SEMによる分割表の分析

荘島宏二郎
早稲田大学 大学院 文学研究科

目的

構造方程式モデリング (SEM, structural equation modeling, Jöregkog & Sörbom (1993)) は汎用統計モデルであり，さまざまな統計モデルをその下位モデルとして含んでいる．例えば，因子分析モデル，パス解析モデル，時系列解析における自己回帰・移動平均モデル (Van Buuren, 1997)，2パラメタ項目反応モデル (Takane & DeLeeuw, 1987) などである．また，SEMは相関データの分析だけでなく，実験データの分析にも用いることができる (Kano, 2000)．すなわち，ANOVAやANCOVA，MANOVAもSEMの下位モデルである (e.g., Bagozzi, 1977; Bagozzi & Yi, 1989; Kano, 2000; Kühel, 1988; 豊田, 2000).

一方で，対数ー線形モデルは多元分割表を分析する統計手法であり，分割表のどこに交互作用があるかなどを分析することができ，その構造模型はANOVAのそれと酷似している．

本研究では，対数ー線形モデルをSEMの枠組みの中で実行することが可能であることを示すと同時に，その問題点について論じることが目的である．

方法

P 次元分割表において対数ー線形モデルの構造模型は，p 番目の要因 A_p には水準 l_p に属する同時確率 $p_{l_1\cdots l_P}$ を

$$\begin{aligned}\log p_{l_1\cdots l_P} = \mu &+ a_{l_1} + \cdots + a_{l_p} + \cdots + a_{l_P}\\ &+ a^{(1)}_{l_1 l_2} + \cdots + a^{(1)}_{l_{P-1} l_P}\\ &+ \cdots + a^{(s)}_{l_q\cdots l_r} + \cdots + a^{(P-1)}_{l_1\cdots l_P}\end{aligned} \tag{1}$$

のように表現するモデルである．ここで，a_{l_p} は要因 A_p の l_p 番目の水準の主効果であり，$a^{(1)}_{l_1 l_2}$ は要因 $A_1 \times A_2$ の水準がそれぞれ l_1, l_2 における1次の交互作用効果であり，$a^{(s)}_{l_q\cdots l_r}$ は $A_q \times \cdots \times A_r$ の水準がそれぞれ $l_q, \cdots, l_r$ における s 次の交互作用効果であり，$a^{(P-1)}_{l_1\cdots l_P}$ は，$A_1 \times \cdots \times A_P$ の水準がそれぞれ $l_1, \cdots, l_P$ における最も高次の交互作用効果である．ここで

$$\sum_{l_1=1}^{L_1} \cdots \sum_{l_P=1}^{L_P} p_{l_1\cdots l_P} = 1 \tag{2}$$

である．L_1 は要因 A_1 の水準数である．

また

$$\sum_{l_1=1}^{L_1} a_{l_1} = \cdots = \sum_{l_P=1}^{L_P} a_{l_P} = 0 \tag{3}$$

$$\sum_{l_1=1}^{L_1} a^{(1)}_{l_1 l_2} = \sum_{l_2=1}^{L_2} a^{(1)}_{l_1 l_2} = \cdots = \sum_{l_P=1}^{L_P} a^{(1)}_{l_{P-1} l_P} = 0 \tag{4}$$

$$\vdots$$

$$\sum_{l_q=1}^{L_q} a^{(s)}_{l_q\cdots l_r} = \cdots = \sum_{l_r=1}^{L_r} a^{(s)}_{l_q\cdots l_r} = 0 \tag{5}$$

$$\vdots$$

$$\sum_{l_1=1}^{L_1} a^{(P-1)}_{l_1\cdots l_P} = \cdots = \sum_{l_P=1}^{L_P} a^{(P-1)}_{l_1\cdots l_P} = 0 \tag{6}$$

という制約が入る．これを Birch(1963) の制約式という．

(1) 式は全ての交互作用項を検討対象としているので飽和モデルという．しかし，実際の分析において，対数ー線形モデルは，しばしば高次の交互作用効果を考慮しない階層モデル (Bishop, Fienberg & Holland, 1975) が採用される．また，(1) 式は分散分析モデルにおいて誤差項を仮定しないモデルであるということが見てとれる．

ここで，$\log p_{l_1\cdots l_P}$ を縦に並べたベクトル

$$y = [\log p_{1\cdots 1} \cdots \log p_{l_1\cdots l_P} \cdots \log p_{L_1\cdots L_P}]' \tag{7}$$

