(10)「統計的推測理論とその応用—正則から非正則へ—」に関する研究報告

Missing, pooling を伴う離散サンプリングにおける MLE と その許容性

関東学院大経済学部　布能英一郎

1.　本報告では、次の２つの定理を示し、更にこれに関連した問題を考察した。なお、以下、母集団からのランダムサンプリングを「観測」と略記する。

定理１

第１回目の観測　$(X_{10}, X_{11}, ..., X_{1k-1}, X_{1k}) \sim Multinomial(N_1, \theta_0, \theta_1, ..., \theta_{k-1}, \theta_k)$

第２回目の観測　$(X_{20}, X_{21}, ..., X_{2k-1}) \sim Multinomial(N_2, \frac{\theta_0}{\theta_0 + .. + \theta_{k-1}}, ..., \frac{\theta_{k-1}}{\theta_0 + .. + \theta_{k-1}})$

$\vdots$

第 k-1 回目の観測　$(X_{k-1\,0}, ..., X_{k-1\,2}) \sim Multinomial(N_{k-1}, \frac{\theta_0}{\theta_0 + \theta_1 + \theta_2}, ..., \frac{\theta_2}{\theta_0 + \theta_1 + \theta_2})$

第 k 回目の観測　$(X_{k0}, X_{k1}) \sim Binomial(N_k, \frac{\theta_0}{\theta_0 + \theta_1}, \frac{\theta_1}{\theta_0 + \theta_1})$

第１回目の観測で $X_{10} = X_{11} = \ldots = X_{1\,k-1} = 0$ の場合、第２回目以降の観測は行わない。第２回目までの観測 で $X_{10} = X_{20} = \ldots = X_{1\,k-2} = X_{2\,k-2} = 0$ の場合、第３回目以降の観測は行わない。以下、同様の状況を仮定する。そうすると、θ_i (i=0,1,2,...,k) の MLE は、自乗損失下で許容的。

定理２

第１回目の観測　$(X_{1\,0}, X_{1\,1}, ..., X_{1\,k-1}, X_{1k}) \sim Multinomial(N_1, \theta_0, \theta_1, ..., \theta_{k-1}, \theta_k)$

第２回目の観測　$(X_{2\,0}, X_{2\,1}, ..., X_{2\,k-1}) \sim Multinomial(N_2, \theta_0, \theta_1, ..., \theta_{k-2}, \theta_{k-1} + \theta_k)$

$\vdots$

第 k-1 回目の観測　$(X_{k-1\,0}, X_{k-1\,1}, X_{k-1\,2}) \sim Multinomial(N_{k-1}, \theta_0, \theta_1, \theta_2 + \theta_3 + ... + \theta_k)$

第 k 回目の観測　$(X_{k\,0}, X_{k\,1}) \sim Binomial(N_k, \theta_0, \theta_1 + ... + \theta_k)$

このとき、θ_i, i=0,1,2,...,k の MLE は、自乗損失下で許容的。

2.　定理１および定理２の証明には、stepwise Bayesian procedure を用いた。

3.　定理１，定理２とも、多項サンプリングにおいてカテゴリーが単調に消滅してゆくものである。定理１は、あるカテゴリーが消滅後、他のカテゴリーに比例再配分されるというモデルであり、定理２は、別のカテゴリー吸収されるというモデルである。これらの拡張として、(1) 先に比例再配分、後に吸収が行われる。またその逆, (2) 比例再配分と吸収が同時に

行われる, (3) 比例再配分と吸収が convex combination で行われる，といったモデルが考えられる。以下、簡単のため、一般論ではなくて、例で考察した。

例 1 第 1 回目の観測 $(X_0, X_1, X_2, X_3) \sim Multinomial(N_1, \theta_0, \theta_1, \theta_2, \theta_3)$
第 2 回目の観測 $(Y_1, Y_2, Y_3) \sim Multinomial(N_2,\ \theta_0 + \theta_1, \theta_2, \theta_3)$
第 3 回目の観測 $(Z_1, Z_2) \sim Multinomial(N_3,\ (\theta_0 + \theta_1)/(\theta_0 + \theta_1 + \theta_2), \theta_2/(\theta_0 + \theta_1 + \theta_2))$

例 2 第 1 回目の観測 $(X_0, X_1, X_2, X_3, X_4) \sim Multinomial(N_1, \theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$
第 2 回目の観測 $(Y_1, Y_2, Y_3) \sim Multinomial(N_2,\ \zeta_1, \zeta_2, \zeta_3)$ where

$$\zeta_1 = \theta_0 + \theta_1, \quad \zeta_2 = (\theta_2 + \theta_3 + \theta_4)\frac{\theta_2}{\theta_2 + \theta_3}, \quad \zeta_3 = (\theta_2 + \theta_3 + \theta_4)\frac{\theta_3}{\theta_2 + \theta_3}.$$

これらのモデルの下での θ_i の MLE は 容易に求められる。更に、自乗損失下での許容性が、定理 1 、定理 2 と同様の証明法で出来た。

例 3 第 1 回目の観測 $(X_0, X_1, X_2, X_3) \sim Multinomial(N_1, \theta_0, \theta_1, \theta_2, \theta_3)$
第 2 回目の観測 $(Y_0, Y_1, Y_2) \sim Multinomial(N_2,\ \zeta_0, \zeta_1, \zeta_2)$

但し $$\zeta_0 = \xi\frac{\theta_0}{\theta_0 + \theta_1 + \theta_2} + (1 - \xi)\theta_0, \quad \zeta_1 = \xi\frac{\theta_1}{\theta_0 + \theta_1 + \theta_2} + (1 - \xi)\theta_1,$$
$$\zeta_2 = \xi\frac{\theta_2}{\theta_0 + \theta_1 + \theta_2} + (1 - \xi)(\theta_2 + \theta_3).$$

このモデルで ξ がパラメーターのとき、MLE は

$$\hat{\theta}_0 = \frac{x_0 + x_1}{N_1}\frac{x_0 + y_0}{x_0 + x_1 + y_0 + y_1} \quad \hat{\theta}_1 = \frac{x_0 + x_1}{N_1}\frac{x_1 + y_1}{x_0 + x_1 + y_0 + y_1} \quad \hat{\theta}_2 = \frac{x_2}{N_1} \quad \hat{\theta}_3 = \frac{x_3}{N_1}$$

$$\hat{\xi} = \frac{(x_0 + x_1 + x_2)\{(x_2 + x_3)(y_0 + y_1) - (x_0 + x_1)y_2\}}{x_3(x_0 + x_1)(y_0 + y_1 + y_2)}$$

である。この MLE が自乗損失下で許容的か否かは、まだわかっていない。 他方、$0 < \xi < 1$ が fixed number のとき、θ_i の MLE は explicit に求められない。許容性についても、全くわかっていない。

4. 定理 1 , 2 では、多項分布に従うことを仮定した。しかし、多項分布を含むような更に広い分布族に対しても、定理 1 , 2 と同様の結果が得られた。ポアソン分布に従う場合の考察も、若干行った。

References

Asano, C. (1965). On estimating multinomial probabilities by pooling incomplete samples, *Annals of the Institute of Statistical Mathematics* **17**, 1-13.

Ghosh, M. and Meeden, G. (1997). *Bayesian methods for finite population sampling,* Chapman and Hall.

Meeden, G., Ghosh, M., Srinivasan. C., and Vardeman, S. (1989). The admissibility of the Kaplan-Meyer and other maximum likelihood estimators in the presence of censoring, *Annals of Statistics* **17**, 1509-1531.

対照漁具の選択性を考慮した選択性実験について

東京水産大学　北門利英

1. はじめに

漁具の選択性とは通常，漁具に遭遇した魚がその漁具で漁獲される確率で定義される．この確率を魚体長の関数とみなした場合，これを選択性曲線と呼ぶ．典型的な選択性実験では，実験漁具と対照漁具を同時に用い，体長階級毎の漁獲され易さの違いが情報となって実験漁具の選択性曲線が推定される．例えば曳網のカバーネット試験ではカバーネットが対照漁具となるが，この対照漁具には実験漁具に比べて目合の小さいものが用いられ，非選択的，すなわち選択率が1と仮定される事が多い．一方で，対照漁具の目合を小さくしすぎると実際の漁業とは異なる作用をもたらす可能性もある．従って，実際の実験では適当な目合サイズの対照漁具を用いる為，その選択性を必ずしも無視できない状況もある．しかしながら，対照漁具に選択性を仮定する状況ではパラメータを増やす事になり，必ずしも安定して実験漁具の選択性曲線のパラメータを推定できるとは限らない．そこで本研究では曳網の選択性実験において，対照漁具に選択性がある場合の統計モデル，推定方法，そしてその効率等に注目し考察した．

2. 対照漁具に選択性を考慮した統計モデル

実験漁具と対照漁具の(遭遇)選択性曲線を $s_1(l), s_2(l)$ とし，選択性曲線にはロジスティック曲線 $s_i(l) = \exp(a_i + b_i l)/(1 + \exp(a_i + b_i l))$ $(i = 1, 2)$ を仮定する．ただし，$-\infty < a_i < 0, 0 < b_i < \infty$ $(i = 1, 2)$ とする．どの様な目合の小さな網を用いても，それが網である以上は本来少なからず選択性を有する．従って，厳密な意味での真のモデルは，対照漁具の選択性を考慮したモデルである．ここでは，対照漁具に選択性を考慮する場合をモデル1，考慮しない場合をモデル2とする．

<u>例1. 曳網のカバーネット試験</u>　実験漁具に進入した魚が実験漁具の目合へ遭遇する確率（遭遇確率とよぶ）を τ, 実験漁具の（遭遇）選択性曲線を $s_1(l)$ とおき，さらに対照漁具であるカバーネットの（遭遇）選択性曲線を $s_2(l)$ とおく（カバーネットの目合への遭遇確率は1と仮定する）．また，n_{l1}, n_{l2} をそれぞれ実験漁具，対照漁具での漁獲尾数とする．この時，$n_{l1}|n_{l+} \sim Bin(n_{l+}, \phi(l))$ を仮定する．ただし，

$$\phi(l) = \frac{1 - \tau + \tau\, s_1(l)}{1 - \tau(1 - s_1(l))(1 - s_2(l))} \ \text{（モデル1）} \qquad \phi(l) = 1 - \tau + \tau\, s_1(l) \ \text{（モデル2）}.$$

<u>例2. 曳網の比較操業試験</u>　並行して操業している漁具で漁獲対象となった魚のうち，実験漁具に進入する確率を分割率とよび p とおく．また実験漁具と対照漁具への（遭遇）選択性曲線を $s_1(l), s_2(l)$ とする（分割後の魚の各目合への遭遇確率は1と仮定する）．さらに，n_{l1}, n_{l2} をそれぞれ実験漁具，対照実験漁具での漁獲尾数とする．

$$\phi(l) = \frac{p\, s_1(l)}{p\, s_1(l) + (1-p)s_2(l)} \ \text{（モデル1）} \qquad \phi(l) = \frac{p\, s_1(l)}{p\, s_1(l) + (1-p)} \ \text{（モデル2）}.$$

ここでは，カバーネット試験において $\tau = 1$ を既知とする場合（ケース１），カバーネット試験において $\tau = 0.8$ でそれを未知とする場合（ケース２），また比較操業試験において $p = 0.6$ でそれを未知とする場合（ケース３）をそれぞれ真のモデルとした．本来興味の対象は実験漁具の選択性であるから，モデル１における (a_2, b_2) は局外パラメータとし，上記３つのケースについて $\theta = (a_1, b_1)$, $\theta = (a_1, b_1, \tau)$ および $\theta = (a_1, b_1, p)$ の推定に注目した．

3. 推定量の評価

対照漁具からの逃避がある場合に，誤ってモデル２を仮定すれば推定量にはバイアスをもたらす．逆に，逃避がほとんどない場合に，無理にモデル１を想定し実験漁具と対照漁具の選択性を同時に推定すると，実験漁具の推定に悪影響を及ぼす可能性もある．また，状況によっては両者の主観的な判断が不可能な場合も容易に想像される．この様な場合，逃避を考慮に入れるかどうかを，データに選択させる意味で，AIC などの情報量規準の利用が考えられる．この手続きを利用した推定は，いわば予備検定を用いた推定方法とみなす事ができる．すなわち，モデル１，２における最尤推定量を $\hat{\theta}_1$, $\hat{\theta}_2$ とし，それぞれのモデルに対する AIC の値を AIC_1, AIC_2 とする時，AIC 最小化推定量 (MAICE) は，

$$\hat{\theta}_3 = \begin{cases} \hat{\theta}_1 & \text{if} \quad AIC_1 \le AIC_2 \\ \hat{\theta}_2 & \text{if} \quad AIC_1 > AIC_2 \end{cases}$$

と書ける．この推定量は不連続であるが，$w_i = \exp(-AIC_i/2)/(\exp(-AIC_1/2)+\exp(-AIC_2/2))$ $(i = 1, 2)$ を重みとした推定量 $\hat{\theta}_4 = w_1\hat{\theta}_1 + w_2\hat{\theta}_2$ も同時に考える．$\hat{\theta}_4$ の導出については Burnham and Anderson (1998) を参照．

これらの推定量の性質を幾つかの設定の下でシミュレーションにより評価した．AIC を利用した２つの推定量 $\hat{\theta}_3, \hat{\theta}_4$ は，モデルを固定して得られる２つ推定量 $\hat{\theta}_1, \hat{\theta}_2$ を含めた４つの推定量の中でミニマックス的な性質を持つことがシミュレーションから示唆された．また，３つのケースにおいて，$\hat{\theta}_3, \hat{\theta}_4$ は，$\hat{\theta}_1$ を多くの場合，不偏性および MSE の意味で改善した．

上記のシミュレーションに加え，モデル２を誤って想定した場合の漸近バイアス等についても評価した．今後，ここで考察したモデル選択を考慮した推定量に対する標準誤差の評価方法に加えて，overdispersion のある観測値に対しても同種の議論が可能かどうか検討する余地があると考えられる．

参考文献

[1] Burnham, K.P. and Anderson, D.R. (1998). *Model Selection and Inference : A Practical Information-Theoretic Approach*. Springer-Verlag, New York.

