

CIRJE-J-127

経験ベイズ信頼区間の漸近補正と  
小地域推定への応用

東京大学大学院経済学研究科

笹瀬吉隆  
久保川達也

2005年3月

CIRJE ディスカッションペーパーの多くは  
以下のサイトから無料で入手可能です。

[http://www.e.u-tokyo.ac.jp/cirje/research/03research02dp\\_j.html](http://www.e.u-tokyo.ac.jp/cirje/research/03research02dp_j.html)

このディスカッション・ペーパーは、内部での討論に資するための未定稿の段階にある論文草稿である。著者の承諾なしに引用・複写することは差し控えられるたい。

**Asymptotic correction of empirical Bayes confidence intervals and its application  
to small area estimation**

Y. Sasase and T. Kubokawa

*Abstract*

This paper addresses the issue of constructing a confidence interval of a small area mean in a random effect or mixed effects linear model. A crude confidence interval based on the empirical Bayes method has the drawback that its coverage probability is much less than a nominal confidence coefficient. For improving on this confidence interval, the paper provides the procedure of adjusting the critical value, and the resulting confidence interval has a coverage probability which is identical to the nominal confidence coefficient in second order asymptotics when the number of small areas is large. The proposed confidence interval is numerically investigated based on simulation experiments and applied to posted land price data. These numerical studies illustrate the practical usefulness of the proposal.

# 経験ベイズ信頼区間の漸近補正と小地域推定への応用

笹瀬 吉隆\* 久保川 達也\* †

平成 17 年 3 月 9 日

## 概要

変量効果モデルや混合線形モデルを利用した小地域推定について、経験ベイズ法に基づいた信頼区間で、群の数を大きくしたときに被覆確率が信頼係数に 2 次漸近近似するものを構成する。これによって、従来得られてきた信頼区間は、より現実的で一般的なモデルへ拡張されることになり、シミュレーションによる数値的性質を調べるとともに、地価公示価格の小地域推定と信頼区間の作成への応用が示される。

キーワード: 経験ベイズ法, 信頼区間, 2 次漸近補正, 変量効果モデル, 混合線形モデル, 小地域推定

## 1 はじめに

本稿で扱う問題は、標本調査における小地域の特性値の推定問題に関係している。通常は調査区全体の特性を調べるために標本調査が行われるが、そのデータから地域ごとの特性値を推定したい状況がしばしば生ずる。例えば、得られたデータから各地域への予算の配分の仕方を決めたり、政策を決定したりする場合がある。そのとき、狭い地域や人口が粗な地域に対しては十分なデータがとられていないため、その地域のデータだけでは特性値の十分な推測ができない。このような状況での推定問題を、特に小地域推定という。

小地域の平均値の推定において推定精度を高めるために用いられる方法が経験ベイズ法であり、Efron-Morris (1975), Fay-Herriot (1979), Prasad-Rao (1990) などによって、その有用性が示されてきた。通常考えられる推定量は調査区全体についての総平均と小地域の標本平均であり、前者は安定しているが地域の差異をみることはできない。一方、後者は地域の差異をみることはできても、小地域のデータ数が少なければ信頼しうる推定値を与えることができない。これらの両者の欠点を補う方法が経験ベイズ法であり、小地域の標本平均を全体の総平均の方向へ縮小することによって、地域差を考慮しつつより安定した推定値を与えることができる。言い換えると、小地域の周辺のデータをプールすることによって推定精度を高めることを意図する方法である。ここで 1 つの問題は、推定精度を高めるために用いた経験ベイズ推定量が実際どの程度の推定誤差を伴っているのかについて見積もることである。そのための 1 つの方法は、経験ベイズ推定量の平均 2 乗誤差についてその推定量を求めることで、Prasad-Rao (1990), Datta-Lahiri (2000) などによって研究されてきた。もう 1 つの方向性は、経験ベイズに基づいた信頼区間を構成すること

---

\*113-0033 東京都文京区本郷 7-3-1 東京大学大学院経済学研究科

†E-Mail: tatsuya@e.u-tokyo.ac.jp

であり, Datta-Ghosh-Smith-Lahiri (2002), Basu-Ghosh-Mukerjee (2003) によって最近研究されてきたが, 彼らは誤差分散が既知という簡単な場合を扱っており, 実用的な結果とはいえない。

そこで, 本稿では, 誤差分散が未知の場合に小地域平均の信頼区間を経験ベイズ法に基づいて求めることを目的とする。第 2 節では, 一元配置の变量効果モデルと経験ベイズ推定法について基本的な考え方を説明する。第 3 節では, 経験ベイズ法に基づいた信頼区間の構成を行う。ベイズ信頼区間に超母数の推定量を代入して得られる素朴な経験ベイズ信頼区間の問題点は, 被覆確率が名目信頼係数よりも小さくなってしまふ点である。第 5 節で与えられるシミュレーション実験の結果からは, 小地域の個数  $a$  が小さいときには, 名目信頼係数よりもかなり小さくなってしまい, 実用的ではないことがわかる。そこで, 信頼区間の両端を適当な関数で調整することによって, 被覆確率が名目信頼係数に近づけることを考える。第 3 節では, このような補正を行って, その被覆確率が  $a$  についての 2 次漸近理論において名目信頼係数に一致する信頼区間を求める。

第 4 節では, 応用上有用な混合線形モデルへの拡張を行う。最後に第 5 節では, シミュレーション実験を通して, 提案された補正済み経験ベイズ信頼区間の性質を調べるとともに, 神奈川県私鉄沿線沿いの地価公示価格データへ適用する。これらの数値結果から, 補正済み経験ベイズ信頼区間が標本平均と  $t$ -分布に基づいた素朴な信頼区間に比べてより有用であることがわかる。

## 2 变量効果モデルと経験ベイズ法

この節では次のようなモデルを扱う。全体で  $a$  個の小地域があり,  $i$  番目の小地域から  $n_i$  個の標本  $y_{i1}, \dots, y_{in_i}$  がとられており,  $i = 1, \dots, a; j = 1, \dots, n_i$  に対して

$$\begin{aligned} y_{ij} &= \theta_i + \varepsilon_{ij}, \\ \theta_i &= \mu + \alpha_i \end{aligned} \quad (2.1)$$

なる一元配置モデルを考える。ここで, 誤差項  $\varepsilon_{ij}$  と地域効果  $\alpha_i$  は互いに独立な確率変数で, それぞれ正規分布

$$\begin{aligned} \varepsilon_{ij} &\sim \mathcal{N}(0, \sigma_e^2), \\ \alpha_i &\sim \mathcal{N}(0, \sigma_a^2) \end{aligned}$$

に従っているとする。また  $\mu$  は全体平均,  $\sigma_e^2, \sigma_a^2$  はそれぞれ群内分散, 群間分散であり, 未知母数  $\mu, \sigma_e^2, \sigma_a^2$  の組を  $\omega = (\mu, \sigma_e^2, \sigma_a^2)$  で表す。

通常の標本調査では, 全体平均  $\mu$  の推定に興味があり, 総平均

$$\hat{\mu} = \sum_{i=1}^a \sum_{j=1}^{n_i} y_{ij} / N \equiv \bar{y}$$

によって推定される。ただし,

$$N = \sum_{i=1}^a n_i$$

である。しかし, 同じデータを用いて地域ごとの平均を推定したい状況があり, これを小地域推定といって近年その問題についての関心が高まり, 理論面と実用面からの研究が活発になされてき

た。例えば，地域の差異に基づいて予算配分や政策を決定する場面では各小地域の推定を行う必要がある。この場合， $i$  番目の小地域の平均  $\theta_i$  の推定に興味があり，通常はその小地域の標本平均

$$\bar{y}_i = \sum_{j=1}^{n_i} y_{ij}/n_i$$

で推定される。しかし，小地域によっては標本のサイズ  $n_i$  がきわめて小さいため， $\bar{y}_i$  の推定誤差が大きくなってしまう。そこで，(2.1) のような変量効果モデルを考えることによって，より精度の高い  $\theta_i$  の推定を行うことを考える。

一元配置の変量効果モデル (2.1) は，次のようなベイズ的モデルとして解釈することができる。 $\theta_i$  を与えたときの  $y_{ij}$  の条件付き分布が

$$y_{ij}|\theta_i \sim \mathcal{N}(\theta_i, \sigma_e^2)$$

に従い，母数  $\theta_i$  の事前分布  $\pi$  が

$$\pi : \theta_i \sim \mathcal{N}(\mu, \sigma_a^2)$$

に従っているときに，このベイズモデルは変量効果モデル (2.1) に一致する。確率変数の組  $\{y_{ij}|i = 1, \dots, a; j = 1, \dots, n_i\}$ ,  $\{\alpha_i|i = 1, \dots, a\}$  の同時確率密度において，その指数部分は次のように変形できる。

$$\begin{aligned} & \sum_{i=1}^a \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i)^2 / \sigma_e^2 + \sum_{i=1}^a \alpha_i^2 / \sigma_a^2 \\ &= \sum_{i=1}^a \left( \frac{n_i}{\sigma_e^2} + \frac{1}{\sigma_a^2} \right) \left\{ \theta_i - \hat{\theta}_i^\pi(\mu, \sigma_e^2, \sigma_a^2) \right\}^2 + \sum_{i=1}^a \frac{1}{\sigma_e^2/n_i + \sigma_a^2} (\bar{y}_i - \mu)^2 + S_1 / \sigma_e^2 \end{aligned}$$

ここで，

$$\begin{aligned} \hat{\theta}_i^\pi(\mu, \sigma_e^2, \sigma_a^2) &= \mu + \frac{\sigma_a^2}{\sigma_e^2/n_i + \sigma_a^2} (\bar{y}_i - \mu), \\ S_1 &= \sum_{i=1}^a \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2, \end{aligned}$$

である。 $\hat{\theta}_i^\pi(\mu, \sigma_e^2, \sigma_a^2)$  が  $\theta_i$  のベイズ推定量になっている。このことから， $\bar{y}_i$  を与えたときの  $\theta_i$  の事後分布は

$$\theta_i|\bar{y}_i \sim \mathcal{N}\left(\hat{\theta}_i^\pi(\mu, \sigma_e^2, \sigma_a^2), \frac{1}{n_i/\sigma_e^2 + 1/\sigma_a^2}\right), \quad (2.2)$$

で与えられ， $\bar{y}_1, \dots, \bar{y}_a, S_1$  の周辺分布は， $S_1$  の自由度が  $\sum_{i=1}^a (n_i - 1) = N - a$  となることから，

$$\begin{aligned} \bar{y}_i &\sim \mathcal{N}(\mu, \sigma_e^2/n_i + \sigma_a^2), \quad i = 1, \dots, a, \\ S_1 &\sim \sigma_e^2 \chi_{N-a}^2, \end{aligned} \quad (2.3)$$

となり，互いに独立に分布することがわかる。

いま,  $\mu, \sigma_e^2, \sigma_a^2$  は未知母数なので,  $\bar{y}_1, \dots, \bar{y}_a, S_1$  の周辺分布に基づいて推定する必要がある。 $\mu, \sigma_a^2$  の不偏推定量は,

$$\begin{aligned}\hat{\mu} = \bar{y} &= \sum_{i=1}^a \sum_{j=1}^{n_i} y_{ij} / N, \\ \hat{\sigma}_e^2 &= S_1 / (N - a),\end{aligned}\tag{2.4}$$

で与えられる。 $\sigma_a^2$  の不偏推定量を求めるために,  $E[\sum_{i=1}^a n_i (\bar{y}_i - \bar{y})^2]$  を計算すると,

$$E\left[\sum_{i=1}^a n_i (\bar{y}_i - \bar{y})^2\right] = (a - 1)\sigma_e^2 + \left(N - \sum_{i=1}^a n_i^2 / N\right) \sigma_a^2$$

となるので,

$$S_2 = \sum_{i=1}^a n_i (\bar{y}_i - \bar{y})^2\tag{2.5}$$

とおくと,  $\sigma_a^2$  の1つの不偏推定量として

$$\hat{\sigma}_a^2 = \{S_2 - (a - 1)\hat{\sigma}_e^2\} / \left(N - \sum_{i=1}^a n_i^2 / N\right)$$

が得られる。ここで  $\hat{\sigma}_a^2$  について注意すべき点が2つある。1つは, 正の確率で  $\hat{\sigma}_a^2$  が負の値をとってしまうことで, それを補うために零で打ち切る推定量

$$\hat{\sigma}_a^{2+} = \max\{\hat{\sigma}_a^2, 0\}$$

が使われる。もう1つは,  $n_i$  が異なるときには  $\hat{\sigma}_a^2$  は一様最小分散不偏推定量にはなっていない点である。これは,  $n_i$  が異なるときには統計量の完備性が失われるため一様最小分散不偏推定量を求めることができないことに由来する。しかし,  $n_1 = \dots = n_a = n$  のときには,  $\hat{\sigma}_a^2$  が一様最小分散不偏推定量になる。

小地域平均  $\theta_i$  の経験ベイズ推定量は, ベイズ推定量  $\hat{\theta}_i^\pi(\mu, \sigma_e^2, \sigma_a^2)$  の未知母数  $\mu, \sigma_e^2, \sigma_a^2$  のところへそれらの推定量を代入することによって得られる。すなわち, 経験ベイズ推定量は

$$\begin{aligned}\hat{\theta}_i^{EB} &= \bar{y} + \frac{\hat{\sigma}_a^2}{\hat{\sigma}_e^2 / n_i + \hat{\sigma}_a^2} (\bar{y}_i - \bar{y}) \\ &= \bar{y}_i - \frac{\hat{\sigma}_e^2}{\hat{\sigma}_e^2 + n_i \hat{\sigma}_a^2} (\bar{y}_i - \bar{y})\end{aligned}$$

で与えられ, 推定精度を高めるために小地域の標本平均  $\bar{y}_i$  を総平均  $\bar{y}$  の方向へ縮小する操作を行っている。縮小関数  $\hat{\sigma}_e^2 / (\hat{\sigma}_e^2 + n_i \hat{\sigma}_a^2)$  は小地域の標本サイズ  $n_i$  に依存しており,  $n_i$  が小さければ大きく縮小していることがわかる。

推定精度を高めるために経験ベイズ推定量  $\hat{\theta}_i^{EB}$  が用いられることになるが, 実際の推定誤差がどの程度に押さえられるのかを見積もることは, データ解析の現場においては重要なことである。そのための1つの方法は, 経験ベイズ推定量の平均2乗誤差についてその推定量を求めることで, Prasad-Rao (1990), Datta-Lahiri (2000) などによって研究されてきた。もう1つの方向性は, 経験ベイズ法に基づいた信頼区間を構成することであり, 本稿では次の節でこの問題を扱うことにする。

### 3 信頼区間の構成

#### 3.1 経験ベイズ信頼区間の2次漸近補正

さて、2節の設定のもとで  $i$  番目の小地域の平均  $\theta_i$  の信頼区間を構成する問題を考えよう。ここでは、一般性を失うことなく、 $\theta_1$  の信頼区間を求めることにする。

まず、 $B, V$  を次のように定義する。

$$B = \frac{\sigma_e^2}{\sigma_e^2 + n_1\sigma_a^2}, \quad V = \sigma_a^2 B = \frac{\sigma_e^2\sigma_a^2}{\sigma_e^2 + n_1\sigma_a^2} \quad (3.1)$$

(2.2) から、 $\bar{y}_1$  を与えたときの  $\theta_1$  の事後分布は、

$$\theta_1|\bar{y}_1 \sim \mathcal{N}((1-B)\bar{y}_1 + B\mu, V)$$

となるので、分散成分  $\sigma_e^2, \sigma_a^2$  及び全体平均  $\mu$  が既知のときにはモデル (2.1) のもとでの信頼係数  $1-\alpha$  の信頼区間は、

$$I_B(\mu, \sigma_e^2, \sigma_a^2) = \left[ (1-B)\bar{y}_1 + B\mu - z_{\alpha/2}V^{1/2}, (1-B)\bar{y}_1 + B\mu + z_{\alpha/2}V^{1/2} \right]$$

で与えられる。ただし、 $z_{\alpha/2}$  は正規分布の上側  $\alpha/2$  点を表している。これを、以降

$$I_B(\mu, \sigma_e^2, \sigma_a^2) : (1-B)\bar{y}_1 + B\mu \pm z_{\alpha/2}V^{1/2}$$

