

同研究を報告した。ここでは Hirano et al. (1991) の要旨を述べる。

Ling のオーダー k の離散分布

成功の確率 p ($0 < p < 1$) を持つ、大きさ n のベルヌーイ試行を X_1, X_2, \dots, X_n とし、 $Y_i = \prod_{j=i}^{i+k-1} X_j$ 、 $M_n^{(k)} = Y_1 + Y_2 + \dots + Y_{n-k+1}$ とするとき、 $M_n^{(k)}$ の従う分布をオーダー k のタイプ II の二項分布といい、 $B_k^{\text{II}}(n, p)$ とかく (Ling (1988))。 r を正の整数とし、試行数 n を固定しないで $M_n^{(k)}$ が r の値をとるまでに要する試行数を $N_r^{(k)}$ とする。 $N_r^{(k)}$ の従う分布をオーダー k のタイプ III の負の二項分布といい、 $NB_k^{\text{III}}(r, p)$ とかく (Ling (1989))。 $N_r^{(k)} - (k+r-1)$ の分布を $NB_k^{\text{III}}(r, p)$ 、確率生成母関数 (pgf) を $\bar{\phi}_{N_r^{(k)}}^{\text{NB}}(t)$ とかく。

Ling (1988) は $B_k^{\text{II}}(n, p)$ の pgf の漸化式を与えた。我々はそれを陽の形で解いた (Th. 2.2)。この解を用いて確率関数を陽の形で与えた (Th. 2.3)。 $M_n^{(k)}$ の漸近的性質は、ある定常過程がウィナー過程に収束する例となっている (Th. 3.1, 3.2)。

$NB_k^{\text{III}}(r, p)$ の pgf を Ling (1989) より使いやすい形で与え、分散もより単純な形で与えた。また $N_r^{(k)}$ はオーダー k の幾何分布に基づく確率変数の r 個の和として表される (Th. 4.2)。この事実は、この分布のモーメントがオーダー k の幾何分布のモーメントから容易に求まることを示している。

$q \rightarrow 0$ と $rq \rightarrow \lambda (> 0)$ を保って、 $r \rightarrow \infty$ とすると、 $\bar{\phi}_{N_r^{(k)}}^{\text{NB}}(t)$ は $\exp\{-\lambda(1-t^k)\}$ に収束する ($p+q=1$)。これを pgf として持つ分布は平均 λ のポアソン分布の確率を $\{kj : j=0, 1, 2, \dots\}$ 上にとる分布である。

参 考 文 献

- Hirano, K., Aki, S., Kashiwagi, N. and Kuboki, H. (1991). On Ling's binomial and negative binomial distributions of order k , *Statist. Probab. Lett.*, **11**, 503-509.
 岩瀬晃盛, 平野勝臣 (1990). ベキ逆ガウス型分布とその応用, *応用統計学*, **19**, 163-176.
 Ling, K.D. (1988). On binomial distributions of order k , *Statist. Probab. Lett.*, **6**, 247-250.
 Ling, K.D. (1989). A new class of negative binomial distributions of order k , *Statist. Probab. Lett.*, **7**, 371-376.

Multivariate Familial Data の統計解析

小 西 貞 則

多変量統計解析における各種分析手法の適用に際しては、通常、母集団から抽出された互いに独立な観測値に基づいて構成される、ある種の確率行列の分解が基本となる。しかし、医学、疫学、遺伝学等の分野においては、統計的独立性の仮定が満たされないデータに基づく分析をしばしば必要とする。その一例が、遺伝の生物統計学的な研究にみられる、親とその不特定多数の子 (同胞) からなる一つの家族を単位として観測されるデータである。

このようなデータを家族データ (familial data) と呼び、また、各個体がいくつかの特性に関して特徴付けられた多変量データとして観測されるとき、これを多変量家族データ (multivariate familial data) と呼ぶ。ここでは、何組かの多変量家族データに基づいて、種々の家族内特性を探るための統計手法開発を目的として研究を行った。

いま、 N 組の家族データを観測し、そのうち α 番目の家族のデータを

$$z_\alpha = (y'_\alpha, x'_{1\alpha}, x'_{2\alpha}, \dots, x'_{k\alpha, \alpha})' \quad \alpha = 1, 2, \dots, N$$

とおく。ここに、 $y_\alpha = (y_{1\alpha}, y_{2\alpha}, \dots, y_{p\alpha})'$ は p 個の特性に関する母親のデータ、 $x_{j\alpha} = (x_{1j, \alpha}, \dots, x_{qj, \alpha})'$ は q 個の特性に関する j 番目の子のデータとする。例えば、 α 番目の親は k_α 匹の子を同時に生み、生まれた子の間には順序を考慮する必要がないという設定を考えてみる。このとき、 z_α は、平均ベクトル $\mu_\alpha =$