[2] Cadigan, N.G. and Millar, R.B. (1992). *Can. J. Fish. Aquat. Sci.*, **49**, 1624-1632.

[3] Millar, R.B. (1992). *J. Amer. Statist. Soc.*, **87**, 962-968.

[4] Mituhasi, T., Tokai, T., Ercoli, R., García, J.C., Salvini, L., Bartozzetti, J. and Roth, R. (2000). *Fish. Sci.*, **66**, 327-333.

Efron の反例について

大阪大学基礎工 熊谷悦生

Efron(1975) は, 最尤推定量の漸近的情報量損失が, 曲指数型分布族において統計的曲率による表現を持つことを示したが, 一方, 曲指数型分布族に属する 3 項分布で irregular な場合がある例も示した. その反例に関しては, Kumagai and Inagaki(1996), Kass and Vos(1997) を参照するとして, ここでは 2 次元指数型分布族の期待値がある微分方程式によって定義された曲指数型分布族を考察する. この微分方程式は, 一般の指数型分布族に対してではないが 3 項分布において Efron が等式 (9.9) で示したものと同じものである.

$$
\text{''}\quad (9.9) \qquad \boldsymbol{\lambda}(\theta_0) \;\equiv\; \boldsymbol{\lambda}(0) + \int_0^{\theta_0} \mu(\theta)\;\{\boldsymbol{\Sigma}(\theta)\,\boldsymbol{\phi}(\theta)\}\;d\theta \quad\text{''}
$$

但し,

$$
\mu(\theta) \equiv \frac{\|\boldsymbol{\lambda}(\theta) - \mathbf{c}\|}{\|\boldsymbol{\Sigma}(\theta)\boldsymbol{\phi}(\theta)\| \sin B(\theta)}
$$

で, 記号 $\|\cdot\|$ はベクトルのノルムを表している. これを 2 次元指数型分布族における Efron のパラメトリゼイションと呼ぶことにする. 注意すべき点は, Efron のパラメトリゼイションを 3 項分布でのパラメトリゼイションとみなさないで, 一般の 2 次元曲指数型分布族でのものとみなす点にある.

まず, 2 次元曲指数型分布族における平均パラメータ $\boldsymbol{\beta}(\theta)$ を使って, 先の (9.9) 式を書き直すと

$$
\boldsymbol{\beta}(\theta) \equiv \boldsymbol{\beta}(0) + \int_0^{\theta} \|\boldsymbol{\beta}(\tau) - \mathbf{c}\| \, \frac{\dot{\boldsymbol{\beta}}(\tau)}{\left\|\dot{\boldsymbol{\beta}}(\tau)\right\| \, \sin B(\tau)} \, d\tau
$$

となり, 両辺を θ で微分して

$$
\dot{\boldsymbol{\beta}}(\theta) \;=\; \frac{\|\boldsymbol{\beta}(\theta) - \mathbf{c}\|}{\left\|\dot{\boldsymbol{\beta}}(\theta)\right\| \, \sin B(\theta)} \, \dot{\boldsymbol{\beta}}(\theta)
$$

を得る. これが任意の θ に対して成り立つことから,

$$
\|\boldsymbol{\beta}(\theta) - \mathbf{c}\| = \left\|\dot{\boldsymbol{\beta}}(\theta)\right\| \, \sin B(\theta)
$$

となるので, その結果

$$
\|\boldsymbol{\beta}(\theta) - \mathbf{c}\| \;=\; \left\|\dot{\boldsymbol{\beta}}(\theta)\right\| \sqrt{1 - \left(\frac{\langle \dot{\boldsymbol{\beta}}(\theta), \boldsymbol{\beta}(\theta) - \mathbf{c} \rangle}{\left\|\dot{\boldsymbol{\beta}}(\theta)\right\| \, \|\boldsymbol{\beta}(\theta) - \mathbf{c}\|} \right)^2}
$$

が考えるべき微分方程式となる.

この微分方程式を解くことによって, ある C_1 級の正函数 $r(\theta)$ を使い以下のように明示的な解が得られた:

$$\beta(\theta) - \boldsymbol{c} \ = \ r(\theta)\, e(\theta + A_0).$$

但し, A_0 は積分定数で $\boldsymbol{e}(\theta + A_0)$ は単位ベクトル $\boldsymbol{e}(\theta + A_0) = (\cos(\theta + A_0), \sin(\theta + A_0))'$ である. この結果から, Efron パラメトリゼイションは, らせん曲線の回転部分だけを規定しているが, 長さ部分は $r(\theta)$ としてしか規定していないことが判明する.

次に Efron パラメトリゼイションの特徴を考察する. ある定ベクトル $\boldsymbol{c}$ が尤度方程式による部分空間 $\mathcal{L}(\theta)$ に属していると仮定する:

$$c \in \ \mathcal{L}(\theta) = \{x \, : \, \langle \dot{\alpha}(\theta), x - \beta(\theta) \rangle = 0\}.$$

この仮定の下では,

$$c \ \in \ \mathcal{L}(\theta), \qquad \dot{\beta}(\theta) \ = \ \boldsymbol{\Sigma}(\theta)\, \dot{\alpha}(\theta)$$

を満たす Efron の微分方程式の解として Efron パラメトリゼイションを把握することが出来, その特徴を以下に示す.

(1) らせん曲線の長さ函数が定数 $r(\theta) = r_0$, 即ち, 期待値パラメータは中心が $\boldsymbol{c}$, 半径が r_0 の円周上に存在する:

$$\beta(\theta) \ = \ \boldsymbol{c} + r_0\, \boldsymbol{e}(\theta + A_0).$$

(2) 曲指数型分布族の共分散行列は単位行列に比例する:

$$\boldsymbol{\Sigma}(\theta) \ = \ \sigma^2(\theta)\, \mathbf{I}_2 \ = \ \begin{pmatrix} \sigma^2(\theta) & 0 \\ 0 & \sigma^2(\theta) \end{pmatrix}.$$

参考文献

B.Efron(1975), Defining the Curvature of a Statistical Problem (with Applications to Second Order Efficiency), *Ann. Statist.*, **3**, 1189–1242.

R.E.Kass and P.W.Vos(1997), *Geometrical Foundations of Asymptotic Inference*, Wiley.

E.Kumagai and N.Inagaki(1996), Comment on Efron's Counterexample, *Mathematica Japonica*, **44**, 449–454.

E.Kumagai, N.Inagaki and K.Inoue(1999), On Efron's Parameterization, *Research Reports on Statistics No.42*, Osaka University.

N.Inagaki and E.Kumagai(1996), Exact Information Loss in Fisher's Circle Model, *Mathematica Japonica*, **44**, 455–467.

Pitman 確率分割における推測問題
—非正則の一例—

渋谷政昭（高千穂商科大学）、大和元（鹿児島大学）

有限集合および数の確率的分割は、応用確率論・統計学の基本的概念である。数の確率的分割のなかで基本的なのが Ewens 確率分割系列族である Johnson et al. (1997) 。Jim Pitman は 1995 年に始まる一連の論文で、1 パラメータの Ewens 確率分割を 2 パラメータに拡張し、本質的な諸概念を明らかにした。本稿では Pitman のモデルの推測問題を議論する。非正則のため良い推定量は得られない。

1 自然数の順序のない分割

自然数 n の、順序のない、自然数の和への分割の全体を $\mathcal{C}_n^u$ とする。$\mathcal{C}_n^u$ 上の確率分布が確率分割である。$\{c_1,\ldots,c_k\}\in\mathcal{C}_n^u$の寸法指標を $s=(s_1,\ldots,s_n)s_j=\#\{i;c_i=j\}$ とする。$\sum_{j=1}^n s_j=k;\sum_{j=1}^n js_j=n.$

$$p(s;\theta,\alpha)=\frac{n!\theta(\theta+\alpha)\cdots(\theta+(k-1)\alpha)}{\theta^{[n]}}\prod_{j=1}^{n}\left(\frac{(1-\alpha)^{[j-1]}}{j!}\right)^{s_j}\frac{1}{s_j!}, \tag{1}$$

$$s=(s_1,\ldots,s_n)\in\mathcal{C}_n^u,\quad k=\sum_{j=1}^{n}s_j.$$

のとき、この確率分割を $\mathcal{P}(\theta,\alpha;\mathcal{C}_n^u)$ で表す。$\alpha=0$ のときは Ewens sampling formula である。

パラメータ (θ,α) の範囲は次のように制限される。

$$\alpha=1 \text{ だと確率 } 1 \text{ で } s_1=n,\quad \theta=-\alpha \text{ だと確率 } 1 \text{ で } s_n=1.$$

これらが退化の条件である。そうでないとき $0\le\alpha<1$ ならば $-\alpha<\theta$ に限られる。

$\alpha<0$ ならば $\theta=-M\alpha$, $M=1,2,\ldots,$ のとき $1\le k\le M<\infty$ に限られ、$s_{M+1}=\cdots=s_n=0$.

2 パラメータ推定

モーメント法、最小距離法

諸標本モーメントが、その期待値の一致推定量ではないため、モーメント推定量も一致推定量ではない。

$$R_0:=\sum_{j=1}^{n}S_j=K_n,\quad R_2:=\sum_{j=1}^{n}\frac{j^{(2)}S_j}{n^{(2)}},\quad R_3:=\sum_{j=1}^{n}\frac{j^{(3)}S_j}{n^{(3)}},$$

とすると（$R_1:=\sum_{j=1}^n jS_j/n\equiv 1$ は利用できない）、

$$E(R_0)=\frac{\theta}{\alpha}\left(\frac{(\theta+\alpha)^{[n]}}{\theta^{[n]}}-1\right),\quad E(R_2)=\frac{1-\alpha}{\theta+1},\quad E(R_3)=\frac{(1-\alpha)^{[2]}}{(\theta+1)^{[2]}},$$

であるから、R_2,R_3 を用いると、

$$\hat{\alpha}=\frac{R_3/R_2-2R_2+R_3}{R_3/R_2-R_2},\quad \hat{\theta}=\frac{1+R_2-2R_3/R_2}{R_3/R_2-R_2}. \tag{2}$$

あるいは R_2 によりパラメータを1個とし、R_0 または S_1 の期待値の式に代入して非線形方程式を解く。

$E(jS_j), j=1,\ldots,n,$ とその期待値との距離を最小にする。2乗距離、Hellinger 距離が考えられる。たとえば2乗距離を最小にするならば、

$$(\hat{\theta},\hat{\alpha}) = \arg\min_{(\theta,\alpha)} \sum_{j=1}^{n}(jS_j - e_j)^2, \quad e_j = E(jS_j) = n\binom{n-1}{j-1}\frac{(1-\alpha)^{[j-1]}(\theta+\alpha)^{[n-j]}}{(\theta+1)^{[n-1]}} \tag{3}$$

最尤法

最尤方程式は

$$\begin{aligned}
\frac{\partial L}{\partial \theta} &= \frac{1}{\theta} + \cdots + \frac{1}{\theta+(k-1)\alpha} - \left(\frac{1}{\theta} + \cdots + \frac{1}{\theta+n-1}\right) = 0, \\
\frac{\partial L}{\partial \alpha} &= \frac{1}{\theta+\alpha} + \cdots \frac{k-1}{\theta+(k-1)\alpha} - \sum_{j=2}^{n} s_j\left(\frac{1}{1-\alpha} + \cdots + \frac{1}{j-1-\alpha}\right) = 0.
\end{aligned} \tag{4}$$

$$\begin{aligned}
\frac{\partial^2 L}{\partial \theta^2} &= -\left(\frac{1}{(\theta+\alpha)^2} + \cdots + \frac{1}{(\theta+(k-1)\alpha)^2}\right) + \left(\frac{1}{(\theta+1)^2} + \cdots + \frac{1}{(\theta+n-1)^2}\right), \\
\frac{\partial^2 L}{\partial \theta \partial \alpha} &= -\left(\frac{1}{(\theta+\alpha)^2} + \cdots + \frac{k-1}{(\theta+(k-1)\alpha)^2}\right), \\
\frac{\partial^2 L}{\partial \alpha^2} &= -\left(\frac{1}{(\theta+\alpha)^2} + \cdots + \frac{(k-1)^2}{(\theta+(k-1)\alpha)^2}\right) - \sum_{j=2}^{n} s_j\left(\frac{1}{(1-\alpha)^2} + \cdots + \frac{1}{(j-1-\alpha)^2}\right).
\end{aligned}$$

Fisher 情報量の計算では $(s_1,\ldots,s_n), \sum_{j=1}^{n} s_j = k,$ を確率変数として期待値をとることになる。

Proposition 1 $\alpha > 0$ のとき、情報行列 $I = I(\theta,\alpha)$ の各要素について

$$I_{\theta\theta} = O(1), \quad I_{\theta\alpha} = O(\log n), \quad I_{\alpha\alpha} = O(n^{\alpha}).$$

α 既知の場合

$\alpha \neq 0$ が既知で θ だけを推定する場合、推定方程式、情報量は上式の一部であるから、$I_{\theta\theta} = O(1)$ に変わりはない。$\alpha = 0$ の特別の場合は、Ewens sampling formula と呼ばれ多くのことが知られている。分割数 $k = \sum_{j=1}^{n} s_j$ が十分統計量である。その確率分布は離散指数型（べき級数型）である。標準理論がそのまま成り立ち、Fisher 情報量は $I(\theta) = O(\log n)$ である。

参考文献

[1] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997) *Discrete Multivariate Distributions*, Wiley

[2] Pitman, J. (1995) Exchangeable and partially exchangeable random partitions, *Probability Theory and Relatd Fields*,**12**, 145–158.

