

臨床研究のための統計学
——V. 多群の比較——

高 木 廣 文

臨 床 泌 尿 器 科

第40巻 第11号 別刷
1986年11月20日 発行

医学書院

●講座/臨床研究のための統計学

V. 多群の比較*

高 木 廣 文**

1. 一元配置分散分析

例題 5・1

第1表に, 急性膀胱炎9例, 急性腎盂腎炎8例, および急性副睾丸炎8例の尿中細菌数を調査した結果を示した。疾患により尿中細菌数が異なるといえるだろうか。

第1表 尿中細菌数 (/ml) のデータ

急性膀胱炎	急性腎盂腎炎	急性副睾丸炎
5.5×10^6	4.6×10^5	3.1×10^5
6.3×10^6	8.3×10^6	0
5.7×10^7	6.7×10^4	5.3×10^6
8.1×10^7	2.9×10^5	3.6×10^4
3.2×10^5	7.4×10^5	1.2×10^3
1.5×10^7	7.9×10^5	4.7×10^4
7.7×10^7	8.9×10^4	0
8.8×10^6	1.3×10^6	0
4.4×10^5		

【検定法 5・1】

説明を簡単にするため, 急性膀胱炎をA群, 急性腎盂腎炎をB群, 急性副睾丸炎をC群と以後呼ぶことにする。3群全体での標本数は25である。まず, 各群の平均値 $\bar{x}_A$, $\bar{x}_B$, $\bar{x}_C$ および全体の平均値を求める。第1表のデータから,

$$\bar{x}_A \div 2.79289 \times 10^7, \bar{x}_B \div 1.50450 \times 10^6,$$

$$\bar{x}_C \div 7.11775 \times 10^5$$

$$\text{および } \bar{x} \div 1.07636 \times 10^7$$

となる。これらの値をもとに, まず(5.3)式から全変動Tを求めると,

$$T \div 1.60961 \times 10^{16}$$

となり, 級間変動Bは(5.4)式から,

$$B = 9 \times (\bar{x}_A - \bar{x})^2 + 8 \times (\bar{x}_B - \bar{x})^2 + 8 \times (\bar{x}_C - \bar{x})^2 \div 4.14599 \times 10^{15}$$

となる。残差変動Wは(5.6)式の各変動間の関係から,

$$W = T - B \div 1.19501 \times 10^{16}$$

のように求められる。これらの値から, 級間分散 V_B は(5.7)式を用いて,

$$V_B = B / (3 - 1) \div 2.07299 \times 10^{15}$$

となり, 残差分散 V_W は(5.8)式から,

$$V_W = W / (25 - 3) \div 5.43186 \times 10^{14}$$

となる。

仮説「各群の母平均値は等しい」を検定するには, V_B と V_W から,

$$F_0 = V_B / V_W \div 3.816$$

を求めればよい。 F_0 は自由度 (2, 22) のF分布に従う。自由度 (2, 22) のF分布の上側5%点は数値表から3.443であり, 1%点は5.719であることがわかる。 F_0 は5%点の値より大きい, 1%点の値よりは小さい。したがって, 5%水準で仮説は棄却される。この結果, 少なくとも1つは尿中細菌数の母平均値の異なる群があるものと考えてよいことになる。

【解説 5・1】

ある変数Xについて, k個のグループでそれぞれ独立に測定を行つたとしよう。第i群での標本数を n_i とし, 各データを x_{ij} とする ($i=1, 2, \dots, k, j=1, 2, \dots, n_i$)。各群のデータは, 第

* Statistics in Clinical Studies: V. Methods of the Comparison among Several Groups

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

第2表 多群のデータ

1 群	2 群	...	i 群	...	k 群
x_{11}	x_{21}	...	x_{i1}	...	x_{k1}
x_{12}	x_{22}	...	x_{i2}	...	x_{k2}
$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_{1n_i}	x_{2n_i}	...	x_{in_i}	...	x_{kn_i}

2表で示したように整理することができる。問題は、各群間の母平均値に差があるかを統計的に検定することである。帰無仮説として、「各群の母平均値は等しい」とする。検定の手順は以下のようになればよい。

第 i 群での平均値 $\bar{x}_i$ は、

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \quad (5.1)$$

となり、全体の総平均値 $\bar{x}$ は、

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij} = \frac{1}{N} \sum_{i=1}^k n_i \bar{x}_i \quad (5.2)$$

