$. 一方, $g(\theta) = \theta^3$ の一樣最適推定量は存在しない.

$\Theta = \mathbb{C}$ の場合, $g(\theta)$ が, 正則ならば R 型下限, 反正則ならば L 型下限, 実数値ならば R 型と L 型の両方, を一樣に達成する推定量が存在して, 次のようになる:

$$(正則) \quad \hat{T} = g(a + b^\dagger) \quad (b \text{ は独立な真空状態 } |0\rangle\langle 0| \text{ をもつ系の消滅作用素}),$$

$$(反正則) \quad \hat{T} = \bar{g}(a^\dagger + b) \quad (\bar{g}(z) = g(\bar{z})),$$

$$(実数値) \quad \hat{T} = \sum_{m,n} c_{n,m} (N+1)^n \sum_{r=0}^{\min(m,n)} (-1)^{\min(m,n)-r} \binom{\max(m,n)}{\min(m,n)-r} \frac{\min(m,n)!}{r!} \\ \times \left(\frac{a}{N+1}\right)^{r+\max(0,n-m)} a^{\dagger r+\max(0,m-n)} \quad (g(\theta) = \sum_{m,n} c_{m,n} \theta^m \bar{\theta}^n).$$

$g(a + b^\dagger)$, $\bar{g}(a^\dagger + b)$ は, それぞれ, $g(a)$, $\bar{g}(a^\dagger)$ の正規拡大である.

注意 Brody & Hughston (1998,2001) は量子情報幾何の中で Bhattacharyya 不等式を別の形で量子版へ拡張しているが, S 型/R 型/L 型の区別, 一般化オブザーバブルの正規拡大, 小標本での達成可能性等の議論はなく, Bhattacharyya 不等式を量子系へ有効な形で拡張したのは我々が最初である. また, Tsuda & Matsumoto (2005) の量子 Koike 下限は $\Theta = \mathbb{R}$ の S 型下限を差分の極限として含む.

Measures of multivariate kurtosis in elliptical distributions

丸山 芳人（東京理科大・理研）

非正規母集団において、3次キュムラントと4次キュムラントの推定問題を考える。一般に、多変量統計解析における基本的な検定統計量（極限分布がカイ二乗分布になる Hotelling's T^2 , MANOVA 検定統計量, 多変量線型仮説検定統計量など）の分布の漸近展開式（厳密には、極限項に続く第2項までの展開）には、母集団の3次キュムラントと4次キュムラントが含まれ、実際の使用においてはそれらの推定が問題になっている。現在常用されている推定量の一つは、Mardia [Ma70] が定義した多変量歪度, 尖度の標本尺度 (sample measures of skewness and kurtosis, 略して SMS, SMK) を基にした一致推定量である。この他に、Srivastava [S84] は主成分を用いて異なる SMS, SMK を提案しており、いずれも正規性の検定に用いられる。一方、推定量の性質に関して、SMS, SMK の極限分布はそれぞれ、カイ二乗分布, 正規分布になることが知られている。また、正規母集団のみならず、非正規性の仮定を少し緩めた Elliptical 分布の下での議論も行われてきた ([Ma74], [BB90], [H94], [ST96])。しかし、[Ma70] を基にした SMS, SMK に関して数多くの結果があるものの、[S84] による SMS, SMK の性質についてはよく分かっていない。もう少し詳しく述べると、母集団が正規分布の場合に、Mardia [Ma74] は SMS, SMK の正確な平均と分散を導出した。また、SMK を基に Berkane and Bentler [BB90] は Elliptical 分布において4次キュムラントと本質的に同等な尖度パラメータの推定について論じた。そしてそのようなときに、Seo and Toyama [ST96] は尖度パラメータの推定量の漸近展開について考察した。更に、Henze [H94] は一般分布の下で SMK の漸近分散を導出した。

本報告においては、まず、非正規分布の典型である Elliptical 分布の下で、Srivastava's SMK の期待値に関する漸近展開及び漸近分散を導出し、4次キュムラントの推定量の漸近的性質を調べる。また、Mardia's SMK の漸近結果 ([ST96] を参照) と理論的、数値的に比較し、推定の精度を考察する。これらの議論は尖度パラメータを一般化したモーメントパラメータに拡張できる (Maruyama [M05] を参照)。なお、3次キュムラントの推定、及び母集団がより一般の非正規分布の場合は、摂動展開後の期待値の計算と式整理が複雑でやや困難と予想されるが、現在考察中である。

参考文献

- [BB90] Berkane, M. and Bentler, P. M. (1990). Mardia's coefficient of kurtosis in elliptical populations, *Acta Math. Appl. Sinica* (English Ser.), **6**, 289-294.
- [H94] Henze, N. (1994). On Mardia's kurtosis test for multivariate normality. *Comm. Statist. Theory Methods* **23**, 1031-1045.
- [Ma70] Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika* **57**, 519-530.
- [Ma74] Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies, *Sankhya*, **B36**, 115-128.
- [M05] Maruyama, Y. (2005). Asymptotic properties for measures of multivariate kurtosis in elliptical distributions, to appear.
- [MS03] Maruyama, Y. and Seo, T. (2003). Estimation of moment parameter in elliptical distributions, *J. Japan Statist. Soc.*, **33**, 215-229.
- [S84] Srivastava, M. S. (1984). A measure of skewness and kurtosis and a graphical method for assessing multivariate normality, *Statist. Probab. Lett.*, **2**, 263-267.
- [ST96] Seo, T. and Toyama, T. (1996). On the estimation of kurtosis parameter in elliptical distributions. *J. Japan Statist. Soc.* **26**, 59-68.

Worcester's log-linear model for four or more dimensions

関東学院大学経済学部 布能 英一郎

1. Introduction Worcester(1971) は、 $2 \times 2 \times 2$ の 3 次元分割表 $\{m_{ijk}\}_{i,j,k=0,1}$ に対し、Worcester's log-linear model とよばれる統計モデルを考察した。近年、本講演者はこのモデルを一般の 3 次元分割表へ次のように拡張した：

$$\begin{aligned} \ln m_{ijk} = & w + \sum_{a=1}^I \delta_i^{[a]} w_1^{[a]} + \sum_{b=1}^J \delta_j^{[b]} w_2^{[b]} + \sum_{c=1}^K \delta_k^{[c]} w_3^{[c]} + \sum_{a=1}^I \sum_{b=1}^J \delta_{ij}^{[ab]} w_{12}^{[ab]} \\ & + \sum_{a=1}^I \sum_{c=1}^K \delta_{ik}^{[ac]} w_{13}^{[ac]} + \sum_{b=1}^J \sum_{c=1}^K \delta_{jk}^{[bc]} w_{23}^{[bc]} + \sum_{a=1}^I \sum_{b=1}^J \sum_{c=1}^K \delta_{ijk}^{[abc]} w_{123}^{[abc]} \end{aligned} \quad (1)$$

where

$$\begin{aligned} \delta_i^{[a]} &= \begin{cases} 1 & \text{if } a \leq i, \\ 0 & \text{else} \end{cases}, \quad \delta_{ij}^{[ab]} = \begin{cases} 1 & \text{if } a \leq i \text{ and } b \leq j \\ 0 & \text{else,} \end{cases} \\ \delta_{ijk}^{[abc]} &= \begin{cases} 1 & \text{if } a \leq i, b \leq j, \text{ and } c \leq k \\ 0 & \text{else} \end{cases} \end{aligned}$$

そして、いくつかのパラメーターがゼロのとき、 m_{ijk} の MLE が exact に求められることを示した。主な結果は次のとおりである：

Theorem 1.1 $w_{13} = 0, w_{23} = 0$ ならば

$$\hat{m}_{ijk} = x_{ijk} \quad \text{if } i, j \geq 1, \quad \hat{m}_{ijk} = \frac{x_{ij+}(x_{+0k} + x_{0+k} - x_{00k})}{x_{+0+} + x_{0++} - x_{00+}} \quad \text{if } i \text{ or } j = 0$$

Theorem 1.2 $w_{23} = 0$ ならば

$$\hat{m}_{ijk} = x_{ijk} \quad \text{if } i \geq 1, \quad \hat{m}_{ijk} = \frac{x_{0j+}x_{0+k}}{x_{0++}} \quad \text{if } i = 0$$

Theorem 1.3 Theorem 1.1 の条件に加え、更に $w_{123}^{[abc]} = 0$ for all $(a, b) \neq (1, 1)$ を仮定すると、 i または j がゼロの場合は、Theorem 1.1 に同じで

$$\hat{m}_{ijk} = \frac{x_{ij+} \sum_{a \geq 1} \sum_{b \geq 1} x_{abk}}{\sum_{a \geq 1} \sum_{b \geq 1} x_{ab+}}, \quad \text{if } i, j \geq 1.$$

Theorem 1.4 Theorem 1.1 の条件に加え、更に $w_{12}^{[ab]} = 0$ for all $a \geq 2, b \geq 2$, および $w_{123}^{[abc]} = 0$ for all (a, b, c) except $(a, b) \neq (1, 1)$ を仮定すると、 i または j がゼロの場合は、Theorem 1.1 に同じで

$$\hat{m}_{ijk} = \frac{(\sum_{b \geq 1} x_{ib+})(\sum_{a \geq 1} x_{aj+})x_{++k}}{(\sum_{a \geq 1} \sum_{b \geq 1} x_{ab+})^2}, \quad \text{if } i, j \geq 1.$$

今回は、これらを 4 次元以上の分割表に拡張したものである。

2. 本論 4 次元（およびそれ以上）の分割表に対して、(1) を拡張して Worcester's log-linear model を定義するのは、形式的な代数操作で出来る。統計学としての興味があるのは、4 次元（以上）に対して、Theorem 1.1 – 1.4 の拡張が出来るか？であろう。

2.1 4次元分割表に対して 4次元分割表に対して Worcester's log-linear model をあてはめたとき、次のような結果が得られた：

Theorem 2.1 $w_{12} \neq 0, w_{34} \neq 0, w_{123} \neq 0, w_{124} \neq 0, w_{1234} \neq 0$ を仮定すると

$$\hat{m}_{ijkl} = \begin{cases} \frac{x_{ij++} \times (x_{+0kl} + x_{0+k-l} - x_{00kl})}{x_{+0++} + x_{0+++} - x_{00++}} & \text{if } i = 0, \text{ or } j = 0 \\ x_{ijkl} & \text{if } i, j \geq 1. \end{cases}$$

