（7）「統計的推測理論とその応用」に関する研究報告

On the admissibility for the MLE under squared error loss in certain discrete distribution model

関東学院大経済　布能英一郎

Summary For estimation problems, an interesting question is whether the maximum likelihood estimator(MLE) is admissible or not under squared error loss. It is somewhat surprising that in many situations the admissibility of the MLE is still an open question. The aim of this research is to study the admissibility of the MLE for certain discrete probability problems. Several years ago, the author of this paper proved the theorem(Theorem 1) which showed the admissibility of the MLE under squared error loss when $P(x\,|\,\theta) \propto (1-\theta)^{a(x)}\theta^{b(x)}$ where $0 \le \theta \le 1$. This paper presents an extended theorem of the above result.

1. Preparation Throughout of this paper, we assume the following:
(1) Sample space $\boldsymbol{\mathcal{X}}$ is countable. (2) Parameter space Θ is $[0,1]$. (3) Loss function $L(\theta, \delta(x))$ is squared error. (4) For each $x \in \boldsymbol{\mathcal{X}}$, there exists at least one $\theta \in \Theta$ such that $P(x\,|\,\theta) > 0$.

Theorem 1 Assume the discrete probability distribution is given by $P(x \mid \theta) = c(x)(1-\theta)^{a(x)}\theta^{b(x)}$, $\theta \in \Theta = [0,1]$ where non-negative integers $a(x)$, $b(x)$ satisfy (1) There exists $x_* \in \boldsymbol{\mathcal{X}}$ that satisfies $a(x_*) = 0,\ b(x_*) > 0$ *and* $c(x_*) = 1$, (2) There exists $x^* \in \boldsymbol{\mathcal{X}}$ that satisfies $a(x^*) > 0,\ b(x_*) = 0$ *and* $c(x^*) = 1$. Then, $\theta_{MLE}(x) = b(x)/(a(x)+b(x))$ is admissible.

2. Main result Consider the follwing discrete probability model : $P(X = x \mid \theta) = c(\theta, x)(1-\theta)^{a(x)}\theta^{b(x)}$, $\theta \in \Theta = [0,1]$ where $a(x)$, $b(x)$ are non-negative integers and the function $c(\theta, x) \ge 0$ can be divided by neither θ nor $1-\theta$. Define $\boldsymbol{\mathcal{X}}(a) = \{x \in \boldsymbol{\mathcal{X}} \mid a(x) > 0 \ \textit{and}\ b(x) = 0\}$, $\boldsymbol{\mathcal{X}}(b) = \{x \in \boldsymbol{\mathcal{X}} \mid a(x) = 0 \ \textit{and}\ b(x) > 0\}$, $\boldsymbol{\mathcal{X}}(c) = \{x \in \boldsymbol{\mathcal{X}} \mid a(x) > 0 \ \textit{and}\ b(x) > 0,\ c(\theta, x) = c(x)\}$.
We assume that the above probability model satisfies the following two conditions:

Condition 1 For all $x \in \boldsymbol{\mathcal{X}}(a)$, $P(x\,|\,\theta) = c(\theta, x)(1-\theta)^{a(x)}$ is maximized when $\theta = 0$.

Condition 2 For all $x \in \boldsymbol{\mathcal{X}}(b)$, $P(x\,|\,\theta) = c(\theta, x)\theta^{b(x)}$ is maximized when $\theta = 1$.

Under these conditions, the MLE of θ is $\theta_{MLE}(x) = b(x)/(a(x)+b(x))$.

Theorem 2 Consider the probability model discribed in this section. Assume $\boldsymbol{\mathcal{X}}(a) \ne \emptyset$, $\boldsymbol{\mathcal{X}}(b) \ne \emptyset$, $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}(a) \cup \boldsymbol{\mathcal{X}}(b) \cup \boldsymbol{\mathcal{X}}(c)$, $\psi_a(\theta) = \sum_{x \in \mathcal{X}(a)} c(\theta, x)(1-\theta)^{a(x)}$ and $\psi_b(\theta) = \sum_{x \in \mathcal{X}(b)} c(\theta, x)\theta^{b(x)}$ are polynomials of θ in finite degree, and $\psi_a(0) = \psi_b(1) = 1$. Then, $\theta_{MLE}(x)$ is admissible.

3. Examples

We present several examples whose admissibility can be proved by Theorem 2 are given. Admissibility of the MLE in these examples cannot be shown by Theorem 1.

Example 1 Let k be the known positive integer.

$$P(x\,|\,\theta) = \begin{cases} 1-\theta^{k-1} & \text{if } x=-1, \\ (1-\theta)^x\theta^k & \text{if } x=0,\ 1,\ldots. \end{cases}$$

In this probability model, $\mathcal{X} = \{-1, 0, 1, 2, ...\}$ and the MLE of θ is

$$\theta_{MLE}(x) = \begin{cases} 0 & \text{if } x=-1, \\ k/(x+k) & \text{if } x=0,\ 1,\ 2,\ldots. \end{cases}$$

Example 2

$$P(x\,|\,\theta) = \begin{cases} (1-\theta)^n + n\theta(1-\theta)^{n-1} & \text{if } x=0, \\ \binom{n}{x}\theta^x(1-\theta)^{n-x} & \text{if } x=2,\ 3,\ ...,\ n-2, \\ n\theta^{n-1}(1-\theta)+\theta^n & \text{if } x=n. \end{cases}$$

In this probability model, $\mathcal{X} = \{0, 2, 3, ..., n-2, n\}$ and $\theta_{MLE}(x) = x/n$.

Example 3

$$P(X=x\,|\,\theta) = \begin{cases} \binom{x+k-1}{x}\theta^k(1-\theta)^x & \text{if } x=0,\ 1,\ ...,\ n-1, \\ \sum_{t=0}^{k-1}\binom{n+k-1}{t}\theta^t(1-\theta)^{n+k-1-t} & \text{if } x=n. \end{cases}$$

In this probability model, $\mathcal{X} = \{0, 1, 2, ..., n-1, n\}$ and the MLE of θ is

$$\theta_{MLE}(x) = \begin{cases} k/(x+k) & \text{if } x=0,\ 1,\ 2,\ ...,\ n-1, \\ 0 & \text{if } x=n. \end{cases}$$

Example 4

$$P(x\,|\,\theta) = \begin{cases} (1-\theta)^x\theta & \text{if } x=0,\ 1,\ 2,\ ...,\ 2m, \\ \dfrac{(1-\theta)^x}{2-\theta} & \text{if } x=2m+1,\ 2m+2. \end{cases}$$

In this probability model, $\mathcal{X} = \{0, 1, ..., 2m, 2m+1, 2m+2\}$ and the MLE of θ is

$$\theta_{MLE}(x) = \begin{cases} \dfrac{1}{x+1} & \text{if } x=0,\ 1,\ 2,\ ...,\ 2m, \\ 0 & \text{if } x=2m+1,\ 2m+2. \end{cases}$$

Deficiency of limit of Bayes estimates with respect to U-statistics and V-statistics (of degree two)

野町俊文（都城高専）、大和　元（鹿児島大学· 理学部）

ノンパラメトリックベイズ推測において、Ferguson (1973) は、分布関数、平均、メディアン、分位点、分散と共分散、$X \leq Y$の確率等について、Dirichlet プロセスを事前分布とするベイズ推定量を与えた。さらに、Dirichlet プロセスのパラメータを0に近づけたとき、ノンベイズ的立場から扱うことができるベイズ推定量の極限も与えた。一方、Dirichlet プロセスに基づくランダム汎関数の期待値は、大和 (1977) によって、Ferguson (1973) と Antoniak (1974) の結果を用いて求められている。その応用として、任意の次数の推定可能な母数 θ について、Dirichlet プロセスを事前分布とする、2乗誤差損失によるベイズ推定量と、そのベイズ推定量の極限は、与えられている。この推定量を便宜的に LB統計量と呼ぶことにする。さらに、固定した分布からの大きさ n の標本に基づいて、最小分散不偏推定量である U統計量と LB統計量の差の2乗平均が、$O(n^{-2})$ であることが示されている。他方、我々 (to appear) は、Sethuraman (1994) の構成法による Dirichlet プロセスを用いてランダム汎関数の期待値をもとめている。また、カーネルの次数が2である場合、その2つの推定量 LB統計量と U統計量の差の2乗平均を求め、さらに、LB統計量の平均2乗誤差（ＭＳＥ）、U統計量の分散、V統計量のＭＳＥを求め、それらの差を求めている。

ところで、推定量の比較の方法として、漸近相対効率を求める方法がある。しかし、これらの推定量に対して、退化する場合 (Example 4) を除くと、1である。そこで、漸近 deficiency (Hodges and Lehmann 1970) を用いて、カーネルの次数が2である場合について比較する。

漸近 deficiency: ある推定量のクラスに属していて、標本数 n と k_n に基づく推定量をそれぞれ $\delta_n, \delta^*_{k_n}$ とするとき、 δ_n と $\delta^*_{k_n}$ とが何らかの意味において、同等になるように k_n を定め、$d = \lim_{n\to\infty}(k_n - n)$ が存在するとき d を δ_nに対する$\delta^*_{k_n}$ の漸近 deficiency という。(赤平 1981)。

漸近 deficiency に関して次の定理が成り立つ。

Theorem (Lehmann 1983; p.350)　推定量 $\delta_n, \delta^*_{k_n}$ のリスク (R_{in}) が、それぞれ、

$$R_{in} = \frac{a}{n^r} + \frac{b_i}{n^{r+1}} + o(\frac{1}{n^{r+1}}) \qquad i = 1, 2$$

であるとき、漸近 deficiency は、次のようになる。

$$d = \frac{b_2 - b_1}{ar}$$

危険関数として2乗誤差平均を用いると、degree= 2 の場合、次のようになる。

$$\begin{aligned}
MSE(LB) &= \frac{4}{n}\{E[g(X_1,X_2)g(X_1,X_3)] - \theta^2\} \\
&\quad + \frac{2}{n^2}\{-10E[g(X_1,X_2)g(X_1,X_3)] + E[g(X_1,X_2)^2] + 4E[g(X_1,X_2)g(X_1,X_1)] \\
&\quad -8\theta E[g(X_1,X_1)] + 2[Eg(X_1,X_1)]^2 + 11\theta^2\} + o(n^{-2}), \\
Var(U) &= \frac{4}{n}\{E[g(X_1,X_2)g(X_1,X_3)] - \theta^2\} \\
&\quad + \frac{2}{n^2}\{-2E[g(X_1,X_2)g(X_1,X_3)] + E[g(X_1,X_2)^2] + \theta^2\} + o(n^{-2}), \\
MSE(V) &= \frac{4}{n}\{E[g(X_1,X_2)g(X_1,X_3)] - \theta^2\} \\
&\quad + \frac{1}{n^2}\{-12E[g(X_1,X_2)g(X_1,X_3)] + 2E[g(X_1,X_2)^2] + 4E[g(X_1,X_2)g(X_1,X_1)] \\
&\quad -6\theta E[g(X_1,X_1)] + [Eg(X_1,X_1)]^2 + 11\theta^2\} + o(n^{-2})。
\end{aligned}$$

