

《紹介と展望》

多次元分割表の取扱いについて†

長谷川 実*

1. 要 約

分割表についての研究は、歴史的に見ても統計学上の基本的な問題として古くから取り扱われていますが、このところ応用面を強調したいいくつかの論文が発表されています。とくにここで強調したいのは、最近の傾向として取り扱われている手法が計算機の利用と結びついて実際に利用されてきていることと、この使用が多くなるにつれて理論的に整えなければならない問題がはつきりしてきたことです。こうした理論と応用との相互関係は好ましいものですし、これからの発展が期待されます。

この小論では、多次元分割表の研究の最も基本的部分を最近の傾向に従って取り扱い、この手法をどのように使用するかについて私が最近取り扱ったデータを中心に述べてみたいと思います。とくに行動科学の分野での利用を期待しています。

2. モデルの構成

2・1 ここでは3次元分割表について取り扱います。理論的には4次元以上の分割表でも3次元での議論と変わらないのですが、最後に述べるように、4次元以上の分割表を実際に解析する際にモデルの取扱いを巡ってかなり重要な問題が生じます。

さて、 $I \times J \times K$ 分割表を考えてみます。与えられたデータは

$$x_{ijk}(i=1, 2, \dots, I; j=1, 2, \dots, J; k=1, 2, \dots, K)$$

とします。この与えられたデータから理論上最も好ましい真の値—母数—

$$m_{ijk}(i=1, 2, \dots, I; j=1, 2, \dots, J; k=1, 2, \dots, K)$$

を構成することが問題です。何が最も好ましい性質なのかについて種々の議論がありますが、ここでは最尤法（Method of maximum likelihood）を中心にこの推測の問題を取り扱います。

最尤推定量の基本的な性質、とくに大標本論の立場に立つと、近似的に望ましい性質を持つことが理由の一つとなって、最近の傾向として最尤法を中心としたこの問題の解析が行なわれます。Birch の基本的な研究 [1], Bishop [2], Goodman [11] [12], Fienberg [7] のいわゆ

† 1974年1月11日受理。

* Control Data Institute, Toronto.

る Quasi-independence に関する研究が代表的な例です. 2章においては Birch の研究を中心として, これらの研究の基本的部分を紹介することにしします.

2・2 まず与えられたデータ, $X=(x_{ijk})$ より可能な推定量 $M=(m_{ijk})$ の形 (モデル) を構成します. どのようなモデルを考察の対象とするかは, 個々のデータとそれを扱う研究者の判断によって定まることですが, まず次のような各因子間の相互関係を表わす代表的なモデルを考えてみます.

- (1) 各因子 I, II, III は互いに独立なモデル:

$$m_{ijk} = m_{i..} m_{.j.} m_{..k} / (m_{...})^2$$

- (2) 因子 III が因子 I, II と互いに独立なモデル:

$$m_{ijk} = m_{ij.} m_{..k} / m_{...}$$

- (3) 因子 I と因子 II とが周辺分布を考えると互いに独立なモデル:

$$m_{ij.} = m_{i..} m_{.j.} / m_{...}$$

- (4) 因子 III のどの値に対しても因子 I と因子 II とが互いに独立なモデル:

$$m_{ijk} = m_{i.k} m_{.jk} / m_{..k}$$

- (5) 3 因子間の相互作用が見られないモデル

表 0 $2 \times 2 \times 2$ 分割表の例					
1	1	2	6	2	8
1	1	2	3	1	4
2	2	4	9	3	12

モデル(5)については別に詳しく述べることにします. 上から順を追ってモデルは複雑になっていきますが, ここで注意しておきたいのは, 各モデルの間の関係はそれほど明白ではありません. 例題は Simpson [21] と Birch [1] と

に見られます. (4), (5)が成立し(3)が成立しない簡単な例として表 0 をあげておきます.

ここではまず m_{ijk} を八つの要素に分解して, いわゆる loglinear model を考えます.

$$\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)}$$

$$\sum_i u_{1(i)} = \sum_j u_{2(j)} = \sum_k u_{3(k)} = 0$$

$$\sum_i u_{12(ij)} = \sum_j u_{12(ij)} = 0$$

$$\sum_i u_{13(ik)} = \sum_k u_{13(ik)} = 0$$

$$\sum_j u_{23(jk)} = \sum_k u_{23(jk)} = 0$$

$$\sum_i u_{123(ijk)} = \sum_j u_{123(ijk)} = \sum_k u_{123(ijk)} = 0$$

ここで u は $\{m_{ijk}\}$ に固有な一定量, $u_{1(i)}$ は因子 I に固有な量, $u_{12(ij)}$ は因子 I と因子 II との相互関係に固有な量, そして $u_{123(ijk)}$ は 3 因子 I, II, III の相互作用を代表する量です.

この loglinear model の導入は Birch によるものですが, こうすることによってそれまでに得られた結果 (たとえば Roy and Kastenbaum [20]) がかなり見通しよく見られるようになりました. ここで考える対象とするモデルは, 次の八つのいわゆる non-hierarchical model です.

$$(L1) \quad u_{123(ijk)} = 0, \quad u_{12(ij)} = 0, \quad u_{23(jk)} = 0, \quad u_{13(ik)} = 0$$

$$(L2) \quad u_{123(ijk)} = 0, \quad u_{23(jk)} = 0, \quad u_{13(ik)} = 0$$

$$(L3) \quad u_{123(ijk)}=0, \quad u_{12(ij)}=0$$

$$(L4) \quad u_{123(ijk)}=0$$

(L2) と (L3) はそれぞれ三つの場合があります。簡単な計算によって (L1) から (1), (L2) から (2), (L3) から (4) を導くことができます。モデル (L4) は (5) に対応する Roy-Kastenbaum, Birch によるモデルで, 「 $u_{123(ijk)}=0$, すなわち 3 因子による相互作用のないモデル」あるいは「各 2 因子間の部分的な相互関係を代表するモデル」ということができます。これらのモデルの中でどのモデルが与えられたデータに対して適当であるかを判断する一つの基準を与えようという試みが以下の目的です (これらをまとめて Quasi-independence の研究ということがあります)。