となる。ただし、 $N = n_1 + n_2 + \dots + n_k$ 、すなわち総標本数である。各データの総平均値からの偏差の2乗和を T とすると、

$$T = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2 \quad (5.3)$$

により求められるが、これは「全変動」と呼ばれる。各群のデータがすべて群の平均値にあるとすれば、その場合の変動は、

$$B = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2 \quad (5.4)$$

により求められる。Bは「級間変動」と呼ばれるものである。また、各群内の変動を足し合わせたものを W とすると、

$$W = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 \quad (5.5)$$

により求められ、 W は「級内変動（残差変動、誤差変動）」と呼ばれる。

各変動の間には、

$$T = B + W \quad (5.6)$$

の関係が成り立つので、どれか2つを求めれば、他の変動は簡単に求められる。

ところで、級間変動は各群の平均値と総平均値の間のばらつきを示すものと考えられ、残差変動は群内のばらつきを示すものと考えられる。したがって、 B と W を用いて、群間の差と残差のどちらが大きいかを調べることができるだろう。

級間変動と残差変動をその自由度で割り、各不偏分散（平均平方和）を求めると、級間分散 V_B は、

$$V_B = B / (k - 1) \quad (5.7)$$

となり、残差分散は、

$$V_W = W / (N - k) \quad (5.8)$$

により推定される。

仮説「各群の母平均値は等しい」を検定するためには、

$$F_0 = V_B / V_W \quad (5.9)$$

を求める。 F_0 が自由度 $(k-1, N-k)$ の F 分布に従うことを用いて検定を行う。すなわち、自由度 $(k-1, N-k)$ の F 分布の上側 α 点の値を $F_\alpha(k-1, N-k)$ としたとき、

$$F_0 > F_\alpha(k-1, N-k) \quad (5.10)$$

ならば、有意水準 α （5%か1%など）で仮説を棄却し、少なくとも1つのグループの母平均値は他と異なるものと考えてよいことになる。有意でなければ、消極的にはあるが仮説を採択し、差があるとはいえないことになる。なお、上記の方法では、どのグループ間に有意差があるかは特定できない。

2. ライアン (Ryan) の方法

例題 5・2

第1表に示した尿中細菌数に関して、分散分析による検定では有意差が認められた。ここでは、同様のデータをもとに特定の2グループ間の母平均値に差があるか否かを検定しよう。

【検定法 5・2】

ここでは、検定全体の有意水準を0.05(5%)とし、「ライアンの方法」を用いて特定の2グループ間の母平均値の差の検定を行う。

すでに求めたように3群の平均値の間には、 $\bar{x}_A > \bar{x}_B > \bar{x}_C$ の順序がある。したがって、まずA群

とC群の母平均値の差を検定する。

(5.11) 式の各数値を代入すると、

$$t_0 = \frac{\bar{x}_A - \bar{x}_C}{\sqrt{V_w(1/9 + 1/8)}} \div 2.403$$

となる。 t_0 は自由度 $25-3=22$ の t 分布に従う。また、この検定の名義的な有意水準は (5.12) 式から、

$$\alpha' = 2 \times 0.05 / \{3 \times (3-1)\} = 0.017$$

となる。自由度22の t 分布について、 t_0 の有意確率を計算すると、0.025 すなわち 2.5% となり、1.7% より大きくなる。したがって、2群の母平均値に差があるとはいえないことになる。

検定結果が有意でなかったため、すべての検定はここで終了し、上記以外のグループの組合せについての検定を行う必要はない。

この結果から、一元配置分散分析により3群の母平均値に少なくとも1つは異なるものがあるとはいえるが、特定の2群間で母平均値に差があるとはいえないことになる。

【解説 5・2】

一元配置分散分析の結果が有意である場合、どの群間の母平均値に有意差があるのかを検討する必要があるだろう。ここでは、各2群間の母平均値の差について検定を行うことで、どのグループ間に有意差があるかを調べる方法、いわゆる「対比較 (対比)」について説明する。

グループ数が k 個ある場合、各2群間の母平均値の差をすべての組合せについて検定すると、 k 個のうち2群を選ぶ組合せの個数は、 $k(k-1)/2$ 個ある。したがって、各検定の有意水準を α とすれば、検定全体での有意水準はその約 $k(k-1)/2$ 倍となってしまう。仮に、グループ数が5の場合でも、組合せの個数は10通りあり、各検定の有意水準を5%としても、全体では、 $1-(1-0.05)^{10} = 0.40$ 、すなわち40%程度の水準となり、極めて大きな有意水準を設定することになってしまう。