と表記することにする。

いま、 $\sigma_e^2, \sigma_a^2, \mu$  は未知母数であるから、それらの推定量  $\hat{\sigma}_e^2, \hat{\sigma}_a^2, \bar{y}$  を  $I_B(\mu, \sigma_e^2, \sigma_a^2)$  に代入した信頼区間

$$I_{EB} = I_B(\bar{y}, \hat{\sigma}_e^2, \hat{\sigma}_a^2) : (1-\hat{B})\bar{y}_1 + \hat{B}\bar{y} \pm z_{\alpha/2}\hat{V}^{1/2} \quad (3.2)$$

が考えられる。ここで、 $\hat{B}, \hat{V}$  はそれぞれ  $B, V$  の推定量で

$$\hat{B} = \frac{\hat{\sigma}_e^2}{\hat{\sigma}_e^2 + n_1\hat{\sigma}_a^2}, \quad \hat{V} = \hat{\sigma}_a^2\hat{B} = \frac{\hat{\sigma}_e^2\hat{\sigma}_a^2}{\hat{\sigma}_e^2 + n_1\hat{\sigma}_a^2} \quad (3.3)$$

で与えられる。しかし、これは  $\mu, B, V$  をそれらの推定量で置き換えているので、被覆確率  $P[\theta_1 \in I_{EB}]$  は  $1-\alpha$  にはならない。Datta *et al.* (2002), Basu *et al.* (2003) では、 $n_1 = \dots = n_a$  で  $\sigma_e^2$  が既知のときに、 $z_{\alpha/2}$  を  $t$  で置き換えて、被覆確率が  $a$  について漸近的に  $1-\alpha$  になるように  $t$  を調節することを考えた。ここでは、 $n_1, \dots, n_a$  が等しくはなく、 $\sigma_e^2$  が未知というより一般的なモデル (2.1) において、Basu *et al.* (2003) の論法を用いて (3.2) を修正することによって、 $a \rightarrow \infty$  において被覆確率が  $1-\alpha + O(a^{-3/2})$  となるような信頼区間を構成しよう。

信頼区間 (3.2) において、 $z_{\alpha/2}$  の代わりに

$$z_{\alpha/2} \left( 1 + \frac{1}{2a}\hat{h} \right)$$

を用いることを考える。ここで、 $\hat{h}$  は、 $\hat{h} = \hat{h}(\bar{y}_1, \dots, \bar{y}_a, S_1) = O_p(1)$  であるような  $\bar{y}_1, \dots, \bar{y}_a, S_1$  の関数であり、被覆確率が  $1 - \alpha + O(a^{-3/2})$  となるように、この関数を特定することが目標となる。 $z_{\alpha/2}$  の代わりに  $z_{\alpha/2}\{1 + (2a)^{-1}\hat{h}\}$  を代入した信頼区間を

$$I_{AEB}: (1 - \hat{B})\bar{y}_1 + \hat{B}\bar{y} \pm z_{\alpha/2}\{1 + (2a)^{-1}\hat{h}\}\hat{V}^{1/2} \quad (3.4)$$

とおくことにする。以降、簡単のために  $z = z_{\alpha/2}$  とおいて議論を進める。信頼区間  $I_{AEB}$  の被覆確率は、

$$\begin{aligned} P[\theta_1 \in I_{AEB}] &= P\left[-z + G(-z) \leq V^{-1/2}\{\theta_1 - (1 - B)\bar{y}_1 - B\bar{y}\} \leq z + G(z)\right] \\ &= E[\Phi(z + G(z)) - \Phi(-z + G(-z))] \end{aligned} \quad (3.5)$$

と表される。ただし  $G(z)$  は  $\bar{y}_1, \dots, \bar{y}_a, S_1, z$  の関数で、

$$\begin{aligned} G(z) &= V^{-1/2} \left\{ (1 - \hat{B})\bar{y}_1 + \hat{B}\bar{y} - (1 - B)\bar{y}_1 - B\mu + z\{1 + (2a)^{-1}\hat{h}\}\hat{V}^{1/2} \right\} - z \\ &= V^{-1/2} \left\{ (B - \hat{B})(\bar{y}_1 - \mu) + \hat{B}(\bar{y} - \mu) \right\} \\ &\quad + z \left[ \{1 + (2a)^{-1}\hat{h}\}\hat{V}^{1/2}/V^{1/2} - 1 \right] \end{aligned} \quad (3.6)$$

で定義される。

(3.6) の右辺の第 1 項を評価するために、正の整数  $i, j$  に対して、 $(\hat{\sigma}_e^2 - \sigma_e^2)^i (\hat{\sigma}_a^2 - \sigma_a^2)^j = O_p(a^{(i+j)/2})$  であり、また  $\bar{y} - \mu = O_p(a^{-1/2})$  であることに注意する。これは、 $\hat{\sigma}_e^2 - \sigma_e^2, \hat{\sigma}_a^2 - \sigma_a^2, \bar{y} - \mu$  を適当に変形すると、いずれも Lindeberg の条件がみたされて中心極限定理が成り立つことからわかる。特に、 $\hat{\sigma}_a^2$  については、 $S_2$  が  $S_2 = \sum_{j=1}^{a-1} (\sigma_e^2 + \gamma_j \sigma_a^2) \chi_{1,j}^2$  と表されることを用いると Lindeberg の条件を確かめることができる。ここで  $\chi_{1,1}^2, \dots, \chi_{1,a-1}^2$  は、互いに独立な確率変数でそれぞれ自由度 1 のカイ 2 乗分布に従っている。また  $\gamma_j$  は正の実数であり、すべての  $i$  に対して  $\gamma_i \leq \max_j \gamma_j$  なる不等式が成り立っているので、 $\gamma_i$  は  $a$  に関して有界であることが注意される。これらの点については、補遺において若干の説明が与えられているので参照されたい。 $\hat{B}$  を  $B$  の周りでテーラー展開すると、

$$\tau = \frac{1}{\sigma_e^2 + n_1 \sigma_a^2}$$

に対して

$$\begin{aligned} \hat{B} - B &= n_1 \tau^2 \left\{ \sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2) \right\} \\ &\quad + n_1 \tau^3 \left\{ n_1 \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2)^2 - \sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2)^2 \right. \\ &\quad \left. + (\sigma_e^2 - n_1 \sigma_a^2) (\hat{\sigma}_a^2 - \sigma_a^2) (\hat{\sigma}_e^2 - \sigma_e^2) \right\} + O_p(a^{-3/2}) \end{aligned} \quad (3.7)$$

となるので、第 1 項は、

$$V^{-1/2} \left\{ (B - \hat{B})(\bar{y}_1 - \mu) + \hat{B}(\bar{y} - \mu) \right\} = \frac{1}{\sqrt{a}} \hat{u}_1 + \frac{1}{a} \hat{u}_2 + O_p(a^{-3/2}) \quad (3.8)$$

なる形で評価される。ただし、 $\hat{u}_1, \hat{u}_2$  は  $O_p(1)$  の関数で、

$$\begin{aligned} \hat{u}_1 &= \sqrt{a} V^{-1/2} \left[ B(\bar{y} - \mu) - n_1 \tau^2 \{ \sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2) \} (\bar{y}_1 - \mu) \right], \\ \hat{u}_2 &= a n_1 \tau^2 V^{-1/2} \left[ \{ \sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2) \} (\bar{y} - \mu) \right. \\ &\quad \left. - \tau \{ n_1 \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2)^2 - \sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2)^2 + (\sigma_e^2 - n_1 \sigma_a^2) (\hat{\sigma}_a^2 - \sigma_a^2) (\hat{\sigma}_e^2 - \sigma_e^2) \} (\bar{y}_1 - \mu) \right] \end{aligned} \quad (3.9)$$

で与えられる。また，(3.6) の右辺の第 2 項を評価するために， $\hat{V}$  を  $V$  の周りで展開すると，

$$\begin{aligned} (\hat{V}/V)^{1/2} - 1 &= \frac{1}{2V}(\hat{V} - V) - \frac{1}{8V^2}(\hat{V} - V)^2 + \frac{1}{16V^3}(\hat{V} - V)^3 \\ &\quad - \frac{5}{128V^4}\left(\frac{\epsilon}{V}\right)^{-7/2}(\hat{V} - V)^4 \end{aligned}$$

となり，さらに  $\hat{V}$  を  $(\sigma_e^2, \sigma_a^2)$  のまわりで展開すると，

$$\begin{aligned} \hat{V} - V &= \tau^2 \{ \sigma_e^4(\hat{\sigma}_e^2 - \sigma_e^2) + \sigma_a^4(\hat{\sigma}_e^2 - \sigma_e^2) \} \\ &\quad - n_1 \tau^3 \{ \sigma_a^4(\hat{\sigma}_e^2 - \sigma_e^2)^2 + \sigma_e^4(\hat{\sigma}_a^2 - \sigma_a^2)^2 - 2\sigma_a^2\sigma_e^2(\hat{\sigma}_e^2 - \sigma_e^2)(\hat{\sigma}_a^2 - \sigma_a^2) \} \\ &\quad + O_p(a^{-3/2}) \end{aligned}$$

とかける。従って，

$$\hat{V}^{1/2}/V^{1/2} = 1 + \frac{1}{\sqrt{a}}\hat{v}_1 + \frac{1}{a}\hat{v}_2 + O_p(a^{-3/2}) \quad (3.10)$$

なる確率展開が得られる。ただし， $\hat{v}_1, \hat{v}_2$  は  $O_p(1)$  の関数で，

$$\hat{v}_1 = \sqrt{a} \frac{\tau^2}{2V} \{ \sigma_a^4(\hat{\sigma}_e^2 - \sigma_e^2) + \sigma_e^4(\hat{\sigma}_a^2 - \sigma_a^2) \}, \quad (3.11)$$

$$\begin{aligned} \hat{v}_2 &= -an_1 \frac{\tau^3}{2V} \{ \sigma_a^4(\hat{\sigma}_e^2 - \sigma_e^2)^2 + \sigma_e^4(\hat{\sigma}_a^2 - \sigma_a^2)^2 - 2\sigma_a^2\sigma_e^2(\hat{\sigma}_e^2 - \sigma_e^2)(\hat{\sigma}_a^2 - \sigma_a^2) \} \\ &\quad - a \frac{\tau^4}{8V^2} \{ \sigma_a^4(\hat{\sigma}_e^2 - \sigma_e^2) + \sigma_e^4(\hat{\sigma}_a^2 - \sigma_a^2) \}^2 \end{aligned} \quad (3.12)$$

で与えられる。(3.8), (3.10) を (3.6) に代入すると，

$$\begin{aligned} G(z) &= \frac{1}{\sqrt{a}}\hat{u}_1 + \frac{1}{a}\hat{u}_2 + z \left[ \left\{ 1 + \frac{1}{2a}\hat{h} \right\} \left\{ 1 + \frac{1}{\sqrt{a}}\hat{v}_1 + \frac{1}{a}\hat{v}_2 \right\} - 1 \right] + O_p(a^{-3/2}) \\ &= \frac{1}{\sqrt{a}}\hat{u}_1 + \frac{1}{a}\hat{u}_2 + z \left[ \frac{1}{\sqrt{a}}\hat{v}_1 + \frac{1}{a} \left\{ \hat{v}_2 + \frac{1}{2}\hat{h} \right\} \right] + O_p(a^{-3/2}) \end{aligned} \quad (3.13)$$

と表現されることになる。

まず， $\Phi(z + G(z))$  をテーラー展開すると，

$$\begin{aligned} \Phi(z + G(z)) &= \Phi(z) + G(z)\phi(z) + \frac{1}{2}G^2(z)\phi'(z) + \frac{1}{2} \int_z^{z+G(z)} (z + G(z) - x)^2 \phi''(x) dx \end{aligned} \quad (3.14)$$

となる。ここで， $\phi(z)$  は標準正規分布の確率密度関数で， $\phi'(z) = -z\phi(z)$ ， $\phi''(z) = (z^2 - 1)\phi(z)$  となることに注意する。これより，右辺の第 4 項については，

$$\begin{aligned} &\frac{1}{2} \int_z^{z+G(z)} (z + G(z) - x)^2 \phi''(x) dx \\ &\leq \phi(3^{1/2}) \int_z^{z+G(z)} (z + G(z) - x)^2 dx \\ &= 3^{-1} \phi(3^{1/2}) G^3(z) \end{aligned}$$

と評価できるが, (3.13) より  $G^3(z) = O_p(a^{-3/2})$  となる。従って,

$$\Phi(z + G(z)) = \Phi(z) + \left\{ G(z) - \frac{z}{2} G^2(z) \right\} \phi(z) + O_p(a^{-3/2})$$

とかけるので, これを (3.5) に代入すると,

$$\begin{aligned} P[\theta_1 \in I_{AEB}] &= E \left\{ \Phi(z) + \left[ G(z) - \frac{z}{2} G^2(z) \right] \phi(z) \right\} \\ &\quad - \left\{ \Phi(-z) + \left[ G(-z) + \frac{z}{2} G^2(-z) \right] \phi(-z) \right\} + O(a^{-3/2}) \\ &= \Phi(z) - \Phi(-z) \\ &\quad + \phi(z) E \left[ G(z) - G(-z) - \frac{z}{2} \{ G^2(z) + G^2(-z) \} \right] + O(a^{-3/2}) \end{aligned}$$

とかけることがわかる。ここで,  $z = z_{\alpha/2}$  であるから,  $\Phi(z) - \Phi(-z) = 1 - \alpha$  であり, また (3.13) より,

$$\begin{aligned} G(z) - G(-z) - \frac{z}{2} \{ G^2(z) + G^2(-z) \} \\ &= 2z \left\{ \frac{1}{\sqrt{a}} \hat{v}_1 + \frac{1}{a} (\hat{v}_2 + 2^{-1} \hat{h}) \right\} - \frac{z}{a} (\hat{u}_1^2 + z^2 \hat{v}_1^2) + O_p(a^{-3/2}) \\ &= \frac{2z}{\sqrt{a}} \hat{v}_1 + \frac{z}{a} [\hat{h} + 2\hat{v}_2 - \hat{u}_1^2 - z^2 \hat{v}_1^2] + O_p(a^{-3/2}) \end{aligned}$$

と評価できる。 $\hat{v}_1$  は (3.11) の形から  $E[\hat{v}_1] = 0$  となるので,

$$h(\sigma_e^2, \sigma_a^2) = \lim_{a \rightarrow \infty} \{ E[-2\hat{v}_2 + \hat{u}_1^2 + z^2 \hat{v}_1^2] \}$$

とおくと,

$$P[\theta_1 \in I_{AEB}] = 1 - \alpha + \frac{z}{a} \phi(z) \left\{ E[\hat{h}] - h(\sigma_e^2, \sigma_a^2) \right\} + O(a^{-3/2}) \quad (3.15)$$

となることがわかる。従って,  $h(\sigma_e^2, \sigma_a^2)$  に推定量  $\hat{\sigma}_e^2, \hat{\sigma}_a^2$  を代入して

$$\hat{h} = h(\hat{\sigma}_e^2, \hat{\sigma}_a^2)$$

とおくと,  $P[\theta_1 \in I_{AEB}] = 1 - \alpha + O(a^{-3/2})$  となり, 要求した条件をみたま信頼区間が得られることになる。

信頼区間  $I_{AEB}$  の具体的な形を与えるためには,  $h(\sigma_e^2, \sigma_a^2)$  を計算する必要がある。これは次の命題で与えられるが, その導出には多少計算が必要になるので次の節で与えることにする。

**命題 3.1**  $k = 1, 2$  に対して  $N_k$  を

$$N_k = \lim_{a \rightarrow \infty} \sum_{i=1}^a n_i^k / a$$

で定義し、これらの値が存在して零でないとは定する。このとき、関数  $h(\sigma_e^2, \sigma_a^2)$  は、

$$\begin{aligned} h(\sigma_e^2, \sigma_a^2) &= \lim_{a \rightarrow \infty} \left\{ E \left[ -2\hat{v}_2 + \hat{u}_1^2 + z^2 \hat{v}_1^2 \right] \right\} \\ &= \frac{4n_1}{\sigma_e^2 \sigma_a^2 (\sigma_e^2 + n_1 \sigma_a^2)^2} \left\{ \left( \frac{1}{N_1 - 1} + 2 \frac{N_2}{N_1^2} \right) \sigma_e^4 \sigma_a^4 + \frac{1}{N_1(N_1 - 1)} (\sigma_e^2 + N_1 \sigma_a^2) \sigma_e^6 \right\} \\ &\quad + \frac{1 + z^2}{2\sigma_e^4 \sigma_a^4 (\sigma_e^2 + n_1 \sigma_a^2)^2} \left\{ \frac{1}{N_1} \left( \frac{N_2}{N_1} - \frac{2}{N_1 - 1} \right) \sigma_e^8 \sigma_a^4 \right. \\ &\quad \quad \left. + \left( \frac{\sigma_e^8}{N_1 - 1} + \frac{\sigma_e^8}{N_1(N_1 - 1)} + \frac{2}{N_1} \sigma_e^6 \sigma_a^2 \right) \right\} \\ &\quad + \frac{1}{\sigma_e^2 \sigma_a^2 (\sigma_e^2 + n_1 \sigma_a^2)} \left( \frac{\sigma_e^2}{N_1} + \frac{N_2}{N_1^2} \sigma_a^2 \right) \sigma_e^4 \end{aligned}$$