Theorem 2.2 $w_{23} \equiv w_{24} \equiv w_{34} \equiv w_{234} \equiv 0$ を仮定すると

$$\hat{m}_{ijkl} = \begin{cases} x_{ijkl} & \text{if } i \geq 1, \\ \frac{x_{0j++}x_{0+k+l}x_{0++l}}{x_{0+++}^2} & \text{if } i = 0. \end{cases}$$

Theorem 2.3 $w_{12} \neq 0, w_{34} \neq 0$ に加え、更に $w_{12}^{[ab]} \neq 0$ if $a = 1$ or $b = 1$, $w_{123}^{[11c]} \neq 0$ for all $c = 1, 2, \dots$, $w_{1234}^{[11cd]} \neq 0$ for all $c, d = 1, 2, \dots$ そうすると、 $i, j \geq 1$ とき、

$$\hat{m}_{ijkl} = \frac{(\sum_{b \geq 1} x_{ib++})(\sum_{a \geq 1} x_{aj+})x_{++kl}}{T_1^2} \quad k = 0, 1, 2, \dots, K, \quad l = 0, 1, 2, \dots, L$$

i, j の少なくとも1つがゼロのとき、

$$\hat{m}_{ijkl} = \frac{x_{ij+}(x_{+0k} + x_{0+k-l} - x_{00kl})}{T_2}, \quad k = 0, 1, 2, \dots, K, \quad l = 0, 1, 2, \dots, L$$

where

$$T_1 = \sum_{a \geq 1} \sum_{b \geq 1} x_{ab++}, \quad T_2 = \sum_{k,l} (x_{+0kl} + x_{0+k-l} - x_{00kl}) = x_{+0++} + x_{0+++} - x_{00++}$$

2.2 5次元以上の分割表に対して

Theorem 2.1, Theorem 2.2 に関しては、5次元以上の分割表に対しても、4次元の場合とほとんど平行に成り立つような定理が得られた。

2.3 MLE の直接解が存在しないとき

3次元分割表においても、MLE の直接解が存在しないことがある。たとえば、 $w_{12} \equiv w_{23} \equiv w_{13} \equiv 0, w_{123} \neq 0$ の場合である。このような場合、iterating scaling method を用いることで MLE が求められることがある。 $w_{12} \equiv w_{23} \equiv w_{13} \equiv 0, w_{123} \neq 0$ なる3次元分割表の場合は、iterating scaling method で MLE が求められる。同様に、4次元分割表で $w_{ij} \equiv w_{ijk} \equiv 0, w_{1234} \neq 0$, for all $1 \leq i < j < k \leq 4$ の場合も、iterating scaling method で MLE が求められることがわかった。他方、 $w_{12} \neq 0, w_{123} \neq 0, w_{1234} \neq 0$ の場合は、まだ未解決である。

参考文献 [1] Bishop, Fienberg and Holland. (1975). *Discrete Multivariate Analysis : Theory and Practice*. The MIT Press. [2] Worcester, J. (1971). The relative odds in the 2^3 contingency tables. *American Journal of Epidemiology* **93** 145-149.

クロスバリデーションを用いた混合分布モデルの推定について

— 統計的推測理論とその応用 —

田中 研太郎

平成 17 年 11 月 15 日 (火)

1 混合分布の母数空間

成分の数が M 個で各成分が位置尺度分布である混合分布の密度関数は $f(x; \theta) = \sum_{j=1}^M \alpha_j f_j(x; a_j, b_j)$ と表される. ここで, α_j は重みで, $f_j(x; a_j, b_j)$ は位置母数 a_j と尺度母数 b_j をもつ位置尺度分布の密度関数である. 位置尺度分布の混合分布の母数空間は

$$\Theta = \{ \theta = (\alpha_1, a_1, b_1, \dots, \alpha_M, a_M, b_M) \in \mathbb{R}^{3M} \mid \sum_{j=1}^M \alpha_j = 1, \alpha_j \geq 0, b_j > 0 \ (j = 1, \dots, M) \}$$

である. 真のモデルがこのモデルに含まれる場合に, 最尤推定によって母数を推定しようとしても, ある成分の位置母数を標本のある値と等しくとり, 同時にその尺度母数を 0 に近づけることにより, 尤度が発散する点列をとることができて, うまくいかないことが知られている.

ここで, パラメータ空間として尺度母数 b_j の大きさを $b_j \geq c$ (c は定数) と制限した母数空間を考えると, その母数空間においては尤度は有界になる. さらに, 制限された母数空間が真の母数を含めばその母数空間における最尤推定量は一致推定量になる. この条件を満たす c の値は真の母数に依存しているが, これをクロスバリデーションを用いて選択することを考える.

2 クロスバリデーションによる母数空間の選択

c によって制限された母数空間を Θ_c とする. c を固定し, 標本の h 番目を抜いたときの最尤推定量を $\hat{\theta}_{-h,c} = \operatorname{argsup}_{\theta \in \Theta_c} \sum_{i \neq h} \log f(x_i; \theta)$ とする. このとき, $CV(c) = \sum_{h=1}^n \log f(x_h; \hat{\theta}_{-h,c})$ を最大にする c の値を選択することにする.

このようにクロスバリデーションにより選択された c によって制限される母数空間において, 最尤推定量が一致性を持てば, その選択は妥当だといえる. この c の値は, 標本サイズ n に依存し得る.

いくつかの正則条件のもとで以下が成立する.

Theorem 2.1. (Tanaka and Takemura (2003))

真のモデルは M 個の成分を持つモデルのみによって表されるとする. $c_0 > 0$ と $0 < d < 1$ を固定し, c_n を $c_n = c_0 \cdot \exp(-n^d)$ とおく. Θ_n を

$$\Theta_n = \{ \theta \in \Theta \mid b_k \geq c_n, (k = 1, \dots, M) \} \quad (2.1)$$

とおく. このとき, Θ_n における最尤推定量は強一貫性を持つ. □

Theorem 2.2. 真のモデルは M 個の成分を持つモデルのみによって表されるとする. Θ'_n を

$$\Theta'_n = \{\theta \in \Theta \mid b_k \geq c'_n, (k = 1, \dots, M)\} \quad (2.2)$$

とおく. このとき, Θ'_n における最尤推定量は強一致性を持つ. \square

よって, \hat{c}_n で制限される母数空間 $\Theta_{\hat{c}_n}$ において最尤推定量の一致性が成り立つためには, 確率 1 で以下が成立すればよいことがわかる.

- \hat{c}_n が $\exp(-n^d)$, ($0 < d < 1$) よりも速く 0 に近づかない.
- $\limsup_{n \rightarrow \infty} \hat{c}_n \leq \min_{k=1, \dots, M} b_{k0}$ (b_{k0} は真の密度関数の各成分の尺度パラメータ.) つまり, 制限された母数空間は n が大きくなれば真のパラメータを含む (または近づく) ようになる.

3 シミュレーション

混合正規分布において, クロスバリデーションによって選択される \hat{c} の値がどのような挙動をするのかは, 理論的には分かっていない. そこで, 2 成分混合正規分布において簡単なシミュレーションを行った. 真の密度関数を

$$0.5 \cdot \phi(x; 0, 1) + 0.5 \cdot \phi(x; 2.5, 0.25) \quad (3.1)$$

とし, データを発生させた. サンプルサイズを 25, 50, 75, 100, 150, 200, 250, 500, 750, 1000 と増やして, そのときの \hat{c}_n をプロットしたのが図 1 である. 図 1 では \hat{c}_n が真のパラメータの分散の最小値 0.25 のまわりを動いている.

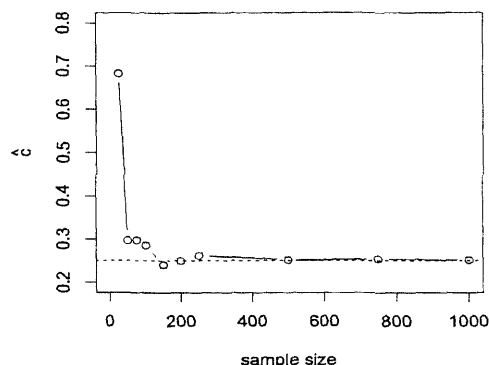


図 1: サンプルサイズに対する \hat{c}_n の挙動

4 今後の課題

制約付きの最尤推定量においてクロスバリデーションを用いた場合に, 一致性が成り立つかどうかは理論的に示されていない. (数値実験の結果を見ると, 一致性が成り立ちそうだと予想される.)

また, 今回の結果は, 罰則付き尤度においてクロスバリデーションを用いた場合の話に拡張できると考えられる.

参考文献

TANAKA, K., AND A. TAKEMURA (2003): "Strong consistency of MLE for finite mixtures of location-scale distributions when the scale parameters are exponentially small," University of Tokyo. Available at <http://www.keisu.t.u-tokyo.ac.jp/Research/techrep.0.html>.

非線形モデルにおける最深回帰推定量

大阪大学大学院基礎工学研究科 藤木 美江
大阪大学大学院基礎工学研究科 白旗 慎吾

1 はじめに

最深回帰推定量 (Deepest Regression Estimator : DRE) は, Aelst, S. V. et al. (2002) で提案された推定量である. これは回帰 depth から導き出され, 従来のロバスト回帰推定量よりも, 外れ値の影響を受けにくく, 効率の低下が小さいという性質をもっている. (Rousseeuw, P. J. and Aelst, S. V., 2000) また, 藤木・白旗 (2004) でシミュレーション実験を行った結果, 他のロバスト回帰推定量に比べ, DRE は平均 2 乗誤差の値が小さくなり, 外れ値を含む場合とデータ数が大きい場合に強い推定量であることが確認できた.

これらについては線形モデルを考えてきたわけだが, 回帰 depth の定義から, より一般的なモデルへの拡張が可能であることもわかっている. 非線形モデルにおける回帰 depth が単調不変性を持ち, 一般化線形モデルにおいて回帰 depth の定義が与えられ, 最深回帰推定量を求めることができる. また, 多項式回帰などへの適用もあるため, ここではその理論面や性質について明らかにするとともに, 実際どのように用いられるかを検証する.

