短　報

聖マリアンナ医科大学雑誌
Vol. 37, pp. 45–49, 2009

2×2 分割表における片側検定と両側検定について

上野　隆彦　　立浪　忍

（受付：平成 21 年 2 月 12 日）

抄　録

2×2 分割表における独立性の検定，あるいは比率の検定は，医学分野でしばしば使われる統計手法である。しかしながらこの独立性の検定については Pearson により提案された χ^2 検定と Fisher により提案された exact test の 2 つの流儀がある。Pearson の方法は両側検定であるが，Fisher の方法は片側検定，両側検定の双方の考え方があり，さらに，その両側検定には 2 通りの方法が提案されている。今般，複数の統計プログラムを調べたところ，いずれにおいても χ^2 検定と，Fisher’s exact test による双方の確率が計算されていたが，Fisher’s exact test に関しては片側確率のみを計算するものと，片側確率および両側確率の両方を計算するものが混在していた。ここで計算される両側確率の値は Finney の提案に基づくものである。Fisher’s exact test における両側確率は Finney の方法から得られる確率と，Fisher の提案による片側確率の 2 倍の値とがある。この 2 つの方法を比較したところ，多くの場合は片側確率の 2 倍の方が，Finney の方法による確率よりも大きな値であった。例数が少ない等の理由で χ^2 検定の代わりに Fisher’s exact test を採用する場合には両側検定を用いるべきであるが，両側確率を算出する 2 つの方法の性質について，検定を行う以前に認識しておくことが必要である。

索引用語

2×2 分割表，片側検定，両側検定，Fisher’s exact test

はじめに

医学的研究の解析方法の一つとして，分割表の検定がしばしば用いられる。その中でも 2×2 分割表は使われる頻度が高い。2×2 分割表の検定方法については，Pearson の χ^2 検定，次に Fisher’s exact test[1)]，Yates の補正をした χ^2 検定[2)]という順に提案されている。この Yates の方法は Fisher’s exact test の片側確率の 2 倍の近似値を計算する方法である。どれが最適な方法かということに関しては，統計学者の間で半世紀以上にわたって議論され続けている問題である。この問題に対する一つの指針は 1950 年代の Cochran の提言[3)]「全ケース数が 20 以下のとき，または全ケース数が 20～40 のときで期待値のうち最も小さいものが 5 以下のときには Fisher’s exact test を，全ケース数が 40 以上の場合には Yates の方法を用いること」である。その後に出版された統計学のテキストの多くは Cochran の意見を推奨しており，生物学や医学分野においても，ほぼそれに従って実際の検定作業が行われている[4)]。

しかし，χ^2 検定は意味としては両側検定であるが棄却域を片側にとっているので，あたかも片側検定であるかのような印象を持つことがある。一方，Fisher の方法はもともと片側検定として提案された手法である。この点を踏まえずに Cochran の提

聖マリアンナ医科大学　医学教育文化部門医学統計学分野

言に従うと，本来，両側検定を計画していた研究データを誤って片側検定してしまう等の誤用や曲解を生じる可能性もあるだろう。

そこで，今般は複数のプログラムにおける2×2分割表の取り扱い方法についての調査を行った。とくに留意した点は Fisher's exact test における片側検定と両側検定の取り扱い方法についてである。

次に，Fisher's exact test における両側検定の方法として提案されている2つの方法[5]，すなわち Finney が Fisher に提案した方法，すなわち実際に観測された分割表より生起確率の低い全ての分割表の生起確率の和を用いる方法と，それに対する Fisher の回答である片側確率の2倍の値を用いる方法を比較し，その適応性について検討した。

対象および方法

調査した汎用プログラムあるいは procedure は，SPSS (version 17.0, SPSS Japan, 東京)，Kyplot (version 4a, KyensLab，東京)，jmp (version 7.0.1, SAS Institute Japan, 東京)，Statcel 2 (4 steps エクセル統計第2版[6]，東京)，および R（http://www.r-project.org/）中の procedure である chisq. test, fisher. test である。