このことにより、LB統計量の U統計量に対する漸近 deficiency と LB統計量の V統計量に対する漸近 deficiency は、それぞれ、次のようになる。

$$d(LB,U) = \frac{-4E[g(X_1,X_2)g(X_1,X_3)] + 2E[g(X_1,X_2)g(X_1,X_1)] - 4\theta E[g(X_1,X_1)] + [Eg(X_1,X_1)]^2 + 5\theta^2}{E[g(X_1,X_2)g(X_1,X_3)] - \theta^2}$$

$$d(LB,V) = \frac{-8E[g(X_1,X_2)g(X_1,X_3)] + 4E[g(X_1,X_2)g(X_1,X_1)] - 10\theta E[g(X_1,X_1)] + 3[Eg(X_1,X_1)]^2 + 11\theta^2}{4E[g(X_1,X_2)g(X_1,X_3)] - \theta^2}$$

Example 1. $\theta = P\{X+Y \leq 0\}$ の推定;
対称なカーネル $g(x_1,x_2) = I(x_1+x_2 \leq 0)$ である。原点について対称な連続分布からの標本に対して、

$$d(LB,U) = -1, \quad d(LB,V) = -\frac{1}{2}$$

を得る。

Example 2. Probability weighted moments $\theta = E[\frac{1}{2}\mathrm{Max}(X,Y)]$ の推定;
原点について対称な一様分布からの標本に対して、

$$d(LB,U) = 1, \quad d(LB,V) = \frac{13}{16}$$

を得る。

Example 3. 分散 $\theta = \sigma^2 = \int \frac{1}{2}(x_1-x_2)^2 dF(x_1)dF(x_2)$ の推定;
４次のモーメントを持つ原点について対称な連続分布からの標本に対して、

$$d(LB,U) = \frac{-4(\mu_4 - 2(\mu_2)^2}{\mu_4 - (\mu_2)^2}, \quad d(LB,V) = \frac{-2\mu_4 + 5(\mu_2)^2}{\mu_4 - (\mu_2)^2}$$

を得る。特に、 N(0, 1) からの標本に対して、次のようになる。

$$d(LB,U) = -2, \quad d(LB,V) = -\frac{1}{2}$$

Example 4. 平均の２乗 $\theta = \mu^2 = (\int x dF(x))^2$の推定;
４次のモーメントを持つ原点について対称な連続分布からの標本に対して、次のようになる。

$$\begin{aligned} MSE(LB) &= \frac{6}{n^2}(\mu_2)^2 + O(n^{-3}), \\ Var(U) &= \frac{2}{n^2}(\mu_2)^2 + O(n^{-3}),\ MSE(V) = \frac{3}{n^2}(\mu_2)^2 + O(n^{-3}) \text{ より、} \\ ARE(LB,U) &= \frac{1}{3},\ ARE(LB,V) = \frac{1}{2} \end{aligned}$$

[参考文献]

Akahira, M. (1981), 数理科学, **219**, 24–32
Antoniak, C. E. (1974), *The Annals of Statistics*, **2**, 1152–1174.
Ferguson, T. S. (1973), *The Annals of Statistics*, **1**, 209-230.
Hodges, J. L. and Lehmann, E. L. (1970) *The Annals of Mathematical Statistics*, **41**, 783–801.
Lehmann, E. L. (1983), *Theory of point estimation*, John Wiley & Sons.
Nomachi and Yamato (to appear)
Sethuraman, J. (1994), *Statistica Sinica*, **4**, 639–650.
Yamato, H. (1977), *Journal of the Japan Statistical Society*, **7**, 57-66.

退化型 U-統計量に関する漸近展開について

金沢大学工学部　金川秀也

Let $\{\xi_j, j\geq 1\}$ be i.i.d. random variables with a probability distribution μ. Suppose that $u(x,y)$ is a real valued symmetric function on $\mathbf{R}\times\mathbf{R}$. Furthermore assume that u is square integrable with respect to $\mu\times\mu$ and degenerate (or canonical), i.e., for any real x

$$E\big(u(\xi_1,x)\big)=0. \tag{1}$$

Let L^2 be the space of all square integrable functions with respect to μ. Then, according to Serfling (1980), we see that the kernel u induces a bounded linear operator $T_u: L^2\to L^2$ (trace class) defined by $T_u f(x) := E\big[u(\xi_1,x)f(\xi_1)\big]$, $f\in L^2$ which has eigenvectors $\{g_i\}$ and eigenvalues $\{\lambda_i\}$ satisfying for each $i\geq 1$

$$\begin{cases} E\big(g_i(\xi_1)\big)=0, \quad E\big(g_i^2(\xi_1)\big)=1 \\ E\big(g_i(\xi_1)g_j(\xi_1)\big)=0\ (i\neq j), \quad E\big(h(\xi_1,x)g_i(\xi_1)\big)=\lambda_i g_i(x) \end{cases}. \tag{2}$$

Then u can be represented by

$$u\big(\xi_i,\xi_j\big)=\sum_{k=1}^{\infty}\lambda_k g_k(\xi_i)g_k\big(\xi_j\big). \tag{3}$$

We introduce a separable Hilbert space H equipped with the inner product $\langle\cdot,\cdot\rangle$ and the norm $\|\cdot\|$ as follows:

$$H:=\left\{\boldsymbol{x}=(x_1,x_2,\cdots)\in\mathbf{R}^\infty : \sum_{k=1}^{\infty}|\lambda_k|\boldsymbol{x}_k^2<\infty\right\},$$

$$\langle\boldsymbol{x},\boldsymbol{y}\rangle:=\sum_{k=1}^{\infty}|\lambda_k|x_k y_k, \quad \|\boldsymbol{x}\|:=\left(\sum_{k=1}^{\infty}|\lambda_k|x_k^2\right)^{1/2}.$$

If we assume that

$$\sum_{k=1}^{\infty}|\lambda_k|<\infty, \tag{4}$$

then from (5.1)

$$E\left(\sum_{k=1}^{\infty}|\lambda_k|g_k^2(\xi_i)\right)=\sum_{k=1}^{\infty}|\lambda_k|E\big(g_k^2(\xi_i)\big)=\sum_{k=1}^{\infty}|\lambda_k|<\infty,$$

which implies that we can define H-valued random variables $G_i := \big(g_1(\xi_i), g_2(\xi_i), g_3(\xi_i),\cdots\big)$ for each $i\geq 1$. Let $\{U_n, n\geq 1\}$ and $\{V_n, n\geq 1\}$ be U-statistics and V-statistics with degree 2 defiende by

$$U_n := \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n} u\big(\xi_i,\xi_j\big).$$

We have the Edgeworth expansion for U_n.

Theorem. *Suppose that* $\{\xi_j, j \ge 1\}$ *is a sequence of i.i.d. random variables with law* μ. *Assume that* u *is square integrable symmetric function with respect to* $\mu \times \mu$. *Furthermore suppose (1), (4) and* $E\|G_1\|^s < \infty$ *for some* $s \ge 3$. *Put*

$$F_n(r) := P\left\{\sqrt{n(n-1)}U_n \le r\right\}, \quad G_n(r) := P\{f(\eta) \le r\} + \sum_{k=1}^{s-2} U_k(r)$$

where η *is a random variable with* Gussian distribution Φ *on* H. *Then we have*

$$F_n(r) = G_n(r) + R_n(r) + O\left(n^{-(s-2)/2}\right) \quad as\ n \to \infty.$$

If $F_n(r)$ *satisfies the Lipschtz condition, then*

$$G_n(r) = O\left(n^{-1}\right)\ as\ n \to \infty.$$

Remark The Cramér-von Mises-Smirnov statistic

$$W_n := \int_0^1 n w(x)(F_n(x) - 1)^2 dx$$

is a typical example satisfying (2.4), where $w(x)$ is a weight function and $F_n(x)$ is the empirical distribution function of uniformly distributed random variables. The kernel u of W_n is

$$u(x, y) := \int_0^1 w(u)\left(I_{\{x \le u\}}(x) - u\right)\left(I_{\{y \le u\}}(x) - u\right)du.$$

When $w(x) = \{x(1-x)\}^{-1}$, the eigenvalues of u are

$$\lambda_k = \frac{1}{k(k+1)}, \quad k \ge 1$$

which satisfies (4) (see Csáki, 1980 or Borovskikh, 1985 for more details).

References

F. Götze, On Edgeworth expansions in Banach spaces, Ann. Probability, **9** (1981) 852-859.

S. Kanagawa and K. Yoshihara, The almost sure invariance principles of degenerate U-statistics of degree two for stationary random variables, Stochastic Processes and their Applications, **49** (1994) 347 - 356.

R. J. Serfling, Approximation Theorems of Mathematical Statisitics, JohnWiley and Sons, New York, 1980.

群逐次検定方式の開発と検討

東海大学・理学部　　道家暎幸

1. はじめに

群逐次検定方式は臨床医学の分野での2処置間の効果の差の検定に関し、倫理性あるいは逐次検定に対し試験実施上の利便性の観点から研究開発されてきた。この検定方式はPocock(1977)が初めて提案し、正規反応または2値反応の群観測値をもとに2つの母平均の差や2つの母比率の差の検定についての手順である。その後、O'Brien-Fleming(1979)、Lan-DeMet(1983)等が繰り返し信頼区間の設定に関し、異なる接近方法を議論している。

この群逐次検定方式の主な目的は、2処置間の効果の差の検定に関し、第1種の過誤の確率を保ちながら、早い結論を得る、または観測個数を期待的に少なくする検定方式の開発が考えられる。そこで検出力を上げるための新しい統計量を提案したり、繰り返し信頼限界を調整する方式の考案が必要である。更に2処置間の差の検定よりも2処置間の優劣比較検定へ、1変量応答から多変量応答をもとにした群逐次検定方式への展開が考えられる。今回は次の様なテーマについて研究を行う。

