



Title	対数線形モデルによる3次元分割表データの分析法
Author(s)	吉田, 光雄
Citation	大阪大学人間科学部紀要. 1980, 6, p. 101-142
Version Type	VoR
URL	<a href="https://doi.org/10.18910/12373">https://doi.org/10.18910/12373</a>
rights	
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

対数線形モデルによる 3 次元  
分割表データの分析法

吉 田 光 雄

## 対数線形モデルによる3次元 分割表データの分析法

### I 序

実験、調査等によって蒐集されたデータに対して、その内部構造を検討するために要因（項目、変量）間相互の関連の分析を行なうのは常套的手段である。データが連続量（quantitative data）の場合には、積率に基づく Pearson, K. の相関係数の拡張としての多変量相関係数、例えば重相関係数、偏相関係数、或いは正準相関係数等を目的に応じて算出し、検討することによってデータの構造を多面的に吟味することができる。

一般に、連続量データは加減算（場合によっては乗除算を加えての四則演算）が可能であり、その故に低次の積率を求めることによってデータの持つ情報を少数の指標に圧縮することができる。又、分布形が既知であるか、或いはその仮定が許容されるときには、分布形とその母数の推定値をもってほぼ完全に縮約することが可能である。多変量の場合であっても2次以下の積率、即ち平均値ベクトル、分散・共分散行列さえ求めておけば十分な場合が多い。更に一般化分散を求めれば、分散・共分散はスカラーにさえ圧縮される。

ところが、属性データ（categorical data）の場合には、加減算が許されないために積率が計算できず、分析の期間を通じて常にカテゴリーを残しておかねばならない。統計的分析は連続量データに比し格段に煩瑣なものである。統計学の発展の歴史においても先ず連続量データに関する理論と方法が研究され、属性データについての研究は後続であった<sup>21)</sup>。その間の事情は、例えば検定の理論が先ずパラメトリックなものから研究され、ノンパラメトリックなものは遅れて開始されたことと軌を一にしているし、又多変量解析の理論は連続量たる多変量正規分布を前提としたものから研究され、それらについてはほぼ体系化された今、属性データを主とした離散量多変量解析の研究は Bishop, Y. M. et al.<sup>7)</sup>によれば、コンピュータの普及と共にここ10年あまり前から急速に始められたばかりという。

心理学で扱うデータがすべて連続量であると限らないのは周知の事実である。Stevens, S. S. の分類を借りれば加減算の許される間隔尺度、比例尺度で与えられるデータはむしろ稀であって、尺度構成の領域における研究は、換言すれば測定水準の向上を目指すものといえよう。即ち、心理学的実験や調査で得られる粗データの多くは名義尺度或いは順序尺度の水準のものであり、これを事後の処理により間隔尺度以上の水準に引き上げるのをねらいとしている。そして少くともデータが順序づけられたカテゴリーの場合には、尺度構成の領域で開発され

た手法を用いることにより加減算の演算の可能性が保証され、或いは順序という制限のもとで許容される統計的処理の範囲が拡大されるが、順序の保証すらない属性データでは、処理法に対する制約は厳しい。連続量データに適用可能な効率的な積率の計算法は、カテゴリーの粋が外せない以上、ここでは用いることができない。

こうした属性データの連関の分析に際して、従来多くなされて来た方法は  $\chi^2$ -統計量を用いる独立性の検定である。項目間のクロス集計結果は2次元の分割表 (contingency table) にまとめられ、 $\chi^2$ -値が計算され両項目を独立と見なして良いかどうかの検定がなされる。しかし、適用されているのは多くは項目が2要因の場合までであって、情報理論を適用する方法以外は3次元以上の分割表は分析法が普及していないためもあり、限局されていた。そもそも、計算法、表示法の煩雑さからクロス集計すらなされることが少ないのが現状のようである。或いは分析法として分散分析の手段が適用されているケースも散見されるが、この場合には、分散分析の前提である(1)データが分散の算出可能な測定値であり、それが例えば二元配置の場合、 $X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$  の単純効果、交互作用効果、誤差項の各成分に分解できること、(2)誤差項が正規分布に従うと仮定できること、即ち  $\varepsilon \sim N(0, \sigma^2)$  の点についての吟味が必要である。属性データ (多くの場合は頻度のデータであるが) に対して、これらの検討を欠く機械的な適用は危険といわざるを得ない。

こうした属性データの処理法のひとつとして、本稿では近年、統計学や社会学の領域で注目を集めるに至った Log-Linear Model (対数線形モデル) による多次元分割表データの分析法を紹介する。

Log-Linear Model は先ず Birch, M. W.<sup>4)</sup> (1963) によって提案され、Goodman, L. A.<sup>12-15)</sup> (1970~72) によって社会学の領域に紹介され理論の発展と多方向への適用が試みられた。そして Haberman, S. J.<sup>16), 17)</sup> (1973, 74), Plackett, R. L.<sup>22)</sup> (1974), Bishop, Y. M. M.<sup>5), 6)</sup> (1969), Bishop, Y. M. M. et al.<sup>7)</sup> (1975) らによって数理統計学的に詳細に吟味され、体系化された理論となっている。又同モデルの入門的解説は Everitt, B. S.<sup>9)</sup> (1977), Payne, C.<sup>19)</sup> (1977), Upton, G. J. G.<sup>26)</sup> (1977) 等に詳しい。

以下、Ⅱでは同モデルの考え方を2次元分割表の場合について紹介し、Ⅲではそれを3次元分割表の場合に拡張する。同様にして4次元以上の場合への拡張も可能であるが詳細な方法の検討は今後の課題とする。更にⅣでは3次元分割表の場合についてのコンピュータ・プログラミングについて言及し、Ⅴでは数値例をもとに同モデルを具体的に述べる。そして最後にⅥでは情報理論で用いられる情報量との関連について吟味しておく。

本稿における紹介は上記の論文や書物に負っており、最終適合モデル発見のアルゴリズム、結果の図示、AIC によるモデルの検討、解法の比較、情報理論との関連、及びコンピュータプログラミングは筆者の試論、試作である。

## II 2次元分割表

### 1, モデル

2 要因  $A, B$  が各々  $I, J$  個のカテゴリに分割されているとき, 同時観測度数を  $f_{ij}$ , 同時期待度数 (モデル  $H_1$  のもとでの) を  $F_{ij}$  とする。このとき, 次の5種類のモデルが考えられる。但しこの場合のモデルとは対数線形モデルというときのモデルではなく, その下位モデル即ち仮説という意味である。

#### (1) $H_4$ : 独立モデル (Independence Model)

$A, B$  が独立で細胞の同時母比率を  $\pi_{ij}$ , 周辺比率を  $\pi_{i\cdot}, \pi_{\cdot j}$  とするとき,  $H_4$  のもとでは  $\pi_{ij} = \pi_{i\cdot} \pi_{\cdot j}$  であり, 従って期待度数の間には周辺度数を  $F_{i\cdot}, F_{\cdot j}$ , 全観測度数を  $n$  とすれば

$$F_{ij} = F_{i\cdot} F_{\cdot j} / n \quad (2.1)$$

の関係がある場合である。このとき両辺の対数をとれば

$$\log F_{ij} = \log F_{i\cdot} + \log F_{\cdot j} - \log n \quad (2.2)$$

となり, 今  $\mu$ -項として

$$\begin{aligned} \mu &= \frac{1}{IJ} \sum_i \sum_j \log F_{ij} \\ \mu_1(i) &= \frac{1}{J} \sum_j \log F_{ij} - \mu \\ \mu_2(j) &= \frac{1}{I} \sum_i \log F_{ij} - \mu \end{aligned} \quad (2.3)$$

$$\text{但し } \sum_i \mu_1(i) = \sum_j \mu_2(j) = 0 \text{ (束縛条件)}$$

とおけば (2.2) 式は (注)

$$\log F_{ij} = \mu + \mu_1(i) + \mu_2(j) \quad (2.4)$$

と表現することができる。記号として Bishop<sup>7)</sup> らは  $u_1(i)$ , Goodman<sup>12)</sup> は  $\lambda_i$  を用いているが本稿では Payne<sup>19)</sup> にならい  $\mu_1(i)$  を用いることとする。 $\mu_1(i), \mu_2(j)$  は要因  $A, B$  の単純効果,  $\mu$  は全体的修正項である。期待度数の対数度数が各効果の線形結合として表わされるとするものであり, Log-Linear Model (以下 LLM と略記する) と呼ばれるゆえんである。

(注) (2.2)式から(2.4)式への代数的展開は容易である。(2.2)式を  $j$  について和をとり  $J$  で除し, 又  $i, j$  について和をとり  $IJ$  で除したものの差をとればよい。展開の過程で

$$\begin{aligned} \mu_1(i) &= \frac{1}{J} \sum_j \log F_{ij} - \frac{1}{IJ} \sum_i \sum_j \log F_{ij} \\ &= \log F_{i\cdot} - \frac{1}{I} \sum_i \log F_{i\cdot} \end{aligned} \quad (2.5)$$

$$\mu_2(j) = \frac{1}{I} \sum_i \log F_{ij} - \frac{1}{IJ} \sum_i \sum_j \log F_{ij}$$

$$=\log F_{.j} - \frac{1}{J} \sum_j \log F_{.j}$$

も得られる。(2.5)式を  $i, j$  について和をとることにより  $\sum_j \mu_1(i) = \sum_j \mu_2(j) = 0$  も容易に確かめることもできる。

## (2) $H_5$ : 飽和モデル (Saturation Model)

独立モデルの場合にはモデルが適合しないとき  $f_{ij}$  と  $F_{ij}$  との間にはずれが生じているが、飽和モデルではそれを  $\mu$ -項に組み込み  $\mu_{12}(ij)$  とし、 $f_{ij} = F_{ij}$  とおく。このことの故に飽和モデルと呼ばれている。このとき LLM は

$$\log F_{ij} = \mu + \mu_1(i) + \mu_2(j) + \mu_{12}(ij) \quad (2.6)$$

となり、 $\mu, \mu_1(i), \mu_2(j)$  は先の場合と同様単純効果、 $\mu_{12}(ij)$  は  $A, B$  の交互作用効果であり

$$\mu_{12}(ij) = \log F_{ij} - \frac{1}{I} \sum_i \log F_{ij} - \frac{1}{J} \sum_j \log F_{ij} + \mu \quad (2.7)$$

$$\text{但し } \sum_i \mu_{12}(ij) = \sum_j \mu_{12}(ij) = 0 \text{ (束縛条件)}$$

である。

このモデルで  $\mu_{12}(ij) = 0$  と見なしうるとき、先の独立モデルに帰着し  $A, B$  は独立と結論できる。又  $\mu_{12}(ij) \neq 0$  のとき、独立モデルは適合せず  $A, B$  は関連あり（帰無仮説採択の意味において）と結論づけることができる。従来2次元の分割表において  $\chi^2$ -統計量を用いた独立性の検定がなされて来たが、同じ結論を LLM において  $H_4, H_5$  のいずれのモデルを採用するかによって得ることができる。

更に今、 $\log F_{ij} = v_{ij}$  とおくと、 $\mu = \bar{v}_{..}$ ,  $\mu_1(i) = \bar{v}_{i.} - \bar{v}_{..}$ ,  $\mu_2(j) = \bar{v}_{.j} - \bar{v}_{..}$ ,  $\mu_{12}(ij) = v_{ij} - \bar{v}_{i.} - \bar{v}_{.j} + \bar{v}_{..}$  となり、 $X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$  を想定し  $\alpha_i = \bar{X}_{i.} - \bar{X}_{..}$ ,  $\beta_j = \bar{X}_{.j} - \bar{X}_{..}$ ,  $\varepsilon_{ij} = X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..}$  と分解する二元配置分散分析のモデルと類似している。従って  $\mu_{12}(ij) \sim N(0, \sigma^2)$  の先の(2)の仮定を導入できるとき分散分析の処理が可能である。(1)の仮定については、(2.6)式から対数を外せば

$$\begin{aligned} F_{ij} &= \exp(\mu + \mu_1(i) + \mu_2(j) + \mu_{12}(ij)) \\ &= \tau \cdot \tau_1(i) \cdot \tau_2(j) \cdot \tau_{12}(ij) \end{aligned} \quad (2.8)$$

$$\text{但し } \tau = e, \quad \tau_1(i) = \exp(\mu_1(i)), \quad \tau_2(j) = \exp(\mu_2(j)),$$

$$\tau_{12}(ij) = \exp(\mu_{12}(ij))$$

となり、LLM は頻数  $F_{ij}$  を各効果の成分の積に分解して分析するモデルといえよう。頻数データを比率に分解（どの要因が何%関与しているか）することができるというのは無理のない仮定であり、LLM の根拠を特に否定する理由はない。

## (3) $H_1 \sim H_3$ : 縮減モデル (Reduced Model)

データの形式は  $A, B$  2次元の分割表で与えられていても構造としては1次元以下の場合

も考えられ、これらは縮減モデルと呼ばれている。

$H_1$ : 細胞の母比率が  $\pi_{ij} = \frac{1}{IJ}$  ですべて等しい場合,  $\log F_{ij} = \mu$  と書ける。

$H_2$ :  $\pi_{ij} = \pi_{i.}/J$ ,  $\log F_{ij} = \mu + \mu_1(i)$  の場合。

$H_3$ :  $\pi_{ij} = \pi_{.j}/I$ ,  $\log F_{ij} = \mu + \mu_2(j)$  の場合。

$H_2, H_3$  は1要因の単純効果のみしか存在しない場合である。但しこれらはいずれも trivial であり、重要なのは  $H_4, H_5$  であるので以下両モデルを中心に解法を述べることにする。

## 2, 解 法

### 第1法

飽和モデルでは  $\hat{F}_{ij} = f_{ij}$  であるので  $P = I \times J$  元の連立方程式を解いて  $\mu$ -項を求めることができる。(  $F_{ij}$  の推定値の意味で, 又後述 (2.11) 式では  $\mu$  の解といった意味で  $\hat{\mu}$  を冠する。)  $\mu$ -項の未知数は  $(1+I+J+IJ)$  個であるが束縛条件があるため  $1+(I-1)+(J-1)+(I-1)(J-1)=IJ$  個に減少し, 連立方程式の次元数と一致する。

今, 簡単のため  $I=2, J=3, P=IJ=6$  の場合について例示する。求める  $\mu$ -項は  $\mu, \mu_1(1), \mu_2(1), \mu_2(2), \mu_{12}(11), \mu_{12}(12)$  の6個であり, 残りは  $\mu_1(2) = -\mu_1(1), \mu_2(3) = -(\mu_2(1) + \mu_2(2)), \mu_{12}(13) = -(\mu_{12}(11) + \mu_{12}(12)), \mu_{12}(21) = -\mu_{12}(11), \mu_{12}(23) = -\mu_{12}(13)$  より求めることができる。これを行列で表現して

$$\mathbf{f}' = (\log F_{11}, \log F_{12}, \log F_{13}, \log F_{21}, \log F_{22}, \log F_{23})'$$

$$\boldsymbol{\mu}' = (\mu, \mu_1(1), \mu_2(1), \mu_2(2), \mu_{12}(11), \mu_{12}(12))'$$

とおき, 更に計画行列 (design matrix) として

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & -1 & 1 & 1 \end{pmatrix} \quad (2.9)$$

を選べば連立方程式は

$$\mathbf{f} = \mathbf{X}\boldsymbol{\mu} \quad (2.10)$$

と書くことができる。計画行列の性質から  $I, J$  が共に2の場合には  $\mathbf{X}'\mathbf{X} = \mathbf{PI}$  であるので

$$\hat{\boldsymbol{\mu}} = \frac{1}{P} \mathbf{X}'\mathbf{f} \quad \text{但し } P=4 \quad (2.11)$$

より  $\mu$  の解を求めることができる。一般の  $I, J$  の場合には  $\mathbf{X}'\mathbf{X}$  を求め

$$\hat{\boldsymbol{\mu}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{f} \quad (2.12)$$

とすればよい。

計画行列の第  $l$  列 ( $l=2, \dots, P$ ) を  $P$  で除したもの (一般の場合には  $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}$  の第  $l$  行) を  $\mathbf{a}_l$  とすれば

$$\hat{\mu}_l = \mathbf{a}_l' \mathbf{f} = \sum_{r=1}^P a_{lr} \log F_r \quad (2.13)$$

