

〔論説〕

統計学入門書にみられる Fisher の直接確率法の両側確率と片側確率をめぐる混乱

竹森 幸一¹⁾

Confused understanding in statistical handbooks on the two-sided and one-side probability for the Fisher's exact probability test

Koichi Takemori¹⁾

(J. Aomori Univ. Health Welf. 7(2): 187-190, 2006)

キーワード：Fisher の直接確率法、正確両側有意確率、カイ二乗検定

Key words: the Fisher's exact probability test, the exact two-tailed significance level, chi-square test

1. はじめに

カテゴリーデータの分析において、2 × 2 表は判定結果が期待度数に対して各セルの観察度数がどちらに片寄っているかを明確に示すことができる点で、よく用いられる。4つのセルの期待度数のいずれかが5未満の場合は、カイ二乗検定を用いると誤った判断を下すおそれがある。この場合は Fisher の直接確率法によって独立性の検定が行われる。

さきに、「Fisher の直接確率法をめぐる混乱」と題して、この検定の帰無仮説、対立仮説、両側確率、片側確率をめぐる、市販の解説書が混乱していることを報告した¹⁾。その後ほぼ10年経過したので、1997年から2006年に発行された市販の和書（訳本を含む）を中心に整理してみた。

2. 両側確率、片側確率をめぐるの混乱

市販の解説書について両側確率、片側確率の計算方法に注目すると次の4群に分けられる。

1群) 両側確率は観察パターンの確率とそれより片寄りが大きくなる方向へずらしたパターンの確率の和（片側確率）に、反対側のパターンの中で観察パターンの確率以下の確率を加えたもの、すなわち全てのパターンの中で、観察パターンの確率以下の確率の総和とするというもの²⁻⁴⁾。

なお、観察パターンとは検定すべき 2 × 2 表のことである。片寄りが大きくなる方向とは4つのセルのどれかに注目し、帰無仮説のもとにおけるそのセルの期待度数より観察度数が小さい場合は減らす方向、逆に期待度数より観察度数が大きい場合は増やす方向のことである。反対側のパターンとは片寄りが大きくなる方向の逆方向に度数を変化させた場合のパターンのことである。

2 × 2 表の観察度数を a, b, c, d とすると、帰無仮説「行と列は互いに独立」のもとで、観察された観察度数を得る確率 (p) は

$$p = \frac{(a+b)! (c+d)! (a+c)! (b+d)!}{(a+b+c+d)! a! b! c! d!} \dots\dots (1)$$

と表される。

検定方式は一般の仮説検定と同様に、片側検定と両側検定がある。片側検定は帰無仮説のもとにおける期待度数からみて観察パターンの方向へさらに片寄ったパターンの起こる確率を加えた確率（片側確率）が有意水準より小さい場合、帰無仮説を棄却するものである。両側検定はデータが取り得る全てのパターンの確率の中で、観察パターンの起こる確率と同等あるいはそれより小さい確率を全て合計した確率（両側確率）を有意水準と比較して判定するものである。

1) 青森県立保健大学健康科学部看護学科

Department of Nursing, Faculty of Health Sciences, Aomori University of Health and Welfare

例えば、観察パターンが表1のような度数からなる場合、データが取り得る全てのパターンと(1)式で計算したそれぞれの確率は表2のようになる(行と列の和は全て表1の観察パターンと同じ)。

表1 2×2表の例

	a	b	合計
A	1 (4.3)	8 (4.7)	9
B	10 (6.7)	4 (7.3)	14
合計	11	12	23

()内は期待度数(以下同様)

表2 データが取り得る全てのパターンとその確率

パターン	0	9	1	8	2	7	3	6	4	5
確率	0.00027	0.00666	0.05330	0.18657	0.31983					
パターン	5	4	6	3	7	2	8	1	9	0
確率	0.27985	0.12438	0.02665	0.00242	0.00007					

片側確率=0.00666+0.00027=0.00693
両側確率=0.00666+0.00027+0.00242+0.00007=0.00942

「観察パターンより片寄りが大きくなる方向へずらしたパターン」とは4つのセルのどれかに注目し、そのセルの観察度数が期待度数より小さい場合はそのセルの度数を小さくする方向へ、反対にそのセルの観察度数が期待度数より大きい場合はそのセルの度数を大きくする方向へずらしたパターンということである。文献4では「片寄りが多くなる方向へ」と「最も小さい度数を0の方向へ」と同等に扱っているがこれは正しくない。例えば4(3)、6(7)、5(6)、15(14)(カッコ内は期待度数)の場合、最も小さい度数は4であるが、期待度数3より大きいので、0の方向ではなく大きくなる方向へずらさなければいけない。