2. 群逐次検定における各段階での標本数の決定手順

この群逐次検定において、Pocock(1977)は事前に指定された最大検定回数、有意水準と検出力の下で試験に必要な標本数（最大標本数）の決定手順を示しているが、一般に母平均の差が小さいとき相当大きな最大標本数が必要となる。また、この手順で得られる各段階での標本数は全て同一と仮定しているため、あまり実際的でない。このような状況に中で検定の途中段階で次段階の検定に必要な標本数を決定する手順を提案する。これはLan-DeMet法(1983)でのアルファ消費関数を用い、有意水準と検出力を各段階へ振り分け、次の段階での標本数を決定するものである。その決定に際し、標本数を減らす方策として、これまでの段階で得られた2処置の標本平均をもとに、2処置の優劣を表す比率を用い、優位な処置が採択されやすいような信頼限界を設定する。これより、従来の群逐次検定方式と異なり、繰り返し信頼限界線は非対称であり、各段階毎の標本数は異なる。シミュレーションにより、最大標本数に関しPocock法と本手順の有効性について比較検討する。

3. 2変量正規応答データに基づく群逐次検定

従来の群逐次検定方式は1変量応答の群観測値をもとに逐次的に2処置間の効果の差を検定する方式である。しかし、実際の臨床試験では2処置間の効果の差よりも2処置間の優劣比較のための群逐次検定方式の方が利用価値は高い。また、処置に対する反応を1変量で測るよりも2変量で測ったほうがより有効な場合も

多い。Jennison-Turnbull(1993) は安全性と有効性の２変量正規応答データに基づいた優位性のための群逐次検定の研究を行っている。その群逐次検定での検定統計量の導出に関し、Kudo(1963) の複合仮説検定方式を参考にしている。本研究では２変量正規応答データに基づく２処置間の優劣判定のための群逐次検定を考える。まず、Inada(1978) の方法を参考に尤度比を用いて統計量を導出し、繰り返し信頼限界を設定する。更に相関係数と検出力との関連を調べる。

4. 群逐次 χ^2 統計量と検出力

最近、２処置間の効果の差を検定するための群逐次検定方式に関し、１変量よりも多変量観測値をもとにした群逐次検定方式の研究が見られる。Jennison-Turnbull(1993) は、２変量正規応答データをもとに優位性のための群逐次検定の研究を行い、Lui(1993) は２つ以上の処置を同時に扱った群逐次検定方式を提案している。Jennison-Turnbull(1991) は１標本問題での多変量群逐次検定において、群逐次検定の特徴を使い新しい群逐次統計量を提案している。この研究では多変量観測値をもとに２処置間の平均ベクトルの差の検定を行う際、分散共分散行列を既知とし、検定統計量として群逐次 χ^2 統計量を用いる。更に、検出力を上げる目的で、Jennison-Turnbull(1991) の考え方を利用して新しい修正群逐次 χ^2 統計量を提案する。この統計量の分布が非心 χ^2 分布に従うことを確かめる。次に２つの統計量に関して繰り返し信頼区間を設定した後、その統計量をもとにした群逐次決定方式の有効性を検出力に関して比較する。

参考文献

[1]Inada,K.(1978).Some Bivariate Tests of Composite Hypotheses with Restricted Alternatives. Reports of the Faculty of Science, Kagoshima University, No.11,25-31.
[2] Jennison,C. and Turnbull,B.W.(1991). Exact Calculations for Sequential t, χ^2 and F Tests, Biometrika, 78,1, 133-141.
[3] Jennison,C. and Turnbull,B.W.(1993).Group Sequential Test for Bivariate Response: Interim Analyses of Clinical Trials with Both Efficacy and Safety Endpoints, Biometrics 49, 741-752.
[4]Kudo,A.(1963).A Multivariate Analogue of the One-sided Test. Biometrika 50,403-418.
[5] Lan,K.K.G. and DeMets,D.L.(1983).Discrete Sequential Boundaries for Clinical Trials, Biometrika 70, 659-663.
[6] Lui,K.J.(1993). A Simple Generalization of the O'Brien and Fleming Group Sequential Test Procedure to More Than Two Treatment Groups, Biometrics 49, 1216- 1219.
[7] O'Brien,P.C. and Fleming,T.R.(1979).A Multiple Testing Procedure for Clinical Trials, Biometrics 35, 549-556.
[8] Pocock,S.J.(1977). Group Sequential Methods in the Design and Analysis of Clinical Trials, Biometrika 64, 191-199.

Nonparametric sequential tests and change-point detection problems (ノンパラメトリックな逐次検定と変化点の探索の問題)

長崎大学・経　永井　圭二

1 序

Woodroofe(1983) は、Savage-Sethuraman(1966) による二標本問題でレーマン対立仮説を検定する順位による逐次確率比検定に対し、非線形更新定理が用いられるかどうかという問題を提起した。この報告では今までの研究者とは異なる方法で順位の対数尤度比を Chernoff-Savage 的に高次に展開し、その問題を解決する。さらに、その応用として変化点の探索の問題 Nagai(1998) を考える。

2 Lehmann 対立仮説の順位による尤度比検定

独立な確率変数列 $(X_1,\dots,X_n,Y_1,\dots,Y_n)$ が観測されるものとする。ここで X_i は分布 F に従い、Y_i は分布 G に従うものとする。ここでレーマン対立仮説の検定 $H_0: G=F$ vs $H_1: G=F^{\Delta}, \Delta>0$ を考える。Savage (1956) によれば、順位の対数尤度比はつぎのように書ける。

$$l_n = n\log\Delta + n\int_{-\infty}^{\infty}\log\left(\frac{F_n+G_n}{F_n+\Delta G_n}\right)d(F_n+G_n). \tag{2.1}$$

Lai(1975) は l_n を Chernoff-Savage 統計量と見なして、ランダムウォーク（独立同一な確率変数の和）と残余項の和に書いた。すなわち、

$$\begin{aligned} l_n &= S_n+\xi_n, \\ S_n &= nS(\Delta,F,G)+n(1-\Delta)\int_{-\infty}^{\infty}\frac{G_n-G}{F+\Delta G}dF+n(\Delta-1)\int_{-\infty}^{\infty}\frac{F_n-F}{F+\Delta G}dG, \end{aligned} \tag{2.2}$$

ここで $S(\Delta,F,G)$ は Kulback-Leibler 情報量で

$$S(\Delta,F,G)=\log(\Delta)+\int_{-\infty}^{\infty}\log\left(\frac{F+G}{F+\Delta G}\right)d(F+G).$$

残余項に対して Lai はどのような小さな $\mu>0$ に対しても $n^{-\mu}\xi_n\to 0$ という漸近的な結果を得た。これに対し、U 統計量の理論を用いることにより、次の結果が得られる。

レーマン対立仮説 $G=F^{\Delta}$ を検定する順位の対数尤度比 l_n はランダムウォークと残余項の和に書ける；$l_n=S_n+\xi_n$. ここで S_n は (2.2) で定義されたものと同じである。残余項 ξ_n は次の関係を満

足する。

$$E\Big[\max_{n\le k\le 2n}\Big(\xi_k - c^*(k)\Big)^2\Big] = O(\log n), \tag{2.3}$$

$c^*(n)$ は次のような数列である。$H=\frac{F+G}{2}$ として、

$$\begin{aligned} c^*(n) &= \int_{H>n^{-1}}\left(\frac{1}{F+G}-\frac{1}{F+\Delta G}\right)(1-F)dF+\int_{H>n^{-1}}\left(\frac{1}{F+G}-\frac{\Delta}{F+\Delta G}\right)(1-G)dG \\ &\quad +\int_{H>n^{-1}}\left\{\left(\frac{-1}{(F+G)^2}+\frac{1}{(F+\Delta G)^2}\right)F(1-F)+\left(\frac{-1}{(F+G)^2}+\frac{\Delta^2}{(F+\Delta G)^2}\right)G(1-G)\right\}dH. \end{aligned}$$

さらに G が実際にレーマン対立仮説であるとすると、すなわち $A\neq 1$ に対して $G=F^A$ とすると、

$$\{\max_{0\le j\le n}|\xi_{n+j}|, n\ge 1\} \quad \text{は一様可積分、、} \quad \xi_n \to V \quad \text{弱収束、}(\ n\to\infty)\ . \tag{2.4}$$

ここで V はある可積分な確率変数である。

3 順位による逐次確率比検定の期待標本数の漸近展開

Savage and Sethuraman (1966) は順位による逐次確率比検定 $N=\inf\{n\ge 1: l_n<a \text{ or } l_n>b\}$ $(a<0<b)$ を定義した。Berk (1973) によれば、その期待標本数の 1 次の漸近近似は

$$\begin{aligned} EN &= \frac{b}{S(\Delta,F,G)}(1+o(1)) \quad \text{if } S(\Delta,F,G)>0\,, \\ EN &= \frac{|a|}{|S(\Delta,F,G)|}(1+o(1)) \quad \text{if } S(\Delta,F,G)<0, \end{aligned}$$

によって与えられる。$(\min(|a|,b)\to\infty)$ Woodroofe (1983) は、この順位による逐次確率比検定で第一種と第二種の誤りを犯す確率の 2 次の近似を求めたが、同時に非線形更新定理を期待標本数の 2 次の漸近展開に使えるかどうかという問題を提起した。その解答として次の結果を得る。

もし $S(\Delta,F,G)>0$ ならば、

$$E(N)=\frac{b-c^*(b)}{S(\Delta,F,G)}+o(\log b) \qquad (\ |a|,b\to\infty). \tag{3.1}$$

$G=F$ の場合、

$$EN=\frac{|a|}{|S(\Delta,F,F)|}+\frac{(c^{**}+o(1))}{|S(\Delta,F,F)|}\log|a| \quad (|a|,b\to\infty). \tag{3.2}$$

ここで $c^{**}=\frac{1}{2}(1-\Delta)^2/(1+\Delta)^2$。さらに $G=F^A$、$A\neq 1$、かつ $S(\Delta,F,F^A)>0$ とすると、

$$E(N)=\frac{b+r-EV}{S(\Delta,F,F^A)}+o(1) \quad (\ |a|,b\to\infty). \tag{3.3}$$

ここで、$r=ES^2_{\tau_+}/(2ES_{\tau_+})$、 $\tau_+=\inf\{n; S_n>0\}$ である。

また、報告では 2 標本の変化点の探索についての Nagai(1998) の結果についても考察された。

The Bayes Sequential Test fot Normal Variance (Preliminary Report)

Herman Chernoff (Harvard University)
Hajime Takahashi (Hitotsubashi University)

The object of this paper is to carry out an experiment on the ability to apply the optimal sequential test for determining the sign of a normal mean as an approximation to the solution of a large class of sequential testing problems. Two potential advantages of such an approximation, if it is effective, are that it is easy to calculate and that it can be expressed simply in terms of a single easily interpreted curve, determining the optimal stopping time. Incidentally one variation of that curve is one where a nominal significance level for stopping is expressed as a function of the information obtained to date.