2 回帰 depth と 最深回帰推定量の定義

回帰 depth と最深回帰推定量の定義を与える. データ集合 $Z_n = \{(x_{i1}, \dots, x_{i,p-1}, y_i) : i = 1, \dots, n\} \subset \mathbb{R}^p$ に対して, y_i を $\theta_1 x_{i1} + \dots + \theta_{p-1} x_{i,p-1} + \theta_p = (\mathbf{x}_i, 1)\boldsymbol{\theta}$, すなわち \mathbb{R}^p のアフィン超平面をあてはめる. $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)^t \in \mathbb{R}^p$ で, $\mathbf{x}_i = (x_{i1}, \dots, x_{i,p-1})^t \in \mathbb{R}^{p-1}$ とする. 残差は $r_i = r_i(\boldsymbol{\theta}) = y_i - (\theta_1 x_{i1} + \dots + \theta_{p-1} x_{i,p-1} + \theta_p)$ である. 定義 1 は \mathbf{x}_i は同点がある場合も有効であり, 分布に関する仮定はしない.

定義 1 データ集合 $Z_n \subset \mathbb{R}^p$ に対して, $\boldsymbol{\theta} \in \mathbb{R}^p$ の $rdepth(\boldsymbol{\theta}, Z_n)$ は,

$$rdepth(\boldsymbol{\theta}, Z_n) = \min_{\mathbf{u}, v} \{ \#(r_i(\boldsymbol{\theta}) \geq 0 \ \& \ \mathbf{x}_i^t \mathbf{u} < v) + \#(r_i(\boldsymbol{\theta}) \leq 0 \ \& \ \mathbf{x}_i^t \mathbf{u} > v) \}$$

である. ただし, $(\mathbf{x}_i^t, y_i) \in Z_n$ に対して, 最小値は $\mathbf{x}_i^t \mathbf{u} \neq v$ を満たす, すべての単位ベクトル $\mathbf{u} = (u_1, \dots, u_{p-1})^t \in \mathbb{R}^{p-1}$ と, $v \in \mathbb{R}$ でとられるものとする.

定義 2 p 次元における最深回帰推定量 $DR(Z_n)$ は, $rdepth(\boldsymbol{\theta}, Z_n)$ を最大にする $\boldsymbol{\theta}$ である.

$$DR(Z_n) = \arg \max_{\boldsymbol{\theta}} rdepth(\boldsymbol{\theta}, Z_n)$$

ただし, $rdepth(\boldsymbol{\theta}, Z_n)$ を最大にする $\boldsymbol{\theta}$ はただ一つとは限らない.

3 非線形モデルにおける回帰 depth の定義

定義 1 より, データ集合 $Z_n = \{(\mathbf{x}_i, y_i) : i = 1, \dots, n\}$ に対して, $\boldsymbol{\theta}$ の $rdepth(\boldsymbol{\theta}, Z_n)$ は \mathbf{x}_i と残差 $r_i(\boldsymbol{\theta})$ の符号に依存しているだけなので, より一般的なモデルに適用することができる. ある実関数

f に対して, $y = f(\mathbf{x}) = f(x_1, \dots, x_{p-1})$ を回帰 fit とし, 残差は $r_i(f) = y_i - f(x_{i1}, \dots, x_{i,p-1})$ とする. そこで f の回帰 depth を次のように定義する. また, 非線形モデルにおける回帰 depth は命題 1 で単調不変性をもつ.

定義 3 任意のデータ集合 $Z_n = \{(\mathbf{x}_i, y_i) : i = 1, \dots, n\}$ で, \mathbb{R}^{p-1} 上での任意の実関数 f に対して, f の $rdepth(f, Z_n)$ は,

$$rdepth(f, Z_n) = \min_{\mathbf{u}, v} \{ \#(r_i(f) \geq 0 \ \& \ \mathbf{x}_i^t \mathbf{u} < v) + \#(r_i(f) \leq 0 \ \& \ \mathbf{x}_i^t \mathbf{u} > v) \}$$

である. ただし, $(\mathbf{x}_i^t, y_i) \in Z_n$ に対して, 最小値は $\mathbf{x}_i^t \mathbf{u} \neq v$ を満たす, すべての単位ベクトル $\mathbf{u} = (u_1, \dots, u_{p-1})^t \in \mathbb{R}^{p-1}$ と, $v \in \mathbb{R}$ でとられるものとする.

命題 1 (Aelst, S.V. et al., 2002) データ集合 $Z_n = \{(\mathbf{x}_i, y_i) : i = 1, \dots, n\}$ で狭義の単調実関数を g とする. $y'_i = g(y_i)$ とおくと, データ集合 $Z'_n = \{(\mathbf{x}_i, y'_i) : i = 1, \dots, n\}$ とする. このとき任意の関数 f に対して, 次が成り立つ.

$$rdepth(f, Z_n) = rdepth(g(f), Z'_n)$$

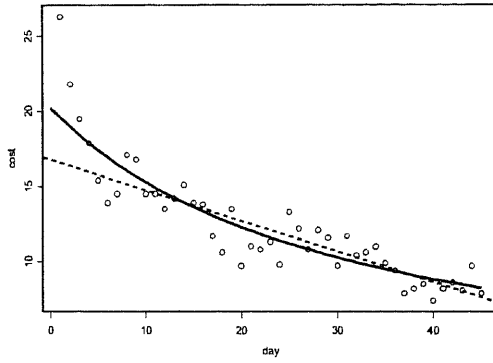


図 1: cost データ (一般化線形モデル)

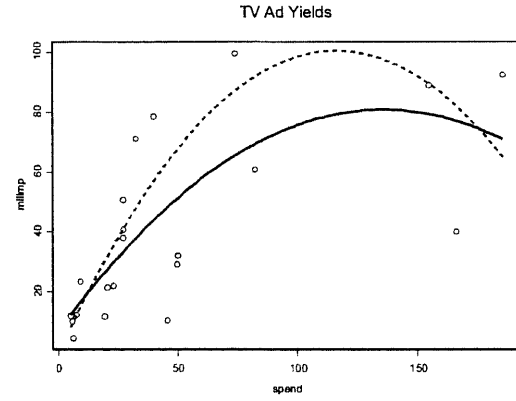


図 2: TV 広告のデータ (多項式回帰)

上の図は, <http://lib.stat.cmu.edu/DASL/>からのデータを分析した結果である. 図 1 は, 実線と点線のどちらも最深回帰推定量の結果であるが, 実線は逆変換をして求めたものである. 図 2 は多項式回帰を適用した. 実線は最小 2 乗推定量, 点線は最深回帰推定量の結果である.

参考文献

- [1] Aelst, S. V., Rousseeuw, P. J., Hubert, M., and Struyf, A. (2002). The Deepest Regression Method, *Journal of Multivariate Analysis*, 81, 138-166.
- [2] Rousseeuw, P. J. and Aelst, S. V. (2000). Robustness of Deepest Regression, *Journal of Multivariate Analysis*, 73, 82-106.
- [3] Rousseeuw, P. J. and Hubert, M. (1999). Regression Depth, *Journal of the American Statistical Association*, 94, 388-402.
- [4] 藤木美江・白旗慎吾, (2004). シミュレーションによる最深回帰推定量の性能評価, 日本統計学会第 72 回大会 講演報告集, pp.494-495.

Nonparametric inference for Lévy processes by continuous observation: a martingale approach

西山陽一（統計数理研究所）

1. 序.

本報告では、Nishiyama (2005) の結果を用いて、Lévy 過程のノンパラメトリック推測問題を考察する。 $t \rightsquigarrow Z_t$ はジャンプのみによって増減する 1 次元の時間的に一様でない Lévy 過程であるとする。その出発点は Z_0 であるとし、Lévy 測度を

$$L(dt, dz) = dt\alpha(t, dz)$$

と表す。ここで $\int_{\mathbb{R}} \alpha(t, dz)$ および $\int_{\mathbb{R}} |z|\alpha(t, dz)$ は無限大になりえるが、 $\int_{\mathbb{R}} (z^2 \wedge 1)\alpha(t, dz) < \infty$ であることに注意されたい。この事実に基づき、 \mathbb{R} 上の非負有界可測関数 w を導入して、

$$A(t, u) = \int_{[0, t] \times \mathbb{R}} w(z) 1_{(-\infty, u]}(z) ds \alpha(s, dz), \quad (t, u) \in [0, 1] \times \mathbb{R}$$

を推定することを考える。もしも Z が複合 Poisson 過程であれば、重み関数は $w \equiv 1$ とすべきである。しかしながら、我々の一般的な状況設定では、

$$w(z) = z^2 \wedge 1 \quad \text{or} \quad \frac{z^2}{1 + z^2}$$

によって定義される重み関数が自然であろう。もしも我々が何らかの理由で $\int_{\mathbb{R}} (|z| \wedge 1)\alpha(t, dz) < \infty$ であることを事前に知っていれば、重み関数は $w(z) = |z| \wedge 1$ とすべきである。無限個の和を計算するのは現実的ではないので、推定量の構成においては、我々は切り取られた重み関数

$$w_n(z) = w(z) 1_{(-\infty, -c_n] \cup [c_n, \infty)}(z)$$

を用いる。ここに $c_n \downarrow 0$ は正の定数列である。我々は $t \rightsquigarrow Z_t$ の連続な観測を得なければならないが、 $|\Delta Z_t| < c_n$ であるようなジャンプは推定量の構成には使用しないことを強調しておく。我々は $(t, u) \rightsquigarrow A(t, u)$ に対する推定量の $\ell^\infty([0, 1] \times \mathbb{R})$ -空間における漸近正規性および漸近有効性を証明する。ここに、 $\ell^\infty(T)$ は T の上で定義された有界関数の空間を表すものとし、それに一様距離を添付する。

2. 推定.