表1の分割表において，(1) 式の χ^2 を統計量とする方法を Pearson の方法，(2) 式の χ_c^2 を統計量とするものを Yates の補正法と分類した。

$$\chi^2=\frac{(a-\hat{a})^2}{\hat{a}}+\frac{(b-\hat{b})^2}{\hat{b}}+\frac{(c-\hat{c})^2}{\hat{c}}+\frac{(d-\hat{d})^2}{\hat{d}}=\frac{n(ad-bc)^2}{efgh} \quad (1)$$

$$\chi_c^2=\begin{cases}\dfrac{n(|ad-bc|-n/2)^2}{efgh}, & |ad-bc|-n/2>0\\ 0, & |ad-bc|-n/2\leq 0\end{cases} \quad (2)$$

ここで a, b, c, d は0以上の整数で，$\hat{a}, \hat{b}, \hat{c}, \hat{d}$ は観測値から得られる期待値である。

一方，Fisher's exact test は，周辺度数 e, f, g, h を定めたときに**表1**の分割表が得られる確率が超幾何分布，

$$\frac{e!\times f!\times g!\times h!}{n!\times a!\times b!\times c!\times d!} \quad (3)$$

で与えられることを利用する方法である。

表1の分割表において $n=40$，$e=15$ という条件の下で考えられる分割表は全部で416個，同様に $n=40$ とし $e=16, 17, 18, 19, 20$ の場合に考えられる分割表の個数はそれぞれ 425, 432, 437, 440, 441 となる。Fisher's exact test における両側確率の検討は，上記2591個の分割表について，Finney の提案法による両側確率と Fisher の提案法による両側確率をそれぞれ計算して比較・検討を行った。

Table 1. Explanation of a 2×2 Contingency Table

	Outcome 1	Outcome 2	Total
Factor 1	*a*	*b*	*e*
Factor 2	*c*	*d*	*f*
Total	*g*	*h*	*n*

Marginal frequencies e, f, g, h, are defined as follows: e=a+b, f=c+d, g=a+c, h=b+d, and n=e+f=g+h
Row: Predicting variable Factor 1 and Factor 2.
Column: Outcome 1 and Outcome 2

結　果

調査を行った諸プログラムについて，2×2分割表の取り扱い方法の一覧を**表2**に示した。いずれのプログラムにおいても，Pearson の方法あるいは Yates の補正法による両側確率，および，Fisher's exact test による結果も表示されていた。もちろん，例数が大きくなると Fisher's exact test の計算は困難になってくるので，このような場合には結果の出力に*記号などを表示して適宜回避するようになっていた。

Fisher's exact test の片側確率と両側確率については，Statcel2 と Kyplot においては，片側確率のみが出力される形式であった。一方，SPSS と jmp については片側確率および両側確率の双方が同時に出力されていた。R の Fisher. test については，default が両側確率であるが，指定によって片側確率が表示されるようになっていた。また両側確率を表示するプログラムでは Finney の提案法による両側確率が出力されていた。片側確率が与えられれば，Fisher の提案した片側確率の2倍は容易に求められる点を考慮するなら，Fisher's exact test における両側確率に Finney の提案を採用するか，片側確率の2倍を採用するかは研究者自身の判断に委ねられていると考えてよいだろう。

表1で $n=40$，$e=15$ とした2×2分割表の Fisher's exact test について，2種の計算による両側確率の値を比較したところ，ほとんどの場合におい

Table 2. Specifications of Computation for the Probability Regarding a 2×2 Contingency Table in Several Statistical Softwares

	Pearson' s method	Yates correction	Fisher' s exact test
Statcel (ver.2)	two-sided	two-sided	single-sided
Kyplot (ver.5)	two-sided	two-sided	single-sided
SPSS (ver.15)	two-sided	two-sided	single-sided, two -sided
jmp (ver.7)	two-sided	not in service	single-sided, two -sided
R (ver.2.8.1)	two-sided	two-sided	single-sided, two -sided*

