経済研究
Vol. 52, No. 3, Jul. 2001

# 小地域統計の推定手法と応用

美 添 泰 人

本稿で取り上げる小地域統計の推定とは，全体としての標本の大きさは十分であっても，個々の小地域ごとに考えると標本の大きさは小さくなるような状況で，各地域ごとの推論にどのような方法を適用すべきかという問題であり，市区町村の死亡率が典型的な例である．この問題に対する有効な方法として，地域間の類似性に関する事前情報を活用することが考えられる．実際にも小地域の死亡数を二項分布とし，その母数に事前分布を想定するという単純なベイズの手法や，さらに複雑な経験的ベイズ推定法を用いる手法などが提案されている．これに加えて，小地域の情報以外でも隣接する年齢階級の死亡率が類似しているという情報を利用すれば，さらに精度の高い推定を実現することが可能である．ここでは，そのような手法の例として，できるだけ主観を排除するようなベイジアンの手法を提案し，小地域における死亡率推定への応用を検討する．

## 1. 問題の所在

小地域統計における推定問題としてここで取り上げるものは，次のような問題である．すなわち，全体としての標本の大きさは十分であっても，個々の小地域ごとに考えると標本の大きさは小さくなる．そのような状況で，各地域ごとに何らかの推論が要求される場合に，どのような方法を適用すべきか，ということである．具体的な例として，以下のものは，密接な関連がある．

**例：市区町村の死亡率**　$M_i$ を第 $i$ 年齢階層の基準人口，$Q_i$ を基準人口における死亡率とし，さらに指標作成対象地域の第 $i$ 年齢階層について，$m_i$ を人口，$d_i$ を観測された死亡数，$q_i=d_i/m_i$ を死亡率とするとき，訂正死亡率（または年齢調整死亡率，DAR：Directly age-Adjusted mortality Rate）および標準化死亡比（SMR：Standardized Mortality Ratio）は次式で定義される．

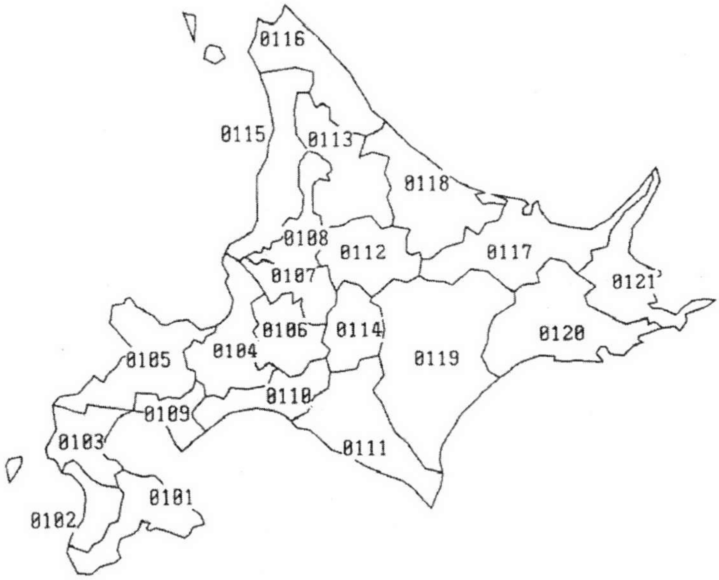
$$\mathrm{DAR}=\frac{\sum M_i q_i}{\sum M_i},\quad \mathrm{SMR}=\frac{\sum d_i}{m_i Q_i}$$

市部，郡部さらに市区町村などの小地域においては，毎年発生する死亡数 $d$ の変動が大きく，実際にも死亡率（SMR など）を求めると，極めて不安定となることが知られている．この問題に対するひとつの試みとして，厚生省の研究の一環として作成されている市町村別死亡率では，ベイズの手法を利用する工夫がなされている．具体的に厚生省大臣官房統計調査部（1990）で利用された手法は次のようなものである．

いま対象となる地域の性・年齢階級別死亡数は互いに独立に2項分布 $d_i \sim B(m_i, \theta_i)$ に従うと仮定し，その母数である死亡率 $\theta$ に事前分布を導入する．事前分布はベータ分布 $\theta_i \sim \beta(\alpha_i, \beta_i)$ と想定し，超母数を定めるために，当該市町村を含む，より大きな地域（医療圏）から得られる観測値を用いて定めている．医療圏については北海道の例を図1に示す．このようなモデルから，各地域の $\theta$ の事後分布が求められる．SMR などの評価にあたっては観察された死亡数 $d_i$ の代りに事後分布の期待値から計算された $m_i E(\theta_i \mid d_i)$ を利用している．

