名義尺度の分割表に対する多重比較法

松 田 眞 一

E-Mail: matsu@nanzan-u.ac.jp

名義尺度の分割表に対する独立性の検定は古くから知られている。しかし，独立性が棄却された場合にその内部構造を調べるための方法論はあまり提案されていない。本論文では閉検定手順を適用して多重比較を行う方法を提案し，その性質について述べる。

1　はじめに

名義尺度の分割表は社会科学を中心としてデータ解析で広く用いられる方法論である。多くの場合はクロス集計表と呼んで，ただその表を視覚的に捉えて偏りを見るためだけに使われているが，2 つの分類の間の関連を客観的に捉えるために独立性の検定を行うのが望ましい。分割表に対する独立性の検定としてはピアソンのカイ 2 乗検定やフィッシャーの正確確率検定などいくつかの方法が提案されており，初等的な統計学の範疇に入るものである。特にカイ 2 乗検定は特別な統計ソフトを用いずとも簡単に計算できるため容易に適用できる。

しかし，実際に独立性の検定を利用した場合，ある問題に突き当たる。それは，検定で棄却され独立ではないという結論が得られたとしても元の表が 2×2 でない限りその意味を理解するのが困難なことである。そのため一般的には，カイ 2 乗統計量を計算するのに用いた各セルの乖離度を使って判断する方法が取られている。「乖離度が最大となるセルは全体が棄却される要因となった」という判断は正しいかもしれないが，2 番目や 3 番目のセルとなると主観的に判断せざるを得ないところがある。

一方，Hirotsu (1983) は主に順序尺度に対する分割表の多重比較法を提案し，名義尺度においても適用可能であることを示している。この方法により検定という客観的な判断でどこに原因があるかつきとめることができるようになったのであるが，名義尺度の場合は順序尺度と違って効率が悪いため検出力の低下を免れないものであった。

そこで，本論文では閉検定手順を用いることによって名義尺度の分割表に対する多重比較法を提案し，その性質を調べる。

2　名義尺度の分割表

まず，本論文で用いる記号を定義しておく。

2 つの要因 A, B はそれぞれ m, n 個の名義尺度で分類されており，$A_1, \ldots, A_m$ と $B_1, \ldots, B_n$ で表されるとする。2 つの分類の組み合わせ（セルと呼ぶことにする）として (A_i, B_j) の性質をもつ個体の観測数を X_{ij} で表す。また，

$$X_{i.} = \sum_{j=1}^{n} X_{ij},\ X_{.j} = \sum_{i=1}^{m} X_{ij},\ N = \sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}$$

とおく。N は定数であるとする。

以上を表にまとめると表 1 のようになる。これが 2 次元の $m \times n$ 分割表である。
本論文の議論は多次元にも同様に拡張可能であるが，ここでは特に触れないでおく。

表 1: 名義尺度の分割表

	B_1	$\cdots$	B_n	計
A_1	X_{11}	$\cdots$	X_{1n}	$X_{1.}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
A_m	X_{m1}	$\cdots$	X_{mn}	$X_{m.}$
計	$X_{.1}$	$\cdots$	$X_{.n}$	N

さて，この分割表に対する独立性の検定とは以下のような帰無仮説に対して行われる検定のことである。

$$H_0 : p_{ij} = p_{i.} \cdot p_{.j} \quad \text{for all } i,j$$

ただし，$p_{ij} = E(X_{ij})/N$，$p_{i.} = \sum_{j=1}^{n} p_{ij}$ ，$p_{.j} = \sum_{i=1}^{m} p_{ij}$ である。
独立性が棄却された場合，いくつかのセルの間にだけ部分的な連関があることになるが，それを交互作用（association）と呼ぶ。（分散分析における交互作用（interaction）とは意味合いが違うので注意する。）分割表の多重比較ではどんな交互作用が存在するのかが知りたいことであるが，本論文では行または列で分割した部分分割表を中心に考え，最終的に 2×2 の部分分割表の交互作用をつきとめることを模索する。
ここで，部分分割表とは元の分割表で特定の列や行を抜き出して構成したもののことである。抜き出した行が $\{i_1, \ldots, i_p\} \subseteq \{1, 2, \ldots, m\}$ で，列が $\{j_1, \ldots, j_q\} \subseteq \{1, 2, \ldots, n\}$ の場合は表 2 のようになる。