上記のような弊害を避けるためには、各検定の水準を組合せの個数で割つたものとすればよい。すなわち、全体の検定水準を α としたければ、個々の検定の水準を $2\alpha / \{k(k-1)\}$ とすればよい (Bonferroni の不等式の利用による方法)。しかし、この基準を用いると、グループ数が大きくなるに

つれ、個々の検定の有意水準は極めて小さくなり、検定結果は有意となりにくくなる。このような有意水準の設定は厳し過ぎるといえよう。ここでは、以下のような手順により群間の母平均値の差の検定を行う。

まず、各群の標本平均を大きさの順に、

$$\bar{x}_1 \geq \bar{x}_2 \geq \dots \geq \bar{x}_i \geq \dots \geq \bar{x}_j \geq \dots \geq \bar{x}_k$$

のように並べる。 i 番目と j 番目の平均値の比較には、それぞれの標本数を n_i, n_j としたとき、

$$t_0 = \frac{\bar{x}_i - \bar{x}_j}{\sqrt{V_w(1/n_i + 1/n_j)}} \quad (5.11)$$

を求める。ただし、 V_w は (5.8) 式による残差分散である。 t_0 が自由度 $(N-k)$ の t 分布に従うことを用いて検定できる。検定全体の有意性を α (5% か 1% など) とする場合、個々の検定の有意水準 α' は、

$$\alpha' = \frac{2\alpha}{k(j-i)} \quad (5.12)$$

と設定する。したがって、自由度 $(N-k)$ の t 分布の両側 α' 点の値 $t_{\alpha'/2(N-k)}$ と t_0 を比べ、 t_0 の方が大きければ有意差ありとすればよい。

実際の計算は、平均値が最大と最小のペアから始め、有意ならば最大の平均値と二番目に最小の平均値の検定を行う。同様な手順で順に差が小さくなるペアに対して検定を繰り返す。最終的に、有意差が認められなくなった時点で、最大の平均値についての検定を終了する。同様な手順で、二番目に最大の平均値と最小のものを検定し、有意ならば二番目に最大の平均値と二番目に最小のもの、のように以後繰り返し検定を行い、有意でない点で検定を終了する。以下同様であるが、このような繰り返しの始めから有意差が認められなければ、すべての検定はそこで終了する。

上記の手順により、全体の有意性は α に保たれるが、この方法は「ライアンの方法」と呼ばれるものである。

3. シェッフエ (Scheffé) の方法

例題 5・3

第1表に示した尿中細菌数について、「シェッフエの方法」により各群の平均値の差を検定してみよう。

【検定法 5・3】

ここでは、各2群間の母平均値の差を検定するだけでなく、特定の群と他の2群をこみにした場合についての検定も行うことにする。

検定全体の有意水準は0.05、すなわち5%とする。自由度(2, 22)のF分布の上側5%点の値は、数値表より3.443である。これを(5.18)式に代入することで、

$$d = \sqrt{(3-1) \times 3.443} \div 2.642$$

であることがわかる。このdは以下のすべての検定で共通に用いられる。

①まず、A群とB群の母平均値の差を検定する。すでに計算した値を用いて平均値の差は、

$$\mu_{A \cdot B} = |\bar{x}_A - \bar{x}_B| \div 2.642 \times 10^7$$

であることがわかる。A群とB群にかける重みはそれぞれ1, -1であるから、dと(5.17)式を用いて、検定のための差の有意限界は、

$$\begin{aligned} d_{A \cdot B} &= d \times \sqrt{V_W \times \{1^2/9 + (-1)^2/8\}} \\ &\div 2.972 \times 10^7 \end{aligned}$$

となる。 $\mu_{A \cdot B}$ は $d_{A \cdot B}$ より小さいので有意とはいえない。

②同様の方法でA群とC群について検定する。

$$\mu_{A \cdot C} = |\bar{x}_A - \bar{x}_C| \div 2.722 \times 10^7$$

が2群の平均値の差であり、有意限界は、

$$\begin{aligned} d_{A \cdot C} &= d \times \sqrt{V_W \times \{1^2/9 + (-1)^2/8\}} \\ &\div 2.972 \times 10^7 \end{aligned}$$

となり、 $\mu_{A \cdot C}$ の方が小さく、有意ではない。

③B群とC群については、

$$\mu_{B \cdot C} = |\bar{x}_B - \bar{x}_C| \div 7.927 \times 10^4$$

$$\begin{aligned} d_{B \cdot C} &= d \times \sqrt{V_W \times \{1^2/8 + (-1)^2/8\}} \\ &\div 3.058 \times 10^7 \end{aligned}$$