容易に想像できるように，この方法によると事前分布から得られる情報が相対的に大きくなる小地域に関しては安定的な推定が可能となっている．例えば図2では通常の方法（Non

図1. 北海道の医療圏




Bayesと記されているもの)とベイジアンの方法を対比したものである．横軸に人口規模，縦軸にSMRが記されている．この図から明らかなように，人口規模の小さな市区町村においてNon Bayesに見られる極端な数値はBayesでは解消されている．全く同じ手順で，医療圏の情報の代りに市・郡の情報を事前分布として利用して市区町村の死亡率を推定したり，県全域の情報を事前分布として利用して広域医療圏の死亡率を推定することが可能である．

地域間の類似性に関する事前情報を活用する方法は，上記の方法以外にも経験的ベイズ推定法による丹後(1988)などがあるが，さらに小地域の情報以外でも隣接する年齢階級の死亡率が類似しているという情報を利用する手法を導入することは，それほど難しいことではない．

以下では，この目的のために自然な手法と考えられる回帰モデルに関するベイジアンの手法を紹介する．なお，一般的なベイジアンの手法の評価としては，たとえば美添(1983)を参照されたい．

## 2. 回帰モデルにおける適応的手法と問題点

回帰モデルにおいて，応答関数に線形性を仮定しない，最も簡単と思われるモデルは

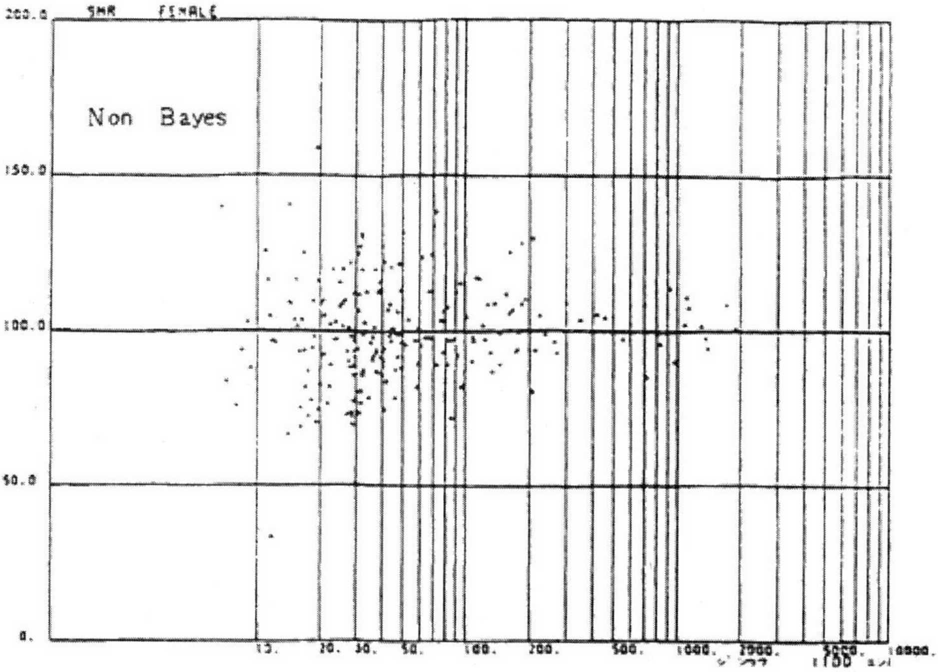
$$E(y\,|\,x_1,\cdots,x_p)=F(\beta_0+\beta_1x_1+\cdots+\beta_px_p),$$
$$\mathrm{var}(y\,|\,x_1,\cdots,x_p)=\sigma^2 \tag{1}$$