1 節で述べたような Lévy 過程が独立に n 個（すなわち $Z^k, k = 1, \dots, n$ ）観測される状況を考える。ただし各 Z_t^k はランダム時間区間 $[\sigma^k, \tau^k] \subset [0, 1]$ の上でのみ観測されるものとする。典型的な例としては、 $Z_0^k > 0$ であって $\sigma^k = 0$ かつ $\tau^k = \inf\{t \geq 0 : Z_t^k \leq 0\} \wedge 1$ であるようなものを想定している。あるいは、右センサーが入った $\tau^k = \inf\{t \geq 0 : Z_t^k \leq 0\} \wedge C^k \wedge 1$ を考えることもできる。しかしながら、我々の状況設定はより一般的であり、仮定するの

は後に述べる条件 (1) のみである。我々は整数値ランダム測度

$$N^k(\omega; dt, dz) = \sum_{t \in (\sigma^k(\omega), \tau^k(\omega)]} 1_{\{\Delta Z_t^k(\omega) \neq 0\}} \varepsilon_{(t, \Delta Z_t^k(\omega))}(dt, dz)$$

を導入する。ただし ε_a は点 a における Dirac 測度である。このとき、予測可能カンペンセイターは

$$1_{]]\sigma^k, \tau^k]]}(\omega, t) dt \alpha(t, dz)$$

によって与えられる。 k に関して和をとることにより、整数値ランダム測度

$$\mu^n(\omega; dt, dz) = \sum_{k=1}^n N^k(\omega; dt, dz)$$

と、その予測可能カンペンセイター

$$\nu^n(\omega; dt, dz) = Y_t^n(\omega) dt \alpha(t, dz)$$

を得る。ただし

$$Y_t^n(\omega) = \sum_{k=1}^n 1_{]]\sigma^k, \tau^k]]}(\omega, t)$$

である。これらに対し、Nishiyama (2005) の理論を適用する。我々は α に対する推定量として、Nelson-Aalen 推定量（を一般化したもの）を採用する；

$$(t, u) \rightsquigarrow \hat{A}^n(t, u) = \int_{[0, t] \times \mathbb{R}} w_n(z) 1_{(-\infty, u]}(z) Y_s^{n-} \mu^n(ds, dz).$$

ただし

$$Y_s^{n-} = \begin{cases} 1/Y_s^n, & \text{if } Y_s^n \geq 1, \\ 0, & \text{otherwise} \end{cases}$$

である。

Theorem 1 ある真に正な可測関数 $y(t)$ が存在して

$$(1) \quad \sup_{t \in [0, 1]} |nY_t^{n-} - 1/y(t)| \longrightarrow 0 \quad \text{in probability, as } n \rightarrow \infty$$

が成り立つことを仮定する。このとき、 $\sqrt{n}(\hat{A}^n - A)$ は $\ell^\infty([0, 1] \times \mathbb{R})$ -空間の中で、平均ゼロのガウス過程 G であって共分散

$$E(G(t_1, u_1)G(t_2, u_2)) = \int_{[0, t_1 \wedge t_2] \times (-\infty, u_1 \wedge u_2]} \frac{|w(z)|^2}{y(s)} ds \alpha(s, dz)$$

をもつものに弱収束する。

さらに、この推定量の ℓ^∞ -空間の中での漸近有効性を証明することもできる。

連続型閾値モデルの離散観測について

大阪大学 基礎工 藤井孝之

1 はじめに

本講演では拡散過程の定数ドリフト係数がある閾値により変化するシンプルな連続型閾値モデルにおいて、閾値パラメータに関する尤度比統計量の漸近的性質を考える。閾値過程は経済時系列の分野において盛んに研究されている非線形モデルのひとつであり、連続閾値過程についても近年いろいろな結果が報告されている。また、連続過程に対する離散観測も実データへの理論の応用を考えると重要であり、ここで連続閾値過程を離散近似した場合についても同等な問題を考え、その観測時間幅を十分小さくしたときの様子を調べる。さらに、離散的な観測から推定量を得るアルゴリズムも与える。

2 連続閾値過程

シンプルな連続閾値過程

$$dX(t) = -a \cdot \text{sign}(X(t) - \theta) dt + \sigma dW(t), \quad t \in [0, T] \quad (1)$$

において、閾値パラメータ θ の推定を考える。ここで、ドリフトおよび拡散係数 a, σ は既知とする。そのとき、Weiner 積分についての測度変換の定理から尤度関数は、

$$L_T(\theta) = \exp \left\{ -\frac{a}{\sigma^2} \int_0^T \text{sign}(X(t) - \theta) dX(t) - \frac{a^2}{2\sigma^2} T \right\}$$

と表現される。 $h \in \mathbf{R}$ に対して、対数尤度比過程

$$Y_T(h) = \log \frac{L_T(\theta + \frac{h}{T})}{L_T(\theta)} = \frac{2a}{\sigma^2} \int_0^T 1_{\{\theta \leq X(t) < \theta + \frac{h}{T}\}} dX(t)$$

で、 $T \rightarrow \infty$ とすれば、局所時間の性質等から非退化な正規確率過程

$$Y(h) = \frac{2a}{\sigma} \cdot W_* \left(\frac{a}{\sigma^2} h \right) - \frac{2a^2}{\sigma^2} \cdot \frac{a}{\sigma^2} |h| \quad (2)$$

に弱収束することが示される。ここで、 $W_*(\cdot)$ は両側 Wiener 過程をあらわす (Kutoyants(2000) 参照)。

3 連続閾値過程の離散観測

連続閾値過程を時間幅 Δ の点ごとに離散的に観測する。観測点の数を n とし、 $T = n\Delta$ を満たすものとする。また、 k 番目の観測時点を $t_k = k \cdot \Delta$ とあらわす。オイラー法により、閾値過程 (1) のパスが

$$X(t_k) - X(t_{k-1}) = \begin{cases} a\Delta + \sigma \cdot \Delta W, & \text{if } X(t_{k-1}) < \theta \\ -a\Delta + \sigma \cdot \Delta W, & \text{if } X(t_{k-1}) \geq \theta \end{cases} \quad (3)$$

と近似される. ここで, $\Delta W \sim N(0, \Delta)$ を満たす. 近似 (3) の推移密度

$$f(X(t_k) | X(t_{k-1})) \sim \begin{cases} N(X(t_{k-1}) + a\Delta, \sigma^2\Delta), & \text{if } X(t_{k-1}) < \theta \\ N(X(t_{k-1}) - a\Delta, \sigma^2\Delta), & \text{if } X(t_{k-1}) \geq \theta \end{cases}$$

より, θ に関する離散擬似尤度

$$L_n(\theta) \propto \exp \left(\frac{2a}{\sigma^2} \sum_{X(t_{k-1}) < \theta} (X(t_k) - X(t_{k-1})) \right) \quad (4)$$

を得る. 擬似尤度 (4) から θ についての最尤推定量を求めるアルゴリズムを考える. (4) は基準値 θ より下にある観測値に関する変動の総和なので, θ の関数として捕えやすくするため観測値を小さい順に並べ替え, 順序統計量 $X_{[1]}, X_{[2]}, \dots, X_{[n]}$ を構成する (図 2). そして, 順次その番号まで変動の和をとり (図 3), 和が最大になる $X_{[i]}$ を θ の推定量として取り出す. 講演では, 以上のアルゴリズムを用いて行ったシミュレーションの結果もあたえる.

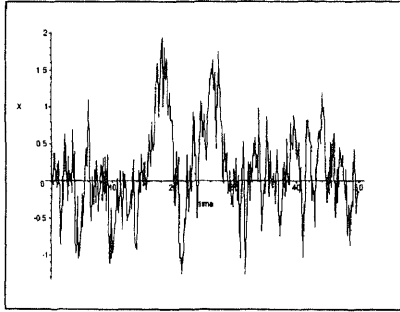


図 1: サンプルパス

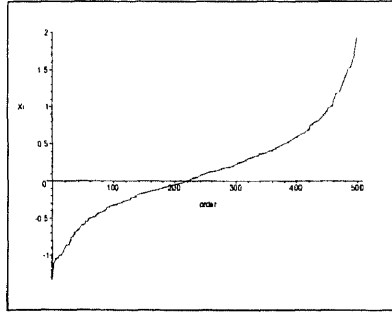


図 2: 順序統計量

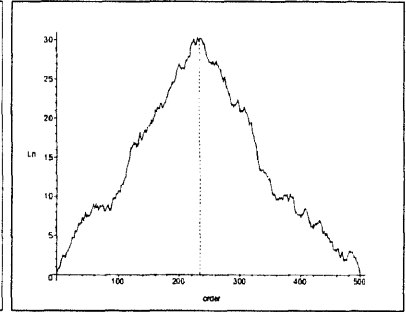


図 3: 擬似尤度

連続モデルと同様に, 対数尤度比過程を考える. 擬似尤度 (4) より, $h > 0$ に対して,

$$Y_n(h) = \log \frac{L_n(\theta + \frac{h}{n\Delta})}{L_n(\theta)} = \frac{2a}{\sigma^2} \sum_{X(t_{\ell-1}) \in A_n(h)} (X(t_{\ell}) - X(t_{\ell-1})).$$

また $A_n(h) = \{x | \theta \leq x < \theta + \frac{h}{n\Delta}\}$ とする. この過程は純粹跳躍過程であり, 定常性および跳躍とその跳躍サイズの独立性から, $n \rightarrow \infty$ としたとき, Y_n は強度 $\frac{a}{\sigma^2\Delta} \cdot h$, 跳躍サイズ分布 $\mu_\theta \sim N\left(-\frac{2a^2}{\sigma^2}\Delta, \frac{4a^2}{\sigma^2}\Delta\right)$ の複合 Poisson 過程 $Y^\Delta(h)$ に弱収束する.

$Y^\Delta(h)$ について, $\Delta \rightarrow 0$ とすれば, 連続観測における対数尤度比過程の極限 (2) に一致する. $h < 0$ の場合も同様なことがいえる.

参考文献

- Kutoyants, Yu. A.(2000), On parameter estimation for switching ergodic diffusion processes, *CRAS ser.1* **330**, 925-330
- Fujii, T(2005), Discrete-time observations for the continuous-time threshold model, preprint.

