

とする。 x_k, y_k は大きさが各 n_k, m_k で母比率が p_{1k}, p_{2k} の 2 項分布に従うものとする。第 k 層のオッズ比は、 $\phi_k = [p_{1k}(1-p_{2k})] / [p_{2k}(1-p_{1k})]$ と定義される。通常オッズ比が層によらず一定として、共通オッズ比 ψ の推定が行なわれる。従来、共通オッズ比に関する条件付尤度に基づく推定や検定は計算時間がかかり、かつ標本数が大きくなると計算困難であるとされていた。しかし、この問題はプログラムのアルゴリズムを工夫することにより解決できる。条件付尤度に基づく推定が可能ならば、オッズ比に関する複数の仮説に対して、それぞれの AIC を求め、最良のモデルを選択することができるだろう。

$t_k = x_k + y_k$ が与えられているものとする。 x_k は非心度 ϕ_k の非心超幾何分布に従う。 K 個の層別四分表の条件付対数尤度 L は、

$$L = \sum_{k=1}^K \left[x_k \ln \phi_k - \ln \left(\sum_u w_{ku} \phi_k^u \right) \right] + K_0,$$

となる。ここで、 $K_0 = \sum_k \ln w_{kx_k}$ であり、

$$w_{ku} = \binom{n_k}{u} \binom{m_k}{t_k - u},$$

である。なお、 w_{ku} は、 $n_k < u$ もしくは $u < t_k - m_k$ の場合、0 とする。

条件付最尤推定量 $\hat{\phi}_c$ は、 $\phi_k = \psi$ として Newton-Raphson 法などを用いて求めればよい。通常は、以下の 3 つの仮説について AIC を比較すればよい。

- (1) $H_0: \phi_1 = \dots = \phi_K = \psi = 1$. 推定すべきパラメータはないので、

$$\text{AIC } 0 = 2 \sum_{k=1}^K \ln \binom{N_k}{t_k},$$

となる。ここで、AIC の比較に無関係な定数 K_0 は式から除いている（以下同様）。

- (2) 均一性の仮説 $H_1: \phi_1 = \dots = \phi_K = \psi (\neq 1)$. 推定すべきパラメータ数は 1 であるから、

$$\text{AIC } 1 = -2 \sum_{k=1}^K \left[x_k \ln \hat{\phi}_c - \ln \left(\sum_u w_{ku} \hat{\phi}_c^u \right) \right] + 2.$$

- (3) 不均一性の仮説 $H_2: \phi_i \neq \phi_j (i \neq j; i, j = 1, \dots, K)$. 推定すべきパラメータ数は K 個あるので、

$$\text{AIC } 2 = -2 \sum_{k=1}^K \left[x_k \ln \hat{\phi}_k - \ln \left(\sum_u w_{ku} \hat{\phi}_k^u \right) \right] + 2K,$$

となる。層の個数を 2~20、 $n_k = m_k$ として各 4~512 例の範囲内で変化させ、発生率 p_{2k} とオッズ比を適当に変えて、各 1 万回のシミュレーションを行ない、AIC による方法と他の方法による結果を比較した。 $\psi = 1$ の場合、AIC による方法での H_0 の採択率は、層の個数が 4 程度の場合 80% 弱であり、層が多くなるにつれやや増加するが 80~85% 程度であった。 $\psi = 3$ とした場合、Mantel-Haenszel 検定を 5% 水準で行なった結果と比較すると、 H_0 の棄却率は AIC による方法が常に高かった。また、各層のオッズ比を 2~5 と不均一にした場合、層の個数が 8 程度で Breslow-Day 検定等の検出力と、AIC による H_2 の採択率はほぼ等しく、それ以下では AIC による方が、それ以上では従来の方法の方が検出力は高かった。

一般次元の分配率空間と集中モメント —— 経済統計の幾何学的解析法 ——

田口時夫

昨年度は標記の研究課題に関して、二つの標識をもつ統計解析の場合について成果を報告した。本年

度はその成果を一般に n 標識をもつ統計解析の場合に拡張して報告した。
 n 個の標識に対応した n 次元の統計的ベクトル変量を

$$\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

とした時、 \mathbf{X} の分配率空間は、 \mathbf{X} の規模分布と \mathbf{X} の各要素に関するシエラ分布により規定される $n+1$ 次のベクトル $\mathbf{D}(\mathbf{x})$ である。従って

$$(1) \quad \mathbf{D}(\mathbf{x}) \equiv \begin{pmatrix} D_0(\mathbf{x}) \\ D_1(\mathbf{x}) \\ \vdots \\ D_n(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} F(\mathbf{x}) \\ S(\mathbf{x}) \end{pmatrix}$$

で表した時、 $\mathbf{D}(\mathbf{x})$ の各成分は

$$(2) \quad D_0(\mathbf{x}) \equiv F(\mathbf{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(\boldsymbol{\xi}) d\xi_1 \cdots d\xi_n$$