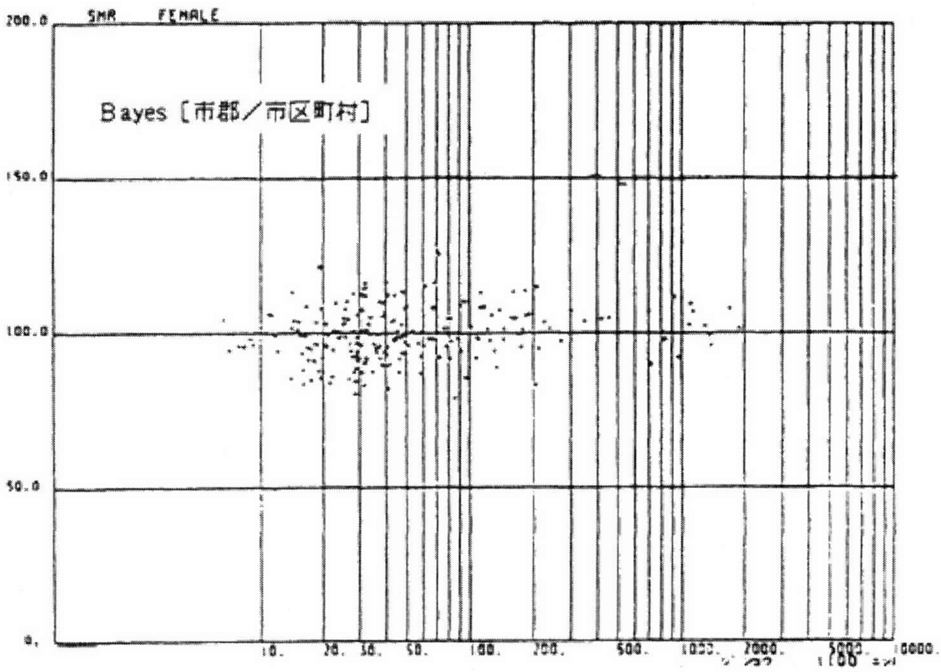
で与えられる．$F$は任意の関数とする．(1)式において応答関数$F$が未知のとき，とくにその非線形性が疑われるときには，$F$を推定するための手法としてカーネル法，スプライン関数による方法の他，ベイジアンによるいくつかの方法がこれまでに提案されているなかで，最も優れていると思われるものはO'Hagan(1978)である．彼のモデルは，多少記法を変えると次のように表される．

$$E(y\,|\,x)=\eta(x)=\beta(x)^Tx,\quad \mathrm{var}(y\,|\,x)=\sigma^2$$

ここで母数の$p$次元ベクトル$\beta=\beta(x)$は説明変数$x$の空間で定義された関数であり，$\beta$が一定でない可能性を明示的に認めている．O'Haganは，$\beta$に次のような正規分布，より正確にはGauss過程を想定した．

図 2.　市町村別の推計値と SMR（北海道・女）
上記研究報告書 p. 15



$$E[\beta(x)] = b,$$
$$\mathrm{cov}[\beta(x), \beta(x')] = \rho(|x-x'|)D \quad (2)$$

ただし $b$ は既知の $p$ 次元ベクトル，$D$ は既知の $p\times p$ 正値定符号行列，また $\rho=\rho(d)$ は非負の $d$ に対して定義された関数であり $\rho(0)=1$ を満たす．ここで共分散関数が正当なものであるための必要十分条件は，$\rho=\rho(x-x')$ と表すとき $\rho(\cdot)$ が $p$ 変量分布の特性関数となることである．なお O'Hagan の論文ではこの点が不正確で，1 変量の議論をそのまま用いている．幸い例で用いたのが正規分布だったため，結果は正しいものが得られている．

$\beta$ あるいは $\eta$ の事後分布は正規分布となり，その母数も簡単に求められる．このモデルが優れているのは，応答関数 $\eta(x)$ が滑らかに変化するという以外に制約が少ない点である．特に(2)式によって $x$ と $x'$ とが近ければ $\eta(x)$ も $\eta(x')$ に近いということが保証される．また，多変数回帰モデルにも容易に適用することができる．