2・3 初めに 3 次元の行列がどのような仕組で得られた (標本) 行列であるかについて仮定を設ける必要があります。ここで考えるのは次の四つの場合です。

(1) ポアソン分布型: 行列の各要素 x_{ijk} が互いに独立なポアソン型の分布 (平均 m_{ijk}) に従う場合

(2) 多項分布型: 与えられた N 個の標本が $I \times J \times K$ 分割表に多項分布 $M(\{p_{ijk}\}; N)$ に従って分類された場合, このときの確率関数は ($m_{ijk}=Np_{ijk}$ として)

$$N! \prod_{ijk} (m_{ijk}/N)^{x_{ijk}} / \prod_{ijk} x_{ijk}!$$

(3) 因子 III の任意の値 k に対して, 周辺度数 $x_{..k} = \sum_{ij} x_{ijk}$ を求め, これにより与えられた $x_{..k}$ 個の標本 x_{ijk} が $I \times J$ 分割表に多項分布 $M(\{p_{ijk}\}; x_{..k})$ に従って分類された場合, このときの確率関数は ($m_{ijk}=x_{..k}p_{ijk}$ として)

$$p = \prod_k (x_{..k}/N)^{x_{..k}} / \prod_{ij} x_{ijk}! \left(\prod_{ij} m_{ijk}/x_{..k} \right)^{x_{ijk}}$$

(2) に従う行列 (x_{ijk}) について, $x_{..k}$ の分布を考えるとよく知られているように

$$q = N! \prod_k (m_{..k}/N)^{x_{..k}} / \prod_k x_{..k}!$$

したがって x_{ijk} の $m_{..k}=x_{..k}$ という条件付きの確率 p/q が (3) に相当することは明らかです。

(4) 混合型: (1) と (2) とを混合した場合が最尤法を論ずる上では (1) に帰着することに注意します。確率変数列

$$x_1, x_2, \dots, x_{\alpha_1}, x_{\alpha_1+1}, \dots, x_{\alpha_2}, x_{\alpha_2+1}, \dots, x_{\alpha_s}, x_{\alpha_s+1}, \dots, x_t$$

を考えます。このうち s 個の組 $(x_{\alpha_{j-1}+1}, \dots, x_{\alpha_j})$ は他と独立な多項分布 $M(m_{\alpha_{j-1}+1}^{(\theta)}, \dots, m_{\alpha_j}^{(\theta)}, N_j)$ ($N_j = \sum_{l=\alpha_{j-1}+1}^{\alpha_j} x_l$) に従うものとし, $(x_{\alpha_s+1}, \dots, x_t)$ は平均 $m_i^{(\theta)}$ ($i=\alpha_s+1, \dots, t$) の他と独立なポアソン分布に従うものとします。このモデルより得られた θ の最尤推定量を $\hat{\theta}^*$ とします。これに対して (1) のモデル, すなわち確率変数列 $x_1, x_2, \dots, x_{\alpha_1}, \dots, x_t$ が平均 $m_l(\theta)$ ($l=1, 2, \dots, \alpha_1, \dots, t$) の互いに独立なポアソン分布に従う場合を考え, この (1) のモデルにより得られた θ の最尤推定量を $\hat{\theta}$ とします。このとき $\hat{\theta} = \hat{\theta}^*$ かつ $m_l(\hat{\theta}) = m_l(\hat{\theta}^*)$ が成立する条件は

$$\sum_{l=\alpha_{j-1}+1}^{\alpha_j} m_l(\hat{\theta}) = N_j \quad (j=1, \dots, s)$$

です。この証明は次の尤度関数を比較することにより得られます。

(1)のモデルの尤度関数を l_1 , (4)のモデルの尤度関数を l_4 とします.

$$l_1 = \prod_{l=1}^t e^{-m_l m_l^{x_l}} / x_l!$$

$$l_4 = \prod_{j=1}^s \{ (N_j! / \prod_{l=\alpha_{j-1}+1}^{\alpha_j} x_l!) \prod_{l=\alpha_{j-1}+1}^{\alpha_j} (m_l / N_j)^{x_l} \} \prod_{k=\alpha_s+1}^t e^{-m_k m_k^{x_k}} / x_k!$$

が得られますが, 簡単な計算により

$$l_4 = \prod_{j=1}^s (N_j! e^{N_j} / N_j^{N_j}) \cdot l_1$$

が得られ, $N_1, \dots, N_s$ が θ に無関係なことに注意して証明が終わります.

いずれにせよここで述べた仕組みで得られた標本は, その尤度関数の本質的な部分が

$$\begin{aligned} & \sum_{ijk} x_{ijk} \log m_{ijk} \\ &= x_{...}u + \sum_i x_{i..}u_1(i) + \sum_j x_{.j.}u_2(j) + \sum_k x_{..k}u_3(k) + \sum_{ij} x_{ij.}u_{12}(ij) + \sum_{ik} x_{i.k}u_{13}(ik) + \sum_{jk} x_{.jk}u_{23}(jk) \\ & \quad + \sum_{ijk} x_{ijk}u_{123}(ijk) \end{aligned}$$

であることがわかります. 以下の議論では, これらの仕組みで得られた標本行列について論ずることにします. また m_{ijk} はすべて 0 でないことを仮定することにした.

2・4 Birch の定理

まず, 尤度関数の形から最小十分統計量は

(L1) においては $x_{i..}, x_{.j.}, x_{..k}$

(L2) においては $x_{ij.}, x_{..k}$

(L3) においては $x_{i.k}, x_{.jk}$

(L4) においては $x_{ij.}, x_{i.k}, x_{.jk}$

です. 簡単のために標本行列の要素 x_{ijk} はすべて 0 でないと仮定します. このとき各モデル (L1), (L2), (L3), (L4) の下で m_{ijk} に対する最尤推定量 $\hat{m}_{ijk}$ は次の条件を満足します.

(1) 最小十分統計量に対応する m_{ijk} の周辺量に対する最尤推定量は標本のそれと一致する.

すなわち (L4) のもとでは $\hat{m}_{ij.} = x_{ij.}$, $\hat{m}_{i.k} = x_{i.k}$, $\hat{m}_{.jk} = x_{.jk}$ が成り立つ.

(2) 最尤推定量 $\hat{m}_{ijk}$ は, (1)によって得られる等式の唯一つの解となっている.

証明は Birch [1] を参照していただきたいのですが, ここでは(1)についてごく簡単に記しておきます.

