

# 実際のデータ解析の方法を考える

高木 廣文

平成12年 2 月  
超音波検査技術 Vol. 25 No. 1

## 統計講座 (第7回)

## 実際のデータ解析の方法を考える

高木 廣文

## はじめに

実際のデータは、人数も多く質問項目も多いため、分析をする場合に、適切な方法がわからないことも多いだろう。そこで今回は、新しい方法の説明とともに、これまで説明した方法を用いて、簡単な例題の解析方法を考えてみよう。

表1は、心不全の患者に対して経皮経管的僧帽弁拡張術 (PTMC) を実施したときの手術前後の心不全の程度 (NYHA分類による) と、僧帽弁逆流の程度のデータを示したものである (例題のための架空例)。実際のデータに比べれば、表1のデータははるかに小さなものであるが、データ解析における幾つかの重要な手順を示すことができるので、以下に解析の手順を考えていこう。

## クロス表とその検定

まず、表1を見てわかることをあげてみよう。すぐに分かることであるが、

(1) データは、いわゆる質的なデータである。すなわち、心不全の程度も僧帽弁逆流の程度も、幾つかのカテゴリーに分類されてデータが収集されている。

また、PTMCという手術の前後でのデータであり、

(2) それぞれのデータが各ケースについて、繰り返し収集されている。すなわち、対応のあるデータ、ということになる。

普通考えると、質的なデータであり、2度ずつ

表1 PTMC (経皮経管的僧帽弁拡張術) 前後の比較

| ID | 心不全の程度 |       | 僧帽弁逆流の程度 |       |
|----|--------|-------|----------|-------|
|    | PTMC前  | PTMC後 | PTMC前    | PTMC後 |
| 1  | IV     | III   | II       | III   |
| 2  | I      | I     | なし       | I     |
| 3  | I      | I     | I        | I     |
| 4  | II     | I     | II       | II    |
| 5  | III    | II    | なし       | I     |
| 6  | II     | III   | I        | III   |
| 7  | IV     | II    | II       | II    |
| 8  | III    | IV    | II       | IV    |
| 9  | I      | I     | なし       | I     |
| 10 | II     | I     | なし       | I     |
| 11 | IV     | II    | II       | II    |
| 12 | III    | II    | I        | I     |
| 13 | III    | I     | なし       | なし    |
| 14 | II     | II    | I        | II    |
| 15 | II     | I     | I        | I     |
| 16 | IV     | II    | II       | I     |
| 17 | III    | II    | I        | I     |
| 18 | II     | I     | I        | I     |
| 19 | I      | I     | なし       | なし    |
| 20 | III    | I     | なし       | I     |

1) 心不全の程度 (NYHA分類)

I : 日常生活に制限がなく、心不全症状がない。

II : 日常生活がわずかに制限される程度で日常の労作で動悸、息切れが生じる。

III : 日常生活が制限され、普通以下の労作で動悸、息切れが生じる。

IV : 日常生活ができず、安静時にも心不全症状がある。

2) 僧帽弁逆流の程度 (心エコー図法による分類)

なし : 逆流なし, I : 極少量, II : 少量,

III : 中等量, IV : 大量。

Hirofumi TAKAGI : 新潟大学医学部保健学科

〒951-8518 新潟県新潟市旭町通2番町746番

e-mail : takagi@clg.niigata-u.ac.jp

URL : <http://www.clg.niigata-u.ac.jp/~takagi/>

データが収集されているので、とりあえず前後の関係を見るために、クロス集計を行うことが多い

かもしれない．実際にクロス集計を行うと，表 2，表 3 のようになる．このような集計を行う場合，実際には解析のための統計学ソフトを使用する前に，各項目のカテゴリで示されたデータを，あらかじめ 1，2，3 などの数値に変換しておくことと便利である．最近のコンピュータのソフトでは，表 1 のような記号や漢字などを，そのまま処理を行えるものもあるが，解析が単なる集計などに限られてしまう．後で説明するような解析方法を行うためには，あらかじめ I，II，III などの記号は，1，2，3 などの数値に変換しておくこととよい．また，僧帽弁逆流の程度のカテゴリに「なし」があるが，これについては，0 という数値をデータとして用いればよいだろう．

表 2，表 3 は 2 つの項目のカテゴリを組み合わせて集計した結果を示したもので，「クロス表」と呼ばれるものである．クロス表は，2 つの項目の関連を調べるために用いられるものである．

表 2 を見てわかることは，表の左下側に数値が多く固まっていることである．すなわち，PTMC 前に比べて PTMC 後では，心不全の程度が数値の小さくなっていることがわかる．また表 3 では，表 2 とは逆に，PTMC 後の方で数値