しかし，このモデルには大きな欠陥がある．それは，極限の場合として $\beta(x)=const.$ の場合，すなわち通常の線形回帰モデルに帰着させようとする場合，このモデルで diffuse prior を採用して得られる結果は通常の最小 2 乗法から得られる結果とは異なるということである．O'Hagan は極限の場合として最小 2 乗法が得られるとしているが，実は本質的に異なった結果しか得ることはできない．その意味は，正確には次のようなことである．簡単のために $p=1$(定数項のないモデル)とし，観測値$(x_i, y_i)$，$(i=1, \cdots, n)$が得られたものとしよう．誤差項の分布を $N(0, \sigma^2 I_n)$ とし，$\beta$ の事前分布を(2)式で $b=0$，$D=\tau^2$，$\rho(d)=\exp(-\xi d^2/2)$ とした Gauss 過程とする．$\beta$ が一定であるという条件は $\xi\to 0$ で表現される．観測値の $n$ 次元ベクトルを $x$，$y$ とするとき，観測値の一つと一致する $x_0$ に対応する $\beta$ について，その事後分布における平均は

$$E[\beta(x_0)\,|\,y]\to(x^Tx+\sigma^2\sigma^{-2}\tau^{-2})^{-1}x^Ty \quad (\xi\to 0)$$

分散は

$$\mathrm{var}[\beta(x_0)\,|\,y]\to(\tau^{-2}+\sigma^{-2}x^Tx)^{-1} \quad (\xi\to 0)$$

となる．$\beta$ に関する diffuse prior では $\tau\to\infty$ となるが，このとき観測値と一致する $x_0$ に対応する $\beta$ では，事後分布の平均，分散ともに通常の回帰分析による結果と一致することは容易に確かめられる．これが O'Hagan の主張である．

一方，観測値と一致しない $x_0$ に対応する $\beta$ については，若干の計算を実行することにより次のように事後分布の平均，分散が導かれる．

$$E[\beta(x_0)\,|\,y] = \tau^2\rho' D_x(\sigma^2 I_n+\tau^2 D_x R D_x)^{-1}y$$
$$\mathrm{var}[\beta(x_0)\,|\,y] = \tau^2[1-\tau^2\rho' D_x(\sigma^2 I_n+\tau^2 D_x R D_x)^{-1}D_x\rho] \quad (3)$$

ただし $R$ は $\{\beta(x_1), \cdots, \beta(x_n)\}$ の相関係数行列，$\rho$ は $\beta(x_0)$ と $\{\beta(x_1), \cdots, \beta(x_n)\}$ の相関係数ベクトル，$D_x=\mathrm{diag}(x_1, \cdots, x_n)$ である．これから

$$\mathrm{var}[\beta(x_0)\,|\,y] \to\tau^2[1-\tau^2x^T(\sigma^2 I_n+\tau^2 xx^T)^{-1}x] \quad (\xi\to 0)$$

となり，ここで $\tau\to\infty$ とすればやはり通常の最小 2 乗法による結論と一致することを示すのはそれほど難しくない．しかし，事後分布の平均については

$$E[\beta(x_0)\,|\,y]\to\tau^2 x^T(\sigma^2 I_n+\tau^2 xx^T)^{-1}y \quad (\xi\to 0)$$

となるから，

$$x^T(\tau^{-2}\sigma^2 I_n+xx^T)^{-1}y\to w^T(xx^T)^- y \quad (\tau\to\infty)$$

と一般化逆行列を含む形でしか表現できない．このことは，実際例の評価においては数値的に

不安定となること，すなわち観測値に近い $x$ についてさえ $\beta(x)$ や $\hat{y}=\beta(x)^T x$ が推定できないということを意味する．

さらに極限をとる順序を変えて先に $\tau\to\infty$ とすることを考えると，(3)式の分散において $\tau^2$ の係数であるかっこの中は $1-\rho' R^{-1}\rho$ となり，これは一般に正であるから分散は無限に大きくなってしまう．なお平均については，先と同じ一般化逆行列の形が導かれる．

以上の結果は，O'Hagan の手法の限界を示したものといえる．すなわち類似の $x$ に対応する $\beta$ は大きく違わないという情報以外には個々の $\beta$ に関する情報がほとんどない場合に不適切な結果を導くため，通常の線形回帰モデルを含んだ一般的なモデルを表現できていないことになる．

## 3. 新しい手法の導入

この節では美添(1987)で紹介した手法を適用し，滑らかな応答関数を表現する事前分布を導く．その内容は前節で紹介した O'Hagan の手法を拡張して難点を取り除いたものと位置づけることができる．死亡指標への応用を念頭におき，人口 $m$ 人の特定地域・年齢階級集団の死亡数 $y$ が2項分布 $y\sim B(m,\pi)$ となるモデルにおいて母数 $\pi$ が

$$\pi(x) = F(\beta(x)^T x) \tag{4}$$

で与えられるものとする．ここで $\beta(x)$ と $x$ は $p$ 次元ベクトルであり，$F$ はロジスティック分布の分布関数 $F(x)=1/(1+e^{-x})$ を表す．$F$ としてロジスティック分布を選ぶのは，この $F$ はある意味で最も簡単なモデルを与えることによる．