ポアソン型の仕組みで得られた場合の尤度関数 L が

$$\log L = \sum_{ijk} \{ x_{ijk} \log m_{ijk} - m_{ijk} - \log(x_{ijk}!) \}$$

で与えられますから, θ を任意のパラメータとすると

$$\frac{\partial \log L}{\partial \theta} = \sum_{ijk} \{ (x_{ijk} - m_{ijk}) \frac{\partial}{\partial \theta} \log m_{ijk} \}$$

が成り立つことに注意します. したがって

$$\frac{\partial \log L}{\partial u} = x_{...} - m_{...}$$

$$\frac{\partial \log L}{\partial u_{1(i)}} = \{(x_{i..} - m_{i..}) - (x_{1..} - m_{1..})\}$$

$$\frac{\partial \log L}{\partial u_{12(ij)}} = (x_{ij.} - m_{ij.}) - (x_{i1.} - m_{i1.}) - (x_{1j.} - m_{1j.}) + (x_{11.} - m_{11.})$$

これから, $\hat{m}_{...} = x_{...}$, $\hat{m}_{i..} - x_{i..} = \hat{m}_{1..} - x_{1..}$ すなわち $\hat{m}_{i..} - x_{i..} = (\hat{m}_{...} - x_{...})/I$ から $\hat{m}_{i..} = x_{i..}$, 同様に $\hat{m}_{.j.} = x_{.j.}$, また

$$\begin{aligned} \hat{m}_{ij.} - x_{ij.} &= (\hat{m}_{i1.} - x_{i1.}) + (\hat{m}_{1j.} - x_{1j.}) - (\hat{m}_{11.} - x_{11.}) \\ &= (\hat{m}_{i..} - x_{i..})/J + (\hat{m}_{.j.} - x_{.j.})/I - (\hat{m}_{...} - x_{...})/IJ \end{aligned}$$

すなわち, $\hat{m}_{ij.} = x_{ij.}$

他についても同様に証明できます.

ここでモデル (L1), (L2), (L3) についての最尤推定量はそれぞれ $\hat{m}_{ijk} = x_{i..}x_{.j.}x_{...k}/x_{...}$, $\hat{m}_{ijk} = x_{ij.}x_{..k}/x_{...}$, $\hat{m}_{ijk} = x_{i.k}x_{.jk}/x_{..k}$ であることを注意します.

3. 逐次近似法による解の構成

3・1 2・4において, モデル (L4) の下では最尤推定量 $\{m_{ijk}\}$ は

$$(*) \begin{cases} m_{ij.} = x_{ij.} \\ m_{i.k} = x_{i.k} \\ m_{.jk} = x_{.jk} \end{cases}$$

を満足しなければならないことを示したのですが, ここでは解の存在と一意性を逐次近似法で証明します. しかもこの方法はあとで述べるように, そのまま計算機の使用が可能ですから実際に使われています.

まず第1段として任意な正の初期値を選びます. ここでは $m_{ijk}^{(0)} = 1$ とします. 次に $x_{ij.}$ を使って

$$m_{ijk}^{(1)} = m_{ijk}^{(0)} x_{ij.} / m_{ij.}^{(0)}$$

を作ります. $m_{ij.}^{(1)} = x_{ij.}$ が成り立つことに注意します. 次に $x_{i.k}$ を使って

$$m_{ijk}^{(2)} = m_{ijk}^{(1)} x_{i.k} / m_{i.k}^{(1)}$$

を作ります. $m_{i.k}^{(2)} = x_{i.k}$ が成り立つことに注意します. 最後に $x_{.jk}$ を使って

$$m_{ijk}^{(3)} = m_{ijk}^{(2)} x_{.jk} / m_{.jk}^{(2)}$$

を作ります. $m_{.jk}^{(3)} = x_{.jk}$ が成り立つことに注意します. この方法を $(r-1)$ 回くり返したとき $\{m_{ijk}^{(3r-2)}, m_{ijk}^{(3r-1)}, m_{ijk}^{(3r)}\}$ が次のように作られています.

$$(1) \quad m_{ijk}^{(3r-2)} = m_{ijk}^{(3r-3)} x_{ij.} / m_{ij.}^{(3r-3)}$$

$$(2) \quad m_{ijk}^{(3r-1)} = m_{ijk}^{(3r-2)} x_{i.k} / m_{i.k}^{(3r-2)}$$

$$(3) \quad m_{ijk}^{(3r)} = m_{ijk}^{(3r-1)} x_{.jk} / m_{.jk}^{(3r-1)}$$

ここで $m_{ij.}^{(3r-2)} = x_{ij.}$, $m_{i.k}^{(3r-1)} = x_{i.k}$, $m_{.jk}^{(3r)} = x_{.jk}$. したがって $m_{...}^{(s)} = x_{...}$ ($s=3r, 3r-1, 3r-2$) が成り立つことに注意します. さて(1), (2), (3)より

$$m_{ijk}^{(3r)} = m_{ijk}^{(3r-3)} (x_{i,jk} / m_{i,jk}^{(3r-1)}) (x_{i,k} / m_{i,k}^{(3r-2)}) (x_{ij} / m_{ij}^{(3r-3)})$$

が得られますが、これをくり返すと

$$f_{jk}^{(r)} = x_{i,jk} / (m_{i,jk}^{(3r-1)} m_{i,jk}^{(3r-4)} \dots m_{i,jk}^{(2)})$$

$$g_{ik}^{(r)} = x_{i,k} / (m_{i,k}^{(3r-2)} m_{i,k}^{(3r-5)} \dots m_{i,k}^{(1)})$$

$$h_{ij}^{(r)} = x_{ij} / (m_{ij}^{(3r-3)} m_{ij}^{(3r-6)} \dots m_{ij}^{(0)})$$

と置いて

$$m_{ijk}^{(3r)} = f_{jk}^{(r)} g_{ik}^{(r)} h_{ij}^{(r)}$$

が最終的に求まります。また

$$m_{ijk}^{(3r-1)} = f_{jk}^{(r-1)} g_{ik}^{(r)} h_{ij}^{(r)}$$

$$m_{ijk}^{(3r-2)} = f_{jk}^{(r-1)} g_{ik}^{(r-1)} h_{ij}^{(r)}$$

が成り立ちます。

$f_{jk}^{(r)}$ について

$$\sum_{jk} x_{i,jk} \log f_{jk}^{(r)} = \sum_{jk} x_{i,jk} \sum_{s=1}^r (\log x_{i,jk} - \log m_{i,jk}^{(3s-3)}) = \sum_{s=1}^r \sum_{jk} x_{i,jk} (\log x_{i,jk} - \log m_{i,jk}^{(3s-1)})$$

を考えますと、まず

$$y_{23}^{(s)} = \sum_{jk} x_{i,jk} (\log x_{i,jk} - \log m_{i,jk}^{(3s-1)}) \geq 0$$