が大きくなっているようである．ただし，この結果は表を見て主観的に考えたものであり，これが偶然の変動によって起きるかどうかを，統計学的に検討する必要があるだろう．その方法は，通常はクロス表のカイ 2 乗検定によることが多い．以下に簡単にその方法を用いてみよう．

クロス表のカイ 2 乗検定は，集計に用いた 2 つの項目が互いに独立，つまり無関係であるという仮定のもとで，観察されたクロス表が偶然できる確率を求めることで行われる．クロス表の 2 項目を，ここでは A，B と呼ぼう．この 2 つの項目 A と B の各カテゴリの組み合わせによって集計された，個々の数値が記載されているクロス表の各箇所は「セル」と呼ばれる．各セルの頻度は，2 項目 A，B が独立ならば，A の各カテゴリごとの単純集計  $n_1, n_2, \dots, n_K$  と，B の各カテゴリごとの単純集計  $m_1, m_2, \dots, m_L$  とから，全体の人数を  $N$  とした場合，

$$E_{ij} = \frac{n_i}{N} \times \frac{m_j}{N} \times N = \frac{n_i \times m_j}{N}$$

となることが期待されるだろう ( $i=1,2,\dots,K$ ;  $j=1,2,\dots,L$ )．この  $E_{ij}$  は，そのセルの「期待値」と呼ばれている．この期待値  $E_{ij}$  と実際に観察された頻度  $f_{ij}$  を用いて，カイ 2 乗値が求められる．

$$\text{カイ2乗値} = \frac{(\text{期待値} - \text{実際値})^2}{\text{期待値}} \text{の全てのセルの合計}$$

である．

上の計算式でわかるように，期待値と実際値が近いほどカイ 2 乗値は小さくなる．逆に，期待値と実際値の差が大きくなるにつれ，カイ 2 乗値は大きくなり，その有意確率 ( $p$  値) は小さくなる．一般に， $p$  値が 5 % 以下ならば，2 つの項目に統計学的に有意な関係があると考えることが多い．

検定のための  $p$  値の計算は，カイ 2 乗分布をもとに計算される．ここで，カイ 2 乗分布とは，標準正規分布に従う変数から，無作為に抽出したデータの 2 乗和の分布として定義されるものである．したがって，何個のデータの 2 乗和の分布かによって，その分布の形状が変わってくるし，当然，同じカイ 2 乗値でもその確率が異

表 2 PTMC前後の心不全の程度(NYHA分類)の比較

| PTMC前 | PTMC後     |          |          |         |           |
|-------|-----------|----------|----------|---------|-----------|
|       | I (%)     | II (%)   | III (%)  | IV (%)  | 合計 (%)    |
| I     | 4 (20.0)  | 0        | 0        | 0       | 4 (20.0)  |
| II    | 4 (20.0)  | 1 (5.0)  | 1 (5.0)  | 0       | 6 (30.0)  |
| III   | 2 (10.0)  | 3 (15.0) | 0        | 1 (5.0) | 6 (30.0)  |
| IV    | 0         | 3 (15.0) | 1 (5.0)  | 0       | 4 (20.0)  |
| 合計    | 10 (50.0) | 7 (35.0) | 2 (10.0) | 1 (5.0) | 20 (100.) |

表 3 PTMC前後の僧帽弁逆流の程度(心エコー図法による)の比較

|    | PTMC後   |          |         |         |        |          |
|----|---------|----------|---------|---------|--------|----------|
| 前  | なし(%)   | I(%)     | Ⅱ(%)    | Ⅲ(%)    | Ⅳ(%)   | 合計(%)    |
| なし | 2(10.0) | 5(25.0)  | 0       | 0       | 0      | 7(35.0)  |
| I  | 0       | 5(25.0)  | 1(5.0)  | 1(5.0)  | 0      | 7(35.0)  |
| Ⅱ  | 0       | 1(5.0)   | 3(15.0) | 1(5.0)  | 1(5.0) | 6(30.0)  |
| 合計 | 2(10.0) | 11(55.0) | 4(20.0) | 2(10.0) | 1(5.0) | 20(100.) |

なってくる。標準正規分布から何個のデータを抽出したものによるカイ 2 乗分布なのかを示すために、「自由度」という用語を使用する。クロス表から計算されるカイ 2 乗値の自由度は、(A のカテゴリ数 - 1) × (B のカテゴリ数 - 1) となる。検定のための有意確率 (p 値) は、統計学ソフトを用いて計算すれば、通常は出力されるので、自由度のことはあまり気にする必要はない。