であり,  $\mu$  の第 2 項以下は  $\sum_{r=1}^P a_{lr} = 0$  (注<sup>1)</sup>) ( $l=2, \dots, P$ ) であるので  $\log F_r$  の対比(contrast)として求めることができる。但しここで  $r$  は添字  $i, j$  を第 2 添字  $j$  から先に変化させたものとする。このとき

$$E(\hat{\mu}_l) = \mu_l \quad (2.14)$$

$$V(\hat{\mu}_l) = \sum_{r=1}^P a_{lr}^2 / F_r$$

であることが証明されており注<sup>2)</sup>

$$\hat{\mu}_s = (\hat{\mu}_l - E(\hat{\mu}_l)) / \sqrt{V(\hat{\mu}_l)} \quad (2.15)$$

で  $\hat{\mu}$  を標準化すれば  $n$  が大のとき漸近的に  $\hat{\mu} \sim N(0, 1)$  となり, これを用いて  $H_0: \mu_l = 0$  の検定をすることができる。即ち, 標準正規分布の両側  $100\alpha\%$  点を  $K_\alpha$  とすれば  $|\hat{\mu}_s| > K_\alpha$  で  $H_0$  を棄却すればよい。

独立モデルの場合には  $f_{ij}$  と  $\hat{F}_{ij}$  は一致しないので

$$\mathbf{f} = \mathbf{X}\boldsymbol{\mu} + \boldsymbol{\varepsilon}$$

と重回帰分析のモデルとなり,  $\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = \min$  として最少 2 乗解  $\boldsymbol{\mu} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{f}$  が求まるがこれは先の飽和モデルの場合と一致する。

種々のデータに適用して解を求めたところ筆者の経験では  $n$  は少なくとも 1000 以上でないとい次に述べる第 2 法の解と一致しないことが多く, 正規近似が十分でないようである。従って LLM の一般解としては次の第 2 法の方が望ましいように思える。

(注1) Bishop et al<sup>1)</sup> (p. 497) は  $\mu$  項の算出公式を次式で与えている。

$$\begin{aligned} \mu_1(i) &= \sum_{l \neq i}^I \sum_m^J \left(-\frac{1}{IJ}\right) \log f_{lm} + \sum_m^J \frac{1}{J} \left(1 - \frac{1}{I}\right) \log f_{im} \\ \mu_2(j) &= \sum_l^I \sum_{m \neq j}^J \left(-\frac{1}{IJ}\right) \log f_{lm} + \sum_l^I \frac{1}{I} \left(1 - \frac{1}{J}\right) \log f_{lj} \\ \mu_{12}(ij) &= \left(1 - \frac{1}{I}\right) \left(1 - \frac{1}{J}\right) \log f_{ij} + \sum_{m \neq j}^J \left(-\frac{1}{J}\right) \left(1 - \frac{1}{I}\right) \log f_{im} \\ &\quad + \sum_{l \neq i}^I \left(-\frac{1}{I}\right) \left(1 - \frac{1}{J}\right) \log f_{lm} + \sum_{l \neq i}^I \sum_{m \neq j}^J \left(\frac{1}{IJ}\right) \log f_{lm} \end{aligned} \quad (2.16)$$

例えば  $\mu_1(i)$  の係数の和をとると

$$\left(-\frac{1}{IJ}\right)(I-1)J + \frac{1}{J} \left(1 - \frac{1}{I}\right)J = 0$$



となっており、他も同様である。

(注2) Bishop et al<sup>17)</sup> (p. 141, 487~) に証明が与えられているがその骨子は次のごとくである。

$\hat{Q}_n$  をパラメータ  $\theta$  の推定量とし、 $\sqrt{n}(\hat{Q}_n - \theta)$  が  $n \rightarrow \infty$  のとき  $N(0, \sigma^2(\theta))$  に、即ち  $\hat{Q}_n$  が平均  $\theta$ 、分散  $\sigma^2(\theta)/n$  の正規分布に法則収束するものとする。このとき  $\hat{Q}_n$  の微分可能な実数値関数  $f(\hat{Q}_n)$  は漸近的に平均  $f(\theta)$ 、分散  $\sigma^2(\theta)[f'(\theta)]^2/n$  の正規分布に従う (同書定理 14.6-1)。

LLM の場合、 $\mathbf{X} = (X_1, \dots, X_r, \dots, X_p)$  が多項分布  $M_p(n, \pi)$  に従うと考えれば、 $\pi_r$  の MLE は  $\hat{\pi}_r = X_r/n$  であり、 $E(\hat{\pi}_r) = \pi_r$ 、 $V(\hat{\pi}_r) = \pi_r(1 - \pi_r)/n$ 、 $Cov(\hat{\pi}_r, \hat{\pi}_s) = -\pi_r\pi_s/n$  である。 $X_r$  の対数変換を  $Y_r = \log X_r = \log n \hat{\pi}_r$  とすれば上の定理より

$$V(Y_r) = \frac{\pi_r(1 - \pi_r)}{n} \left( \frac{n}{\hat{\pi}_r} \right)^2 = \frac{1 - \pi_r}{n\pi_r} = \frac{1}{n\pi_r} - \frac{1}{n} \rightarrow \frac{1}{n\pi_r} = \frac{1}{F_r} \quad (n \rightarrow \infty). \quad (2.17)$$

更に LLM の  $\mu$ -項の第  $l$  項は  $\hat{\mu}_l = \sum_r a_{lr} Y_r = \sum_r a_{lr} \log X_r$  で求められるので

$$V(\hat{\mu}_l) = \sum_r a_{lr} V(Y_r) + 2 \sum_{r \neq s} a_{lr} a_{ls} Cov(Y_r, Y_s) = \sum_r a_{lr}^2 \frac{1}{F_r} \quad (2.18)$$

となる。ここで共分散は定理の類推から  $Cov(Y_r, Y_s) = -\frac{\pi_r\pi_s}{n} \left( \frac{1}{\hat{\pi}_r\hat{\pi}_s} \right) = -\frac{1}{n} \rightarrow 0$  と考えてよい。

## 第2法

$H_1 \sim H_4$  の各モデルのもとでの期待度数  $\hat{F}_{ij}$  を求め、観測度数  $f_{ij}$  とのずれを検討する解法である。 $H_5$ : 飽和モデルの場合には  $\hat{F}_{ij} = f_{ij}$ ,  $df = 0$  となり検定はできないので  $H_1 \sim H_4$  の場合について行なえばよい。各モデル下での  $\hat{F}_{ij}$  の算出公式、自由度については表 1 に示す。適合度検定について従来は Pearson, K. の  $\chi^2$ -統計量を用いる方法が多用されて来たが、ここでは尤度比統計量  $G^2$  (又は  $\chi^2_L$ ) を使用することとする。階層モデル (後述) のと

表 1 2次元分割表のモデル

番号	LLM $\log F_{ij}$	構造 $\pi_{ij}$	期待度数 $\hat{F}_{ij}$	自由度 $df$	束縛
H <sub>1</sub>	$\mu$	$\frac{1}{IJ}$	$\frac{n}{IJ}$	$IJ-1$	{n}
H <sub>2</sub>	$\mu + \mu_1$	$\frac{\pi_{i.}}{J}$	$\frac{f_{i.}}{J}$	$I(J-1)$	{A}
H <sub>3</sub>	$\mu + \mu_2$	$\frac{\pi_{.j}}{I}$	$\frac{f_{.j}}{I}$	$J(I-1)$	{B}
H <sub>4</sub>	$\mu + \mu_1 + \mu_2$	$\pi_{i.} \pi_{.j}$	$\frac{f_{i.} f_{.j}}{n}$	$(I-1)(J-1)$	{A}{B}
H <sub>5</sub>	$\mu + \mu_1 + \mu_2 + \mu_{12}$	—	$f_{ij}$	0	{AB}

き,  $G^2$ -統計量には加算性が成立するので便利である。検定の結果,  $H_1 \sim H_4$  のすべてが棄却されれば  $H_5$  を採用すればよい。 $H_1 \sim H_4$  の中に採択されるものがあればそれらを用いて総合的判断により最適モデルを1つ結論づける。この際のアゴリズムについてはⅢ, 3次元分割表の場合に述べる。

2次元分割表の場合, 関心は結局のところ  $H_4$  か  $H_5$  のいずれかを採用することであるので, その意味では従来の独立性の検定を行なうことで十分であるが, LLM の方法を適用すれば  $H_1 \sim H_5$  についても吟味することができるし, 加えて各効果を定量的に抽出することもできる。又3次元以上への拡張も容易であり, 考えうる種々のモデルを総合的, 体系的に検討できる点が同法の特徴といえよう。但し計算量は龐大となるのでコンピュータの使用を俟って始めて適用可能な方法である。

### Ⅲ 3次元分割表

#### 1. $\mu$ -項

3要因を  $A, B, C$  とし, カテゴリー数を各  $I, J, K, P=IJK$  とする。単純効果, 2次の交互作用効果は2次元の場合の拡張である。

$$\begin{aligned}\mu &= \frac{1}{IJK} \sum_i \sum_j \sum_k \log F_{ijk} - \bar{v} \dots \\ \mu_1(i) &= \frac{1}{JK} \sum_j \sum_k \log F_{ijk} - \mu = \bar{v}_{i..} - \bar{v} \dots \\ \mu_2(j) &= \frac{1}{IK} \sum_i \sum_k \log F_{ijk} - \mu = \bar{v}_{.j.} - \bar{v} \dots \\ \mu_3(k) &= \frac{1}{IJ} \sum_i \sum_j \log F_{ijk} - \mu = \bar{v}_{..k} - \bar{v} \dots\end{aligned}\tag{3.1}$$

$$\text{但し } \sum_i \mu_1(i) = \sum_j \mu_2(j) = \sum_k \mu_3(k) = 0 \text{ (束縛条件)}$$

$$\begin{aligned}\mu_{12}(ij) &= \frac{1}{K} \sum_k \log F_{ijk} - \frac{1}{JK} \sum_j \sum_k \log F_{ijk} - \frac{1}{IK} \sum_i \sum_k \log F_{ijk} + \mu \\ &= \bar{v}_{ij.} - \bar{v}_{i..} - \bar{v}_{.j.} + \bar{v} \dots\end{aligned}\tag{3.2}$$

同様に

$$\begin{aligned}\mu_{13}(ik) &= \bar{v}_{i.k} - \bar{v}_{i..} - \bar{v}_{..k} + \bar{v} \dots \\ \mu_{23}(jk) &= \bar{v}_{.jk} - \bar{v}_{.j.} - \bar{v}_{..k} + \bar{v} \dots \\ \sum_i \mu_{12}(ij) &= \sum_j \mu_{12}(ij) = \sum_i \mu_{13}(ik) = \dots = 0 \text{ (束縛条件)}\end{aligned}$$

3次の交互作用は次のように導入される。

$$\begin{aligned}\mu_{123}(ijk) &= \log F_{ijk} - \frac{1}{I} \sum_i \log F_{ijk} - \frac{1}{J} \sum_j \log F_{ijk} - \frac{1}{K} \sum_k \log F_{ijk} \\ &\quad + \frac{1}{IJ} \sum_i \sum_j \log F_{ijj} + \frac{1}{IK} \sum_i \sum_k \log F_{ijk} + \frac{1}{JK} \sum_j \sum_k \log F_{ijk} - \mu \\ &= v_{ijk} - \bar{v}_{ij.} - \bar{v}_{i.k} - \bar{v}_{.jk} + \bar{v}_{i..} + \bar{v}_{.j.} + \bar{v}_{..k} - \bar{v} \dots\end{aligned}\quad (3.3)$$

$$\sum_i \mu_{123}(ijk) = \sum_j \mu_{123}(ijk) = \sum_k \mu_{123}(ijk) = 0 \quad (\text{束縛条件})$$

2次元分割表の場合と同じく  $v$  で表現すれば三元配置分散分析  $X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \alpha\beta_{ij} + \alpha\gamma_{ik} + \beta\gamma_{jk} + \alpha\beta\gamma_{ijk}$  の場合と類似しており統一的理解が容易となる。

## 2. モデル及び解法

$\mu$ -項をすべて書き下すと

$$\log F_{ijk} = \mu + \mu_1(i) + \mu_2(j) + \mu_3(k) + \mu_{12}(ij) + \mu_{13}(ik) + \mu_{23}(jk) + \mu_{123}(ijk) \quad (3.3)$$

となり、これは飽和モデルである。以後自明の場合には添字  $i, j, k$  を省略し

$$\begin{aligned}\log F_{ijk} &= \mu + \mu_1 + \mu_2 + \mu_3 + \mu_{12} + \mu_{13} + \mu_{23} + \mu_{123} \\ &= \mu. + \mu_{12} + \mu_{13} + \mu_{23} + \mu_{123}\end{aligned}\quad (3.4)$$

等と書くことにする。(3.1)~(3.3) 式の  $\mu$ -項を用いて (3.4) 式が表現できること、及び束縛条件が満たされていることは先の2次元分割表の場合と同様である。

第1の解法では  $P = IJK$  元の連立方程式を解き、検定により  $H_0: \mu_i = 0$  のものを落して LLM のモデルを決めればよい。第2法で解くために考えるモデルをすべて列挙すれば表2のごとくとなり、合計19組のモデルについて適合度検定を反復すればよい。 $H_1 \sim H_{10}$  は Reduced Model であり、3次元分割表の本来の興味は  $H_{11} \sim H_{19}$  についての検討である。

### (1) 縮減モデル (Reduced Model), $H_1 \sim H_{10}$

データの形式は3次元で与えられていても構造としては2次元以下に縮約される場合であって、 $H_2 \sim H_4$  は1次元のモデル、 $H_5 \sim H_{10}$  は2次元のモデルである。

### (2) 相互独立モデル (Mutual Independence Model), $H_{11}$

2次元の独立の場合の素直な3次元への拡張であって、 $A, B, C$  が相互に独立である場合である。即ち  $P(A \cap B \cap C) = P(A)P(B)P(C)$  であり母比率間には  $\pi_{ijk} = \pi_{i..}\pi_{.j.}\pi_{..k}$  の関係が見られる場合である。期待度数の算出に周辺度数  $\hat{F}_{i..} = f_{i..}$ ,  $\hat{F}_{.j.} = f_{.j.}$ ,  $\hat{F}_{..k} = f_{..k}$  が用いられるので拘束条件から見てモデルを  $\{A\}, \{B\}, \{C\}$  又は  $(A \perp\!\!\!\perp B \perp\!\!\!\perp C)$  等の記号で端的に表現することにする。ここで  $\{ \}$  は Payne<sup>19)</sup> による表記方法であり、1次元の周辺度数を例えば  $\{A\}$ , 2次元の周辺度数を  $\{AB\}$  と表わすことにすれば  $\{A\} = f_{i..} = \sum_i \sum_k$