正規分布の平均に関する検定の重ね合わせによる多重決定方式

筑波大・数理物質科学 舞原 寛祐 (Hirosuke Maihara)

1 はじめに

仮説検定問題は受容か棄却かの2値の決定問題ととらえることができるが、臨界値の近傍に落ちた検定統計量の値の処理について問題がある。データに含まれる少数の外れ値の影響などによって、有意水準以下の確率で検定統計量の観測値が母数の真値とは反対側の棄却域に落ちてしまった場合、致命的な過誤が生じてしまう。仮説検定問題に決定論の枠組みを導入し、過誤の際のリスクの評価などを行うことができる (e.g. 舞原・赤平 [MA04])。特に医学統計などの分野ではそのような致命的な過誤は望まれないものであり、帰無仮説の棄却に対しては保守的である仮説検定よりもなお一層保守的な決定方式として多重決定方式がそのような分野で使われることがある。多重決定問題では、万が一真実からずれた決定を下してしまったとしても、その決定は本来下されるべき正しい決定と似通ったものになり、真実と全く正反対の決定を下してしまう過誤は生じない。また、1母数についての検定問題では、Neyman-Pearsonの基本定理などから一様最強力 (UMP) 検定や一様最強力不偏 (UMP) 検定など好ましい検定を構成することができるが、多母数についての同時の検定問題、すなわち複数の母集団の母数に関する同時の検定や比較の問題においては、UMP 検定ないしは UMPU 検定を構成することは容易ではなく、多母数に関する同時の決定問題においてある種の高い検出力をもつ決定方式として多重決定方式が用いられる。

竹内 [T73] は、1変量および2変量正規分布の平均の符号決め問題などについて多重決定問題を論じた。本論では、[T73] における設定および定理を再検討した上で、シミュレーションによって多重決定方式の挙動を調べるとともに、3変量正規分布の平均の符号決め問題に関する多重決定問題について考察する (舞原 [M05a])。

2 設定

$\mathcal{X} (\subseteq \mathbb{R}^k)$, $\Theta (\subseteq \mathbb{R}^p)$ をそれぞれ標本空間、母数空間とし、 $\mathcal{P} = \{P_\theta | \theta \in \Theta\}$ を \mathcal{X} のボレル集合族 $\mathcal{B}(\mathcal{X})$ 上に定義された確率分布族、 $\mathbf{X}_1, \dots, \mathbf{X}_n$ を \mathcal{P} に属する同一分布 P_θ に従う k 次元確率ベクトルとし、 $\mathbf{X} := (\mathbf{X}_1, \dots, \mathbf{X}_n)$ とおく。 Θ の被覆 $\mathcal{S}_e := \{\Omega_1, \dots, \Omega_m\}$ は部分被覆をもたず、 $\Omega_i \not\subseteq \Omega_j$ ($i \neq j$) を満たすとする。 \mathcal{S} を $\Omega_1, \dots, \Omega_m$ の和集合からなる Θ の部分集合族、すなわち $\mathcal{S} := \{\bigcup_{i \in I} \Omega_i | I \in \mathcal{M}\}$ とする。ただし、 $\{1\}, \dots, \{m\} \in \mathcal{M} \subseteq 2^{\{1, \dots, m\}}$, $\emptyset \notin \mathcal{M}$ とする。このとき、 \mathcal{S}_e の要素 $\Omega_1, \dots, \Omega_m$ を \mathcal{S} の第1次要素といい、 \mathcal{X}^n から \mathcal{S} への写像 φ を多重決定方式という。各 $\Omega \in \mathcal{S}$ に対して定数 α_Ω ($0 < \alpha_\Omega < 1/2$) が存在して $P_\theta \{\theta \in \varphi(\mathbf{X})\} \geq 1 - \alpha_\Omega$ ($\theta \in \Omega$) を満たすとき、 $\{1 - \alpha_\Omega | \Omega \in \mathcal{S}\}$ を φ の信頼系という。

次の定理から、仮説検定にもとづいて多重決定方式を構成することができる。この定理は仮説検定と信頼区間の関係を表わす、よく知られた定理の系にあたる。

定理 1 \mathcal{S} が \mathcal{S}_e の元の任意の和集合を含むとする。帰無仮説 $H_i: \theta \in \Omega_i$, 対立仮説 $K_i: \theta \notin \Omega_i$ に対する有意水準 α_i の検定の受容域を \mathcal{A}_i とし ($i = 1, \dots, m$), $\mathcal{X}^n = \bigcup_{i=1}^m \mathcal{A}_i$ であるとするとき、

$$\varphi(\mathbf{x}) := \bigcup_{\substack{1 \leq i \leq m \\ \mathbf{x} \in \mathcal{A}_i}} \Omega_i \quad (*)$$

は信頼系 $\{1 - \alpha_\Omega | \Omega \in \mathcal{S}\}$ をもつ多重決定方式となる。ただし、 $\alpha_\Omega := \sup_{\substack{1 \leq i \leq m \\ \Omega_i \subseteq \Omega}} \alpha_i$ とする。とくに、 \mathcal{S}_e の元が互いに素であるとき、逆に任意の多重決定方式 $\varphi(\mathbf{x})$ は (*) のように表現することができる。

第1次要素のあらゆる和集合を含む \mathcal{S} に対して定理 1 を適用して多重決定方式を構成した場合、多数の集合の和集合で表現される複雑な決定までも導かれてしまうので、決定空間 \mathcal{S} を縮小し、結果として得たい決定を限定して、簡易な多重決定方式を構成することもできる。

3 3つの母平均の符号の多重決定

次に、3つの母平均の符号についての多重決定問題を考える。 $\mathbf{X}_1, \dots, \mathbf{X}_n$ を3変量正規分布 $N_3(\mu, I_3)$ からの無作為標本とする。ただし、 $\mu := {}^t(\mu_1, \mu_2, \mu_3) \in \mathbb{R}^3$ とし、 I_3 は3次の単位行列とする。仮説

$$\begin{aligned} H_1: \mu_1 = 0, \mu_2 = 0, \mu_3 = 0, \\ H_2: \mu_1 = 0, \mu_2 = 0, \mu_3 > 0, \quad H_3: \mu_1 = 0, \mu_2 = 0, \mu_3 < 0, \quad H_4: \mu_1 = 0, \mu_2 > 0, \mu_3 = 0, \\ H_5: \mu_1 = 0, \mu_2 < 0, \mu_3 = 0, \quad H_6: \mu_1 > 0, \mu_2 = 0, \mu_3 = 0, \quad H_7: \mu_1 < 0, \mu_2 = 0, \mu_3 = 0, \end{aligned}$$

$$\begin{aligned}
H_8 : \mu_1 = 0, \mu_2 > 0, \mu_3 > 0, & H_9 : \mu_1 = 0, \mu_2 > 0, \mu_3 < 0, & H_{10} : \mu_1 = 0, \mu_2 < 0, \mu_3 > 0, \\
H_{11} : \mu_1 = 0, \mu_2 < 0, \mu_3 < 0, & H_{12} : \mu_1 > 0, \mu_2 = 0, \mu_3 > 0, & H_{13} : \mu_1 > 0, \mu_2 = 0, \mu_3 < 0, \\
H_{14} : \mu_1 < 0, \mu_2 = 0, \mu_3 > 0, & H_{15} : \mu_1 < 0, \mu_2 = 0, \mu_3 < 0, & H_{16} : \mu_1 > 0, \mu_2 > 0, \mu_3 = 0, \\
H_{17} : \mu_1 > 0, \mu_2 < 0, \mu_3 = 0, & H_{18} : \mu_1 < 0, \mu_2 > 0, \mu_3 = 0, & H_{19} : \mu_1 < 0, \mu_2 < 0, \mu_3 = 0, \\
H_{20} : \mu_1 > 0, \mu_2 > 0, \mu_3 > 0, & H_{21} : \mu_1 > 0, \mu_2 > 0, \mu_3 < 0, & H_{22} : \mu_1 > 0, \mu_2 < 0, \mu_3 > 0, \\
H_{23} : \mu_1 > 0, \mu_2 < 0, \mu_3 < 0, & H_{24} : \mu_1 < 0, \mu_2 > 0, \mu_3 > 0, & H_{25} : \mu_1 < 0, \mu_2 > 0, \mu_3 < 0, \\
H_{26} : \mu_1 < 0, \mu_2 < 0, \mu_3 > 0, & H_{27} : \mu_1 < 0, \mu_2 < 0, \mu_3 < 0,
\end{aligned}$$

それぞれに対して尤度比検定を構成し、定理 1 を適用して、それらの受容域から多重決定方式を構築する。これら 27 個の仮説は、原点 $\mu = \mathbf{0}$ に関する対称性から、 $\{H_1\}$, $\{H_2, \dots, H_7\}$, $\{H_8, \dots, H_{19}\}$, $\{H_{20}, \dots, H_{27}\}$ の 4 つの族に分けることができる。 H_1, H_2, H_8, H_{20} に対する受容域を求め、それらを原点中心に回転させれば、残りの仮説に対する受容域が求めることができる。 H_1, H_2, H_8, H_{20} に対する有意水準 α の尤度比検定の受容域はそれぞれ

$$\begin{aligned}
A_1 &:= \{\bar{\mathbf{x}} \mid n\{(\bar{x}_1)^2 + (\bar{x}_2)^2 + (\bar{x}_3)^2\} \leq c_1^2\}, \\
A_2 &:= \{\bar{\mathbf{x}} \mid n\{(\bar{x}_1)^2 + (\bar{x}_2)^2\} \leq c_2^2, \bar{x}_3 > 0\} \oplus \{\bar{\mathbf{x}} \mid n\{(\bar{x}_1)^2 + (\bar{x}_2)^2 + (\bar{x}_3)^2\} \leq c_2^2, \bar{x}_3 \leq 0\}, \\
A_8 &:= \{\bar{\mathbf{x}} \mid |\sqrt{n}\bar{x}_1| \leq c_8, \bar{x}_2 > 0, \bar{x}_3 > 0\} \oplus \{\bar{\mathbf{x}} \mid n\{(\bar{x}_1)^2 + (\bar{x}_3)^2\} \leq c_8^2, \bar{x}_2 > 0, \bar{x}_3 \leq 0\} \\
&\quad \oplus \{\bar{\mathbf{x}} \mid n\{(\bar{x}_1)^2 + (\bar{x}_2)^2\} \leq c_8^2, \bar{x}_2 \leq 0, \bar{x}_3 > 0\} \\
&\quad \oplus \{\bar{\mathbf{x}} \mid n\{(\bar{x}_1)^2 + (\bar{x}_2)^2 + (\bar{x}_3)^2\} \leq c_8^2, \bar{x}_2 \leq 0, \bar{x}_3 \leq 0\}, \\
A_{20} &:= \{\bar{\mathbf{x}} \mid \bar{x}_1 > 0, \bar{x}_2 > 0, \bar{x}_3 > 0\} \oplus \{\bar{\mathbf{x}} \mid \bar{x}_1 > 0, \bar{x}_2 > 0, -c_{20} \leq \sqrt{n}\bar{x}_3 \leq 0\} \\
&\quad \oplus \{\bar{\mathbf{x}} \mid \bar{x}_1 > 0, -c_{20} \leq \sqrt{n}\bar{x}_2 \leq 0, \bar{x}_3 > 0\} \oplus \{\bar{\mathbf{x}} \mid -c_{20} \leq \sqrt{n}\bar{x}_1 \leq 0, \bar{x}_2 > 0, \bar{x}_3 > 0\} \\
&\quad \oplus \{\bar{\mathbf{x}} \mid n\{(\bar{x}_2)^2 + (\bar{x}_3)^2\} \leq c_{20}^2, \bar{x}_1 > 0, \bar{x}_2 \leq 0, \bar{x}_3 \leq 0\} \\
&\quad \oplus \{\bar{\mathbf{x}} \mid n\{(\bar{x}_1)^2 + (\bar{x}_3)^2\} \leq c_{20}^2, \bar{x}_1 \leq 0, \bar{x}_2 > 0, \bar{x}_3 \leq 0\} \\
&\quad \oplus \{\bar{\mathbf{x}} \mid n\{(\bar{x}_1)^2 + (\bar{x}_2)^2\} \leq c_{20}^2, \bar{x}_1 \leq 0, \bar{x}_2 \leq 0, \bar{x}_3 > 0\} \\
&\quad \oplus \{\bar{\mathbf{x}} \mid n\{(\bar{x}_1)^2 + (\bar{x}_2)^2 + (\bar{x}_3)^2\} \leq c_{20}^2, \bar{x}_1 \leq 0, \bar{x}_2 \leq 0, \bar{x}_3 \leq 0\}
\end{aligned}$$