を用意し，推定すべき母数を縦に並べたベクトル

$$\theta = [\mu\ a_1 \cdots a_{l_p} \cdots a_{L_P}\ a^{(1)}_{11} \cdots a^{(s)}_{l_q\cdots l_r} \cdots a^{(P-1)}_{L_1\cdots L_P}]' \tag{8}$$

と，さらにそれに対応するデザイン行列 X を用意すると

$$y = X\theta \tag{9}$$

と表現することができる．ここで，X は，たとえば，$P = 3, L_1 = L_2 = L_3 = 2$ のとき

$$X = \begin{bmatrix} 110101010001000100010000000 \\ 110100110000100010001000000 \\ 110011001001000001000100000 \\ 110010101000100000100010000 \\ 101101000100010100000001000 \\ 101100100100001010000000100 \\ 101011000010010001000000010 \\ 101010100010001000100000001 \end{bmatrix} \tag{10}$$

となっている．

また，(9) 式は，一般線形モデルの構造模型において誤差変数ベクトルを仮定しないものとなっている．

ここで，X において，それぞれの列ベクトルを

$$X = [x_{\mu}, x_{a_1}, \cdots, x_{a_{L_P}}, \cdots, x_{a^{(1)}_{11}}, \cdots, x_{a^{(s)}_{l_q\cdots l_r}}, \cdots, x_{a^{(P-1)}_{L_1\cdots L_P}}] \tag{11}$$

と書くことにし，観測変数と見なす．SEM においては，しばしば観測変数の横ベクトルをただのイタリック体で表記することが多いので，ここでもそれに従って $x'_{\mu} \to x_{\mu}, x'_{a^{(s)}_{l_q\cdots l_r}} \to x_{a^{(s)}_{l_q\cdots l_r}}$ と表記することにする．また，y' も y と表記することにする．

つぎに，$x_{a_1} \cdots x_{a^{(P-1)}_{L_1\cdots L_P}}$ と y を縦に並べた構造変数ベクトル

$$t = [x_{a_1} \cdots x_{a^{(s)}_{l_q\cdots l_r}} \cdots x_{a^{(P-1)}_{L_1\cdots L_P}}\ y]' \tag{12}$$

を用意する．また

$$\mu_t = [L_1^{-1} \cdots (L_q \times \cdots \times L_r)^{-1} \cdots (L_1 \times \cdots \times L_P)^{-1}\ \mu]' \tag{13}$$

$$K = \begin{array}{c} O \\ a_1 \quad \cdots \quad a^{(s)}_{l_q\cdots l_r} \quad \cdots \quad a^{(P-1)}_{L_1\cdots L_P} \quad 0 \end{array} \tag{14}$$

$$\left[\boldsymbol{O} \text{ は } \left(\prod_{p=1}^{P}(L_P+1)-1\right) \times \left(\prod_{p=1}^{P}(L_P+1)\right) \text{ の大きさ}\right]$$

とおき，さらに

$$u = \underset{0}{\begin{bmatrix} x_{\alpha_1} - L_1^{-1} \\ \vdots \\ x_{\alpha^{(s)}_{l_q \cdots l_r}} - (L_q \times \cdots \times L_r)^{-1} \\ \vdots \\ x_{\alpha^{(P-1)}_{L_1 \cdots L_P}} - (L_1 \times \cdots \times L_P)^{-1} \end{bmatrix}} \tag{15}$$

とおけば，(9) 式は

$$t = \mu_t + K t + u \tag{16}$$

と表記することができる．

つぎに，$(I - K)^{-1}$ の存在を仮定し，これを T とおくと

$$E[t] = T \mu_t \tag{17}$$

$$V[t] = T \Sigma_u T' \tag{18}$$

となる．ここで $\Sigma_u = V[u]$ である．(17), (18) 式が対数－線形モデルの平均・共分散構造である．平均・共分散構造が導かれたことによって，SEM の枠組みの中で実行することが理論上は可能であることが示された．

しかし，実際には，t には互いに線形従属している観測変数が存在しているので，分析をする際には無駄な観測変数に対して以下のような操作を行なわなければならない．そうしないと多重共線を引き起こし，尤度が計算できなくなることが知られている (豊田, 2000)．多重共線を避ける方法を論じる上で P 次元分割表における一般的な議論を行いにくいので，以下からは $P = 2, L_1 = 3, L_2 = 2$ の状況を考えることにする．