表2 3次元分割表のモデル

番号	LLM $\log F_{ijk}$	構造 $\pi_{ijk}$	期待度数 $\hat{F}_{ijk}$	自由度 $df$	束縛	パラメータ数 $r$
(1) 縮減モデル						
H <sub>1</sub>	$\mu$	$1/IJK$	$n/IJK$	$IJK-1$	$\{n\}$	1
H <sub>2</sub>	$\mu + \mu_1$	$\pi_{i..}/JK$	$f_{i..}/JK$	$I(JK-1)$	$\{A\}$	$I$
H <sub>3</sub>	$\mu + \mu_2$	$\pi_{.j.}/IK$	$f_{.j.}/IK$	$J(K-1)$	$\{B\}$	$J$
H <sub>4</sub>	$\mu + \mu_3$	$\pi_{..k}/IJ$	$f_{..k}/IJ$	$K(IJ-1)$	$\{C\}$	$K$
H <sub>5</sub>	$\mu + \mu_1 + \mu_2$	$\pi_{i..}\pi_{.j.}/K$	$f_{i..}f_{.j.}/Kn$	$IJK-I-J+1$	$\{A\}\{B\}$	$I+J-1$
H <sub>6</sub>	$\mu + \mu_1 + \mu_3$	$\pi_{i..}\pi_{..k}/J$	$f_{i..}f_{..k}/In$	$IJK-I-K+1$	$\{A\}\{C\}$	$I+K-1$
H <sub>7</sub>	$\mu + \mu_2 + \mu_3$	$\pi_{.j.}\pi_{..k}/I$	$f_{.j.}f_{..k}/In$	$IJK-J-K+1$	$\{B\}\{C\}$	$J+K-1$
H <sub>8</sub>	$\mu + \mu_1 + \mu_2 + \mu_{12}$	$\pi_{ij.}/K$	$f_{ij.}/K$	$IJ(K-1)$	$\{AB\}$	$IJ$
H <sub>9</sub>	$\mu + \mu_1 + \mu_3 + \mu_{13}$	$\pi_{i.k}/I$	$f_{i.k}/I$	$IK(J-1)$	$\{AC\}$	$IK$
H <sub>10</sub>	$\mu + \mu_2 + \mu_3 + \mu_{23}$	$\pi_{.jk}/I$	$f_{.jk}/I$	$JK(I-1)$	$\{BC\}$	$JK$
(2) 相互独立モデル						
H <sub>11</sub>	$\mu + \mu_1 + \mu_2 + \mu_3$	$\pi_{i..}\pi_{.j.}\pi_{..k}$	$f_{i..}f_{.j.}f_{..k}/n^2$	$IJK-(I+J+K)+2$	$\{A\}\{B\}\{C\}$	$I+J+K-2$
(3) 多重独立モデル						
H <sub>12</sub>	$\mu + \mu_{12}$	$\pi_{ij.}\pi_{..k}$	$f_{ij.}f_{..k}/n$	$(K-1)(IJ-1)$	$\{AB\}\{C\}$	$IJ+K-1$
H <sub>13</sub>	$\mu + \mu_{13}$	$\pi_{i.k}\pi_{.j.}$	$f_{i.k}f_{.j.}/n$	$(J-1)(IK-1)$	$\{AC\}\{B\}$	$IK+J-1$
H <sub>14</sub>	$\mu + \mu_{23}$	$\pi_{.jk}\pi_{i..}$	$f_{.jk}f_{i..}/n$	$(I-1)(JK-1)$	$\{BC\}\{A\}$	$JK+I-1$
(4) 条件付独立モデル						
H <sub>15</sub>	$\mu + \mu_{12} + \mu_{13}$	$\pi_{ij.}\pi_{i.k}/\pi_{i..}$	$f_{ij.}f_{i.k}/f_{i..}$	$I(J-1)(K-1)$	$\{AB\}\{AC\}\{A\}$	$IJ+IK-I$
H <sub>16</sub>	$\mu + \mu_{12} + \mu_{23}$	$\pi_{ij.}\pi_{.jk}/\pi_{.j.}$	$f_{ij.}f_{.jk}/f_{.j.}$	$J(I-1)(K-1)$	$\{AB\}\{BC\}\{B\}$	$IJ+JK-J$
H <sub>17</sub>	$\mu + \mu_{13} + \mu_{23}$	$\pi_{i.k}\pi_{.jk}/\pi_{..k}$	$f_{i.k}f_{.jk}/f_{..k}$	$K(I-1)(J-1)$	$\{AC\}\{BC\}\{C\}$	$IK+JK-K$
(5) 対連関モデル						
H <sub>18</sub>	$\mu + \mu_{12} + \mu_{13} + \mu_{23}$	—	—	$(I-1)(J-1)(K-1)$	$\{AB\}\{AC\}\{BC\}$	$IJ+IK+JK-I-J-K+1$
(6) 飽和モデル						
H <sub>19</sub>	$\mu + \mu_{12} + \mu_{13} + \mu_{23} + \mu_{123}$	—	$f_{ijk}$	0	$\{ABC\}$	$IJK$
$\mu_{.} = \mu + \mu_1 + \mu_2 + \mu_3$			$df+r=IJK$			

$f_{ijk}$ ,  $\{AB\}=f_{ij.}=\sum_k f_{ijk}$  である。同様の意味で Bishop et al<sup>7)</sup> は  $C_1, C_{12}$  の記号を用いている。

### (3) 多重独立モデル (Multiple Independence Model), H<sub>12</sub>~H<sub>14</sub>

要因  $A, B$  と  $C$  が独立性を示す場合であって,  $P(A \cap B \cap C) = P(A \cap B)P(C)$ ,  $\pi_{ijk} = \pi_{ij.}\pi_{..k}$  の構造を示す。3次元の分割表を  $A$  と  $B$  のカテゴリーを合併して1要因とし, それと  $C$  との2次元分割表に還元したとき独立性が示される場合である。従って周辺度数を用いて  $\{AB\}$ ,  $\{C\}$ , 又は独立性の構造から  $(AB \perp\!\!\!\perp C)$  等の記号で略記される。 $A, B, C$  を循環させて  $(AC \perp\!\!\!\perp B)$ ,  $(BC \perp\!\!\!\perp A)$  の3通りが存在する。

(4) 条件付独立モデル (Conditional Independence Model),  $H_{15} \sim H_{17}$ 

要因  $B, C$  間の独立性が  $A$  の条件付のもとで示される場合であって  $P(A \cap B \cap C) = P(A)P(B|A)P(C|A) = P(A \cap B \cap C)P(A \cap C)/P(A)$ ,  $\pi_{ijk} = \pi_{ij\cdot} \pi_{i\cdot k} / \pi_{i\cdot\cdot}$  の構造を示す。

3次元の分割表でカテゴリー  $A, B$  を合併, カテゴリー  $A, C$  を合併して2次元の分割表に還元したとき独立性が示される場合である。周辺度数を用いれば  $\{AB\} \{AC\} \{A\}$  であり, 独立性の構造からは  $(B \perp\!\!\!\perp C | A)$  で示される。 $A, B, C$  を循環させて  $(A \perp\!\!\!\perp C | B), (A \perp\!\!\!\perp B | C)$  の3通りが存在する。

(5) 対連関モデル (Pair-Wise Association Model),  $H_{18}$ 

3要因  $A, B, C$  が相互に関連する場合, 即ち  $A$  と  $B$  が関連,  $A$  と  $C$  が関連,  $B$  と  $C$  が関連する場合である。周辺度数  $\{AB\}, \{AC\}, \{BC\}$  の拘束はあるものの  $\pi_{ijk}$  を周辺比率の積では表現できない。従って反復法により求める。

Newton-Raphson 法による反復のテクニックは次のように進めればよい。任意の初期値, 例えば  $F_{ijk}^{(0)} = \{1, 1, \dots, 1\}$  から出発し

$$F_{ijk}^{(p+1)} = \frac{F_{ijk}^{(p)} f_{\cdot jk}}{F_{\cdot jk}} \quad (3.5)$$

より  $(p+1)$  サイクルの  $F_{ijk}$  を求める。収束の判定条件は  $|F_{ijk}^{(p+1)} - F_{ijk}^{(p)}| < \varepsilon$  でもよいし,  $|F_{\cdot jk}^{(p)} - f_{\cdot jk}| < \varepsilon$  でもよい。但し添字  $\cdot jk$  に続いて  $i \cdot k, ij \cdot$  についても試みなければならない。収束は速く, 経験的には 5~10 サイクルでは十分のようである。

(注) (3), (4)の機械的拡張から  $P(A \cap B \cap C) = P(B|A)P(C|B)P(A|C) = P(C|A)P(B|C)P(A|B)$

$$\pi_{ijk} = \frac{\pi_{ij\cdot} \pi_{i\cdot k} \pi_{\cdot jk}}{\pi_{i\cdot\cdot} \pi_{\cdot j\cdot} \pi_{\cdot\cdot k}} \quad (3.6)$$

の構造を想定したとき,  $\mu$ -項から LLM は代数的に導くことはできた。従ってその限りにおいては, この構造式は肯定しうが, 周辺度数の拘束条件  $\{AB\}, \{AC\}, \{BC\}$  が満足されないのので (3.6) 式は容認できない。Everitt<sup>9)</sup>, Payne<sup>19)</sup> らも周辺度数の積で表現できないので反復法により求めることを提案している。

(6) 飽和モデル (Saturation Model),  $H_{19}$ 

2次元分割表の場合と同じく  $\hat{F}_{ijk} = f_{ijk}$ ,  $\{ABC\}$  の場合である。残差部分を  $\mu_{123}(ijk)$  としてモデルに組込んでいるので独立性の検定は不要である。この場合, 自由度も 0 である。

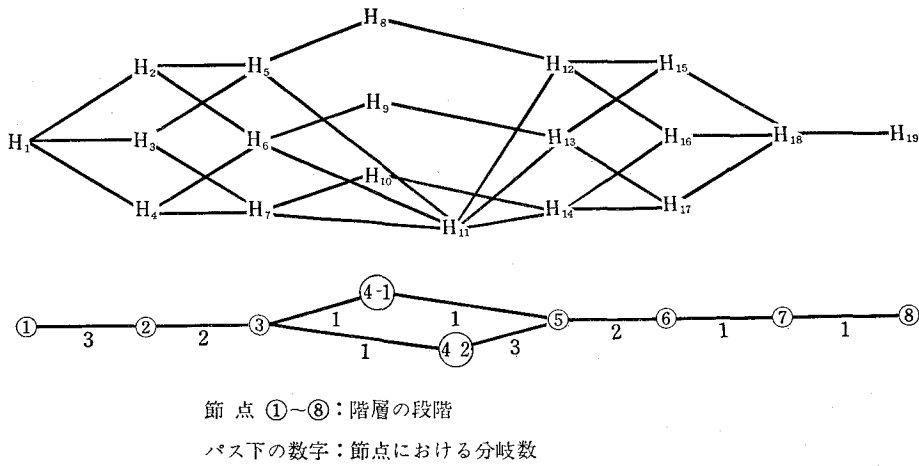
以上の各モデルの各々の場合について, 求められた  $\hat{F}_{ijk}$  と  $f_{ijk}$  との適合度検定を行なう。その際検定統計量として尤度比統計量

$$G^2 = -2 \log \sum_i \sum_j \sum_k f_{ijk} \log \frac{\hat{F}_{ijk}}{f_{ijk}} \quad (3.7)$$

を用い、これが当該モデルのもとで近似的に自由度=df の  $\chi^2$ -分布に従うことを利用する。自由度は表 2 中の第 5 欄に示すごとく  $\mu$ -項で 0 と見なしたものの数である。例えば  $H_{17}$  において、 $\mu$ -項は  $(1+I+J+IJ+IK+JK+1)$  のうち、 $\mu_{12}(ij)$ ,  $\mu_{13}(ik)$ ,  $\mu_{23}(jk)$ ,  $\mu_{123}(ijk)$  は 0 と見なしているものでその個数は  $(I-1)(J-1)+(I-1)(K-1)+(J-1)(K-1)+(I-1)(J-1)(K-1)=IJK-I-J-K+2$  である。 $\mu$ -項で 0 と見なしていないものは  $\mu, \mu_1(i), \mu_2(j), \mu_3(k)$  であり、その個数は  $1+(I-1)+(J-1)+(K-1)=I+J+K-2$  である。これと自由度を加えれば  $IJK$  となり、細胞数  $P$  に一致する。この関係は  $H_1 \sim H_{19}$  のすべてのモデルについて当てはまる。

3. 最終適合モデルの決定

第 2 法で解くとき、 $H_1 \sim H_{19}$  の中からあてはめのよいモデルを拾い出し最終的に適合するモデル (Final Fitted Model) を決定しなければならない。その判定に際して LLM が階層構造 (Hierarchy) をなしていることに注目しておけば最終判定が容易となる。



節点 ①～⑧：階層の段階  
パス下の数字：節点における分岐数

図 1 モデルの階層構造

(3.4) 式の LLM には下位の  $\mu$ -項が 0 のとき、それに関わる上位の  $\mu$ -項も 0 となる性質があり、その故に図 1 のごとき階層構造を示している。例えば  $\mu_1(i)=0$  のとき、上位の  $\mu_{12}(ij)=\mu_{13}(ik)=\mu_{123}(ijk)=0$  である。又  $\mu_1(i), \mu_2(j)$  に関していえば図 1 より

$$H_1 \subset H_2 \subset H_5 \subset \begin{matrix} H_8 \\ H_{11} \end{matrix} \subset H_{12} \subset H_{15} \subset H_{18} \subset H_{19}$$

のごとき包摂関係が示されることになる。

この性質を用いれば、最終モデルの発見は次の 4 つのステップを踏んで探して行けばよいことがわかる (表 3)。

ステップ1  $H_1 \sim H_{19}$  のすべてのモデルが適合するとき、 $H_1$  が最終適合モデルである。

表3 最終適合モデル選択のためのアルゴリズム

番 号	ステップ 1	ステップ 2	ステップ 3	ステップ 4	ステップ 5	ステップ 6
$H_1$	○	×	×	×	×	その他
$H_2$ } $H_{10}$	○	○又は× 表4参照	×	×	×	
$H_{11}$	○	○ or ×	○	×	×	
$H_{12}$ } $H_{18}$	○	○	○	○又は× 表5参照	×	
$H_{19}$	○	○	○	○	○	
最 終	$H_1$	$H_2 \sim H_{10}$	$H_{11}$	$H_{12} \sim H_{18}$	$H_{19}$	$H_{20}$

○: モデル適合 (5%水準)

×: モデル不適合 (5%水準)

ステップ2  $H_1$  が不適合,  $H_{11} \sim H_{19}$  がすべて適合,  $H_2 \sim H_{10}$  の中に適, 不適のものがあるとき,  $H_2 \sim H_{10}$  の中から探せばよい。

ステップ3  $H_1 \sim H_{10}$  がすべて不適合,  $H_{11} \sim H_{19}$  がすべて適合のとき,  $H_{11}$  が最終適合モデルである。

ステップ4  $H_1 \sim H_{11}$  がすべて不適合,  $H_{12} \sim H_{18}$  の中に適, 不適のモデルがあるとき,  $H_{12} \sim H_{18}$  の中から探せばよい。

ステップ5  $H_1 \sim H_{18}$  が不適合,  $H_{19}$  のみ適合のとき  $H_{19}$  が最終適合モデル。

ステップ6 その他の場合, 機械(論理)的判定は困難で他の情報を加味して総合的に行なう。

ステップ1, 3, 5における判定は容易であり, ステップ2, 4の場合に判定のアルゴリズムが複雑となる。このとき手順を一覧して表示すれば, 表4, 表5, である。コンピュータ・プログラミングに際しては  $H_1 \sim H_{19}$  の適合, 不適合の結果のみを記憶させておけばよく, 表4, 表5のアルゴリズムは FORTRAN 語でも容易に書ける。

問題は  $H_1 \sim H_{19}$  のいずれにも判定できない場合である。この場合にはケースに応じて総合的になさざるを得ないのでプログラミングは容易ではなく, 判定不能という意味で  $H_{20}$  とし, 事後他の情報を用いつつ人間が判断することとする。そのような場合は稀であったが散見されたケースについて, 次のごとき基準を導入すれば判定は可能であった。

(1) 判定に矛盾をきたした  $H_i$  の有意水準を調べ  $\alpha=0.05$  を  $\alpha=0.01$  に上げて  $H_i$  を適合(○)にするにするか, 又は  $\alpha=0.10$  に落として不適合(×)にすることにより判定の矛

表 4 ステップ 2 のアルゴリズム (H<sub>2</sub>~H<sub>10</sub>)

番 号	LLM	モデルの適合状況								
H <sub>2</sub>	$\mu + \mu_1$	○	×	×	×	×	×	×	×	×
H <sub>3</sub>	$\mu + \mu_2$	×	○	×	×	×	×	×	×	×
H <sub>4</sub>	$\mu + \mu_3$	×	×	○	×	×	×	×	×	×
H <sub>5</sub>	$\mu + \mu_1 + \mu_2$	○	○	×	○	×	×	×	×	×
H <sub>6</sub>	$\mu + \mu_1 + \mu_3$	○	×	○	×	○	×	×	×	×
H <sub>7</sub>	$\mu + \mu_2 + \mu_3$	×	○	○	×	×	○	×	×	×
H <sub>8</sub>	$\mu + \mu_1 + \mu_2 + \mu_{12}$	○	○	×	○	×	×	○	×	×
H <sub>9</sub>	$\mu + \mu_1 + \mu_3 + \mu_{13}$	○	×	○	×	○	×	×	○	×
H <sub>10</sub>	$\mu + \mu_2 + \mu_3 + \mu_{23}$	×	○	○	×	×	○	×	×	○
H <sub>11</sub>	$\mu + \mu_1 + \mu_2 + \mu_3$	○	○	○	○	○	○	×	×	×
最 終		H <sub>2</sub>	H <sub>3</sub>	H <sub>4</sub>	H <sub>5</sub>	H <sub>6</sub>	H <sub>7</sub>	H <sub>8</sub>	H <sub>9</sub>	H <sub>10</sub>