[3] Pitman, J. (1996) Random discrete distributions invariant under size-biased permutation, *Journal of Applied Probability*,**28**, 525–539.

[4] 渋谷政昭、大和元 (2000) Jim Pitman's 2-parameter random partition,（覚書）

[5] 星野伸明 (2000) Pitman Sampling Formula の個票開示リスク評価への応用、第68回日本統計学会講演報告集、札幌（北海道大学）、365-366

制約付き最尤推定量の集中確率による優位性について

神戸商大・商経　森谷 義哉
熊本大・工　岩佐 学

1　はじめに

n 次元正規確率変数 $X \sim N(\theta, I_n)$ を考える. 平均 θ に関する制約条件の下で θ を推定するとき, 制約の下での最尤推定量 (RMLE) と制約を考慮しない通常の最尤推定量 (MLE) の優劣を決定をする. 平均 2 乗誤差による優劣について積極的に研究されてきたが, Rueda, Salvador and Fernández (1997) は集中確率による優劣を考えた. 彼らは θ に関する制約が 1 つの線形不等式で与えられる場合に θ の RMLE が MLE より優れていることを示した.

そこで, 本報告では θ に関する制約が 2 つの線形不等式で与えられる場合に集中確率の意味で θ の RMLE が MLE より優れていることを示す. そして, この結果の多次元 (3 次元以上) へのいくつかの拡張を行う.

以後の説明のために記号を準備する. $|x| = \sqrt{x'x}$ は $\mathbf{R}^n$ のユークリッドノルムを表す. $\mathcal{A}, \mathcal{C}$ をそれぞれ $\mathbf{R}^n$ の原点対称 $(A = -A)$ な集合, 閉凸集合の全体とし, 制約条件は $\theta \in C$ の形で与える. さらに, MLE X の C への射影を $\pi(X|C)$ で表すと, $\pi(X|C)$ は制約条件 $\theta \in C$ の下での RMLE となる.

2　2 次元での優位性

定義 1 任意の $A \in \mathcal{A} \cap \mathcal{C}$ に対して, 集中確率に関する不等式

$$P_\theta[X \in A + \theta] \leq P_\theta[\pi(X|C) \in A + \theta] \quad \forall \theta \in C \tag{1}$$

が成り立つとき, RMLE $\pi(X|C)$ は MLE X より集中確率の意味で優れているという.

初等的な議論により次の定理が証明される. この定理は, 多次元への拡張を考えるとき基本となる結果である.

定理 1 $n = 2$ のとき, 任意の $C \in \mathcal{C}$ に対して, RMLE $\pi(X|C)$ は MLE X より集中確率の意味で優れている.

3　多次元への拡張

定理 1 を 3 次元以上へ拡張するために, 線形空間 L によって定まる次のような対称性と凸性を考える.

定義 2 A が L-凸であるとは, すべての $x \in L^\perp$ に対して $A \cap (x + L)$ が凸であることをいい, L-対称であるとは, すべての $x \in L^\perp$ に対して $A \cap (x + L)$ が x に関して対称であることをいう.

$n-2$ 次元線形空間を含む凸集合 C を考えると, 次の系は定理 1 より簡単に導かれる.

系 1 C が $n-2$ 次元線形空間 L を含むとき, 可測集合 A が $L^{\perp}$-対称かつ $L^{\perp}$-凸であれば, 不等式 (1) がすべての $\theta \in C$ に対して成り立つ.

定義 3 $\mathbf{R}^n$ の集合 A と線形空間 L に対して,

$$A^L = \bigcup_{x_{L^\perp} \in L^\perp} \{x_L + x_{L^\perp} \mid x_L \in L, |x_L| \leq \gamma(x_{L^\perp})\}$$

と定義し, A の L-対称化と呼ぶ. ただし, $\gamma(x_{L^\perp}) \geq 0$ は次の等式を満たすように定める.

$$P_0[X_L \in (A^L - x_{L^\perp}) \cap L] = P_0[X_L \in (A - x_{L^\perp}) \cap L].$$

明らかに A^L は L-凸 かつ L-対称であり, さらに A^L について次の 2 つの補題が成り立つ.

補題 1 $A \in \mathcal{A} \cap \mathcal{C}$ ならば, A^L は $L^{\perp}$-凸かつ $L^{\perp}$-対称である.

補題 2 C が線形空間 L を含むならば,

$$P_\theta[\pi(X|C) \in A + \theta] = P_\theta[\pi(X|C) \in A^L + \theta] \quad \forall\theta.$$

C が線形空間 L を含む凸集合であれば, A の L-対称化 A^L を考えることにより, 系 1 と補題 1, 2 を経て次の定理を得る.

定理 2 C が $n-2$ 次元線形空間を含むとき, RMLE $\pi(X|C)$ は MLE X より集中確率の意味で優れている.

定理 2 の結果は, 特別な場合として C が 2 つの線形不等式によって定まる閉凸錐

$$C = \{x \in R^n \mid a'x \geq 0, b'x \geq 0\} \quad (a, b \in \mathbf{R}^n)$$

の場合を含んでおり, 結局 2 個の線形不等式制約までは集中確率の意味で RMLE が MLE を改良していることがわかる. これは, Rueda *et al.* (1997) の拡張になっている.

また, 定理 1 の別な形の拡張として次の定理を得る.

定理 3 C が 2 次元線形空間に含まれるとき, RMLE $\pi(X|C)$ は MLE X より集中確率の意味で優れている.

$n = 3$ の場合は肯定的な解決を期待しているが解決には至っていない. $n \geq 4$ であれば, 集中確率の意味で $\pi(X|C)$ が X を改良しないような C を具体的に構成できる.

参考文献

[1] Iwasa, M. and Moritani, Y. Concentration probabilities for restricted and unrestricted MLEs, to appear in *J. Multivariate Anal.*

[2] Rueda, C., Salvador, B. and Fernández, M. A. (1997). Simultaneous estimation in a restricted linear model. *J. Multivariate Anal.*, **61**, 61-66.

ある制約を設けた Gaussian 素子の学習誤差と汎化誤差

三重大学 工学部 物理工学科 萩原克幸

1 はじめに

階層型ニューラルネットや Radial Basis Function(RBF) を用いた非線形回帰モデルでは, モデル選択規準 AIC(Akaike Information Criterion)[1] で仮定されているような, モデルが真の分布を含む状況において, 真の結合重み (ネットワーク・パラメータ) が識別不能となり, モデル選択規準を導く際, 通常の漸近展開が適用できないという問題が生じる. モデル選択規準は, 学習誤差に基づく汎化誤差の不偏推定量であるため, 階層型ニューラルネットのモデル選択の問題を解決するためには, 真の結合重みが識別可能でない場合について, 汎化誤差および学習誤差のデータの分布に関する期待値を知る必要がある. 本研究では, ある制約を設けた Gaussian 素子について, 学習誤差と汎化誤差の期待値のオーダーを導く.

2 問題設定

$Q(x,y) = Q(y|x)Q(x)$, $x \in \mathbf{R}^D$, $y \in \mathbf{R}$ から独立にサンプルされた L 個のデータの組みを $D = \{(x_l, y_l), l = 1, \ldots, L\}$ とする. y_l は x_l に対し, $y_l = h(x_l) + \xi_l$ という規則により生成されるものとする. ただし, ξ_l は適当な分布からの独立なサンプルである. 以下では, ネットワーク・パラメータの識別可能性が失われる例として,

仮定 1 $y_l = \xi_l \sim N(0, \sigma^2) = Q(y)$, $l = 1, \ldots, L$

を考え, このデータを, Gaussian 素子

$$f(x) := \alpha g(x),\ g \in \mathcal{G},\ \mathcal{G} := \left\{ g : g(x) = e^{\left(-\frac{1}{\tau}(x-\mu)'(x-\mu)\right)}, \tau \in \mathbf{R}^+, \mu \in \{x_1, \ldots, x_L\} \right\}. \tag{1}$$

によりフィッティングする状況を考える. ただし, $\mathbf{R}^+ = \{x : 0 < x < \infty\}$ であり, $'$ は行列の転置を表す.

いま, $r_{\mathrm{emp}}(\alpha, g) = \frac{1}{L}\sum_l (y_l - f(x_l))^2$ として, 固定された g に対し, $r_{\mathrm{emp}}(\widehat{\alpha}, g) = \min_\alpha r_{\mathrm{emp}}(\alpha, g)$ とする. 学習誤差を, $\widehat{r}_{\mathrm{emp}} = \inf_{g \in \mathcal{G}} r_{\mathrm{emp}}(\widehat{\alpha}, g)$ と定義する. このとき, 簡単な計算により,

$$\widehat{r}_{\mathrm{emp}} = \frac{1}{L}\sum_l y_l^2 + \frac{1}{L}\sup_{g \in \mathcal{G}} K(g),\ K(g) = \left(\sum_l g(x_l) y_l\right)^2 / \sum_l g^2(x_l) \tag{2}$$

となる.

3 学習誤差および汎化誤差の期待値のオーダー

3.1 学習誤差の期待値のオーダー

学習誤差の期待値を $R_{\mathrm{emp}}(L) = \mathbf{E}_D\{\widehat{r}_{\mathrm{emp}}\}$ で定義する. ただし, $\mathbf{E}_D$ はデータ D の分布に関する期待値である. これを計算するためには, 式 (2) より,

$$\mathbf{E}_D\left\{\sup_{g \in \mathcal{G}} K(g)\right\} \tag{3}$$

を計算すれば良い. これは, g を固定すると $K(g)$ が自由度 1 の χ^2 分布に従う確率変数であることから, χ^2-process の最大値の期待値を計算する問題に帰着される. 式 (3) の下界は, 指標関数によるフィッティングを経由する際に極値理論を適用する形で計算されているので [2], ここでは, 上界を求める. まず, ϵ–covering の概念を導入して, χ^2-process の有限な部分列として, χ^2 分布に従う確率変数列を作り, その最大値 $K(g_*)$ を用いて, 確率 1 で $K(g) \le K(g_*) + C(L)$ が成り立つことを示す. ここで, $C(L)$ は期待値の計算できる適当な確率変数である. ここで, $\mathbf{E}_D\{K(g_*)\}$ および $\mathbf{E}_D\{C(L)\}$ が $O(\log L)$ で上から抑えられることが示せるので, 結局, 式 (3) は $O(\log L)$ で上から抑えられる. したがって, [2] の結果と合わせて次の定理を得る.