### 3.1 手法の概要

滑らかな応答関数 $\pi(x)$ を導くために，$\beta(x)$ は局所的に変動が小さいことを要求する．そのために次のような確率的 Lipschitz 条件を想定する．

$$\begin{aligned} &Pr\{\|\beta(x)-\beta(x')\|<K\|x-x'\|\} \\ &\quad =1-\varepsilon \quad (\forall x,x') \end{aligned} \tag{5}$$

ただし，$\|x-x'\|=d(x,x')$ は $p$ 次元空間において適当に定義された距離である．(5)式のために，以下の条件を導入する．

(A) $\mathrm{var}[\beta(x)-\beta(x')]$ は $d(x,x')^2$ に比例する

(B) $\mathrm{var}[\beta(x)]=\xi^{-1}\Omega$（のちに $\xi\to 0$ とする）

(C) $\beta(x)$ は正規分布に従う

条件(A)は(5)式のための十分条件であり，$\mathrm{var}[\beta(x)-\beta(x')]$ は有限であることを要求する．一方，(B)は単独の $\beta(x)$ に関しては情報がなく diffuse prior を適用することを表している．(C)は簡単化のための条件である．以上から $\beta(x)$ は Gauss 過程に従い，$\rho$ を $\rho(x,x)=1$ となるような適当な関数として，次の形の母数をもつものと想定できる．なお，$E[\beta(x)]=0$ としているのは $\xi\to 0$ とするため，平均は任意とできることによる．

$$\begin{aligned} &E[\beta(x)]=0, \\ &\mathrm{cov}[\beta(x),\beta(x')]=\xi^{-1}\rho(x,x')\Omega \end{aligned} \tag{6}$$

ここで $\Omega$ は $p\times p$ の正値定符号行列である．$\beta$ の成分間の関係について十分な知識がなければ，$\Omega$ を $p$ 次元単位行列とおくのが簡単である．以下，Yoshizoe(1986)から，主要な定理をまとめておく．

**定理1** 任意の $n$ と $(x_1,\cdots,x_n)$ に対して $\beta(x_1),\cdots,\beta(x_n)$ の共分散行列 $(np\times np)$ が正値定符号となるための必要十分条件は

$$\rho(x,x')=\rho(x-x')$$

と表され，それがある $p$ 変数確率分布の特性関数となることである．

この定理を多変量正規分布の特性関数に対して用いると

$$\rho(x, x') = \exp[-\xi a(x, x')] \quad (7)$$

と書き換えられる．ここで

$$a(x, x') = (\lambda/2)(x-x')^T \Omega^{-1}(x-x') \quad (8)$$

である．$\lambda$ と $\xi$ が正であれば，$\rho$ は定理の条件を満たす．もうひとつの母数 $\lambda$ は Lipschitz 条件の強さを与えるものである．

(6)，(7)式が実際に(5)式の条件を満たしていることは，簡単に確かめられる．すなわち $d(x, x') = [(x-x')^T \Omega^{-1}(x-x')]^{1/2}$ を Mahalanobis 距離とすれば，任意の $K>0$ に対して

$$\lim_{\xi\to 0} Pr\{\|\beta(x)-\beta(x')\| < Kd(x, x')\} = Pr(\chi_p^2 < \lambda^{-1}K^2) = const.$$

となる．

したがって $\lambda\to 0$ のときには $\beta(x)=\beta(x')$ となって，(4)式のモデルは通常のロジットモデルと一致する．また $\lambda\to\infty$ のときは左辺の確率は 0 に近づく．つまり応答関数に何の制約もないモデルも含まれることになる．このように $\lambda$ はモデルの可変性を表す母数である．