まず，2 要因の対数－線形モデルは，2 つの要因をそれぞれ A, B としたとき，要因 A には水準 i，要因 B には水準 j に属する同時確率 p_{ij} を

$$\log p_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} \tag{19}$$

と表現する．

このとき，t は

$$t = [x_{\alpha_1}\ x_{\alpha_2}\ x_{\alpha_3}\ x_{\beta_1}\ x_{\beta_2}\ x_{(\alpha\beta)_{11}}\ x_{(\alpha\beta)_{12}}\ x_{(\alpha\beta)_{21}}\ x_{(\alpha\beta)_{22}}\ x_{(\alpha\beta)_{31}}\ x_{(\alpha\beta)_{32}}\ y]'$$

$$= \begin{array}{c} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\ \begin{matrix} \log p_{11} & \log p_{12} & \log p_{21} & \log p_{22} & \log p_{31} & \log p_{32} \end{matrix} \end{array} \tag{20}$$

となっており，また，μ_t, K, u はそれぞれ

$$\mu_t = [3^{-1}\ 3^{-1}\ 3^{-1}\ 2^{-1}\ 2^{-1}\ 6^{-1}\ 6^{-1}\ 6^{-1}\ 6^{-1}\ 6^{-1}\ 6^{-1}\ \mu]' \tag{21}$$

$$K = \begin{bmatrix} O_{11 \times 12} \\ \alpha_1\ \alpha_2\ \alpha_3\ \beta_1\ \beta_2\ (\alpha\beta)_{11}\ (\alpha\beta)_{12}\ (\alpha\beta)_{21}\ (\alpha\beta)_{22}\ (\alpha\beta)_{31}\ (\alpha\beta)_{32}\ 0 \end{bmatrix} \tag{22}$$

$$
u = \begin{bmatrix} x_{a_1} - 3^{-1} \\ x_{a_2} - 3^{-1} \\ x_{a_3} - 3^{-1} \\ x_{b_1} - 2^{-1} \\ x_{b_2} - 2^{-1} \\ x_{(ab)_{11}} - 6^{-1} \\ x_{(ab)_{12}} - 6^{-1} \\ x_{(ab)_{21}} - 6^{-1} \\ x_{(ab)_{22}} - 6^{-1} \\ x_{(ab)_{31}} - 6^{-1} \\ x_{(ab)_{32}} - 6^{-1} \end{bmatrix}' \tag{23}
$$

となっている．ここで，(20) 式において $x_{a_1}; x_{b_1}; x_{(ab)_{11}}; x_{(ab)_{21}}; x_{(ab)_{31}}; x_{(ab)_{32}}$ は他の観測変数によって，それぞれ順に，例えば

$$
\begin{aligned}
f_{a_1} &= 1 - x_{a_2} - x_{a_3} && (24)\\
f_{b_1} &= 1 - x_{b_2} && (25)\\
f_{(ab)_{11}} &= f_{a_1} - x_{(ab)_{12}} && (26)\\
f_{(ab)_{21}} &= x_{a_2} - x_{(ab)_{22}} && (27)\\
f_{(ab)_{31}} &= x_{a_3} - f_{(ab)_{32}} && (28)\\
f_{(ab)_{32}} &= f_{b_1} - f_{(ab)_{11}} - f_{(ab)_{21}} && (29)
\end{aligned}
$$

のように表現することができる．これは 1 例である．(24) から (29) 式のようにすると，$t; \mu_t; K; u$ はそれぞれ

$$
t_{[} = [f_{a_1}\ x_{a_2}\ x_{a_3}\ f_{b_1}\ x_{b_2}\ f_{(ab)_{11}}\ x_{(ab)_{12}}\ f_{(ab)_{21}}\ x_{(ab)_{22}}\ f_{(ab)_{31}}\ f_{(ab)_{32}}\ y]' \tag{30}
$$

$$
\mu_{[} = [1\ 3^{-1}\ 3^{-1}\ 1\ 2^{-1}\ 0\ 6^{-1}\ 0\ 6^{-1}\ 0\ 0\ \mu]' \tag{31}
$$

$$
K'_{[} = \begin{array}{c} \begin{bmatrix}
0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0
\end{bmatrix} \\
\begin{array}{cccccccccccc} a_1 & a_2 & a_3 & b_1 & b_2 & (ab)_{11} & (ab)_{12} & (ab)_{21} & (ab)_{22} & (ab)_{31} & (ab)_{32} & ' \end{array}
\end{array} \tag{32}
$$

$$
u_{[} = \begin{bmatrix} 0 \\ x_{a_2} - 3^{-1} \\ x_{a_3} - 3^{-1} \\ 0 \\ x_{b_2} - 2^{-1} \\ 0 \\ x_{(ab)_{12}} - 6^{-1} \\ 0 \\ x_{(ab)_{22}} - 6^{-1} \\ 0 \\ 0 \end{bmatrix}' \tag{33}
$$