定理 1 *式 (1) の素子において, 仮定 1 の下で, 十分大きな L に対し, $\sigma^2 - R_{\mathrm{emp}}(L) = O\left(\frac{\log L}{L}\right)$ である.*

3.2 汎化誤差の期待値のオーダー

未学習データ $z_1,\ldots,z_L$ を $Q(y)$ からの独立なサンプルとし，それらは $y_1,\ldots,y_L$ のいずれとも独立であるものとする．汎化誤差を条件付き期待値 $R(f,L,\vec{x}) := \mathbf{E}_{z_1,\ldots,z_L}\left\{\frac{1}{L}\sum_l (z_l - f(x_l))^2 \middle| \vec{x}\right\}$ により定義する．ただし，$\vec{x} = (x_1,\ldots,x_L)$ である．$\widehat{r}_{\mathrm{emp}}$ を与える関数 f を $\widehat{f}$ として，汎化誤差の期待値を $R(L) := \mathbf{E}_D\left\{R(\widehat{f},L,\vec{x})\right\}$ で定義する．汎化誤差は，簡単な計算により $R(\widehat{f},L,\vec{x}) = \sigma^2 + \frac{1}{L}\sup_{g\in\mathcal{G}} K(g)$ とできることが知れるので，これと定理 1 から，次の定理が導かれる．

定理 2 *式 (1) の素子において，仮定 1 の下で，十分大きな L に対し，$R(L) - \sigma^2 = O\left(\frac{\log L}{L}\right)$ である．*

いま，漸近展開が適用できる状況では，仮定 1 の下で，学習誤差の期待値および汎化誤差の期待値と雑音分散との差は $O(1/L)$ である [3]．一方，定理 1 および定理 2 の結果は，その差が $O(\log L/L)$ であることを述べており，本稿のモデルに対しては，漸近展開による評価より，学習誤差を小さく，汎化誤差を大きく見積もる必要があることを示している．また，この差はモデル選択規準のペナルティ項を直接決める量であり，したがって，汎化を測る入力点と学習の際の入力点が同じ場合，本稿のモデルに対するモデル選択規準のペナルティ項は $O(\log L/L)$ となることを意味している．

4 学習誤差と汎化誤差の確率的オーダー

確率 1 で成り立つ不等式 $K(g) \le K(g_*) + C(L)$ と正規分布に従う確率変数の絶対値の列の最大値に対する概収束の結果 [4] を用いて，$\sup_{g\in\mathcal{G}} K(g)$ を評価できて，これを用いると，学習誤差について，次の定理が得られる．

定理 3 *式 (1) の素子において，仮定 1 の下で，$\sigma^2 - \widehat{r}_{\mathrm{emp}} = O_P\left(\frac{\log L}{L}\right)$ が成り立つ．ここで，O_P は確率的なオーダーを意味する．*

一方，これと同じ評価を用いて，汎化誤差については，前節の議論より，次のことが言える．

定理 4 *式 (1) の素子において，仮定 1 の下で，$L \to \infty$ のとき，$R(\widehat{f},L,\vec{x}) - \sigma^2 = O_P\left(\frac{\log L}{L}\right)$ が成り立つ．*

5 おわりに

これまでに，階層型ニューラルネットや RBF が識別可能性を失う場合について，学習誤差の期待値の上界および汎化を測る入力点が学習の際の入力点と同じ場合の汎化誤差の期待値の下界が得られていた [2]．本稿では，ϵ–covering の概念と極値理論を組み合わせる形で，平均値パラメータを離散化した Gaussian 素子の学習誤差と汎化誤差の期待値のオーダーを解析し，それらと雑音分散の差が $\log L/L$ のオーダーであることを示した．さらに，学習誤差と汎化誤差の雑音分散からの隔たりの確率的なオーダーも $\log L/L$ となることを示した．今後の課題は，まず，本稿での結果を素子数に関して一般化する必要がある．また，ここでの汎化誤差の解析では，汎化を測る入力点が学習の際の入力点と同じであることが本質的であったが，これを仮定しない場合の解析も残されている．

参考文献

[1] Akaike H. : In *2nd International Symposium on Information Theory*, B.N.Petrov and F.Csáki eds., Akadémia Kiado, Budapest, 267–281, (1973).

[2] 萩原克幸, 久野和宏, 臼井支朗 : 「統計的推測理論とその情報論的側面」予稿集, 95–102, (1998).

[3] Murata N., Yoshizawa S., Amari S. : *IEEE Trans. on Neural Networks*, **5**, 6, 865–872 (1994).

[4] Deo C. M. : *Sankhyā Series A*, **34**, pp.289–292, (1972).

THE BIOEQUIVALENCE PROBLEMS AND TESTING HYPOTHESES PROBLEMS FOR MULTINOMIAL DISTRIBUTIONS

筑波大・数学　津田 美幸

多変量離散指数型分布族に対する検定

自然数 p, q $(p \geq q \geq 2)$ を固定する. 集合 Θ は p 次元 Euclid 空間の開集合で, 原点を含むとする. p 次元整数値確率ベクトル $\boldsymbol{X} := {}^t(X_1, \ldots, X_p)$ は $\boldsymbol{\theta} := {}^t(\theta_1, \ldots, \theta_p)$ $(\boldsymbol{\theta} \in \Theta)$ を自然母数とする指数型分布 $P_{\boldsymbol{\theta}}$ に従うとする. p 次元整数値ベクトル $\boldsymbol{x} := {}^t(x_1, \ldots, x_p)$ に対し, $\boldsymbol{X}$ の確率関数を

$$P_{\boldsymbol{\theta}}\{\boldsymbol{X} = \boldsymbol{x}\} = a(\boldsymbol{\theta}) b(\boldsymbol{x}) \exp\left(\sum_{i=1}^{p} \theta_i x_i\right)$$

で定義する. ただし, $b(\mathbf{0}) > 0$ とし, $x_i < 0$ となる i $(1 \leq i \leq p)$ が存在するならば $b(\boldsymbol{x}) = 0$ とする. Θ の部分集合 Θ_0, Θ_1 を

$$\Theta_0 := \{\boldsymbol{\theta} \in \Theta \mid \theta_i \geq 0 \text{ となる } i \ (1 \leq i \leq q) \text{ が存在する.}\},$$

$\Theta_1 := \Theta \backslash \Theta_0$ で定義し, 未知の $\boldsymbol{\theta}$ に対する仮説

$$H_0 : \boldsymbol{\theta} \in \Theta_0 \text{ vs. } H_1 : \boldsymbol{\theta} \in \Theta_1 \tag{0.1}$$

を有意水準 α $(0 < \alpha < 1/2)$ で検定する問題を考える.

ある i $(1 \leq i \leq q)$ に対して, $\theta_i = 0$ とする. $\boldsymbol{\eta}_i := {}^t(\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_p)$ は局外母数であり, $\boldsymbol{Y}_i := {}^t(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_p)$ は $\boldsymbol{\eta}_i$ に対する完備十分統計量である. 仮に, $\phi_1(\boldsymbol{x})$ が不偏検定ならば, $\phi_1(\boldsymbol{x})$ は相似検定であり, Neyman 構造を持つ. すなわち, $P_{\{\theta_i = 0\}}\{\boldsymbol{Y}_i = \boldsymbol{y}_i\} > 0$ となる全ての $\boldsymbol{Y}_i$ の実現値 $\boldsymbol{y}_i$ に対して,

$$E_{\{\theta_i = 0\}}\{\phi_1(\boldsymbol{X}) \mid \boldsymbol{Y}_i = \boldsymbol{y}_i\} = \alpha \tag{0.2}$$

が成り立つ. しかし, 自明な検定 $\phi_1(\boldsymbol{x}) :\equiv \alpha$ 以外には (0.2) を満たす検定は存在しない (Lehmann (1952)). そこで, bioequivalence 問題に対する Tsuda (2000) の手法を拡張して, 次の性質を満たす検定 $\phi_1(\boldsymbol{x})$ を考える.

性質 0.1. 自然数 i $(1 \leq i \leq q)$ の関数 $s_*(i)$ が存在して, $\sum_{j \neq i} x_j \leq s_*(i)$ かつ $P_{\boldsymbol{\theta}}\{\boldsymbol{Y}_i = \boldsymbol{y}_i\} > 0$ ならば (0.2) が成り立つ.

本研究では, 再帰的な手順によって ϕ_1 構成し, ある検定の族 $\mathcal{C}$ の中で許容的であることを示した. また, 応用例として, 多項分布の局所最頻値に関する仮説検定を挙げ, 具体的に検定を構成した.

各成分が独立な多変量分布の検定

各 $i = 1,\dots,p$ に対して，$\Lambda^{(i)}$ を母数空間ととし，母数 $\lambda^{(i)} \in \Lambda^{(i)}$ に対して，$(\Omega^{(i)}, \mathcal{F}^{(i)}, P^{(i)}_{\lambda^{(i)}})$ を確率空間とする．また, $\Lambda^{(i)}$ の空でない真の部分集合 $\Lambda^{(i)}_*$ と，統計量

$$\varphi^{(i)}_{j^{(i)}} : \Omega^{(i)} \longrightarrow \Re \qquad (j^{(i)} = 0,\dots,m-1)$$

が存在して，任意の $\lambda^{(i)} \in \Lambda^{(i)}_*$ に対して，

$$E_{\lambda^{(i)}}[\varphi^{(i)}_{j^{(i)}}] \;=\; 1/m$$

とする．ただし, m は 2 以上の自然数とする．また，任意の $\lambda^{(i)} \in \Lambda^{(i)}$ に対して，

$$\sum_{j^{(i)}=0}^{m-1} E_{\lambda^{(i)}}[\varphi^{(i)}_{j^{(i)}}] \;=\; 1$$

であるとする．さらに, $\Omega := \Omega^{(1)} \times \cdots \times \Omega^{(p)}$ とし，$\{A^{(1)} \times \cdots \times A^{(p)} \mid A^{(1)} \in \mathcal{F}^{(1)},\dots, A^{(p)} \in \mathcal{F}^{(p)}\}$ を含む最小の可算加法族を $\mathcal{F}$ とし, $\Lambda := \Lambda^{(1)} \times \cdots \times \Lambda^{(p)}$, $\lambda := (\lambda^{(1)},\dots,\lambda^{(p)}) \in \Lambda$, $A^{(1)} \in \mathcal{F}^{(1)},\dots,A^{(p)} \in \mathcal{F}^{(p)}$ に対して確率測度 P_λ を，

$$P_\lambda(A^{(1)} \times \cdots \times A^{(p)}) := \prod_{i=1}^{p} P^{(i)}_{\lambda^{(i)}}(A^{(i)})$$

で定め，確率空間 $(\Omega, \mathcal{F}, P_\lambda)$ を構成する．また，

$$\Lambda_* := \left\{(\lambda^{(1)},\dots,\lambda^{(p)}) \in \Lambda \mid \text{ある } i \text{ が存在して}, \lambda^{(i)} \in \Lambda^{(i)}_*\right\}$$

とする．L を order m の $p-1$ 次元 latin hypercube とする．$\boldsymbol{j} := {}^t(j^{(1)},\dots,j^{(p)})$ の集合 J を

$$J := \{\boldsymbol{j} \mid L \text{ の } {}^t(j^{(1)}+1,\dots,j^{(p-1)}+1) \text{ 成分の値は } j^{(p)}\}$$

で定義する．

$$\phi_2 := \sum_J \prod_{i=1}^{p} \varphi^{(i)}_{j^{(i)}}$$

とすると, 任意の $\lambda \in \Lambda_*$ に対して，

$$E_\lambda[\phi_2] \;=\; 1/m$$

である．

この手法は, Iwasa (1991,1994) でも論じられている．本研究では, 多変量 bioequivalence 問題に応用し, Wang, Hwang and Dasgupta (1999) の検定を一様に改良した．

ミキシング過程に対する検定統計量の漸近展開

名大・工　阪本雄二

時間間隔 $[0,T]$ 上の観測値が，ある確率空間上で定義された幾何的混合性を持つマルコフ過程 $Y^T = (Y_t)_{t\in[0,T] or [0,1,\dots,T]}$ の実現値であるとし，その確率分布が適当な測度に関して正の密度 $p_{T,\theta}$ を持つものとする．ここで，θ は p-次元ユークリッド空間の凸で有界な開集合 Θ に値をとる未知母数であるとする．たとえば，確率差分方程式

$$Y_t = V_0(Y_{t-1},\dots,Y_{t-m},\theta) + V(Y_{t-1},\dots,Y_{t-m})\varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}, \ \ Y_0 \sim \nu_\theta$$

を満たす (非線形) 自己回帰時系列 $(Y_t)_{t\in[0,1,\dots,T]}$ であるとか，確率微分方程式

$$dY_t = V_0(Y_t,\theta)dt + V(Y_t)dw_t, \quad w_t \ : \text{ the standard Wiener}, \ \ Y_0 \sim \nu_\theta$$

を満たす拡散過程などが，例として上げられる (これらの例のマルコフ性は明らかであるが，幾何的混合性に関しては適当な条件を必要とする)．本報告では，これらの統計モデルに対する検定統計量の漸近展開 (T が大きいときの) を求めることにする．

対立仮説 $\theta \neq \theta_0$ に対して帰無仮説 $\theta = \theta_0$ を検定するときの検定統計量について考えよう．まず，記号を用意する．対数尤度 $\ell(\theta) = \log p_{T,\theta}$ の母数による微分を添え字を用いて表すことにし:$\ell_A(\theta) = \delta_{a_1}\cdots\delta_{a_k}\ell(\theta)$, $\delta_a = \partial/\partial\theta^a$, $A = a_1\cdots a_k$, それらを正規化して，$Z_A = (\ell_A(\theta_0) - E[\ell_A(\theta_0)])/\sqrt{T}$, さらに，直交化して，$\tilde{Z}_A = Z_A - E[Z_A Z_{a'}]g^{a'a}Z_a$, $\tilde{Z}^{ab} = g^{aa'}g^{bb'}\tilde{Z}_{a'b'}$, $\tilde{Z}^{abc}\ g^{aa'}g^{bb'}g^{cc'}\tilde{Z}_{a'b'c'}$ とおく．ただし，$g_{ab} = E[Z_a Z_b]$, $(g^{ab}) = (g_{ab})^{-1}$. これらの記号を用いると，対数尤度 $\ell(\theta)$ に基づく多くの検定統計量 V_T (尤度比，Wald, Rao) は，帰無仮説 $\theta = \theta_0$ のもとで，次の様な確率展開をもつことが知られている:

$$\begin{aligned} V_T =& g_{ab}\tilde{Z}^a\tilde{Z}^b + \frac{1}{\sqrt{T}}\left(a^{ab}_{1;cd}\tilde{Z}_a\tilde{Z}_b\tilde{Z}^{cd} + a^{abc}_2\tilde{Z}_a\tilde{Z}_b\tilde{Z}_c\right) \\ &+ \frac{1}{T}\left(b^{ab}_1\tilde{Z}_a\tilde{Z}_b + b^{ab}_{2;cd,ef}\tilde{Z}_a\tilde{Z}_b\tilde{Z}^{cd}\tilde{Z}^{ef} + b^{abcd}_3\tilde{Z}_a\tilde{Z}_b\tilde{Z}_c\tilde{Z}_d + b^{abc}_{4;de}\tilde{Z}_a\tilde{Z}_b\tilde{Z}_c\tilde{Z}^{de}\right. \\ &\left. + b^{abc}_{5;def}\tilde{Z}_a\tilde{Z}_b\tilde{Z}_c\tilde{Z}^{def} + b^{abcdef}_6\tilde{Z}_a\tilde{Z}_b\tilde{Z}_c\tilde{Z}_d\tilde{Z}_e\tilde{Z}_f\right) + R_T. \end{aligned}$$