に注目します。これは与えられた $\mathbf{a} = \{a_i\}$ に対して $\sum_i a_i \log b_i$ ($\sum_i b_i = \sum_i a_i$) が $\mathbf{b} = \mathbf{a}$ で最大値をとることからただちにわかります。この事実は以下でくり返し使われます。したがって

$$\sum_{jk} x_{i,jk} \log f_{jk}^{(r)} = \sum_{s=1}^r y_{23}^{(s)}$$

より、 $\sum_{jk} x_{i,jk} \log f_{jk}^{(r)}$ は r に対する増加列であることがわかります。同様に $g_{ik}^{(r)}$, $h_{ij}^{(r)}$ について

$$y_{13}^{(s)} = \sum_{ik} x_{i,k} (\log x_{i,k} - \log m_{i,k}^{(3s-2)})$$

$$y_{12}^{(s)} = \sum_{ij} x_{ij} (\log x_{ij} - \log m_{ij}^{(3s-3)})$$

と定義すると

$$\sum_{ik} x_{i,k} \log g_{ik}^{(r)} = \sum_{s=1}^r y_{13}^{(s)}$$

$$\sum_{ij} x_{ij} \log h_{ij}^{(r)} = \sum_{s=1}^r y_{12}^{(s)}$$

より $y_{13}^{(s)} \geq 0$, $y_{12}^{(s)} \geq 0$ を考え合わせてこれが r に対する増加列であることがわかります。ここで $m_{ijk}^{(3r)} = x_{ijk} \dots$ より

$$\sum_{ijk} x_{ijk} (\log x_{ijk} - \log m_{ijk}^{(3r)}) \geq 0$$

に注意し

$$\sum_{ijk} x_{ijk} \log x_{ijk} \geq \sum_{jk} x_{i,jk} \log f_{jk}^{(r)} + \sum_{ik} x_{i,k} \log g_{ik}^{(r)} + \sum_{ij} x_{ij} \log h_{ij}^{(r)}$$

から $\sum_{s=1}^{\infty} y_{23}^{(s)} < \infty$, $\sum_{s=1}^{\infty} y_{13}^{(s)} < \infty$, $\sum_{s=1}^{\infty} y_{12}^{(s)} < \infty$ が得られます。このうちとくに

$$0 = \lim_{r \rightarrow \infty} y_{23}^{(r)} = \lim_{r \rightarrow \infty} \sum_{jk} x_{i,jk} (\log x_{i,jk} - \log m_{i,jk}^{(3r-1)})$$

に注意します. $\sum_{jk} m_{jk}^{(3r-1)} = x...$ を使い $\sum_{jk} x_{jk} \log u_{jk}$ ($\sum_{jk} u_{jk} = x...$) 形の関数が最大値を $\mathbf{u} = \mathbf{x}(1)$ ($\mathbf{x}(1) = \{x_{jk}\}$) でとることに注意すると

$$\lim_{r \rightarrow \infty} m_{jk}^{(3r-1)} = x_{jk}$$

が得られます. 同様に $\lim_{r \rightarrow \infty} y_{13}^{(r)} = 0$, $\lim_{r \rightarrow \infty} y_{12}^{(r)} = 0$ より

$$\lim_{r \rightarrow \infty} m_{i,k}^{(3r-2)} = x_{i,k}$$

$$\lim_{r \rightarrow \infty} m_{ij}^{(3r-3)} = x_{ij}.$$

が得られます.

さてここで $m_{ijk}^{(r)}$ の収束を示すことにします. いま適当な $\{r\}$ の部分列 $\{s\}$, $\{t\}$ が存在して

$$\lim_{s \rightarrow \infty} m_{ijk}^{(s)} = c_{ijk}$$

$$\lim_{t \rightarrow \infty} m_{ijk}^{(t)} = d_{ijk}$$

が成り立つと仮定します. 上に述べたことから

$$c_{jk} = d_{jk} = x_{jk}$$

$$c_{i,k} = d_{i,k} = x_{i,k}$$

$$c_{ij} = d_{ij} = x_{ij}.$$

がすでにわかっています. また

$$\sum_{ijk} c_{ijk} \log m_{ijk}^{(t)} = \sum_{jk} x_{jk} \log f_{jk}^{(u)} + \sum_{ik} x_{i,k} \log g_{ik}^{(v)} + \sum_{ij} x_{ij} \log h_{ij}^{(w)}$$

(ここで $t=3t'$ に対して $u=v=w=t'$; $t=3t'-1$ に対して $u=t'-1$, $v=w=t'$; $t=3t'-2$ に対して $u=v=t'-1$, $w=t'$) を考えますと

$$\sum_{ijk} c_{ijk} \log d_{ijk} = \lim_{t \rightarrow \infty} \sum_{ijk} c_{ijk} \log m_{ijk}^{(t)} = \sum_{s=1}^{\infty} y_{23}^{(s)} + \sum_{s=1}^{\infty} y_{13}^{(s)} + \sum_{s=1}^{\infty} y_{12}^{(s)}$$

が成り立っています. 同様に

$$\sum_{ijk} c_{ijk} \log c_{ijk} = \lim_{s \rightarrow \infty} \sum_{ijk} c_{ijk} \log m_{ijk}^{(s)} = \sum_{s=1}^{\infty} y_{23}^{(s)} + \sum_{s=1}^{\infty} y_{13}^{(s)} + \sum_{s=1}^{\infty} y_{12}^{(s)}$$

がわかっています. よって

$$\sum_{ijk} c_{ijk} (\log c_{ijk} - \log d_{ijk}) = 0$$

が成り立ちますから, $\mathbf{c} = \mathbf{d}$ が得られ $m_{ijk}^{(r)}$ の収束性が示されたことになります.

$$m_{ijk}^{(\infty)} = \lim_{r \rightarrow \infty} m_{ijk}^{(r)}$$

とおくことにします. すでに

$$m_{ij}^{(\infty)} = \sum_k m_{ijk}^{(\infty)} = \lim_{r \rightarrow \infty} m_{ij}^{(r)} = x_{ij}.$$

$$m_{i,k}^{(\infty)} = x_{i,k}, \quad m_{jk}^{(\infty)} = x_{jk}$$

は証明されています. したがって $m_{ijk}^{(\infty)}$ は方程式 (*) の解であることがわかりました. さらに $\{m_{ijk}^{(\infty)}\}$ がモデル (L4) を満足することは適当な部分列 $\{s\}$ を選び

$$\begin{aligned}\lim_{s \rightarrow \infty} f_{jk}^{(s)} &= f_{jk}^{(\infty)} \\ \lim_{s \rightarrow \infty} g_{ik}^{(s)} &= g_{ik}^{(\infty)} \\ \lim_{s \rightarrow \infty} h_{ij} &= h_{ij}^{(\infty)}\end{aligned}$$

とすることができますから

$$\log m_{ijk}^{(\infty)} = \lim_{s \rightarrow \infty} \log m_{ijk}^{(3s)} = \log f_{jk}^{(\infty)} + \log g_{ik}^{(\infty)} + \log h_{ij}^{(\infty)}$$