となる。ただし、 $\bar{\mathbf{x}} := {}^t(\bar{x}_1, \bar{x}_2, \bar{x}_3) = (1/n) \sum_{i=1}^n \mathbf{x}_i$ とし、 c_1, c_2, c_8, c_{20} はそれぞれ

$$\begin{aligned}
1 - \alpha &= 2\Phi(c_1) - 2c_1\phi(c_1) - 1 = \Phi(c_2) - \left(c_2 + \sqrt{\pi/2}\right)\phi(c_2) \\
&= \Phi(c_8) - (1/2)\left(c_8 + 2\sqrt{\pi/2}\right)\phi(c_8) = \Phi(c_{20}) - (1/4)\left(c_{20} + 3\sqrt{\pi/2}\right)\phi(c_{20})
\end{aligned}$$

を満たすただ 1 つの正の定数である。 $\alpha = 0.05$ の場合、 $c_1 \doteq 2.795$, $c_2 \doteq 2.654$, $c_8 \doteq 2.501$, $c_{20} \doteq 2.331$ となる。以上の 4 種類の受容域を原点に関して回転させることにより、残り 23 個の仮説それぞれに対する受容域も求めることができる。

定理 1 に基づいて、これら 27 個の受容域から多重決定方式を構成することができる。直方体型の受容域から定理 1 に基づいて同様に多重決定方式を構成するなど、いくつかの手法で多重決定方式を構成し、シミュレーションなどによってそれらの挙動の比較を行う。また、仮説検定方式の検出力の定義を拡張した形で多重決定方式の“検出力”を定義し、尤度比検定の受容域、直方体型の受容域それぞれに基づく多重決定方式の検出力の比較を行う。

参考文献

- [H03] 広津千尋 (2003). 多重比較法と多重決定方式. 「統計学の基礎 II – 統計学の基礎概念を見直す」(甘利俊一 他 編), 55–112, 岩波書店.
- [M05a] 舞原寛祐 (2005). 正規分布の平均の符号に関する多重決定問題. 京都大学 数理解析研究所講究録, **1439**, 151–184.
- [M05b] MAIHARA, H. (2005). Multiple decision problem on signatures of mean vector in three-variate normal case. In revision in *J. Japan Statist. Soc.*.
- [MA04] MAIHARA, H. AND AKAHIRA, M. (2004). Weighted loss functions for set estimation and testing hypotheses. *J. Japan Statist. Soc.*, **34** (2), 189–206.
- [T73] 竹内啓 (1973). 数理統計学的方法的基礎. 東洋経済新報社.

順序制約のもとでの予測量の改良

長春大学・理 肖 玉山
熊本大学・工 高田 佳和

1 序

統計的推測問題において、未知母数に関する事前情報やその問題の数学的構造によって、未知母数に関して順序が仮定できる場合がある。順序制約を用いることにより、通常の推定方法(順序制約を用いない)を改良できることがある。ここでは、未知母数に関する順序制約のもとでの予測量(順序制約を用いない)の改良について述べる。

X と Y は確率変数で、その同時分布は未知母数 ξ に依存するとする。 X の値をもとに Y の値を予測したい。このとき、 X 以外で確率変数 Z の値も利用可能とする。ただし、 Z の確率分布は未知母数 η に依存し、 $\eta \geq \xi$ とする。 Z の値をどのように利用すれば、 X の値だけによる予測量を改良できるか、位置分布族、尺度分布族の場合について述べる。

主要な結果を得るために、 TP_2 (Total Positivity of order 2) 関数が基本的な役割を演じる。関数 $K(x, y)$ が TP_2 であるとは

$$\begin{vmatrix} K(x_1, y_1) & K(x_1, y_2) \\ K(x_2, y_1) & K(x_2, y_2) \end{vmatrix} \geq 0$$

がすべての $x_1 < x_2$ 、 $y_1 < y_2$ に対して成立するときを言う。

2 位置分布族

(X, Y, Z) の同時確率密度関数は

$$f(x - \xi, y - \xi, z - \eta)$$

とする。ここで、 f の関数型は既知、 ξ と η は未知母数で、 $\xi \leq \eta$ とする。 X だけによる予測量

$$\delta_c = X + c$$

を、 Z も用いた予測量

$$\delta_\phi = X + \phi(Z - X)$$

によって改良する方法について述べる。ここで、 c は定数、 ϕ は関数とする。 $Y = y$ の値を d によって予測するときの損失を

$$L(y - d)$$

とする。ただし、 $L(t)$ は $t < 0$ のとき狭義の単調減少関数、 $t > 0$ のとき狭義の単調増加関数で、 $L(0) = 0$ とする。 $U = Z - X$ 、 $V = Y - X$ とおくと、 δ_c と δ_ϕ のリスク関数は

$$R(\theta, \delta_c) = E_\theta\{L(V - c)\}, \quad R(\theta, \delta_\phi) = E_\theta\{L(V - \phi(U))\}$$

と表せれる。又、 U と V の同時確率密度関数は、 $g(u - \lambda, v)$ となる。ここで、 $\lambda = \eta - \xi \geq 0$ であり

$$g(u, v) = \int_{-\infty}^{+\infty} f(t, t + v, t + u) dt$$

このとき

$$G(u, v) = \int_0^{+\infty} g(u - t, v) dt$$

と定義すると、次の結果を得た。

定理 1

(i) $G(u, v)$ は TP_2 関数

(ii) $\phi(u)$ は非減少関数、 $\lim_{u \rightarrow \infty} \phi(u) = c$

(iii) 任意の u に対して、 $\int_{-\infty}^{+\infty} L'(v - \phi(u))G(u, v)dv \leq 0$

を仮定すると

$$R(\theta, \delta_\phi) \leq R(\theta, \delta_c)$$

3 尺度分布族

(X, Y, Z) の同時確率密度関数を

$$\xi^{-2}\eta^{-1}f(x/\xi, y/\xi, z/\eta), \quad x > 0, y > 0, z > 0$$

とする。ここで、 f の関数型は既知、 ξ, η は未知母数で、 $0 < \xi \leq \eta$ とする。 X だけの予測量

$$\delta_c = cX$$

を Z も用いた予測量

$$\delta_\phi = \phi(Z/X)X$$

で改良する方法について述べる。ここで、 c は正の定数、 ϕ は正の関数。 $Y = y$ の値を d で予測するときの損失を

$$L(d/y)$$

とする。ただし、関数 $L(t)$ ($t > 0$) は $t < 1$ のとき、狭義の単調減少関数、 $t > 1$ のとき狭義の単調増加関数で、 $L(1) = 0$ とする。 $U = Z/X, V = Y/X$ とおくと、 δ_c と δ_ϕ のリスク関数は

$$R(\theta, \delta_c) = E_\theta\{L(c/V)\}, \quad R(\theta, \delta_\phi) = E_\theta\{L(\phi(U)/V)\}$$

と表される。又、 U と V の同時確率密度関数は $\lambda^{-1}g(u/\lambda, v)$ となる。ここで、 $\lambda = \eta/\xi \geq 1$ であり

$$g(u, v) = \int_0^\infty t^2 f(t, tv, tu) dt$$

このとき

$$G(u, v) = u \int_1^\infty t^{-2} g(u/t, v) dt$$

と定義するとき、次の結果を得た。

定理 2

(i) $G(u, v)$ は TP_2 関数

(ii) $\phi(u)$ は非減少関数、 $\lim_{u \rightarrow \infty} \phi(u) = c$

(iii) 任意の u に対して、 $\int_0^\infty v^{-1} L'(\phi(u)/v) G(u, v) dv \geq 0$

を仮定すると

$$R(\theta, \delta_\phi) \leq R(\theta, \delta_c)$$

外れ値の割合が多い場合にも バイアスが小さいロバスト推定

藤澤 洋徳 (統計数理研究所)

過去のロバスト推定は、外れ値の割合が多い場合には、外れ値に引きずられて、無視できないバイアスを生じてきた。例えば、メディアンは典型的で、外れ値が片側に 20% も存在すれば、完全にバイアスを生じる。本報告では、外れ値の割合が多い場合でさえも、バイアスが小さいロバスト推定が可能である、ということを議論する。

まずは次の相互エントロピーを用意しておこう：

$$d_{\gamma}(g, f) = -\frac{1}{\gamma} \log \int g(x)f(x)^{\gamma} dx + \frac{1}{1+\gamma} \log \int f(x)^{1+\gamma} dx \quad (\gamma > 0).$$

これに誘導されるダイバージェンスは $D_{\gamma}(g, f) = d_{\gamma}(g, f) - d_{\gamma}(g, g)$ で定義される。経験密度関数を \bar{g} で表すことにして、ロバスト推定量を次で定義することにしよう：

$$\hat{\theta}_{\gamma} = \arg \min_{\theta} d_{\gamma}(\bar{g}, f_{\theta}).$$

このようにして作られるロバスト推定量は、外れ値の割合が多くてもバイアスを生じにくい。その理由を以下で概観する。

まずは問題の設定を整理することにしよう。目的とする密度関数を $f(x)$ で表し、外れ値の密度関数を $\delta(x)$ で表す。例えば、外れ値の密度関数としては、外れ値 $x = x^*$ でのディラック関数 $\delta_{x^*}(x)$ が典型的である。ここで注意すべきは、外れ値のもつ本質的な性質は $f(x^*)$ が十分に小さいということである。データを生成する密度関数は、外れ値の割合が ε であるとする、次の形で表現できる：

$$g(x) = (1 - \varepsilon)f(x) + \varepsilon\delta(x).$$

パラメトリック密度関数を $f_{\theta}(x)$ とする。パラメトリック密度関数 $f_{\theta}(x)$ によって目的密度関数 $f(x)$ を当てようとするわけである。以下では簡単のために目的密度関数はパラメトリック密度関数のクラスに入っているとしよう： $f(x) = f_{\theta^*}(x)$ 。このときは目的パラメータは θ^* になる。