となり，改めて

$$t_{[} = ñ_{[} + K_{[}t_{[} + u_{[} \tag{34}$$

と表記することが可能である．ここで，$(I\ Ä\ K_{[})^{Ä1}$ の存在を仮定し，これを $T_{[}$ とおけば

$$t_{[} = T_{[}(ñ_{[} + u_{[}) \tag{35}$$

となる．

さらに，$t_{[}$ から観測変数のみを取り出す選択行列

$$G = \begin{bmatrix} 0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&1 \end{bmatrix} \tag{36}$$

を用いれば，$t_{[}$ から $x_{a_2}; x_{a_3}; x_{b_2}; x_{(ab)_{21}}; x_{(ab)_{22}}; y$ を取り出すことができ

$$\begin{aligned} v &= [x_{a_2}\ x_{a_3}\ x_{b_2}\ x_{(ab)_{21}}\ x_{(ab)_{22}}\ y]^0 \\ &= Gt_{[} \\ &= GT_{[}(ñ_{[} + u_{[}) \end{aligned} \tag{37}$$

となる．したがって

$$E[v] = GT_{[}ñ_{[} \tag{38}$$

$$V[v] = GT_{[}Ü_{u_{[}}T_{[}^{0}G^{0} \tag{39}$$

を得る．ここで $Ü_{u_{[}} = V[u_{[}]$ である．これが，多重共線を避けるために必要最低限の観測変数を用いたときの対数ー線形モデルの平均・共分散構造であり，したがって，SEM の枠組みの中で実行することが可能であることが示された．なお，上で導いた平均・共分散構造は，RAM(Reticular Action Model, McArdle, 1980; McArdle & McDonald, 1984) と呼ばれる表現方法である．

以上までは $P = 2; L_1 = 3; L_2 = 2$ の状況を考えてきたわけであるが，どんな要因数，水準数の対数ー線形モデルでも多重共線に配慮し選択行列を適切に用いれば，SEM の枠組みの中で実行することが可能である．

一般に，要因数が P でそれぞれの要因の水準数が $L_1; ÅÅÅ; L_p; ÅÅÅ; L_P$ のとき，全ての p に対して，主効果に対応する $x_{a_1}; ÅÅÅ; x_{a_{l_p}}; ÅÅÅ; x_{a_{L_p}}$ のうち，1 つは不要である．また，s 次の交互作用項に対応する $x_{a^{(s)}_{1ÅÅÅ1}}; ÅÅÅ; x_{a^{(s)}_{l_qÅÅÅl_r}}; ÅÅÅ; x_{a^{(s)}_{L_qÅÅÅL_r}}$ のうち，必要なのは $(L_q\ Ä\ 1)\ Ç\ ÅÅÅÇ\ (L_r\ Ä\ 1)$ 個だけである．

必要な数だけ観測変数として残しておき，不必要な観測変数は残りの観測変数によって表現し，それらを適切に組み合わせることによって，P 次元分割表における対数ー線形モデルを既存の SEM のソフトウェアで実行することが可能である．

分析

Table 1 はタイタニック号のデータである．2201 人に関して，生存した人，死亡した人をその人の持つ属性に関してカウントしたものである．CLASS として，乗務員，1 等船客，2 等船客，3 等船客の 4 水準が有り，SEX として男性，女性の 2 水準が有る．また，ALIVE として生存，死亡したかについて 2 水準が有り，そのときの各条件における人数が FREQ であり，基準変数である．

Table 1 に対して，SAS の CATMOD プロシジャ(付録 1) で飽和モデルを分析した結果は Table 2 の通りである．

次に，Table 1 のデータに対して，以上で提案された手続きによって，対数ー線形モデルを SAS の CALIS プロシジャ(Hatcher, 1994) を用いて分析した (付録 2)．なお，CALIS では，RAM，COSAN，LINEQS

