

●講座/臨床研究のための統計学

III. クロス表の検定*

高 木 廣 文**

1. 一般のクロス表の検定

例題 3・1

第1表は、2種の薬剤A、Bを慢性前立腺患者に無作為に投与し、その効果を調べたものである。これは、連載1回目の例題1.4（順位と検定）と同一の例であるが、ここではカイ2乗検定により、薬効の差を検定しよう。

第1表 A薬とB薬の症状改善について

	著効	有効	不変	悪化	計
A薬	8	10	6	1	25
B薬	3	9	10	4	26
計	11	19	16	5	51

【検定法 3・1】

まず、(3・3)式を用いて、カイ2乗値 χ_0^2 を計算する。

$$\begin{aligned}\chi_0^2 &= 51 \times \left(\frac{8^2}{25 \times 11} + \frac{10^2}{25 \times 19} + \frac{6^2}{25 \times 16} \right. \\ &\quad + \frac{1^2}{25 \times 5} + \frac{3^2}{26 \times 11} + \frac{9^2}{26 \times 19} + \frac{10^2}{26 \times 16} \\ &\quad \left. + \frac{4^2}{26 \times 5} - 1 \right) \div 5.108\end{aligned}$$

となる。 χ_0^2 の自由度は(3・4)式より、

$$(2-1) \times (4-1) = 3$$

となる。自由度3のカイ2乗分布の上側5%点の値は数値表から7.815であることがわかる。 χ_0^2 の値はこれより小さいので、有意ではない。したがって、A薬とB薬の効果に差があるとはいえない

第2表 変数AとBのクロス表

A \ B	B ₁	B ₂ ...	B _j ...	B _l	計
A ₁	n ₁₁	n ₁₂ ...	n _{1j} ...	n _{1l}	n _{1.}
A ₂	n ₂₁	n ₂₂ ...	n _{2j} ...	n _{2l}	n _{2.}
⋮	⋮	⋮	⋮	⋮	⋮
A _i	n _{i1}	n _{i2} ...	n _{ij} ...	n _{il}	n _{i.}
⋮	⋮	⋮	⋮	⋮	⋮
A _k	n _{k1}	n _{k2} ...	n _{kj} ...	n _{kl}	n _{k.}
計	n _{.1}	n _{.2} ...	n _{.j} ...	n _{.l}	N

$$N = n_{1.} + n_{2.} + \dots + n_{k.} = n_{.1} + n_{.2} + \dots + n_{.l}$$

ことになる。

【解説 3・1】

クロス表の検定は、2つに大別して考えることができる。すなわち、クロス表の2つの変数の一方がグループを示すもので、他方の変数の分布にグループ間で差があるか否かを検定する場合は、「一様性の検定」と呼ばれる。また、2変数間の関連の有無を検定する場合は、「独立性の検定」と呼ばれる。しかし、これらの検定を行うための実際の計算手順は、まったく同一であり、以下に説明する方法により検定することができる。

第2表を用いて、カイ2乗検定の方法を説明しよう。2変数A、Bに関してクロス表が作成されているものとし、変数Aは表に示したように、k個のカテゴリからなり、同様にBはl個のカテゴリからなるものとする。したがって、クロス表は(k×l)個のマス目をもつが、各マス目は「セ

* Statistics in Clinical Studies: Ⅲ. Methods of the Test for the Cross-table

** 聖路加看護大学 Takagi Hirofumi: St. Luke's College of Nursing

第3表 カテゴリの併合

	著効	有効	その他	計
A薬	8	10	7	25
B薬	3	9	14	26
計	11	19	21	51

ル (cell)」と呼ばれる。

表に示したように、A変数についてカテゴリ A_i という特性をもち、かつBについて B_j という特性をもつものの度数 (人数、個数など) を n_{ij} のように表すことにする。一般的にいうと、Aについて A_i で、かつBについて B_j というセルの度数を n_{ij} と表すことにする。また、カテゴリ A_i の度数の合計を $n_{i\cdot}$ とし、カテゴリ B_j の度数の合計を $n_{\cdot j}$ とする。すなわち、

$$n_{i\cdot} = n_{i1} + n_{i2} + \cdots + n_{ij} = \sum_{j=1}^l n_{ij} \quad (3 \cdot 1)$$

$$n_{\cdot j} = n_{1j} + n_{2j} + \cdots + n_{kj} = \sum_{i=1}^k n_{ij} \quad (3 \cdot 2)$$