○：モデル 適合（5％水準）  
×：モデル不適合（5％水準）

表 5 ステップ 4 のアルゴリズム (H<sub>12</sub>~H<sub>18</sub>)

番 号	LLM	モデル適合状況							
H <sub>12</sub>	$\mu. + \mu_{12}$	○	×	×	×	×	×	×	×
H <sub>13</sub>	$\mu. + \mu_{13}$	×	○	×	×	×	×	×	×
H <sub>14</sub>	$\mu. + \mu_{23}$	×	×	○	×	×	×	×	×
H <sub>15</sub>	$\mu. + \mu_{12} + \mu_{13}$	○	○	×	○	×	×	×	×
H <sub>16</sub>	$\mu. + \mu_{12} + \mu_{23}$	○	×	○	×	○	×	×	×
H <sub>17</sub>	$\mu. + \mu_{13} + \mu_{23}$	×	○	○	×	×	○	×	×
H <sub>18</sub>	$\mu. + \mu_{12} + \mu_{13} + \mu_{23}$	○	○	○	○	○	○	○	○
最 終		H <sub>12</sub>	H <sub>13</sub>	H <sub>14</sub>	H <sub>15</sub>	H <sub>16</sub>	H <sub>17</sub>	H <sub>18</sub>	

$\mu. = \mu + \mu_1 + \mu_2 + \mu_3$   
○：モデル 適合（5％水準）  
×：モデル不適合（5％水準）

循を解消する。

(2) 矛盾をきたす 2 つの H<sub>i</sub> の上側確率 P<sup>注)</sup> を比較し、予め設定しておいた有意水準内であれば（例えば  $\alpha < 0.30$  あたりが限度であろうが）P の小さい方を不適合(×)にするか、又は P の大きい方を適合(○)に書きかえて調整する。

(3) n が大のとき、第 1 法の結果を参考にする。

(4) 類似の他の項目（データ）の結果を参考にする。

具体例で示せば次のごとくである。

今、H<sub>12</sub>~H<sub>18</sub> の適合状況が(×××○○×○)であったとする。表 5 によればこれは判



定不能である。このとき矛盾克服の方法は2つ考えられる。先ず  $H_{12}$  の (×) が (○) であれば  $H_{12}$  を最終適合モデルに決定できるので  $H_{12}$  の上側確率を調べ、これが  $P > 0.01$  であれば解決する。他の方法は  $H_{15}, H_{16}$  の適合 (○) のうちいずれが不適合 (×) であればよいので、 $H_{15}$  と  $H_{16}$  の  $P$  を調べ小さい方を、例えば  $\alpha < 0.30$  の範囲内であれば不適合 (×) として処理する。

こうした柔軟な判断はプログラミングが容易ではないので最終判定を保留しておく。出現のケースも少ないし、さほど問題とはならない。

(注) 適合度検定は通常なされるごとく実現値  $\chi^2_0$  と数表からの上側  $100\alpha\%$  点  $\chi^2_{af}(\alpha)$  とを比較する方式をとれば  $\chi^2_0$  分布の数表を予め記憶させておかねばならず、非能率的なので実現値からその上側確率  $P = \Pr\{\chi^2_{af} > G^2_0\}$  を求め、 $P < \alpha$  のとき棄却、 $P > \alpha$  のとき採択するという方式をとった。従って各モデル毎に  $P$  が算出されている。

#### 4. 結果の図示と解釈

LLM で最終的に1つのモデルが決定されれば、それを  $\mu$ -項で書き下し、次いでその解釈がなされなければならない。縮減モデルの場合はさておき、3次元のモデルの場合についての興味の中心は要因間相互の関連であり、 $\mu_{12}(ij), \mu_{13}(ik), \mu_{23}(jk)$ , 及び  $\mu_{123}(ijk)$  の存在とその量の大小である。LLM で残された  $\mu$ -項は要因間の有意な関連が存在することを示唆している。

##### (1) 相互独立モデル $H_{11}$

$\mu_{12} = \mu_{13} = \mu_{23} = \mu_{123} = 0$  であるから、 $A, B, C$  は相互に独立であり、この場合解釈は容易である。

##### (2) 多重独立モデル, $H_{12} \sim H_{14}$

$H_{12}$  の場合について示せば  $\mu_{12} \neq 0, \mu_{13} = \mu_{23} = \mu_{123} = 0$  であり、 $A$  と  $B$  のみが関連を示し他はすべて独立である。

##### (3) 条件付独立モデル, $H_{15} \sim H_{17}$

$H_{17}$  は  $\mu_{12} \neq 0, \mu_{13} \neq 0, \mu_{23} = \mu_{123} = 0$  を示している。 $A-B, A-C$  間に関連があり、 $BC$  間は独立である。データに見られる  $BC$  間の関連は  $A$  を介在することによって生じた関連である。

##### (4) 対連関モデル, $H_{18}$

2次の交互作用がすべて有意で  $\mu_{12}, \mu_{13}, \mu_{23}$  はすべて0でない。3要因  $A, B, C$  が相互に関連を示す場合である。

以上の結果を端的に図示するためには、例えば図2のごときものを考えればよい。

しかしながら解釈を更に深めるためには関連の存在と共にその強度の情報も必要であろう。

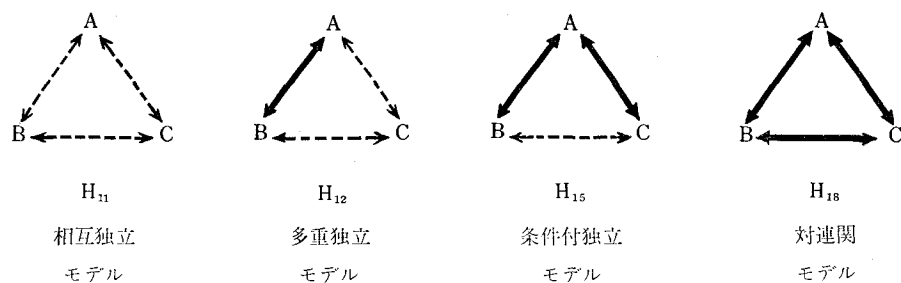


図2 3次元モデルの図示 ( $H_{11} \sim H_{18}$ )

その指標として、ばらつきの程度を試験的に採り上げてみる。

$\mu_{12}(ij)$  に関してそのパラメータ数は  $IJ$  個であり、分散を

$$S_{12}^2 = \frac{1}{IJ} \sum_i \sum_j \mu_{12}^2(ij) \quad (3.8)$$

で求める。束縛条件があるので独立に動きうるのは  $(I-1)(J-1)$  個であるが、平均を 0 とすることと、カテゴリー数が 2 つの場合にも適用可能とするため、 $IJ$  個の方を用いることとする。他の  $\mu_{13}(ik)$ ,  $\mu_{23}(jk)$  に比し  $\mu_{12}(ij)$  の分散が大のとき、 $\log F_{ijk} = \mu_{12}(ij) + \mu_{13}(ik) + \mu_{23}(jk) + \mu_{123}(ijk)$  の加算に際して  $\mu_{12}(ij)$  が大きく効いていることを示している。端的には範囲 (レンジ)

$$R_{12} = \max_{i,j} \mu_{12}(ij) - \min_{i,j} \mu_{12}(ij) \quad (3.9)$$

を求め、 $R_{13}$ ,  $R_{23}$  と比較することも可能であり、ちなみに数量化理論 (林) ではアイテムの予測への貢献度として簡便的に範囲が用いられている。しかし範囲よりも分散の方が  $\mu_{12}(ij)$  のパラメータすべてを用いるので望ましいと思われる。

そしてここでの興味は 2 次の交互作用 3 種間の強度の比較であるので各分散を相対化して新たに  $S_{12}^2, S_{13}^2, S_{23}^2$  とし、それを図 3 のごとき三角座標 (方眼紙) で図示することにする。同図では、正三角形 ABC の高さを 1 とし、点 P から各辺への垂線の長さが分散の相対値と等しくなるよう点 P を選べば、例えば三角形 ABP の面積を大小をもって  $S_{12}^2$  即ち  $\mu_{12}(ij)$  の項の強度を示すことができる。  $S_{13}^2, S_{23}^2$  についても同様であり、LLM の結論から得られた有意な項に斜線を施すことによって、最終適合モデルを直観的に理解することができよう。

## 5. AIC によるモデルの比較

第 2 法で  $H_1 \sim H_{18}$  の各モデルについて適合度検定を行ない最終適合モデルを探索する方法について述べたが、AIC (Akaike Information Criterion) を用いて最適のモデルを発見することも可能である。Akaike, H.<sup>1)</sup>, Sakamoto, Y. and Akaike, H.<sup>24)</sup> によれば AIC は

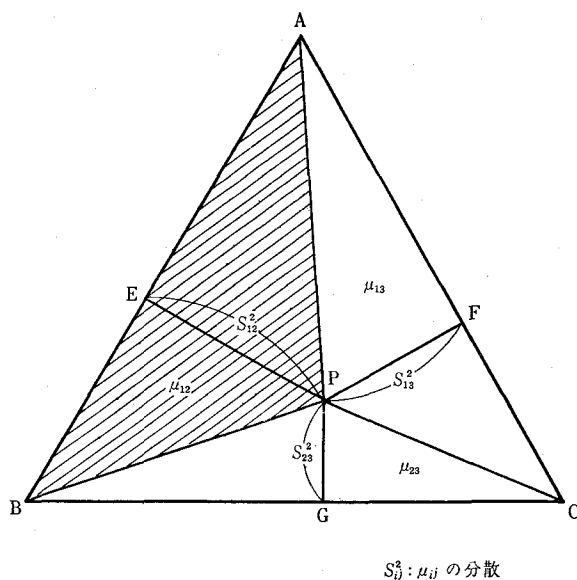


図3 三角座標による交互作用項強度の図示

$$\text{AIC} = -2 \log (\text{最大尤度}) + 2(\text{パラメータ数}) \quad (3.10)$$

で定義され、モデルのあてはめの悪さの指標 (AIC が小さい程あてはめがよい) として用いることができるという。上式の第1項は以下の変形によりあてはめられたモデル下で予測値と観測値とのずれの程度を表す量に還元でき、パラメータを多く用いることによりあてはめを良くし、いくらでも小さくすることができる (データ数だけパラメータを用いればあてはめは完璧である)。しかし、パラメータを多くすることはそれだけあてはめの効率を悪くすることであり、第2項を加えることにより相対的にバランスを企り、より小の労力でより大の精度を確保せんとする、いわば "parsimony の原理" に基づいたモデル評価の方式といえよう。

3次元分割表データに AIC を適用するには以下のごとくに考えればよい。まず、各細胞の頻度  $X_{ijk}$  は多項分布に従うと仮定する。これは無理のない仮定である。そのとき実現値  $f_{ijk}$  を得る確率は、母比率を  $\pi_{ijk}$  とすると

$$P(X_{ijk} = f_{ijk}) = \frac{n!}{\prod_{i,j,k} f_{ijk}!} \prod_{i,j,k} (\pi_{ijk})^{f_{ijk}} \quad (3.11)$$

$$\sum_{i,j,k} f_{ijk} = n, \quad \sum_{i,j,k} \pi_{ijk} = 1$$

であり、これより  $\pi_{ijk}$  の最尤推定量を求めそれを上式に代入して AIC の第1項を求めればよい。今一例として相互独立モデル,  $H_{11}: \pi_{ijk} = \pi_{i..} \pi_{.j.} \pi_{..k}$  の場合を選べば、各周辺母比率の最尤推定量は容易に  $\hat{\pi}_{i..} = f_{i..}/n$ ,  $\hat{\pi}_{.j.} = f_{.j.}/n$ ,  $\hat{\pi}_{..k} = f_{..k}/n$  として求められる。更に  $H_{17}$

のとき,

$$\hat{\pi}_{ijk} = \hat{\pi}_{i..} \hat{\pi}_{.j.} \hat{\pi}_{..k} = \frac{f_{i..} f_{.j.} f_{..k}}{n^3} = \frac{\hat{F}_{ijk}}{n} = \frac{f_{ijk}}{n} \times \frac{\hat{F}_{ijk}}{f_{ijk}} \quad (3.12)$$

と変形できるので AIC の第1項は

$$\begin{aligned} -2 \log(\text{最大尤度}) &= -2 \left( \log \frac{n!}{\prod f_{ijk}!} + \sum f_{ijk} \log \hat{\pi}_{ijk} \right) \\ &= -2 \left( \log \frac{n!}{\prod f_{ijk}!} + \sum f_{ijk} \log \frac{f_{ijk}}{n} \right) - 2 \sum f_{ijk} \log \frac{\hat{F}_{ijk}}{f_{ijk}} \\ &= C + G^2 \end{aligned} \quad (3.13)$$

となる。上の前項 ( ) 内は選択されたモデルに関係しない量のなのでモデル相互の比較に際しては常数  $C$  としてよく、後項は (3.7) 式より  $G^2$  そのものである。そして  $\hat{\pi}_{ijk} = \hat{F}_{ijk}/n$  の変形は  $H_{18}$  以外のすべてのモデルの場合にも可能である。 $H_{18}$  については  $\hat{\pi}_{ijk}$  を代数式で表現できないが反復法で求められた  $F_{ijk}$  を用いればよい。AIC の第2項は  $\hat{F}_{ijk}$  の推定に際して用いたパラメータ数  $r = IJK$ —自由度 (表2, 第7欄) の2倍である。従って  $H_1 \sim H_{19}$  の全モデルを通じて定数  $C$  を除き

$$AIC = G^2 + 2r \quad (3.14)$$

として求めてもよいことがわかる。

前述のごとく  $H_1 \sim H_{19}$  は階層デルであり、その間の包含関係は図1の樹形図に示す通りである。4—1 ( $H_8 \sim H_{10}$ ) を通るパスは  $3 \times 2 \times 1 \times 1 \times 2 \times 1 \times 1 = 12$  通り、4—2 ( $H_{11}$ ) を通るものは  $3 \times 2 \times 1 \times 3 \times 2 \times 1 \times 1 = 36$  通り、併せて 48 通りの階層構造の組合わせが得られることになる。例えば

$$\begin{array}{ccccccccccc} H_1 & \text{---} & H_2 & \text{---} & H_5 & \text{---} & H_8 & \text{---} & H_{12} & \text{---} & H_{15} & \text{---} & H_{18} & \text{---} & H_{19} \\ \mu + \mu_1 & & + \mu_2 & & + \mu_{12} & & + \mu_3 & & + \mu_{13} & & + \mu_{23} & & + \mu_{123} & & \\ & & \vdots & & & & & & & & & & & & \\ & & \vdots & & & & & & & & & & & & \\ H_1 & \text{---} & H_4 & \text{---} & H_7 & \text{---} & H_{11} & \text{---} & H_{14} & \text{---} & H_{17} & \text{---} & H_{18} & \text{---} & H_{19} \\ \mu + \mu_3 & & + \mu_2 & & + \mu_1 & & + \mu_{23} & & + \mu_{13} & & + \mu_{12} & & + \mu_{123} & & \end{array}$$

等であり、これらは加える順序の相違であり、いずれもすべての  $\mu$ -項を含んでいる。各々の場合についてモデル番号がふえるにつれて次々と  $\mu$ -項が加わり (即ちパラメータ数が増加し)、それに応じて  $G^2$  は減少して予測の精度は向上することになる。AIC の最少値 (minimum AIC estimate) を今 MAICE と書けば、各々の場合の MAICE をその階層構造の中での最適モデルとして選択すればよい。

本来 AIC は包含関係のあるモデルの中での最適モデルの発見に用いられるべき性質のものであるが、48 通りの全組合せを通じて MAICE を選ぶことは結局のところ  $H_1 \sim H_{19}$  の

AICを比較してその中で最小のものを最適モデルとして結論づけることと同じである。後にVIで数値例を呈示するが、第2法で選択されたモデルはAICを最少にしている。ただ例外も見られたが、詳細に検討すれば第2法の結論よりも望ましいようであり、AICによる方法を第3法とし、この点については後に考察することとしたい。