Table 1: TITANIC DATA

FREQ	CLASS	SEX	ALIVE
3	乗務員	女性	死亡
20	乗務員	女性	生還
670	乗務員	男性	死亡
192	乗務員	男性	生還
4	1 等船客	女性	死亡
141	1 等船客	女性	生還
118	1 等船客	男性	死亡
62	1 等船客	男性	生還
13	2 等船客	女性	死亡
93	2 等船客	女性	生還
154	2 等船客	男性	死亡
25	2 等船客	男性	生還
106	3 等船客	女性	死亡
90	3 等船客	女性	生還
422	3 等船客	男性	死亡
88	3 等船客	男性	生還

Table 2: CATMOD による分析結果

Parameter	Estimate	S.E.	Pr
$\hat{c}_0$	Ä0:1684	0.1244	
$\hat{c}_1$	Ä0:3250	0.1099	< :01
$\hat{c}_2$	Ä0:2948	0.0865	< :01
$\hat{s}_0$	Ä0:7969	0.0573	< :01
$\hat{a}_0$	Ä0:1240	0.0573	< :05
$(\widehat{cs})_{00}$	Ä1:1207	0.1244	< :01
$(\widehat{cs})_{10}$	0:1562	0.1099	
$(\widehat{cs})_{20}$	0:5073	0.0865	< :01
$(\widehat{ca})_{00}$	Ä0:0378	0.1244	
$(\widehat{ca})_{10}$	Ä0:6057	0.1099	< :01
$(\widehat{ca})_{20}$	0:0866	0.0865	
$(\widehat{sa})_{00}$	Ä0:7839	0.0573	< :01
$(\widehat{csa})_{000}$	Ä0:0028	0.1244	
$(\widehat{csa})_{100}$	Ä0:2676	0.1099	< :01
$(\widehat{csa})_{200}$	0:1625	0.0865	

という 3 つのスタイルによって SEM を実行することができる．COSAN は，McDonald(1978) の COSAN モデルという表現方法を利用し，LINEQS は，Bentler & Weeks(1980) の EQS モデルという表現方法を利用している．今回は，LINEQS ステートメントにより，母数の推定をおこなった．平均・共分散構造の表現方法は異なっていても，モデルの表現力という観点からは，3 者は同一のモデルである (豊田, 1991).

用いた推定法は非重み付き最小 2 乗法 (ULS) であった．ULS を用いた理由は，飽和モデルは観測変数の数が標本数よりも大きくなってしまうという標本数条件 (豊田, 1992) に抵触してしまうからである．母数の推定値は Table 3 のようになった．

Table 2 と Table 3 を比較して，SEM のソフトウェアで対数ー線形モデルの母数を推定できていることが分かった．Table 3 において標準誤差が記載されていないのは，ULS を用いたことにもよるが，それ

Table 3: CALIS による分析結果

Parameter	Estimate
$\hat{c}_0$	Ä0:1684
$\hat{c}_1$	Ä0:3250
$\hat{c}_2$	Ä0:2948
$\hat{s}_0$	Ä0:7969
$\hat{a}_0$	Ä0:1240
$(\widehat{cs})_{00}$	Ä1:1207
$(\widehat{cs})_{10}$	0:1562
$(\widehat{cs})_{20}$	0:5073
$(\widehat{ca})_{00}$	Ä0:0378
$(\widehat{ca})_{10}$	Ä0:6057
$(\widehat{ca})_{20}$	0:0866
$(\widehat{sa})_{00}$	Ä0:7839
$(\widehat{csa})_{000}$	Ä0:0028
$(\widehat{csa})_{100}$	Ä0:2676
$(\widehat{csa})_{200}$	0:1625

に関連した議論は考察の②で行なった.

考察

SEM のソフトウェアで対数ー線形モデルが実行可能であることを示したが，課題がいくつか残っていると思われる．それらを挙げると

① プログラムが容易でない.

② 標準誤差を検討できない.

であろう.

①について，P 次元分割表において，P が大きくなればなるほど，SEM のソフトウェアでプログラミングするのが大変になる．まず，多重共線を避けるために，冗長な観測変数を便宜上，潜在変数と見たててプログラムを書かなくてはならない．その作業は，P が大きいときに煩雑である．つぎに，Birch(1963) の制約式をプログラムに盛り込むのも P が大きいときには大変である．ただし，CALIS だと，若干，その労が省略される.

