

臨床研究のための統計学
— IV. 特殊な形式のデータの検定 —

高 木 廣 文

臨 床 泌 尿 器 科

第 40 卷 第 10 号 別刷
1986年 10 月 20 日 発行

医学書院

● 講座/臨床研究のための統計学

IV. 特殊な形式のデータの検定*

高 木 廣 文**

1. 対応のあるデータからなる2標本の比較

例題 4・1

高尿酸血症患者に対して、無作為に9例にA薬を、8例にB薬をそれぞれ投与した。投与前と投与1週間後の尿酸値を調べたところ第1表のようなデータが得られた。これから、両薬間に効果の差があるといえるだろうか。

【検定法 4・1】

2群の各標本についてのデータは、投与前後の対応のあるものとなつている。各薬剤の効果は、投与前後の測定値の差で示せるものと考えられるので、まず第2表のように測定値の差を各標本について求める。なお、第2表では投与前の値から投与後の値を引いたものを示している。

第2表のような形式にデータを変換することにより、すでに連載1回目で取り上げた「2つの母平均値の差の検定」の問題として扱えるようになる。解説1・1, 1・2を参考に検定を行う。

まず、A薬群の平均値 $\bar{x}_A$ と不偏分散 s^2_A を求める。第2表に示した前後の差の総和と2乗和から、
 $\bar{x}_A = 26.3/9 = 2.9222$

および、

$$s^2_A = 90.61/8 - 2.9222^2 \times 9/8 = 1.7196$$

のように求められる。同様に、B薬群の平均値 $\bar{x}_B$ と不偏分散 s^2_B は、

$$\bar{x}_B = 36.2/8 = 4.525$$

および、

$$s^2_B = 181.36/7 - 4.525^2 \times 8/7 = 2.5079$$

となる。

第1表 2群での尿酸値 (単位 mg/dl)

A 薬 群		B 薬 群	
前	後	前	後
11.0	6.5	13.5	7.5
10.2	8.0	10.7	6.6
9.8	7.5	11.2	9.3
10.9	8.1	13.1	6.8
9.0	7.2	10.4	5.9
9.7	8.9	9.6	6.8
10.3	6.8	14.0	8.0
13.0	9.5	9.9	5.3
11.6	6.7		

第2表 前後の差のデータ

A 薬 群		B 薬 群	
4.5		6.0	
2.2		4.1	
2.3		1.9	
2.8		6.3	
1.8		4.5	
0.8		2.8	
3.5		6.0	
3.5		4.6	
4.9			
総 和 26.3		36.2	
2 乗和 90.61		181.36	

すでに、説明したように「2つの母平均値の差の検定」は2群の母分散が等しいか否かにより用いる方法が異なる。したがって、まず「等分散の検定」を行う。

2つの不偏分散の比 F_0 は、

$$F_0 = 2.5079/1.7196 = 1.458$$

である。 F_0 は自由度 (8, 7) のF分布に従う。ま

* Statistics in Clinical Studies: IV. Methods of the Test for the Special Samples

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

た、数値表から自由度(8, 7)のF分布の上側2.5%点の値は4.899であることがわかる。 F_0 の値はこれより小さいので、有意とはならず、等分散を仮定することができる(解説1・2参照のこと)。したがって、ここでは等分散の場合の2つの母平均値の差の検定を行う。

解説1・1で示した(1.3)式を用いて、

$$t_0 = \frac{|2.9222 - 4.525|}{\sqrt{\frac{(8 \times 1.7196 + 7 \times 2.5079)}{(9+8-2)} \times \frac{(9+8)}{9 \times 8}}} \\ \div 2.283$$

を求める。 t_0 はt分布に従い、その自由度は(1.4)式から、 $9+8-2=15$ と求められる。自由度15のt分布の両側5%点の値は、数値表から2.131、1%点は2.602であることがわかる。 t_0 の値は5%点より大きい、1%点よりは小さい。したがって、5%水準で有意差があるものと認められる。

この結果、A薬とB薬には尿酸値の減少に関する効果に差があり、B薬での効果がより大きいものと考えられる。

【解説4・1】

第1表に示したデータは、2群の各標本がそれぞれ対応のある場合のものである。このような場合、各群での薬効の有無を検定することもできるし、例題のように2群間での効果の差を検定することもできる。

いま、標本*i*のある処理の前後での特定の変数の測定値を x_i, y_i としよう。仮定されるモデルは、

$$x_i = \mu + \varepsilon_i \quad (4.1)$$

というものである。ここで μ は、母平均値である。また、 ε_i は誤差項であり、分布として平均0、分散 σ^2 の正規分布に従うものと仮定するのが普通である。

y_i については、上記のモデルに処理の効果を加えて、

$$y_i = \mu + \delta + \varepsilon'_i \quad (4.2)$$

とする。ここで、 μ は(4.1)式の場合と同様であり、 δ は処理の平均的な効果を示す定数である。 ε'_i は同様に誤差項であり、分布として平均0、分散 σ^2 の正規分布を仮定する。