(注) 今、 $\pi_{ijk}$ ,  $f_{ijk}$  をまとめて  $\pi$ ,  $f$  と書き、(3.11) 式より尤度関数の対数をとる

$$l(\pi, f) = \log \frac{n!}{\prod f_{ijk}} + \sum_{ijk} f_{ijk} \log \pi_{ijk} - \lambda (\sum_{ijk} \pi_{ijk} - 1) \quad (3.15)$$

但し  $\lambda$ : ラグランジュの乗数

とおく。

$H_{11}$  の場合  $\pi_{ijk} = \pi_{i..} \pi_{.j.} \pi_{..k}$  の関係があるので、これを用いれば

$$\begin{aligned} l(\pi; f) &= c + \sum_i \sum_j \sum_k f_{ijk} (\log \pi_{i..} + \log \pi_{.j.} + \log \pi_{..k}) - \lambda (\sum_i \sum_j \sum_k \pi_{i..} \pi_{.j.} \pi_{..k} - 1) \\ &= c + \sum_i f_{i..} \log \pi_{i..} + \sum_j f_{.j.} \log \pi_{.j.} + \sum_k f_{..k} \log \pi_{..k} - \lambda (\sum_i \pi_{i..} \sum_j \pi_{.j.} \sum_k \pi_{..k} - 1) \end{aligned} \quad (3.16)$$

$$\frac{\partial l}{\partial \pi_{i..}} = \frac{\partial l}{\partial \pi_{.j.}} = \frac{\partial l}{\partial \pi_{..k}} = \frac{\partial l}{\partial \lambda} = 0 \text{ より}$$

$$\begin{cases} \frac{f_{i..}}{\pi_{i..}} - \lambda = \frac{f_{.j.}}{\pi_{.j.}} - \lambda = \frac{f_{..k}}{\pi_{..k}} - \lambda = 0 \\ \sum_i \pi_{i..} = \sum_j \pi_{.j.} = \sum_k \pi_{..k} = 1 \end{cases} \quad (3.17)$$

$$\text{の方程式を解いて } \lambda = n, \hat{\pi}_{i..} = \frac{f_{i..}}{n}, \hat{\pi}_{.j.} = \frac{f_{.j.}}{n}, \hat{\pi}_{..k} = \frac{f_{..k}}{n} \quad (3.18)$$

が得られる。

$H_{11}$  の他のモデルの場合も同様で

$$\hat{\pi}_{ij.} = \frac{f_{ij.}}{n}, \hat{\pi}_{i..k} = \frac{f_{i..k}}{n}, \hat{\pi}_{.jk} = \frac{f_{.jk}}{n} \text{ も容易に求められる。} \quad (3.19)$$

## 6. 階層モデルの加算性

$H_m \subset H_M$  なるモデルに対して

$$G^2(m-M) = G^2(m) - G^2(M) \quad (3.20)$$

$$df(m-M) = df(m) - df(M)$$

なる関係があるとき、 $G^2$  の加算性という。AIC についても同様に  $G^2, r$  について加算性が成立しているので

$$AIC(m-M) = AIC(m) - AIC(M) \quad (3.21)$$

である。今これを具体例で示せば  $\mu_{12}$  を加えることによって得られるモデルの包含関係は  $H_{11} - H_{12}$ ,  $H_{13} - H_{15}$ ,  $H_{14} - H_{16}$ ,  $H_8 - H_5$ ,  $H_{17} - H_{18}$  (いずれも前者に  $\mu_{12}$  を加えれば後者となる) である。但し、 $H_{18}$  を用いる場合には周辺度数から反復法により  $\hat{F}_{ijk}$  を求め

ているので加算性は成立しない。これらの関係は  $\mu_{13}$  に関して  $H_{11}-H_{13}$ ,  $H_{12}-H_{15}$ ,  $H_{14}-H_{17}$ ,  $H_9-H_6$  から,  $\mu_{23}$  に関しては  $H_{11}-H_{14}$ ,  $H_{12}-H_{16}$ ,  $H_{13}-H_{17}$ ,  $H_{10}-H_7$ , から求めたものがすべて一致するところから確認することができる。

$G^2$  は  $\chi^2$ -分布に従い,  $\chi^2$ -変数の加法性から  $G^2(m-M)$  もまた自由度  $df(m-M)$  の  $\chi^2$ -分布に従うと考えてよい。従ってこの  $G^2(m-M)$  を用いて試論ではあるが  $\mu_i$  を 0 と見なしてよいかどうかの検定が可能のように思える。即ち  $G^2(m-M)$  は  $-2 \log L(\hat{\pi}_{ijk}|H_m)/L(\hat{\pi}_{ijk}|H_M)$  と変形できるのでモデル  $H_m$  と  $H_M$  の尤度比統計量に関連した統計量であり, 直観的には両モデル間の距離を表わす量と考えられる。そしてその距離を発生せしめているのはモデル  $H_m$  に加えられたパラメータ  $\mu_i$  であり,  $\mu_i=0$  のとき両モデルは重なり  $G^2(m-M)$  は 0 に近づく。以上のことより, 棄却域として  $R=\{G^2|G^2>\chi^2_{df}(\alpha)\}$  を選べば有意水準  $\alpha$  なる  $H_0: \mu_i=0$ ,  $H_1: \mu_i \neq 0$  (対立仮説) の検定方式を定めることができる。この方式から最適モデルを選択する方法を今後第4法とし, 第2法による結果と比較すればほぼ平行関係が見られた。但しデータ4, 5についてはずれを見たが詳細については後に数値例で考察する。

#### IV コンピュータ・プログラミング

LLM による解法は前述のごとく第1法と第2法があり, いずれもカテゴリー数がふえるとき計算量は尨大となるのでコンピュータの活用を俟ってはじめて可能な方法である。

2次元分割表の場合には伝統的な独立性の検定を実施し, 独立性が否定されたとき, 細胞の有意差の検討を行なえば  $\mu_{12}(ij)$  の分析が簡単にできるので, コンピュータによる処理は3次元分割表データの場合とする。

処理の流れの概要は図4であり, 同図に従い要点を解説する。なお, 計算結果のプリントは常時行なわれるので特に出力の表示はしていない。

##### 1. メインルーチン

メインルーチンはデータの読み込みであり, カテゴリー数  $I, J, K$  と頻数  $f_{ijk}$  を入力した後, 周辺度数  $f_{i..}$ ,  $f_{ij.}$  もここで求めておく。1次元周辺度数は  $f_{i..}$  の他  $i$  を循環させて  $f_{.j.}$ ,  $f_{..k}$  があり, 同様に2次元周辺度数も  $f_{ij.}$  の他  $f_{i.k}$ ,  $f_{.jk}$  がある。そして  $f_{ijk}$  の対数値を  $\mathbf{f}$  とし, 配列に格納しておく。

2つの解法はサブルーチンとして独立しているので解法を選択することができるが, 両法を併用し, 解を比較することももちろん可能である。

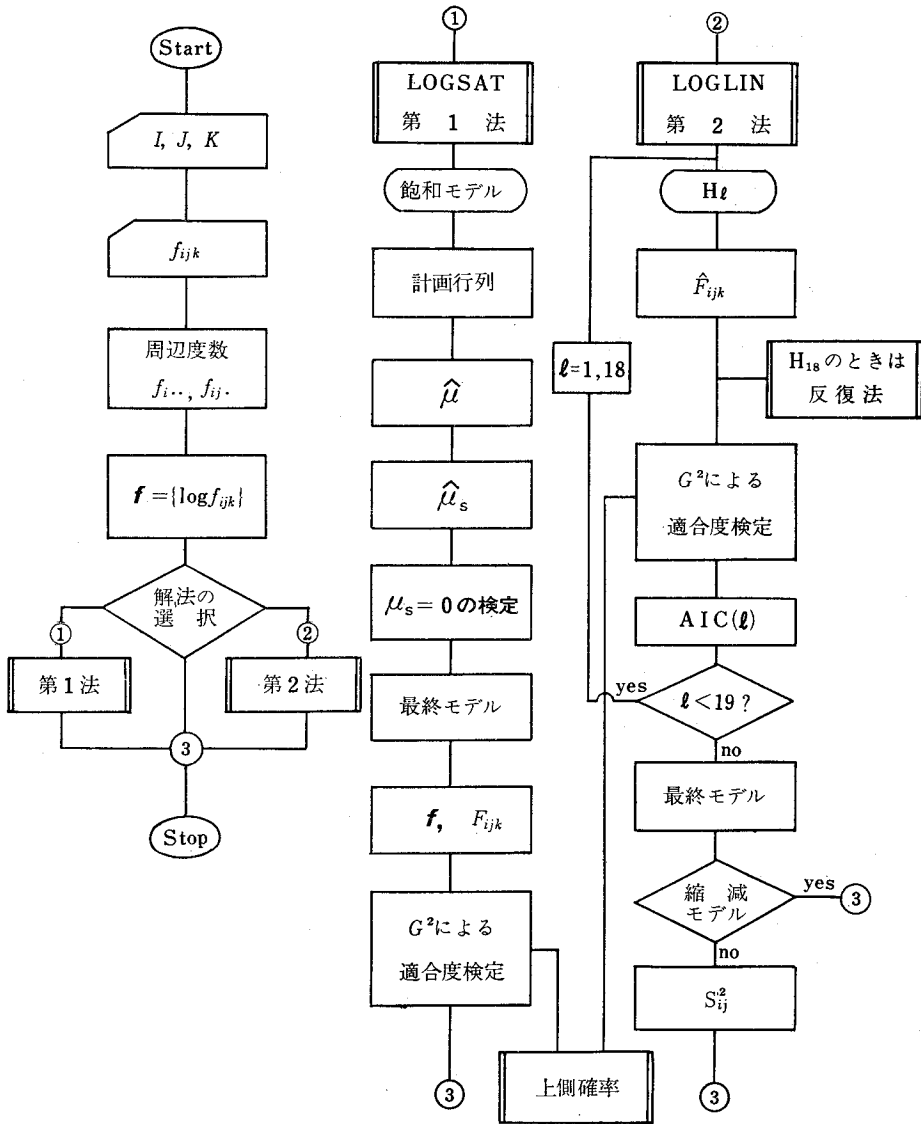


図4 LLM解法のフローチャート

## 2. 第1法, サブルーチン LOGSAT

飽和モデルのもとで計画行列を作成し, 連立方程式を解いて解 $\hat{\mu}$ を求める。カテゴリー数は一般の場合であるので  $\hat{\mu} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\mathbf{f}$  より求める。 $\hat{\mu}$ を標準化して $\hat{\mu}_s$ とし, 検定の結果, 有意のものを残して適合モデルを決定する。

最終モデルより $\hat{\mathbf{f}}$ を求め, 対数を外して $\hat{F}_{ijk}$ を算出し, 観測度数 $f_{ijk}$ との適合度検定を行なっておく。 $\chi^2$ -値と $G^2$ -値(尤度比統計量)を計算するが, 検定に使用するのは $G^2$ -値である。そしてモデルの適合状況を検討するため, 所与の自由度を用いて $\chi^2$ -分布の上側確率 $P$ を求め,  $P > 0.05$ を確認しておく。

## 3. 第2法, サブルーチン LOGLIN

第2法では $H_l (l=1\sim 19)$ の19種数のモデルが考えられるので, 各々の場合について $\hat{F}_{ijk}$ を求める。 $H_{18}$ (対連関モデル)の場合には反復法によるため, 計算が複雑であるのでサブルーチンとしておくが, その他の場合には周辺度数の積から簡単に求められる。

$H_l$ の適合状況から最終適合モデルを決定しなければならないが, これは縮減モデル(1, 2次元)と3次元モデルに分け, 表3～表5のアルゴリズムを用いて行なう。その結果はAIC( $l$ )-値により確認することができる。

最終モデルが縮減モデルでないとき, 2次元の交互作用 $\mu_{ij}$ の分散を求め, 図示の準備をする。周辺装置にX-Yプロッターが使用できるとき, 作図も可能であるが, 値さえ算出されていれば手書きも困難でない。

FORTRAN語でプログラミングした結果では第1法のサブルーチンが約420語, 第2法が約550語であった。

## V 数 値 例

プログラムチェック及びデバッグ用として成書に紹介されている以下のデータを用いたが, 本節では出力結果に基づいて更に手法の検討を行なうことにする。データはいずれも $A, B, C$ の3次元分割表である。

データ 1 (Payne<sup>19)</sup>, p. 106),  $I=2, J=2, K=3$ 。

$A$  投票:  $A_1$  保守党,  $A_2$  労働党,

$B$  性:  $B_1$  男,  $B_2$  女,

$C$  階級:  $C_1$  上流,  $C_2$  中流,  $C_3$  労働者。



観測度数  $f_{ijk}$  を第3添字から先に変化させ、1次元の配列で示す。

(82, 79, 118, 96, 101, 155, 30, 53, 252, 30, 34, 227),  $n=1257$ 。

データ 2 (Everitt<sup>9)</sup>, p. 67),  $I=2, J=2, K=3$ 。

A 集団行動:  $A_1$  適応,  $A_2$  不適応,

B リスク指標:  $B_1$  低,  $B_2$  高,

C 学校環境:  $C_1$  不良,  $C_2$  中,  $C_3$  良,

(16, 15, 5, 7, 34, 3, 1, 3, 1, 1, 8, 3),  $n=97$ 。

データ 3 (Bishop<sup>7)</sup> p. 87),  $I=2, J=2, K=2$ 。

A 植物の生育:  $A_1$  生,  $A_2$  枯,

B 植えた時期:  $B_1, B_2$ ,

C 切断の長さ:  $C_1, C_2$ ,

(156, 107, 84, 31, 84, 133, 156, 209),  $n=960$ 。

データ 4 (Bishop<sup>7)</sup>, p. 90),  $I=2, J=2, K=2$ 。

A 食 中 毒:  $A_1$  症状あり,  $A_2$  症状なし,

B ポテトサラダ:  $B_1$  食べた,  $B_2$  食べない,

C か に の 肉:  $C_1$  食べた,  $C_2$  食べない,

(120, 22, 4, 0, 80, 24, 31, 23),  $n=308$ 。

データ 4 (Everitt<sup>9)</sup>, p. 95),  $I=4, J=4, K=2$ 。

A 血 圧:  $A_1 \sim 122, A_2 127 \sim 146, A_3 147 \sim 166, A_4 167 \sim$ ,

B 血清コレステロール:  $B_1 \sim 200, B_2 200 \sim 219, B_3 220 \sim 259, B_4 260 \sim$ ,

C 冠 状 動 脈 異 状:  $C_1$  あり,  $C_2$  なし,

(2, 117, 3, 121, 3, 47, 4, 22, 3, 85, 2, 98, 1, 43, 3, 20, 8, 119, 11, 209, 6, 68, 6, 43, 7, 67, 12, 99, 11, 46, 11, 33),  $n=1330$ 。

# 1. 第1法による解

入力データから計画行列を成作し、連立方程式を解いて  $\hat{\mu}$  を求めたのか表6である。 $\hat{\mu}$  を標準化したものが  $\hat{\mu}_s$  であり、検定の結果有意なもの(\*印を付した)を残して LLM を書くことができる。

データ1の場合、B の単純効果 ( $\mu_2$ ) が有意とならないが2次の交互作用 ( $\mu_{12}$ ) が有意

表 6 第 1 法結果 ( $\mu_s$ )