が得られて、 $u_{123}=0$ が成り立つことがわかります。以上で条件 (L4) を満足する方程式 (*) の解 $\{m_{ijk}^{(\infty)}\}$ が構成されました。

よって 2・4 と組み合わせますと、モデル (L4) を満足する最尤推定量がこの方法によって得られたことになり、Birch の結果の一つを再び証明したことになっています。ここで注意しておきたいのは、得られた $\{m_{ijk}^{(\infty)}\}$ は初期値のとり方に関係しないということです。このことは Birch の結果(2)より明らかですが、実際計算の上でも簡単な場合には次のようにしてわかります。

正の初期値 $\{m_{ijk}^{(0)}\}$ として $m_{ijk}^{(0)} = a_{ij}b_{ik}$ を選びますと

$$\begin{aligned}m_{ijk}^{(1)} &= b_{ik}x_{ij} / \left(\sum_k b_{ik} \right) \\ m_{ijk}^{(2)} &= \{b_{ik}x_{ij} / \left(\sum_k b_{ik} \right)\} x_{i \cdot k} / \sum_j \{b_{ik}x_{ij} / \left(\sum_k b_{ik} \right)\} = x_{ij} \cdot x_{i \cdot k} / x_{i \cdot \cdot}\end{aligned}$$

したがって $m_{ijk}^{(2)}$ は、すでに $m_{ijk}^{(0)} = 1$ とした場合のそれと一致しています。

3・2 ここで Deming-Stephan の方法について述べます。3・1 で述べた逐次近似の方法についてはいくつもの研究がありますが (たとえば Norton [18], Kastenbaum-Lamphiear [15], Brown [4], Ireland-Kullback [14]), この方法を実用的に使い始めたのは Deming-Stephan [6] です。研究の動機は、1940 年センサスで得られた分割表 $Y=(y_{ijk})$ をたとえば 1947 年センサスで得られた同じ種類の分割表 $X=(x_{ijk})$ とどのように比較するかということにあるようです。すなわち、 $Y=(y_{ijk})$ を試料として $X=(x_{ijk})$ をどのように調整するかということです。Deming-Stephan のこれに対する一つの解答は次のとおりです。

X より一つの行列 $M=(m_{ijk})$ を

$$\begin{aligned}m_{\cdot jk} &= x_{\cdot jk}(y_{\cdot \cdot \cdot} / x_{\cdot \cdot \cdot}) \\ m_{i \cdot k} &= x_{i \cdot k}(y_{\cdot \cdot \cdot} / x_{\cdot \cdot \cdot}) \\ m_{ij \cdot} &= x_{ij \cdot}(y_{\cdot \cdot \cdot} / x_{\cdot \cdot \cdot})\end{aligned}$$

を満足し、かつ

$$s = \sum_{ijk} (m_{ijk} - y_{ijk})^2 / y_{ijk}$$

を最小にするように求め、この M を Y との比較に使用する。

これを通常の Lagrange 法を使って解きますと、Lagrange multiplier を λ_{jk} , $\lambda_{i \cdot k} (\lambda_{i \cdot K} = 0)$, $\lambda_{ij \cdot} (\lambda_{iJ \cdot} = 0, \lambda_{iJ \cdot} = 0)$ として

$$\begin{aligned}m_{ijk} &= y_{ijk}(x_{ij \cdot} / y_{ij \cdot} + \lambda_{jk} - k \text{ AVE } \lambda_{jk} + \lambda_{i \cdot k} - k \text{ AVE } \lambda_{i \cdot k}) \\ &= y_{ijk}(x_{\cdot jk} / y_{\cdot jk} + \lambda_{ij \cdot} - i \text{ AVE } \lambda_{ij \cdot} + \lambda_{i \cdot k} - i \text{ AVE } \lambda_{i \cdot k})\end{aligned}$$

$$=y_{ijk}(x_{i,k}/y_{i,k}+\lambda_{ij}-j\text{AVE}\lambda_{ij}+\lambda_{jk}-j\text{AVE}\lambda_{jk})$$

が得られます。ここで $k\text{AVE}\lambda_{jk}=\sum_k y_{ijk}\lambda_{jk}/y_{ij}$ で他についても同様に定義します。ここで Deming-Stephan は、この逐次解として 3・1 で述べた方法を提案しています。まず X を Y の因子Ⅲの方向に単純に比を使って調整すると $M_1=(m_{ijk}^{(1)})$

$$m_{ijk}^{(1)}=y_{ijk}x_{ij}/y_{ij}$$

が得られます。ここで $m_{ij}^{(1)}=x_{ij}$ ですが、一般にこれからは $m_{jk}^{(1)}=x_{jk}$, $m_{i,k}^{(1)}=x_{i,k}$ は得られません。そこで次に M_1 を Y の因子Ⅱの方向に単純に比を使って調整します。

$$m_{ijk}^{(2)}=m_{ijk}^{(1)}x_{i,k}/m_{i,k}^{(1)}$$

より $M_2=(m_{ijk}^{(2)})$ が得られます。ここで $m_{i,k}^{(2)}=x_{i,k}$ ですが、 $m_{ij}^{(2)}=x_{ij}$, $m_{jk}^{(2)}=x_{jk}$ に一般に成り立っていません。そこで M_2 を Y の因子Ⅰの方向に同様に調整すると

$$m_{ijk}^{(3)}=m_{ijk}^{(2)}x_{jk}/m_{jk}^{(2)}$$

より $M_3=(m_{ijk}^{(3)})$ が得られます。ここで $m_{jk}^{(3)}=x_{jk}$ です。この方法をくり返すことによって、直観的には各因子に対する調整がよくまざり合って求める解に近づいていきます。すなわち初期値 $m_{ijk}^{(0)}$ を y_{ijk} とした場合の 3・1 で述べた方法にほかなりません。Deming-Stephan は、こうして得られた $M_r=(m_{ijk}^{(r)})$ は収束して $M_\infty=(m_{ijk}^{(\infty)})$ が得られるが、これがこの問題に関する解であると簡単に記しています。

3・3 次に実際計算上の注意をします。3・1 で述べましたモデル (L4) を満足する最尤推定量の構成は、そのまま計算機のプログラムに書くことができます。ここでは実際にこのプログラムを組むときの注意を述べます。