である。また、全体の標本数を N とする。

2変数A, Bの関連の有無を検定するためには、統計量として、

$$\chi_0^2 = N \left(\frac{n_{11}^2}{n_{1\cdot} n_{\cdot 1}} + \frac{n_{12}^2}{n_{1\cdot} n_{\cdot 2}} + \cdots + \frac{n_{kl}^2}{n_{k\cdot} n_{\cdot l}} - 1 \right) = N \left(\sum_{j=1}^l \frac{1}{n_{i\cdot} n_{\cdot j}} \sum_{i=1}^k n_{ij}^2 - 1 \right) \quad (3 \cdot 3)$$

を求める。 χ_0^2 は、カイ2乗分布に従い、その自由度 ν は、

$$\nu = (k-1) \cdot (l-1) \quad (3 \cdot 4)$$

となる。

χ_0^2 の値が、自由度 ν のカイ2乗分布の上側5% (もしくは1%) 点の値 $\chi_{0.05}^2(\nu)$ より大きければ、変数間に有意な関連があるものとし、小さければ関連が認められないものとして、検定を行えばよい。

ところで、変数のカテゴリ数が大きくなるにつれ、セルの個数も急激に増加し、標本数があまり大きくない場合、特定のセルの度数が極めて小さくなることもある。そのような場合、 χ_0^2 のカイ

2乗分布へのあてはまりが悪くなることが知られている。正確にいうと、セルの度数ではなく、2変数が独立であると仮定した場合の各セルの期待値が問題となる。変数Aのカテゴリ A_i かつBのカテゴリ B_j に対応するセルの期待値 e_{ij} は、

$$e_{ij} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{N} \quad (3 \cdot 5)$$

となる。通常、期待値が5以下の場合、カイ2乗分布による近似はよくないといわれている。第1表について、「A薬」で「悪化」のセルの期待値は、上式を用いて計算すると、

$$\text{期待値} = \frac{25 \times 5}{51} \div 2.41$$

となり、極めて小さ過ぎる。このような場合、可能ならば隣合うカテゴリ同士を併合して、検定を行うとよい。

第3表は、「不変」と「悪化」を併合し、「その他」としてまとめたものである。この表から χ_0^2 の値を求めると、4.641となる。また、自由度は $(3-1) \times (2-1) = 2$ である。数値表から自由度2のカイ2乗分布の上側5%点は5.991であることがわかる。併合した場合も、前述の場合と同様に有意ではない。この点から、カテゴリの併合は無意味であるかのように思えるかもしれないが、有意確率を計算すると、併合前は0.164、併合後は0.098となり、5%より大きいものの確率的に小さくなっていることがわかる。ただし、カテゴリの併合により、常に有意確率が小さくなる訳ではなく、重要なことはカイ2乗分布への適合がよくなるという点に注目すべきである。

連載1回目に、Wilcoxonの順位和検定を用いて第1表について検定したところ、2群間に有意差が認められた。この点からいえば、2群間の差に関して、変数のカテゴリ間に順位を仮定できる場合には、順位和検定を用いて、2群での順位の分布の差 (一定方向へのずれ) を検定すればよいだろう。カイ2乗検定は群間での一定方向への分布のずれについてではなく、特定カテゴリの分布の差 (頻度の差) を検定したい場合に用いればよい。

2. 4分表のカイ2乗検定

例題 3・2

慢性前立腺患者に、無作為にA薬とプラセボを各15例ずつ投与し、自覚症状について効果を調べたところ、第4表に示したような結果が得られた。このデータから、A薬に効果があるといえるだろうか。

第4表 A薬とプラセボの投与結果

	自覚症状		計
	改善	不変	
A薬	11	4	15
プラセボ	6	9	15
計	17	13	30

【検定法 3・2】

2変数のどちらも2カテゴリからなるクロス表の場合、検定に前述した計算式を用いるのではなく、より簡単な後述の計算式(3・6)を用いると、

$$\chi_0^2 = \frac{30 \times (|11 \times 9 - 4 \times 6| - 30/2)^2}{15 \times 15 \times 17 \times 13}$$

$$\div 2.172$$

となる。 χ_0^2 は自由度1のカイ2乗分布に従う。したがって、自由度1のカイ2乗分布の上側5%点の値3.84と比較すると、 χ_0^2 の値は小さいので有意とはいえない(有意確率を求めると14.1%になる)。