The experiment consists of comparing the approximation to the optimal Bayes solution for the sequential test that the standard deviation of a normal distribution exceeds one. The optimal solution is obtained numerically by the backward induction. In this problem the unknown value of $\theta = \sigma^{-2}$ is assumed to have a Gamma(α_0, β_0), where

$$f(\theta \mid \alpha, \beta) = [\alpha^{\beta}/\Gamma(\beta)]\, \theta^{\beta-1} e^{-\alpha\theta}$$

Then the posterior distribution of θ given $X_1, \ldots, X_{n+1}$ is the gamma with $\alpha = \alpha_0 + S_n/2$ and $\beta = \beta_0 + n/2$, where $S_n = \Sigma_{i=1}^{n+1}(X_i - \bar{X}_{n+1})^2$, $\bar{X}_{n+1} = \Sigma_{i=1}^{n+1} X_i$.

We relate our problem to the normal problem as follows. First we note that the Fisher Information per observation for estimating σ is $2\sigma^{-2}$. This corresponds to σ^{-2} in estimating normal mean μ (cf. Chernoff (1972)). Second, if θ has the Gamma(α, β) distribution,

(1) $$Y^* = E\{\sigma - 1\} = E\{\theta^{-1/2} - 1\} = \alpha^{-1/2}\{\Gamma(\beta - 1/2)/\Gamma(\beta)\} - 1$$

which is comparable to Y(s) process in the normal mean problem (Chernoff(1972)). The variance of Y^* is

(2) $$\mathrm{Var}\{Y^*\} = \alpha\{\Gamma(\beta-1)/\Gamma(\beta) - [\Gamma(\beta-1/2)/\Gamma(\beta)]^2\}.$$

When $\sigma = 1$ and β is large α tends to be close to β and we may approximate the variance of the distribution of $\sigma - 1$ by

(3) $$s^* = s^*(\beta) = \beta\{\Gamma(\beta-1)/\Gamma(\beta) - [\Gamma(\beta-1/2)/\Gamma(\beta)]^2\}$$

which has the advantage of depending only on β which is deterministic in this sequential problem. Now the time between each observatrions δs^* is easily calculated by

(4) $$\delta s^* = s^*(\beta) - s^*(\beta + 1/2).$$

Given α_0, β_0, c,k = 1, we can approximate σ^{*2} by 1/2 and the analogue of a is $a^* = (2/c)^{1/3}$, (see Chernoff(1972)). We may express our normal approximation to the solution of the problem by the boundaries

(5) $\tilde{y}^* = \tilde{y}(s^*) = y(a^{*2}s^*)$

or

(6) $y^* = (1/a^*)\tilde{y}^* = \pm(1/a^*)\tilde{y}(a^{*2}s^*)$

where $\tilde{y}$ is the tabulated solution for the normalized normal problem. It is convenient to stay in the (α, β) plane, and as above,

(7) $\alpha = [\Gamma(\beta)(1 \pm y^*)/\Gamma(\beta - 1/2)]^2$

With the continuity correction, we have intersecting boundaries

(8) $\alpha = [\Gamma(\beta)(1 \pm (y^* - 0.58(\delta s^*)^{1/2})/\Gamma(\beta - 1/2)]^2$

This approximation has several shortcomings. We have used the approximations $\alpha = \beta$ and $\sigma = 1$. When the boundaries are relatively far apart, as they tend to be for small β, these approximations are of questionable value. The correction for discontinuity, which is related to the distribution of the excess in crossing the boundary (Hogan) depends on the distribution. For example, if the increments had the discrete binomial distribution with $X_i = \pm 1$ with probability 1/2, the correction would be 0.5. In our case, with a heavily skewed chi-square increment, there is likely to be a better correction for discontinuity with different corrections for the lower and upper boundary.

These difficulties are alleviated somewhat by the fact that there is usually very little loss when the optimal Bayes strategy is modified somewhat, leading to improvement for some values of the parameter which compensate for deficits for other values of the parameter. In the process of calculating the optimal Bayes strategy we shall compare the Bayes risks for the optimal with that of the normal approximation. The main difference in the backward induction calculation for the normal approximation is that the rules for stopping and continuing are prescribed in advance, and not by that of minimizing between the two alternatives.

In addition, we propose using the above normal approximation for a first step in deriving a second better approximation. Instead of approximating α by β in deriving s^* we will use a different α_u and α_l selected from (8) for our first approximation for the upper and lower boundaries. These give rise to distinct s^*_u and s^*_l functions as well as δs^*_u and δs^*_l to substitute in (8) together with adjusted corrections for continuity . We also consider the Edgeworth type approximation to α.

多変量分散成分の推定について

東京大学 経済学部　久保川達也

多変量分散成分モデルもしくは多変量混合線形モデルにおける分散成分の推定は，Bock and Vandenberg (1968), Bock and Petersen (1975), Klotz and Putter (1969) らによって ML や REML 推定に関して古くから研究されてきた問題であり，最近 Amemiya (1985), Amemiya and Fuller (1984), Anderson *et al.* (1986) らの論文によって再認識され，理論研究の新たな展開をみせている。

多変量の一元配置モデル

$$\boldsymbol{y}_{ij} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{e}_{ij}, \quad i = 1, \ldots, k,\ j = 1, \ldots, r,$$

において $\boldsymbol{\alpha}_i$，$\boldsymbol{e}_{ij}$ は独立な確率変数で，$\boldsymbol{\alpha}_i \sim \mathcal{N}_p(\boldsymbol{0}, \Sigma_A)$, $\boldsymbol{e}_{ij} \sim \mathcal{N}_p(\boldsymbol{0}, \Sigma_1)$ に従う。$\boldsymbol{\mu} \in \boldsymbol{R}^p$ は未知の共通平均，Σ_A，Σ_1 は未知の共分散行列とする。$\overline{\boldsymbol{y}}_{i.} = r^{-1}\sum_{j=1}^r \boldsymbol{y}_{ij}$, $\overline{\boldsymbol{y}}_{..} = (rk)^{-1}\sum_{i=1}^k\sum_{j=1}^r \boldsymbol{y}_{ij}$ として，$\boldsymbol{S}_1 = \sum_{i=1}^k\sum_{j=1}^r(\boldsymbol{y}_{ij} - \overline{\boldsymbol{y}}_{i.})(\boldsymbol{y}_{ij} - \overline{\boldsymbol{y}}_{i.})'$, $\boldsymbol{S}_2 = r\sum_{i=1}^k(\overline{\boldsymbol{y}}_{i.} - \overline{\boldsymbol{y}}_{..})(\overline{\boldsymbol{y}}_{i.} - \overline{\boldsymbol{y}}_{..})'$ とおくと，$\overline{\boldsymbol{y}}_{..}$, $\boldsymbol{S}_1$，$\boldsymbol{S}_2$ は最小十分統計量で，互いに独立に

$$\begin{aligned}
\overline{\boldsymbol{y}}_{..} &\sim \mathcal{N}_p(\boldsymbol{\mu}, (rk)^{-1}\Sigma_2), \quad \Sigma_2 = \Sigma_1 + r\Sigma_A, \\
\boldsymbol{S}_1 &\sim \mathcal{W}_p(\Sigma_1, n_1), \quad n_1 = k(r-1), \\
\boldsymbol{S}_2 &\sim \mathcal{W}_p(\Sigma_2, n_2), \quad n_2 = k-1,
\end{aligned}$$

に従う。共分散行列 Σ_1, Σ_2 の間には $\Sigma_1 \le \Sigma_2$ なる不等式制約が入っている。ただし，この不等式は $\Sigma_2 - \Sigma_1$ が非負定値であることを表わす。推定精度を高めるためにこの不等式制約の情報を使いたいのであるが，有限標本での推定量の優越性を主張することは技術的に困難であると思われてきた。しかし最近 Srivastava and Kubokawa (1999) は，この問題を解決するのに成功したので，彼らの主要部分を以下に紹介する。

まず群内分散 Σ_1 の推定問題を扱う。$\boldsymbol{S}_2^{1/2}$ を $\boldsymbol{S}_2 = (\boldsymbol{S}_2^{1/2})^2$ なる対称行列とし，$\boldsymbol{P}$ を

$$\boldsymbol{P}'\boldsymbol{S}_2^{-1/2}\boldsymbol{S}_1\boldsymbol{S}_2^{-1/2}\boldsymbol{P} = \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p),$$

なる $p \times p$ 直交行列とする。ここで固有値は $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_p$ をみたすとする。非負関数 $\psi_i(\Lambda)$ に対して，$\Psi(\Lambda) = \mathrm{diag}(\psi_1(\Lambda), \ldots, \psi_p(\Lambda))$ とおいて，Σ_1 の推定量の一般形

$$\widehat{\Sigma}_1(\Psi) = \boldsymbol{S}_2^{1/2}\boldsymbol{P}\Psi(\Lambda)\boldsymbol{P}'\boldsymbol{S}_2^{1/2}$$

を考えよう。この推定量を $\boldsymbol{S}_2$ の情報を用いて改善するために，打ち切りルール $[\Psi(\Lambda)]^{TR}$ を

$$\begin{aligned}
[\Psi(\Lambda)]^{TR} &= \mathrm{diag}(\psi_1^{TR}(\Lambda), \ldots, \psi_p^{TR}(\Lambda)), \\
\psi_i^{TR}(\Lambda) &= \min\left\{\psi_i(\Lambda), \frac{\lambda_i + 1}{n_1 + n_2}\right\}, \quad i = 1, \ldots, p,
\end{aligned}$$

で定義する。そのとき打ち切り推定量は

$$\widehat{\Sigma}_1([\Psi]^{TR}) = \boldsymbol{S}_2^{1/2}\boldsymbol{P}\mathrm{diag}(\psi_1^{TR}(\Lambda), \ldots, \psi_p^{TR}(\Lambda))\boldsymbol{P}'\boldsymbol{S}_2^{1/2},$$

と表わされる。Srivastava and Kubokawa (1999) は，エントロピー損失のもとで，推定量 $\widehat{\Sigma}_1(\Psi)$ が $\widehat{\Sigma}_1([\Psi]^{TR})$ によって改良されることを証明した。

$\boldsymbol{S}_1 = \boldsymbol{S}_2^{1/2}\boldsymbol{S}_2^{-1/2}\boldsymbol{S}_1\boldsymbol{S}_2^{-1/2}\boldsymbol{S}_2^{1/2} = \boldsymbol{S}_2^{1/2}\boldsymbol{P}\Lambda\boldsymbol{P}'\boldsymbol{S}_2^{1/2}$ であるから，不偏推定量 $\widehat{\Sigma}_1^{UB} = n_1^{-1}\boldsymbol{S}_1$ は

$$\widehat{\Sigma}_1^{UB} = \widehat{\Sigma}_1(\Psi^{UB}), \quad \Psi^{UB} = \mathrm{diag}(n_1^{-1}\lambda_1, \ldots, n_1^{-1}\lambda_p)$$