ここで、2つの測定値の差をとると、

$$d_i = y_i - x_i = \delta + \varepsilon''_i \quad (4.3)$$

となる。誤差の分布としては、上記と同様に正規

分布を仮定する。

(4.3)式を仮定して、処理の効果について仮説「 $\delta=0$ 」を検定するのが、解説2・1で説明した「対応のある場合の母平均値の差の検定」である。

例題のように、対応のあるデータが2組ある場合には、第2表のような形式にすることで、通常の独立な2標本の母平均値の差を検定する問題として扱えるようにできる。これは、(4.1)式と(4.3)式を比べるとわかるように、2つの母平均 μ_1 と μ_2 について仮説「 $\mu_1 = \mu_2$ 」を検定する問題と、処理の効果についての仮説「 $\delta_1 = \delta_2$ 」を検定する問題が、形式的に同一であることによる。すなわち、誤差の分布に正規分布を仮定することにより、解説1・1、1・2で説明したように、t分布を利用した検定が可能となる。また、誤差の分布に正規分布を仮定するのが適当でない場合には、解説1・3で説明した「Wilcoxonの順位和検定」を用いることもできる。詳しい方法については、それぞれの解説を参照して欲しい。

2. 測定値の比で表されたデータ

例題4・2

急性副睾丸炎患者に対し無作為に、それぞれ6例にA薬を、5例にB薬を投与し、投与前に対する投与開始3日目の末梢白血球数の比を求めた。第3表は各薬群ごとに投与前の数値を100とした場合の、投与後の白血球数の比を示したものである。このデータから、両群間の効果に差があるといえるだろうか。

第3表 投与前後の白血球数の比

A 薬 群	B 薬 群
92	80
85	75
88	78
95	82
84	88
74	

(数値は投与前の値を100とした場合のもの)

【検定法4・2】

このようなデータでは、単純に比の平均値を求めて、2つの母平均値の差の検定を行うのは適当

第4表 白血球数の比の順位

A 薬 群	B 薬 群
10	4
7	2
8	3
11	5
6	9
1	
順位和 43	23

ではない。後述するように、2つの変数の比をとつたものの（誤差の）分布に正規分布を仮定するには、かなり無理があるためである。

ここでは、投与前後の測定値の比の大きさの順序に意味があるものとして、すでに解説1・3で説明した「Wilcoxon の順位和検定」を行う。第4表は、第3表のデータについて、数値の小さい順に1から11まで順位を付けたものである。順位和を求めると、表中に示したようにA薬群43、B薬群23となる。すでに説明したように、数値表から有意となる限界の順位和を求め、それとの比較で検定が可能である。

数値表から、2つの標本数が6と5の場合の下側2.5%点の順位和は18であることがわかる。小さい方の順位和は23であり、これより大きいので有意とはいえない。したがって、両側5%水準で両薬に有意差があるとは言えないことになる。

【解説 4・2】

第3表のようなデータを扱う場合、よく犯す誤りはデータが「比率（百分率）」を示すものとして分析することである。

たとえば、ある薬物を50名に服薬させたところ30名の症状の改善がみられたとしよう。この場合、改善率は $30/50=0.6$ 、すなわち60%であるが、このように全体の個数に対するある部分の個数の割合を「比率（rate）」という。

これに対して、男の人数が50、女の人数が40の場合、 $50/40=1.25$ であるが、これは「比（ratio）」という。比は、単に2つの変数の測定値の除算の結果である。したがって、第3表に示したような、投与前後の測定値について、一方を他方で除したものは「比」である。

上記のように「比率」と「比」はよく混同され

第5表 尿酸値の比

A 薬 群	B 薬 群
0.591 (7)	0.556 (3)
0.784 (14)	0.617 (8)
0.765 (13)	0.830 (16)
0.743 (12)	0.519 (1)
0.800 (15)	0.567 (4)
0.916 (17)	0.708 (10)
0.660 (9)	0.571 (5)
0.731 (11)	0.535 (2)
0.578 (6)	