で与えられる。

分散比を  $\rho = \sigma_e^2 / \sigma_a^2$  とおくと、 $h(\sigma_e^2, \sigma_a^2)$  は実は  $\rho$  の関数になっていて、

$$\begin{aligned} h^*(\rho) = h(\sigma_e^2, \sigma_a^2) &= \frac{4n_1}{\rho(1 + n_1\rho)^2} \left\{ \left( \frac{1}{N_1 - 1} + 2 \frac{N_2}{N_1^2} \right) \rho^2 + \frac{1}{N_1(N_1 - 1)} (1 + N_1\rho) \right\} \\ &\quad + \frac{1 + z^2}{2\rho^2(1 + n_1\rho)^2} \left\{ \frac{1}{N_1} \left( \frac{N_2}{N_1} - \frac{2}{N_1 - 1} \right) \rho^2 \right. \\ &\quad \quad \left. + \left( \frac{\rho^4}{N_1 - 1} + \frac{1}{N_1(N_1 - 1)} + \frac{2}{N_1} \rho \right) \right\} \\ &\quad + \frac{1}{\rho(1 + n_1\rho)} \left( \frac{1}{N_1} + \frac{N_2}{N_1^2} \rho \right) \end{aligned} \quad (3.16)$$

というより簡単な形で表されることがわかる。分散成分の不偏推定量  $\hat{\sigma}_e^2, \hat{\sigma}_a^2$  を用いた  $\rho$  の推定量は  $\hat{\rho} = \hat{\sigma}_a^2 / \hat{\sigma}_e^2$  であり、これを  $h^*(\rho)$  に代入すれば信頼区間  $I_{AEB}$  が構成できたことになる。 $I_{AEB}$  中の  $\hat{B}$  は  $\hat{B}(\hat{\rho}) = (1 + n_1\hat{\rho})^{-1}$  であることを思い出すと、以上の結果は次の命題としてまとめられる。

命題 3.2 (3.4) で与えられる信頼区間  $I_{AEB}$  は、

$$I_{AEB} : \frac{1}{1 + n_1\hat{\rho}} \{n_1\hat{\rho}\bar{y}_1 + \bar{y}\} \pm z_{\alpha/2} \sqrt{\frac{\hat{\sigma}_e^2 \hat{\rho}}{1 + n_1\hat{\rho}}} \left( 1 + \frac{1}{2a} h^*(\hat{\rho}) \right) \quad (3.17)$$

と表されて、 $a \rightarrow \infty$  のときに  $P[\theta_1 \in I_{AEB}] = 1 - \alpha + O(a^{-3/2})$  となる。

さて、分散成分  $\sigma_a^2$  の不偏推定量  $\hat{\sigma}_a^2$  の欠点は正の確率で負の値をとってしまうことである。そこで、 $\hat{\sigma}_a^2$  を 0 で打ち切った推定量  $\hat{\sigma}_a^{2+} = \max\{\hat{\sigma}_a^2, 0\}$  が考えられる。しかし、命題 3.1 の中で与えられた関数  $h(\sigma_e^2, \sigma_a^2)$  の形をみると、 $\sigma_a^2$  が分母にきてしまう項があり、 $\hat{\sigma}_a^{2+} = 0$  となった場合のことを考慮すると、 $\sigma_a^2$  を  $\hat{\sigma}_a^{2+}$  で推定したのでは不十分である。そこで、Kubokawa (2000)、Kubokawa-Saleh-Konno (2000) で提案された  $\sigma_e^2, \sigma_a^2$  の打ち切り推定量  $\hat{\sigma}_e^{2TR}, \hat{\sigma}_a^{2TR}$  を用いるこ

とにする。これらの推定量は

$$\hat{\sigma}_e^{2TR} = \min \left\{ \hat{\sigma}_e^2, \frac{(N-a)\hat{\sigma}_e^2 + (a-1)(a+1)^{-1}S_2}{N-1} \right\}, \quad (3.18)$$

$$\begin{aligned} \hat{\sigma}_a^{2TR} &= \frac{(a-1)N}{N^2 - \sum_{i=1}^a n_i^2} \max \left\{ \frac{S_2}{(a-1)\hat{\sigma}_e^2} - 1, \frac{2}{a-1} \right\} \\ &\quad \times \min \left\{ \hat{\sigma}_e^2, \frac{(N-a)\hat{\sigma}_e^2 + (a-1)(a+1)^{-1}S_2}{N-1} \right\} \end{aligned} \quad (3.19)$$

与えられるが、Kullback-Leibler 損失関数のもとで  $(\sigma_e^2, \sigma_a^2)$  を同時に推定する問題においては、打ち切り推定量  $(\hat{\sigma}_e^{2TR}, \hat{\sigma}_a^{2TR})$  が不偏推定量  $(\hat{\sigma}_e^2, \hat{\sigma}_a^2)$  を優越することが Kubokawa (2000) によって示されている。 $\hat{\sigma}_a^{2TR}$  はほとんど至るところで正の推定値を与えることができる。分散成分の比  $\rho$  は、

$$\hat{\rho}^{TR} = \hat{\sigma}_a^{2TR} / \hat{\sigma}_e^{2TR} = \frac{(a-1)N}{N^2 - \sum_{i=1}^a n_i^2} \min \left\{ \frac{S_2}{(a-1)\hat{\sigma}_e^2} - 1, \frac{2}{a-1} \right\}$$

で推定される。上述の不偏推定量は若干複雑な形をしているようにみえるが、実は場合分けは 1 通りしかなく、 $S_2/\hat{\sigma}_e^2 > a+1$  のときには、 $(\hat{\sigma}_e^{2TR}, \hat{\sigma}_a^{2TR})$  は不偏推定量  $(\hat{\sigma}_e^2, \hat{\sigma}_a^2)$  に一致し、 $S_2/\hat{\sigma}_e^2 \leq a+1$  のときには、 $\hat{\sigma}_e^{TR} = \{(N-a)\hat{\sigma}_e^2 + (a-1)(a+1)^{-1}S_2\}/(N-1)$ 、 $\hat{\sigma}_a^{2TR} = \hat{\sigma}_e^{TR}\hat{\rho}^{TR}$ 、 $\hat{\rho}^{TR} = 2N/(N^2 - \sum_{i=1}^a n_i^2)$  となっている。信頼区間  $I_{AEB}$  において、 $\hat{\sigma}_e^2, \hat{\rho}$  の代わりに  $\hat{\sigma}_e^{2TR}, \hat{\rho}^{TR}$  を用いたものを考えると、漸近的には  $I_{AEB}$  と同等になることが次の命題で示される。

**命題 3.3**  $\sigma_e^2, \sigma_a^2, \rho$  の打ち切り推定量を用いた信頼区間  $I_{TR}$  を、

$$I_{TR}: \frac{1}{1 + n_1\hat{\rho}^{TR}} \{n_1\hat{\rho}^{TR}\bar{y}_1 + \bar{y}\} \pm z_{\alpha/2} \sqrt{\frac{\hat{\sigma}_e^{2TR}\hat{\rho}^{TR}}{1 + n_1\hat{\rho}^{TR}}} \left(1 + \frac{1}{2a}h^*(\hat{\rho}^{TR})\right) \quad (3.20)$$

とおくと、 $a \rightarrow \infty$  のときに  $P[\theta_1 \in I_{TR}] = 1 - \alpha + O(a^{-3/2})$  となる。

この命題は、次のようにして確かめることができる。まず、 $S_2$  は、 $S_2 = \sum_{j=1}^{a-1} (\sigma_e^2 + \gamma_j \sigma_a^2) \chi_{1,j}^2$  と表されることに注意する。ここで  $\chi_{1,1}^2, \dots, \chi_{1,a-1}^2$  は、互いに独立な確率変数でそれぞれ自由度 1 のカイ 2 乗分布に従っている。また  $\gamma_j$  は正の実数である。 $\gamma_0 = \max_i \gamma_i$  とおくと、 $\chi_{a-1}^2 = \sum_{j=1}^{a-1} \chi_{1,j}^2$  に対して  $S_2 \leq (\sigma_e^2 + \gamma_0 \sigma_a^2) \chi_{a-1}^2$  となる。この不等式を用いると、

$$\begin{aligned} &P[\theta_1 \in I_{AEB}] - P[\theta_1 \in I_{TR}] \\ &= E[I(\theta_1 \in I_{AEB}) - I(\theta_1 \in I_{TR})] \\ &= E[\{I(\theta_1 \in I_{AEB}) - I(\theta_1 \in I_{TR})\} I(S_2/\hat{\sigma}_e^2 < a+1)] \\ &\leq 2P[S_2/\hat{\sigma}_e^2 < a+1] \\ &\leq 2P[(1 + \gamma_0 \rho) \chi_{a-1}^2 < (a+1) \chi_{N-a}^2 / (N-a)] \end{aligned}$$

と評価できる。ただし、集合  $A$  に対して  $I(A)$  は定義関数を表しており、 $A$  が成り立つときには  $I(A) = 1$ 、そうでないときには  $I(A) = 0$  である。不等式  $(1 + \gamma_0 \rho) \chi_{a-1}^2 < (a+1) \chi_{N-a}^2 / (N-a)$  を

$$\begin{aligned} T &\equiv \frac{a+1}{N-a} \{\chi_{N-a}^2 - (N-a)\} - (1 + \gamma_0 \rho) \{\chi_{a-1}^2 - (a-1)\} \\ &> (a-1) \{\gamma_0 \rho - 2/(a-1)\} \end{aligned}$$

と変形し，任意に固定した  $\rho$  に対して  $\gamma_0\rho - 2/(a-1) > 0$  となるような十分大きい  $a$  をとる。  
任意の正の実数  $k$  に対して，チェビシエフ型の不等式を用いると，

$$\begin{aligned} & P[(1 + \gamma_0\rho)\chi_{a-1}^2 < (a+1)\chi_{N-a}^2/(N-a)] \\ &= P[|T| > (a-1)\{\gamma_0\rho - 2/(a-1)\}] \\ &\leq E\left[\left\{\frac{|T|}{(a-1)\{\gamma_0\rho - 2/(a-1)\}}\right\}^k\right] \\ &= O(a^{-k/2}) \end{aligned}$$

と評価される。従って， $P[\theta_1 \in I_{TR}] = P[\theta_1 \in I_{AEB}] + O(a^{-k/2})$  となるので， $k \geq 3$  にとれば， $O(a^{-1})$  までの次数においては2つの信頼区間の被覆確率は同等になることがわかる。

### 3.2 命題 3.1 の証明

ここで，命題 3.1 の証明を与えよう。 $h(\sigma_e^2, \sigma_a^2) = \lim_{a \rightarrow \infty} E[-2\hat{v}_2 + \hat{u}_1^2 + z^2\hat{v}_1^2]$  を計算することになる。(3.9), (3.11), (3.12) から

$$\begin{aligned} & -2\hat{v}_2 + \hat{u}_1^2 + z^2\hat{v}_1^2 \\ &= \frac{an_1\tau^3}{V} \{ \sigma_a^4(\hat{\sigma}_e^2 - \sigma_e^2)^2 + \sigma_e^4(\hat{\sigma}_a^2 - \sigma_a^2)^2 - 2\sigma_e^2\sigma_a^2(\hat{\sigma}_e^2 - \sigma_e^2)(\hat{\sigma}_a^2 - \sigma_a^2) \} \\ & \quad + \frac{a\tau^4}{4V^2} \{ \sigma_a^4(\hat{\sigma}_e^2 - \sigma_e^2) + \sigma_e^4(\hat{\sigma}_a^2 - \sigma_a^2) \}^2 \\ & \quad + \frac{a}{V} \{ B(\bar{y} - \mu) - n_1\tau^2 [\sigma_a^2(\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2(\hat{\sigma}_a^2 - \sigma_a^2)] (\bar{y}_1 - \mu) \}^2 \\ & \quad + z^2 \frac{a\tau^4}{4V^2} \{ \sigma_a^4(\hat{\sigma}_e^2 - \sigma_e^2) + \sigma_e^4(\hat{\sigma}_a^2 - \sigma_a^2) \}^2 \end{aligned}$$

となるので， $h(\sigma_e^2, \sigma_a^2)$  は，

$$h(\sigma_e^2, \sigma_a^2) = \frac{n_1\tau^3}{V} I_1 + \frac{(1+z^2)\tau^4}{4V^2} I_2 + \frac{B^2}{V} I_3 + \frac{n_1^2\tau^4}{V} I_4 - 2\frac{n_1B\tau^2}{V} I_5 \quad (3.21)$$

と表される。ここで，

$$\begin{aligned} I_1 &= \lim_{a \rightarrow \infty} a \{ \sigma_a^4 \text{Var}(\hat{\sigma}_e^2) + \sigma_e^4 \text{Var}(\hat{\sigma}_a^2) - 2\sigma_e^2\sigma_a^2 \text{Cov}(\hat{\sigma}_e^2, \hat{\sigma}_a^2) \}, \\ I_2 &= \lim_{a \rightarrow \infty} a \{ \sigma_a^8 \text{Var}(\hat{\sigma}_e^2) + \sigma_e^8 \text{Var}(\hat{\sigma}_a^2) + 2\sigma_e^4\sigma_a^4 \text{Cov}(\hat{\sigma}_e^2, \hat{\sigma}_a^2) \}, \\ I_3 &= \lim_{a \rightarrow \infty} a E[(\bar{y} - \mu)^2], \\ I_4 &= \lim_{a \rightarrow \infty} a E[\{ \sigma_a^2(\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2(\hat{\sigma}_a^2 - \sigma_a^2) \}^2 (\bar{y}_1 - \mu)^2], \\ I_5 &= \lim_{a \rightarrow \infty} a E[\{ \sigma_a^2(\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2(\hat{\sigma}_a^2 - \sigma_a^2) \} (\bar{y} - \mu)(\bar{y}_1 - \mu)] \end{aligned} \quad (3.22)$$

で与えられる。以下では，これらの値を個々に評価していこう。

まず， $I_1, I_2$  を評価するために，推定量  $\hat{\sigma}_e^2, \hat{\sigma}_a^2$  の分散および共分散を計算する必要がある。

$S_1 \sim \sigma_e^2 \chi_{N-a}^2$  より,  $\hat{\sigma}_e^2$  の分散は

$$\begin{aligned} \text{Var}(\hat{\sigma}_e^2) &= \frac{\sigma_e^4}{(N-a)^2} E[\{S_1/\sigma_e^2 - (N-a)\}^2] \\ &= \frac{2}{N-a} \sigma_e^4 \end{aligned}$$

となる。また (2.3), (2.5) から  $S_1$  と  $S_2$  が独立となることがわかるので,  $\hat{\sigma}_a^2$  の分散は,  $C = N/(N^2 - \sum_{i=1}^a n_i^2)$  に対して,

$$\begin{aligned} \text{Var}(\hat{\sigma}_a^2) &= E\left[\{C(S_2 - (a-1)\hat{\sigma}_e^2) - \sigma_a^2\}^2\right] \\ &= E\left[\{C\{S_2 - (a-1)\sigma_e^2 - \sigma_a^2/C\} - (a-1)C(\hat{\sigma}_e^2 - \sigma_e^2)\}^2\right] \\ &= C^2 E\left[\{S_2 - (a-1)\sigma_e^2 - \sigma_a^2/C\}^2\right] + C^2(a-1)^2 \text{Var}(\hat{\sigma}_e^2) \end{aligned} \quad (3.23)$$

となる。ここで,

$$\begin{aligned} &E\left[\{S_2 - (a-1)\sigma_e^2 - \sigma_a^2/C\}^2\right] \\ &= E[S_2^2] - \{(a-1)\sigma_e^2 + \sigma_a^2/C\}^2 \\ &= 2 \sum_{i=1}^a (\sigma_e^2 + n_i \sigma_a^2)^2 + \frac{2}{N^2} \left\{ \sum_{i=1}^a n_i (\sigma_e^2 + n_i \sigma_a^2) \right\}^2 - \frac{4}{N} \sum_{i=1}^a n_i (\sigma_e^2 + n_i \sigma_a^2)^2 \end{aligned}$$