表 2: 部分分割表

	B_{j_1}	$\cdots$	B_{j_q}
A_{i_1}	$X_{i_1 j_1}$	$\cdots$	$X_{i_1 j_q}$
$\vdots$	$\vdots$		$\vdots$
A_{i_p}	$X_{i_p j_1}$	$\cdots$	$X_{i_p j_q}$

3　閉検定手順

多重比較法を構成するためには閉検定手順という方法論の説明が不可欠であるのでここで述べる。
Marcus et al. (1976) は次のような閉検定手順という定理を導出した。（Hochberg and Tamhane (1987) や永田・吉田 (1997) を参照。なお，原著では定理にしておらず，以下の形は主に Hochberg and Tamhane (1987) による。）

定理 1 (Marcus et al. (1976))　仮説の族 $\{H_i\ (1 \leq i \leq m)\}$ があるとき，すべての $P \subseteq \{1, 2, \ldots, m\}$ に対して，以下のような仮説を構成する。

$$H_P = \bigcap_{i \in P} H_i$$

そのとき，H_P の棄却を次のように定めた検定の手順はその族の *Type I FWE* を有意水準 α に強く抑制する。

「H_P が棄却される必要十分条件は $P \subseteq Q$ を満たすすべての Q に対して H_Q が同手順の中で有意水準 α で棄却されていること」

ここで，Type I FWE（Familywise Error Rate）とは,「仮説の族の中で真に正しい仮説がどれか一つでも誤って棄却される確率」のことで，正しい仮説の様々な組み合わせに対するその確率の最大値が多重比較法での第 1 種の過誤の確率に相当する。この定理で「強く抑制する[1]」とは，その最大値が抑えられることを意味する。

この定理により仮説の族を明確にするだけで通常の検定を拡張して多重比較法を構成することができるようになった。また，今まで知られていた多重比較法の検出力を上げることも可能である。その例としては Tukey の方法にそれを適用した Tukey-Welcsh の方法がある。

4　提案する分割表の多重比較法

$2 \times n$ 分割表（$n \geq 3$）に対して，提案する方法を述べる。

次のような仮説の族を考える。

$$S = \{p_{1i}/p_{2i} = p_{1j}/p_{2j} \mid 1 \leq i < j \leq n\}$$

すべての仮説が成り立つ場合は独立性の帰無仮説に一致する。

提案する方法を作るために，まず S を拡張する。

S の要素から構成される仮説はいくつかの列の集まりにおいて比率 p_{1i}/p_{2i} が等しいというものになるが，そのうち以下のように一つの集まりでのみ構成される仮説の族を S' とおく。

$$p_{1i_1}/p_{2i_1} = \cdots = p_{1i_k}/p_{2i_k}, \quad i_1, \ldots, i_k \subseteq \{1, 2, \ldots, n\}$$

ここで，k を仮説の大きさと呼ぶことにする。

この S' に閉検定手順を適用し，個々の仮説の検定に対して Tukey-Welsch の方法と同様の有意水準を当てはめる。この方法は独立性の検定を拡張した多段階の多重比較法と考えられる。

なお，独立性の検定にはカイ 2 乗検定を用いても Fisher の正確確率検定を用いてもよい。（形式的には混在して用いてもよいが，後で述べるコンソナンスがさらに悪くなる可能性がある。）

[1]単に「抑制する」方法というものは，仮説の族に含まれるすべての仮説が正しい場合だけ Type I FWE を抑えるため，現在推奨されていない。

手順 1 有意水準 α を定める。

手順 2 分割表全体に関して有意水準 α で独立性の検定を行う。棄却されれば $k = n - 1$ とおいて次へ進む。保留されれば S' のすべての仮説を保留して終了する。

手順 3 S' でまだ保留となっていない大きさ k のすべての仮説について，対応する列のみを取り出した部分分割表に対する有意水準 α_k の独立性の検定を行う。ただし，

$$\alpha_{n-1} = \alpha, \quad \alpha_k = 1 - (1 - \alpha)^{k/n} \ (k = 2, \ldots, n - 2)$$