	(1) データ 1		(2) データ 2		(3) データ 3		(4) データ 4	
$\mu$ -項	$\hat{\mu}$	$\hat{\mu}_s$	$\hat{\mu}$	$\hat{\mu}_s$	$\hat{\mu}$	$\hat{\mu}_s$	$\hat{\mu}$	$\hat{\mu}_s$
$\mu$	4.419	123.81*	1.495	8.340*	4.663	123.97*	2.864	14.781*
$\mu_1(1\cdot\cdot)$	0.210	5.876*	0.782	4.363*	-0.266	-7.065*	-0.685	-3.534*
$\mu_2(\cdot 1\cdot)$	-0.011	-0.298	-0.130	-0.725	0.098	2.615*	1.009	5.209*
$\mu_3(\cdot\cdot 1)$	-0.475	-8.489*	-0.315	-1.128	0.078	2.067*	0.670	3.457*
$\mu_3(\cdot\cdot 2)$	-0.298	-5.689*	0.858	4.107*	—	—	—	—
$\mu_{12}(11\cdot)$	-0.102	-2.859*	0.217	1.208	0.366	9.734*	0.764	3.945*
$\mu_{13}(1\cdot 1)$	0.332	5.939*	0.398	1.423	0.266	7.065*	0.299	1.544
$\mu_{13}(1\cdot 2)$	0.162	3.095*	-0.018	-0.086	—	—	—	—
$\mu_{23}(\cdot 11)$	-0.029	-0.514	0.337	1.205	-0.098	-2.615*	0.047	0.244
$\mu_{23}(\cdot 21)$	0.060	1.148	-0.320	-1.530	—	—	—	—
$\mu_{123}(111)$	0.063	1.119	-0.010	-0.035	-0.057	-1.505	-0.177	-0.913
$\mu_{123}(112)$	-0.070	-1.342	-0.176	-0.842	—	—	—	—
分散			$\mu_{ij}$ : not significant					
$S_{12}^2$	0.102	(0.204)	0.217	(0.270)	0.336	(0.501)	0.764	(0.688)
$S_{13}^2$	0.357	(0.712)	0.318	(0.396)	0.266	(0.364)	0.299	(0.269)
$S_{23}^2$	0.043	(0.085)	0.268	(0.334)	0.098	(0.135)	0.047	(0.043)
最終モデル	$H_{15}$ $\mu_{\cdot} + \mu_{12} + \mu_{13}$		$H_6$ $\mu + \mu_1 + \mu_3$		$H_{18}$ $\mu_{\cdot} + \mu_{12} + \mu_{13} + \mu_{23}$		$H_{12}$ $\mu_{\cdot} + \mu_{12}$	

$\mu_{\cdot}$ :  $\mu + \mu_1 + \mu_2 + \mu_3$

\*: 5%水準で有意

( ): 分散の相対値

データ 5 の結果については省略

であるので、階層構造には矛盾するがここでは残すことにする。

$\mu_{ij}$  間の強度を比較するために、その分散  $S_{ij}^2$  を求めたのが同表の下欄である。 $\hat{\mu}$  の検定結果とはほぼ平行関係にあることがわかる。例えばデータ 1 で有意な  $\mu_{12}$ ,  $\mu_{13}$  は分散も大きく、有意性を示さない  $\mu_{23}$  は小さい値でその相対値も 8.5% にしかすぎない。しかし両者の平行関係は分散の検討をもって検定に代えうることを意味しない。分散はあくまでも同一データ内の比較に用いるべき性質のものである。例えばデータ 1 の  $S_{12}^2$  は 20.4% で有意であり、データ 4 の  $S_{13}^2$  は 26.9% でも有意でないが、これは矛盾ではない。又データ 2 は縮減モデルとの結論が出ているので、そのことを考慮に入れて分散の検討を行うべきである。

データ 5 は  $\mu$ -項が  $4 \times 4 \times 2 = 32$  にも及ぶので結果の作表は省略するが、 $\mu_{12}$  が有意でなく  $H_{17}$ :  $\mu_{\cdot} + \mu_{13} + \mu_{23}$  の結論であった (表 9)。

次に  $\mu$ -項を用いて求めた  $\hat{F}_{ijk}$  が  $f_{ijk}$  と一致するかを確かめてみる。LLM より一例として  $\hat{F}_{111}$  を求めると

$$\begin{aligned}\log \hat{F}_{111} &= \hat{\mu} + \hat{\mu}_1(1) + \hat{\mu}_2(1) + \hat{\mu}_3(1) + \hat{\mu}_{12}(11) + \hat{\mu}_{13}(11) + \hat{\mu}_{23}(11) + \hat{\mu}_{123}(111) \\ &= 4.419 + 0.210 - 0.011 - 0.475 - 0.102 + 0.332 - 0.029 + 0.063 = 4.407\end{aligned}$$

であるから  $\hat{F}_{111} = e^{4.407} = 82.023$  となり  $f_{ijk} = 82$  と試算誤差の範囲内で一致している。更に飽和モデル以外の場合の例として  $H_{11}$  : 相互独立モデルを考えてみる。LLM は

$$\log F_{ijk} = \mu + \mu_1(i) + \mu_2(j) + \mu_3(k)$$

であるから

$$\begin{aligned}\log \hat{F}_{111} &= \hat{\mu} + \hat{\mu}_1(1) + \hat{\mu}_2(1) + \hat{\mu}_3(1) \\ &= 4.419 + 0.210 - 0.011 - 0.475 = 4.143\end{aligned}$$

となり  $\hat{F}_{111} = 62.99$  である。すべての細胞について繰返し、かつ周辺度数を求めると  $\hat{F}_{1..} = 723.15$ ,  $\hat{F}_{.1.} = 592.56$ ,  $\hat{F}_{..1} = 211.08$ ,  $\hat{F}_{...} = 1198.3$  となるから

$$\hat{F}_{111} = \frac{723.15 \times 592.56 \times 211.08}{1198.3^2} = 62.99$$

であり両者は一致している。つまり  $H_{11}$  の構造式を満たしている。

しかしながら注意を要するのは、第2法では最尤推定量を用いて  $\hat{F}_{i..} = f_{i..}$  等としているので

$$\hat{F}_{111} = \frac{f_{1..} \times f_{.1.} \times f_{..1}}{n^2} = \frac{631 \times 614 \times 238}{1257^2} = 58.36$$

とは一致しないことである。

$\log \hat{F}_{ijk}$  は上記のごとく手計算では面倒であるがコンピュータでは計画行列を用いて(2.9)式から  $\mathbf{f} = \mathbf{X}'\hat{\boldsymbol{\mu}}$  で一度に求められる。第2法で解く場合には構造式を問題にするよりもLLMで $\mu$ -項の加算的性質を検討の方が簡明である。

## 2. 第2法による解

第2法の適合度検定では  $G^2$  を用いるがその値は表7に示す通りである。但しデータ2～5については数字の検討が目的ではないので煩雑さを避け  $H_{11} \sim H_{18}$  のみを示すにとどめる。いずれも  $H_1 \sim H_{10}$  は不適合であった。表5のアルゴリズムを用いた最終判定結果は表7, 下欄である。これらのデータの場合についてはステップ6を用いることなく機械的判定が可能であった。

第2法の結果を MAICE による結論と比較するとき、データ1～3については合致したものの、データ4,5においてくい違いを示した。これは後にV-5で考察する。

表7における  $P$  は  $G^2 \sim \chi^2_{df}$  として実理値  $G^2_0$  を得たときの上側確率である。 $P > 0.05$  のときモデル適合として○,  $P < 0.05$  のとき不適合として×に記号で示されている。第2法では○か×かの結果しか用いないが、ステップ6では  $P$  を参考に用いなければならないし、

表7 第2法 ( $G^2$ ) および第3法 (ATC) の結果

(1) データ1

H	$G^2_0$	df	Fit	P	AIC	r
1	535.51	11	×	0 <sup>1)</sup>	537.51	1
2	535.49	10	×	0	539.49	2
3	534.84	10	×	0	542.84	2
4	165.73	9	×	0	171.73	3
5	534.82	9	×	0	540.82	3
6	165.71	8	×	0	173.71	4
7	165.06	8	×	0	173.06	4
8	523.93	8	×	0	531.93	4
9	14.32	6	×	0.026	155.16	6
10	164.68	6	×	0	176.68	6
11	165.04	7	×	0	175.04	5
12	154.15	6	×	0	166.15	6
13	13.65	5	×	0.018	27.65	7
14	164.66	5	×	0	178.66	7
15	2.76	4	○	0.599	△18.76	8
16	153.77	4	×	0	169.77	8
17	13.27	3	×	0.004	31.27	9
18	1.87	2	○		21.87	10
最 終	H <sub>15</sub>				H <sub>15</sub>	

1) 0:  $P < 0.001$     △: MAICE

(2) データ2

H	$G^2_0$	df	Fit	P	AIC	r
11 <sup>2)</sup>	16.42	7	×	0.022	26.42	5
12	14.98	6	×	0.020	26.98	6
13	12.76	5	×	0.026	26.76	7
14	5.56	5	○	0.351	△19.56	7
15	11.32	4	×	0.023	27.32	8
16	4.12	4	○	0.390	20.12	8
17	1.90	3	○	0.593	19.90	9
18	0.94	2	○	0.624	20.94	10
最 終	H <sub>14</sub>				H <sub>14</sub>	

2) データ2～5については H<sub>1</sub>～H<sub>10</sub> の結果は省略

## (3) データ 3

H	$G^2_0$	df	Fit	P	AIC	r
11	151.02	4	×	0	159.02	4
12	53.44	3	×	0	63.44	5
13	105.18	3	×	0	115.18	5
14	151.02	3	×	0	161.02	5
15	7.60	2	×	0.022	19.60	6
16	53.44	2	×	0	65.44	6
17	105.18	2	×	0	117.18	6
18	2.29	1	○	0.130	△16.29	7
最 終	$H_{18}$				$H_{18}$	

## (4) データ 4

H	$G^2_0$	df	Fit	P	AIC	r
11	69.82	4	×	0	77.82	4
12	15.69	3	×	0.001	25.69	5
13	60.58	3	×	0	70.58	5
14	58.82	3	×	0	68.82	5
15	6.45	2	×	0.040	18.45	6
16	4.69	2	○	0.096	16.69	6
17	49.58	2	×	0	61.58	6
18	1.13	1	○	0.289	△15.13	7
最 終	$H_{16}$				$H_{18}$	

## (5) データ 5

H	$G^2_0$	df	Fit	P	AIC	r
11	78.96	24	×	0	94.96	8
12	54.51	15	×	0	68.51	17
13	48.51	21	×	0.001	70.51	11
14	54.85	21	×	0	76.85	11
15	24.06	12	×	0.020	64.06	20
16	30.41	12	×	0.002	70.41	20
17	24.40	18	○	0.142	52.40	14
18	4.77	9	○	0.854	△50.77	23
最 終	$H_{17}$				$H_{18}$	

モデルの適合状況を詳細に検討する際にはこの  $P$  値が有効な情報をもたらすことになる。

## 3. 階層構造及び AIC (第3法)

図1の階層構造をデータ1の場合の数値例で示せば図5のごとくである。同図より

$H_{11}-H_{12}-H_{15}-H_{18}$ ,  $H_{11}-H_{12}-H_{16}-H_{18}$ , ……………,  $H_{11}-H_{14}-H_{17}-H_{18}$ 等の併せて



ですべて一致しているが

$$G^2(17) - G^2(18) = 13.27 - 1.87 = 11.4$$

の結果はこれらとくい違っている。自由度についても

$$df(11) - df(12) = df(13) - df(15) = df(14) - df(16) = 1$$

である。AIC についても同様に

$$AIC(11) - AIC(12) = 175.04 - 166.15 = 8.89$$

$$AIC(13) - AIC(15) = 27.65 - 18.76 = 8.89$$

$$AIC(14) - AIC(16) = 178.66 - 169.77 = 8.89$$

で共に同じ値である。そしてこれらの関係は  $\mu_{13}, \mu_{23}$  についても、又他のデータについても妥当している。

然るに伝統的に用いられて来た  $\chi^2$ -統計量では  $\mu_{12}$  の場合（コンピュータ出力は省略しているが）、

$$\chi^2(11) - \chi^2(12) = 159.82 - 149.38 = 10.44$$

$$\chi^2(13) - \chi^2(15) = 13.60 - 2.74 = 10.86$$

$$\chi^2(14) - \chi^2(16) = 159.17 - 149.59 = 9.58$$

となり加算性は近似的にしか成立していない。この故に本稿では適合度検定に際しては  $G^2$ -統計量を用いることとした。

表8  $G^2(m-M)$  による  $\mu_{ij}$  の検定結果

	データ 1	データ 2	データ 3	データ 4	データ 5
$\mu_{12}$ $G^2$	10.89*	1.44	97.58*	54.13*	24.45*
$df$	1	1	1	1	9
$\mu_{13}$ $G^2$	151.39*	3.66	45.84*	9.24*	30.45*
$df$	2	2	1	1	3
$\mu_{23}$ $G^2$	0.38	10.86*	0.00	11.00*	24.11*
$df$	2	2	1	1	3

$\mu_{12} : H_{11} - H_{12}, H_{13} - H_{15}, H_{14} - H_{16}, H_8 - H_5$

$\mu_{13} : H_{11} - H_{13}, H_{12} - H_{15}, H_{14} - H_{17}, H_9 - H_6$

$\mu_{23} : H_{11} - H_{14}, H_{12} - H_{16}, H_{13} - H_{17}, H_{10} - H_7$

\*:  $G^2(m-M) \in R$ , 5%水準で  $H_0 : \mu_{ij} = 0$  棄却

## 5. $G^2(m-M)$ による検定 (第4法)

前項の方法によりデータ毎に2次の交互作用項について  $G^2(m-M)$  を求め一覧表にしたのが表8である。\*印を付したものは  $G_0^2 \in R$  となり、 $\mu_i \neq 0$  と見なしうるもの、即ちモデル  $H_m$  と  $H_M$  間の距離が遠く交互作用を無視できないものである。第2法とほぼ同様の結論が得られているが、第1～4法を結合したものが表9である。

表9 第1～4法の結果比較

$\mu_{ij}$	方法	データ 1 $n=1257$	データ 2 $n=97$	データ 3 $n=960$	データ 4 $n=308$	データ 5 $n=1330$
$\mu_{12}$	1	*	0	*	*	0 ⑤
	2	*	0	*	*	0
	3	*	0	*	*	*
	4	*	0	*	*	*
$\mu_{13}$	1	*	0	*	0 ③	*
	2	*	0	*	*	*
	3	*	0	*	0	*
	4	*	0	*	*	*
$\mu_{23}$	1	0	*	*	0 ③	*
	2	0	*	*	*	*
	3	0	*	*	*	*
	4	0	*	0 ②	*	*
最終モデル	1	H <sub>15</sub>	H <sub>6</sub> ①	H <sub>18</sub>	H <sub>12</sub> ④	H <sub>17</sub>
	2	H <sub>15</sub>	H <sub>14</sub>	H <sub>18</sub>	H <sub>16</sub>	H <sub>17</sub>
	3	H <sub>15</sub>	H <sub>14</sub>	H <sub>18</sub>	H <sub>18</sub>	H <sub>18</sub>
	4	H <sub>15</sub>	H <sub>14</sub>	H <sub>15</sub>	H <sub>18</sub>	H <sub>18</sub>

\*:  $\mu_{ij} \neq 0$  5%水準で棄却0:  $\mu_{ij} = 0$  5%水準で採択第1法:  $\mu_s = 0$  第2法: 適合度検定 第3法: AIC 第4法:  $G^2(m-M)$ 

①～⑤については本文中に記述。

## 6. 4法の比較

表9より第1～第4法間の解でずれを生じた個所について考察を加える。①～⑤は表中の注番号である。

データ2において  $\mu_{ij}$  に関する限り4者は一致しているが、①第1法では H<sub>6</sub>: 縮減モデルとなっている。これはサンプル数が  $n=97$  と少なく第1法の漸近理論の適用上の無理に起因すると思われる。

データ3では②第4法の  $G^2(m-M)$  が0となり矛盾する。従って最終モデルは H<sub>15</sub> となり、第4法のみ異なっている。第2種の過誤なのか検定方式そのものに由来するのか不明である。

データ4では③第1法で  $\mu_{13}, \mu_{23} = 0$  となり H<sub>12</sub> であるが、これは  $n=308$  によるものと思われる。④第2法では H<sub>16</sub>、第3、4法では H<sub>18</sub> で両者はくい違っているが、これは  $\mu_{12}$  の判断の相違による。第2法では適合、不適合の並べ方から機械的に最終モデルを選択するが、今上側確率 P を参考に用いれば H<sub>16</sub> のとき  $P=0.096$ 、H<sub>18</sub> のとき  $P=0.289$  で P から判断する限り H<sub>18</sub> の方が採択され易い。H<sub>16</sub>、H<sub>18</sub> の両者は接近しており第2法ではこうした情報は得られない。しかし第2法では  $AIC(16)=16.69$ 、 $AIC(18)=15.13$  であるから H<sub>18</sub> が選択されることになるが、H<sub>16</sub> の可能性も示唆しており、この意味で AIC



を用いる方法が望ましいともいえよう。

データ5においてもデータ4の場合と同じく⑤ $\mu_2$ の判断の相違から、第2法では $H_{17}$ が結論づけられているが、 $H_{17}$ のとき $P=0.142$ 、 $H_{18}$ のとき $P=0.854$ で $H_{18}$ の方が採択されやすい。 $AIC(17)=52.40$ 、 $AIC(18)=50.77$ であるから第3法ではこの辺の事情の想像がつく。