文献2では、観察データが落ちている方の分布の裾で確率を求め、両側検定にするためにはこの値を2倍する方法も紹介しており、ほとんどの場合はこの方法で求めた値は上述した値と差は顕著でないが、たまには顕著になることがあると述べている。

2群)片側検定(確率)、両側検定(確率)という表現がある場合とない場合があるが、両側検定では(両側確率は)片側確率を2倍するというもの⁵⁻⁹⁾。

普通、両側には相異なる確率があり、例外的に2行か2列の合計が等しい場合、両側確率は単純に片側確率を

2倍にすると1群の方法で求めた両側確率と同じになる。2群の解説書では両側のp値は単に片側p値の2倍であると述べているが、行か列が等しいときを除いて2倍にならない。

表3は2行の合計が等しくともに5である。この場合の全ての取りうるパターンとその確率は下記のとおりで、左右対称となり、両側確率は片側確率の2倍になる。

パターン	4	1	3	2	2	3	1	4	0	5
確率	0.0238	0.2381	0.4762	0.2381	0.0238					

片側確率=0.0238+0.2381=0.2619
両側確率=0.0238+0.2381+0.2381+0.0238=0.5238

表3 2行の合計が等しくなる2×2表の例

	a	b	合計
A	3 (2)	2 (3)	5
B	1 (2)	4 (3)	5
合計	4	6	10

表4 2列の合計が等しくなる2×2表の例

	a	b	合計
A	3 (3.5)	4 (3.5)	7
B	2 (1.5)	1 (1.5)	3
合計	5	5	10

表4は2列の合計が等しくともに5である。この場合の全ての取りうるパターンとその確率は下記のとおりで、左右対称となり、両側確率は片側確率の2倍になる。

パターン	2	5	3	4	4	3	5	2
確率	0.0833	0.4167	0.4167	0.0833				

片側確率=0.0833+0.4167=0.5
両側確率=0.0833+0.4167+0.4167+0.0833=1.0

表5 2行、2列の合計とも等しくならない2×2表の例

	a	b	合計
A	1 (1.3)	5 (4.7)	6
B	2 (1.7)	6 (6.3)	8
合計	3	11	14

表5は2行も2列も合計が等しくならない場合で、こ

の場合の全ての取りうるパターンとその確率は下記のとおりで、左右対称でなく、両側確率は片側確率の2倍とにならない。2倍すると両側確率は1を越える。

パターン	0 6	1 5	2 4	3 3
	3 5	2 6	1 7	0 8
確率	0.1538	0.4615	0.3297	0.0549
片側確率=0.1538+0.4615=0.6153				
両側確率=0.1538+0.4615+0.3297+0.0549=1.0				

文献5の例題は2行の和が50、50と等しく両側確率は片側確率の2倍になる。文献6の例題は2列の和が9、9と等しく両側確率は片側確率の2倍になる。文献7の例題は2行、2列の合計とも等しくないので、両側確率は片側確率の2倍にならない。文献8の例題も2行、2列の合計とも等しくなく、期待度数からみて観察パターンの方向へさらに片寄ったパターンの起こる確率を加えた確率を片側確率とし、両側確率はそれを2倍して求めているが、観察パターンより小さい確率は反対側にはない。この例では片側確率と両側確率は等しくなる。文献9は「標本人数分布が理論人数分布よりかけ離れる表の全てについて足し算する。両側検定では2倍する」と述べている。SPSSの出力が掲載されていて、正確有意確率(両側)が0.348、正確有意確率(片側)が0.199となっている。0.199の2倍は0.398となる。

なおSPSSのSignificance Levels for Fisher's Exact Testの両側確率のアルゴリズム¹⁰⁾はThe exact two-tailed significance level p_2 is defined as the sum of the one-tailed significance level p_1 and the probabilities of all points in the other side of the sample space of N_i which are not greater than the probability of $N_i=n_i$ としてあり、1群の方法と同様である。

3群)片側検定(確率)という表現がある場合とない場合があるが、片側検定(確率)についてのみ解説しているもの¹¹⁻¹³⁾。

文献11の例題では「抗生物質の効果があるかどうかをみたいので片側検定となる」と述べて、片側検定をしている。文献12は片側検定という表現はないが、片側検定をしている。表6が例題として載せてあるが、各パターンの起こる確率の計算は間違っているようである。4パターンの確率を計算し、 $p=0.161+0.052+0.012+0.002=0.227$ としているが、 $p=0.0791+0.0256+0.0049+0.0004=0.1100$ となるはずである。

表6 文献12の例題

	a	b	合計
A	197	192	389
B	3	8	11
合計	200	200	400

文献13は表7に示した例題について、片側検定をするとして、「最も小さい例数4が0になるまで1ずつ減じていき」、5組のパターンの確率を計算し、片側確率(0.896)を計算している。しかし観察度数4は期待度数3より大きいので、より片寄った方向は5、6と増やす方向である。そうすると、観察パターンを入れて6つのパターンができ、その確率はそれぞれ0.2276、0.0853、0.0167、0.0016、0.0001、0.0000となり片側確率はこれらの合計で0.3313となる。