まずいつ計算を終了するかについてですが

$$u_r=\max_{i,j,k} |\hat{m}_{ijk}^{(r)}-\hat{m}_{ijk}^{(r-1)}|$$

を考えて、 u_r が適当な与えられた数 (たとえば 1/1000) より小さくなったときに $\hat{m}_{ijk}^{(r)}$ を $\hat{m}_{ijk}$ の近似解とする方法が考えられます。あるいは

$$u_r=\max \{ |\hat{m}_{ij}^{(r)}-x_{ij}|, |\hat{m}_{i,k}^{(r)}-x_{i,k}| \}$$

を計算の終了の指示量とすることも考えられます。もっと計算に結びついて

$$w_r=|\sum_{ijk} x_{ijk} \log \hat{m}_{ijk}^{(r)} - \sum_{ijk} x_{ijk} \log \hat{m}_{ijk}^{(r-1)}|$$

を考えることもできます。いずれの場合にしても本質的な差はありません。一般的にいて (のちに述べる行列の要素が 0 を含む場合を除き)、この逐次近似の収束は早くて $m_{ijk}^{(30)}$ 以上になることは誤差を適当にとるかぎりあまりありません。そこで $3r=45$ 程度で収束しない場合はそれまでで計算を終了して、もう一度モデルについて検討し直すほうが普通だと思われます。

次にこの逐次近似法では x_{ijk} が全部正であることを使っていないことに注意します。すなわち計算上では標本行列の要素のいくつかは 0 であっても方程式 (*) の解は一応構成されています。それではこの場合でもこの逐次近似法による方程式 (*) の解はモデル (L4) を満足する最尤推定量でしょうか? Birch の定理を述べるにあたって、 $x_{ijk} > 0$ は $m_{ijk} > 0$ とともに仮定されていますから、これは考えておかなければならない問題です。

一般的にこの問題を次のように整理してみます. 与えられた $I \times J \times K$ 標本行列 $X = (x_{ijk})$ に対して

$$x_{ijk}^{(n)} = x_{ijk} \quad (x_{ijk} > 0)$$

$$x_{ijk}^{(n)} = 1/n \quad (x_{ijk} = 0)$$

と定義して行列 $X_n = (x_{ijk}^{(n)})$ を考えることにします. X_n に対しては Birch の結果が使えますから, モデル (L4) を満足する最尤推定量 $(\hat{m}_{ijk}^{(n)})$ を考えることができます. そこで問題は

(1) $\hat{m}_{ijk}^{(n)}$ は n に関する収束列か

(2) (1)が成り立つとき, $\hat{m}_{ijk} = \lim_{n \rightarrow \infty} \hat{m}_{ijk}^{(n)}$ とおくと $(\hat{m}_{ijk})$ はモデル (L4) を満足する最尤推定量かということになります. この問題は一般的な条件の下で肯定的です. 標本行列 (x_{ijk}) の要素のいくつかが本質的に得られない場合 (これに対して 0 とおいて解析された場合があります) についての取り扱いに関しては, たとえば Fienberg [9] があります. また Fienberg-Holland [10] はこれらの問題を詳しく研究しています.

4. モデルの選択

4・1 まず簡単に, 一般的に以下で使われる統計量について説明します. 与えられたデータ (この場合 $I \times J \times K$ 分割表 $X = (x_{ijk})$) に対して 2, 3 章で述べた方法によって, 一般に八つの non-hierarchical モデルを考え, このおのおののモデルに対する最尤法による推定量を求めることができます.

さて第二段階として, おのおののモデルが実際のデータ X とどのくらいよく適合しているかを測る必要があります. このような場合に通常使われている統計量は

$$X^2 = \sum_{ijk} (m_{ijk} - x_{ijk})^2 / m_{ijk}$$

または

$$G^2 = 2 \sum_{ijk} x_{ijk} (\log x_{ijk} - \log m_{ijk})$$

です. この統計量 X^2 および G^2 についての標本数が多い場合の近似的な性質についてはよく研究されています (たとえば Rao [19] 第 6 章を参照してください). さしあたってここで必要なことは, X^2 および G^2 がモデルが持つ条件に対応した自由度のカイ 2 乗分布に従うと見なしてよいということです. 自由度については

自由度 = (分割表の要素の数) - (モデルの中のパラメータの数) = (モデルの下で 0 とおいたパラメータの数) が成り立っています. したがってよく知られているように, 右のような自由度についての表 1 が得られます.

表 1 各モデルに対応する自由度

モデル	自 由 度
(L 1)	$IJK - I - J - K + 2$
(L 2)	$(IJ - 1)(K - 1)$
(L 3)	$(I - 1)(J - 1)K$
(L 4)	$(I - 1)(J - 1)(K - 1)$

4・2 次にモデルの選択にあたって有用な統計量 G^2 の分割について述べます. ここで述べる方法は Bishop [2], Goodman [11] によっていますが, この方法は Lancaster [17], Ku-Kullback [16] らによって導入されました.

いま、一つのモデルが他のモデルの特別な場合になっているとします。たとえば (L1) は (L2) の特別な場合になっています。与えられた標本行列 $X=(x_{ijk})$ に対して $(\hat{m}_{ijk}(1))$, $(\hat{m}_{ijk}(2))$ をそれぞれモデル (L1) の下での最尤法による推定量, (L2) の下での最尤推定量とします。ここで

$$G^2(1)=2\sum_{ijk} x_{ijk}(\log x_{ijk}-\log \hat{m}_{ijk}(1))$$

$$G^2(2)=2\sum_{ijk} x_{ijk}(\log x_{ijk}-\log \hat{m}_{ijk}(2))$$

とおきますと, $\hat{m}_{ijk}(2)=x_{ij\cdot}x_{\cdot\cdot k}/x_{\cdot\cdot\cdot}$ ですから, $\hat{m}_{ij\cdot}(2)=x_{ij\cdot}$ が成り立つことに注意し

$$\begin{aligned} G^2(1)-G^2(2) &= 2\sum_{ijk} x_{ijk}(\log \hat{m}_{ijk}(2)-\log \hat{m}_{ijk}(1)) \\ &= 2\sum_{ij} x_{ij\cdot}u_{12}(ij) \\ &= 2\sum_{ij} \hat{m}_{ij\cdot}(2)\cdot u_{12}(ij) \\ &= 2\sum_{ijk} \hat{m}_{ijk}(2)(\log \hat{m}_{ijk}(2)-\log \hat{m}_{ijk}(1)) \end{aligned}$$

したがって $G^2(1)-G^2(2)$ は, 与えられた行列 $M_2=(m_{ijk}(2))$ に対してモデル (L1) がどれほど適当かどうかを調べる統計量とみなすことができます。すなわち統計量 $G^2(1)-G^2(2)$ は近似的に χ^2 -分布に従い, その自由度は $G^2(1)$ の自由度と $G^2(2)$ の自由度の差で与えられます。