*Choice by command

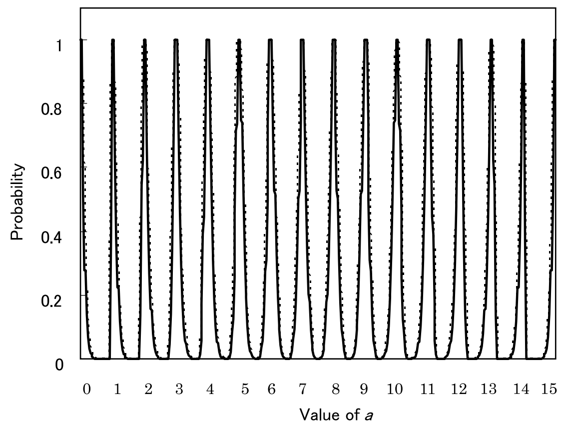

Figure 1. Comparison of the two-sided probability resulting from two different methods: probability from Finney's method (—) and twice of Fisher's single-sided probability(----).
—: Probability computed by Finney's method
----: Probability computed by Fisher's method

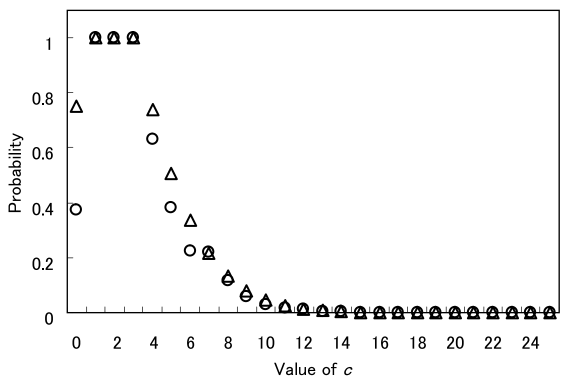

Figure 2. Differences in the values from two different methods in a specific range: $a=1$, $b=14$ in the contingency table illustrated in Table 1.
○: Probability computed by Finney's method
△: Probability computed by Fisher's method

Table 3. Summary of 2×2 Tables Which Resulted in That the Probability from Finney's Method was Smaller than 0.05 Although the Twice of Fisher's Single-sided Probability was Larger than 0.05.

e=15 (a, c)	e=16 (a, c)	e=17 (a, c)	e=18 (a, c)	e=19 (a, c)
(0, 7)	(1, 9)	(0, 6)	(3, 11)	(0, 5)
(2, 12)	(2, 11)	(2, 10)	(4, 0)	(1, 7)
(3, 0)	(3, 13)	(3, 12)	(5, 14)	(3, 10)
(3, 14)	(5, 1)	(4, 0)	(8, 3)	(4, 0)
(6, 19)	(6, 2)	(13, 23)	(9, 18)	(5, 13)
(7, 20)	(7, 19)	(14, 11)	(9, 4)	(6, 1)
(8, 5)	(8, 20)	(15, 13)	(10, 19)	(9, 3)
(9, 6)	(8, 4)	(17, 17)	(13, 8)	(9, 17)
(12, 11)	(9, 5)		(14, 22)	(10, 4)
(12, 25)	(10, 22)		(15, 11)	(10, 18)
(13, 13)	(11, 23)			(13, 20)
(15, 18)	(13, 11)			(14, 8)
	(14, 13)			(15, 21)
	(15, 15)			(16, 11)
				(18, 14)
				(19, 16)

て Finney の方法による確率の値の方が片側確率の2倍より小さいものであった (**図1**)。なお，**図2**には図1の一部を拡大して，両者の差が顕著である部分を示した。

片側確率の2倍で有意差ありと結論される分割表は194個あり，これら194個の分割表は Finney の方法でも有意差ありとなる。Finney の方法では更に12個の分割表について有意差ありと結論される。e=16, 17, 18, 19 についても同様の結果であり，片側確率の2倍により得られる両側確率の方がわずかに保守的であると考えられる。また e=20 の場合，2つの両側確率は完全に一致するが，これは (3) 式の超幾何分布の対称性によるものである。

表3には，検討した全ての2×2分割表において，

Finney の方法では $p<0.05$ となるものの片側確率の2倍では $p>0.05$ となる60例を列挙した。その逆に片側確率の2倍では $p<0.05$ となるものの，Finney の方法では $p>0.05$ となった例はなかった。また，双方ともに $p<0.05$ となるのは1224例であった。

考　察

2×2分割表の検定では Fisher's exact test は χ^2 検定よりも保守的であるとされているが，それは両側確率同士で比較した場合である。計算に使用したプログラムが Fisher's exact test については片側確率を提示するように設計されている場合には，Fisher's exact test の結果の方がより低い確率を提示する場合もある。