以上の理由により、第3法はAICを求めてMAICEを唯一つの採用するのではなく、相互に値を比較することにより柔軟な判断が端的に可能となるので(第2法によるときはPと表3～5によるアルゴリズムを用いねばならないのでやや複雑である)、第2法の結果の確認という以上に、望ましい方法であるといえよう。

第4法の結果はほぼ第3法と平行しており、特に計算を追加してこの方法を積極的に用いる根拠は乏しい。加算性はAICにも見られるので両者は同根の方法である。

## VI 情報理論による分析とLLM

情報理論(Information theory)の心理学への紹介は、McGill, W. J.<sup>18)</sup>(1954), Garner, W. R.<sup>19)</sup>(1958), Attneave, F.<sup>20)</sup>(1959), Binder, A. and Wolin<sup>21)</sup>, B. R. (1964)らをはじめ古くから試みられており、属性データに分析可能ということで心理学のさまざまな領域で広く用いられて来た。又分散分析との関連や、尺度水準別の適用法等、多方面からの研究もなされている(12, 20, 23, 25)。

LLMの紹介を終えるあたり、本節では情報理論で用いられる諸概念とLLMの $\mu$ -項との関連を整理し、尤度比統計量を用いる限り両分析法は同一のものに帰着することを示しておく。

### 1. 2変量情報量

事象 $X_i$ の生起の確率を $\pi_i$ とすると、情報量は

$$I(\pi_i) = -\log_2 \pi_i \quad (6.1)$$

で定義され、 $I(\pi_i)$ の期待値

$$H(X) = -\sum \pi_i \log_2 \pi_i \quad (6.2)$$

は $X$ の(平均)情報量又はエントロピーと呼ばれている。 $H(X)$ は $X$ が確定的であるとき小さい値を、不確定的であるとき大きい値を示すので不確実性の指標(uncertainty)として用いられる。情報理論では通常2を底とする対数がとられ、単位はビットと呼ばれているが、

ここでは  $G^2$  との関連を問題にするため自然対数を用いることとする。両者は比例関数にあるので概念の本質を損うことはない。

2変量を  $X, Y$  ( $\Pi$  節の  $A, B$  に対応する) とするとき,  $H(X), H(Y)$  の他同時事象の情報量  $H(X, Y)$ , 条件付情報量  $H(X|Y), H(Y|X)$  等が用いられる。それらは同時比率を  $\pi_{ij}$ , 周辺比率を  $\pi_{i\cdot}, \pi_{\cdot j}$  とおけば以下のごとくである。

$$\begin{aligned} H(X) &= -\sum_i \pi_{i\cdot} \log \pi_{i\cdot} \\ H(Y) &= -\sum_j \pi_{\cdot j} \log \pi_{\cdot j} \\ H(X, Y) &= -\sum_i \sum_j \pi_{ij} \log \pi_{ij} \\ H(X|Y) &= -\sum_i \sum_j \pi_{ij} \log \pi_{ij} / \pi_{i\cdot} = H(X, Y) - H(Y) \\ H(Y|X) &= -\sum_i \sum_j \pi_{ij} \log \pi_{ij} / \pi_{\cdot j} = H(X, Y) - H(X) \end{aligned} \quad (6.3)$$

又伝達情報量 (transmitted information) は

$$\begin{aligned} T(X:Y) &= H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y) \\ &= -\sum_i \sum_j \pi_{ij} \log \pi_{i\cdot} \pi_{\cdot j} / \pi_{ij} \end{aligned} \quad (6.4)$$

であり, これは  $X$  をあるシステムへの入力,  $Y$  を出力と考えると,  $X$  から  $Y$  に伝わる部分でありその意味で伝達情報量と呼ばれている。関連の分析に際しては,  $X, Y$  が独立のとき  $T(X:Y) = 0$  であるので  $T(X:Y)$  の大小の検討をすればよい。

データから情報量を計算するときは, 母比率の最尤推定量として前部までのごとく  $\hat{\pi}_{ij} = f_{ij}/n, \hat{\pi}_{i\cdot} = f_{i\cdot}/n, \hat{\pi}_{\cdot j} = f_{\cdot j}/n$  を用いればよい。又尤度比統計量  $G^2$  との関連も次のごとく導くことができる。注)

$$\begin{aligned} \hat{H}(X, Y) &= -\sum_i \sum_j \hat{\pi}_{ij} \log \hat{\pi}_{ij} = \log n - \frac{1}{n} \sum_i \sum_j f_{ij} \log f_{ij} \\ &= \log IJ - \frac{1}{2n} G^2(1) \end{aligned} \quad (6.5)$$

$$\hat{H}(Y|X) = -\frac{1}{n} \sum_i \sum_j f_{ij} \log \frac{f_{ij}}{f_{i\cdot}} = \log J - \frac{1}{2n} G^2(2) \quad (6.6)$$

$$\hat{H}(X|Y) = -\frac{1}{n} \sum_i \sum_j f_{ij} \log \frac{f_{ij}}{f_{\cdot j}} = \log I - \frac{1}{2n} G^2(3) \quad (6.7)$$

$$\hat{H}(X) = \log n - \frac{1}{n} \sum_i f_{i\cdot} \log f_{i\cdot} = \log I - \frac{1}{2n} (G^2(1) - G^2(2)) \quad (6.8)$$

$$\hat{H}(Y) = \log n - \frac{1}{n} \sum_j f_{\cdot j} \log f_{\cdot j} = \log J - \frac{1}{2n} (G^2(1) - G^2(3)) \quad (6.9)$$

$$\hat{T}(X:Y) = \log n - \frac{1}{n} \sum_i \sum_j f_{ij} \log \frac{f_{i\cdot} f_{\cdot j}}{f_{ij}} = \frac{1}{2n} G^2(4) \quad (6.10)$$

$T(X:Y)$  による  $X, Y$  間の関連は  $G^2(4) = 2n \hat{T}(X:Y) \sim \chi^2_{(I-1)(J-1)}$  の性質を用いて検

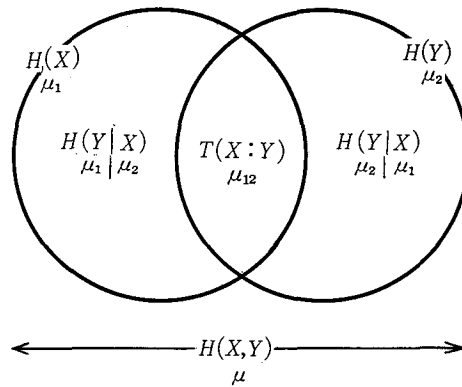


図6 2変量情報量とLLM

討することができる。情報理論でも Pearsom, K. の  $\chi^2$  統計量を用いる方法が多く紹介されているようであるが、既に考察したごとく加算性が成立するので  $G^2$  を使用の方が望ましい。 $\chi^2$ への近似も  $G^2$ の方がよいようである。

そして  $T(X:Y)$  の他  $H(X)$ ,  $H(X,Y)$  等の情報理論の概念はすべて  $G^2$  に置き換えることが可能である。今,  $H(X,Y)$  を  $\mu$  に,  $H(X)$  を  $\mu_1$ ,  $H(Y)$  を  $\mu_2$ ,  $H(Y|X)$  を  $\mu_2|\mu_1$  (又は  $\mu - \mu_1$ ),  $H(X|Y)$  を  $\mu_1|\mu_2$  (又は  $\mu - \mu_2$ ),  $T(X:Y)$  を  $\mu_{12}$  に対応させる。LLM の  $\mu$ -項は, 例えば  $\mu$ ,  $\mu_1$ ,  $\mu_2$  は互いに素であるが情報理論では  $H(X)$  は  $T(X:Y)$  を含み,  $H(X,Y)$  は  $H(X|Y)$ ,  $T(X:Y)$ ,  $H(Y|X)$  を含んでいる (図6)。従って上述の対応はゆるい意味においてである。この対応づけから, 例えば (6.10) より  $G^2(4) = 2n \hat{T}(X:Y)$  の関係を用いて  $\chi^2$  検定を行ない,  $H_4$  のモデルが棄却されたとき  $\mu_{12}$  が大きく, 従って  $\mu_{12} \approx 0$  の結論を下すことができる。

(注) 次のごとく考えれば代数的展開はいずれも容易である。例えば (6.5) 式において

$$\log f_{ij} = -\log \frac{n/IJ}{f_{ij}} \times \frac{IJ}{n} \text{ と変形すれば表1及び } G^2 \text{ の公式より}$$

$$\hat{H}(X,Y) = \log IJ - \frac{1}{2n} (-2 \sum_i \sum_j f_{ij} \log \frac{n/IJ}{f_{ij}}) = \log IJ - \frac{1}{2n} G^2(1) \quad (6.11)$$

である。又 (6.8) 式も上と同様に

$$\log f_{i.} = \log \frac{f_{i.}/J}{f_{ij}} \times \frac{f_{ij}}{n/IJ} \times \frac{n}{I} \text{ と変形すれば}$$

$$\begin{aligned} \hat{H}(X) &= \log n - \frac{1}{n} \sum_i \sum_j f_{ij} (\log \frac{f_{i.}/J}{f_{ij}} - \log \frac{n/IJ}{f_{ij}} + \log n - \log I) \\ &= \log I - \frac{1}{2n} (G^2(1) - G^2(2)) \end{aligned} \quad (6.12)$$

である。

## 2. 3変量情報量

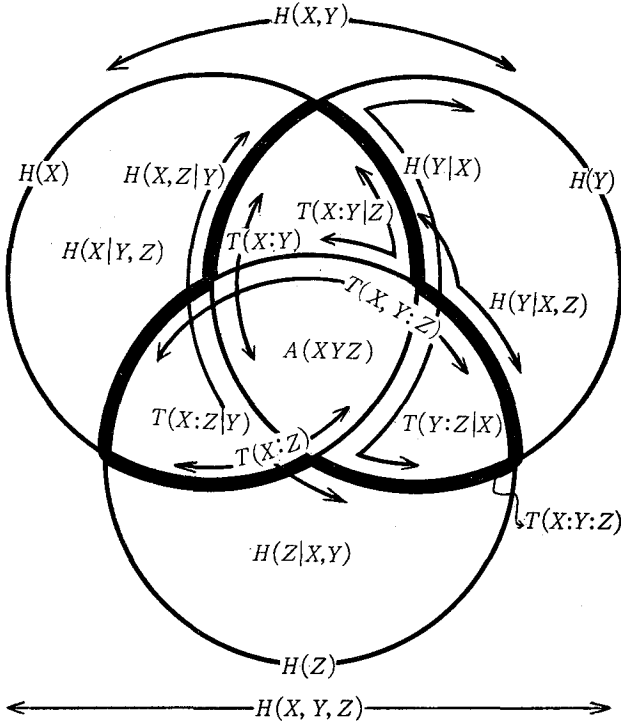


図7 3変量情報量

2変量の場合の情報量は3変量の場合にも拡張することができる。但し3変量にすることにより条件付情報量，伝達情報量ともに図7のごとく多種のものが定義され複雑である。

先ず  $H(X, Y, Z)$ ,  $H(X)$ ,  $H(X, Y)$  等は2変量からの単純な拡張である。

$$\hat{H}(X, Y, Z) = \log n - \frac{1}{n} \sum_i \sum_j \sum_k f_{ijk} \log f_{ijk} = \log IJK - \frac{1}{2n} G^2(1) \quad (6.13)$$

$$\begin{aligned} \hat{H}(X) &= \log n - \frac{1}{n} \sum_i f_{i..} \log f_{i..} \\ &= \log n - \frac{1}{n} \sum_i \sum_j \sum_k f_{ijk} \log \frac{f_{i..}/JK}{f_{ijk}} \cdot \frac{f_{ijk}}{n/IJK} \cdot \frac{n}{I} \\ &= \log I - \frac{1}{2n} (G^2(1) - G^2(2)) \end{aligned} \quad (6.14)$$

$$\hat{H}(X, Y) = \log n - \frac{1}{n} \sum_i \sum_j f_{ij.} \log f_{ij.} = \log IJ - \frac{1}{2n} (G^2(1) - G^2(8)) \quad (6.15)$$

この他  $X, Y, Z$  を循環させて  $\hat{H}(Y)$ ,  $\hat{H}(Z)$ ,  $\hat{H}(X, Z)$ ,  $\hat{H}(Y, Z)$  もあるが，これらは表10にまとめて示す。LLM との関連は， $H(X, Y, Z)$  は全情報量で  $\mu$  に， $H(X)$  は  $\mu_1$  に， $H(X, Y)$  は  $\mu_1 + \mu_2$  に対応させておく。

条件付情報量については3種のものが考えられる。

表10 3変量情報量と LLM

情報量 <sup>1)</sup>	$G^2$	$\mu$ -項
$H(X, Y, Z)$	$\log I/K - G^2(1)/2n$	$\mu$
$H(Y, Z X)$	$\log JK - G^2(2)/2n$	$\mu_2 + \mu_3   \mu_1$
$H(X, Z Y)$	$\log IK - G^2(3)/2n$	$\mu_1 + \mu_3   \mu_2$
$H(X, Y Z)$	$\log IJ - G^2(4)/2n$	$\mu_1 + \mu_2   \mu_3$
$H(X)$	$\log I - (G^2(1) - G^2(2))/2n$	$\mu_1$
$H(Y)$	$\log J - (G^2(1) - G^2(3))/2n$	$\mu_2$
$H(Z)$	$\log K - (G^2(1) - G^2(4))/2n$	$\mu_3$
$H(X, Y)$	$\log IJ - (G^2(1) - G^2(8))/2n$	$\mu_1 + \mu_2$
$H(X, Z)$	$\log IK - (G^2(1) - G^2(9))/2n$	$\mu_1 + \mu_3$
$H(Y, Z)$	$\log JK - (G^2(1) - G^2(10))/2n$	$\mu_2 + \mu_3$
$H(Z X, Y)$	$\log K - G^2(8)/2n$	$\mu_3   \mu_1 + \mu_2$
$H(Y X, Z)$	$\log J - G^2(9)/2n$	$\mu_2   \mu_1 + \mu_3$
$H(X Y, Z)$	$\log I - G^2(10)/2n$	$\mu_1   \mu_2 + \mu_3$
$H(Z X)$	$\log K - (G^2(8) - G^2(15))/2n$	$\mu_3   \mu_1$
$H(Z Y)$	$\log K - (G^2(8) - G^2(16))/2n$	$\mu_3   \mu_2$
$H(Y X)$	$\log J - (G^2(9) - G^2(15))/2n$	$\mu_2   \mu_1$
$H(Y Z)$	$\log J - (G^2(9) - G^2(17))/2n$	$\mu_2   \mu_3$
$H(X Y)$	$\log I - (G^2(10) - G^2(16))/2n$	$\mu_1   \mu_2$
$H(X Z)$	$\log I - (G^2(10) - G^2(17))/2n$	$\mu_1   \mu_3$
$T(X:Y)$	$(G^2(5) - G^2(8))/2n$	$\mu_{12}$
$T(X:Z)$	$(G^2(6) - G^2(9))/2n$	$\mu_{13}$
$T(Y:Z)$	$(G^2(7) - G^2(10))/2n$	$\mu_{23}$
$T(X:Y:Z)$	$G^2(11)/2n$	$\mu_{12} + \mu_{13} + \mu_{23}$
$T(X, Y:Z)$	$G^2(12)/2n$	$\mu_{13} + \mu_{23}$
$T(X, Z:Y)$	$G^2(13)/2n$	$\mu_{12} + \mu_{23}$
$T(Y, Z:X)$	$G^2(14)/2n$	$\mu_{12} + \mu_{13}$
$T(Y:Z X)$	$G^2(15)/2n$	$\mu_{23}   \mu_1$
$T(X:Z Y)$	$G^2(16)/2n$	$\mu_{13}   \mu_2$
$T(X:Y Z)$	$G^2(17)/2n$	$\mu_{12}   \mu_3$
$A(XYZ)$	$G^2(18)/2n$	$\mu_{123}$