②については，2 つの主な理由があると思われる．それらは

ｉ 0,1 で構成される外生的な観測変数を確率変数と見なさざるを得ない.

ⅱ 総度数 N が考慮されていない.

である.

ｉについて，(38), (39) 式は外生的な観測変数を確率変数として見なした上での平均・共分散構造である．しかし，実際には，0,1 のみで構成される観測変数は，分析者が定めたデザイン行列であって，非確率変数である．すなわち，本当ならば 2 値の外生的観測変数を所与としたもとで y の条件付き平均・共分散構造を導くべきである．しかし，それには問題がある．まず，ñ(= ñ^0)，x = [x$_{a_1}$ ÅÅÅx$_{a^{(P\,Ä\,1)}_{L_1 ÅÅÅ L_P}}$]^0 と

$$å = [a_1 \ ÅÅÅa^{(s)}_{l_q ÅÅÅ r} \ ÅÅÅa^{(P\,Ä\,1)}_{L_1 ÅÅÅ L_P}]^0 \tag{40}$$

を用いると

$$y|x = ñ + å^0 x \tag{41}$$

と表現できる．ここで，x は非確率変数である．すると，(41) 式から明らかなように，$y|x$ には誤差項がない (対数ー線形モデルには誤差項がない) ために，条件付き平均・(共) 分散構造が

$$E[y|x] = \mu + \beta' x \tag{42}$$

$$V[y|x] = 0 \tag{43}$$

となってしまう．共分散構造が O であると，最適化すべき目的関数が生成できない．したがって，SEM の枠組みの中で対数ー線形モデルを実行しようと思えば，やむをえず，x を確率変数と見なさざるを得ない．外生的な観測変数を確率変数と見なそうが，非確率変数と見なそうが，母数の点推定値は変わらないことが知られている (狩野, 1999, p.54; Kano, 2000) ものの，標準誤差は一致するとは限らず，参考程度になるだけである (狩野, 1999, p.56).

ii について，対数ー線形モデル分析において，N の大きさはパラメタの推定値の標準誤差に決定的な影響を与える．たとえば，Table 2 において，$\hat{c}_0$ の標準誤差は 0.1244 であるが，試みに総度数を 100 倍にして N = 220100 のもとで分析を行なうと標準誤差は 10 分の 1 の 0.01244 となる．

総度数 N を考慮して標本数条件を回避した階層モデルの CALIS プログラムを書き，最尤推定法や一般化最小二乗法を用いることは可能である．その際，母数の標準誤差は出力される．しかし，N = 2201 であろうが N = 220100 であろうが，標準誤差の値は変わらない．というより，対数ー線形モデルにおける母数の標準誤差と SEM における標準誤差は性質の異なるものである．というのは，対数ー線形モデルの対数尤度関数は

$$f(\theta) = \sum_{l_1=1}^{L_1} \cdots \sum_{l_P=1}^{L_P} N_{l_1 \cdots l_P} \log p_{l_1 \cdots l_P} \tag{44}$$

であり，SEM における対数尤度関数とは異なる．対数尤度関数が異なれば，フィッシャー行列も異なるので，標準誤差も違ってしまう．なお，$N_{l_1 \cdots l_P}$ は，要因 $A_1 \times \cdots \times A_P$ のそれぞれの水準 $l_1, \cdots, l_P$ に属する観測度数である．

結論

本研究によって，対数ー線形モデルの母数の推定値を SEM のソフトウェアで得ることができることが示された．ただし，考察で論じた理由により，SEM のソフトウェアで対数ー線形モデルを実際に分析するメリットは現時点ではない．最も本質的な理由はパラメタの推定値の標準誤差が SEM では評価できないことにあるだろう．しかし，本研究は，SEM の汎用性の 1 つの傍証を示したと言える．

引用文献

Bagozzi, R.P. (1977). Structural equation models in experimental research. Journal of Marketing Research, 14, 271–284.

Bagozzi, R.P. & Yi, Y. (1989). On the use of structural equation models in experimental design. Journal of Marketing Research, 26, 271–284.

Bentler, P.M. & Weeks, D.G. (1980). Linear structural equations with latent variables. Psychometrika, 45, 289–308.

Birch, M.W. (1963). Maximum likelihood in three-way contingency table. Journal of the Royal Statistical Society Series B, 25, 220–233.

Bishop, Y., Fienberg, S.E. & Holland, P.W. (1975). Discrete Multivariate Analysis : Theory and Practice. Cambridge, MA : MIT Press.