と書けるので,

$$\begin{aligned} \text{Var}(\hat{\sigma}_a^2) &= 2C^2 \left\{ \sum_{i=1}^a (\sigma_e^2 + n_i \sigma_a^2)^2 + \frac{1}{N^2} \left\{ \sum_{i=1}^a n_i (\sigma_e^2 + n_i \sigma_a^2) \right\}^2 \right. \\ &\quad \left. - \frac{2}{N} \sum_{i=1}^a n_i (\sigma_e^2 + n_i \sigma_a^2)^2 + \frac{(a-1)^2}{N-a} \sigma_e^4 \right\} \\ &= \frac{2N}{N^2 - \sum n_i^2} \left[ \frac{N(N-1)(a-1)}{(N-a)(N^2 - \sum n_i^2)} \sigma_e^4 + 2\sigma_e^2 \sigma_a^2 \right. \\ &\quad \left. + \frac{N^2 \sum n_i^2 + (\sum n_i^2)^2 - 2N \sum n_i^3}{N(N^2 - \sum n_i^2)} \sigma_a^4 \right] \end{aligned}$$

と表される。同様にして,  $\hat{\sigma}_e^2$  と  $\hat{\sigma}_a^2$  の共分散を計算すると,

$$\begin{aligned} \text{Cov}(\hat{\sigma}_e^2, \hat{\sigma}_a^2) &= -C(a-1) \text{Var}(\hat{\sigma}_e^2) \\ &= -\frac{2(a-1)N}{(N-a)(N^2 - \sum n_i^2)} \sigma_e^4 \end{aligned} \quad (3.24)$$

となる。  $N_k = \lim_{a \rightarrow \infty} \sum_{i=1}^a n_i^k / a$  とおくとき,  $\lim_{a \rightarrow \infty} aC = 1/N_1$  となることに注意すると,

$$\begin{aligned} \lim_{a \rightarrow \infty} a \text{Var}(\hat{\sigma}_e^2) &= \frac{2}{N_1 - 1} \sigma_e^4, \\ \lim_{a \rightarrow \infty} a \text{Var}(\hat{\sigma}_a^2) &= \frac{2}{N_1} \left\{ \frac{1}{N_1 - 1} \sigma_e^4 + 2\sigma_e^2 \sigma_a^2 + \frac{N_2}{N_1} \sigma_a^4 \right\}, \\ \lim_{a \rightarrow \infty} a \text{Cov}(\hat{\sigma}_e^2, \hat{\sigma}_a^2) &= -\frac{2}{N_1(N_1 - 1)} \sigma_e^4 \end{aligned} \quad (3.25)$$

が得られる。これらの値を (3.22) に代入すると、

$$I_1 = 2 \left( \frac{1}{N_1 - 1} + 2 \frac{N_2}{N_1^2} \right) \sigma_e^4 \sigma_a^4 + \frac{2}{N_1(N_1 - 1)} (\sigma_e^2 + N_1 \sigma_a^2) \sigma_e^6 \quad (3.26)$$

となる。同様にして、

$$I_2 = \frac{2}{N_1} \left( \frac{N_2}{N_1} - \frac{2}{N_1 - 1} \right) \sigma_e^8 \sigma_a^4 + 2 \left( \frac{\sigma_a^8}{N_1 - 1} + \frac{\sigma_e^8}{N_1(N_1 - 1)} + \frac{2}{N_1} \sigma_e^6 \sigma_a^2 \right) \quad (3.27)$$

が得られる。 $I_3$  については、 $\bar{y} = \sum_{i=1}^a n_i \bar{y}_i / N \sim \mathcal{N}(\mu, \sum_{i=1}^a n_i^2 (\sigma_e^2 / n_i + \sigma_a^2) / N^2)$  であるから、

$$I_3 = \frac{1}{N_1} \sigma_e^2 + \frac{N_2}{N_1^2} \sigma_a^2 \quad (3.28)$$

となる。

次に、 $I_4$  を評価しよう。このためには若干の計算が必要になる。一般性を失うことなく  $\mu = 0$  としてよい。2乗を展開すると、 $I_4 = \sigma_a^4 I_{41} + \sigma_e^4 I_{42} - 2\sigma_a^2 \sigma_e^2 I_{43}$  と書け、 $I_{41}$ 、 $I_{42}$ 、 $I_{43}$  は、それぞれ  $I_{41} = \lim_{a \rightarrow \infty} aE[(\hat{\sigma}_e^2 - \sigma_e^2)^2 \bar{y}_1^2]$ 、 $I_{42} = \lim_{a \rightarrow \infty} aE[(\hat{\sigma}_a^2 - \sigma_a^2)^2 \bar{y}_1^2]$ 、 $I_{43} = \lim_{a \rightarrow \infty} aE[(\hat{\sigma}_e^2 - \sigma_e^2)(\hat{\sigma}_a^2 - \sigma_a^2) \bar{y}_1^2]$  で与えられる。 $\hat{\sigma}_e^2$  と  $\bar{y}_1$  の独立性から、

$$I_{41} = \lim_{a \rightarrow \infty} a \text{Var}(\hat{\sigma}_e^2) \text{Var}(\bar{y}_1) \quad (3.29)$$

となることは容易にわかる。次に  $I_{42}$  を評価したいが、 $\hat{\sigma}_a^2$  と  $\bar{y}_1$  は独立ではないので期待値を個々に評価する必要がある。 $I_{42}$  を

$$\begin{aligned} I_{42} &= \lim_{a \rightarrow \infty} aE[(\hat{\sigma}_a^2 - \sigma_a^2)^2] E[\bar{y}_1^2] + \lim_{a \rightarrow \infty} aE[(\hat{\sigma}_a^2 - \sigma_a^2)^2 (\bar{y}_1^2 - E[\bar{y}_1^2])] \\ &= J_1 + J_2 \end{aligned} \quad (3.30)$$

と分解すると、 $J_1$  は (3.25) から

$$J_1 = \lim_{a \rightarrow \infty} a \text{Var}(\hat{\sigma}_a^2) \text{Var}(\bar{y}_1) \quad (3.31)$$

となることがわかる。 $J_2$  については、(3.23) より、

$$\begin{aligned} &E[(\hat{\sigma}_a^2 - \sigma_a^2)^2 (\bar{y}_1^2 - E[\bar{y}_1^2])] \\ &= E \left[ \left\{ C \{ S_2 - (a-1)\sigma_e^2 - \sigma_a^2 / C \} - (a-1)C \{ \hat{\sigma}_e^2 - \sigma_e^2 \} \right\}^2 (\bar{y}_1^2 - E[\bar{y}_1^2]) \right] \\ &= C^2 E \left[ \{ S_2^2 - 2[(a-1)\sigma_e^2 + \sigma_a^2 / C] S_2 \} (\bar{y}_1^2 - E[\bar{y}_1^2]) \right] \end{aligned} \quad (3.32)$$

とかける。 $E[S_2(\bar{y}_1^2 - E[\bar{y}_1^2])]$  については、 $\bar{y}_1$  とそれ以外の項に分けて、

$$\begin{aligned} &E[S_2(\bar{y}_1^2 - E[\bar{y}_1^2])] \\ &= E \left[ \sum_{i=1}^a n_i (\bar{y}_i - \bar{y})^2 (\bar{y}_1^2 - E[\bar{y}_1^2]) \right] \\ &= E \left[ \left\{ n_1 \bar{y}_1^2 + \sum_{j=2}^a n_j \bar{y}_j^2 - N^{-1} \left( n_1 \bar{y}_1 + \sum_{j=2}^a n_j \bar{y}_j \right)^2 \right\} (\bar{y}_1^2 - E[\bar{y}_1^2]) \right] \\ &= n_1 (1 - n_1 / N) E[\bar{y}_1^4 - \bar{y}_1^2 E[\bar{y}_1^2]] \\ &= 2(1 - n_1 / N) (\sigma_e^2 + n_1 \sigma_a^2)^2 / n_1 \end{aligned} \quad (3.33)$$

となる。同様にして，

$$\begin{aligned}
& E[S_2^2(\bar{y}_1^2 - E[\bar{y}_1^2])] \\
&= E\left[\left\{n_1\bar{y}_1^2 + \sum_{j=2}^a n_j\bar{y}_j^2 - N^{-1}\left(n_1\bar{y}_1 + \sum_{j=2}^a n_j\bar{y}_j\right)^2\right\}^2(\bar{y}_1^2 - E[\bar{y}_1^2])\right] \\
&= n_1^2(1 - n_1/N)^2 E[\bar{y}_1^4(\bar{y}_1^2 - E[\bar{y}_1^2])] + 4(n_1^2/N^2)E\left[\left(\sum_{j=2}^a n_j\bar{y}_j\right)^2 E[\bar{y}_1^2(\bar{y}_1^2 - E[\bar{y}_1^2])]\right] \\
&\quad + 2n_1(1 - n_1/N)E\left[\sum_{j=2}^a n_j\bar{y}_j^2 + N^{-1}\left(\sum_{j=2}^a n_j\bar{y}_j\right)^2\right]E[\bar{y}_1^2(\bar{y}_1^2 - E[\bar{y}_1^2])] \\
&= 12n_1^2(1 - n_1/N)^2(\sigma_e^2/n_1 + \sigma_a^2)^3 + 8(n_1^2/N^2)\sum_{j=2}^a n_j(\sigma_e^2 + n_j\sigma_a^2)(\sigma_e^2/n_1 + \sigma_a^2)^2 \\
&\quad + 4n_1(1 - n_1/N)\sum_{j=2}^a (1 + n_j/N)(\sigma_e^2 + n_j\sigma_a^2)(\sigma_e^2/n_1 + \sigma_a^2)^2 \\
&= 4n_1(1 - n_1/N)\sum_{i=1}^a (1 - n_i/N)(\sigma_e^2 + n_i\sigma_a^2)(\sigma_e^2/n_1 + \sigma_a^2)^2 \\
&\quad + 8(n_1/N)\sum_{i=1}^a n_i(\sigma_e^2 + n_i\sigma_a^2)(\sigma_e^2/n_1 + \sigma_a^2)^2 \\
&\quad + 8n_1^2(1 - 2n_1/N + n_1^2/N^2 - n_1^3/N)(\sigma_e^2/n_1 + \sigma_a^2)^3 \tag{3.34}
\end{aligned}$$

と書くことができる。(3.33), (3.34) を (3.32) に代入すると，

$$\begin{aligned}
& 2[(a - 1)\sigma_e^2 + \sigma_a^2/C]E[S_2(\bar{y}_1^2 - E[\bar{y}_1^2])] \\
&= 4n_1(1 - n_1/N)\sum_{i=1}^a (1 - n_i/N)(\sigma_e^2 + n_i\sigma_a^2)(\sigma_e^2/n_1 + \sigma_a^2)^2
\end{aligned}$$

となることから，これと (3.34) の右辺の第 1 項とがキャンセルすることがわかる。残りの項はすべて  $O(a^{-3/2})$  であるから，結局  $J_2 = 0$  となるので，

$$I_{42} = J_1 = (3.31) \tag{3.35}$$

となる。 $I_{43}$  についても (3.23) を用いて展開すると，(3.24) より

$$E[(\hat{\sigma}_e^2 - \sigma_e^2)(\hat{\sigma}_a^2 - \sigma_a^2)\bar{y}_1^2] = Cov(\hat{\sigma}_e^2, \hat{\sigma}_a^2)Var(\bar{y}_1)$$

となるので，

$$I_{43} = \lim_{a \rightarrow \infty} aCov(\hat{\sigma}_e^2, \hat{\sigma}_a^2)Var(\bar{y}_1) \tag{3.36}$$

と評価される。以上より，(3.29), (3.35), (3.36) を  $I_4$  に代入すると， $I_4$  については

$$\begin{aligned}
I_4 &= \lim_{a \rightarrow \infty} a\{\sigma_a^4 Var(\hat{\sigma}_e^2) + \sigma_e^4 Var(\hat{\sigma}_a^2) - 2\sigma_e^2\sigma_a^2 Cov(\hat{\sigma}_e^2, \hat{\sigma}_a^2)\} Var(\bar{y}_1) \\
&= I_1 \times (\sigma_e^2/n_1 + \sigma_a^2) \tag{3.37}
\end{aligned}$$

と評価することができる。

最後に、 $\mu = 0$  として  $I_5$  を計算しよう。 $\hat{\sigma}_e^2$  と  $(\bar{y}_1, \bar{y})$  の独立性と (3.23) とから、

$$\begin{aligned} & E \left[ \{ \sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2) \} \bar{y} \bar{y}_1 \right] \\ &= -CE \left[ \{ S_2 - (a-1)\sigma_e^2 - \sigma_a^2/C \} \bar{y} \bar{y}_1 \right] \sigma_e^2 \\ &= -C \frac{n_1}{N} E \left[ \{ S_2 - (a-1)\sigma_e^2 - \sigma_a^2/C \} \bar{y}_1^2 \right] \sigma_e^2 - C \frac{1}{N} E \left[ S_2 \bar{y}_1 \sum_{j=2}^a n_j \bar{y}_j \right] \sigma_e^2 \end{aligned}$$

と書き換えることができる。(3.33) の結果を用いると、

$$\begin{aligned} E \left[ \{ S_2 - (a-1)\sigma_e^2 - \sigma_a^2/C \} \bar{y}_1^2 \right] &= E[S_2(\bar{y}_1^2 - E[\bar{y}_1^2])] \\ &= 2(1 - n_1/N)(\sigma_e^2 + n_1\sigma_a^2)^2/n_1 \end{aligned} \quad (3.38)$$

と評価できる。(3.33) と同様にして  $S_2$  を展開して期待値を評価すると、

$$\begin{aligned} E \left[ S_2 \bar{y}_1 \sum_{j=2}^a n_j \bar{y}_j \right] &= -2 \frac{n_1}{N} E[\bar{y}_1^2] E \left[ \left( \sum_{j=2}^a n_j \bar{y}_j \right)^2 \right] \\ &= -\frac{2}{N} (\sigma_e^2 + n_1\sigma_a^2) \sum_{j=2}^a n_j (\sigma_e^2 + n_j\sigma_a^2) \end{aligned} \quad (3.39)$$

となる。(3.38), (3.39) とから、

$$\begin{aligned} & E \left[ \{ \sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2) \} \bar{y} \bar{y}_1 \right] \\ &= 2 \frac{C}{N} \left\{ -2(1 - n_1/N)(\sigma_e^2 + n_1\sigma_a^2)^2 + \frac{2}{N} (\sigma_e^2 + n_1\sigma_a^2) \sum_{j=2}^a n_j (\sigma_e^2 + n_j\sigma_a^2) \right\} \end{aligned}$$

と表されるが、これは  $O(a^{-2})$  であるので、 $I_5 = 0$  となってしまう。

以上の計算結果 (3.26), (3.27), (3.28), (3.37) と  $I_5 = 0$  を (3.21) に代入すると、命題 3.1 の式を得ることができる。

## 4 混合線形モデルへの拡張

さて、前節で求めた結果を混合線形モデルへ拡張することを考えよう。前節と同様、全体で  $a$  個の小地域があり、 $i$  番目の小地域から  $n_i$  個の標本  $y_{i1}, \dots, y_{in_i}$  がとられており、各  $y_{ij}$  には共変量  $x_{ij}$  が存在していて、 $i = 1, \dots, a; j = 1, \dots, n_i$  に対して

$$\begin{aligned} y_{ij} &= \theta_{ij} + \epsilon_{ij}, \\ \theta_{ij} &= \mathbf{x}_{ij}^t \boldsymbol{\beta} + \alpha_i \end{aligned} \quad (4.1)$$