上記の検定結果から、A薬による自覚症状改善の程度が、プラセボより効果があるとはいえないことになる。

【解説 3・2】

2変数がそれぞれ2カテゴリからなる場合、クロス表は「4分表」ともよばれ、一般のクロス表に比べ検定のための計算手順は簡単である。すなわち、第5表のような4分表がある場合、

$$\chi_0^2 = \frac{N(|ad - bc| - N/2)^2}{n_1 n_2 m_1 m_2} \quad (3 \cdot 6)$$

を求める。ここで、 n_1 , n_2 , m_1 , m_2 は表に示したように各行と各列の合計であり、Nは全標本数である。

第5表 4分表の検定

A \ B	B ₁	B ₂	計
A ₁	a	b	n ₁
A ₂	c	d	n ₂
計	m ₁	m ₂	N

$$n_1 = a + b, n_2 = c + d, m_1 = a + c, m_2 = b + d$$

$$N = a + b + c + d$$

上記の χ_0^2 は自由度1のカイ2乗分布に従うので、上側5%点の値3.84、もしくは1%点の値6.63と比べ、 χ_0^2 の方が大きい場合に有意とすればよい。なお、(3・6)式の分子中の $-N/2$ は、「連続性の補正(イエーツの補正)」と呼ばれるもので、人数のように整数値しかとらない(離散的)データから得た統計量が、連続的なカイ2乗分布に対して、あてはまりがより良くなるようにするためのものである。ただし、連続性の補正は、

$$|ad - ab| > N/2 \quad (3 \cdot 7)$$

の場合に行い、そうでない場合には、 $\chi_0^2 = 0$ とする。

3. フィッシャーの直接確率

例題 3・3

第4表で示したA薬とプラセボの自覚症状に関するクロス表から、A薬の効果の有無をカイ2乗検定ではなく、フィッシャーの直接確率を求めて検定する。

【検定法 3・3】

まず、「計」の欄の数値が一定の場合に、第4表のようなクロス表が偶然できる確率を求める。(3・8)式より、

$$p_0 = \frac{15! \cdot 15! \cdot 17! \cdot 13!}{30! \cdot 11! \cdot 4! \cdot 6! \cdot 9!} \div 0.057046$$

となる。

検定のためには、第4表の「計」の欄の数値を一定のまま、対角成分の積の大きいセルの数値を1つつ多くし、逆に積の小さいセルの数値は1つつ減らしていき、各クロス表の確率をすべて求める必要がある。第6表が最も偏りが大きな状態であるが、第4表からこの状態までには、表には示していないがA薬で「不変」のセルが3, 2, 1となる3つのクロス表がある。これら、残り4つ

第6表 最も偏りの大きい場合

	自覚症状		計
	改善	不変	
A 薬	15	0	15
プラセボ	2	13	15
計	17	13	30

のクロス表の確率を(3・9)式を用いて計算する。

A薬で不変のセルの度数が3の場合、

$$p_1 = \frac{4 \times 6}{12 \times 10} \times p_0 \div 0.011409$$

となり、度数が2の場合、

$$p_2 = \frac{4 \times 5}{13 \times 11} \times p_1 \div 0.001197$$

となる。また、度数が1の場合と0の場合は、

$$p_3 = \frac{2 \times 4}{14 \times 12} p_2 \div 0.00005699$$

および、

$$p_4 = \frac{1 \times 3}{15 \times 13} \times p_3 \div 0.000000088$$

となる。これらの確率の和Pを求めると、

$$p = p_0 + p_1 + p_2 + p_3 + p_4 \div 0.0697$$

となる。これが、「フィッシャーの直接確率」とよばれるものである。ただし、上記の確率は一方方向の偏りのみを考えて計算されており、カイ2乗検定と対応させるためには、この値を2倍して用いばいだろう。

$$2p \div 0.139$$

であり、有意確率は13.9%となり、5%水準に達しないことがわかる。したがって、A薬の効果があるとはいえないことになる。なお、すでに述べたカイ2乗検定で求めた χ^2_0 値の有意確率は14.1%であり、フィッシャーの直接確率と大差ないことがわかる。

【解説 3・3】

フィッシャーの直接確率は、すでに例題で示したように、検定のための有意確率を直接求める方法である。t検定やカイ2乗検定などではデータからある統計量を計算し、特定の分布から5%点や1%点の値を数値表を用いて求め、計算された

第7表 偏らせた4分表

A \ B	B ₁	B ₂	計
A ₁	a + i	b - i	n ₁
A ₂	c - i	d + i	n ₂
計	m ₁	m ₂	N

統計量との大小比較で検定を行うのが通常である。この点からいえば、検定のための有意確率が、直接計算できるという利点はあるが、そのための計算がフィッシャーの直接確率ではかなり面倒であることも事実である。