と表わされる。上述の打ち切りルールを適用すると $\widehat{\Sigma}_1^{REML} = \widehat{\Sigma}_1([\Psi^{UB}]^{TR})$,

$$[\Psi^{UB}]^{TR} = \mathrm{diag}\left(\min\left\{\frac{\lambda_1}{n_1}, \frac{\lambda_1+1}{n_1+n_2}\right\}, \ldots, \min\left\{\frac{\lambda_p}{n_1}, \frac{\lambda_p+1}{n_1+n_2}\right\}\right)$$

なる推定量が得られ不偏推定量を改良することがわかるが，これは REML 推定量になっている。

上で取り上げられた不偏推定量がミニマクスにならないことは，共分散行列の推定においてよく知られた結果である。我々の問題において 1 つのミニマクスな推定量は Stein 型推定量で $\widehat{\Sigma}_1^{ST} = \widehat{\Sigma}_1(\Psi^{ST})$,

$$\Psi^{ST}(\Lambda) = \mathrm{diag}(d_1\lambda_1, \ldots, d_p\lambda_p), \quad d_i = (n_1+p+1-2i)^{-1}, \quad i = 1, \ldots, p.$$

で与えられる。この推定量に打ち切りルールを適用すると $\widehat{\Sigma}_1^{STTR} = \widehat{\Sigma}_1([\Psi^{ST}]^{TR})$,

$$[\Psi^{ST}(\Lambda)]^{TR} = \mathrm{diag}\left(\min\left\{d_1\lambda_1, \frac{\lambda_1+1}{n_1+n_2}\right\}, \ldots, \min\left\{d_p\lambda_p, \frac{\lambda_p+1}{n_1+n_2}\right\}\right),$$

となる。推定量 $\widehat{\Sigma}_1^{STTR}$ がリスクの意味で優れていることが Srivastava and Kubokawa (1999) によって数値的に示されている。

一方，同様な打ち切りルールが群間分散 Σ_A の推定においても得られ，不偏推定量に対して

$$\widehat{\Sigma}_A^{REML} = r^{-1}\boldsymbol{S}_2^{1/2}\boldsymbol{P}\mathrm{diag}\left(\max\left\{\frac{1}{n_2} - \frac{\lambda_i}{n_1}, 0\right\}, i = 1, \ldots, p\right)\boldsymbol{P}'\boldsymbol{S}_2^{1/2},$$

や，Stein 型ミニマクス推定量に対して

$$\widehat{\Sigma}_A^{STR} = r^{-1}\boldsymbol{S}_2^{1/2}\boldsymbol{P}\mathrm{diag}\left(\max\left\{e_i - d_i\lambda_i, 0\right\}, i = 1, \ldots, p\right)\boldsymbol{P}'\boldsymbol{S}_2^{1/2}$$

などの，非負定値な打ち切り推定量が導かれる。ここで $e_i = 1/\{n_2 - (p+1-2i)\}$ である。

本研究成果の詳しい内容については，以下の文献を参照して下さい。

Srivastava, M.S. and Kubokawa, T. (1999). Improved nonnegative estimation of multivariate components of variance. *Ann. Statist.*, **27**, No. 6, to appear.

Admissible minimax estimators of a mean vector of scale mixtures of multivariate normal distributions

Yuzo Maruyama
Graduate School of Mathematics, Kyushu University

1 Introduction

Since Stein(1956) showed that the usual minimax best equivariant estimator of a p-dimensional normal mean is inadmissible for $p \geq 3$, much study of improving upon the best invariant estimator have been given. Brown(1966) substantially extended this result, the Stein phenomenon, to a very wide class of distributions and loss functions. Brandwein and Strawderman(1991) found and discussed classes of minimax estimators of a mean vector of spherically symmetric distributions. In view of statistical decision theory, however, we are interested in characterizing a class of estimators satisfying not only minimaxity but also admissibility. So far as I know, explicit results of constructing admissible minimax estimators of a mean vector were restricted to the case of the normal distribution. In the normal case, see Strawderman(1971), Alam(1973), Berger(1976), Fourdrinier *et al.*(1998) and Maruyama(1998). Here we consider scale mixtures of multivariate normal distributions as follows. Let X have density $f(\|x-\theta\|^2)$ where

$$f(\|x-\theta\|^2) = \int_0^\infty (2\pi)^{-p/2} v^{p/2} \exp(-\frac{\|x-\theta\|^2 v}{2}) g(v) dv,$$

where $g(v)$ is a known probability density function. The object is to estimate θ with the quadratic loss function $\|\delta-\theta\|^2$. Admissible minimax estimators of the mean vector of scale mixtures of multivariate normal distributions will be found.

2 The main result

First of all, we derive the class of generalized Bayes estimators for the following prior distribution. Let the conditional distribution of θ given λ, $0 < \lambda < 1$, be normal with mean 0 and covariance matrix $\lambda^{-1}(1-\lambda)I_p$ and a density function of λ is proportional to $\lambda^{-a}(1-\lambda)^b I_{(0,1)}(\lambda)$. the (generalized) density function $h_{a,b}(\theta)$ is

$$h_{a,b}(\theta) = C\int_0^1 \left(\frac{t}{1-t}\right)^{p/2} \exp\left(-\frac{t}{2(1-t)}\|\theta\|^2\right) t^{-a}(1-t)^b dt.$$

We have the generalized Bayes estimator $\delta_{a,b}(x) = (1-\phi_{a,b}(\|x\|^2)/\|x\|^2)x$, where

$$\phi_{a,b}(w) = w\frac{\int_0^1\int_0^\infty \lambda^{p/2-a+1} v^{p/2-a+1}(1-\lambda)^b(1-\lambda+v\lambda)^{a-b-2}\exp(-wv\lambda/2)g(v)dvd\lambda}{\int_0^1\int_0^\infty \lambda^{p/2-a} v^{p/2-a+1}(1-\lambda)^b(1-\lambda+v\lambda)^{a-b-2}\exp(-wv\lambda/2)g(v)dvd\lambda}.$$

We show the two theorems which characterize the properties of the behavior of $\phi_{a,b}(w)$.

Theorem 1 *Assume the following conditions:*

1. $b-a+2 \geq 0$ *and* $b \geq 0$.

2. *for* $s_1 \leq s_2$ *and* $t_1 \leq t_2$, $g(s_1t_1)g(s_2t_2) \leq g(s_1t_2)g(s_2t_1)$.

Then $\phi_{a,b}(w)$ *is monotone increasing.*

Theorem 2 $\lim_{w\to\infty} \phi_{a,b}(w) = 2(p/2-a+1)\int_0^\infty v^{-1}g(v)dv$.

For the shrinkage estimator $\delta_\phi(x) = (1-\phi(\|x\|^2)/\|x\|^2)x$, the new minimaxity condition is given. It is noted that the known minimax conditions by Strawderman(1974) and Berger(1975) is following:

1. $\phi(w)$ is nondecreasing.

2. $\phi(w)/w$ is nonincreasing.

3. $0 \leq \phi(w) \leq 2(p-2)/\int_0^\infty vg(v)dv$.

For $\phi_{a,b}(w)$, however, their conditions are not useful because the range of values (a,b) for satisfying the decrease of $\phi_{a,b}(w)/w$ is not suitably derived. Therefore we derive the conditions of minimaxity which do not required the decrease of $\phi_{a,b}(w)/w$.

Theorem 3 *Assume the following conditions:*

1. $\phi(w)$ *is nondecreasing.*

2. $0 \leq \phi(w) \leq 2(p-2)\int_0^\infty v^{p/2-1}g(v)dv/\int_0^\infty v^{p/2}g(v)dv$.

Then $\delta_\phi(x)$*is minimax.*

From Theorem 1, 2 and 3, we have the theorem about minimaxity and admissibility of $\delta_{a,b}(x)$.

Theorem 4 *(admissibility and minimaxity of* $\delta_{a,b}(x)$*)*

- $\delta_{a,b}(x)$ *is admissible when* $a < 1$ *and* $b > -1$.

- $\delta_{a,b}(x)$ *is minimax when* $a_g \leq a < p/2+1$ *and* $b \geq \max(a-2, 0)$, *where*

$$a_g = p/2+1-(p-2)\frac{\int_0^\infty v^{p/2-1}g(v)dv}{\int_0^\infty v^{-1}g(v)dv\int_0^\infty v^{p/2}g(v)dv}.$$

- $\delta_{a,b}(x)$ *is admissible and minimax when* $a_g \leq a < 1$ *and* $b \geq 0$.

Remark 1 *(multivariate-t distribution)*
We easily check that $g(v) = Cv^{m/2-1}\exp(-mv/2)$, *which corresponds to* $g(v)$ *in the case of Multivariate-t distribution, satisfies the condition 2 of Theorem 1. Moreover* a_g *for this distribution is* $p/2+1-(p-2)(m-2)/(p+m-2)$.

一標本モデルでの解析手法の選択法

——分布の探索とブートストラップ法

横浜市立大学理学部数理科学　　白石高章

1　序

対をなすデータ $(y_1, z_1), \cdots, (y_n, z_n)$ から新しい観測値を $x_i \equiv y_i - z_i$ とすれば, $x_1, \cdots, x_n$ は互いに独立で同一の対称な分布に従っていると考えて良い. 一般に, 観測値 $x_1, \cdots, x_n$ は互いに独立で同一の対称な分布に従っているとする. このとき, 分布が正規分布か全く未知か, 正規分布の近傍であるかのそれぞれの場合について, 最良の手法を紹介した. パラメトリック法, ノンパラメトリック法, セミパラメトリック法のいずれを選択したらよいかを正規性の検定法, 分布の探索法, ブートストラップ法などにより解説した.

2　手法の比較

t 検定は観測値が正規分布に従うとき有意確率 (p 値) を t 分布表を使って求めることができる. 観測値が正規分布に従っていることがわからなければ有意確率を計算できず, 検定がおこなえない. しかしながら, $\bar{Y} \equiv \frac{1}{n}\sum_{i=1}^n Y_i = \frac{1}{n}\sum_{i=1}^n \mathrm{sign}(Y_i)|Y_i|$ とおけば,

$$T_S = \sqrt{\frac{n-1}{n}} \cdot \frac{\sum_{i=1}^n \mathrm{sign}(Y_i)|Y_i|}{\sqrt{\sum_{i=1}^n \{\mathrm{sign}(Y_i)|Y_i| - \bar{Y}\}^2}}$$

より, T_S も $\mathrm{sign}(Y_i)$ と $|Y_i|$ の関数. [5] の M 検定の場合と同様に, H_0 の下で $|\boldsymbol{Y}| = |\boldsymbol{y}|$ を与えたときの T_S の条件付き分布 $P_0(T_S \le t \,|\, |\boldsymbol{Y}| = |\boldsymbol{y}|)$ によって有意確率を計算すれば観測値が未知の連続分布に従っていても検定が行える. また, $|\boldsymbol{Y}| = |\boldsymbol{y}|$ を与えたとき T_S は $T_S' = \sum_{i=1}^n \mathrm{sign}(Y_i)|Y_i|$ の狭義増加関数であるので, 条件付き分布による t 検定は, T_S' に基づく条件付き分布による検定と同値である.