となるような混合線形モデルを扱う。ここで、誤差項  $\epsilon_{ij}$ 、地域効果  $\alpha_i$  は互いに独立な確率変数で、 $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_e^2)$ 、 $\alpha_i \sim \mathcal{N}(0, \sigma_a^2)$  に従っている。また共変量  $x_{ij}$  は  $m \times 1$  の既知のベクトルであると、 $\mathbf{x}_{ij}^t$  は  $x_{ij}$  の転置を表すものとする。 $\boldsymbol{\beta}$  は回帰係数ベクトルで、すべての小地域を通して共通に設定されている。1 番目の小地域について共変量  $x$  が与えられたときに、

$$\theta_1 = \mathbf{x}^t \boldsymbol{\beta} + \alpha_1$$

の信頼区間を構成する問題を考えよう。特に  $x = \bar{x}_1$  のときには、1 番目の小地域の平均値を意味している。

まず、信頼区間や母数の推定量を与えるために、すべての  $y_{ij}$  と  $\alpha_i$  の同時密度関数の指数の部分の次のように表現しておく。

$$\begin{aligned}
& \sum_{i=1}^a \sum_{j=1}^{n_i} (y_{ij} - \mathbf{x}_{ij}^t \boldsymbol{\beta} - \alpha_i)^2 / \sigma_e^2 + \sum_{i=1}^a \alpha_i^2 / \sigma_a^2 \\
&= \sum_{i=1}^a \left( \frac{n_i}{\sigma^2} + \frac{1}{\sigma_a^2} \right) \left\{ \alpha_i - \frac{n_i \sigma_a^2}{n_i \sigma_a^2 + \sigma^2} (\bar{y}_i - \bar{\mathbf{x}}_i^t \boldsymbol{\beta}) \right\}^2 \\
&+ \frac{1}{\sigma_e^2} \sum_{i=1}^a \sum_{j=1}^{n_i} \left\{ (y_{ij} - \bar{y}_i) - (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)^t \boldsymbol{\beta} \right\}^2 \\
&+ \sum_{i=1}^a \frac{n_i}{n_i \sigma_a^2 + \sigma_e^2} (\bar{y}_i - \bar{\mathbf{x}}_i^t \boldsymbol{\beta})^2 \tag{4.2}
\end{aligned}$$

式 (4.2) の右辺の各項をそれぞれ  $J_A, J_B, J_C$  とおくと、 $J_A$  より、 $\bar{y}_i$  を与えたときの  $\alpha_i$  の条件付き分布は、

$$\alpha_i | \bar{y}_i \sim \mathcal{N} \left( \frac{n_i \sigma_a^2}{n_i \sigma_a^2 + \sigma^2} (\bar{y}_i - \bar{\mathbf{x}}_i^t \boldsymbol{\beta}), \frac{\sigma_e^2 \sigma_a^2}{\sigma_e^2 + n_i \sigma_a^2} \right)$$

となることがわかる。従って、 $(\sigma_e^2, \sigma_a^2, \boldsymbol{\beta})$  が既知のときには、信頼係数  $1 - \alpha$  の  $\theta_1$  の信頼区間は

$$I_B : \mathbf{x}^t \boldsymbol{\beta} + (1 - B)(\bar{y}_1 - \bar{\mathbf{x}}_1^t \boldsymbol{\beta}) \pm z_{\alpha/2} \sqrt{V} \tag{4.3}$$

で与えられることになる。ただし、 $B, V$  は (3.1) で定義されている。

次に、未知母数  $\sigma_e^2, \sigma_a^2, \boldsymbol{\beta}$  を (4.2) から推定して (4.3) に代入する。 $\sigma_e^2$  は  $J_B$  の残差平方和から

$$\hat{\sigma}_e^2 = \frac{1}{N - a - m + \lambda} \sum_{i=1}^a \sum_{j=1}^{n_i} \left\{ (y_{ij} - \bar{y}_i) - (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)^t \hat{\boldsymbol{\beta}}_1 \right\}^2 \tag{4.4}$$

で推定される。ただし、 $m - \lambda$  は行列  $\sum_{i=1}^a \sum_{j=1}^{n_i} (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)^t$  のランクを表しており、通常は線形モデル (4.1) が定数項をもつときには  $\lambda = 1$ 、もたないときには  $\lambda = 0$  となる。また  $\hat{\boldsymbol{\beta}}_1$  は  $J_B$  における  $\boldsymbol{\beta}$  の最小 2 乗推定量を表している。(4.2) から、 $\hat{\sigma}_e^2$  は  $\bar{y}_1, \dots, \bar{y}_a$  と独立に分布し、

$$\begin{aligned}
(N - a - m + \lambda) \hat{\sigma}_e^2 / \sigma_e^2 &\sim \chi_{N - a - m + \lambda}^2, \\
\bar{y}_i &\sim \mathcal{N}(\bar{\mathbf{x}}_i^t \boldsymbol{\beta}, \sigma_e^2 / n_i + \sigma_a^2), \quad i = 1, \dots, a,
\end{aligned}$$

に従うことがわかる。 $\sigma_a^2$  については、2 次形式  $\sum_{i=1}^a n_i (\bar{y}_i - \bar{\mathbf{x}}_i^t \boldsymbol{\beta})^2$  の残差平方和に基づいて推定することが考えられる。そのとき、 $\boldsymbol{\beta}$  の最小 2 乗推定量は、 $\sum_{i=1}^a n_i \bar{\mathbf{x}}_i \bar{\mathbf{x}}_i^t$  を正則行列と仮定すると、

$$\hat{\boldsymbol{\beta}}_2 = \left( \sum_{i=1}^a n_i \bar{\mathbf{x}}_i \bar{\mathbf{x}}_i^t \right)^{-1} \sum_{i=1}^a n_i \bar{\mathbf{x}}_i \bar{y}_i$$

で与えられ、残差平方和を

$$S_2 = \sum_{i=1}^a n_i (\bar{y}_i - \bar{\mathbf{x}}_i^t \hat{\boldsymbol{\beta}}_2)^2 \tag{4.5}$$

とおくことにする。\$S\_2\$ の期待値は、

$$\begin{aligned} E[S_2] &= E \left[ \sum_{i=1}^a n_i (\bar{y}_i - \bar{\mathbf{x}}_i^t \boldsymbol{\beta})^2 - (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta})^t \sum_{i=1}^a n_i \bar{\mathbf{x}}_i \bar{\mathbf{x}}_i^t (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}) \right] \\ &= (a\sigma_e^2 + N\sigma_a^2) - \left\{ m\sigma_e^2 + \sigma_a^2 \text{tr} \left( \sum_{i=1}^a n_i^2 \mathbf{x}_i \bar{\mathbf{x}}_i^t \right) \left( \sum_{i=1}^a n_i \mathbf{x}_i \bar{\mathbf{x}}_i^t \right)^{-1} \right\} \\ &= (a - m)\sigma_e^2 + N_*\sigma_a^2 \end{aligned}$$

と書ける。ただし、\$N = \sum\_{i=1}^a n\_i\$ に対して、

$$N_* = N - \text{tr} \left( \sum_{i=1}^a n_i^2 \mathbf{x}_i \bar{\mathbf{x}}_i^t \right) \left( \sum_{i=1}^a n_i \mathbf{x}_i \bar{\mathbf{x}}_i^t \right)^{-1}$$

である。従って、\$\sigma\_a^2\$ の1つの不偏推定量は、

$$\hat{\sigma}_a^2 = \frac{1}{N_*} \{ S_2 - (a - m)\hat{\sigma}_e^2 \}$$

で与えられることがわかる。これらの推定量を (4.3) の信頼区間 \$I\_B\$ に代入すると、被覆確率が信頼係数 \$1 - \alpha\$ より小さくなってしまいうので、(3.4) と同様にして \$z\_{\alpha/2}\$ を \$z\_{\alpha/2}(1 + (2a)^{-1}\hat{h})\$ で調整した信頼区間

$$I_{AEB} : \mathbf{x}^t \hat{\boldsymbol{\beta}}_2 + (1 - \hat{B})(\bar{y}_1 - \bar{\mathbf{x}}_1^t \hat{\boldsymbol{\beta}}_2) \pm z_{\alpha/2} \{ 1 + (2a)^{-1}\hat{h} \} \hat{V}^{1/2} \quad (4.6)$$

を考えよう。ここで、\$\hat{B}, \hat{V}\$ はそれぞれ \$B, V\$ の推定量で (3.3) で定義された形をしている。以降、簡単のために \$z = z\_{\alpha/2}\$ とおいて議論を進める。

(3.5) と同様にして、信頼区間 \$I\_{AEB}\$ の被覆確率は、

$$P[\theta_1 \in I_{AEB}] = E[\Phi(z + G(z)) - \Phi(-z + G(-z))]$$

となり、ここでは \$G(z)\$ は

$$\begin{aligned} G(z) &= V^{-1/2} \left\{ (B - \hat{B})(\bar{y}_1 - \bar{\mathbf{x}}_1^t \boldsymbol{\beta}) + \{ \mathbf{x} + (\hat{B} - 1)\bar{\mathbf{x}}_1 \}^t (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}) \right\} \\ &\quad + z \left[ \{ 1 + (2a)^{-1}\hat{h} \} \hat{V}^{1/2} / V^{1/2} - 1 \right] \end{aligned}$$

で与えられる。\$\hat{B}\$ のテーラー展開 (3.7) を用いると、第1項は

$$\begin{aligned} &V^{-1/2} \left\{ (B - \hat{B})(\bar{y}_1 - \bar{\mathbf{x}}_1^t \boldsymbol{\beta}) + \{ \mathbf{x} + (\hat{B} - 1)\bar{\mathbf{x}}_1 \}^t (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}) \right\} \\ &= \frac{1}{\sqrt{a}} \hat{u}_1 + \frac{1}{a} \hat{u}_2 + O_p(a^{-3/2}) \end{aligned}$$

と評価される。ただし、\$\tau = 1/(\sigma\_e^2 + n\_1\sigma\_a^2)\$ に対して、

$$\begin{aligned} \hat{u}_1 &= \sqrt{a} V^{-1/2} \left[ \{ \mathbf{x} + (B - 1)\bar{\mathbf{x}}_1 \}^t (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}) \right. \\ &\quad \left. - n_1 \tau^2 \{ \sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2) \} (\bar{y}_1 - \bar{\mathbf{x}}_1^t \boldsymbol{\beta}) \right], \\ \hat{u}_2 &= a n_1 \tau^2 V^{-1/2} \left[ \{ \sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2) \} \bar{\mathbf{x}}_1^t (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}) \right. \\ &\quad \left. - \tau \{ n_1 \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2)^2 - \sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2)^2 \right. \\ &\quad \left. + (\sigma_e^2 - n_1 \sigma_a^2) (\hat{\sigma}_a^2 - \sigma_a^2) (\hat{\sigma}_e^2 - \sigma_e^2) \} (\bar{y}_1 - \bar{\mathbf{x}}_1^t \boldsymbol{\beta}) \right] \end{aligned}$$

である。 $\hat{V}^{1/2}/V^{1/2}$  に対しては, (3.11), (3.12) で与えられる  $\hat{v}_1, \hat{v}_2$  を用いて, (3.10) と同じ形の評価式

$$\hat{V}^{1/2}/V^{1/2} = 1 + \frac{1}{\sqrt{a}}\hat{v}_1 + \frac{1}{a}\hat{v}_2 + O_p(a^{-3/2})$$

が成り立つ。従って,

$$G(z) = \frac{1}{\sqrt{a}}\hat{u}_1 + \frac{1}{a}\hat{u}_2 + z \left[ \frac{1}{\sqrt{a}}\hat{v}_1 + \frac{1}{a} \left\{ \hat{v}_2 + \frac{1}{2}\hat{h} \right\} \right] + O_p(a^{-3/2})$$

と表現されることになる。関数  $h(\sigma_e^2, \sigma_a^2)$  を

$$h(\sigma_e^2, \sigma_a^2) = \lim_{a \rightarrow \infty} \{ E[-2\hat{v}_2 + \hat{u}_1^2 + z^2\hat{v}_1^2] \}$$

で定義すると, (3.14) から (3.15) までの議論から,

$$P[\theta_1 \in I_{AEB}] = 1 - \alpha + \frac{z}{a}\phi(z) \left\{ E[\hat{h}] - h(\sigma_e^2, \sigma_a^2) \right\} + O(a^{-3/2})$$

となることがわかる。 $h(\sigma_e^2, \sigma_a^2)$  に推定量  $\hat{\sigma}_e^2, \hat{\sigma}_a^2$  を代入して  $\hat{h} = h(\hat{\sigma}_e^2, \hat{\sigma}_a^2)$  とおくと,  $P[\theta_1 \in I_{AEB}] = 1 - \alpha + O(a^{-3/2})$  となり, 要求した条件をみたま信頼区間が得られることになる。

命題 4.1  $k = 1, 2$  に対して,  $N_k = \lim_{a \rightarrow \infty} \sum_{i=1}^a n_i^k / a$ ,

$$M_k = \lim_{a \rightarrow \infty} \sum_{i=1}^a n_i^k \bar{\mathbf{x}}_i \bar{\mathbf{x}}_i^t / a$$

と定義し, これらの値が存在し,  $N_1$  が正であり  $M_1$  が正則であることが仮定される。このとき, 関数  $h(\sigma_e^2, \sigma_a^2)$  は,

$$\begin{aligned} h(\sigma_e^2, \sigma_a^2) &= \lim_{a \rightarrow \infty} \left\{ E[-2\hat{v}_2 + \hat{u}_1^2 + z^2\hat{v}_1^2] \right\} \\ &= \frac{4n_1}{\sigma_e^2 \sigma_a^2 (\sigma_e^2 + n_1 \sigma_a^2)^2} \left\{ \left( \frac{1}{N_1 - 1} + 2 \frac{N_2}{N_1^2} \right) \sigma_e^4 \sigma_a^4 + \frac{1}{N_1(N_1 - 1)} (\sigma_e^2 + N_1 \sigma_a^2) \sigma_e^6 \right\} \\ &\quad + \frac{1 + z^2}{2\sigma_e^4 \sigma_a^4 (\sigma_e^2 + n_1 \sigma_a^2)^2} \left\{ \frac{1}{N_1} \left( \frac{N_2}{N_1} - \frac{2}{N_1 - 1} \right) \sigma_e^8 \sigma_a^4 \right. \\ &\quad \left. + \left( \frac{\sigma_a^8}{N_1 - 1} + \frac{\sigma_e^8}{N_1(N_1 - 1)} + \frac{2}{N_1} \sigma_e^6 \sigma_a^2 \right) \right\} \\ &\quad + \frac{\sigma_e^2 + n_1 \sigma_a^2}{\sigma_e^2 \sigma_a^2} \{ \mathbf{x} + (B - 1)\mathbf{x}_1 \}^t (M_1^{-1} \sigma_e^2 + M_1^{-1} M_2 M_1^{-1} \sigma_a^2) \{ \mathbf{x} + (B - 1)\mathbf{x}_1 \} \end{aligned}$$

で与えられる。

この関数  $h(\sigma_e^2, \sigma_a^2)$  は分散比  $\rho = \sigma_a^2 / \sigma_e^2$  を用いて

$$\begin{aligned} h^*(\rho) &= h(\sigma_e^2, \sigma_a^2) \\ &= \frac{4n_1}{\rho(1 + n_1\rho)^2} \left\{ \left( \frac{1}{N_1 - 1} + 2 \frac{N_2}{N_1^2} \right) \rho^2 + \frac{1}{N_1(N_1 - 1)} (1 + N_1\rho) \right\} \\ &\quad + \frac{1 + z^2}{2\rho^2(1 + n_1\rho)^2} \left\{ \frac{1}{N_1} \left( \frac{N_2}{N_1} - \frac{2}{N_1 - 1} \right) \rho^2 \right. \\ &\quad \left. + \left( \frac{\rho^4}{N_1 - 1} + \frac{1}{N_1(N_1 - 1)} + \frac{2}{N_1} \rho \right) \right\} \\ &\quad + \frac{1 + n_1\rho}{\rho} \left\{ \mathbf{x} - \frac{n_1\rho}{1 + n_1\rho} \mathbf{x}_1 \right\}^t (M_1^{-1} + M_1^{-1} M_2 M_1^{-1} \rho) \left\{ \mathbf{x} - \frac{n_1\rho}{1 + n_1\rho} \mathbf{x}_1 \right\} \end{aligned}$$

と表される。分散成分の不偏推定量  $\hat{\sigma}_e^2, \hat{\sigma}_a^2$  を用いた  $\rho$  の推定量は  $\hat{\rho} = \hat{\sigma}_a^2 / \hat{\sigma}_e^2$  であり, これを  $h^*(\rho)$  に代入すれば信頼区間  $I_{AEB}$  が構成できたことになる。  $I_{AEB}$  の中の  $\hat{B}$  は  $\hat{B}(\hat{\rho}) = (1 + n_1 \hat{\rho})^{-1}$  であることを思い出すと, 以上の結果は次の命題としてまとめられる。

命題 4.2 (4.6) で与えられる信頼区間  $I_{AEB}$  は,

$$I_{AEB}: \mathbf{x}^t \hat{\beta}_2 + \frac{n_1 \hat{\rho}}{1 + n_1 \hat{\rho}} (\bar{y}_1 - \bar{\mathbf{x}}_1^t \hat{\beta}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{\sigma}_e^2 \hat{\rho}}{1 + n_1 \hat{\rho}}} \left(1 + \frac{1}{2a} h^*(\hat{\rho})\right) \quad (4.7)$$

と表されて,  $a \rightarrow \infty$  のときに  $P[\theta_1 \in I_{AEB}] = 1 - \alpha + O(a^{-3/2})$  となる。