すでに4分表のカイ2乗検定の説明で用いた第5表により、実際に直接確率を求める計算方法を説明する。

全標本数をNとし、変数AのカテゴリA₁、A₂の度数は、表に示したようにそれぞれn₁、n₂であるとする。同様に、変数BのカテゴリB₁、B₂の度数はm₁、m₂とする。直接確率は、このような場合に偶然に4つのセルの度数がa、b、c、dとなる確率、およびそれ以上に特定方向に偏つた状態の確率を求めるものである。

計の欄の数値を固定した場合、変数AのカテゴリA₁かつ変数BのカテゴリB₁のセルの度数がaとなる場合を考えよう。

全標本数Nのうち、変数AについてカテゴリA₁の度数の合計がn₁となる場合の数は、

$${}_N C_{n_1} = \frac{N!}{n_1! \cdot (N - n_1)!} = \frac{N!}{n_1! \cdot n_2!}$$

通りある。また、n₁のうち変数BについてaがB₁に、bがB₂となる組み合わせの数は、それぞれ、

$${}_{m_1} C_a = \frac{m_1!}{a! \cdot (m_1 - a)!} = \frac{m_1!}{a! \cdot c!}$$

通りと、

$${}_{m_2} C_b = \frac{m_2!}{b! \cdot (m_2 - b)!} = \frac{m_2!}{b! \cdot d!}$$

通りある。したがって、第5表が偶然でできる確率P₀は、

$$P_0 = \frac{{}_{m_1} C_a \cdot {}_{m_2} C_b}{{}_N C_{n_1}} = \frac{n_1! \cdot n_2! \cdot m_1! \cdot m_2!}{N! \cdot a! \cdot b! \cdot c! \cdot d!}$$

(3・8)

により求められる。

検定のためには、第5表の状態がさらに偏つた4分表も考慮する必要がある。

いま、 $ad > bc$ としよう。偏つた状態とは、第7表に示したように、積の大きい方はより大きくなるように、また小さい方はより小さくなるように各セルに1ずつ度数を増減することである。第7表で、 $i = 0$ とすれば第5表と等しく、 $i = 1$ とすれば1だけずらした表となる。 d と c のうち小さい方の値を k とすると、各セルの度数は0以上なので、 i のとり得る値は、 $0, 1, \dots, k$ となる。それらのすべての4分表の確率を計算する必要がある。しかし、(3・8) 式を各表について計算する必要はない。

第7表で $i = 1$ とした場合の確率 p_1 は、 p_0 を用いて、

$$p_1 = \frac{b \cdot c}{(a+1)(d+1)} \cdot p_0$$

となり、同様に $i = 2$ の場合、 p_1 を用いて、

$$p_2 = \frac{(b-1)(c+1)}{(a+2)(d+2)} \cdot p_1$$

により求められる。

一般に、確率 p_i が求めれば、 $(i+1)$ 番目の表の確率は、

$$p_{i+1} = \frac{(b-i)(c-i)}{(a+i+1)(d+i+1)} \cdot p_i \quad (3 \cdot 9)$$

により求められる。

フィッシャーの直接確率 P は、これらの確率をすべて足し合わせたものである。すなわち、

$p = p_0 + p_1 + \dots + p_k$ ただし $k = \text{Min}(b, c)$ である。

上記の手順により求められた確率 P は、特定方向のみの偏りを問題としている。したがって、片側検定でよいという確固とした理由がある場合以外は、そのままの値を用いるのではなく、数値を2倍した値を検定のための有意確率として用いられよう。

ところで、4分表の検定としてフィッシャーの直接確率を用いるか、カイ2乗検定を用いるかという問題があるが、できる限りフィッシャーの直接確率を求めるべきであろう。しかし、直接確率の計算には階乗計算が含まれるので、標本数が大

きくなると若干面倒くさくなるし、計算手順をうまくしないとせつかく求めた確率が正確でなかったりすることもある。パソコンが使える場合には、適当な統計パッケージ（たとえば、現代数学社発行の HALBAU など）を利用すればよいだろう。ただし、すでに例題で示したように (3・6) 式によるカイ2乗検定の結果と、フィッシャーの直接確率によるものに大差はなかった。したがって、標本数がある程度大きければ、どちらを用いて検定しても結果に影響はないだろう。