ここで，$a^{ab}_{1;cd}$, a^{abc}_2, b^{ab}_1, $b^{ab}_{2;cd,ef}$, b^{abcd}_3, $b^{abc}_{4;de}$, $b^{abc}_{5;def}$, b^{abcdef}_6 は $\tilde{Z}_a$, $\tilde{Z}^{ab}$, $\tilde{Z}^{abc}$ などのキュミュラントで表される定数で，統計量ごとに異なる．また，剰余項 R_T はある意味で小さい確率変数であるが，ここでは次の様な裾確率の評価を満たすものとする:ある定数 $C > 0$ と $\epsilon > 0$ に対して，

$$P_{\theta_0}\{|R_T| \geq CT^{(2+\epsilon)/2}\} = o(T^{-1}).$$

この様な確率展開を持つ検定統計量のクラスを $\mathscr{T}$ と表すことにする．

この統計量のクラス $\mathscr{T}$ に含まれる任意の検定統計量 V_T の局所的な検出力は，以下のように展開できる:

Theorem 1. $\theta = \theta_\varepsilon (:= \theta_0 + \varepsilon/\sqrt{T})$ のもとで，

$$P_{\theta_\varepsilon}[V_T < x] = \chi_{p,\delta}(x) + \frac{1}{\sqrt{T}}U_1(x) + \frac{1}{T}(U^{(1)}_2(x) - \partial_x U^{(2)}_2(x) - \partial_x U^{(3)}_2(x) + \frac{1}{2}\partial^2_x U^{(4)}_2(x)) + o(\frac{1}{T}).$$

ただし，$\chi_{p,\delta}(x)$ は，自由度 p, 非心母数 $\delta = g_{ab}\varepsilon^a\varepsilon^b$ の非心 χ^2 分布の分布関数で，

$$U_1(x) = \frac{1}{6}K^{abc}\tilde{H}^{\varepsilon,g}_{abc,p}(x) + \frac{1}{2}J_{ab,c}\varepsilon^{ab}g^{cd}\tilde{H}^{\varepsilon,g}_{d,p}(x) + \frac{1}{6}\bar{\nu}_{abc}\varepsilon^{abc}\chi_{p,\delta}(x) - a^{abc}_2\partial_x H^{\varepsilon,g}_{abc,p}(x)$$

$$\begin{aligned} U^{(1)}_2(x) =& \frac{1}{24}H^{abcd}\tilde{H}^{\varepsilon,g}_{abcd,p}(x) + \frac{1}{72}K^{abc}K^{def}\tilde{H}^{\varepsilon,g}_{abcdef,p}(x) + \frac{1}{12}K^{abc}g^{de}\varepsilon^{fg}J_{fg,e}(\tilde{H}^{\varepsilon,g}_{abcd,p}(x) + 3g_{ad}\tilde{H}^{\varepsilon,g}_{bc,p}(x)) \\ &+ \frac{1}{36}K^{abc}\varepsilon^{def}\bar{\nu}_{def}\tilde{H}^{\varepsilon,g}_{abc,p}(x) + \frac{1}{4}\tilde{N}^{c,d,ef}g_{ae}g_{bf}\varepsilon^{ab}\tilde{H}^{\varepsilon,g}_{cd,p}(x) \end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{24}\varepsilon^{abcd}(3\tilde{M}_{ab,cd}\chi_{p,\delta}(x)+4L_{abc,d}\chi_{p+2,\delta}(x)+\bar{\nu}_{abcd}\chi_{p,\delta}(x)+3J_{ab,e}J_{cd,f}g^{ee'}g^{ff'}H^{\varepsilon,g}_{e'f',p}(x))\\
&+\frac{1}{72}\bar{\nu}_{abc}\varepsilon^{abcdef}(6J_{de,f}\chi_{p+2,\delta}(x)+\bar{\nu}_{def}\chi_{p,\delta}(x))\\
U_2^{(2)}(x)=&\frac{1}{12}a_2^{abc}K^{def}(H^{\varepsilon,g}_{abcdef,p}(x)-3g_{ef}H^{\varepsilon,g}_{abcd,p}(x))+\frac{1}{4}a^{ab}_{1;cd}\tilde{N}^{cd,e,f}(H^{\varepsilon,g}_{abef,p}(x)-g_{ef}H^{\varepsilon,g}_{ab,p}(x))\\
&+\frac{1}{4}a^{ab}_{1,cd}g_{ee'}g_{ff'}\varepsilon^{ef}\tilde{M}^{cd,e'f'}H^{\varepsilon,g}_{ab,p}(x)+\frac{1}{4}a_2^{abc}J_{ef,d'}g^{dd'}\varepsilon^{ef}H^{\varepsilon,g}_{abcd,p}(x)+\frac{1}{6}a_2^{abc}\bar{\nu}_{def}\varepsilon^{def}H^{\varepsilon,g}_{abc,p}(x)\\
U_2^{(3)}(x)=&b_1^{ab}H^{\varepsilon,g}_{ab,p}(x)+b^{ab}_{2;cd,ef}\tilde{M}^{cd,ef}H^{\varepsilon,g}_{ab,p}(x)+b_3^{abcd}H^{\varepsilon,g}_{abcd,p}(x)+b_6^{abcdef}H^{\varepsilon,g}_{abcdef,p}(x)\\
U_2^{(4)}(x)=&a^{ab}_{1;a'b'}a^{cd}_{1;c'd'}\tilde{M}^{a'b',c'd'}H^{\varepsilon,g}_{abcd,p}(x)+a_2^{abc}a_2^{def}H^{\varepsilon,g}_{abcdef,p}(x).
\end{aligned}
$$

上の展開に現れる係数は，以下のように定義される：$K_{abc}=E[Z_aZ_bZ_c]$, $K^{abc}=g^{aa'}g^{bb'}g^{cc'}K_{a'b'c'}$, $J_{ab,c}=E[Z_{ab}Z_c]$, $J^{ab,c}=g^{aa'}g^{bb'}g^{cc'}J_{a'b',c'}$ $M_{ab,cd}=E[Z_{ab}Z_{cd}]$, $M^{ab,cd}=g^{aa'}g^{bb'}g^{cc'}g^{dd'}M_{a'b',c'd'}$, $\tilde{M}^{ab,cd}=M^{ab,cd}-J^{ab,e}g_{ef}J^{f,cd}$, $N_{ab,c,d}=E[Z_{ab}Z_cZ_d]$, $L_{abc,d}=E[Z_{abc}Z_d]$, $\tilde{N}^{a,b,cd}=N^{a,b,cd}-J^{cd,f}g_{ef}K^{abe}$, $H_{abcd}=\mathrm{Cum}[Z_a,Z_b,Z_c,Z_d]$, $H^{abcd}=g^{aa'}g^{bb'}g^{cc'}g^{dd'}H_{a'b'c'd'}$. また，

$$
\begin{aligned}
g^{(\varepsilon)}_{a_1\cdots a_k,j}&=\sum_{(a_1,\cdots,a_{k-2j},a_{k-2j+1}a_{k-2j+2},\cdots,a_{k-1}a_k)}g_{a_1a_1'}\cdots g_{a_{k-2j}a'_{k-2j}}\varepsilon^{a_1'}\cdots\varepsilon^{a'_{k-2j}}g_{a_{k-2j+1}a_{k-2j+2}}\cdots g_{a_{k-1}a_k},\\
\tilde{g}^{(\varepsilon)}_{a_1\cdots a_k,j}&=\sum_{\substack{s+l=j\\0\le l\le s\le[k/2]}}\binom{s}{l}(-1)^l g^{(\varepsilon)}_{a_1\cdots a_k,s},\\
H^{\varepsilon,g}_{a_1\cdots a_k,p}(x)&=\sum_{j=0}^{[k/2]}g^{(\varepsilon)}_{a_1\cdots a_k,j}\chi_{p+2k-2j,\delta}(x),\quad \tilde{H}^{\varepsilon,g}_{a_1\cdots a_k,p}(x)=\sum_{j=0}^{2[k/2]}\tilde{g}^{(\varepsilon)}_{a_1\cdots a_k,j}\chi_{p+2k-2j,\delta}(x),
\end{aligned}
$$

Taniguchi(1991), C.R.Rao-Mukerjee(1997) では，同様の展開を線形時系列などのモデルに対して，母数の次元が 1 の場合について与えている．

この結果を利用すると，様々な統計モデルに対する検定統計量の漸近展開が導くことができるが，特に拡散過程に対しては，次のように与えられる (ここでは，簡単のため，帰無仮説のもとにおける分布のみを与える)．

Theorem 2.

$$P[V_T<x]=\chi_{p,0}(x)+T^{-1}\sum_{i=0}^{3}C_i^*\chi_{p+2i,0}(x)+o(T^{-1}).$$

ただし，

$$
\begin{aligned}
C_0^*=&\frac{1}{24}H^*-\frac{1}{72}K^{**}+\frac{1}{4}a^{ab}_{1;ef}\tilde{N}^{*c,d,ef}\rho_{ab}\rho_{cd}-\frac{1}{2}(b_1^{ab}+b^{ab}_{2;cd,ef}\tilde{M}^{*cd,ef})\rho_{ab}\\
&+\frac{1}{8}a^{ab}_{1;ef}a^{cd}_{1;gh}\tilde{M}^{*ef,gh}\rho^{[3]}_{abcd},\\
C_1^*=&-\frac{1}{12}H^*+\frac{1}{24}K^{**}-\frac{1}{4}a^{ab}_{1;ef}\tilde{N}^{*c,d,ef}(\rho_{ab}\rho_{cd}+\rho^{[3]}_{abcd})+\frac{1}{4}a_2^{abc}K^{*def}\rho_{ef}\rho^{[3]}_{abcd}\\
&+\frac{1}{2}(b_1^{ab}\rho_{ab}+b^{ab}_{2;cd,ef}\tilde{M}^{*cd,ef}\rho_{ab}-b_3^{abcd}\rho^{[3]}_{abcd})+\frac{1}{8}(a_2^{abc}a_2^{def}\rho^{[15]}_{abcdef}-2a^{ab}_{1;ef}a^{cd}_{1;gh}\tilde{M}^{*ef,gh}\rho^{[3]}_{abcd}),\\
C_2^*=&\frac{1}{24}H^*-\frac{1}{24}K^{**}-\frac{1}{12}a_2^{abc}K^{*def}(3\rho^{[3]}_{abcd}\rho_{ef}+\rho^{[15]}_{abcdef})+\frac{1}{4}a^{ab}_{1;ef}\tilde{N}^{*c,d,ef}\rho^{[3]}_{abcd}\\
&+\frac{1}{2}(b_3^{abcd}\rho^{[3]}_{abcd}-b_6^{abcdef}\rho^{[15]}_{abcdef})+\frac{1}{8}(-2a_2^{abc}a_2^{def}\rho^{[15]}_{abcdef}+a^{ab}_{1;ef}a^{cd}_{1;gh}\tilde{M}^{*ef,gh}\rho^{[3]}_{abcd}),\\
C_3^*=&\frac{1}{72}K^{**}+\frac{1}{12}a_2^{abc}K^{def}\rho^{[15]}_{abcdef}+\frac{1}{2}b_6^{abcdef}\rho^{[15]}_{abcdef}+\frac{1}{8}a_2^{abc}a_2^{def}\rho^{[15]}_{abcdef}
\end{aligned}
$$

係数 H^* などは，*Sakamoto-Yoshida(1998, 1999)* を参照．

ランダム極限モデルの漸近展開によるオプション価格の近似と数値実験

東京大学数理 柏倉賢司
東京大学数理、九州大学数理 吉田朋広

1 Random Limit Expansion for Small Diffusion Processes

consider a $d = d(1) + d(2)$-dimensional diffusion processes
$X^{\epsilon} = (X^{(1),\epsilon}, X^{(2),\epsilon})_{t\in[0,T]}$ is defined by the stochastic differential equations:

$$\begin{cases} dX_t^{(1),\epsilon} &= V_0^{(1)}(X_t^{\epsilon},\epsilon)dt + V^{(1)}(X_t^{\epsilon},\epsilon)dw_t^{(1)} \\ dX_t^{(2),\epsilon} &= V_0^{(2)}(X_t^{(2),\epsilon},\epsilon)dt + V^{(2)}(X_t^{(2),\epsilon},\epsilon)dw_t^{(2)}, \\ X_0^{(1),\epsilon} &= x_0^{(1)} \\ X_0^{(2),\epsilon} &= x_0^{(2)}, \end{cases}$$

where :
$w^{(1)}$: $r(1)$-dimensional Wiener processes,
$w^{(2)}$: $r(2)$-dimensional Wiener processes,
coefficients are all in class C_b^{∞},(namely,smooth in (x,ϵ) with bounded $\partial_x^i\partial_\epsilon^j$-derivatives for $i \geq 1$ and $j \geq 0$).
Moreover,we assume that

$$V^{(2)}(\cdot,0) = 0 \quad \text{equivalently.}$$

We here consider a functional Z^{ϵ} defined by

$$Z^{\epsilon} = \int_0^T \beta(X_t^{\epsilon},\epsilon)\nu(dt),$$

where $\beta \in C_{\uparrow}^{\infty}(\mathbf{R}^d \times [0,1];\mathbf{R}^k)$ and ν is a random measure on $[0,T]$.
Functionals Z^{ϵ} have stochastic expansions:

$$Z^{\epsilon} \sim Z^{(0)} + \epsilon Z^{(1)} + \epsilon^2 Z^{(2)} + \cdots \text{in } D_{\infty}(\mathbf{R}^k)$$

as $\epsilon \to 0$ with $Z^{(1)}, Z^{(2)}, \cdots \in D_{\infty}(\mathbf{R}^k)$.
Under the non-degeneracy condition for $Z^{(0)}$,one obtains the asymptotic expansion:

$$P[\mathbf{T}(Z^{\epsilon})] \sim P[\Phi^{(0)}] + \epsilon P[\Phi^{(1)}] + \epsilon^2 P[\Phi^{(2)}] + \cdots$$

as $\epsilon \to 0$ for any measurable function $\mathbf{T}$ of at most polynomial growth.