ここで $n$ 個の観測値$(x_1, \cdots, x_n)$における $\beta(x)$ に関心がある場合について，(6)，(7)式の Gauss 過程で表される事前分布の性質を確認しておこう．簡単のために $\beta_i=\beta(x_i)$ と書くと，$\beta_1, \cdots, \beta_n$ の事前分布は次のようになる．

$$\begin{aligned} p(\beta) &= \lim_{\xi\to 0} c\,|\xi^{-1}R|^{-p/2} \\ &\quad \exp\{-(\xi/2)\sum_i\sum_j \rho^{ij}\beta_i^T\Omega^{-1}\beta_j\} \\ &= \lim_{\xi\to 0} c\,|\xi^{-1}R|^{-p/2} \\ &\quad \exp\{-(\xi/2)\beta^T(R^{-1}\otimes\Omega^{-1})\beta\} \end{aligned}$$

ここで $\beta$ は $\beta_1, \cdots, \beta_n$ を一列に並べた $np$ 次元ベクトル，$R=(\rho_{ij})$ は $\rho_{ij}=\rho(x_i, x_j)$ を成分とする $n\times n$ 行列，$R^{-1}=(\rho^{ij})$ であり，$\otimes$ は Kronecker 積を表す．いま，$A$ を $a_{ij}=a(x_i, x_j)=(\lambda/2)(x_i-x_j)^T\Omega^{-1}(x_i-x_j)$ を要素とする $n$ 次元行列とし，$\mathbf{1}$ を 1 を $n$ 個並べたベクトルとすると，次の 2 つの定理が成り立つ．

**定理 2**　$A$ が非特異のとき $\lim_{\xi\to 0}\xi R^{-1} = -A^{-1} + (\mathbf{1}^TA^{-1}\mathbf{1})^{-1}A^{-1}\mathbf{1}\mathbf{1}^TA^{-1}$ であり，これを $M$ と書くと $M$ は次の性質をもつ：(1) $M^T=M$，(2) $M\mathbf{1}=0$，(3) $\mathrm{rank}\, M=n-1$

**定理 3**　$\xi\to 0$ のとき，$\beta$ の分布は $\{\beta_1=\beta_2=\cdots=\beta_n\}$ で定義される $p$ 次元空間 $M$ 上では一様分布(improper distribution)となり，$M$ と直交する $(n-1)p$ 次元の補空間上で真の密度(proper density)をもつ．

以上の命題から，この事前分布は $(n-1)p$ 次元の部分の空間上でのみ情報を持っていることがわかる．さらに $\lambda\to 0$ のときには $\beta$ の分布は退化して

$$Pr(\beta\in M) = 1$$

となり，空間 $M$ 上では一様分布となる．したがって，この場合の事後分布のモードは最尤推定量，すなわち最小 2 乗法と一致する．この点が O'Hagan のモデルとの違いである．
なお $A$ が非特異であるためにはすべての $x$ が異なることが必要であるが，このときには事前分布は

$$p(\beta) \propto \exp\{-(1/2)\beta^T(M\otimes\Omega^{-1})\beta\}$$

と表される．

### 3.2　母数の推定

以下では，任意の $x$ に対して $\beta(x)$ の事後分布のモードと分散を求めるための計算法の概略を紹介する．まず，データ$(x_i, y_i)$と人口 $m_i$ $(i=1, \cdots, n)$ が与えられたときの対数尤度関数は次のようになる．ただし $n$ は地域・年齢階級を表す．

$$\begin{aligned} \log p(y\,|\,\beta(x)) &= \text{const.} + \sum_{i=1}^n [y_i \log F_i \\ &\quad + (m_i-y_i)\log(1-F_i)] \\ &= \text{const.} + \sum_i \beta_i^T x_i y_i \\ &\quad + \sum_i m_i \log(1-F_i) \quad (9) \end{aligned}$$

ある特定の値 $x_0$ における $\beta_0=\beta(x_0)$ を考えるとき，$(\beta_0, \beta_1, \cdots, \beta_n)$ の事前分布と尤度とを組合わせると，次の対数事後密度関数を得る．ただしベクトルおよび行列の上の波(～)は，それが$(\beta_0, \beta_1, \cdots, \beta_n)$に対する$(n+1)$次元のベクトルないし行列であることを表している．この場合，要素の番号は $0, \cdots 1, \cdots, n$ とつけることにする．なお，事後分布を評価するまで $\xi>0$ としておく．

$$\begin{aligned}\phi(\beta) &= \text{const.}+\sum_i \beta_i^T x_i y_i \\ &\quad +\sum_i m_i \log(1-F_i) \\ &\quad -(1/2)\tilde{\beta}^T(\xi\tilde{R}^{-1}\otimes\Omega^{-1})\tilde{\beta} \qquad (10)\end{aligned}$$