Hatcher, L. (1994). A Step–by–Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling. Cary, NC : SAS Institute.

Jöreskog, K.G. & Sörbom, D. (1993). LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language. Hillsdale, NJ : Lawrence Erlbaum Associates.

狩野裕. (1999,7). 共分散構造分析. 日本統計学会チュートリアルセミナー資料. 於: 岡山理科大学.

Kano, Y. (2000). Structural equation modeling for experimental data. In B. Cudek, S. Toit & D. Sörbom (Eds.), Structural Equation Models : Present and Future. Chicago, IL : SSI. pp.381{402.

Kühel, S.M. (1988). Testing MANOVA designs with LISREL. Sociological Methods and Research, 16, 504{523.

McArdle, J.J. (1980). Causal modeling applied to psychonomic systems simulation. Behavior Research Methods and Instrumentation, 12, 193{209.

McArdle, J.J. & McDonald, R.P. (1984). Some algebraic properties of the reticular action model for moment structures. British Journal of Mathematical and Statistical Psychology, 37, 234{251.

McDonald, R.P. (1978). A simple comprehensive model for the analysis of covariance structures. British Journal of Mathematical and Statistical Psychology, 31, 59{72.

Takane,Y. & DeLeeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. Psychometrika, 52, 393{408.

豊田秀樹. (1991). 共分散構造分析の下位モデルとその適用例. 教育心理学研究, 38, 438{444.

豊田秀樹. (1992). SAS による共分散構造分析. 東京大学出版会.

豊田秀樹. (2000). 共分散構造分析 [応用編] ー構造方程式モデリングー. 朝倉書店.

Van Buuren, A. (1997). Fitting ARMA time series by structural equation models. Psychometrika, 62, 215{236.

付録 1

付録 1 で，本分析で用いた CATMOD(SAS) のプログラムを掲載する.

```
title 'TITANIC by CATMOD';run;
data titanic;
   input freq class sex alive;
cards;
003 0 0 0
020 0 0 1
670 0 1 0
192 0 1 1
004 1 0 0
141 1 0 1
118 1 1 0
062 1 1 1
013 2 0 0
093 2 0 1
154 2 1 0
025 2 1 1
106 3 0 0
090 3 0 1
422 3 1 0
088 3 1 1
;run;
/* ------------------------------------ */
proc catmod data=titanic;
   weight freq;
   model class*sex*alive=_response_;
   loglin class|sex|alive;
run;
```