〔比は、投与後／投与前、() 内は順位〕

るが、明確に区別すべきものである。

一般に、データを比の形で収集するのは避けるべき方法である。比が意味を持つ場合というのは、2変数 X, Y において、標本 i のデータ x_i, y_i の比の間に、

$$y_i/x_i = \rho + \varepsilon_i \quad (4.4)$$

のような関係が成り立つ場合だけだからである。ここで、 ρ は定数、 ε_i は誤差項であり、正規分布に従うものとする。

すでに述べたように、変数の測定値について(4.1)式のようなモデルが妥当であり、誤差が正規分布に従うと仮定できる場合、そのような2変数の測定値の比について、(4.4)式の関係が成り立つとは考えがたい。したがって、比で得られたデータについて t 検定などを用いるべきではない。場合によっては、2変数の比に関して「Fiellerの式」を用いて信頼区間を求めたり、検定も行えるが、これについては文献を参照して欲しい。

例題4・2では、比の大小関係に意味のあるものとして順位和検定を行つたが、この方法が必ずしも妥当なものかどうかは、若干問題があるかもしれない。すなわち、比の大きさの順位といったものが、処理の効果を反映するものかという問題である。これは、研究上設定する仮説の問題であるといえよう。できれば比のデータではなく、第1表のように投与前後の測定値を、そのままデータとして収集した方が分析上好ましい場合が多い。一般に、データが数値で測定される場合、そのままの数値をデータとして保存し、必要に応じて加工した方が研究上便利なが多いものと考えられる。初めから、測定値を適当に区切つて、「正

常」「境界」「異常」のような形や比をとつたもので集めると、もとのデータが必要な場合に、測定値を再現することは不可能である。

第1表のデータが第5表に示すように、投与前後の測定値の比で示されているとして、順位と検定を行つてみよう。表中の()内に順位を示しておいた。A薬群の順位和は104, B薬群の順位和は49である。数値表から、2つの標本数が9と8の場合の下側2.5%点の順位和を求めると51であることがわかる。B薬群の順位和は49であり51より小さいので、両側5%水準で有意差が認められることになる。この結果は、検定法4・1の結果と大差ない。したがって、第1表のデータに関しては、データの差をとるか、比をとるかによつて検定結果に影響はほとんどないことがわかる。差に意味があるのか、比に意味があるのかは、研究上の仮説やモデルに基づいて、吟味されるべきであらう。しかし、第3表のデータでは投与前後の差のデータは得られないため、このような比較は不可能である。特定のデータに対する検定方法は、1つだけとは限らないので、生のデータを保存しておくことが肝要と思われる。

3. 発生率の小さな特性の比較

例題 4・3

人口23,500人のA町と人口6,800人のB村で、5年間に発見された膀胱癌患者数は、それぞれ15名と8名であつた。2つの集団で膀胱癌の発生率に差があるといえるだろうか。

【検定法 4・3】

地域における特定の疾病の罹患率や死亡率を比較する場合、実際には各地域の男女比や年齢別の人口構成の相違などが問題となり、死亡率などの標準化を行う必要が生じるが、ここではその点に関しては、考慮しないことにする。

各地区の5年間での発生率を求めると、A町では、

$$p_A = 15/23,500 \div 6.383 \times 10^{-4}$$

となり、B村では、

$$p_B = 8/6,800 \div 1.177 \times 10^{-3}$$

となる。見かけ上は、B村の方が発生率は高いよ

うである。

2地区の発生率は、上記のように極めて小さく、かつ人口数が大きいため、膀胱癌の発生はポワソン分布に従うものと考えてよいだろう。2地区での発生率に差がないとすると、発生数の差は地区の人口数によるものとなるだろう。2地区全体に対するA町の人口の割合は、

$$p = 23,500/(23,500 + 6,800) = 0.7756$$

である。上記の人口の割合で、患者の発生数の差が生じるとすれば、患者数の分布は2項分布に従うことになる。

したがって、23名中15名以下の患者がA町で発生する確率は(4.11)式を用いて、

$$P_0 = \sum_{k=0}^{15} \frac{23!}{k! \cdot (23-k)!} \cdot p^k (1-p)^{23-k} \div 0.1233$$

と求められる。上記の確率は片側検定のためのものなので、数値を2倍すると0.2466、すなわち約24.7%となり、5%水準に達しない。

この結果から、2地区で膀胱癌の発生率に差があるとはいえないことになる。

【解説 4・3】

ある事象の発生率をPとすると、n人中k人(k=0, 1, 2, …, n)にその事象が発生する確率 prob(k)は、

$$\text{prob}(k) = \frac{n!}{k! \cdot (n-k)!} \cdot p^k (1-p)^{n-k} \quad (4.5)$$

により求められ、その分布は「2項分布」と呼ばれる。

2項分布において、発生率Pが極めて小さく、かつ標本数nが極めて大きい場合、

$$m = np \quad (4.6)$$

としたとき、発生数がkとなる確率は、

$$\text{prob}(k) = \frac{m^k}{k!} \cdot e^{-m} \quad (k=0, 1, \dots) \quad (4.7)$$