条件付き分布による t 検定, ウィルコクソンの順位検定, M 検定を, 検出力により, 比較を行った.

シミュレーションの繰り返し数 1000 回, 有意水準 5%, $n = 15$, $\mu - \mu_0 = 0.6$, $F(x)$ が正規分布 $N(0,1)$, 混合正規分布 $CN \equiv 0.95N(0,1) + 0.05N(0,9)$, 異常値をもつ混合正規分布 $CO \equiv 0.98N(0,1) + 0.02I_5$, ロジステイック分布 $LG(0, \frac{\sqrt{3}}{\pi})$, 両側指数分布 $DE(0, \frac{1}{\sqrt{2}})$ に設定した. この場合の $\psi(x)$ として, $\psi(x) = \max(\min(x, 1.399), -1.399)$ を採用し, 近傍 $\{F(x) = 0.95 \cdot \Phi(x) + 0.05 \cdot H(x) : H(-x) = 1 - H(x)\}$ に対して最良になっている. 表 3 から, 観測値の従う分布が $DE(0, \frac{1}{\sqrt{2}})$ のように正規分布からかなり離れた分布以外はそれほど検出力に差はない.

次に, 正規分布のときの最良推定量 $\tilde{\mu}$, ホッジス • レーマン推定量 $\hat{\mu}$, M 推定量 $\breve{\mu}$ を比較する. $\tilde{\mu}$ に対する $\hat{\mu}$ の相対効率, $\tilde{\mu}$ に対する $\breve{\mu}$ の相対効率, $\hat{\mu}$ に対する $\breve{\mu}$ の相対効率はパラメータ (μ, σ) に依存しない.

$$e(\hat{\mu}, \tilde{\mu}) = \frac{E(\tilde{\mu} - \mu)^2}{E(\hat{\mu} - \mu)^2}, \quad e(\breve{\mu}, \tilde{\mu}) = \frac{E(\tilde{\mu} - \mu)^2}{E(\breve{\mu} - \mu)^2}, \quad e(\breve{\mu}, \hat{\mu}) = \frac{E(\hat{\mu} - \mu)^2}{E(\breve{\mu} - \mu)^2}$$

を繰り返し数 1000 回のシミュレーション実験で求めた. これらにより, 以下の結論を得る.

○ 観測値が正規分布に従っている場合は, 頑健な手法は最良な手法に比べてほんの少し劣り, ノンパラメトリック法が最もよくない.

○ 観測値が混合正規分布 $0.95N(0,1)+0.05N(0,9)$, 異常値をもつ混合正規分布 $0.98N(0,1)+0.02I_5$, ロジステイック分布 $LG(0,\frac{\sqrt{3}}{\pi})$ などの正規分布に近い分布に従っている場合は, 頑健な手法が最も良く, 正規母集団での最良手法は劣る.

○ 観測値が両側指数分布 $DE(0,\frac{1}{\sqrt{2}})$ などの正規分布からかなり離れた分布に従っている場合は, ノンパラメトリック法が最も良く, 正規母集団での最良手法は非常に劣る. 頑健な手法はノンパラメトリック法にくらべれば劣るが, 正規母集団での最良手法よりも非常に良い.

3 正規性の検定と分布の探索

カイ二乗適合度検定やコルモゴロフ正規性適合度検定は観測値が正規分布でないときでも n が大きくないと正規性を棄却できないことが多い. 正規母集団での最良手法, ノンパラメトリック法, 頑健な手法により解析結果が微妙に異なった. また正規性の検定も棄却できず, 3 種の手法のどれを最も信頼すべきかはっきりしない. ブートストラップによる選択法もまだ改良の余地が多いことも指摘できた.

4 一標本解析のソフト

正規母集団での最良手法, ノンパラメトリック法, M 統計量に基づく頑健手法, 正規性の検定法, データの従っている分布に近い分布を探す, 経験分布関数と正規分布関数の重ねがきグラフ, 経験分布関数とロジスティック分布関数の重ねがき, 経験分布関数と両側指数分布関数の重ねがきグラフ, ブートストラップによる一標本手法を選択するソフトの紹介を行った.

現在のところ正規性の検定, 分布の探索, 経験分布と分布関数, ブートストラップによる手法の選択法などの総合判断によりパラメトリック法, ノンパラメトリック法, セミパラメトリック法の解析手法を選択こととなった.

参考文献

[1] 新井親夫 (1998). Fortran90 入門.
[2] 白石高章 (1998). 二標本モデルにおける推測法の比較. 日本計算機統計学会和文誌 11 巻 p13-24
[3] Good, P. (1993). Permutation Tests. Springer.
[4] Greenwood, P.E. and Nikulin, M.S. (1996). A Guide to Chi-Squared Testing. Wiley & Sons, Inc.
[5] Hajek,J , Sidak, Z, and Sen, P.K. (1999). Theory of Rank Tests. Academic Press.
[6] Shiraishi, T. (1991). Hypothesis testing and parameter estimation based on M-statistics in k samples with unequal variances. Metrika, 38, p163-178.
[7] Shiraishi, T. (1993). Statistical procedures based on signed ranks in k samples with unequal variances. Ann. Inst. Statist. Math., 45, p265-278.
[8] Shiraishi, T. (1996). On scale-invariant M-statistics in multivariate k samples. J. Japan Statist. Soc., 26, p241-253.

Problems in Orientational Estimation

香川大学工学部　筑瀬　靖子

1. Introduction

We are concerned with some problems in estimating the orientation parameters of the matrix Langevin distributions defined on the special manifolds of our interest, the Stiefel and the Grassmann manifolds. The matrix Langevin distributions are most commonly used and tractable distributions on the manifolds, corresponding to the normal distributions on the Euclidean spaces.

The Stiefel manifold $V_{k,m}$ is represented by the set of $m \times k$ matrices X such that $X'X = I_k$, the $k \times k$ identity matrix ($k \leqq m$); we have the hypersphere $V_{1,m}$ and the orthogonal group $O(m)=V_{m,m}$ as practical examples.

The matrix Langevin $L(m,k;F)$ distribution on $V_{k,m}$ is defined by the density function

$$\exp\,(tr\, F'X)\ /\ {}_0F_1(\tfrac{1}{2}m;\tfrac{1}{4}F'F), \tag{1}$$

and we let the unique singular value decomposition of the $m \times k$ parameter matrix F be $F = \Gamma\Lambda\Theta'$ of known rank p ($p \leqq k$), where $\Gamma \in \tilde{V}_{p,m}$, $\Theta \in V_{p,k}$ and $\Lambda = \mathrm{diag}\,(\lambda_1,..,\lambda_p)$, $\lambda_1 \geq \ldots \geq \lambda_p > 0$.

2. Estimation of the Orientation Parameters Γ and Θ

2.1. Maximum Likelihood Estimators (M.L.E.'s)

Given a random sample $\underline{X} = (X_1,\ldots,X_n)$ from the $L(m,k;F)$ distribution with $\bar{X} = \sum_{i=1}^{n} X_i / n$, the m.l.e.'s $\hat{\Gamma}$ and $\hat{\Theta}$ of Γ and Θ are given by the matrix of the latent vectors of $\bar{X}\bar{X}'$ and that of $\bar{X}'\bar{X}$, corresponding to the first p largest latent roots of $\bar{X}'\bar{X}$, respectively.

2.2. Bayes Estimators

Employing suitable (joint) conjugate prior density functions of Γ and Θ, the Bayes estimators of Γ and Θ with Λ known are given by the posterior modal orientations, which are also obtained from the minimum –loss criterion with suitable loss functions.

2.3. Optimality Properties

2.3.1. Estimating Γ with Θ and Λ known

We may be interested in a general family of rotationally symmetric distributions around Γ, which are invariant under the simultaneous transformations

$$X \rightarrow HX \text{ and } \Gamma \rightarrow H\Gamma \text{ for } H \in O(m), \tag{2}$$

that is , having the density function of the from $f(\Gamma' X)$. The estimator $\hat{\Gamma}(\underline{X})$ of Γ is said to be equivariant under (2) if $\hat{\Gamma}(H\underline{X}) = H\hat{\Gamma}(\underline{X})$. We consider the problem of invariant estimation under the loss function $L(\Gamma,\hat{\Gamma})$ satisfying $L(H\Gamma,H\hat{\Gamma}) = L(\Gamma,\hat{\Gamma}) = \rho(\Gamma'\hat{\Gamma})$.

The m.l.e. of Γ is equivariant. Note that with the rotationally symmetric distribution we have $E(\hat{\Gamma}) = \Gamma E(\Gamma'\hat{\Gamma})$, with $B = E(\Gamma'\hat{\Gamma})$ a bias of $\hat{\Gamma}$ and the unbiasedness ($B = I_p$) is not a useful optimality property . We discuss the optimality properties of minimum risk equivariance, admissibility, minimaxity and efficiency of the estimation problem.

2 . 3 . 2 . Estimating Γ and Θ with Λ known

We develop the invariant estimation of Γ and Θ with a general family of rotationally symmetric distributions invariant under the simultaneous transformations $X \rightarrow H_1XH_2'$, $\Gamma \rightarrow H_1\Gamma$ and $\Theta \rightarrow H_2\Theta$ for $H_1 \in O(m)$ and $H_2 \in O(k)$, that is , having the density function of the from $f(\Gamma' X \Theta)$.

2. 4 Sufficiency and Ancillarity

When we are interested in the orientation parameters Γ and Θ with Λ known for $p \underset{\neq}{<} k$, we need appropriate conditioning from the conditionality principle. We use the conditional distribution of the m. l. e.'s of the main parameters for given suitable ancillary statistics.

3 . Remarks

The Grassmann manifold consisting of the k-dimensional hyperplanes in $R^m (k \leqq m)$ is represented by the manifold $P_{k,m-k}$ of $m \times m$ orthogonal projection matrices P idempotent of rank k. We can carry out similar kinds of discussions of the estimation problem of the orientation parameter Γ of the matrix Langevin distribution with density function $exp(tr\ BP)/{}_1F_1\ (\frac{1}{2}k;\ \frac{1}{2}m;B)$, for $B = \Gamma\Lambda\Gamma'$ (spectral decomposition) an $m \times m$ symmetric matrix with certain identifiability restrictions on B.