表7 文献13の例題

	a	b	合計
A	4 (3)	6 (7)	10
B	5 (6)	15 (14)	20
合計	9	21	30

4群)両側検定(確率)、片側検定(確率)という表現はないが、両側確率を求めているもの¹⁴⁾

文献14はデータがとりうる全てのパターンの中で、観察パターンの確率以下の確率の和を求めている、1群の両側確率の求め方と同様である。

なお、文献8は例題の両側確率の計算方法、文献9は両側確率についての考え方、文献12は計算結果の確認、文献13は、より片寄った方向の確認についてそれぞれ出版社を通じて著者に確認したところ、文献8は訳本であるが訳者及び原著者から例題についての見解は得られていない。文献9は片側確率を2倍して求めた両側確率がSPSSの出力と一致しない点を改訂するとの回答、文献12は訳本であるが出版社から訳者さらに原著者に確認した結果、本論文で指摘したように改訂する予定との回答があり、出版社のホームページに正誤表として2006年11月1日に掲載された。文献13は本論文で指摘したとおりで、改訂に向けて検討するとの回答があった。

3. まとめ

以上述べたように、市販の入門書にみられる Fisher の直接確率法をめぐる混乱は未だ収まっていない。また間違った記述や例題解答が多くみられる。引用した入門書は、統計学の初心者がみる機会が多いわけで、初心者は書かれている内容を信用して学習することになる。

2×2表の場合、「行と列が独立である」との帰無仮説のもとに、解決すべき問題の対立仮説が片側検定か両側検定かを決定して、どちらかの正確有意確率を計算することになる。従って入門書には片側検定と両側検定の両方の確率計算方法を示しておく必要がある。片側確率は全ての引用文献とも同じ考え方で求めているが、観察パターンより片寄った方向を正しく判断する方法を明示する必要がある。「一番小さい数値を0の方向へ変化させる」という表現では間違えることがあるので、観察度数と期待度数を比較し、正しい方向へ変化させる必要がある。

両側確率の求め方は、1群と2群の2つに大別される。2群の「いつでも2倍する」方法は観察パターンより小さい確率のパターンが一方にだけにあり、反対側は全て観察パターンより大きい確率の場合は2倍する方法では観察パターンより大きい確率も加算することになり、検定の意味からおかしい。また片側確率が0.5を越える場合、2倍して両側確率を求めると確率が1を越えるという結果となる。

以上述べたように、Fisher の直接確率法について市販の統計学入門書が混乱しているように見受けられる。早い時期に、これらの混乱が収まることを期待したい。

文献

- 1) 竹森幸一, 三上聖治: Fisher の直接確率法をめぐる混乱. 医学のあゆみ, 181 (8), 558-560, 1997.
- 2) Douglas G Altman (大船義久、佐久間 昭訳): 医学研究における実用統計学. 209-211, サイエンス・エッセイ社, 1999.
- 3) 新城明久: ノンパラメトリック法. 11-13, 川島書店, 2004.
- 4) 縣 俊彦: 基本医学統計学. 第4版, 50-52, 中外医学社, 2004.
- 5) 宮原英夫, 丹後俊郎編: 医学統計学ハンドブック. 初版第4刷, 173-179, 朝倉書店, 2002.
- 6) 三野大来: 看護師のための統計学. 151-153, 共立出版, 2005.
- 7) 佐藤敏雄, 村松 幸: やさしい医療系の統計学. 第2版, 129-130, 医歯薬出版, 2002.
- 8) TDV Swinscow, MJ Campbell (折笠秀樹監訳): はじめて学ぶ医療統計学. 109-115, 総合医学社, 2003.
- 9) 大櫛陽一他: SPSSによる看護・福祉・医学統計学入門 [改訂版]. 94-99, 福村出版, 2004.
- 10) SPSS Inc.: Appendix 5. Significance levels for Fisher's Exact Test. SPSS 13.0 for Windows and SmartViewer, SPSS Inc., 2004.
- 11) 正井栄一: 医学・保健学のためのやさしい統計学. 131-132, 金原出版, 2004.
- 12) ジェフリー R ノーマン, デビッド L ストレイナー (中野正孝他訳): 論文が読める! 早わかり統計学. 第2版, 121-124, メディカル・サイエンス・インターナショナル, 2005.
- 13) 杉田暉道, 栃久保 修: 統計学入門. 第7版, 94-96, 医学書院, 2001.
- 14) DE Matthews, VT Farewell (宮原英夫, 折笠秀樹監訳): 実践医学統計学. 17-22, 朝倉書店, 2005.