命題 3.2 の下の議論から, 実際の応用場面では,  $\hat{\rho} = \hat{\sigma}_a^2 / \hat{\sigma}_e^2$  の代わりに

$$\hat{\rho}^{TR} = \hat{\sigma}_a^{2TR} / \hat{\sigma}_e^{2TR} = \frac{a - m}{N_*} \min \left\{ \frac{S_2}{(a - m) \hat{\sigma}_e^2} - 1, \frac{2}{a - m} \right\}$$

を用いて  $\rho$  を推定した方がよい。この打ち切り推定量を用いると, 信頼区間は

$$I_{TR}: \mathbf{x}^t \hat{\beta}_2 + \frac{n_1 \hat{\rho}^{TR}}{1 + n_1 \hat{\rho}^{TR}} (\bar{y}_1 - \bar{\mathbf{x}}_1^t \hat{\beta}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{\sigma}_e^2 \hat{\rho}^{TR}}{1 + n_1 \hat{\rho}^{TR}}} \left(1 + \frac{1}{2a} h^*(\hat{\rho}^{TR})\right) \quad (4.8)$$

で与えられる。

最後に, 命題 4.1 の証明のスケッチを与えておこう。これは, 命題 3.1 の証明と全く同じ論法に沿って確かめることができる。まず,

$$\begin{aligned} h(\sigma_e^2, \sigma_a^2) &= \lim_{a \rightarrow \infty} E[-2\hat{v}_2 + \hat{u}_1^2 + z^2 \hat{v}_1^2] \\ &= \frac{n_1 \tau^3}{V} I_1 + \frac{(1 + z^2) \tau^4}{4V^2} I_2 + \frac{1}{V} I_3 + \frac{n_1^2 \tau^4}{V} I_4 - 2 \frac{n_1 \tau^2}{V} I_5 \end{aligned} \quad (4.9)$$

と表されることに注意する。ここで,

$$\begin{aligned} I_1 &= \lim_{a \rightarrow \infty} a \{ \sigma_a^4 \text{Var}(\hat{\sigma}_e^2) + \sigma_e^4 \text{Var}(\hat{\sigma}_a^2) - 2\sigma_e^2 \sigma_a^2 \text{Cov}(\hat{\sigma}_e^2, \hat{\sigma}_a^2) \}, \\ I_2 &= \lim_{a \rightarrow \infty} a \{ \sigma_a^8 \text{Var}(\hat{\sigma}_e^2) + \sigma_e^8 \text{Var}(\hat{\sigma}_a^2) + 2\sigma_e^4 \sigma_a^4 \text{Cov}(\hat{\sigma}_e^2, \hat{\sigma}_a^2) \}, \\ I_3 &= \lim_{a \rightarrow \infty} a E[\{\mathbf{x} + (B - 1)\mathbf{x}_1\}^t (\hat{\beta}_2 - \beta) (\hat{\beta}_2 - \beta)^t \{\mathbf{x} + (B - 1)\mathbf{x}_1\}], \\ I_4 &= \lim_{a \rightarrow \infty} a E[\{\sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2)\}^2 (\bar{y}_1 - \bar{\mathbf{x}}_1^t \beta)^2], \\ I_5 &= \lim_{a \rightarrow \infty} a E[\{\sigma_a^2 (\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2 (\hat{\sigma}_a^2 - \sigma_a^2)\} \{\mathbf{x} + (B - 1)\mathbf{x}_1\}^t (\hat{\beta}_2 - \beta) (\bar{y}_1 - \bar{\mathbf{x}}_1^t \beta)] \end{aligned}$$

で与えられる。これらの値を評価するために, 推定量  $\hat{\sigma}_e^2, \hat{\sigma}_a^2$  の分散および共分散の極限值を計算しておく。まず,  $\text{Var}(\hat{\sigma}_e^2) = 2\sigma_e^4 / (N - a - m + \lambda)$  より,

$$\lim_{a \rightarrow \infty} a \text{Var}(\hat{\sigma}_e^2) = \frac{2}{N_1 - 1} \sigma_e^4$$

となる。また

$$\begin{aligned}
\text{Var}(\hat{\sigma}_a^2) &= E \left[ \{N_*^{-1}(S_2 - (a - m)\hat{\sigma}_e^2) - \sigma_a^2\}^2 \right] \\
&= \frac{1}{N_*^2} E \left[ \{S_2 - (a - m)\sigma_e^2 - N_*\sigma_a^2\}^2 \right] + \left( \frac{a - m}{N_*} \right)^2 \text{Var}(\hat{\sigma}_e^2) \\
&= \frac{1}{N_*^2} E [S_2^2] - \frac{1}{N_*^2} \{(a - m)\sigma_e^2 + N_*\sigma_a^2\}^2 + \left( \frac{a - m}{N_*} \right)^2 \text{Var}(\hat{\sigma}_e^2)
\end{aligned} \tag{4.10}$$

となる。ここで、

$$\begin{aligned}
E[S_2^2] &= E \left[ \left\{ \sum_{i=1}^a n_i \left[ (\bar{y}_i - \bar{\mathbf{x}}_i^t \boldsymbol{\beta}) - \bar{\mathbf{x}}_i^t (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}) \right]^2 \right\}^2 \right] \\
&= E \left[ \left\{ \sum_{i=1}^a n_i (\bar{y}_i - \bar{\mathbf{x}}_i^t \boldsymbol{\beta})^2 - \sum_{i=1}^a n_i \bar{\mathbf{x}}_i^t (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}) \right\}^2 \right] \\
&= E \left[ \left\{ \sum_{i=1}^a n_i (\bar{y}_i - \bar{\mathbf{x}}_i^t \boldsymbol{\beta})^2 \right\}^2 + \left\{ \sum_{i=1}^a n_i \bar{\mathbf{x}}_i^t (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}) \right\}^2 \right. \\
&\quad \left. - 2 \sum_{i=1}^a n_i (\bar{y}_i - \bar{\mathbf{x}}_i^t \boldsymbol{\beta}) \sum_{i=1}^a n_i \bar{\mathbf{x}}_i^t (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}) \right]
\end{aligned}$$

と書ける。  $A = \sum_{i=1}^a n_i \bar{\mathbf{x}}_i \bar{\mathbf{x}}_i^t$ ,  $z_i = \bar{y}_i - \bar{\mathbf{x}}_i^t \boldsymbol{\beta}$  とおくと,  $\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta} = A^{-1} \sum_{i=1}^a n_i \bar{\mathbf{x}}_i z_i$  と書けることに注意すると、

$$\begin{aligned}
E[S_2^2] &= E \left[ \sum_{i,j} n_i n_j z_i^2 z_j^2 + \sum_{i,j,k,\ell} n_i n_j n_k n_\ell \bar{\mathbf{x}}_i^t A^{-1} \bar{\mathbf{x}}_j \bar{\mathbf{x}}_k^t A^{-1} \bar{\mathbf{x}}_\ell z_i z_j z_k z_\ell \right. \\
&\quad \left. - 2 \sum_{i,j} n_i n_j^2 \bar{\mathbf{x}}_j^t A^{-1} \bar{\mathbf{x}}_j z_i^2 z_j^2 \right]
\end{aligned}$$

と書けるので、 $a \rightarrow \infty$  のときの主要項のみを評価すると、

$$\begin{aligned}
&\lim_{a \rightarrow \infty} a \text{Var}(\hat{\sigma}_a^2) \\
&= \frac{1}{N_1^2} \lim_{a \rightarrow \infty} \frac{1}{a} \left\{ 2 \sum_i (\sigma_e^2 + n_i \sigma_a^2)^2 + \left( \sum_i (\sigma_e^2 + n_i \sigma_a^2) \right)^2 \right. \\
&\quad \left. - 2 \sum_i (\sigma_e^2 + n_i \sigma_a^2) \sum_j n_j (\sigma_e^2 + n_j \sigma_a^2) \bar{\mathbf{x}}_j^t A^{-1} \bar{\mathbf{x}}_j \right. \\
&\quad \left. - (a \sigma_e^2 + N \sigma_a^2)^2 + 2(a \sigma_e^2 + N \sigma_a^2)(m \sigma_e^2 + \text{tr} \sum_{i=1}^a n_i^2 \bar{\mathbf{x}}_i \bar{\mathbf{x}}_i^t A^{-1} \sigma_a^2) \right\} \\
&\quad + \frac{2}{N_1^2 (N_1 - 1)} \sigma_e^4 \\
&= \frac{2}{N_1} \left\{ \frac{1}{N_1} \sigma_e^4 + 2 \sigma_e^2 \sigma_a^2 + \frac{N_2}{N_1} \sigma_a^4 \right\} + \frac{2}{N_1^2 (N_1 - 1)} \sigma_e^4 \\
&= \frac{2}{N_1} \left\{ \frac{1}{N_1 - 1} \sigma_e^4 + 2 \sigma_e^2 \sigma_a^2 + \frac{N_2}{N_1} \sigma_a^4 \right\}
\end{aligned}$$

となることがわかる。また,

$$\lim_{a \rightarrow \infty} a \text{Cov}(\hat{\sigma}_e^2, \hat{\sigma}_a^2) = -\frac{2}{N_1(N_1 - 1)} \sigma_e^4$$

が得られる。これらの値は第3節で与えられたもの (3.25) に等しいので,  $I_1, I_2$  はそれぞれ (3.26), (3.27) で与えられるものに一致する。

$I_3$  については,

$$E[(\hat{\beta}_2 - \beta)(\hat{\beta}_2 - \beta)^t] = A^{-1} \sum_{i,j} n_i \bar{\mathbf{x}}_i E[z_i z_j] n_j \bar{\mathbf{x}}_j^t A^{-1} = A^{-1} \sum_i n_i^2 E[z_i^2] \bar{\mathbf{x}}_i \bar{\mathbf{x}}_i^t A^{-1}$$

と書けることから,

$$\begin{aligned} I_3 &= \{\mathbf{x} + (B-1)\bar{\mathbf{x}}_1\}^t \lim_{a \rightarrow \infty} a E[(\hat{\beta}_2 - \beta)(\hat{\beta}_2 - \beta)^t] \{\mathbf{x} + (B-1)\bar{\mathbf{x}}_1\} \\ &= \{\mathbf{x} + (B-1)\bar{\mathbf{x}}_1\}^t (\mathbf{M}_1^{-1} \sigma_e^2 + \mathbf{M}_1^{-1} \mathbf{M}_2 \mathbf{M}_1^{-1} \sigma_a^2) \{\mathbf{x} + (B-1)\bar{\mathbf{x}}_1\} \end{aligned} \quad (4.11)$$

となる。

$I_4$  については,  $I_4 = \lim_{a \rightarrow \infty} a E[\{\sigma_a^2(\hat{\sigma}_e^2 - \sigma_e^2) - \sigma_e^2(\hat{\sigma}_a^2 - \sigma_a^2)\}^2 z_1^2] = \sigma_a^4 I_{41} + \sigma_e^4 I_{42} - 2\sigma_a^2 \sigma_e^2 I_{43}$  と展開し, それぞれについて評価する。ここで,  $I_{41} = \lim_{a \rightarrow \infty} a E[(\hat{\sigma}_e^2 - \sigma_e^2)^2 z_1^2]$ ,  $I_{42} = \lim_{a \rightarrow \infty} a E[(\hat{\sigma}_a^2 - \sigma_a^2)^2 z_1^2]$ ,  $I_{43} = \lim_{a \rightarrow \infty} a E[(\hat{\sigma}_e^2 - \sigma_e^2)(\hat{\sigma}_a^2 - \sigma_a^2) z_1^2]$  である。 $\hat{\sigma}_e^2$  と  $z_1$  の独立性から,

$$I_{41} = \lim_{a \rightarrow \infty} a \text{Var}(\hat{\sigma}_e^2) \text{Var}(z_1)$$

となることが容易にわかる。 $I_{42}$  の評価は若干面倒であるが, まず  $I_{42}$  を

$$\begin{aligned} I_{42} &= \lim_{a \rightarrow \infty} a \text{Var}(\hat{\sigma}_a^2) \text{Var}(z_1^2) + \lim_{a \rightarrow \infty} a E[(\hat{\sigma}_a^2 - \sigma_a^2)^2 (z_1^2 - E[z_1^2])] \\ &= J_1 + J_2 \end{aligned}$$

と分解すると,  $J_2$  については, (4.10) より,  $W_1 = z_1^2 - E[z_1^2]$  に対して

$$\begin{aligned} J_2 &= \lim_{a \rightarrow \infty} \frac{a}{N_*^2} E \left[ \{S_2 - (a-m)\hat{\sigma}_e^2 - N_* \sigma_a^2\}^2 W_1 \right] \\ &= \lim_{a \rightarrow \infty} \frac{a}{N_*^2} E \left[ \{S_2^2 - 2[(a-m)\sigma_e^2 + N_* \sigma_a^2] S_2\} W_1 \right] \end{aligned}$$

と書ける。 $E[S_2 W_1]$  については,  $z_1$  とそれ以外の項に分けて,

$$\begin{aligned} E[S_2 W_1] &= E \left[ \left\{ \sum_i n_i z_i^2 - \sum_{i,j} n_i n_j \bar{\mathbf{x}}_i^t A^{-1} \bar{\mathbf{x}}_j z_i z_j \right\} W_1 \right] \\ &= E \left[ \left\{ n_1 z_1^2 - n_1^2 \bar{\mathbf{x}}_1^t A^{-1} \bar{\mathbf{x}}_1 z_1^2 \right\} W_1 \right] \\ &= 2n_1 (1 - n_1 \bar{\mathbf{x}}_1^t A^{-1} \bar{\mathbf{x}}_1) (\sigma_e^2/n_1 + \sigma_a^2)^2 \end{aligned}$$

となるので,

$$\begin{aligned} &2 \lim_{a \rightarrow \infty} \frac{a}{N_*^2} \{(a-m)\sigma_e^2 + N_* \sigma_a^2\} E[S_2 W_1] \\ &= \frac{4}{N_1^2} \{\sigma_e^2 + N_1 \sigma_a^2\} n_1 (\sigma_e^2/n_1 + \sigma_a^2)^2 \end{aligned} \quad (4.12)$$

と書ける。一方,  $E[S_2^2 W_1]$  は,

$$E[S_2^2 W_1] = E \left[ \left\{ \left( \sum_i n_i z_i^2 \right)^2 + \left( \sum_{i,j} n_i n_j \bar{\mathbf{x}}_i^t A^{-1} \bar{\mathbf{x}}_j z_i z_j \right)^2 - 2 \left( \sum_i n_i z_i^2 \right) \left( \sum_{i,j} n_i n_j \bar{\mathbf{x}}_i^t A^{-1} \bar{\mathbf{x}}_j z_i z_j \right) \right\} W_1 \right]$$

と書け, 同様に評価していくと, 極限をとったときに残る項は

$$\begin{aligned} & \lim_{a \rightarrow \infty} \frac{a}{N_*^2} E[S_2^2 W_1] \\ &= \lim_{a \rightarrow \infty} \frac{a}{N_*^2} 4n_1 (\sigma_e^2/n_1 + \sigma_a^2) \sum_{j=2}^a n_j (\sigma_e^2/n_j + \sigma_a^2) \\ &= \frac{4}{N_1^2} n_1 (\sigma_e^2/n_1 + \sigma_a^2)^2 \{ \sigma_e^2 + N_1 \sigma_a^2 \} \end{aligned}$$

となる。これは, (4.12) と等しいので,  $J_2 = 0$  となり, 結局

$$I_{42} = J_1 = \lim_{a \rightarrow \infty} a \text{Var}(\hat{\sigma}_a^2) \text{Var}(z_1)$$

となる。 $I_{43}$  についても,

$$\begin{aligned} I_{43} &= \lim_{a \rightarrow \infty} \frac{a}{N_*} E [ (\hat{\sigma}_e^2 - \sigma_e^2) \{ S_2 - (a-m)\hat{\sigma}_e^2 - N_*\sigma_a^2 \} z_1^2 ] \\ &= \lim_{a \rightarrow \infty} -\frac{a}{N_*} (a-m) E [ (\hat{\sigma}_e^2 - \sigma_e^2)^2 ] E [ z_1^2 ] \\ &= \lim_{a \rightarrow \infty} a \text{Cov}(\hat{\sigma}_a^2, \hat{\sigma}_e^2 \text{Var}(z_1)) \end{aligned}$$