$(\mu'_m, \mu'_s, \dots, \mu'_s)'$, 分散共分散行列

$$\Sigma_a = \begin{pmatrix} \Sigma_m & e'_{ka} \otimes \Sigma_{ms} \\ e_{ka} \otimes \Sigma'_{ms} & I_{ka} \otimes \Sigma_s + (e_{ka} e'_{ka} - I_{ka}) \otimes \Sigma_{ss} \end{pmatrix}$$

の $(p+qk_a)$ -次元分布に従うとする。ただし、 $e_{ka} = (1, \dots, 1)'$, I_{ka} は単位行列、 $A \otimes B$ は行列 A, B のクロネッカー積とする。モデルに含まれるパラメータは、Konishi and Khatri (1990) の一般化推定量を用いて推定し、これを $\{\hat{\Sigma}_m, \hat{\Sigma}_s, \hat{\Sigma}_{ms}, \hat{\Sigma}_{ss}\}$ とおく。

遺伝的な要因を探るという観点からは、複数の特性に関して、(i) 親とその同胞間の関連性の程度、(ii) 同胞内の類似性の度合を定量的に評価する尺度が必要となる。ここでは (i), (ii) を各々 $\hat{\Sigma}_m^{-1} \hat{\Sigma}_{ms} \hat{\Sigma}_s^{-1} \hat{\Sigma}'_{ms}$ の最大固有値の平方根、 $\hat{\Sigma}_{ss} \hat{\Sigma}_s^{-1}$ の最大固有値を用いて推定することを提唱した (Konishi et al. (1991))。さらに、多変量データの次元縮小という観点から、推定量 $\hat{\Sigma}_s$ に基づく主成分分析を研究し、関連する統計的推測理論の研究を行った。

参 考 文 献

Konishi, S. and Khatri, C.G. (1990). Inferences on interclass and intraclass correlations in multivariate familial data, *Ann. Inst. Statist. Math.*, **42**, 561-580.
 Konishi, S., Khatri, C.G. and Rao, C. R. (1991). Inferences on multivariate measures of interclass and intraclass correlations in familial data, *J. Roy. Statist. Soc. Ser. B*, **53**, 649-659.

Small Diffusion の最尤推定量の漸近展開

吉 田 朋 広

確率過程 X がつぎの確率微分方程式で定まるとせよ。

$$dX_t = V_0(X_t, \theta)dt + \varepsilon V(X_t)dw_t, \quad t \in [0, T], \varepsilon \in (0, 1]$$

$$X_0 = x_0.$$

ここで、 θ は k 次元未知パラメータ、 $V_0, V = (V_1, \dots, V_r)$ はなめらかな既知の関数とする。未知母数 θ を観測 $\{X_t; 0 \leq t \leq T\}$ から推定したい。 $\varepsilon \rightarrow 0$ のとき、最尤推定量 $\hat{\theta}_\varepsilon$ は一致性を持ち、1次漸近有効であることが知られている。Malliavin calculus によって、バイアス修正された最尤推定量 $\hat{\theta}_\varepsilon^*(w; \theta_0) = \hat{\theta}_\varepsilon - \varepsilon^2 b(\hat{\theta}_\varepsilon)$ に対して、その分布の漸近展開はつぎのように与えられる。 θ_0 はパラメータの真値を表す。

定理. バイアス修正された最尤推定量 $\hat{\theta}_\varepsilon^*(w; \theta_0)$ に対してつぎの漸近展開が成り立つ。

$$P \left[\frac{\hat{\theta}_\varepsilon^*(w; \theta_0) - \theta_0}{\varepsilon} \in A \right] = \int_A p_0(x) dx + \varepsilon \int_A p_1(x) dx + \dots, \quad \varepsilon \downarrow 0, \quad A \in \mathbf{B}^k.$$

この展開は $A \in \mathbf{B}^k$ に関して一様。とくに、

$$p_0(x) = \phi(x; 0, I^{-1}),$$

$$p_1(x) = \left[\sum_{i,j,l} I^{ij} A_{ijl} x^l + \sum_{i,j,l} I^{ij} B_{ijl} x^l - \sum_{j,l} b^j(\theta_0) I_{jl} x^l - \sum_{i,j,l} A_{ijl} x^i x^j x^l - \sum_{i,j,l} \frac{1}{2} B_{ijl} x^i x^j x^l \right] \phi(x; 0, I^{-1}), \dots$$

ここで、 A_{ijl}, B_{ijl} は確率微分方程式の係数から定まる定数である。 ϕ は正規分布の確率密度で、 $I = (I_{ij})$ は Fisher 情報量、 $I^{-1} = (I^{ij})$ 。

同様にして、尤度比統計量の漸近展開、contiguous alternative での最尤推定量と尤度比統計量の漸近展開、2次漸近中央値不偏推定量の2次の分布の限界とバイアス修正された最尤推定量の2次漸近有効