今の問題では $x_i \neq x_j$ であるが，一般には $x$ の値に重複がある場合には注意が必要である．たとえば $x_1=x_2$ のときにはすべての $k$ に対して $\rho(x_1, x_k)=\rho(x_2, x_k)$ だから $A$ は特異となり，定理3は適用できない．実際 $x_1=x_2$ とすれば，母数は $\beta_1$，$\beta_2$ の2つではなく $\beta_1$ だけとなる．この点に関しては，上記の事前分布は，たとえば $x_1$ と $x_2$ が一致したときには確率1で $\beta_1=\beta_2$ となり，この部分集合の上で事後分布が最大値をとるような性質を持っているため，以下の計算法は観測値の重複の有無にかかわらず利用可能である．

事後モードを求めるには Newton-Raphson 法と類似の手法を用いて $\phi'(\beta)=0$ を解けばいいが，ここでは近似式

$$\begin{aligned}\phi'(\tilde{\beta}^{(1)}) = 0 &= \phi'(\tilde{\beta}^{(0)})+\phi''(\tilde{\beta}^{(0)})(\tilde{\beta}^{(1)}-\tilde{\beta}^{(0)}) \\ &\quad +o(|\tilde{\beta}^{(1)}-\tilde{\beta}^{(0)}|)\end{aligned}$$

から導かれる次の繰返し法を利用する．

$$\tilde{\beta}^{(1)} = \tilde{\beta}^{(0)}+(-\phi'')^{-1}\phi' \qquad (11)$$

ただし $\phi''$ および $\phi'$ は $\tilde{\beta}_0$ で評価される微分係数で，それぞれ $(n+1)p$ 次元の正方行列およびベクトルである．$\xi\to 0$ とすると $A$ が非特異でない限り $\phi''$ および $\phi'$ が存在しない点に注意が必要である．

ここで以下の記法を導入する．$D_x$ を $p$ 次元の列ベクトル $x_1, \cdots, x_n$ を対角に並べた $(np)\times n$ の行列とする．また $F_i=F(\beta_i^T x_i)$，$f_i=f(\beta_i^T x_i)=F_i(1-F_i)$，$z_i=y_i-m_iF_i$ として $E=\text{diag}(m_1f_1, \cdots, m_nf_n)$ を $n\times n$ 行列，$z=(z_i)$ を $n$ 次元列ベクトルとする．さらに $x_0$ に対応する行または列に0を付加した行列とベクトルに波(～)をつけて表す．たとえば $\tilde{E}=\text{diag}(0, E)$，$\tilde{R}=\text{diag}(0, R)$ はいずれも$(n+1)$次元正方行列，$\tilde{D}_x=\text{diag}(0, D_x)$ は $(n+1)p\times(n+1)$ 次元の行列，$\tilde{z}^T=(0, z^T)$ は$(n+1)$次元ベクトルである．例外として $\tilde{A}$ だけは $i, j=0, \cdots, n$ となる$(n+1)$次元行列 $(a_{ij})$，$\tilde{\mathbf{1}}$ は$(n+1)$次元の1だけからなる列ベクトルとする．

比較的簡単な変形により

$$\begin{aligned}(-\phi'')^{-1} &= [(\xi\tilde{R}^{-1}\otimes\Omega^{-1})+\tilde{D}_x\tilde{E}\tilde{D}_x^T]^{-1} \\ &= (\xi^{-1}\tilde{R}\otimes\Omega)-(\xi^{-1}\tilde{R}\otimes\Omega) \\ &\quad \tilde{D}_x\tilde{L}\tilde{D}_x^T(\xi^{-1}\tilde{R}\otimes\Omega)\end{aligned}$$

と表されることを利用すると，若干の計算の後に(11)式を次のように書き替えることができる．

$$\begin{aligned}\tilde{\beta}^{(1)} &= \tilde{\beta}^{(0)}+(-\phi'')^{-1}\phi' \\ &= \tilde{\beta}^{(0)}+(-\phi'')^{-1}[\tilde{D}_x\tilde{z} \\ &\quad -(\xi\tilde{R}^{-1}\otimes\Omega^{-1})\tilde{\beta}^{(0)}] \\ &= \tilde{\beta}^{(0)}+(-\phi'')^{-1}\tilde{D}_x\tilde{z}-\{(\tilde{1}\otimes 1_p) \\ &\quad -(-\phi'')^{-1}\tilde{D}_x\tilde{E}\tilde{D}_x^T\}\tilde{\beta}^{(0)} \\ &= (-\phi'')^{-1}(\tilde{D}_x\tilde{z}+\tilde{D}_x\tilde{E}\tilde{D}_x^T\beta^{(0)})\end{aligned}$$