もう少し記号を整理することにする。ロバスト推定量の収束先と目的パラメータは次のように表現できる：

$$\theta_{\gamma}^* = \arg \min_{\theta} d_{\gamma}(g, f_{\theta}), \quad \theta^* = \arg \min_{\theta} d_{\gamma}(f, f_{\theta}).$$

この差が潜在的なバイアスである。この潜在的なバイアスが小さいほど、推定量は外れ値に対してロバストであると言える。提案したロバスト推定は、表面的には $d_\gamma(g, f_\theta)$ を小さくしようとしている、言い換えると、 $D_\gamma(g, f_\theta)$ を小さくしようとしている。本来は、 $D_\gamma(f, f_\theta)$ を小さくしたい、言い換えると、 $d_\gamma(f, f_\theta)$ を小さくしたいので、そこにギャップが生じている。提案したロバスト推定は、用意された相互エントロピーのもつ性質によって、このようなギャップを自然に小さくするのである。

ここで次の仮定を設けよう：

$$(*) \quad \nu_f = \left\{ \int \delta(x) f(x)^{\gamma_0} dx \right\}^{1/\gamma_0} \text{ is sufficiently small} \\ \text{for an appropriately large } \gamma_0 > 0.$$

もし外れ値の密度関数が $\delta(x) = \delta_{x^*}(x)$ ならば、 $\nu_f = f(x^*)$ となり、簡単に理解できる仮定となる。そうでなくても、外れ値の密度関数 $\delta(x)$ が目的密度関数 $f(x)$ の確率が低い部分に存在しているというイメージが可能である。本報告では、本質的には上記のタイプの仮定だけを設定し、外れ値の割合が十分に小さいという仮定は設定しない。また $\gamma < \gamma_0$ であるとする。

さらに、次のようなパラメータの存在領域 Ω_{ν^*} が定義されるとしよう：
 $\Omega_{\nu^*} = \{\theta : \nu_{f_\theta} < \nu^*\} \ni \theta_\gamma^*, \theta^*$. ただし ν^* は十分に小さいとする。以下ではパラメータ θ はこのパラメータ領域にあるとする。このとき、パラメトリック密度関数 $f_\theta(x)$ は、目的密度関数 $f(x)$ と同じように、その確率の低い部分に外れ値の密度関数 $\delta(x)$ をおいている、と考えることができる。

上述した二つの仮定の下で次が成立する：

$$\begin{aligned} d_\gamma(g, f_\theta) &= d_\gamma(f, f_\theta) - (1/\gamma) \log(1 - \varepsilon) + O(\varepsilon \nu^\gamma). \\ \Delta(g, f, f_\theta) &= D_\gamma(g, f_\theta) - D_\gamma(g, f) - D_\gamma(f, f_\theta) = O(\varepsilon \nu^\gamma). \end{aligned}$$

ただし $\nu = \max\{\nu_f, \nu^*\}$ である。最初の関係式は、 $\theta_\gamma^* = \theta^* + O(\varepsilon \nu^\gamma)$ を導くので、潜在的なバイアスは十分に小さいと想定され、提案したロバスト推定がきちんと働くことを示唆している。次の関係式は、近似的なピタゴリアン関係を示していて、構成したロバスト推定がある種の射影になっていることを表している。

では逆に、上述したような好ましい性質を提示する相互エントロピーは、他に存在するのだろうか。適当な仮定の下では、そのような相互エントロピーは、本質的に唯一であり、それが $d_\gamma(g, f)$ であることを証明できる。ダイバージェンスについても適当な唯一性を示すことができる。

多変量母集団の平均ベクトルに関する DL 法とボンフェロニ型不等式による同時信頼区間

北海道大学大学院経済学研究科 柿沢 佳秀

1. はじめに

各 $a = 1, \dots, q$ について $\mathbf{X}_1^{(a)}, \dots, \mathbf{X}_{N_a}^{(a)}$ ($N_a > p$) を独立同一な p 次元確率ベクトルとし, $E(\mathbf{X}_t^{(a)}) = \boldsymbol{\theta}^{(a)}$, $\text{Var}(\mathbf{X}_t^{(a)}) = \Sigma^{(a)}$ とする. 必要に応じて高次モーメントの存在も仮定する. 標本平均ベクトルと標本分散共分散行列を $\bar{\mathbf{X}}^{(a)} = N_a^{-1} \sum_{t=1}^{N_a} \mathbf{X}_t^{(a)}$, $S_X^{(a)} = (N_a - 1)^{-1} \sum_{t=1}^{N_a} (\mathbf{X}_t^{(a)} - \bar{\mathbf{X}}^{(a)})(\mathbf{X}_t^{(a)} - \bar{\mathbf{X}}^{(a)})'$ で定義する. また, 全ての標本数を $N = \sum_{a=1}^q N_a$ と書くとき, $S_{\text{pool}, X} = (N - q)^{-1} \sum_{a=1}^q (N_a - 1) S_X^{(a)}$ は等分散 $\Sigma^{(1)} = \dots = \Sigma^{(q)}$ を仮定する場合の共通な分散 Σ に対する不偏推定量となる. なお, 1 元配置モデル $\mathbf{X}_t^{(a)} = \boldsymbol{\theta}^{(a)} + \mathbf{U}_t^{(a)}$, $a = 1, \dots, q$; $t = 1, \dots, N_a$ で書き直した場合, 誤差項 $\mathbf{U}_t^{(a)}$ に対して同じ記号 $\bar{\mathbf{U}}^{(a)}$, $S_U^{(a)}$, $S_{\text{pool}, U}$ を採用する.

本報告では (I) 対照比較 $\ell'(\boldsymbol{\theta}^{(a)} - \boldsymbol{\theta}^{(q)})$, $\ell \in \mathbf{R}^p - \{\mathbf{0}\}$, $a = 1, \dots, q-1$ と (II) 対比較 $\ell'(\boldsymbol{\theta}^{(a)} - \boldsymbol{\theta}^{(b)})$, $\ell \in \mathbf{R}^p - \{\mathbf{0}\}$, $a, b = 1, \dots, q$; $a < b$ の同時信頼区間を再検討する. 以下, a, b を省略することもあるが混乱しないだろう.

1.1. 等分散 $\Sigma^{(1)} = \dots = \Sigma^{(q)}$ の場合

Roy and Bose (1953; AMS) と Siotani (1960; AISM) は等分散の仮定の下で

$$(I) \quad \ell'(\bar{\mathbf{X}}^{(a)} - \bar{\mathbf{X}}^{(q)}) \pm \{T_{\max, C}^2(\alpha) N_{aq}^{-1} (\ell' S_{\text{pool}, X} \ell)\}^{1/2}, \quad \ell \in \mathbf{R}^p - \{\mathbf{0}\}$$

$$(II) \quad \ell'(\bar{\mathbf{X}}^{(a)} - \bar{\mathbf{X}}^{(b)}) \pm \{T_{\max}^2(\alpha) N_{ab}^{-1} (\ell' S_{\text{pool}, X} \ell)\}^{1/2}, \quad \ell \in \mathbf{R}^p - \{\mathbf{0}\}$$

なる同時信頼区間を与えた. ここに, $N_{ab} = N_a N_b / (N_a + N_b)$ とし, $T_{\max, C}^2(\alpha)$ と $T_{\max}^2(\alpha)$ はそれぞれホテリングの 2 標本版の T^2 統計量

$$T_{ab}^2 = N_{ab} (\bar{\mathbf{U}}^{(a)} - \bar{\mathbf{U}}^{(b)})' S_{\text{pool}, U}^{-1} (\bar{\mathbf{U}}^{(a)} - \bar{\mathbf{U}}^{(b)}), \quad a, b = 1, \dots, q; a < b$$

の最大値 $T_{\max, C}^2 = \max_{a=1, \dots, q-1} (T_{aq}^2)$, $T_{\max}^2 = \max_{1 \leq a < b \leq q} (T_{ab}^2)$ の上側 α 点である. しかし最大値分布論は例外を除けば断念せざるをえないため, 通常はボンフェロニ不等式に基づいて

$$P(T_{\max, C}^2 \leq x) \geq 1 - \sum_{a=1}^{q-1} P(T_{aq}^2 > x) = 1 - \alpha$$

および

$$P(T_{\max}^2 \leq x) \geq 1 - \sum_{1 \leq a < b \leq q} P(T_{ab}^2 > x) = 1 - \alpha$$

を満たす解を臨界点 $T_{\max, C}^2(\alpha)$ と $T_{\max}^2(\alpha)$ の代わりに用いた保守的な同時信頼区間を構築する. ここで確率 $P(T_{ab}^2 > x)$ は正規母集団ならば F 分布で数値計算できるが, クラメル条件を満たす一般分布ならば漸近展開 (Seo (2002; JJSS), Okamoto and Seo (2004; JJSCS) は楕円分布を仮定していたから, 報告者は一般分布へ拡張した) で近似計算する. 漸近展開を使う場合のボンフェロニ法の保守性は漸近展開の誤差を無視した意味とするが, ボンフェロニ不等式と漸近展開の組み合わせの問題がある.