Fisher 自身は，2×2分割表の検定は2つの方法の間に優劣があることを前提とした片側検定の方法として考えたようであるが，データ処理の中では計算機から出力された確率の値を，片側，両側の違いを認識することなく参照する場合もあろう。

しかし，**表2**に示したように使用するプログラムによって，Fisher's exact test の確率は片側確率のみを提示するものもあるので注意が必要である。出力される結果をみると，χ^2 検定で得られる確率の方が Fisher's exact test で得られる確率よりも大きいことがある。それは χ^2 検定では両側確率が表示され，Fisher's exact test では片側確率が表示されているというだけのことから生じている可能性もあるので注意すべきである。

研究の結果を2×2分割表にまとめた場合，χ^2 検定と Fisher's exact test のどちらを用いるべきか，ということに関しては，はじめに述べた Cochran の提言にしたがって問題はないと思われるが以下の点に注意したい。Cochran の提言の意味は例数が小さいときは Fisher の方法を用い，例数が大きいときには Fisher の方法の代用として Yates の方法を用いよう，という提言である。はじめにで述べた Fisher, Yates の方法の関係を考えれば，Cochran が推奨した方法は Fisher の方法であると解釈してよいだろう。当時は計算機の性能の面で，$n=40$ に線を引いたに過ぎない。そうした背景を考慮するなら2×2分割表の検定には Fisher's exact test の両側確率を用いることが望ましいと考えられる。

分割表には必ずしも対称性がないから，Finney の方法は，片側確率を2倍するよりもより正確な計算方法のように見える。しかし，片側確率を2倍するという Fisher のコメントからは，検定の保守性を高め安易に有意差ありと結論しない方がよいという考えがうかがえる。実際，今回の調査を行った範囲では Fisher の方法は Finney の方法よりも保守的であった。研究において両側確率を検討する際に，この二つの方法で判定が食い違う珍しいケースでは慎重な検討を行うことが必要であるが，どちらの方法を選択するかは，研究者自身の考え方に依存するといってよいであろう。

なお，本稿では Fisher's exact test における両側確率について述べたが，研究計画の段階から，片側検定によって判定を下すべき事例として進められたものに関しては，研究者自身の考え方に基づいて片側検定を用いることに問題はない。

参考文献

1) Fisher, R. A. On the interpretation of χ^2 from contingency tables, and the calculation of P. Journal of the Royal Statistical Society 1922; 85: 87–94.
2) Yates, F. Contingency table involving small numbers and the χ^2 test. Journal of the Royal Statistical Society (Supplement) 1934; 1: 217–235.
3) Cochran, W. G. Some methods for strengthening the common χ^2 tests. Biometrics 1954; 10: 417–451.
4) Petrie, A. and Sabin, C. 一目でわかる医科統計学，吉田勝美監訳，メディカル・サイエンス・インターナショナル，東京，2001，55–57.
5) Yates, F. Tests of significance for 2×2 contingency tables. Journal of the Royal Statistical Society A 1984; 147: 426–463.
6) 柳井久江．4 steps エクセル統計，第2版，星雲社，東京，2004: 206–210.

Abstract

Two-sided and Single-sided Tests in a 2×2 Contingency Table

Takahiko Ueno and Shinobu Tatsunami

The test of independence for 2×2 table (or the test of a difference of proportions for 2 groups) is popular statistical method in the field of medical research. There are two ways for the handling of the table; one is the χ^2-test proposed by Pearson, the other the Fisher's exact probability test. The χ^2-test is a two-sided test, on the other hand, the Fisher's exact probability test is a single-sided one. However, this is not well known among end users. We found that some of commercial statistical programs provide single-sided probability only, while others provide both of single-sided and two-sided probabilities. We compared the value of two-sided probability from Finney's method with the twice value of Fisher's single-sided one. It was found that the former was always smaller than the latter in the region $p<0.05$. Therefore, the latter was more conservative. We have to recognize the characteristics of the two-sided probabilities from the two different methods in the application of the Fisher's exact probability test.

Key Word

2×2 contingency table, single-sided test, two-sided test, Fisher's exact test

Medical statistics, St. Marianna University School of Medicine