最後に $\xi\to 0$ のときの $(-\phi'')^{-1}$ を評価する．いま $X$ は $p$ 次元行ベクトル $x_i^T$ を並べた $n\times p$ の説明変数行列，$\tilde{X}$ は第0行をゼロベクトルとする $X$ の $(n+1)\times p$ 次元への拡張行列，$\Delta=E^{1/2}[I-E^{1/2}D_x^T(A\otimes\Omega)D_xE^{1/2}]^{-1}E^{1/2}$，$S=X^T\Delta X$，$\tilde{\Delta}=\text{diag}(0, \Delta)$ とする．次の定理はこの手法の核となる命題である．

**定理4** $\xi\to 0$ のとき $\Phi=\lim_{\xi\to 0}(-\phi'')^{-1}$ は次の式で与えられる．

$$\begin{aligned}\Phi &= \lim_{\xi\to 0}(-\phi'')^{-1}\\ &= [(\tilde{1}\otimes I_p)+(\tilde{A}\otimes\Omega)\tilde{D}_x\tilde{\Delta}\tilde{X}]\\ &\quad S^{-1}[(\tilde{1}\otimes I_p)+(\tilde{A}\otimes\Omega)\tilde{D}_x\tilde{\Delta}\tilde{X}]^T\\ &\quad -(\tilde{A}\otimes\Omega)-(\tilde{A}\otimes\Omega)\tilde{D}_x\tilde{\Delta}\tilde{D}_x^T(\tilde{A}\otimes\Omega)\end{aligned} \tag{12}$$

最終的な繰返し計算の手順は，ここで与えられた $\Phi$ を用いて

$$\tilde{\beta}^{(1)} = \Phi\tilde{D}_x\tilde{z}+\tilde{D}_x\tilde{E}\tilde{D}_x^T\tilde{\beta}^{(0)}\} \tag{13}$$

と表されるが，この形であればコンピュータ用のプログラムは比較的容易に作成できる．

### 3.3 死亡指標への適用方法

説明変数 $x$ として，小地域の中心点の緯度，経度，年齢階級の 3 変数を想定することができる．平成 9 年の北海道の例では，市区町村の数は 220，年齢階級は 5 歳階級で(0-4, 5-9, …, 95-99, 100-)と 21 階級となるから，対応するサンプルサイズは $n=220\times 21=4620$ である．これに対する応答変数は，対応する小地域，年齢階級の死亡数であり，その分布を 2 項分布 $y\sim B(m,\pi(x))$ とする．先に示した(9)式の対数尤度関数において $0\leq y_i\leq m_i$ である．計算結果については，出力が膨大となるためここには記さないが，この例のような大きなデータを扱う場合にはアルゴリズムを工夫しないと相当の計算時間がかかることが明らかである．この点に関しては，上記のアルゴリズムにおいて行列の処理方法を工夫することによって，大幅に計算時間を短縮できることも示されている．

（青山学院大学経済学部）

### 注

1）その意味については美添(1986)にやや詳しい説明がある．

### 参 考 文 献


厚生省大臣官房統計調査部（1990）「小地域における死亡率――ベイズ統計学からのアプローチ――」，厚生行政科学研究，研究会報告書．

丹後俊郎(1988)「死亡指標の経験的ベイズ推定量について」『応用統計学』第 17 巻第 2 号，pp. 81-96．

美添泰人(1983)「ベイズの手法による統計分析――部分的なサーベイと今後の展望――」竹内 啓編『計量経済学の新展開』東京大学出版会，第 6 章．

美添泰人(1987)「ベイズの手法による二項回帰モデルの推定法」鈴木雪夫・竹内 啓編『社会科学の計量分析』東京大学出版会，第 3 章．

O'Hagan, A. (1978) "Curve Fitting and Optimal Design for Prediction," *Journal of the Royal Statistical Society*, ser. B, 40, pp. 1-42.

Yoshizoe, Y. (1986) *Analysis of the Binary Regression Model*, Harvard University.