である。

手順 4 $k = 2$ の場合は終了する。

手順 5 S' 内で保留された仮説を含む（列の添え字の集合としては含まれる）大きさ $k - 1$ の仮説をすべて保留する。大きさ k の仮説のうち，一つでも棄却されれば $k - 1$ を新しく k とおき，**手順 3** に戻ってこの手順を繰り返す。

定理 2 上の検定手順は仮説の族 S の *Type I FWE* を有意水準 α に強く抑制する。

証明は簡単で検定統計量の独立性から明らかに Šidák の不等式が満たされることにより，S の要素から構成される仮説で S' に含まれないものも有意水準 α での検定が閉検定手順として矛盾なく行えるからである。

5 提案する方法の拡張の可能性

$m \times n$ 分割表（$m, n \geq 3$）に対してどのような方法を用いればよいか検討する。

列もしくは行のみに関心がある場合は前節の検定手順と全く同様である。（仮説はその列もしくは行内の比率が一致するというものに変わるが，後の手順はそのままである。）

一方，両方に関心がある場合は次の 2 つの場合に分けられる。

A 行と列の分割についてのみ関心がある場合

B さらに細かく 2×2 分割表まで関心がある場合

A については全体の分割表が棄却されているなら独立だと考えて列と行それぞれで前節の検定手順を用いて差し支えないと考えられる。

B に関しては厳密に閉検定手順を用いると大変複雑な手順となる。実行可能な手順は以下の 3 つの方法である。

B-1 行か列のどちらを先に検定するか順序をつけ，先に分割する方で大きさ 2 の仮説まで検定を行った後，もう一方でさらに分割するという手順

B-2 B-1 の手順を行が先の場合も列が先の場合も両方行い，ともに棄却された 2×2 分割表のみ棄却する手順

B-3 元の表に含まれるすべての部分分割表に対して検定を行い，直接閉検定手順を用いて仮説の棄却を決定する手順

ただし、どの方法でも有意水準は前節の決め方を 2 重に用いることで設定する。すなわち，部分分割表の行の数を p，列の数を q とするとその分割表に用いる有意水準 α_{pq} は以下のようになる。

$$\alpha_{pq} = \begin{cases} \alpha & (p = m-1 \text{ and } q = n-1) \\ 1-(1-\alpha)^{q/n} & (p = m-1 \text{ and } q = 2,\ldots,n-2) \\ 1-(1-\alpha)^{p/m} & (p = 2,\ldots,m-2 \text{ and } q = n-1) \\ 1-(1-\alpha)^{pq/(mn)} & (p = 2,\ldots,m-2 \text{ and } q = 2,\ldots,n-2) \end{cases}$$

B-1 では有意水準を厳密な意味で守れないことは確実である。なぜなら、B-2 において食い違いが生じることがあるからである。面倒でも B-2 を用いる方がよいであろう。その B-2 においても実質有意水準がどの程度が近似的に守られるかは詳細に調べる必要がある。

一方，B-3 では有意水準をほぼ満たしているのではないかと期待される。B-3 で検定対象に含まれない仮説のうち問題になるのはいびつな形状のものに限られるが，それらは長方形の仮説すべてを検定するこの手順において間接的に有意水準が守られるのではないかと推測されるからである。

6 数値例

Hirotsu (1983) の例を解析する。

表 3 はガンセンターで取られた職種別ガン患者数で，発見時の重症度ごとに分けられている。

表 3: 職種別ガン患者数

職業	軽度	中程度	重度
1. 専門職・技術職	148	444	86
2. 経営者・公務員	111	352	49
3. 事務系職	645	1911	328
4. 営業職	165	771	119
5. 農林水産業・鉱業従事者	383	1829	311
6. 運輸通信業者	96	293	47
7. 職人	98	330	58
8. 製造業従事者	199	874	155
9. サービス業従事者	59	199	30
10. 定職なし	262	1320	236

職業によって発見時の重症度に違いがあるかという問題を考えてみる。表全体に関して独立性に対するカイ 2 乗統計量を求めてみると 96.39 となり，その p 値は 0.0001 より小さい。すなわち，職業による重症度の違いがあると分かるが，それがどの職業とどの職業の