と評価できるので, 以上をまとめると,

$$\begin{aligned} I_4 &= \lim_{a \rightarrow \infty} a \{ \sigma_a^4 \text{Var}(\hat{\sigma}_e^2) + \sigma_e^4 \text{Var}(\hat{\sigma}_a^2) - 2\sigma_a^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_e^2, \hat{\sigma}_a^2) \} \text{Var}(z_1) \\ &= I_1 \times (\sigma_e^2/n_1 + \sigma_a^2) \end{aligned}$$

と書くことができる。

最後に,  $I_5$  については,  $\hat{\sigma}_e^2$  と  $z_1, \dots, z_a$  との独立性から,

$$\begin{aligned} I_5 &= - \lim_{a \rightarrow \infty} \frac{a\sigma_e^2}{N_*} E \left[ \{ \mathbf{x} + (B-1)\mathbf{x}_1 \}^t (\hat{\beta}_2 - \beta) \{ S_2 - (a-m)\sigma_e^2 - N_*\sigma_a^2 \} z_1 \right] \\ &= - \lim_{a \rightarrow \infty} \frac{a\sigma_e^2}{N_*} E \left[ \{ \mathbf{x} + (B-1)\mathbf{x}_1 \}^t A^{-1} \left( \sum_i n_i \bar{\mathbf{x}}_i z_i \right) \right. \\ &\quad \left. \times \left\{ \sum_i n_i z_i^2 - \sum_{i,j} n_i n_j \bar{\mathbf{x}}_i^t A^{-1} \bar{\mathbf{x}}_j z_i z_j - (a-m)\sigma_e^2 - N_*\sigma_a^2 \right\} z_1 \right] \end{aligned}$$

と書き換えることができる。これを  $z_1$  とそれ以外の項に分けて評価していくと,  $I_5 = 0$  となることが確かめられる。以上の計算結果を (4.9) に代入すると, 命題 4.1 の式を得ることができる。

## 5 数値実験と応用例

### 5.1 数値実験結果

第2節及び第3節で提案された信頼区間に関して、被覆確率と信頼区間の長さについての数値実験の結果を与える。数値実験では、(2.1) で与えられる変量効果モデルにおいて、群の個数  $a$  を 5 から 40 まで増加させたときの信頼区間の挙動を調べている。調べる信頼区間は、(3.2) で与えられた経験ベイズ信頼区間  $I_{EB}$  と、それを漸近補正した信頼区間  $I_{TR}$  であり、 $I_{TR}$  は (3.20) で与えられている。ここでは簡単のために  $I_{EB}$ ,  $I_{TR}$  はそれぞれ EB, TR で表されている。信頼係数は 95%, すなわち  $\alpha = 0.05$  とする。群内標本の数  $n_i$  は、それぞれ平均 10 のポアソン分布に基づいて発生させる。また群内分散は  $\sigma_e^2 = 1$  とし、群間分散については  $\sigma_a^2 = 0.5, 1.0, 2.0$  の場合が調べられ、平均  $\mu$  については一般性を失うことなく  $\mu = 0$  としている。10,000 回のシミュレーション実験に基づいた結果が以下で与えられる図の中で示されている。

図 1, 2, 3 は、それぞれ  $\sigma_a^2 = \rho$  が 0.5, 1.0, 2.0 のときの信頼区間  $I_{EB}$  と  $I_{TR}$  の被覆確率の挙動を示している。これらの図をながめてみると、EB に比べて TR の被覆確率は信頼係数 0.95 に近い値を与えており、漸近補正の効果がみとれる。 $a$  が小さいときには、EB の被覆確率がかなり小さくなっている一方、TR の被覆確率は相当改善されており、特に  $\sigma_a^2 = 0.5$  のときの図 1 からは、TR における漸近補正は有効に働いていることがわかる。

図 4 は、群の個数  $a$  を 30 に固定し、群間分散  $\sigma_a^2$  を 0.5 から 4.0 まで変化させたときの様子を描いたものである。TR による EB の改良は、 $\sigma_a^2$  が小さいときに顕著に現れているが、 $\sigma_a^2$  が大きくなるにつれて TR と EB との差は小さくなっていく。

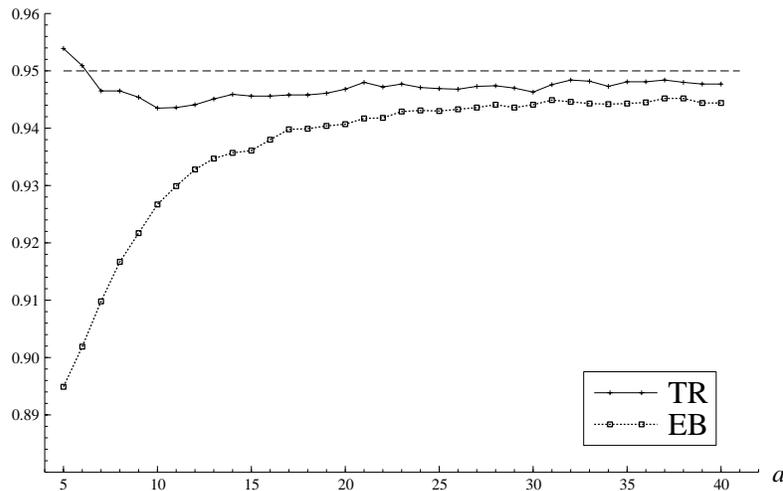


図 1: 経験ベイズ信頼区間 (EB) と補正後の信頼区間 (TR) の被覆確率 ( $\sigma_a^2 = 0.5, \sigma_e^2 = 1$  の場合で、 $a$  は 5 から 40 の範囲)

次に、漸近補正を行った信頼区間  $I_{TR}$  と通常の  $t$ -分布に基づいた信頼区間の長さを比較してみよう。標本平均  $\bar{y}_1$  に基づいた通常の信頼区間は、 $\bar{y}_1 | \theta_1 \sim \mathcal{N}(\theta_1, \sigma_e^2/n_1)$  より、

$$I_0 = \left[ \bar{y}_1 - t_{\alpha/2} \sqrt{\hat{\sigma}_e^2/n_1}, \bar{y}_1 + t_{\alpha/2} \sqrt{\hat{\sigma}_e^2/n_1} \right] \quad (5.1)$$

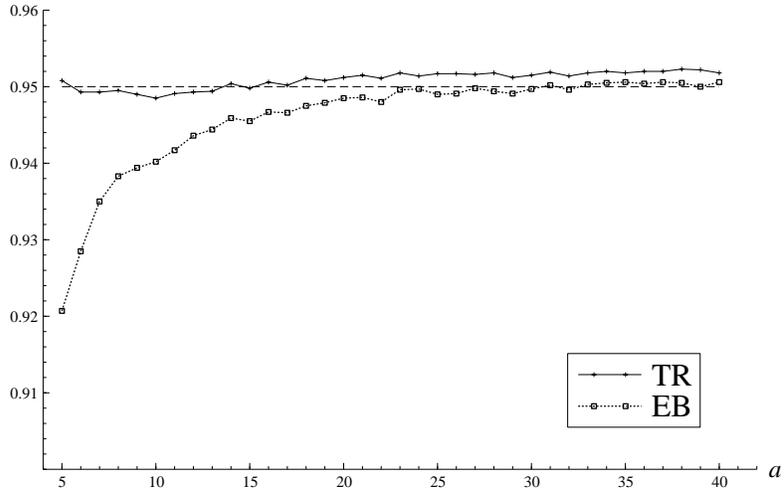


図 2: 経験ベイズ信頼区間 (EB) と補正後の信頼区間 (TR) の被覆確率 ( $\sigma_a^2 = 1, \sigma_e^2 = 1$  の場合で,  $a$  は 5 から 40 の範囲)

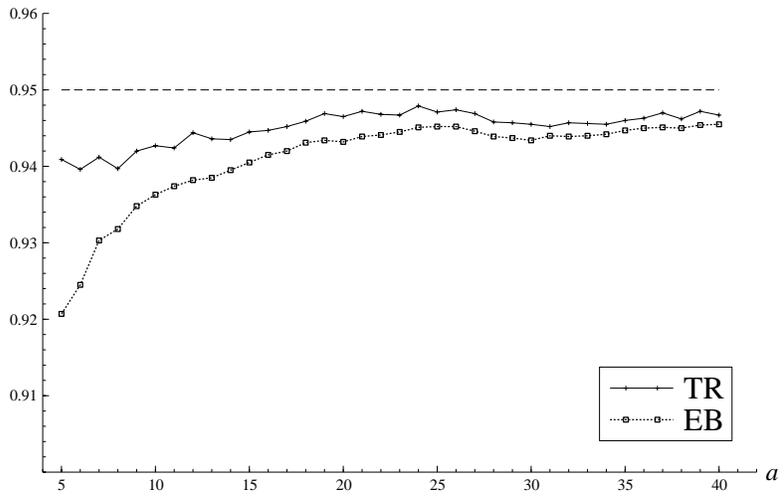


図 3: 経験ベイズ信頼区間 (EB) と補正後の信頼区間 (TR) の被覆確率 ( $\sigma_a^2 = 2, \sigma_e^2 = 1$  の場合で,  $a$  は 5 から 40 の範囲)

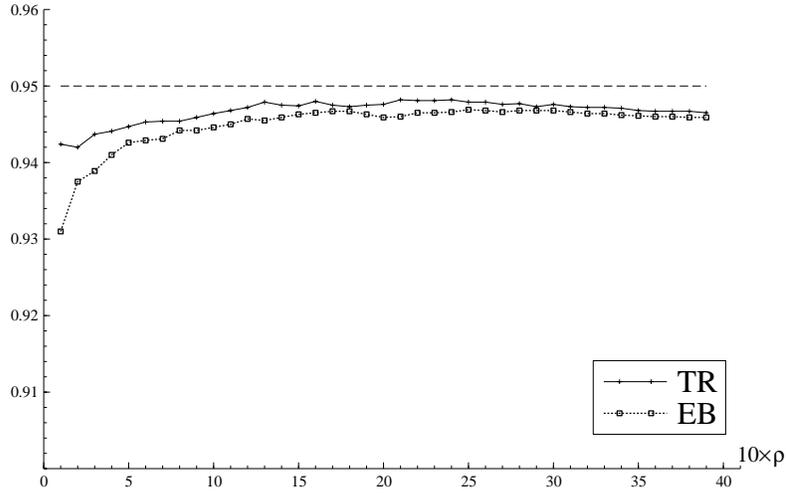


図 4: 経験ベイズ信頼区間 (EB) と補正後の信頼区間 (TR) の被覆確率  
 ( $a = 30, \sigma_e^2 = 1$  の場合で,  $\sigma_a^2 = \rho$  は 0.1 から 4.0 の範囲)

で与えられる。ただし,  $t_{\alpha/2}$  は自由度  $N - a$  の  $t$ -分布の上側  $\alpha/2$  点を表している。信頼区間  $I_{TR}$  は,  $I_0$  に比べてはたして信頼区間の幅は短くなっているだろうか。図 5 は, 信頼区間の長さの比  $|I_{TR}|/|I_0|$  を, 横軸に個数  $a$  の値をとって描いたもので,  $\sigma_a^2 = \rho$  が 0.5, 1.0, 2.0 の 3 通りの場合について示されている。一般に  $a$  が小さくないときには,  $I_{TR}$  の長さは,  $I_0$  より短くなる傾向にあることがわかる。特に  $\sigma_a^2 = 0.5$  の場合には,  $a \leq 13$  のとき  $I_{TR}$  の方が長くなるものの,  $a \geq 14$  に対してはかなり短くなっている。 $\sigma_a^2 = 0.5$  のときに, 信頼区間の長さを短くしてしかも信頼係数を 95% に保つためには  $a \geq 14$  である必要があることがわかる。

## 5.2 応用例

さて, 具体的な小地域についてのデータを用いて, 提案された補正済み経験ベイズ信頼区間の有用性について検討してみよう。ここで用いるデータは, 神奈川県における京浜急行線沿いの宅地物件について 2001 年に公表された  $1m^2$  当たりの地価公示価格である。各駅を 1 つの小地域と考え, また  $i$  番目の駅を最寄り駅とする物件のデータをその小地域からとられたデータと考えて, その個数を  $n_i$  で表す。小地域の総数は  $a = 66$  であり,  $n_i$  は 1 から 19 まで不均一な値をとっているが平均 4 程度である。各地価公示価格を対数変換したものを  $y_{ij}$  とし, (4.1) に対応するモデル

$$y_{ij} = \beta_0 + x_{1i}\beta_1 + x_{2ij}\beta_2 + \alpha_i + \varepsilon_{ij}$$

を想定してみる。ここで, 共変量  $x_{1i}$  は  $i$  番目の駅から東京駅に到着するのに要する時間,  $x_{2ij}$  は物件 ( $ij$ ) から最寄り駅 ( $i$ ) までの距離を表している。 $\mathbf{x}_{ij} = (1, x_{1i}, x_{2ij})^t$ ,  $\bar{\mathbf{x}}_i = (1, x_{1i}, \bar{x}_{2i})^t$  とおくと, 行列  $\sum_{i=1}^a \sum_{j=1}^{n_i} (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)^t$  のランクは 1 となるので, (4.4) の  $\hat{\sigma}_e^2$  における  $\lambda$  は  $\lambda = 2$  となることに注意する。このモデルにおいて, 駅 (小地域) ごとに平均的な地価公示価格

$$\theta_i = \beta_0 + x_{1i}\beta_1 + \bar{x}_{2i}\beta_2 + \alpha_i$$

の 95% 信頼区間を構成してみよう。

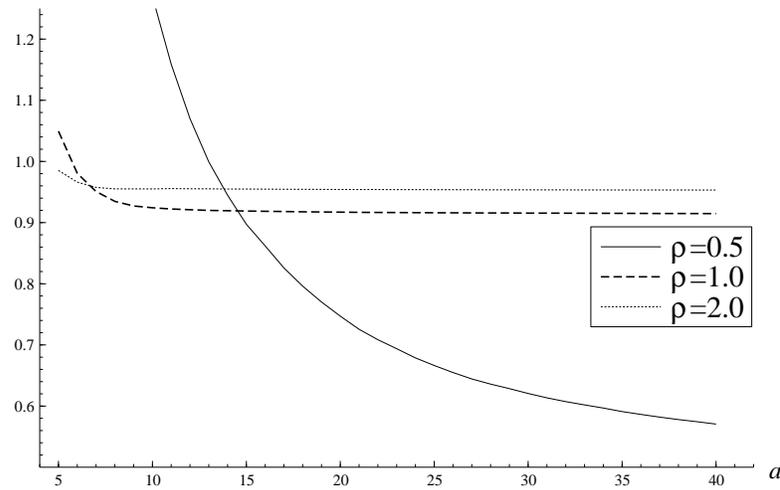


図 5: 補正後の信頼区間 (TR) の幅と  $t$ -分布に基づく信頼区間の長さの比較  
 $(\sigma_e^2 = 1$  で,  $\sigma_a^2 = \rho$  は 0.5, 1.0, 2.0 の場合で,  $a$  は 5 から 40 の範囲)

表 1: 京浜急行線沿線の物件の  $1m^2$  当たりの駅ごとの平均価格  
 $(n_i$  は  $i$  番目の駅を最寄り駅とする物件の個数を表し, 2つの信頼区間の左端, 右端, 中央の値及び信頼区間の長さが以下で与えられる。)

No.	最寄り駅	$n_i$	補正済み経験ベイズ信頼区間 (TR)				$t$ -分布に基づいた信頼区間 (t-CI)			
			左端	右端	中央	長さ	左端	右端	中央	長さ
1	神奈川	1	258151	404497	331324	146346	149560	451440	300500	301880
2	戸部	3	295788	416232	356010	120444	278360	452640	365500	174280
3	日ノ出町	4	275984	376208	326096	100224	252130	403070	327600	150940
4	黄金町	3	270222	380217	325220	109995	249110	423390	336250	174280
5	南太田	2	283480	416607	350044	133126	275940	489400	382670	213460
6	屏風浦	2	239115	351371	295243	112256	193600	407060	300330	213460
7	杉田	2	210259	308970	259615	98710	132270	345730	239000	213460
8	京急富岡	8	207378	262708	235043	55330	174750	281480	228115	106730
9	能見台	11	221305	272011	246658	50705	202240	293260	247750	91020
10	金沢文庫	6	225550	294283	259917	68733	199520	322760	261140	123240
11	追浜	7	181425	232931	207178	51506	140830	254920	197875	114090
12	京急田浦	3	163570	230153	196861	66582	92860	267140	180000	174280
13	安針塚	3	153921	216598	185260	62676	79610	253890	166750	174280
14	逸見	4	156185	212926	184555	56740	102930	253870	178400	150940
15	汐入	2	167644	246372	207008	78728	91270	304730	198000	213460
16	横須賀中央	3	193222	271875	232549	78652	171860	346140	259000	174280
17	京急安浦	6	198054	258409	228232	60354	188240	311480	249860	123240
18	京急大津	3	172784	243141	207963	70357	121110	295390	208250	174280
19	馬堀海岸	2	172542	253596	213069	81053	113600	327060	220330	213460
20	浦賀	10	154894	192047	173471	37153	124000	219460	171730	95460