1.2. 等分散を仮定しない場合

言うまでもなく多変量ベールン・フィッシャー問題に関連づけて同時信頼区間を構成する必要がある、1つの候補として次の形が自然に思いつく。

$$\ell'(\bar{\mathbf{X}}^{(a)} - \bar{\mathbf{X}}^{(b)}) \pm \left\{ A^2(\alpha) \ell' \left(\frac{S_X^{(a)}}{N_a} + \frac{S_X^{(b)}}{N_b} \right) \ell \right\}^{1/2}, \quad \ell \in \mathbf{R}^p - \{\mathbf{0}\}.$$

ここに、臨界点 $A^2(\alpha)$ は Siotani (1960; AISM) の導出に従えばベールン・フィッシャー型統計量

$$T_{BF,ab}^2 = (\bar{\mathbf{U}}^{(a)} - \bar{\mathbf{U}}^{(b)})' \left(\frac{S_U^{(a)}}{N_a} + \frac{S_U^{(b)}}{N_b} \right)^{-1} (\bar{\mathbf{U}}^{(a)} - \bar{\mathbf{U}}^{(b)}), \quad a, b = 1, \dots, q; a < b$$

の最大値から決める. (I) の場合は $T_{BF,\max,C}^2 = \max_{a=1,\dots,q-1} (T_{BF,aq}^2)$ の上側 α 点 $T_{BF,\max,C}^2(\alpha)$, (II) の場合は $T_{BF,\max}^2 = \max_{1 \leq a < b \leq q} (T_{BF,ab}^2)$ の上側 α 点 $T_{BF,\max}^2(\alpha)$ であるが、これらを求めることは断念せざるをえない. 正規母集団の場合に個々の $T_{BF,ab}^2$ の分布を求めることができるが、1.1 節のときの F 分布と比べ極めて複雑であることに注意する (Nel et.al. (1990; Commun.Stat.)). 一方、 $T_{BF,ab}^2$ の分布の漸近展開はクラメル条件を満たす一般分布の下で柿沢・岩下 (1998; 日本統計学会講演) でアナウンスされており、ボンフェロニ不等式を利用すれば (漸近展開の誤差を無視した) 保守的な同時信頼区間を与えるような $\sum_{a=1}^{q-1} P(T_{BF,aq}^2 > x) = \alpha$, $\sum_{1 \leq a < b \leq q} P(T_{BF,ab}^2 > x) = \alpha$ の解 $A_I^2(\alpha)$, $A_{II}^2(\alpha)$ について漸近展開が得られた (柿沢 (2005; 日本統計学会講演)).

2. 等分散の場合のボンフェロニ型不等式の応用と Siotani の修正 2 次近似法

ボンフェロニ不等式の改良 (それをボンフェロニ型不等式と呼ぶことにする) はグラフ理論と密接な関係があり近年も進展しているようであるが、少なくとも完全グラフの全域木で定式化されている Hunter (1976; J.Appl.Prob.)-Worsley (1982; Biometrika) の不等式を応用した 1 次元 ($p = 1$) での多重比較法の先行研究は沢山ある. 多変量ではボンフェロニ型不等式と異なる手法 (Siotani (1959,1960; AISM) の修正 2 次近似法) が Seo and Siotani (1992; JJSS) で数値検討されており、Seo (2002; JJSS) 等による一連の研究で楕円母集団でも研究がなされている. ボンフェロニ型不等式に基づく手法は保守性をもつが、Siotani のオリジナルな修正 2 次近似法の理論的な性質は未解決であることに注意しよう. 本報告の 1 つの目的として、等分散な多変量母集団で平均ベクトルに関する保守性をもつ改善された同時信頼区間をボンフェロニ型不等式から考察し、Siotani (1959,1960; AISM) と異なる修正 2 次近似法も与えた.

3. DL 法の応用

ボンフェロニ型不等式 (ボンフェロニ不等式やその他の類似の確率不等式も含む) と漸近展開の組み合わせの問題がある. 漸近展開の誤差のため理論的な保守性が悪化する、あるいは、偶然に改善されることもあり、保守性が崩れる可能性も否定できない. 非正規母集団を考えれば分布論は漸近論 (漸近展開) になるだろうから、確率不等式の使用を回避できる手法 (例えば、ヘルダー不等式をある種の線形結合の母数の同時信頼区間に応用した DL 法 (Dalal (1978; Biometrika), Hochberg and Tamhane (1987; Wiley)) は 1 つの対抗馬となりうる. DL 法は 1.1 節と 1.2 節のいずれの状況に対しても応用可能で、報告者は付随する 4 つの統計量の分布の漸近展開も導出した.

同等性の検定と信頼区間について

岩佐 学 (熊本大学大学院自然科学研究科)

$m > 0$ を既知の定数とすると、確率変数

$$Y \sim N(\theta, m^2 \sigma^2), s^2 \sim \sigma^2 \chi_n^2 \quad (1)$$

に基づいて

$$H_0: |\theta| \geq \Delta \text{ vs. } H_1: |\theta| < \Delta \quad (2)$$

を検定する一標本問題を考える. この問題はある種の生物学的同等性の統計的仮説検定問題と同等であり, 多く議論されている問題である. その中心的な話題の一つが, 「直観的には最適と思われる二片側 t 検定が大きなバイアスを持ち検出力の意味で最適でない」という事実に関するもので, 不偏検定の構成法など多くの研究が成されてきた.

本講演における我々の関心は, θ の信頼区間に基づいて (2) に対する検定を構成することにある. 信頼区間からは「 θ の信頼区間が同等性を認める限度 $[-\Delta, \Delta]$ に含まれるとき, 同等性を主張 (帰無仮説を棄却) する」という自然な規則を使うことで検定の棄却域が構成できる.

通常使われる θ の信頼区間 (信頼度 $1 - 2\alpha$) は, 自由度 n の t 統計量 $T = \frac{Y - \theta}{ms}$ を使って,

$$-t(\alpha, n) < T < t(\alpha, n), \text{ すなわち } Y - mst(\alpha, n) < \theta < Y + mst(\alpha, n)$$

と表される. ただし, $t(\alpha, n)$ は自由度 n の t 分布の上側 $\alpha\%$ 確率点である. この信頼区間から構成される検定の棄却域はサイズ α の二片側 t 検定に一致する. つまり, あまり検出力の高い検定は構成できない. そこで, Westlake(1976) は原点に関して対称な信頼区間を同等性検定の構成に用いることを発案し, 次のような信頼区間を提案した.

$$Y - k_1 ms < \theta < Y - k_2 ms$$

ただし, k_1, k_2 は

$$2Y = (k_1 + k_2)ms, F_t(k_2) - F_t(k_1) = 1 - 2\alpha$$

を満たす Y, s の関数となる. (F_t は自由度 n の t 分布の分布関数) この信頼区間による検定の棄却域は

$$R_2 = \{(Y, s) \mid Y - k_1 ms > -\Delta \text{ かつ } Y - k_2 ms < \Delta\}$$

となり, 確かにサイズ α の二片側 t 検定より検出力は高くなるが, 実は R_2 の検定のサイズは 2α となることがわかる. 棄却域の形状も不偏検定とはほど遠いものであり, 決して優れた検定方法を構成できているとは到底言えない.

そこで我々は仮説検定の受容域と信頼域の関係を利用した, 新たな対称信頼区間の構成法を提案する. いま, 任意の θ_0 に対して次のような帰無仮説, 対立仮説を考える:

$$H_{00}: \theta = \theta_0 \text{ vs. } H_{aa}: \theta \neq \theta_0. \quad (3)$$

帰無仮説の下で次の t 統計量が得られる:

$$T = \frac{Y - \theta_0}{ms}.$$

T は自由度 n の t 分布に従うので, (3) に対するサイズ 2α の棄却域として次の $R(\theta_0)$ をとる:

$$R(\theta_0) = \left\{ (Y, s) \left| g_2 \leq \frac{Y - \theta_0}{ms} \leq g_1 \right. \right\},$$

ただし g_1, g_2 は $R(\theta_0)$ が Y について原点对称になるよう以下を満たすものとする:

$$g_1 + g_2 = -\frac{2\theta_0}{ms}.$$

さらに, 棄却域 $R(\theta_0)$ の余事象を考えることにより受容域 $A(\theta_0)$ を得る:

$$A(\theta_0) = \left\{ (Y, s) \left| \frac{Y - \theta_0}{ms} < g_2 \text{ かつ } \frac{Y - \theta_0}{ms} > g_1 \right. \right\}.$$

$A = \bigcup_{\theta_0} A(\theta_0)$ を反転させることで θ の信頼区間を構成することが出来る. ただし, g_1, g_2 は θ_0 の関数になっているので反転している式を明示するのは容易ではないが, 反転させたものが θ について原点对称な区間になることは示される.

また, この信頼区間は

$$-t\left(\frac{1}{2} - \alpha, n\right) < \frac{Y}{ms} < t\left(\frac{1}{2} - \alpha, n\right) \quad (4)$$

のとき, 空集合になる. 空集合は $[-\Delta, \Delta]$ に含まれると考える. したがって, (4) が成り立つときは必ず同等性が示されることになる.

信頼区間を明示する式は与えられないが, 同等性の仮説 (2) に対する検定を構成することはでき, その棄却域 R_3 は $R(\Delta)$ と一致することがわかる. そこで

$$R_3 = \{(Y, s) \mid \Delta + g_2^* ms < Y < \Delta + g_1^* ms\}$$

とおく. ただし g_1^*, g_2^* は次の 2 式を満たす:

$$F_t(g_1^*) - F_t(g_2^*) = 2\alpha, \quad (g_1^* + g_2^*)ms = -2\Delta.$$

定理 1 棄却域 R_3 は次のように表現できる.

$$R_3 = \left\{ (Y, s) \left| F_t\left(\frac{-Y - \Delta}{ms}\right) - F_t\left(\frac{Y - \Delta}{ms}\right) < 2\alpha \right. \right\}.$$

定理 1 より, R_3 は Anderson and Hauck(1983) によって提案された (有意水準 2α の) 検定と同一であることがわかる. 彼らは別のアプローチよりこの検定を導出している. この検定の棄却域は非有界であり, 二片側 t 検定の改良として提案されている不偏検定と類似した性質を持っている.

この検定のサイズ, 検出力に関しては Frick(1987) や Müller-Cohrs (1990) などの研究があるが, 数値的な結果が多い. 実際の検定のサイズは 2α より大きく (数 10% 超過) なることが数値計算により示されているが, その値は確定されていない. 前出の 2 つの検定はともに $\sigma^2 \rightarrow \infty$ のとき検出力が 0 に近づくことが欠点としてあったわけであるが, 次の定理よりこの検定ではその欠点が解消されていることがわかる.

定理 2 $\theta = \Delta$ のとき,

- (1) $\lim_{\sigma \rightarrow 0} P((Y, s) \in R_3) = 2\alpha.$
- (2) $\lim_{\sigma \rightarrow \infty} P((Y, s) \in R_3) = 2\alpha.$

本研究により, 不偏検定に類似した検出力の高い検定を構成することが信頼区間を用いた方法でも可能だということが確かめられた. 信頼区間の信頼度と検定のサイズの関係などまだ解決しなければならない問題も多いが, 同等性の検定における不偏検定の構成法の新たな展開を与える興味深い成果である.