間であるかは自明ではない。Hirotsu (1983) はこの問題に関して多重比較法を提案し，職業 (3, 10), (3, 5) の間に 1%で有意な差，職業 (3, 4), (3, 8) との間に 5%で有意な差を検出している。その方法の計算は複雑で特殊な表を必要とするためここでは導出の詳細は割愛する。

この表に対し前節の方法を適用してみよう。(行のみに興味があると仮定する。列には順序があるが，そちらには関心がないなら問題はない。)

1. 有意水準 $\alpha = 0.05$ とする。
2. 分割表全体に関して $\chi^2 = 96.39$ となり，自由度 18 だから棄却される。$k = 9$ とおいて次へ進む。
3. 大きさ 9 のすべての仮説について，対応する行のみを取り出した部分分割表に対する χ^2 値は以下のようになる。

行番号	χ^2 値
2 3 4 5 6 7 8 9 10	89.63
1 3 4 5 6 7 8 9 10	90.80
1 2 4 5 6 7 8 9 10	54.61
1 2 3 5 6 7 8 9 10	89.25
1 2 3 4 6 7 8 9 10	74.59
1 2 3 4 5 7 8 9 10	92.31
1 2 3 4 5 6 8 9 10	95.34
1 2 3 4 5 6 7 9 10	91.73
1 2 3 4 5 6 7 8 10	95.16
1 2 3 4 5 6 7 8 9	73.34

   自由度はすべて 16 なので有意水準 $\alpha_9 = 0.05$ の限界値と比較するとどれも棄却される。(p 値はすべて < 0.0001 である。)
4. 大きさ 8 のすべての仮説について部分分割表の検定を行うとどれも棄却される。(全部で 45 通りあり，詳細は省略する。比較に用いる有意水準は $\alpha_8 = 1-(1-\alpha)^{8/10} = 0.0402$ となることに注意する。p 値は 0.0002 以下である。)
5. 大きさ 7 のすべての仮説についてもすべての部分分割表が棄却される。(有意水準は $\alpha_7 = 1-(1-\alpha)^{7/10} = 0.0353$ となる。p 値は 0.0152 以下である。)
6. 大きさ 6 のすべての仮説について部分分割表の検定を行うと次の 4 つの仮説が保留される。(有意水準は $\alpha_6 = 1-(1-\alpha)^{6/10} = 0.0303$ となる。)

行番号	χ^2 値	p 値
1 2 3 6 7 9	5.14	0.8817
1 2 4 6 7 9	19.81	0.0311
1 2 6 7 8 9	18.05	0.0541
4 5 7 8 9 10	17.37	0.0666

   これらの仮説を含む仮説はすべて検定を行わずに保留する。たとえば，行番号 (1, 2, 3, 6, 7, 9) からなる仮説を含むものとしては (1, 2, 3, 6, 7) や (3, 6, 7) や (1, 3) などからなる仮説である。

7. 大きさ 5 の仮説のうちすでに保留されたもの以外について部分分割表の検定を行うと新たに次の 2 つの仮説が保留される。(有意水準は $\alpha_5 = 1 - (1-\alpha)^{5/10} = 0.0253$ となる。)

行番号	χ^2 値	p 値
2 4 7 8 9	16.80	0.0323
4 6 7 8 9	15.70	0.0469

これらの仮説を含む仮説はすべて検定を行わずに保留する。

8. 大きさ 4 の仮説のうちすでに保留されたもの以外について部分分割表の検定を行うと新たに次の 1 つの仮説が保留される。(有意水準は $\alpha_4 = 1 - (1-\alpha)^{4/10} = 0.0203$ となる。)

行番号	χ^2 値	p 値
4 5 6 8	14.32	0.0262

この仮説を含む仮説はすべて検定を行わずに保留する。

9. 大きさ 3 の仮説のうちすでに保留されたもの以外について部分分割表の検定を行うとすべての仮説が棄却される。(有意水準は $\alpha_3 = 1 - (1-\alpha)^{3/10} = 0.0153$ となる。)
10. 大きさ 2 の仮説のうちすでに保留されたもの以外について部分分割表の検定を行うとすべての仮説が棄却される。(有意水準は $\alpha_2 = 1 - (1-\alpha)^{2/10} = 0.0102$ となる。)
結果として棄却される仮説は以下の表のようになる。