(4.8) で与えられた  $\theta_i$  の補正済み経験ベイズ信頼区間は,

$$\bar{x}_i^t \hat{\beta}_2 + \frac{n_i \hat{\rho}^{TR}}{1 + n_i \hat{\rho}^{TR}} (\bar{y}_i - \bar{x}_i^t \hat{\beta}_2) \pm z_{0.025} \sqrt{\frac{\hat{\sigma}_e^2 \hat{\rho}^{TR}}{1 + n_i \hat{\rho}^{TR}}} \left(1 + \frac{1}{2a} h^*(\hat{\rho}^{TR})\right)$$

であり,これを TR で表すことにする。この信頼区間の有用性を検討するために,標本平均  $\bar{y}_i$  と  $t$ -分布に基づいた通常の信頼区間

$$\bar{y}_i \pm t_{0.025} \sqrt{\hat{\sigma}_e^2/n_i}$$

の結果も与えることにする。これを  $t$ -CI で表すことにする。この節で用いた例では,  $N = 262$ ,  $m = 3$ ,  $\lambda = 1$ ,  $a = 66$  であり,推定値は  $\hat{\beta}_2 = (13.2825, -0.0146, -0.0001)^t$ ,  $\hat{\sigma}_e^2 = 0.0345$ ,  $\hat{\sigma}_a^2 = 0.0181$  となった。 $\beta_1$  が負の値であることから,東京駅から遠くなるにつれて地価公示価格は低くなる傾向にあり,合理的な符号を示していることがわかる。分散比の推定値は  $\hat{\rho} = \hat{\sigma}_a^2/\hat{\sigma}_e^2 = 0.5246$  で与えられる。

表 1 は,神奈川駅から浦賀駅までの 20 駅について各駅のデータ数  $n_i$  と  $\theta_i$  に対する信頼区間 TR,  $t$ -CI の左端,右端の点,中央の点及び信頼区間の長さを与えている。ここで,  $y_{ij}$  の値は地価公示価格を対数変換しているので,  $\theta_i$  の信頼区間を求めた後に指数変換を行って元のスケールに戻した値が,表 1 で与えられる信頼区間の両端の値であることに注意しておく。

表 1 の計算結果をグラフを用いて説明したのが図 6, 7, 8 である。図 6 は, TR と  $t$ -CI の信頼区間の両端の点を,駅 No.1 から No.20 までを横軸にとってプロットし直線で結んだものである。TR の方が  $t$ -CI よりも信頼区間の幅が狭くなるだけでなく,より安定した区間を与えていることがわかる。図 7 は,信頼区間の長さとは各最寄り駅の小地域に属する標本の数  $n_i$  との関係を示したものである。 $t$ -CI の長さは,  $n_i$  が大きければ小さくなるものの,  $n_i$  が小さいときにはかなり大きくなっている。これに対して, TR の長さは,基本的には同じ傾向性があるものの,  $n_i$  が小さくても  $t$ -CI よりもはるかに安定した信頼区間の長さを与えている。図 8 は, 2 つの信頼区間について,それぞれの区間の中点の値と  $n_i$  との関係を示したものである。区間の中点は点推定値として用いられるが,経験ベイズ推定値 (TR の中点) は,  $n_i$  が大きいときには標本平均 ( $t$ -CI の中点) に近い与える一方,  $n_i$  が小さいときには  $\bar{x}_i^t \hat{\beta}_2$  の方向へ縮小することによって安定化させていることが伺える。これは経験ベイズ手法のもっている一般的な性質であり,この例においても同様な性質が確認されたことになる。

## 6 補遺

ここでは,(2.5) もしくは (4.5) で与えられた統計量  $S_2$  について,その分布に関する表現式の説明と注意を与える。まず,スカラーまたは行列  $A_1, \dots, A_a$  に対して記号  $\{dA_i\}$  を  $\{dA_i\} = \text{diag}(A_1, \dots, A_a)$  で定義すると,簡単な計算から (4.5) の  $S_2$  は,

$$\begin{aligned} S_2 &= \sum_{i=1}^a n_i (\bar{y}_i - \bar{x}_i^t \hat{\beta})^2 \\ &= (\bar{y}_1, \dots, \bar{y}_a) \{d\sqrt{n_i}\} \mathbf{K} \{d\sqrt{n_i}\} (\bar{y}_1, \dots, \bar{y}_a)^t \end{aligned}$$

と書き直すことができる。ここで,  $\mathbf{K}$  は,

$$\mathbf{K} = \mathbf{I}_a - (\sqrt{n_1} \bar{x}_1, \dots, \sqrt{n_a} \bar{x}_a)^t \left\{ \sum_{i=1}^a n_i \bar{x}_i \bar{x}_i^t \right\}^{-1} (\sqrt{n_1} \bar{x}_1, \dots, \sqrt{n_a} \bar{x}_a)$$

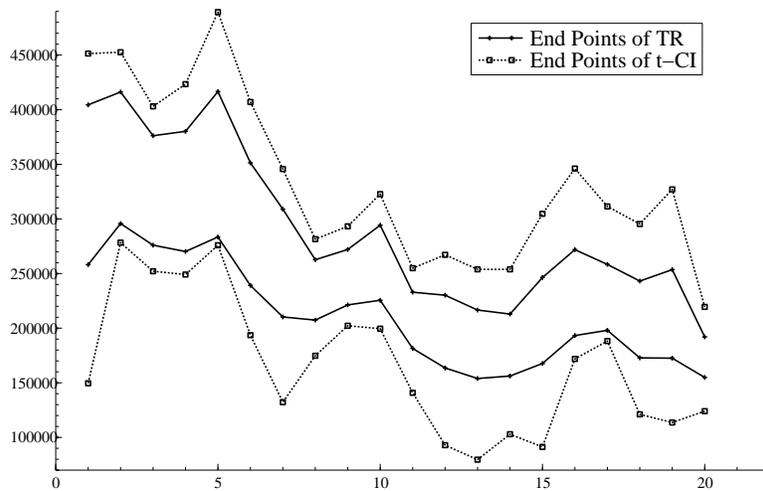


図 6: 補正後の信頼区間 (TR) と  $t$ -分布に基づく信頼区間 (t-CI) の両端の値の比較 (No.1 から No.20 までの各駅を横軸に並べている。)

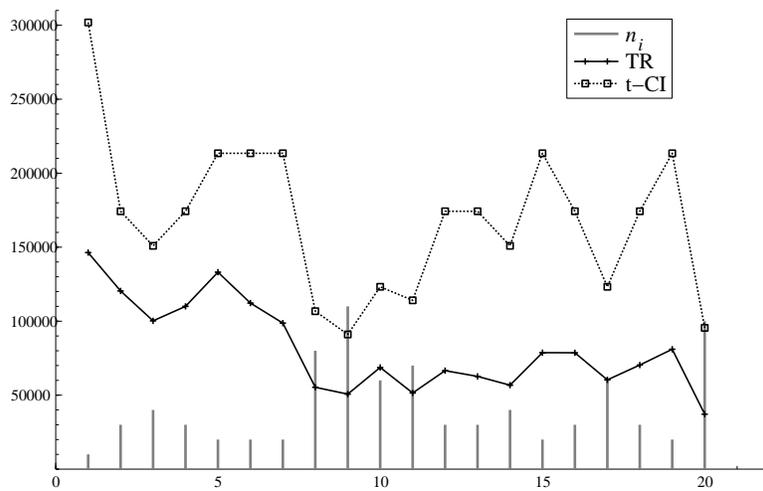


図 7: 補正後の信頼区間 (TR) と  $t$ -分布に基づく信頼区間 (t-CI) の長さと  $n_i$  との関係 ( $n_i$  の値については、縦軸の 1 メモリ (10000 に相当) が  $n_i = 1$  に対応する。)

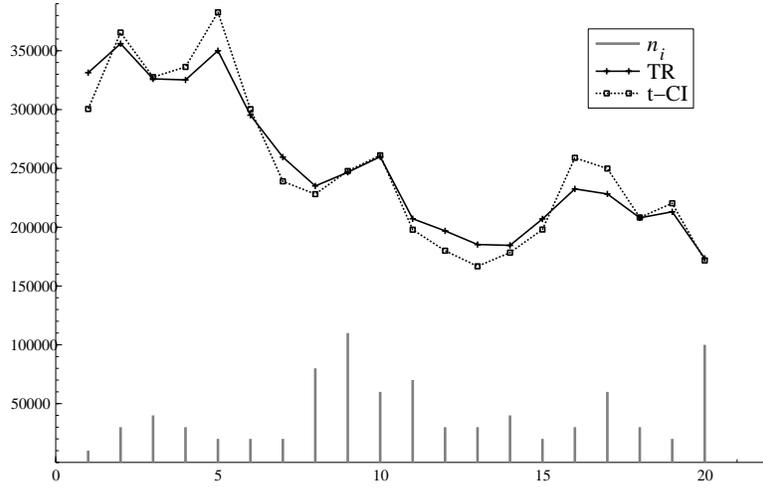


図 8: 補正後の信頼区間 (TR) と  $t$ -分布に基づく信頼区間 (t-CI) の中央値と  $n_i$  との関係 ( $n_i$  の値については, 縦軸の 1 メモリ (10000 に相当) が  $n_i = 1$  に対応する。)

で定義される巾等行列で, そのランクは  $a - m$  である。ここで  $u_i = \sum_{j=1}^{n_i} \varepsilon_{ij} / \sqrt{n_i}$  に対して  $\mathbf{u} = (u_1, \dots, u_a)^t$  とおき,  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_a)^t$  とおくと,

$$\{d\sqrt{n_i}\}(\bar{y}_1, \dots, \bar{y}_a)^t = \mathbf{u} + \{d\sqrt{n_i}\}\boldsymbol{\alpha}$$

と表される。また  $K$  が巾等であることから,  $a \times a$  直交行列  $P$  が存在して

$$K = P^t \begin{pmatrix} I_{a-m} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} P = P_1^t P_1$$

と書ける。ただし  $P^t = (P_1^t, P_2^t)$  であり,  $P_1$  は  $(a - m) \times a$  の行列である。 $\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \sigma_e^2 I_a)$ ,  $\boldsymbol{\alpha} \sim \mathcal{N}(\mathbf{0}, \sigma_a^2 I_a)$  より,  $P_1(\mathbf{u} + \{d\sqrt{n_i}\}\boldsymbol{\alpha})$  の分布は,

$$P_1(\mathbf{u} + \{d\sqrt{n_i}\}\boldsymbol{\alpha}) \sim \mathcal{N}_{a-m}(\mathbf{0}, \sigma_e^2 I_{a-m} + \sigma_a^2 P_1 \{d n_i\} P_1^t)$$

に従うことがわかる。ここで, 互いに直交するランク 1 の巾等行列  $E_1, \dots, E_{a-m}$  を用いて

$$P_1 \{d n_i\} P_1^t = \sum_{i=1}^{a-m} \gamma_i E_i, \quad \sum_{i=1}^{a-m} E_i = I_{a-m}$$

と書けるので, 共分散行列の部分は  $\sigma_e^2 I_{a-m} + \sigma_a^2 P_1 \{d n_i\} P_1^t = \sum_{i=1}^{a-m} (\sigma_e^2 + \gamma_i \sigma_a^2) E_i$  となる。従って,

$$P_1(\mathbf{u} + \{d\sqrt{n_i}\}\boldsymbol{\alpha}) = \{\sigma_e^2 I_{a-m} + \sigma_a^2 P_1 \{d n_i\} P_1^t\}^{1/2} \mathbf{z}$$

とおくと,

$$\begin{aligned} S_2 &= \mathbf{z}^t \{\sigma_e^2 I_{a-m} + \sigma_a^2 P_1 \{d n_i\} P_1^t\} \mathbf{z} \\ &= \sum_{i=1}^{a-m} (\sigma_e^2 + \gamma_i \sigma_a^2) \mathbf{z}^t E_i \mathbf{z} \end{aligned}$$

と表される。  $z \sim \mathcal{N}(\mathbf{0}, I_a)$  に従うことから,  $z^t E_1 z, \dots, z^t E_{a-m} z$  は互いに独立に  $\chi_1^2$  分布に従うことがわかる。

最後に, 固有値  $\gamma_i$  の最大値が  $\max_i n_i$  で抑えられることを示す。  $P^t = (P_1^t, P_2^t)$  は直交行列であるから,  $P\{dn_i\}P^t$  の固有値は  $n_1, \dots, n_a$  で, 最大固有値は  $\max_i n_i$  である。また

$$P\{dn_i\}P^t = \begin{pmatrix} P_1\{dn_i\}P_1^t & P_1\{dn_i\}P_2^t \\ P_2\{dn_i\}P_1^t & P_2\{dn_i\}P_2^t \end{pmatrix}$$

において,  $\gamma_1, \dots, \gamma_{a-m}$  は部分行列  $P_1\{dn_i\}P_1^t$  の固有値となっていることに注意する。従って,  $\max_i n_i \geq \max_j \gamma_j$  を示すためには, 部分行列の最大固有値が元の行列の最大固有値以下になることを示せばよい。実は, このことは Rao-Rao (1998) で与えられた次の命題から従うことがわかる。

命題 6.1 一般に,  $m \times m$  対称行列  $A$  が, 正方行列  $B, D$  を用いて

$$A = \begin{pmatrix} B & C \\ C^t & D \end{pmatrix}$$

と分割されているとする。このとき,  $A$  の最大固有値は  $B$  の最大固有値以上になる。

証明は以下のようにして簡単になされる。  $v_1$  を  $B$  の最大固有値  $\beta_1$  に対応する基準化された固有ベクトルとし,  $m \times 1$  ベクトル  $w_1$  を  $w_1 = (v_1^t, \mathbf{0}^t)^t$  で定義すると,  $w_1^t A w_1 = v_1^t B v_1 = \beta_1$  をみたすことがわかる。一方,  $w_1$  は  $A$  の最大固有値  $\alpha_1$  に対応する固有ベクトルとは限らないので,  $\alpha_1 \geq w_1^t A w_1 = v_1^t B v_1 = \beta_1$  となり, 命題が確かめられる。

## 参考文献

- [1] Basu, R., Ghosh, J.K., and Mukerjee, R. (2003). Empirical Bayes prediction intervals in a normal regression model: higher order asymptotics. *Statist. Prob. Letters*, **63**, 197-203.
- [2] Datta, G. S., Ghosh, M., Smith, D.D. and Lahiri, P. (2002). On an asymptotic theory of conditional and unconditional coverage probabilities of empirical Bayes confidence Intervals. *Scandinavian J. Statist.*, **29**, 139-152.
- [3] Datta, G.S. and Lahiri, P. (2000). A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. *Statist. Sinica*, **10**, 613-627.
- [4] Casella, G. and Hwang, J.T. (1983). Empirical Bayes confidence sets for the mean of a multivariate normal distribution. *J. Amer. Statist. Assoc.*, **78**, 688-698.
- [5] Efron, B. and Morris, C. (1975). Data analysis using Stein's estimator and its generalizations. *J. Amer. Statist. Assoc.*, **70**, 311-319.
- [6] Fay, R.E. and Herriot, R. (1979). Estimates of income for small places: An application of James-Stein procedures to census data. *J. Amer. Statist. Assoc.*, **74**, 269-277.

- [7] Kubokawa, T., Saleh, A.K.Md.E., and Konno, Y. (2000). Bayes, minimax and nonnegative estimators of variance components under Kullback-Leibler loss. *J. Statist. Plan. Inf.*, **86**, 201-214.
- [8] Kubokawa, T. (2000). Estimation of variance and covariance components in elliptically contoured distributions. *J. Japan Statist. Soc.*, **30**, 143-176.
- [9] Morris, C.N. (1983). Parametric empirical Bayes inference: Theory and application. *J. Amer. Statist. Assoc.*, **78**, 47-65.
- [10] Prasad, N.G.N. and Rao, J.N.K. (1990). The estimation of the mean squared error of small-area estimators. *J. Amer. Statist. Assoc.*, **85**, 163-171.
- [11] Rao, C.R. and Rao, M.B. (1998). *Matrix Algebra and Its Applications to Statistics and Econometrics*, World Scientific.
- [12] Searle, S.R., Casella, G., and McCulloch, C.E. (1992). *Variance Components*, Wiley, New York.