このことは (L1), (L2), (L3), (L4) のすべての組についていうことができます。すなわち, モデル (Ll) の下での最尤法による推定量を $M_l=(\hat{m}_{ijk}(l))$ としますと

$$G^2(l)-G^2(l+1) \geq 0$$

は与えられた行列 M_{l+1} に対してモデル (Ll) がどれほど適当かどうかを調べる統計量になっています。こうして一般的に統計量 G^2 はいくつかの項 (意味のある統計量) に分解することができます。たとえば

$$G^2(1)-G^2(3)=(G^2(1)-G^2(2))+(G^2(2)-G^2(3))$$

は 4・3 で詳しく述べるように, モデル (L1) とモデル (L3) とを比較する上で重要な役目をしています。

4・3 ここでは実際に得たデータを元にしてモデルの選択について述べます。表 2 はミネソタ大学の Stoeckeler 教授が, 女子学生, 両親, 学生が選んだ同年配の男子について, その人々の一般の価値に対する基準と実際に住家を選ぶにあたっての価値基準の差について調査したものに基づいています (Stoeckeler-Hasegawa [22])。

一般に (++) , (+) , (±) , (−) , (−−) の順で, 経済性を重んずる立場からおもに審美的な面を強調する立場に移動していますが, ここでは簡単に五つの線型な基準を選んだことにしておきます。そこで因子 I はグループ (1. 女子学生, 2. 母親, 3. 父親, 4. 同年配男子), 因子 II は一般の価値基準, 因子 III は家を選ぶにあたっての価値基準ということになります。このような問題では解析方法を選ぶにあたってさまざまなことが考えられますし, またその中から適当な方法を選ぶことがたいせつなわけですが, ここでは簡単に 2, 3 章で述べた方法をそのまま適用してみることにします。

表2 4 × 5 × 5 分割表の例 (Stoeckeler-Hasegawa [21])

グループⅠ							グループⅡ						
(--)	(-)	(±)	(+)	(++)	計		(--)	(-)	(±)	(+)	(++)	計	
(--)	5	4	4	0	0	13	(--)	1	0	1	0	1	3
(-)	3	6	6	5	3	23	(-)	0	2	0	1	0	3
(±)	1	0	10	9	2	22	(±)	0	2	5	7	4	18
(+)	2	0	4	6	10	22	(+)	0	1	1	12	12	26
(++)	0	0	1	6	13	20	(++)	0	0	2	5	27	34
計	11	10	25	26	28	100	計	1	5	9	25	44	84

グループⅢ							グループⅣ						
(--)	(-)	(±)	(+)	(++)	計		(--)	(-)	(±)	(+)	(++)	計	
(--)	2	1	1	2	1	7	(--)	2	4	0	1	0	7
(-)	0	1	4	1	1	7	(-)	0	6	5	2	2	15
(±)	1	1	3	2	5	12	(±)	2	1	3	8	3	17
(+)	0	0	3	8	9	20	(+)	0	0	4	7	10	21
(++)	0	1	1	8	29	39	(++)	0	0	1	4	9	14
計	3	4	12	21	45	85	計	4	11	13	22	24	74

考えるモデルは8種類あるわけですが、そのうちで次の4種類がこの問題に関して本質的です。

- (1) $\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{23(jk)} + u_{13(ik)}$
- (2) $\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{23(jk)}$
- (3) $\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{23(jk)} + u_{13(ik)}$
- (4) $\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{23(jk)}$

まずどのモデルも u_{23} の項を含んでいますが、これは因子Ⅱと因子Ⅲとの相互関係を表わす量ですから、この問題では除くことはできません。実際に計算してみると、 u_{23} を除いたモデルは完全に不適合の範囲にはいります。さて3で述べた方法によっておのおののモデルに対して最尤法による推定量を求めることができます。また第2段階としてもおのおののモデルが実際のデータとどのくらいよく適合しているかについて G^2 (または X^2) 統計量を使って調べることができます。こうして得られたおのおののモデルに対する G^2 の値は表3に整理されています。

表3 各モデルに対応する統計量 G^2

モデル	G^2	(自由度)
(1) $u + u_1 + u_2 + u_3 + u_{12} + u_{23} + u_{13}$	47.0*	(48)
(2) $u + u_1 + u_2 + u_3 + u_{12} + u_{23}$	59.10	(60)
(3) $u + u_1 + u_2 + u_3 + u_{23} + u_{13}$	70.75	(60)
(4) $u + u_1 + u_2 + u_3 + u_{23}$	120.67	(72)

* 粗い近似値になっています

4・2で述べましたが、実用的に使われている道具は G^2 の分解がおもなものです。この方法はいわゆる nested hierarchy (一方が次のモデルの特別な場合になっている) モデルの列に対してのみ使用が可能なることに注意しま

す。この注意は重要で、この例については

モデルの列 (1), (2), (4)

モデルの列 (1), (3), (4)

が二つの nested-hierarchy モデルの列で、おのおのの列の要素についての比較が可能です。

(2)と(3)との理論上厳密な比較は現在のところ可能ではないようです. したがって分散分析の方法とは異なって, nested-hierarchy モデルの列の取り方によって異なるモデルの選択が可能です.

さて, モデルの列(1), (2), (4)を見ますと, 表 3 より(1), (2)のモデルはよく与えられた $X=(x_{ijk})$ に適合していますが, (4)については有意水準 5 % で不適合になっています. そこで(1)か(2)のモデルのどちらを選ぶかということになります. ここで

$$G^2(\text{モデル}(2)) - G^2(\text{モデル}(1)) = 59.1 - 47.0 = 12.1$$

で, 自由度は $60 - 48 = 12$ に注意しますと有意の差は見られないことになります. ちなみに, 自由度 12 のカイ 2 乗分布の 95 % 点は 21.026, 99 % 点は 26.217, 99.9 % 点は 32.909 です.

そこで次の hierarchy モデルの比較に移りますと, 自由度は $72 - 60 = 12$ であることに注意して

$$G^2(\text{モデル}(4)) - G^2(\text{モデル}(2)) = 102.67 - 59.10 = 43.57$$

で完全に不適合であることがわかります. このような場合はモデルの選択を(2)にすることが適当です. モデル(1)はよく X に適合していますが, より簡単なモデル(2)とに有意の差がみられませんが, 一方, モデル(2)はよく X に適合していますし, かつ同じ nested-hierarchy モデル列の次のモデル(4)との間に有意の差が見られるからです.