表 2 の心不全の程度を PTMC 前後の関係について、その関連の程度を検定すると、カイ 2 乗値は 13.357 となり、自由度は  $(4-1) \times (4-1) = 9$  となり、その有意確率 (p 値) は 0.147 と有意ではない。また、表 3 の僧帽弁逆流の程度の PTMC 前後の関係については、カイ 2 乗値は 13.647 となり、自由度は  $(3-1) \times (5-1) = 8$  となり、その有意確率 (p 値) は 0.091 と同様に有意ではない。

クロス表の検定では 2 変数の間に、それぞれ有意な関係は認められなかった。しかし、ここで示した検定は、実は全く適切な検定ではない。それでは、どのような検定を用いれば、データの特性と研究目的に即した結果を得ることができるのか。

### ■ 適切な解析

心不全の程度も僧帽弁逆流の程度も、質的データとはいえ、カテゴリ間には自然な順序がある。通常のクロス表のカイ 2 乗検定は、カテゴリ間の順序を考慮していない。したがって、情報の有効利用の点から、必要な解析はその順序を考慮したものでなければならないだろう。また、データは PTMC の前後に繰り返し取られているので、このことも考慮する必要がある。

上記の点を考慮すると、適切な解説として少なくとも「対応のあるデータ」の解析方法を使用する必要がある。対応のあるデータの解析法として、前回説明した方法のうち、順序を考慮した「ウィルコクソンの符号付き順位和検定」が、考えられる解析方法の候補といえよう。

表 4 は PTMC 前後での心不全と、僧帽弁逆流の程度の順位の差を示したものである。符号付

表 4 PTMC 前後の順位差

| ID | 心不全 | 僧帽弁逆流 |
|----|-----|-------|
| 1  | 1   | -1    |
| 2  | 0   | -1    |
| 3  | 0   | 0     |
| 4  | 1   | 0     |
| 5  | 1   | -1    |
| 6  | -1  | -2    |
| 7  | 2   | 0     |
| 8  | -1  | -2    |
| 9  | 0   | -1    |
| 10 | 1   | -1    |
| 11 | 2   | 0     |
| 12 | 1   | 0     |
| 13 | 2   | 0     |
| 14 | 0   | -1    |
| 15 | 1   | 0     |
| 16 | 2   | 1     |
| 17 | 1   | 0     |
| 18 | 1   | 0     |
| 19 | 0   | 0     |
| 20 | 2   | -1    |

(注) 差 = PTMC 前 - PTMC 後

き順位和検定では、前後のデータが同一の値の場合、そのデータは検定には使用されない。心不全に関しては、前後が 0 のケースが 5 例あり、標本数は 15 として検定される。僧帽弁逆流のデータは、0 が 10 例あるので、有効な標本数は 10 と半分減ってしまう。

心不全では、差がマイナスのデータは 2 例であり、いずれも -1 である。プラスは 1 が 8 例、2 が 5 例となっている。差の絶対値が 1 のケースに与える中間順位は、 $(1+10)/2=5.5$ 、同様に差が絶対値が 2 のケースには、 $(11+15)/2=13$ 、の中間順位を与える。したがって、マイナスのデータの順位和は、 $-5.5 \times 2 = -11$ 、プラスの順位和は、 $5.5 \times 8 + 13 \times 5 = 109$ 、となる。

検定のためには、p 値を求める必要があるが、データに同順位があるので正規分布近似により求める。そのためには、

$$\text{分子} = |109 - 15 \times (15 + 1) / 4| = 49$$

$$\begin{aligned} \text{分母} &= \sqrt{[15 \times (15 + 1) \times (2 \times 15 + 1) / 24] \\ &\quad - \{(10^3 - 10) + (5^3 - 5)\} / 48]} \\ &= \sqrt{286.875} = 16.937 \end{aligned}$$

を計算し、これらから正規近似値  $z = 49 / 16.937 = 2.893$ , と求められる。

この  $z$  値 2.893 の標準正規分布での両側確率を求めると,  $p = 0.0038$  (0.38%), となる。この結果は、心不全の程度について、PTMC前後で統計学的な有意差が認められたことになる。すなわち、PTMCにより心不全の程度は有意に軽減されたものといえる。

同様に、僧帽弁逆流の程度について正規近似値を求めると,  $z = 2.486$ , と求められ,  $p = 0.0129$  (1.29%) となる。この結果は、僧帽弁逆流の程度についても、PTMC前後では統計学的な有意差があることを示している。符号付き順位和の絶対値は、マイナスの順位の方が大きいので、PTMCにより僧帽弁逆流の程度は増加していることになる。