により近似することができる。ただし、eは自然対数の底2.71828……である。上記の分布は「ポワソン分布」と呼ばれるものである。

例題4・3は、2つのポワソン分布の比較ということになるが、これは2つの母比率の差の検定を行うのと同じことになる。2つの集団での標本数を n_1, n_2 とし、特定の事象の発生数が k_1, k_2 である場合、その発生率はそれぞれ、

$$p_1 = k_1/n_1, \quad p_2 = k_2/n_2 \quad (4.8)$$

となる。 p_1, p_2 が0に近ければ、標本数が極めて大きい場合、その事象の発生は独立にポワソン分布に従うものとみなせる。各発生率が比較的大きい場合、「2つの母比率の差の検定」のためには解説3・2で説明した「4分表の検定」を用いればよい。

発生率が非常に小さい場合、仮説「2つの母比率は等しい」を検定するためには以下の方法を用いればよい。2群での発生数の和を、

$$t = k_1 + k_2 \quad (4.9)$$

とする。発生率に差がなければ、2群での発生数は各標本数にのみ依存するはずである。いま、2つの発生率に関して $p_1 < p_2$ であるとする。標本での発生率の小さい群(群1)の、全体に対する標本数の割合 P は、

$$p = n_1 / (n_1 + n_2) \quad (4.10)$$

となる。全体の発生数 t が与えられた場合、群1での発生数は比率を P とした2項分布に従い、発生数が k_1 以下となる確率 p_0 は、

$$p_0 = \sum_{k=0}^{k_1} \frac{t!}{k! (t-k)!} \cdot p^k (1-p)^{t-k} \quad (4.11)$$

により求められる。 p_0 は片側確率なので、両側検定のためには、値を2倍して用いればよい。例題では上記の式を用いたが、標本数の比率 P が0や1に近い値でなければ、以下の正規分布による近似を用いることができる。すなわち、

$$z_0 = \frac{|k_1 - tp| - 0.5}{\sqrt{tp(1-p)}} \quad (4.12)$$

を求める。 z_0 は標準正規分布に従うので、両側5%点1.96または1%点2.58と比べ、 z_0 が大きい場合に有意差ありとすればよい。なお、(4.12)式の分子の-0.5は人数のような離散的な数値を用いる場合、連続的な分布への適合がより良くなるようにするための「連続性の補正」を行うためのものである。

例題4・3について、(4.12)式を用いて z_0 を求めると、比率 $p = 0.7756$ として、

$$z_0 = \frac{|15 - 23 \times p| - 0.5}{\sqrt{23 \times p \times (1-p)}} \div 1.16866$$

となる。この値は、標準正規分布の両側5%点の値1.96より小さいので、有意差があるとはいえないことになる。また、有意確率を計算すると両側

第6表 2地区での膀胱癌の発生数

地 区	患 者	そ の 他	計
A 町	15	23,485	23,500
B 村	8	6,792	6,800
計	23	30,277	30,300

で約24.2%となり、検定法4・3で求めた確率と大差ないことがわかる。したがって、実際には(4.11)式を用いるよりも、(4.12)式による正規分布近似を用いる方が簡単であろう。

比較のために例題4・3を第6表のように4分表にして示し、これについて解説3・2で説明したカイ2乗検定を行ってみよう。(3.6)式を用いて、 $\chi^2_0 = \frac{30300 \times (|15 \times 6792 - 23485 \times 8| - 30300/2)^2}{23500 \times 6800 \times 23 \times 30277}$

$$\div 1.36680$$

となる。 χ^2_0 は自由度1のカイ2乗分布に従う。

ところで、独立に標準正規分布に従う変数の2乗和はカイ2乗分布に従う。すなわち、標準正規分布に従う1つの統計量を2乗したものは、自由度1のカイ2乗分布に従うことになる。そこで、正規分布近似による検定のために求めた z_0 について、その2乗値 z^2_0 を計算すると、

$$z^2_0 = 1.16866^2 \div 1.36576$$

となる。

これと χ^2_0 の値を比較すると、 χ^2_0 の値が若干大きいものの、両者の数値は小数第2位まで一致し、どちらを用いても検定結果に重大な差を与えないことがわかる。

上記のことから、有意差を検定するためには、ここで説明したどの方法を用いても大差ないものと考えられる。したがって、有意確率を正確に求める必要がなければ、計算の簡便さからいって、(4.12)式の正規分布近似による検定法が扱いやすいものといえるだろう。

なお、2項分布とポワソン分布に関して詳しく知りたい場合には、文献2を参照されたい。

文 献

- 1) 佐久間昭：薬効評価—計画と解析I，東大出版，1977。
- 2) 竹内 啓，藤野和建：2項分布とポワソン分布，東大出版，1981。