LINEX LOSS FUNCTION AND STATISTICAL PREDICTION

肖 玉山 (熊本大・自然) 高田 佳和 (熊本大・理)

1. Introduction

We discuss a prediction problem under a LINEX loss function. A concept of risk unbiasedness is introduced to the prediction problem. Some results about LINEX-unbiased predictor are derived and the best LINEX-unbiased predictor is given. we also consider an application of the concept of the Pitman's measure of closeness (PMC) to a statistical prediction problem. Under some conditions, the Pitman-closest predictor is determined.

We suppose that X is observed random vector and Y a future real random variable, and the joint distribution of X and Y depends on an unknown parameter θ. After observing $X = x$, we want to predict the value of Y. A non-negative loss function $L(d, y)$ represents the loss of predicting $Y = y$ by d. Let $\delta(X)$ be a predictor of Y and

$$R(\theta, \delta) = E_\theta \{L(\delta(X), Y)\}.$$

be the risk function.

2. LINEX unbiasedness

DEFINITION 1 *If a predictor $\delta(X)$ satisfies*

$$E_\theta \{L(\delta(X), Y)\} = \min_c E_\theta \{L(\delta(X), Y + c)\},$$

where c is a real number, then it is called a risk-unbiased predictor. If a risk-unbiased predictor minimizes the risk for all values of θ, then it is called the best risk-unbiased predictor.

Now we shall give results which characterize the risk-unbiased predictor (cf. Lemma 1 of Klebanov, 1974).

THEOREM 1 *Under squared error loss, $\delta(X)$ is risk-unbiased if and only if*

$$E_\theta \delta(X) = E_\theta Y.$$

Under a squared error loss the risk-unbiased predictor is called mean-unbiased. A LINEX loss function is defined by

$$L(d, y) = \exp[\alpha(d - y)] - \alpha(d - y) - 1, \quad \alpha \neq 0.$$

A risk-unbiased predictor with respect to the LINEX loss is called LINEX-unbiased.

THEOREM 2 *A predictor $\delta(X)$ is LINEX-unbiased if and only if*

$$E_\theta \left\{\exp[\alpha(\delta(X) - Y)]\right\} = 1.$$

for all θ.

THEOREM 3 *If δ is a LINEX-unbiased, then it can not be mean-unbiased, unless $P_\theta(\delta(X) = Y) = 1$. The converse is also true.*

By using of an adequate statistic, we give the Rao-Blackwell theorem with regard to a prediction problem.

THEOREM 4 *If $\delta(X)$ is LINEX-unbiased and T is adequate, then*

$$\delta^*(T) = \frac{1}{\alpha} \log E[\exp(\alpha\delta(X))|T]$$

is also LINEX-unbiased and the risk of δ^* is less than that of δ.

THEOREM 5 *If T is adequate and complete, and*

$$E_\theta(\exp(-\alpha Y)|T) = Q(\theta)K(T) \quad a.e.$$

for some non-zero functions Q and K, then δ^* in Theorem 4 is the unique best LINEX-unbiased predictor.

3. Pitman's measure of closeness

Let C be a family of predictors. Then $\delta \in C$ is said to be the Pitman-closest in C with respect to L if δ is better than any other $\delta' \in C$ under PMC with respect to L.

Let δ_M be a median unbiased predictor of Y and T a statistic based on X. Under the Pitman's measure of closeness, we consider the prediction problem of Y within the following family

$$C = \{\delta; \delta(X) = \delta_M(X) + Z(T)\},$$

where $Z = Z(T)$ is any function of T.

THEOREM 6 *Suppose that $Y - \delta_M$ is independent of T. Then δ_M is Pitman-closest in the family C with respect to the LINEX loss.*

多次元exponential familyのAPOruleについて

大阪府立大学・工学部
長尾壽夫

xをp次元ユークリッド空間R^pの点とし、λを非退化sigma-有限測度とし、分布族$F_\omega(x)$, $(\omega \in \Omega^k)$をR^p上の確率分布とすると

$$dF_\omega(x) = \exp\{\omega' b(x) - \psi(\omega)\} d\lambda(x)$$

のように表される分布族を多次元多母数指数分布族ということにする。すると多項分布とか多次元正規分布はこれらのクラスに入っている。
ここで$\omega = (\omega_1, \cdots, \omega_k)'$, $b(x) = (b_1(x), \cdots, b_k(x))'$は$k$次元パラメータとする。また$\Omega^k = \Omega_1 \times \cdots \times \Omega_k$となる$k$次元部分空間とし、$\Omega_i = (\underline{\omega}_i, \bar{\omega}_i)$ $i = 1, \cdots, k$ とする。このとき、

$$\theta_i = \mathrm{E}(b_i(x)|\omega) = \frac{\partial}{\partial \omega_i}\psi(\omega)$$

$$\mathrm{Var}(b(x)|\omega) = (\frac{\partial^2}{\partial \omega_i \partial \omega_j}\psi(\omega)) = (\psi_{ij}).$$

となる。
つぎに、上の密度関数からの標本を$x_1, \cdots, x_n$とする。このとき、尤度関数を$L_n(\omega)$とする。$\mathcal{F}_0$は自明なσ-加法族とし、$\mathcal{F}_n (n \geq 1)$は$x_1, \cdots, x_n$によって生成される$\sigma-$加法族とする。パラメータωは事前ルベーグ密度$\xi_0(\omega)$を持つ確率変数とする。その分布$\pi_0(\omega)$は

$$d\pi_0(\omega) = \xi_0(\omega) d\omega.$$

このとき、$\mathcal{F}_n$が与えられたときのωの事後密度は、

$$d\pi_n(\omega) = c_n^{-1} L_n(\omega) \xi_0(\omega) d\omega = \xi_n(\omega) d\omega$$

で与えられる。ここでc_nは正規化定数である。$p \leq k$ とする。ここで、

$$\mathrm{E}\{(\theta - \hat{\theta}_t)'(\theta - \hat{\theta}_t) + ct\}$$

を最小にさせる$\hat{\theta}_t$ を見つける。ただし、$\theta = (\theta_1, \cdots, \theta_p)'$ とする。するとよく知られたように、$\hat{\theta}_n = \mathrm{E}(\theta|\mathcal{F}_n)$ で与えられる。そこで、

$$R_n(c) = \mathrm{E}\{(\theta - \hat{\theta}_n)'(\theta - \hat{\theta}_n)\} + cn$$

を具体的に求めると、

$$R_n(c) = \frac{1}{n}U_n + \frac{1}{n^2}M_n + cn = 2(cU_n)^{\frac{1}{2}} + n^{-1}(U_n - \sqrt{c}n)^2 + n^{-2}M_n$$

となる。ただし、$U_n = \mathrm{E}(\sum_{i=1}^p \psi_{ii}|\mathcal{F}_n)$ であり、$M_n = \mathrm{E}\{\sum_{i=1}^p (\frac{\partial^2}{\partial \omega_i^2}\xi_0(\omega))/\xi_0(\omega)|\mathcal{F}_n\}$
$- \sum_{i=1}^p \{\mathrm{E}(\frac{\partial}{\partial \omega_i}\log \xi_0(\omega)|\mathcal{F}_n)\}^2$である。APO ruleとして、次を定める。

$$t = \inf\{n | U_n < cn^2\}.$$

$$R_t(c) = 2\sqrt{c}\mathrm{E}((\sum_{i=1}^{p}\psi_{ii})^{1/2}|\mathcal{F}_t) + 2\sqrt{c}\frac{W_t}{\sqrt{U_t + V_t}} + t^{-2}M_t + t^{-1}(\sqrt{U_t} - \sqrt{c}t)^2$$

の各項について考える。

$$\mathrm{E}\{(\mathrm{E}(\sum_{i=1}^{p}\psi_{ii})^{1/2})|\mathcal{F}_t)\} = \mathrm{E}((\sum_{i=1}^{p}\psi_{ii})^{1/2}))$$

となる。

$$\frac{1}{\sqrt{c}}\frac{W_t}{\sqrt{U_t + V_t}} \to \frac{(gradg(\omega))'}{2(\sum_{i=1}^{p}\psi_{ii})}T_k^{-1}(\omega)(gradg(\omega))$$

である。次に$\dfrac{M_t}{ct^2}$の項について考える。

$$\frac{|M_t|}{ct^2} \leq \frac{|M_t|}{U_t} \leq \frac{1}{2}(U_t^{-2} + M_t^2)$$

であるから、上の２項が一様可積を示せばよい。

$$\mathrm{E}(\frac{M_t}{ct^2}) \to \mathrm{E}(\frac{1}{\sum_{i=1}^{p}\frac{\partial^2}{\partial\omega_i^2}\psi(\omega)}\{\sum_{i=1}^{p}(\frac{\partial^2}{\partial\omega_i^2}\xi_0(\omega))/\xi_0(\omega) - \sum_{i=1}^{p}(\frac{\partial}{\partial\omega_i}\log\xi_0(\omega))^2\})$$

$$\frac{1}{ct}(\sqrt{c}t - \sqrt{U_t})^2 \leq 2\{\frac{1}{t} + tU_t^{-2}(U_{(t-1)} - U_t)^2\}$$

ここで$\mathrm{E}\{\sup_n nU_n^{-2}(U_{(n-1)} - U_n)^2\} < \infty$　となるようにpriorを定める。
次に最良のstopping time aについて考察する。
$a = n$のとき、$\mathrm{E}\{R_{n+1}(c)|\mathcal{F}_n\} \geq R_n(c)$ である

$$M_n = M_n^{(1)} + M_n^{(2)}$$

ただし、$M_n^{(1)} = \sum_{i=1}^{p}\mathrm{E}(\frac{\partial^2\rho_0(\omega)}{\partial\omega_i^2}|\mathcal{F}_n)$ であり $M_n^{(2)} = \sum_{i=1}^{p}\mathrm{Var}(\frac{\partial\rho_0(\omega)}{\partial\omega_i}|\mathcal{F}_n)$. よって、

$$c \geq \frac{1}{n(n+1)}U_n + \frac{2n+1}{n^2(n+1)^2}M_n^{(1)} + \frac{2n+1}{n^2(n+1)^2}M_n^{(2)} \geq \frac{1}{n(n+1)}U_n + \frac{2n+1}{n^2(n+1)^2}M_n^{(1)}$$

また$a = n$のとき、$E(R_n(c)|\mathcal{F}_{n-1}) \leq R_{n-1}(c)$であるから、同様に計算し直すことによって、$ca^2 \to \sum_{i=1}^{p}\psi_{ii}$ となる。そこで
定理。tをAPO rule, aをoptimal stopping とすると、sをtかaとすると、

$$\mathrm{E}\{(\theta - \hat{\theta}_s)'(\theta - \hat{\theta}_s) + cs\} = 2\sqrt{c}V_0 + cM_0 + o(c)$$