2 Simurations

○ Example 1

$$\begin{cases} dX_t^{\epsilon} &= (\mu + \beta v_t^{\epsilon})dt + c(v_t^{\epsilon})^{\frac{1}{2}}dw_t \\ dv_t^{\epsilon} &= -\theta(v_t^{\epsilon} - \alpha)dt + \epsilon(v_t^{\epsilon})^{\frac{1}{2}}d\tilde{w}_t \\ X_0^{\epsilon} &= x_0 \\ v_0^{\epsilon} &= v_0 \end{cases}$$

ここで、$\alpha,\beta,\theta,\mu,c$ は定数、$\tilde{w}=\rho w+\sqrt{1-\rho^2}w^*$, ρは定数, w と w^*は独立

◇シミュレーション結果 $(\mu=0.5,\beta=0.4,c=2.5,\theta=2.0,\alpha=1.5,\rho=0.5)$

$Z^\epsilon=X_T^\epsilon, T(z)=(z-K)_+$

$x_0=v_0=10,K=20,\rho=0.5$			
ϵ	0.3	0.7	1.0
(1) MC	0.402530	0.436347	0.462579
(2) I	0.377184	0.377184	0.377184
Difference	0.025346	0.059163	0.085395
Diff.rate %	6.3	13.6	18.5
(3) II	0.400558	0.431724	0.455100
Difference	0.001972	0.004622	0.007479
Diff.rate %	0.48	1.1	1.6
(4) III	0.401728	0.438095	0.468100
Difference	0.000802	-0.001748	-0.005521
Diff.rate %	0.19	0.40	1.2

MC=Monte Caro 10000000 回

I=First order,II=Second order,III=Third order

$Z^\epsilon=\frac{1}{T}\int_0^T X_t^\epsilon dt, T(z)=(z-K)_+$

$x_0=10,v_0=20,K=15,\rho=0.5$			
ϵ	0.3	0.7	1.0
(1) MC	1.162432	1.184568	1.201118
(2) I	1.145018	1.145018	1.145018
Difference	0.017414	0.039550	0.056100
Diff.rate %	1.5	3.3	4.7
(3) II	1.161097	1.182536	1.198616
Difference	0.001335	0.002032	0.002502
Diff.rate %	0.11	0.17	0.20
(4) III	1.160806	1.180949	1.195375
Difference	0.001626	0.003619	0.005743
Diff.rate %	0.13	0.30	0.47

MC=Monte Caro 10000000 回

I=First order,II=Second order,III=Third order

○ Example 2

$$\begin{cases} dX_t^{(1),\epsilon} &= \mu X_t^{(1),\epsilon}dt+\sigma X_t^{(1),\epsilon}(1+\epsilon\Sigma(X_t^{(1),\epsilon}))dw_t^{(1)} \\ X_0^{(1),\epsilon} &= x_0 \end{cases}$$

ここで、μ,σ は定数、Σ は関数

◇シミュレーション結果 $(Z^\epsilon=X_T^\epsilon, T(z)=(z-K)_+)$

$\Sigma(x)=x^{-\frac{1}{2}}$

$\mu=0.05,\sigma=0.5,x_0=K=10,$			
ϵ	0.1	0.3	0.5
(1) MC	2.374894	2.490895	2.616581
(2) I	2.290992	2.290992	2.290992
Difference	0.073902	0.199903	0.325589
Diff.rate%	3.1	8.0	12.4
(3) II	2.352858	2.476691	2.600323
Difference	0.012036	0.014304	0.016258
Diff.rate %	0.50	0.57	0.62

MC=Monte Caro 1000000 回

I=First order,II=Second order,III=Third order

$\epsilon=0.3$

$\mu=0.1,\sigma=0.3,x_0=K=5$			
$\Sigma(x)$	$\exp(-x)$	$x^{\frac{1}{2}}$	$x^{-\frac{1}{2}}$
(1) MC	0.925646	1.331432	1.006538
(2) I	0.924703	0.924703	0.924703
Difference	0.000943	0.406729	0.081835
Diff.rate %	0.09	30.6	8.24
(3) II	0.926075	1.333830	1.005026
Difference	-0.000429	-0.002379	0.001512
Diff.rate %	-0.04	-0.32	0.15

MC=Monte Caro 1000000 回

I=First order,II=Second order,III=Third order

参考文献

[1] Sørensen,M. and Yoshida,N(2000).Random Limit Expansion for Small Diffusion Processes,Unpublished manuscript.

[2] Yoshida,N.(1992).Asymptotic expansion for statistics related to small diffusions,J.J.S.S.22,139-159.

[3] Yosida,N.(1999).Pertubation methods and option pricing,Proc.East Asian Symposium on Statistics.Feb.28-29,2000.University of Tokyo.

時間領域の情報のみを用いた強従属過程における中心極限定理とその応用

早大政経　加藤 剛

1　時間領域の情報だけを用いた推定方法

$\{X(t):t\in\boldsymbol{R}\}$ を $E[X(t)]=0$ なる実数値定常 Gauss 過程とし，その相関関数が，ある定数 $d>0$, $C>0$ について

$$R(\tau):=E[X(t+\tau)X(t)]\ \sim\ C|\tau|^{-d}L(\tau)\quad(|\tau|\to\infty)\tag{1}$$

を満たすとする．ここで，$L(\tau)$ は slowly varing な関数である．$d\in(0,1)$ のとき $R(x)\notin L^1(\boldsymbol{R})$ なので，この場合は，$\{X(t)\}$ は強従属過程または長期記憶モデルと呼ばれる．また，d のことを, memory parameter または fractional differencing parameter という．$\{X(t)\}$ がスペクトル密度 $f(\lambda)$ を持つとき，(1) は，f がある定数 $C'>0$ について

$$f(\lambda)\ \sim C'|\lambda|^{d-1}L(|\lambda|^{-1})\quad(|\lambda|\to0)\tag{2}$$

を満たすことと同値である．

パラメータ d のセミパラメトリック推定法については，これまでにも多くの結果が出されている．それらの論文に共通することは，周波数領域の情報，より具体的にはスペクトル密度に関する情報 (2) を何らかの形で使うことである．

これに対し，Kato and Masry (2000) は，推定量の構成からその漸近的性質の証明までのすべてを，時間領域の情報，つまり共分散関数とその挙動 (1) だけを用いて行うことを試みた．ただし，$L(\tau)$ は定数であるとする．発想は単純で，定数 $q>0$ を1つ固定するとき，(1) から

$$R(q\tau)\ \sim C(q\tau)^{-d}\quad(\tau\to\infty)$$

なので，

$$\frac{1}{\log q}\log\frac{R(\tau)}{R(q\tau)}\ \sim\ d\quad(\tau\to\infty)\tag{3}$$

が成り立つ．これから，$R(\tau)$ と $R(q\tau)$ の推定量から d の推定量が構成できることが示唆される．実際，n とともに発散する数列 $\{\tau_n\}$ をうまく選んで定めた $R(\tau_n)$ と $R(q\tau_n)$ の推定量

$$Y_{1,n}=\frac{1}{n}\sum_{k=1}^{n-\tau_n}X(k)X(k+\tau_n),\qquad Y_{2,n}=\frac{1}{n}\sum_{k=1}^{n-q\tau_n}X(k)X(k+q\tau_n)$$

を用いると，

$$\hat{d}_{q,n}=\frac{1}{\log q}\log\frac{Y_{1,n}}{Y_{2,n}}\quad(q\geq2)$$

が $d>\frac{1}{2}$ の一致推定量になることが証明できる．

一致性の次は，$\hat{d}_{q,n}$ の漸近正規性を示したい．この問題は，$n \times n$ 行列 F_n を

$$F_n := (f_{i-j}^{(n)})_{0\le i,j\le n-1}, \qquad f_k^{(n)} := \begin{cases} q^d & : \ |k| = q\tau_n \\ -1 & : \ |k| = \tau_n \\ 0 & : \ \text{その他} \end{cases}$$

で定めた２次形式を含む確率変数 $Q_n = (X_n' F_n X_n - E[X_n' F_n X_n])/(2n)$ の漸近正規性を示すことに帰着される．

2 時間領域の情報を用いた中心極限定理

実数値定常 Gauss 過程の２次形式に関する中心極限定理は，(2) を利用した Fox and Taqqu (1987) の結果がよく知られている．しかし，行列 F_n の構造から，彼らの結果を Q_n の漸近正規性の証明に適用することはできない．そこで，Kato and Masry (2000) は，もともと議論のすべてを時間領域で行うという方針もあって，以下のような別の証明を与えることにした．

確率変数 X の p 次キュムラントを $\mathrm{cum}_p(X)$ で表すことにすると，$\mathrm{cum}_1(\sqrt{n}\,Q_n) = 0$，および $p = 2, 3, \ldots$ について

$$\begin{aligned}\mathrm{cum}_p(\sqrt{n}\,Q_n) &= \frac{(p-1)!}{2\,n^{p/2}}\,\mathrm{tr}\{(R_n F_n)^p\} \\ &= \frac{(p-1)!}{2\,n^{p/2}} \sum_{j_1=0}^{n-1}\sum_{j_2=0}^{n-1}\cdots\sum_{j_{2p}=0}^{n-1} R(j_1-j_2) f_{j_2-j_3}^{(n)} R(j_3-j_4) f_{j_4-j_5}^{(n)} \cdots R(j_{2p-1}-j_{2p}) f_{j_{2p}-j_1}^{(n)}\end{aligned}$$

が成り立つ．これに (1) から導かれる $|R(\tau)| \le C_1/(1+|\tau|)^d$ ($C_1 > 0$ は定数) と不等式

$$\sum_{k=-\infty}^{\infty} (1+|x-k|)^{-d}(1+|y-k|)^{-d} \le \begin{cases} M(1+|x-y|)^{1-2d} & : \ \frac{1}{2} < d < 1 \\ M(1+|x-y|)^{-1}\log(e+|x-y|) & : \ d = 1 \\ M(1+|x-y|)^{-d} & : \ d > 1 \end{cases}$$

を適用すると ($M > 0$ は定数)，$\{\tau_n\}$ を $\lim_{n\to\infty}\tau_n = \infty$，$\lim_{n\to\infty}\tau_n/n = 0$ を満たすように選ぶとき，

$$\lim_{n\to\infty} \mathrm{cum}_2(\sqrt{n}\,Q_n) = (q^{2d}+1)\sum_{k=-\infty}^{\infty}\{R(k)\}^2, \qquad \lim_{n\to\infty}\mathrm{cum}_p(\sqrt{n}\,Q_n) = 0, \quad (p = 1, 3, 4, \ldots)$$

の成り立つことが示せる．したがって，$\sqrt{n}\,Q_n$ は，平均 0，分散 $(q^{2d}+1)\sum_{k=-\infty}^{\infty}\{R(k)\}^2$ の正規分布に分布収束する．

参考文献

[1] Fox, R. and Taqqu, M.S. (1987). Central limit theorems for quadratic forms in random variables having long-range dependence. *Probab. Th. Rel. Fields.* 74, 213-240.