付録 2

付録 2 で，本分析で用いた CALIS(SAS) のプログラムを掲載する．

```
title 'TITANIC by CALIS';run;
data titanic;
   input @1 freq 3. @5 (xc0-xc3)(1.) @9 (xs0-xs1)(1.) @11 (xa0-xa1)(1.)
         @13 (xcs00-xcs01)(1.) @15 (xcs10-xcs11)(1.)
         @17 (xcs20-xcs21)(1.) @19 (xcs30-xcs31)(1.)
         @21 (xca00-xca01)(1.) @23 (xca10-xca11)(1.)
         @25 (xca20-xca21)(1.) @27 (xca30-xca31)(1.)
         @29 (xsa00-xsa01)(1.) @31 (xsa10-xsa11)(1.)
         @33 (xcsa000-xcsa001)(1.) @35 (xcsa010-xcsa011)(1.)
         @37 (xcsa100-xcsa101)(1.) @39 (xcsa110-xcsa111)(1.)
         @41 (xcsa200-xcsa201)(1.) @43 (xcsa210-xcsa211)(1.)
         @45 (xcsa300-xcsa301)(1.) @47 (xcsa310-xcsa311)(1.);
   prop=freq/2201;
   y=log(prop);
cards;
003 10001010100000001000000010001000000000000000
020 10001001100000000100000001000100000000000000
670 10000110010000001000000000100010000000000000
192 10000101010000000100000000010001000000000000
004 01001010001000000010000010000000100000000000
141 01001001001000000001000001000000010000000000
118 01000110000100000010000000100000001000000000
062 01000101000100000001000000010000000100000000
013 00101010000010000000100010000000000010000000
093 00101001000010000000010001000000000001000000
154 00100110000001000000100000100000000000100000
025 00100101000001000000010000010000000000010000
106 00011010000000100000001010000000000000001000
090 00011001000000100000000101000000000000000100
422 00010110000000010000001000100000000000000010
088 00010101000000010000000100010000000000000001
;run;
/* ----------------------------------------------------------- */
proc calis data=titanic ucov aug short nomod method=uls
           privec maxiter=5000;
     var y xc0-xc2 xs0 xa0
           xcs00 xcs10 xcs20 xca00 xca10 xca20 xsa00
           xcsa000 xcsa100 xcsa200;
   lineqs
     fc3=intercept-xc0-xc1-xc2,
     fs1=intercept-xs0,
     fa1=intercept-xa0,
     fcs01=xc0-xcs00,
     fcs11=xc1-xcs10,
     fcs21=xc2-xcs20,
     fcs30=xs0-xcs00-xcs10-xcs20,
     fcs31=fc3-fcs30,
     fca01=xc0-xca00,
     fca11=xc1-xca10,
```

```
    fca21=xc2-xca20,
    fca30=xa0-xca00-xca10-xca20,
    fca31=fc3-fca30,
    fsa01=xs0-xsa00,
    fsa10=xa0-xsa00,
    fsa11=fs1-fsa10,
    fcsa001=xcs00-xcsa000,
    fcsa010=xca00-xcsa000,
    fcsa011=fcs01-fcsa010,
    fcsa101=xcs10-xcsa100,
    fcsa110=xca10-xcsa100,
    fcsa111=fcs11-fcsa110,
    fcsa201=xcs20-xcsa200,
    fcsa210=xca20-xcsa200,
    fcsa211=fcs21-fcsa210,
    fcsa300=xsa00-xcsa000-xcsa100-xcsa200,
    fcsa301=fcs30-fcsa300,
    fcsa310=fca30-fcsa300,
    fcsa311=fcs31-fcsa310,
    y=mu intercept+c0 xc0+c1 xc1+c2 xc2+c3 fc3
      +s0 xs0+s1 fs1+a0 xa0+a1 fa1
      +cs00 xcs00+cs01 fcs01
      +cs10 xcs10+cs11 fcs11
      +cs20 xcs20+cs21 fcs21
      +cs30 fcs30+cs31 fcs31
      +ca00 xca00+ca01 fca01
      +ca10 xca10+ca11 fca11
      +ca20 xca20+ca21 fca21
      +ca30 fca30+ca31 fca31
      +sa00 xsa00+sa01 fsa01
      +sa10 fsa10+sa11 fsa11
      +csa000 xcsa000+csa001 fcsa001
      +csa010 fcsa010+csa011 fcsa011
      +csa100 xcsa100+csa101 fcsa101
      +csa110 fcsa110+csa111 fcsa111
      +csa200 xcsa200+csa201 fcsa201
      +csa210 fcsa210+csa211 fcsa211
      +csa300 fcsa300+csa301 fcsa301
      +csa310 fcsa310+csa311 fcsa311+ey;
std
  ey=0.0;
/* Birch's constraints */
  c0=0.0-c1-c2-c3;
  s0=0.0-s1;
  a0=0.0-a1;
  cs00=0.0-cs10-cs20-cs30;
  cs01=0.0-cs11-cs21-cs31;
  cs00=0.0-cs01;
  cs10=0.0-cs11;
  cs20=0.0-cs21;
  cs30=0.0-cs31;
  ca00=0.0-ca10-ca20-ca30;
```

```
    ca01=0.0-ca11-ca21-ca31;
    ca00=0.0-ca01;
    ca10=0.0-ca11;
    ca20=0.0-ca21;
    ca30=0.0-ca31;
    sa00=0.0-sa01;
    sa10=0.0-sa11;
    sa00=0.0-sa10;
    sa01=0.0-sa11;
    csa000=0.0-csa100-csa200-csa300;
    csa001=0.0-csa101-csa201-csa301;
    csa010=0.0-csa110-csa210-csa310;
    csa011=0.0-csa111-csa211-csa311;
    csa000=0.0-csa010;
    csa001=0.0-csa011;
    csa100=0.0-csa110;
    csa101=0.0-csa111;
    csa200=0.0-csa210;
    csa201=0.0-csa211;
    csa300=0.0-csa310;
    csa301=0.0-csa311;
    csa000=0.0-csa001;
    csa010=0.0-csa011;
    csa100=0.0-csa101;
    csa110=0.0-csa111;
    csa200=0.0-csa201;
    csa210=0.0-csa211;
    csa300=0.0-csa301;
    csa310=0.0-csa311;
run;
```