ただし、

$$V_0 = \mathrm{E}\{(\sum_{i=1}^{p}\psi_{ii})^{1/2}),\quad M_0 = \mathrm{E}\{\frac{1}{(\sum_{i=1}^{p}\psi_{ii})}(gradg(\omega))'T_k^{-1}(\omega)(gradg(\omega))$$

$$+ \sum_{i=1}^{p}(\frac{\partial^2}{\partial\omega_i^2}\xi_0(\omega))/\xi_0(\omega) - \sum_{i=1}^{p}(\frac{\partial}{\partial\omega_i}\log\xi_0(\omega))^2\})$$

コントロールをもつ場合の最適な二段階法について

筑波大学 数学系　　青木　充
筑波大学 数学系　　青嶋　誠

$k+1$ 個の独立な正規母集団 $\pi_i : N(\mu_i, \sigma^2)$, $i=0,...,k$ を仮定する．ただし，π_0 は control 母集団とし，π_i, $i=1,...,k$ を test 母集団とする．ここで，全てのパラメータは未知で，分散は共通である．$\boldsymbol{\mu}=(\mu_0,\mu_1,...,\mu_k)$ とおく．任意の固定された定数 $\delta_1^\star$ と $\delta_2^\star$ $(0<\delta_1^\star<\delta_2^\star)$ に対して，母集団の集合 $\Omega=(\pi_1,...,\pi_k)$ は，$\Omega=\Omega_B\cup\Omega_I\cup\Omega_G$，ただし，

$$\Omega_B=(\pi_i:\mu_i\le\mu_0+\delta_1^\star),\ \Omega_I=(\pi_i:\mu_0+\delta_1^\star<\mu_i<\mu_0+\delta_2^\star),\ \Omega_G=(\pi_i:\mu_i\ge\mu_0+\delta_2^\star)$$

のように3つの部分集合に分割される．この Ω を，$S_B\supset\Omega_B$ と $S_G\supset\Omega_G$ なる2つの部分集合 $S_B\cup S_G=\Omega$ に分割することが目的である．そのような正しい決定がなされたとき，Correct Decision (CD) と呼ぶことにする．ここでは，この目的に対して，定数 $P^\star$ $(2^{-k}<P^\star<1)$ を指定して，

$$P(CD|\boldsymbol{\mu},\sigma^2)\ge P^\star \quad \text{for} \quad {}^\forall(\boldsymbol{\mu},\sigma^2) \tag{1.1}$$

を最小の標本数で満たす最適な推測法を構築することを考える．

π_0 から大きさ n_0 の，π_i $(i=1,...,k)$ からは大きさ c^2n_0 $(0<c\le1)$ の無作為標本を抽出する．そのとき，標本の総数を n とすると $n_0=A^2n$ と書ける．ただし，$A=A(c)=(1+kc^2)^{-1/2}$ である．標本平均 $\overline{X}_{0(n)}=\sum_{j=1}^{A^2n}X_{0j}/(A^2n)$, $\overline{X}_{i(n)}=\sum_{j=1}^{c^2A^2n}X_{ij}/(c^2A^2n)$, $i=1,...,k$ を計算し，定数 $d\,(>0)$ に対してルール

$$S_B=(\pi_i:\overline{X}_{i(n)}-\overline{X}_{0(n)}<d),\quad S_G=(\pi_i:\overline{X}_{i(n)}-\overline{X}_{0(n)}\ge d) \tag{1.2}$$

を定義する．ここで，(c,d,n) は (1.1) の要求を満足するように決定される．このルールのもと，

$$\max_{c\in(0,1]}\max_{d>0}\inf_{\boldsymbol{\mu},\sigma^2}P(CD|\boldsymbol{\mu},\sigma^2)\ge P^\star \tag{1.3}$$

なる意味で標本数の総和を最小にする最適な推測法 $R(c_0,d_0,N)$ を提案する．

i) $k=1$ $(r=0,1)$ の場合

固定された初期標本数 $m\,(>2)$ に対して，control 母集団と test 母集団から，独立にそれぞれ大きさ A^2m と c^2A^2m の初期標本を抽出して，プールされた不偏標本分散 S_m^2 を計算する．二段階法の標本数を

$$N=\max\left\{m,\left[\frac{4a_m^2}{(\delta_2^\star-\delta_1^\star)^2}S_m^2\right]+1\right\} \tag{1.4}$$

と定義する．ここで，a_m は積分方程式 $F_\nu(ca_m/(1+c^2))=P^\star$ の解である．ただし，$F_\nu(\cdot)$ は自由度 $\nu=m-2$ の t 分布の分布関数である．2段階目として，control 母集団と test 母集団から，独立にそれぞれ大きさ $A^2(N-m)$ と $c^2A^2(N-m)$ の追加標本を抽出する．各々の母集団で，初期標本と追加標本を合わせて標本平均 $\overline{X}_{0(N)}$ と $\overline{X}_{1(N)}$ を計算する．そのとき，n を N に置き換えたルール (1.2) を考える．d の最適値 $d_0=(\delta_1^\star+\delta_2^\star)/2$ と c の最適値 $c_0=1$ を用いて，二段階法 (1.4) にもとづく推測法 $R(1,d_0,N)$ を提案する．

ii) k が偶数の場合

$k=1$ のときと同様に，各々の母集団から初期標本を抽出し，プールされた不偏標本分散 S_m^2 を計算する．二段階法の標本数を，(1.4) にある a_m を b_m に置き換えたもので定義する．各母集団で追加標本を抽出し，$\overline{X}_{0(N)}$ と $\overline{X}_{i(N)}$, $i=1,...,k$, を計算して，n を N で置き換えたルール (1.2) を考える．ここで，b_m は積分方程式

$$\int_0^\infty\int_{-\infty}^\infty\Phi^\ell\Big(c(x+Ab_my)\Big)\Phi^\ell\Big(c(-x+Ab_my)\Big)d\Phi(x)dG_\nu(y)=P^\star \tag{1.5}$$

の解として定め，最適な c 値は

$$\int_0^\infty\int_{-\infty}^\infty\Phi^{\ell-1}(x)\Phi^\ell(-x+2b_mcAy)(x-b_mkc^3A^3y)\phi(x/c-b_mAy)d\Phi(x)dG_\nu(y)=0 \tag{1.6}$$

の解として決定される．ただし，$G_\nu(\cdot)$ は $\sqrt{\chi_\nu^2/\nu}$ の分布関数，χ_ν^2 は自由度 $\nu = m-k-1$ のカイ二乗分布に従う確率変数である．(1.5)-(1.6) の解 (b_m, c_m) を用いた二段階法 (1.4) と d の最適値 $d_0 = (\delta_1^\star + \delta_2^\star)/2$ にもとづく推測法 $R(c_m, d_0, N)$ を提案する．

iii) k が3以上の奇数の場合

二段階法の標本数を，k が偶数の場合と同様に定義する．ただし $b_m = (b_{1m} + b_{2m})/2$ とおく．n を N で置き換えて，$d = d_0 + (\delta_2^\star - \delta_1^\star)(b_{1m} - b_{2m})\{2(b_{1m} + b_{2m})\}^{-1}(\equiv d_m)$ をもつルール (1.2)を考える．ここで，b_m と $d = d_m$の定義式にある (b_{1m}, b_{2m})は，積分方程式

$$\int_0^\infty \int_{-\infty}^\infty \Phi^\ell(c\,h^+)\Phi^{k-\ell}(c\,h^-)d\Phi(x)dG_\nu(y) = P^\star \tag{1.7}$$

ならびに

$$\begin{aligned}&\ell\int_0^\infty \int_{-\infty}^\infty \Phi^{\ell-1}(c\,h^+)\Phi^{k-\ell}(c\,h^-)y\phi(c\,h^+)d\Phi(x)dG_\nu(y)\\ &= (k-\ell)\int_0^\infty \int_{-\infty}^\infty \Phi^\ell(c\,h^+)\Phi^{k-\ell-1}(c\,h^-)y\phi(c\,h^-)d\Phi(x)dG_\nu(y)\end{aligned} \tag{1.8}$$

の連立方程式の解として定める．ただし，$h^+ = x + Ab_{1m}y$，$h^- = -x + Ab_{2m}y$ である．なお，d_mに関して，関係式

$$\frac{d_m}{\delta_2^\star - \delta_1^\star} = d_{m0} + \frac{2-\gamma}{\gamma - 1} \tag{1.9}$$

が任意の $\delta_1^\star$ と $\delta_2^\star$ $(0 < \delta_1^\star < \delta_2^\star)$ に対して成り立つ．ここで，$\gamma = \delta_2^\star/\delta_1^\star$，$d_{m0} = 1.5 + (b_{1m} - b_{2m})\{2(b_{1m} + b_{2m})\}^{-1}$ である．最適な c 値は，$q = -x + cA(b_{1m} + b_{2m})y$ とすると

$$\begin{aligned}&\ell\int_0^\infty \int_{-\infty}^\infty \Phi^{\ell-1}(x)\Phi^{k-\ell}(q)(x - b_{1m}kc^3A^3y)\phi(x/c - b_{1m}Ay)d\Phi(x)dG_\nu(y) + (k-\ell)\times\\ &\int_0^\infty \int_{-\infty}^\infty \Phi^{k-\ell-1}(x)\Phi^\ell(q)(x - b_{2m}kc^3A^3y)\phi(x/c - b_{2m}Ay)d\Phi(x)dG_\nu(y) = 0\end{aligned} \tag{1.10}$$

の解として決定される．(1.7)-(1.8), (1.10) の解 (b_{1m}, b_{2m}, c_m) を用いた二段階法 (1.4) と (1.9) 式から得られる d の最適値 d_m にもとづく推測法 $R(c_m, d_m, N)$ を提案する．

なお，control 母集団と test 母集団の平均の差の同時信頼区間に関しても，標本数を

$$N = \max\left\{m,\ \left[\frac{t_m^2 S_m^2}{d^2}\right] + 1\right\} \tag{1.11}$$

と定義し，t_m を

$$\int_0^\infty \int_{-\infty}^\infty \left\{\Phi\Big(c(x + At_my)\Big) - \Phi\Big(c(x - At_my)\Big)\right\}^k d\Phi(x)dG_\nu(y)dxdy = 1-\alpha \tag{1.12}$$

と定め，最適な c 値を

$$\int_0^\infty \int_{-\infty}^\infty \{\Phi(x) - \Phi(x - 2t_mcAy)\}^{k-1}(x - t_mkc^3A^3y)\phi(x/c - t_mAy)d\Phi(x)dG_\nu(y) = 0$$

の解として決定する二段階法を用いることで，最適な推測法を構築することができる．