[2] Kato, T. and Masry, E. (2000). Gaussian semiparametric estimation in the time-domain for strongly dependent stationary processes. (submitted)

Edgeworth approximation in the O-U process

北海道大学大学院経済学研究科　柿沢　佳秀

1 はじめに

本報告では Ornstein-Uhlenbeck (O-U) 拡散過程

$$dX_t = \theta X_t dt + dW_t \tag{1}$$

の未知母数 $\theta \in (-\infty, 0)$ の２つの推定量

$$\widehat{\theta}_{1,T} = \int_0^T X_t\,dX_t \Big/ \int_0^T X_t^2\,dt\,, \quad \widehat{\theta}_{2,T} = -(T/2) \Big/ \int_0^T X_t^2\,dt \tag{2}$$

を扱う．ここで W_t は標準ブラウン運動とし，初期条件 (I) $X_0 = x$ (定数) または (II) $X_0 \sim N(\,0, 1/(-2\theta)\,)$ を仮定する．よく知られているように $\widehat{\theta}_{1,T}$ と $\widehat{\theta}_{2,T}$ は一致推定量であり，同じ漸近正規性をもつ: $\{T/(-2\theta)\}^{1/2}(\widehat{\theta}_{a,T} - \theta) \xrightarrow{d} N(0,1)$, $a = 1, 2$.

初期条件 $X_0 = 0$ の下で $\widehat{\theta}_{1,T}$ の分布の収束の速さを論じた論文がある．例えば, Mishra and Prakasa Rao (1985; rate $T^{-1/5}$), Bose (1986; rate $T^{-1/2}(\log T)^2$), Bishwal and Bose (1995; rate $T^{-1/2}(\log T)^{1/2}$) である．いずれもよく知られた Berry-Esseen 型評価に現れる rate $T^{-1/2}$ よりも劣っていたが, 最近 Bishwal (2000) は $T^{-1/2}$ へ改善できることを指摘した．

柿沢 (2000 年数学会/早稲田大学) は初期条件 (I) または (II) を仮定して２つの推定量の分布の３次の漸近展開を求めたが, 本報告では正当化の議論も含めて

$$P_\theta\Big[\Big\{\frac{T}{(-2\theta)}\Big\}^{1/2}(\widehat{\theta}_{a,T} - \theta) \le y\Big] = \Phi(y) + T^{-1/2}\Lambda_{a,1}(y)\phi(y) + T^{-1}\Lambda_{a,2}(y)\phi(y) + \cdots \tag{3}$$

を再考察する．

漸近展開の正当化の話題は近年 Malliavin 解析の導入により進展しており, 広汎な確率過程に対して適用可能になっている (楠岡・吉田・阪本氏の一連の論文がある)．拡散過程はその典型例であるが, ここでは Malliavin 解析を使わずに (3) の正当化を試みる．結果として以下の定理が初期条件 (I) または (II) の下で証明される．

定理 1 任意の $r \ge 4$ を固定する．『$\deg\{\Lambda_{a,j}(y)\} \le 3j - 1$ かつ, 奇数 j のとき偶関数, 偶数 j のとき奇関数』であるような多項式 $\Lambda_{a,j}(y)$ が存在し

$$\sup_{y\in R}\Big|P_\theta\Big[\Big\{\frac{T}{(-2\theta)}\Big\}^{1/2}(\widehat{\theta}_{a,T} - \theta) \le y\Big] - \Phi(y) - \sum_{j=1}^{r-3} T^{-j/2}\Lambda_{a,j}(y)\phi(y)\Big| = \mathrm{o}(T^{-(r-3)/2})\,.$$

2 形式的な漸近展開

２つの推定量 $\widehat{\theta}_{1,T}$ と $\widehat{\theta}_{2,T}$ はそれぞれ推定関数

$$G_{1,T}(\theta) = \int_0^T X_t\,dX_t - \theta\int_0^T X_t^2\,dt\,, \quad G_{2,T}(\theta) = -\int_0^T \Big(\theta X_t^2 + \frac{1}{2}\Big)\,dt$$

に対応することに注意しよう．推定量が比の統計量であることから

$$P_\theta\Big[\Big\{\frac{T}{(-2\theta)}\Big\}^{1/2}(\widehat{\theta}_{a,T}-\theta)\le y\Big]=P_\theta\Big[G_{a,T}\Big(\theta+y\Big\{\frac{(-2\theta)}{T}\Big\}^{1/2}\Big)\le 0\Big]$$

の漸近展開を求めればよい．ここで

$$Z_{a,T}(y)=\Big\{\frac{(-2\theta)}{T}\Big\}^{1/2}G_{a,T}\Big(\theta+y\Big\{\frac{(-2\theta)}{T}\Big\}^{1/2}\Big)+y\xrightarrow{d}N(0,1)$$

に注目すれば, $P_\theta[\ Z_{a,T}(y)\le z\ ]$ の漸近展開を求める問題に帰着された．初期条件 (I) または (II) の O-U 拡散過程 (1) に対して積率母関数

$$\begin{aligned}&E_\theta\Big[\exp\Big\{\lambda_1\Big(\int_0^T X_t^2\,dt+\frac{T}{2\theta}\Big)+\lambda_2\Big(\int_0^T X_t\,dX_t+\frac{T}{2}\Big)\Big\}\Big]\\&\quad=\exp\Big\{TF\Big(\frac{\lambda_1}{\theta^2}\Big)\Big\}G\Big(\frac{\lambda_1}{\theta^2},\frac{\lambda_2}{(-\theta)}\Big)\Delta_T\Big(\frac{\lambda_1}{\theta^2},\frac{\lambda_2}{(-\theta)}\Big),\quad(\lambda_1,\lambda_2)'\in\mathbf{R}^2\end{aligned}\tag{4}$$

が少なくとも $\lambda_1<\theta^2/2$, $|\lambda_2|<(-\theta)$ で存在しその解析的表現が利用できる．柿沢 (2000 年数学会/早稲田大学) は特性関数 $E_\theta[\exp\{isZ_{a,T}(y)\}]$ の 3 次まで (原理的には任意オーダーまで可能) の漸近展開を反転して (3) に現れる多項式 $\Lambda_{a,1}(y)$, $\Lambda_{a,2}(y)$ を与えた (具体的な式は省略します)．

3　漸近展開 (3) の正当化

２つの確率変数

$$U_1=\frac{(-\theta)(-2\theta)^{1/2}}{T^{1/2}}\Big(\int_0^T X_t^2\,dt+\frac{T}{2\theta}\Big),\quad U_2=\frac{(-2\theta)^{1/2}}{T^{1/2}}\Big(\int_0^T X_t\,dX_t+\frac{T}{2}\Big)$$

を定義すると

$$\Big\{\frac{T}{(-2\theta)}\Big\}^{1/2}(\widehat{\theta}_{1,T}-\theta)=\frac{U_1+U_2}{1+2U_1/(-2\theta T)^{1/2}},\quad\Big\{\frac{T}{(-2\theta)}\Big\}^{1/2}(\widehat{\theta}_{2,T}-\theta)=\frac{U_1}{1+2U_1/(-2\theta T)^{1/2}}$$

と書け, $Z_{1,T}(y)=\{1-2y/(-2\theta T)^{1/2}\}U_1+U_2$, $Z_{2,T}(y)=\{1-2y/(-2\theta T)^{1/2}\}U_1$ に注意しよう．2 節での形式的な議論を以下の手順で正当化する．

(i) 多項式 $\Gamma_{a,j}(z;y)$ が存在し

$$\sup_{z\in R}\Big|P_\theta[\ Z_{a,T}(y)\le z\ ]-\Phi(z)-\sum_{j=1}^{r-3}T^{-j/2}\Gamma_{a,j}(z;y)\phi(z)\Big|=\mathrm{o}(T^{-(r-3)/2})$$

が $|y|<K(\log T)^{1/2}$ で一様に成立することを証明する．その過程にて多項式 $\Gamma_{a,j}(y;y)=\Lambda_{a,j}(y)$ の性質が明らかになる．

(ii) 裾確率の評価

$$P_\theta\Big[\Big\{\frac{T}{(-2\theta)}\Big\}^{1/2}|\widehat{\theta}_{a,T}-\theta|\ge(r-2)^{1/2}(\log T)^{1/2}\Big]=\mathrm{o}(T^{-(r-3)/2})$$

を証明する．$\Phi(\pm y)$ と $|y|^\ell\phi(y)$ の振る舞いのおかげで $|y|\ge(r-2)^{1/2}(\log T)^{1/2}$ でも定理 1 が成立する．

錐型の特異点を持つモデルにおける最尤推定量の漸近的挙動

統計数理研究所　福水健次

本研究は、真のパラメータが識別不能な場合における最尤推定量の漸近的挙動を解析するための、一般的な枠組みを確立することを目的としている。パラメトリックな統計的推測問題においては、最尤推定量の漸近的挙動を解析する際、その正則条件のひとつに、真のパラメータが一意的に定まるという仮定を課すことが一般的である。例えば、漸近正規性の議論では、真のパラメータが一点であるという仮定が本質的である。しかしながら、有限混合モデル、縮小ランクなどの多くの重要な統計モデルでは、真の確率密度関数が、設定したモデルよりも小さいサイズのモデルに含まれている場合に、モデルのパラメータ空間内で真の密度関数を表すパラメータ集合が高次元連続集合になることが知られている。このような場合には、通常の漸近有効性の議論は適用できず、その最尤推定量の漸近的挙動は未知の問題を多く含んでいる。実際、有限混合分布のモデルサイズの尤度比検定では、一般的な極限分布は知られていないのが実情である。

本研究では、このような識別不能性を持つモデルとして、特に多層ニューラルネットワークを念頭においた枠組みを示す。ニューラルネットワークは、有限混合分布モデルなどと同様、パラメータの識別不能性を持つモデルであることが知られている。

識別不能なモデルの特別な定式化として、「局所錐型モデル」を用いた。d 次元のパラメータ θ を持ったモデルを $\{f(x;\theta)\}$ と書くとき、これが局所錐型モデルであるとは、パラメータが $\theta=(a, b)$ というように d-1 次元のパラメータ a と 1 次元のパラメータ b への分解を持ち、真の確率密度関数 $f_0(x)$ が $b=0$ という条件で表され、かつ、各 a を固定して b をパラメータだと考えた 1 次元部分モデルが識別可能になっていることを意味する。真のパラメータは $\{(a,0)\mid a$ は任意$\}$ という集合になるので、このモデルは識別不能性を持っている。このような局所錐型モデルに対して、真の密度 $f_0(x)$ から独立に発生した n 個のサンプル $X_1, \ldots, X_n$ による最尤推定量の対数尤度比

$$\sup_\theta L_n(\theta)=\sup_\theta\sum_{i=1}^{n}\log\frac{f(X_i;\theta)}{f_0(X_i)}$$

が漸近的にどのように振舞うかを考察した。特に、サンプル数 n に対する対数尤度比のオーダーに関して、後に述べるいくつかの結果を得た。漸近有効性の正則条件が成立する場合には、対数尤度比の漸近的挙動はよく知られており、パラメータの数の自由度を持つ χ^2 分布に法則収束する。ところが、識別不能性を持つ場合にはそのオーダーすら異なる場合がある。

まず、Dacunha-Castelle & Gassiat (1997)に従って、局所錐型モデルの最尤推定量の対

数尤度比が、錐の底の上で定義される、ある経験過程の２乗の sup を用いて表現できることを述べた。ここで、錐の底とは、局所錐型モデルを確率密度関数の空間で考えた際の、真の密度における単位接ベクトルからなる集合である。これはまた、各 a に対する部分モデルのスコア関数からなる集合とも言い換えられる。Dacunha-Castelle & Gassiat (1997)では、この経験過程がガウス過程に法則収束する場合のみを議論しているが、本研究では、これが収束しないような場合の考察も行った。

研究の結果として、上で述べた経験分布がガウス過程によい意味で法則収束する場合には、対数尤度比のオーダーは $O_p(1)$ であること、および、そうでない場合には、$O_p(1)$ よりも真に大きいオーダーが現れうることが明らかとなった。また、Hartigan (1985)のアイデアを拡張することにより、錐の底の中に０に確率収束するような関数列が存在すれば、対数尤度比が $O_p(1)$ よりも真に大きいオーダーを持つという定理を示した。これは、通常より遅いオーダーを持つための非常に有用な十分条件である。

以上で得られた一般的枠組みを多層ニューラルネットの場合に適用した。多層ニューラルネットでは、真の関数が、設定したモデルよりも少ない中間素子数で実現できれば、真のパラメータに識別不能性を持つ。まず、この識別不能性が局所錐型モデルとして表現できることを示した。次に、モデルの持つ中間素子数が、真の関数を表現するために必要な中間素子数より２個以上大きいときには、上で述べた定理の条件が満足され、最尤推定量の対数尤度比は $O_p(1)$ より大きいオーダーを持つことを示した。さらに、このオーダーを詳しく解析し、$O_p(\log n)$ 以上のオーダーを持つことを示した。この結果は、Hagiwara, Kuno, & Usui (2000) の拡張となっている。また、ある特別なタイプのニューラルネットでは、モデルの中間素子数が１個だけ冗長な場合には、最尤推定量の対数尤度比が通常と同様 $O_p(1)$ になることを示した。

以上述べたように、本研究では、識別不能性を持つモデルの最尤推定を解析するために、局所錐型モデルを導入してその理論解析を行い、最尤推定量の対数尤度比のオーダーに関して一般的な結果を得た。また、その結果をニューラルネットに適用した。しかし、識別不能なモデルに対する最尤推定量の対数尤度比の漸近分布に関しては、まったく未知の部分が多い。今後は、さらに詳しい分布論へと研究を展開していきたい。