1)  $G^2$  との対応を考えたとき厳密にはへを冠した方がよいと思われるが煩雑さを避け省略する。

$$\begin{aligned}
 \hat{H}(Z|X) &= \hat{H}(X, Z) - \hat{H}(X) = -\frac{1}{n} \sum_i \sum_j \sum_k f_{ijk} \log \frac{f_{i..k}}{f_{i..}} \\
 &= \log I - \frac{1}{2n} (G^2(8) - G^2(15))
 \end{aligned} \tag{6.16}$$

$$\hat{H}(Y, Z|X) = \hat{H}(X, Y, Z) - \hat{H}(X) = \log JK - \frac{1}{2n} G^2(2) \tag{6.17}$$

$$\hat{H}(Z|X, Y) = \hat{H}(X, Y, Z) - \hat{H}(X, Y) = \log K - \frac{1}{2n} G^2(8) \quad (6.18)$$

$H(Z|X)$  は  $H(Z)$  のうち  $X$  に係る部分を除いたもので LLM では 2 変量の場合と同じく  $\mu_3|\mu_2$  と書くことにする。同様に  $H(Y, Z|X)$  は  $H(Y, Z)$  から  $X$  の部分を除いたもので  $\mu_2 + \mu_3|\mu_1$ ,  $H(Z|X, Y)$  は  $H(Z)$  から  $X, Y$  の部分を除いたものであるので  $\mu_3|\mu_1, \mu_2$  として対応させておく。

伝達情報量は次のごとくであり, これらを用いて変量間の各種関連の検討をすることができる。先ず  $T(X: Y)$  は 2 変量の場合と同じく  $X, Y$  間の関連で

$$\begin{aligned} \hat{T}(X: Y) &= \hat{H}(X) - \hat{H}(Y) - \hat{H}(X, Y) = \log n - \frac{1}{n} \sum \sum \sum f_{ijk} \log \frac{f_{i..} f_{.j.}}{f_{ij.}} \\ &= \frac{1}{2n} (G^2(5) - G^2(8)) \end{aligned} \quad (6.19)$$

である。 $\mu_{12}$  の検討にはこれを用いればよく, 自由度は  $G^2(5) - G^2(8)$  から  $(IJK - I - J + 1) - IJ(K - 1) = (I - 1)(J - 1)$  となり  $2n \hat{T}(X: Y) \sim \chi^2_{(I-1)(J-1)}$  が利用できる。 $\mu_{13}, \mu_{23}$  も同様に  $\hat{T}(X: Z), \hat{T}(Y: Z)$  を用いればよい。

$T(X, Y: Z)$  は 2 変量  $X, Y$  と  $Z$  との関連であり

$$\begin{aligned} \hat{T}(X, Y: Z) &= \hat{T}(Z: X, Y) = \hat{H}(X, Y) + \hat{H}(Z) - \hat{H}(X, Y, Z) \\ &= \log n - \frac{1}{n} \sum \sum \sum f_{ijk} \log \frac{f_{ij.} f_{..k}}{f_{ijk}} = \frac{1}{2n} G^2(12) \end{aligned} \quad (6.20)$$

である。これより  $G^2(12) = 2n \hat{T}(X, Y: Z) \sim \chi^2_{(K-1)(I+J-1)}$  であるから,  $G^2(12)$  の値が大きく  $H_{12}$  が棄却されたとき  $\mu_{12} + \mu_{13}$  が大きく無視できないことを示している。 $X, Y, Z$  を循環させれば多重独立モデルの検討が可能である。

$T(X: Y|Z)$  は先の  $T(X: Y)$  から  $Z$  の部分を除いたものであり,

$$\begin{aligned} \hat{T}(X: Y|Z) &= \hat{H}(X|Z) - \hat{H}(X|Y, Z) = -\frac{1}{n} \sum \sum \sum f_{ijk} \log \frac{f_{i.k}}{f_{..k}} \cdot \frac{f_{.jk}}{f_{ijk}} \\ &= \frac{1}{2n} G^2(17) \end{aligned} \quad (6.21)$$

が導ける。 $X, Y$  の関連を  $Z$  の条件付のもとで検討するのであるから  $G^2(17) = 2n \hat{T}(X: Y|Z) \sim \chi^2_{K(I-1)(J-1)}$  より  $H_{17}$  が棄却されたとき  $\mu_{12}$  を無視できないことがわかり,  $T(X: Z|Y), T(Y: Z|X)$  とあわせて条件付独立モデルの検討をすることができる。

$T(X: Y: Z)$  は束縛 (cnstrain) と呼ばれており, 3 変量  $X, Y, Z$  間の関連の程度を表わしている。

$$\begin{aligned} \hat{T}(X: Y: Z) &= \hat{H}(X) + \hat{H}(Y) + \hat{H}(Z) - \hat{H}(X, Y, Z) \\ &= 2 \log n - \frac{1}{n} \sum \sum \sum f_{ijk} \log \frac{f_{i..} f_{.j.} f_{..k}}{f_{ijk}} = \frac{1}{2n} G^2(11) \end{aligned} \quad (6.22)$$

これより  $G^2(11) = 2n \hat{T}(X: Y: Z) \sim \chi^2_{IJK - (I+J+K) + 2}$  の検定から  $H_{11}$  が棄却されたとき,

$\mu_{12} + \mu_{13} + \mu_{23}$  が大きく無視できないことを示している。これらは相互独立モデルの場合に相当している。

最後に  $A(XYZ)$  は情報理論でも交互作用 (interaction) と呼ばれており

$$\begin{aligned} A(XYZ) &= T(X, Y : Z) - (T(X : Z) + T(Y : Z)) \\ &= T(X : Y : Z) - T(X : Y) - T(X : Z) - T(Y : Z) \end{aligned} \quad (6.23)$$

で定義される (Attneave<sup>2)</sup>, p. 58)。従って (6.19), (6.20) の循環式を用いて変形すれば

$$\hat{A}(XYZ) = -\frac{1}{n} \sum \sum \sum f_{ijk} \log \frac{f_{ij.} f_{i.k} f_{.jk}}{f_{i..} f_{.j.} f_{..k}} \times \frac{n}{f_{ijk}} = -\frac{1}{2n} G^2(18) \quad (6.24)$$

である。 $\hat{A}(XYZ) < 0$  のときは  $-\hat{A}(XYZ) = G^2(18)/2n$  とする。又操作上は分離できないが  $A$  は  $A = A^{(+)} - A^{(-)}$  として  $A^{(+)}$  (交互作用) と  $A^{(-)}$  (伝達重複) とに分けられる。

但しここでは反復法による  $G^2(18)$  の算出ではなく (3.6) 式の構造式より得られる

$$\hat{F}_{ijk} = \frac{n f_{ij.} f_{i.k} f_{.jk}}{f_{i..} f_{.j.} f_{..k}} \quad (6.25)$$

を用いた。これを用いることにより (6.23) 式と (6.24) 式が統一的に説明できるし、3次の交互作用の概念でも情報理論と LLM が対応していることがわかる。 $G^2(18) = 2n \hat{A}(XYZ) \sim \chi^2_{(I-1)(J-1)(K-1)}$  の検定から  $H_{18}$  が棄却のとき  $\mu_{123} \neq 0$  を結論づければよい。

以上見て来たごとく、LLM の3次元の各種モデルを伝達情報量を用いて検討することができ、又縮減モデルについても情報量  $H(\cdot)$  を用いて分析しうが詳細は割愛する。3変量の場合について  $G^2$  との関連を整理し一覧表にすれば表10のごとくである。同表に明きらかなごとく両方法は平行関係にあり、LLM の検討は情報理論の概念(用語)を用いても可能である。数値例は省略するが表10により計算した  $G^2$  の値と表7の  $G^2_0$ -値は計算誤差の範囲内でよく一致していた ( $G^2(18)$  は除いて)。情報量の計算はさほど面倒なものではなく、従って第2法の手順は情報理論を用いれば大幅に簡略化することも可能である。

## VII 結 語

多次元属性データが分割表の形で与えられるとき、変数間の関連を分析する手法とし Log-Linear Model (LLM, 対数線形モデル) による方法が注目されている。本稿では3次元の場合について理論と分析法が紹介された。

3変数の場合、変数間の相互関連のあり方は(1)縮減モデル、(2)相互独立モデル、(3)多重独立モデル、(4)条件付独立モデル、(5)対連関モデル、(6)飽和モデルに大別され、併せて19種数の下位モデル(仮説)について検討されるが、LLM は所与のデータがそのいずれが最適で

あるかを決定する方法であり、名称は細胞の期待度数の対数を

$$\log F_{ijk} = \mu + \mu_1 + \mu_2 + \mu_3 + \mu_{12} + \mu_{13} + \mu_{23} + \mu_{123} \quad (\text{飽和モデル})$$

と  $\mu$ -項の線形結合に分解して分析することに由来している。

解法としては2通りのものが紹介されている。第1法は連立方程式を解いて  $\mu$ -項を求め、検定により  $\mu_i=0$  のものを落とし、残された項から上記(1)~(5)のいずれのモデルが適合するかを探る方法である。

第2法は、各場合についてその構造式から期待度数  $\hat{F}_{ijk}$  を求め、観測度数  $f_{ijk}$  との適合度検定を行ない、それらの結果から最終的に適合モデルを決定する方法である。

両方法の結論は多くの場合一致するが、第1法の検定は正規分布への漸近理論を用いているので、 $n$ (サンプル数)が小の場合には必ずしも一致しない。従って解法としては第2法の方が望ましく、得られた結論の理解には、 $\mu_{ij}$ の項が2変数間の関連の存在を示しているの、第1法の方が便利である。

第2法の適合度検定では尤度比統計量  $G^2$  が用いられたが、このときの結論は適、不適の2値的なものである。そして19種類のモデルの適合状況から論理的に最終モデルが決定され、そのアルゴリズムも考察されたが、結果に矛盾をきたし論理的(機械的)判断が困難の場合には、 $G^2$ の上側確率、或はその他の情報を用いて総合的に判断せざるを得ない。

$G^2$  から AIC を求め MAICE (min AIC) のものを最適モデルとして決定する方法も吟味された。このとき、各モデルの適合状況を連続的に評価でき、MAICEに近い複数個のモデルを選択すれば柔軟な解釈が可能である。

第1法で残された  $\mu_{ij}$ の項は3変数相互の関連を示唆しているが、関連の有無と同時にその強度もデータの考察時に必要である。強度の指標として  $\mu_{ij}$ の分散をとりあげ、それを相対化して  $S^2_{ij}$ として算出し、三角座標で表現する方法が提示された。視覚化された結果から直観的な解釈が可能である。

以上の計算の過程は FORTAN 語でプログラミングされ、チェックデータを用いてデバッグ及び出力結果の検討がなされた。

分割表データの分析法のひとつとして情報理論を応用し、各細胞の出現頻度から平均情報量、変数間伝達情報量を算出する方法も心理学の分野に紹介されているが、LLMの各モデルはこれら情報理論の概念を用いても検討することができる。両方法の平行関係についても整理された。

従来、心理学の領域では名義尺度等の属性データが多く取扱われるにも拘らず、クロス集計による分析は2次元に留まるケースが多いようにも思える。対数線形モデルによる方法は3次元以上の属性データの分析に有効であり、今後一層の普及が期待されるところである。



謝辞 末筆ながら対数線形モデルおよび AIC について御教示いただきました本学人間科学部西田春彦教授, 基礎工学部稲垣宜生助教授, ならびに平素統計学について御指導いただいております教養部高松俊朗教授に深く感謝申し上げます。

## 文 献

- 1) Akaike, H. 1976, On entropy maximization principle, In Krishnaiah, P. R. ed, *Applications of statistics*, North-Holland, 27-41.
- 2) Attneave, F. 1959, *Applications of information theory to psychology*, Holt. 小野茂, 羽生義正訳, 心理学と情報理論, ラティス。
- 3) Binder, A. and Wolin, B. R. 1964, Informational models and their uses, *Psychometrika*, **29**, 29-54.
- 4) Birch, M. W. 1963, Maximum likelihood in three-way contingency tables, *J. Roy. Stat. Soc. B*, **25**, 220-233.
- 5) Bishop, Y. M. M. 1969 a, Incomplete two-dimensional contingency tables, *Biometrics*, **25**, 119-128.
- 6) Bishop, Y. M. M. 1969 b, Full contingency tables, logits and split contingency tables, *Biometrics*, **25**, 383-399.
- 7) Bishop, Y. M. M. et al, 1975, *Discrete multivariate analysis*, MIT-Press.
- 8) Davis, J. A. 1974, Hierarchical models for significance tests in multivariate contingency table: An exegesis of Goodman's recent papers, *Sociol. Methodol.* 189-231.
- 9) Everitt, B. S. 1977, *The analysis of contingency tables*, Chapman and Hall.
- 10) Garner, W. R. and McGill, W. J. 1956, The relation between information and variance analysis, *Psychometrika*, **21**, 219-228.
- 11) Garner, W. R. 1958, Symmetric uncertainty analysis and its implications for psychology, *Psychol. Rev.* **65**, 183-196.
- 12) Goodman, L. A. 1970, The multivariate analysis of qualitative data: Interactions among multiple classifications, *J. Amer. Stat. Assoc.* **65**, 226-256.
- 13) Goodman, L. A. 1971, The analysis of multidimensional contingency tables: Stepwise procedures and direct estimation method for building models for multiple classifications, *Technometrics*, **13**, 33-61.
- 14) Goodman, L. A. 1972 a, A modified multiple regression approach to the analysis of dichotomous variables, *Amer. Sociol. Rev.* **37**, 28-46.
- 15) Goodman, L. A. 1972 b, A general model for the analysis of surveys, *Amer. J. Sociol.* **77**, 1035-1086.
- 16) Haberman, S. J. 1973, Log-Linear model for frequency data: Sufficient statistics and likelihood equations, *Annal. Statist.* **1**, 617-632.
- 17) Haberman, S. J. 1974, *The analysis of frequency data*, Univ. Chicago Press.
- 18) McGill, W. J. 1954, Multivariate information transmission, *Psychometrika*, **19**, 97-116.
- 19) Payne, C. 1977, The log-linear model for contingency tables, In O'Muircheartaigh C. A. ed, *The analysis of survey data*, Vol. 2, John Wiley.
- 20) Pearson, W. H. 1966, Estimation of correlation coefficient from an uncertainty measure, *Psychometrika*, **31**, 421-433.
- 21) Pearson, E. S. and Kendall, M. G. ed, 1970 *Studies in the history of statistics and probability*, Charles Griffin.
- 22) Plackett, R. L. 1974, *The analysis of categorical data*, Charles Griffin.
- 23) Read, C. B. 1978, Tests of symmetry in three-way contingency tables, *Psychometrika*, **43**, 409-420.
- 24) Sakamoto, Y. and Akaike, H. 1978, Analysis of cross classified data by AIC, *Ann. Inst. Statist. Math.* **30 B**, 185-197.
- 25) Särndel, C. E. 1974, A comparative study of association measures, *Psychometrika*, **39**, 165-187.
- 26) Upton, G. J. G. 1977, *The analysis of cross-tabulated data*, Wiley.

ANALYSIS OF THREE-DIMENSIONAL CONTINGENCY  
TABLE BY LOG-LINEAR MODEL

Mitsuo YOSHIDA

Log-Linear Model (LLM) has recently been studied in the area of sociology and mathematical statistics. When multi-dimensional categorical variables are presented in the form of contingency table, LLM provides a comprehensive scheme for the analysis of relationships between those variables, such as mutual-, multiple-, conditional-independence, pair-wise associations, and so on. In that model, those relationships can be described with linear combinations of  $\mu$ -terms which denote simple effects or interactions.

This paper tries to introduce the theory and method of LLM for three-dimensional case, and proposes wider applications of the model in psychology. Generally, there are two different types of analyses in the model. The one is to compute the  $\mu$ -terms to solve simultaneous linear equations under the hypothesis of saturation model and the final fitted model is decided by z-test of  $H_0: \mu_s = 0$ . The other is to test goodness of fit by likelihood ratio statistic ( $G^2$ ) and to detect the most fitted model from nineteen hypotheses. By a rule of thumb, the latter method is better, for the former, using the limiting theory of standard normal distribution, necessitates more than 1000 samples. In case of less samples, the results of two methods are not always coincident.

The following is discussed by the author.

Tentative algorithms of the final model in the second method are offered. The results are confirmed by AIC (Akaike Information Criterion)-values. Intensities of associations are visualized in triangular coordinates of variances of second order interactions.  $G^2$  and  $\mu$ -terms are examined by information theory. Computational procedures are programmed in FORTRAN-languages about 1000 statements.