もう一方の nested-hierarchy 列(1), (3), (4)について考えてみます. まずモデル(3)もよく X に適合していることに注意します. しかし

$$G^2(\text{モデル}(3)) - G^2(\text{モデル}(1)) = 70.75 - 47.0 = 23.75$$

ですからモデル(1)とモデル(3)とは有意水準を 5 % とすると有意な差が見られます. したがってこの nested-hierarchy モデル列においてはモデル(1)を選択することが適当です. したがってこの場合には, 結果としては異なる nested-hierarchy の中のモデルが全部比較できたことになり, モデル(2)が適当なモデルとして選択されました.

モデル(2)は

$$\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{23(jk)}$$

となっていて重要な情報を提供しています. すなわちこのモデルを選択することによって因子 I と因子 II との相互関係, 因子 II と因子 III との相互関係は無視できませんが, 因子 I と因子 III との相互関係および 3 因子の相互関係は無視できるということです. とくにグループと一般の価値基準との関係が強調されて, 一方ではグループと家の選択に対する価値基準との相互関係があまり見られないということは常識的な結論ですが, もっと詳しく考えてみる必要があります. しかしこの節の初めに述べたように, この例に対するこの種の解析の仕方には, 標本の数とともに十分検討する余地があります. この点に関しては, ここでは議論に立ち入らないことにします (Stoeckeler-Hasegawa [22]).

ここでこの方法を公式的に記しておきます. まず考えられるモデルをいくつかの nested-hierarchy 列にわけます. おのおのの列の中で G^2 統計量の分割により適当なモデルを選択します. その方法は

- (A) $G^2(I)$ によっておのおののモデル自体が X に適合しているかを調べる.

- (B) $G^2(I) - G^2(I+1)$ を複雑なモデルから単純なモデルへと考えていく。はじめて有意の差ができた場合のモデル (I) を選択する。

異なった列に対応するそれぞれの適当なモデルについては、詳しいデータの内容の検討でそのうちで最適なものと思われるものを選択するか、または一番簡単なモデルを選択するか、あるいは一応そのままに結論を保留するということです。

この問題に対する Goodman の詳しい研究 (たとえば [11], [12], [13]) がありますが、現在のところ筆者は十分に内容を消化できないでいます。ここで Bishop の多次元分割表の研究に対する見解を伝えておきたいと思います。それはセンサス局で研究にたずさわっていた長い間に、三つ以上の因子の相互関係を項に含んだモデルをとり上げたことはないというのです。モデルを複雑にすれば適合度はよくなりますが、それだけモデルの意味が明白でなくなります。4次元以上の分割表を取り扱う上でこの見解は考えてみる必要があると思います。

なお詳しい分割表に関する研究は、Bishop-Fienberg-Holland-Mosteller [3] にまとめられて出版されることになっています。

最後になりましたが、多次元分割表の研究を紹介して下さったミネソタ大学の Fienberg 教授と、この研究を熱心に勧めて下さった名古屋大学の飛田武幸教授に心から感謝いたします。

参 考 文 献

ここではごく限られた範囲の文献のみを記します。詳しい文献は Goodman [10], [11] を参照してください。

- [1] Birch, M. W., "Maximum Likelihood in Three-way Contingency Tables," *J. Roy. Statist. Soc., B* **25** (1963), 220-233.
- [2] Bishop, Y. M. M., "Full Contingency Tables, Logits, and Split Contingency Tables," *Biometrics*, **25** (1969), 383-400.
- [3] Bishop, Y. M. M., S. E., Fienberg, P. W. Holland, and F. Mosteller, *Discrete Multivariate Analysis: Theory and Practice* (to appear).
- [4] Brown, D. T., "A note on Approximations to Discrete Probability Distributions," *Inform. Control*, **2** (1959), 386-392.
- [5] Darroch, J. N., "Interactions in Multifactor Contingency Tables," *J. Roy. Statist. Soc., B* **24** (1962), 251-263.
- [6] Deming, W. E. and F. F. Stephan, "On a Least Square Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known," *Ann. Math. Statist.*, **11** (1940), 427-444.
- [7] Fienberg, S. E., "Quasi-Independence and Maximum Likelihood Estimation in Incomplete Contingency Tables," *J. Amer. Statist. Assoc.*, **65** (1970), 1610-1616.
- [8] ———, "The Analysis of Multidimensional Contingency Tables," *Ecology*, **51** (1970), 419-433.
- [9] ———, "The Analysis of Incomplete Multiway Contingency Tables," *Biometrics*, **28** (1972), 177-202.
- [10] Fienberg, S. E. and P. W. Holland, "Methods for Eliminating Zero Counts in Contingency Tables," *Random Counts in Scientific Work Volume 1*, Pennsylvania State University Press, 1970, 233-260.
- [11] Goodman, L. A., "On Partitioning Chi-square and Detecting Partial Association in Three-way Contingency Tables," *J. Roy. Statist. Soc., B* **31** (1969), 486-498.
- [12] ———, "The Multivariate Analysis of Qualitative Data: Interactions among Multiple Classifications," *J. Amer. Statist. Assoc.*, **65** (1970), 226-256.

- [13] ———, "Partitioning of Chi-square, Analysis of Marginal Contingency Tables, and Estimation of Expected Frequencies in Multidimensional Contingency Tables," *J. Amer. Statist. Assoc.*, **66** (1971), 339-344.
- [14] Ireland, C. T. and S. Kullback, "Contingency Tables with Given Marginals," *Biometrika*, **55** (1968), 179-188.
- [15] Kastenbaum, M. A. and D. E. Lamphiear, "Calculation of Chi-square to Test the no Three-factor Interaction Hypothesis," *Biometrics*, **15** (1959), 107-115.
- [16] Ku, H. H. and S. Kullback, "Interaction in Multidimensional Contingency Tables: An Information Theoretic Approach," *J. Res. Nat. Bur. Stand.*, **72 B** (1968), 159-199.
- [17] Lancaster, H. O., "Complex Contingency Tables Treated by the Partition of Chi-square," *J. Roy. Statist. Soc., B* **13** (1951), 242-249.
- [18] Norton, H. W., "Calculation of Chi-square from Complex Contingency Tables," *J. Amer. Statist. Assoc.*, **40** (1945), 241-258.
- [19] Rao, C. R., *Linear Statistical Inference and Its Applications*, John Wiley & Sons, Inc., 1965.
- [20] Roy, S. N. and M. A. Kastenbaum, "On the Hypothesis of No 'Interaction' in a Multiway Contingency Table," *Ann. Math. Statist.*, **27** (1956), 749-757.
- [21] Simpson, E. H., "The Interpretation of Interaction in Contingency Tables," *J. Roy. Statist. Soc., B* **13**, 238-241.
- [22] Stoeckeler, H. S. and M. Hasegawa, "A Means of Identifying Values as Behavioral Potential in Making Consumer Housing Decisions" (under review).