相対レニーエントロピーによる非正則モデルの漸近推定について

林 正人 [1]
理化学研究所[2] 脳科学総合研究センター 脳数理研究チーム

本研究では1パラメータの確率分布族のパラメータ推定の問題を後に述べる相対レニーエントロピーを用いて漸近理論の枠組みで扱う．そのために相対レニーエントロピーに関する議論をまとめた．さらに個々のモデルについて相対レニーエントロピーの極限を評価する．このとき通常の Fisher 情報量の定義では発散するモデルで発散しないように Fisher 情報量を再定義する．そして相対レニーエントロピーと帰無仮説及び対立仮説がともに単純である仮説検定との関係について述べる．それらの準備の下で正則なモデルだけでなく異なる未知パラメータに対応する確率分布が互いに絶対連続にならないという意味での非正則[3] なモデルを含む一般的なモデルについて漸近論とりわけ大偏差型の評価や極限分布の視点から考察する．上述の意味での非正則なモデルについてなされた研究は Akahira, Akahira and Takeuchi, Ibragimov and Has'minskii によるものなどがあるが正則な場合に比べると極めて少ない．本研究では相対レニーエントロピーを用いることによって非正則なモデルを含む一般的な状況の下で漸近論を展開する．

誤差を推定値が真値よりも大きくなる誤差（右側誤差）と推定値が真値よりも小さくなる誤差（左側誤差）に別けて議論し, それらを同等な評価のみではなく, ある重みをつけた評価も同時に行う．このような議論を行うことによって両者の trade off を扱うことができる．一般には一様分布族の場合最大値と最小値の中間の値（標本中点）が最適な推定量であると認識されている．しかし本研究で扱った一様分布族などの非正則なモデルにおいては最適な推定量は一意には決まらず最適性が重みの取り方に依存することの方が一般的であることが確かめられた．この現象は一様分布族において顕著にあらわれる．(大偏差, 極限分布双方に評価方法の下で確かめられた．同時に, 正則なモデルもしくは正則なモデルに近い場合は推定量の最適性が重みの取り方に依存しないことが確かめられた．これは, 正則なモデルについては従来の trade off を扱わない定式化で全く問題がないことを示している．

続いて, 右側誤差と左側誤差を同等に評価する枠組みの下で位置共変モデルについてより詳しい考察を行った．ある極限操作の取り替えに関してある病的な現象に注目した．この病的な現象は大偏差型の評価及び極限分布どちらの評価を採用してもモデルが非正則である場合に限って起る．以下この議論を大偏差の場合に限ってこの現象の概略を述べる．大偏差型の評価では一般

[1] e-mail masahito@brain.riken.go.jp
[2] 〒 351-0198 埼玉県和光市広沢 2-1

[3] パラメータ空間が特異点を持つという意味での非正則なモデルに関する研究があるが, 本研究では一貫して上述の意味での非正則なモデルのみ扱うことにする．

に真値から推定値が ϵ 以上ずれる確率の rate を ϵ の関数として考える. そして個々の推定量の rate の $\epsilon \to +0$ の極限を調べる. 一方それとは異り ϵ を固定する毎に推定量についてその rate の上限を取り, その上限の $\epsilon \to +0$ の極限を考えることもできる. 両者は極限 $\epsilon \to +0$ と sup の順序が異るだけである. 従って安易に極限の取り換えを許すのであれば両者は一致する. 実際, 正則なモデルにおいては両者は一致する. しかしながら, 非正則なモデルの場合には必ずしも両者は一致せず, 異る場合の方が一般的である. 本研究で注目した例については推定量を位置共変なものに限定して考えたときに両者が一致しないことが確かめられた. ただし, モデルが非正則である場合に必ずそのような病的な現象が起きるとは限らず正則なモデルに近いときやモデルが対称であるときは両者は一致する.

この相違は以下に述べるように点推定と区間推定の相違という観点から捉えることができる. まず, ϵ を固定したときの大偏差の評価は区間推定に対応していると考えることができ, $\epsilon \to +0$ の極限を考えたときの大偏差の意味で最適な推定量を考えることは点推定に対応していると考えることが出来る. したがって上記の 2 つの量の不一致から点推定が区間推定の極限として捉えることができないと考えられる. また逆に正則なモデルのときのように両者が一致する場合についは点推定を区間推定の極限として捉えることができる. 極限分布における同様の考察においても同種の 2 つの極限式を考えることができる. そして非正則の場合には一般には両者が異ることが同様に確かめられ, 正則なモデルもしくは正則なモデルに近い場合はそのような両者が一致することが確かめられる. しかしながら, 大偏差の場合のように両者の違いを区間推定と点推定の相違に結び付けるという解釈ができない.

そのほか大偏差型の評価に限って様々な議論を扱った. 右側誤差と左側誤差を同等に評価する枠組みで 2 次の漸近理論も扱った. 正則なモデルでの大偏差型評価の下での高次の漸近論は Fu により議論されているが本研究で扱うような非正則なモデルでの大偏差型評価の下での高次の漸近論は著者の知る限りなされていない. しかし, 残念ながらこれについては達成可能な限界をほとんど求めることができず今後の課題となる.

そして大偏差型の評価における超有効性についても触れておいた. モデルが正則であるときには大偏差型の評価については相対エントロピー (Kullback-Leibler の divergence) の単調性から限界を導くことができるため, MSE と異なり超有効性は起きない. しかしながら, モデルが非正則である場合は大偏差型の評価においても超有効な推定量が存在することが Ibragimov and Has'minskii により指摘されている. ここでは本研究で行った定式化と超有効性との関係について議論した.

Asymptotic Expansion and Efficiency of One-Step Estimator

名大・工　阪本雄二

東大・数理　吉田朋広

母数空間 $\Theta \subset \mathbb{R}^p$ に値をとる未知母数 θ_0 に対する推定問題を考える．母数空間 Θ と標本空間 $\mathfrak{X}_T$, $T>0$, の直積空間 $\Theta\times\mathfrak{X}_T$ 上で定義された $\mathbb{R}^p$-値可測関数 ψ_T に対して，(非線形) 方程式 $\psi_T(\hat{\theta}_T; X_T^{\theta_0})=0$ の解 $\hat{\theta}_T$ を，標本 $X_T^{\theta_0}$ に対する M-推定量と呼ぶことが多いが，一般に，このような推定量を構成することは困難である．したがって，実際には，適当な一致推定量 $\hat{\theta}_{T,0}$ を初期値として，線形化した方程式

$$\psi_T(\hat{\theta}_{T,i-1}; X_T^{\theta_0}) + \frac{\partial \psi_T}{\partial \theta}(\hat{\theta}_{T,i-1}; X_T^{\theta_0})(\hat{\theta}_{T,i} - \hat{\theta}_{T,i-1}) = 0. \tag{1}$$

を逐次解くことによって $\hat{\theta}_{T,i}$, $i=1,2,\dots$, を求め，適当な基準で逐次計算をうち切り，そのとき得られた $\hat{\theta}_{T,m}$ を $\hat{\theta}_T$ の近似値として用いることが多い．本報告では，このような逐次計算で求められた $\hat{\theta}_{T,2}$ の漸近展開を求め，その有効性を検証する．

標本が，幾何混合性を持つマルコフ過程 $X_T^{\theta_0}$ の実現値であり，また，推定関数 $\psi_T(\theta_0; X_T^{\theta_0})$ がその加法的汎関数であると仮定する．そのとき，推定関数 $\psi_T(\theta_0; X_T^{\theta_0})$ の母数微分に対するモーメント条件とその分布に関する正則条件の下で (Kusuoka-Yoshida(2000), Sakamoto-Yoshida(1999) を参照)，以下のような漸近展開が得られる．

Theorem 1. 定数 $c>0$, $\tilde{C}>0$, $\tilde{\epsilon}>0$ が存在して，高々多項式の発散のオーダーをもつ可測関数 f に対して，

$$\left| E[f(\sqrt{T}(\hat{\theta}^*_{T,2} - \theta))] - \int dy^{(0)} f(y^{(0)}) q_{T,2}(y^{(0)}) \right| \le c\omega(f, \tilde{C}T^{-(\tilde{\epsilon}+2)/2}, \hat{g}) + o(T^{-1}). \tag{2}$$

ただし，$q_{T,2}$ はある符号付き測度で，ω は f の連続度にかかわるある測度である *(それらの具体的な表現については，Sakamoto-Yoshida(1999) を参照).*

Theorem 2. 母数の次元 p が 1 であるとし，X_T が正の密度 $p_T(X_T;\theta)$ を持ち，$\psi_T = \delta \log p_T$ が上の定理の条件を満たすものとする．そのとき，$\hat{\theta}_{T,2}$ は，最尤推定量と同じバイアス修正を施すと，赤平-竹内の意味で 2 次有効になる．

Example 1 (counting process) 強度 $f_T(t,\theta)$, $\theta \in (\alpha,\beta)$, を持つ計数過程 $X_T=(X_t)_{t\in[0,T]}$ を考える．このとき，対数尤度 $\ell(\theta)$ は，

$$\ell_T(\theta) = \int_0^T \log f_T(t,\theta) dX_t - \int_0^T f_T(t,\theta)dt,$$

で与えられる．推定関数を $\psi_T = \delta \ell_T$ として求められる $\hat{\theta}_{T,2}$ は，

$$\beta(\theta) = -\frac{1}{6T}\int_0^T (\delta \log f_T(t,\theta))^3 f_T(t,\theta)dt \Bigg/ \left(\frac{1}{T}\int_0^T (\delta \log f_T(t,\theta))^2 f_T(t,\theta)dt\right)^2.$$

なるバイアス修正を施すと，つまり，$\hat{\theta}^*_{T,2} = \hat{\theta}_{T,2} - \beta(\hat{\theta}_{T,2})/\sqrt{T}$ と修正すると，2 次有効になる．漸近展開を正当化するための，正則条件などについては，Kutoyants(1998), Sakamoto-Yoshida(2000) などを参照．また，初期推定量として，一致推定量が必要となるが，適当な identifiability 条件の下で，その存在が保証される．たとえば，強度 f が周期 1 の周期関数 p で与えれらるとき，つまり，$f(t,\theta)=p(t-\theta)$, $\Theta=(0,1)$, のとき，母数 $\vartheta := C(\theta) = e^{2\pi i\theta}$ に対する一致推定量として，次のものが考えられる:

$$\hat{\vartheta}_T = \Pi_{S^1}\left(\frac{1}{cT}\int_0^T e^{2\pi i t} dX_t\right).$$

ここで，Π_{S^1} と c は

$$\Pi_{S^1} z = \begin{cases} z/|z| & \text{if } z \ne 0 \\ 1 & \text{if } z = 0 \end{cases}$$

$$c = \int_0^1 e^{2\pi is} p(s)ds,$$

により定義される．これから，θ に対する一致推定量が構成される．

Example 2 (diffusion process) 確率微分方程式

$$dX_t = V_0(\theta, X_t)dt + V(X_t)dw_t,$$

を満たす拡散過程 $X_T = (X_t)_{t\in[0,T]}$ について考える．ここで，X_T は定常で幾何的混合性を持つものと仮定する．また，$\theta \in (\alpha, \beta)$ で，w_t は標準ウィナー過程とする．このとき，条件付き対数尤度関数

$$\psi_T(\theta) = \int_0^T \frac{\delta V_0(\theta, X_t)}{V(X_t)^2} dX_t - \int_0^T \frac{V_0(\theta, X_t)\delta V_0(\theta, X_t)}{V(X_t)^2} dt.$$

によって与えられる推定関数 ψ を用いると，$\hat{\theta}_{T,2}$ は，バイアス修正

$$\beta(\theta) = -\frac{1}{6}\int_{-\infty}^{\infty} [\delta b(\theta, \cdot))^2](x)\ \delta b(\theta, x)\nu_\theta(x)dx \Bigg/ \left(\int_{-\infty}^{\infty} (\delta b(\theta, x))^2 \nu_\theta(x)dx,\right)^2$$

により，2 次有効になる．つまり，$\hat{\theta}^*_{T,2} = \hat{\theta}_{T,2} - \beta(\hat{\theta}_{T,2})/\sqrt{T}$ は，2 次有効である．ただし，$b = V_0/V^2$，ν_θ は定常分布，$[\cdot]$ は，

$$[f](x) = \frac{1}{V(x)n_{\theta_0}(x)} \int_x^{\infty} 2(f(u) - \nu_{\theta_0}(f))n_{\theta_0}(u)du, \quad \nu_{\theta_0}(f) = \int_{-\infty}^{\infty} f(x)\nu_{\theta_0}(x)dx.$$

により定義される．

初期推定量としての一致推定量の構成方法に関しては，たとえば，つぎのようなモーメント推定量が考えられる．ある関数 F に対して，

$$\gamma = \rho(\theta) = \int_{-\infty}^{\infty} F(x)\nu_\theta(x)dx$$

で定義される母数の変換を考える．この母数 γ に対して，モーメント推定量

$$\hat{\gamma}_T = \frac{1}{T}\int_0^T F(X_t)dt$$

が一致性を持つことが示せる (Kutoyants-Yoshida(1999))．もし，上の母数の変換が一対一であるように $(0 < \inf_\theta |\rho'(\theta)| < \sup_\theta |\rho'(\theta)| < \infty)$ 構成できるなら，明らかに，$\hat{\theta}_{T,0} = \rho^{-1}(\hat{\gamma}_T)$ は，θ の一致推定量となる．