最後に、これまで説明した方法を用いて、いくつかの例題を解いてみよう。

例題 3・4

第8表は、女性の腹圧性尿失禁患者にA薬を投与し、3つの年齢群別にその効果を調べたものである。このデータから、年齢による薬効差があるといえるだろうか。

第8表 年齢別によるA薬の効果

年 齢	改善	不変	計
20歳未満	13	11	24
20～49歳	21	15	36
50歳以上	43	21	64
計	77	47	124

【検定法 3・4】

すでに示したように (3・3) 式を用いて、

$$\begin{aligned} \chi_0^2 &= 124 \times \left(\frac{13^2}{24 \times 77} + \frac{11^2}{24 \times 47} + \frac{21^2}{36 \times 77} \right. \\ &\quad \left. + \frac{15^2}{36 \times 47} + \frac{43^2}{64 \times 77} + \frac{21^2}{64 \times 47} - 1 \right) \\ &\doteq 1.563 \end{aligned}$$

となる。自由度は $(3-1) \times (2-1) = 2$ である。

自由度2のカイ2乗分布の上側5%点の値は5.991であり、 χ_0^2 の値はこれより小さいので有意ではない。したがって、A薬の効果に年齢差があるとはいえないことになる。

例題 3・5

第8表に示した例では、50歳以上の患者群の標本数が他の群に比べ多くなっている。そこで、第9表に示したように49歳以下の患者を1つのグループにまとめて、2群間の改善の割合に有意差があるか検定してみよう。

第9表 2群間のA薬の効果の差

年齢群	改善	不変	計
49歳以下	34	26	60
50歳以上	43	21	64
計	77	47	124

【検定法 3・5】

49歳以下での「改善」の割合は34/60で、これは約56.7%であり、50歳以上では同様に約67.2%となり、2群間に10%以上の差がみられる。このように2群での比率の差を検定する場合、「2つの母比率の差の検定」と呼ばれる方法を用いることもあるが、それは(3・6)式による4分表のカイ2乗検定と完全に一致する。ここでは、(3・6)式を用いて検定する。

$$\chi_0^2 = \frac{124 \times (|34 \times 21 - 26 \times 43| - 124/2)^2}{60 \times 64 \times 77 \times 47}$$

$$\div 1.044$$

となる。この値は、自由度1のカイ2乗分布の上側5%点の値3.84より小さいので有意とはいえない。なお、 χ_0^2 の値の有意確率を計算すると約30.7%である。したがって、A薬の効果に年齢差があるとはいえないことになる。

例題 3・6

第9表について、フィッシャーの直接確率を求めて、A薬に年齢差があるか否かを検定する。

【検定法 3・6】

第9表の標本数は、フィッシャーの直接確率を求めるにはやや大きすぎるが、カイ2乗検定との

第10表 直接確率の計算

度数*	確率	度数	確率
21	7.1620×10^{-2}	10	8.8307×10^{-8}
20	4.3044×10^{-2}	9	9.7180×10^{-9}
19	2.2547×10^{-2}	8	8.8103×10^{-10}
18	1.0276×10^{-2}	7	6.4918×10^{-11}
17	4.0668×10^{-3}	6	3.8219×10^{-12}
16	1.3939×10^{-3}	5	1.7583×10^{-13}
15	4.1247×10^{-4}	4	6.1335×10^{-15}
14	1.0499×10^{-4}	3	1.5539×10^{-16}
13	2.2888×10^{-5}	2	2.6734×10^{-18}
12	4.2506×10^{-6}	1	2.7675×10^{-20}
11	6.6833×10^{-7}	0	1.2881×10^{-22}

* 度数は、「50歳以上」かつ「不変」のセルの度数を示す。

比較という点で、実際の計算を示してみたい。

まず、(3・8)式を用いて第9表が偶然できる確率 p_0 を求める。

$$p_0 = \frac{60! \cdot 64! \cdot 77! \cdot 47!}{124! \cdot 34! \cdot 26! \cdot 43! \cdot 21!} \div 0.071620$$

である。この p_0 を用いて、「50歳以上」かつ「不変」のセルを21から0まで変化させた場合の各確率を(3・9)式を利用して計算すればよい。セルの度数が20の場合の確率は、

$$p_1 = \frac{34 \times 21}{(43+1) \times (26+1)} \times p_0 \div 0.043044$$

となる。以後、順に確率を求めていけばよいが、それらの確率を第10表にまとめて示した。

すべての確率を加えると、フィッシャーの直接確率 P は、

$$P = p_0 + p_1 + p_2 + \dots + p_{21} \div 0.15349$$

となる。この値を2倍すると約0.307、すなわち30.7%となる。これは、先の例題で示したように、カイ2乗検定により求めた χ_0^2 値の有意確率と一致している。

上記のことから明らかなように、ある程度標本数が大きければ、苦勞してフィッシャーの直接確率を求める必要はなく、カイ2乗検定を行えばよい。