上記のように、単純なクロス表のカイ 2 乗検定では、2 変数間の関係の有無を調べるために、統計学的に有意な結果は得られなかった。しかし、カテゴリ間の順位を考慮し、さらに対応のあるデータという、特殊なデータに対する検定方法を採用することで、極めて統計学的に有意な結果を得ることができた。このように、データのもつ情報を有効に利用するかしないかでは、結果に大きな違いを生ずることになる。

## ■ やりすぎの解析

ところで、カテゴリに与えた自然な順位 (1, 2, 3 など) を通常の数値データとして、以前説明した「対応のある場合の母平均値の差の検定」を用いることも考えられる。

表 4 に示した PTMC 前後の差のデータから、心不全については、差の平均値 0.8, 差の標準偏差 0.951 となり、これより検定のための  $t$  値は、

$$t = 0.8 \times \sqrt{20} / 0.951 = 3.76$$

となる。この  $t$  値は自由度  $20 - 1 = 19$  の  $t$  分布に従うことで検定が行える。  $p$  値を求めると,  $p = 0.0013$  (0.13%) となり、高度に有意である。

同様に、僧帽弁逆流の程度に関しても,  $t = 2.939$  となり,  $p = 0.0084$  (0.84%) とやはり高度に有意である。

$t$  検定による結果は、心不全と僧帽弁逆流の程度に関して、どちらも最も統計学的に有意な結果 ( $p$  値の小さい) となった。もっとも、これらの  $t$  検定の結果がノンパラメトリック検定である符号付き順位和検定の結果と符合することに気づいただろうか。  $t$  検定は、変数の分布が正規分布に従わないと、理論的には利用できないことになっている。実際には、少しくらい正規性からはずれた分布であっても、データの和 (平均値) に関する検定のための統計量を用いているので、そのようなずれに対して頑健 (ロバスト) になっている。したがって、測定データが少しくらい正規分布していないとしても、重大な誤りを生じることは少ないと考えられている。

しかし、この検定方法はデータが少なくとも間隔尺度 (数値の間隔が等間隔) であることが必要とされる。心不全の程度にしても、僧帽弁逆流の程度にしても、この前提は表 1 のデータでは怪しいものであり、順序尺度のデータに関して、データの数値を等間隔のスケールとして扱うので、やはり問題であろう。分析のために有意になりやすい検定方法を選べばよいというものではない。

## ■ 心不全と僧帽弁逆流の程度の関係は

これまでの分析では、PTMCの前後では、心不全の程度はカテゴリの小さい方向、すなわち軽快しているが、僧帽弁逆流の程度は大きい方向、すなわち逆流量は増加の傾向があることが示された。それでは、PTMC前後の心不全のと僧帽弁逆流の程度の関係はどうなっているのだろうか。当然、一方は減少、もう一方は増加であるから、負の相関があるものと予想されるだろう。実際に表 4 のデータから、両者のスピアマンの順位相関係数を求めてみよう。計算は省

略するが、スピアマンの順位相関係数は0.565と正の相関を示す。これは、全く予想に反する結果といえよう。なぜこのように矛盾する結果となったのだろうか。

一つには、符合付き順位和検定では、PTMC前後でデータが変わらなかったケースは、分析に使用されていないが、対応のある母平均値の差の検定では使用されているということである。すなわち、符合付き順位和検定では、心不全については5例の標本が分析に使用されず、僧帽弁逆流については標本の半分に当たる10例のデータが使用されていない。一方、t検定では20例全てのデータを使用している。

例題のように標本数が小さい場合、個別の分析と、変数をまとめた分析の結果が矛盾する場合が、このようにあり得る。このような場合、早急な解釈はさけ、もう少し標本数が増加して、安定した結果が得られるまで、結論は保留すべきであろう。

## ■ おわりに

今回は、前回の復習のような形式になったが、データの形式から分析方法を考えてみた。実際のデータとはかなり異なるかもしれないが、データを見て、それに相応しい解析方法を選ぶのは、日常的にデータ解析を行っている研究者以外は、結構難しいのではないだろうか。Q & Aのような形式の統計講座も、読者にはわかりやすいかもしれないので、今後、今回のような形式の講座とともに、少し工夫をしたものもやってみたいと思っている。

## 参考文献

- 1) 高木廣文：相関係数に関する統計学的方法，超音波検査技術，24(3)，208-215，1999.
- 2) 高木廣文：対応のあるデータの比較，超音波検査技術，24(6)，424-429，1999.