行番号	χ^2 値	p 値
1 5	17.99	0.0001
1 10	20.12	< 0.0001
2 5	14.60	0.0007
2 10	17.81	0.0001
3 4	22.27	< 0.0001
3 5	45.12	< 0.0001
3 8	20.13	< 0.0001
3 10	45.50	< 0.0001
6 10	15.61	0.0004

以上の計算は C でプログラムを書いて実行した。もっと小さな表の場合は手計算でも可能であろう。

なお，有意水準 $\alpha = 0.01$ とした場合，最終的に棄却される仮説は次のようになる。(比較する有意水準は $\alpha_2 = 1 - (1-\alpha)^{2/10} = 0.0020$ となる。)

行番号	χ^2 値	p 値
1 5	17.99	0.0001
1 10	20.12	< 0.0001
3 5	45.12	< 0.0001
3 10	45.50	< 0.0001

Hirotsu (1983) と比較してどちらの有意水準でもより多くの交互作用を検出することができた。

7 提案する方法の性質

多段階の多重比較法には次のような満たすべき 2 つの性質が知られている。（ここでいう仮説の族とは，元の族から構成される仮説も含めて拡張されたものである。）

- コヒーレンス（coherence）
  仮説の族の中の一つの仮説が棄却された場合，それに含まれる仮説はすべて棄却されるという性質
- コンソナンス（consonance）
  仮説の族の中の一つの仮説が棄却された場合，それを含む仮説のうちどれか一つは棄却されるという性質

この 2 つの性質のうちコヒーレンスは必ず満たさなければならない性質とされており，コンソナンスは満たすことが推奨される性質となっている。

本論文で提案した方法は閉検定手順を用いて構成されているので，コヒーレンスを満たすことは明らかである。一方，コンソナンスを満たすかどうかは分からないので，具体的なデータを調べた結果，コンソナンスを満たさない例が発見された。そのためコンソナンスは成り立っていないことが分かった。

表 4: コンソナンスを満たさない例

21	12	22
34	16	17
28	31	19

表 4 において全体でのカイ 2 乗統計量の p 値を求めると 0.042 となり，有意水準 5%で棄却される。しかし，行方向での部分分割表を調べると p 値が (1,2) に対して 0.209，(1,3) に対して 0.055，(2,3) に対して 0.097 となり，どれも棄却されない。（ちなみに，列方向でもどれも棄却されない。）

提案した方法がコンソナンスを満たさない最大の理由は，部分分割表での検出力の低下である。これは，データの総数が減ることによる。しかし，この例の場合では元々三すくみのような状態になっているせいもあるので，検出できないのはある程度は仕方ないのかもしれない。

8 おわりに

本論文では，名義尺度の分割表における多重比較法を提案し，それなりの効果があることを確認した。提案した方法は簡便であり，統計ソフトや表計算ソフトですぐに計算でき

る。（分割表が大きい場合は計算の繰り返しの労力が増大するのでマクロを組むなど何らかの工夫は必要である。）独立性の検定の後にその中身について知りたいと考える人は多いので，より詳細な分析が簡単にできるのだということを広めていきたい。

一方，2 つの問題点が残っている。その一つは，2×2 までのより詳細な分析をする場合に有意水準がどの程度まで保てるのかの解明であり，もう一つは分割表の元々の仮説構造に近い周辺合計を固定した多重比較法の構築である。今後はそれらに取り組んでいきたいと思っている。

参考文献

Hirotsu, C. (1983): Defining the pattern of association in two-way contingency tables, *Biometrika*, **70**, 3, 579-589.

Hocheberg, Y, Tamhane, A. C. (1987): *Multiple Comparison Procedures*, John Wiley & Sons, New York.

Marcus, R., Peritz, E., Gabriel, K. R. (1976): On closed testing procedures with special reference to ordered analysis of variance, *Biometrika*, **63**, 3, 655-660.

永田靖・吉田道弘 (1997): “統計的多重比較法の基礎”, サイエンティスト社.