及び

$$(3) \quad \begin{pmatrix} D_1(\mathbf{x}) \\ \vdots \\ D_n(\mathbf{x}) \end{pmatrix} \equiv S(\mathbf{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} \boldsymbol{\eta} f(\boldsymbol{\xi}) d\xi_1 \cdots d\xi_n, \quad \boldsymbol{\eta} \equiv \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} = \begin{pmatrix} \xi_1/\mu_1 \\ \vdots \\ \xi_n/\mu_n \end{pmatrix}$$

で与えられる。

分配関係 $\mathbf{D}(\mathbf{x})$ の解析においては、 $n+1$ 箇の異なる観測値

$$\mathbf{X}_j = \begin{pmatrix} X_{1j} \\ \vdots \\ X_{nj} \end{pmatrix}; \quad j=1, \dots, n+1$$

のクラスを単位とし、さらに成分 $X_{0j}=1$ を加え

$$(4) \quad \hat{\mathbf{X}}_j \equiv \begin{pmatrix} X_{0j} \\ X_{1j} \\ \vdots \\ X_{nj} \end{pmatrix}; \quad j=1, \dots, n+1$$

を解析単位とするのが妥当である。さらにその測度成分として

$$(5) \quad D(i_1 \cdots i_\nu, j_1 \cdots j_\nu) = \begin{vmatrix} X_{i_1 j_1} & \cdots & X_{i_1 j_\nu} \\ \vdots & & \vdots \\ X_{i_\nu j_1} & \cdots & X_{i_\nu j_\nu} \end{vmatrix}; \quad \nu=1, 2, \dots, n$$

を用いることができる。ここで (i_1, \dots, i_ν) 及び (j_1, \dots, j_ν) はそれぞれ $(0, 1, \dots, n)$ 及び $(1, 2, \dots, n)$ の部分集合とする。

この時、 \mathbf{X} に関する ν 次の集中積率 $c_{j_q i_r}^{(\nu)}$ は、一般に

$$(6) \quad c_{j_q i_r}^{(\nu)} = (-1)^{r+1} E_{X_1} \cdots E_{X_\nu} \left[\left\{ \prod_{p=1, \dots, q, \dots, \nu} \text{sgn}(X_{i_p j_p} - X_{i_r j_q}) \right\} \times D \begin{pmatrix} i_0 & i_1 & \cdots & i_r & \cdots & i_\nu \\ j_1 & j_2 & \cdots & j_\nu & \cdots & j_\nu \end{pmatrix} \right]$$

で規定できる。その個数 $N^{(m)}$ は $n \geq 3$ の場合

$$(7) \quad N^{(n)} = \left(\frac{n^2}{4} + \frac{n}{4} + 1 \right) 2^n - 2$$

である。集中積率間の単純な演算は各種の記述統計量を与え、その全体は独自の体系と解析方法を形成する事ができる。

疾病構造の変遷に関する研究

駒澤 勉

動脈硬化性疾患の予防を目的とした集団検診の1975年、1985年の被検診者に対して、発症追跡調査を実施した検査データについて、10年の間に検査データ構造がどのように変遷したかを数量化分析によって検討した。

対象は健康に関する共同研究を行っている(財)日本労働文化協会・動脈硬化疫学研究所が実施した集団検診の被検診者である。これらの対象に対して検診後1年半以内に動脈硬化性の脳・心血管疾患の発症追跡調査を実施した。2年次別の調査対象数、回収数、発症数等は表1、発症状況は表2の通りであった。

虚血性心疾患(狭心症、心筋梗塞)の2年次別の発症割合は1975年次調査43.2%、1985年次53.5%であり、脳卒中(脳梗塞、脳出血)は1975年次39.2%、1985年次29.2%と、多くの報告の通り、心疾患と脳疾患の割合が逆転していることをこの調査研究でも示した。

この2年次別の発症割合の変遷を検査データ構造から分析するために数量化Ⅲ類分析で検討した。分析に用いた検査項目は、血清総コレステロール、眼底検査(Scheie 硬化性所見)、心電図、血圧値、大動脈脈波速度値(一つの動脈硬化度の指標)である。疾病危険因子の2年次別の変化は血圧値、眼底所見、大動脈脈波速度値などの高血圧動脈硬化因子において1985年次の方が明らかに各項目・区分の数量が臨地的見地からもリスクの高い順序に与えられ、かつ虚血性心疾患の危険因子である心電図のST・T

表1. 対象数と内訳。

	1975年次	1985年次
調査数	82,060名	100,352名
回収数	37,457	34,895
発症数	199	301

表2. 発症疾患別内訳。

	1975年次	1985年次
狭心症	50名 (25.1%)	110名 (36.6%)
心筋梗塞	36 (18.1%)	51 (16.9%)
TIA・RIND	25 (12.6%)	41 (13.6%)
脳梗塞	64 (32.2%)	69 (22.9%)
脳出血	14 (7.0%)	19 (6.3%)
くも膜下出血	10 (5.0%)	11 (3.7%)
合計	199 (100%)	301 (100%)