となり、 $\mu_{B \cdot C}$ の方が小さく、有意ではない。

④A群とB, C群の2群を一緒にした場合の母平均値の差を検定する。A群には重みとして1を、B, C群には-1/2を重みとする。

$$\mu_{A \cdot BC} = |\bar{x}_A - (\bar{x}_B + \bar{x}_C)/2| \div 2.682 \times 10^7$$

となり、有意限界は、

$$\begin{aligned} d_{A \cdot BC} &= d \times \sqrt{V_W \times \{1^2/9 + (-1/2)^2 \times 2\}} \\ &\div 2.548 \times 10^7 \end{aligned}$$

となる。 $\mu_{A \cdot BC}$ は $d_{A \cdot BC}$ より大きいので、有意差が認められる。

⑤B群とA, C群について、同様に検定する。

$$\mu_{B \cdot AC} = |\bar{x}_B - (\bar{x}_A + \bar{x}_C)/2| \div 1.282 \times 10^7$$

$$\begin{aligned} d_{B \cdot AC} &= d \times \sqrt{V_W \times \{1^2/8 + (-1/2)^2 \times 2\}} \\ &\div 2.624 \times 10^7 \end{aligned}$$

となり、 $\mu_{B \cdot AC}$ は $d_{B \cdot AC}$ より小さいので、有意とはいえない。

⑥最後にC群とA, B群について検定を行う。

$$\mu_{C \cdot AB} = |\bar{x}_C - (\bar{x}_A + \bar{x}_B)/2| \div 1.400 \times 10^7$$

$$\begin{aligned} d_{C \cdot AB} &= d \times \sqrt{V_W \times \{1^2/8 + (-1/2)^2 \times 2\}} \\ &\div 2.624 \times 10^7 \end{aligned}$$

となり、 $\mu_{C \cdot AB}$ は $d_{C \cdot AB}$ より小さいので、有意とはいえない。

上記の結果から、尿中細菌数について急性膀胱炎と他の2群、すなわち急性腎盂腎炎と急性副睾丸炎を一緒にして平均値を比べた場合にのみ有意差が認められることがわかった。各2群間での検定では有意でなくとも、複数のグループをまとめた場合に有意となることもあり、注意が必要である。特に各群の標本数が小さい場合にそうである。なお、検定全体の有意水準は5%となることが保証される。

【解説 5・3】

ライアンの方法は、適当な2群を一对として順に母平均値の差を検定する方法であつた。しかし場合によつては、いくつかのグループの平均に差がなく、それをまとめて他の複数のグループと比較したい場合もあるだろう。そのような場合、すでに説明した対比較ではなく、いくつかの平均値の線形結合(重み付して、加算したもの)により、新たに適当な統計量を求め、それを用いて比較することができる。そのような方法は「線形比較」と呼ばれ、ここでは「シェッフエの方法」を説明する。

k個のグループでの変数Xの母平均値をそれぞれ、 $\mu_1, \mu_2, \dots, \mu_k$ とする。各平均値の線形結合を、

$$\mu = w_1\mu_1 + w_2\mu_2 + \dots + w_k\mu_k \quad (5.13)$$

とする。ただし、 $w_i (i=1, 2, \dots, k)$ は各平均値にかかる重みであり、

$$w_1 + w_2 + \dots + w_k = 0 \quad (5.14)$$

となるように定めるものとする。

μ の期待値 $E(\mu)$ と分散 $V(\mu)$ は、

$$E(\mu) = w_1\bar{x}_1 + w_2\bar{x}_2 + \dots + w_k\bar{x}_k = \mu \quad (5.15)$$

および,

$$V(\mu) = \sigma^2 \left\{ \frac{w_1^2}{n_1} + \dots + \frac{w_k^2}{n_k} \right\} \quad (5.16)$$

となる。ここで, $n_1, \dots, n_k$ は各群の標本数である。 σ^2 は誤差分散であるが, これを (5.8) 式による V_w により推定すると, 分散は,

$$\hat{\sigma} = V_w \sum_{i=1}^k w_i^2 / n_i \quad (5.17)$$

により推定できる。

上記の結果を用いて, μ の $100(1-\alpha)$ 信頼区間を求めることができる。自由度 $(k-1, N-k)$ の F 分布の上側 α 点の値を $F_\alpha(k-1, N-k)$ とし,

$$d = \sqrt{(k-1) \cdot F_\alpha(k-1, N-k)} \quad (5.18)$$

とすると, μ の信頼区間の上側の点を $\hat{\mu}_u$, 下側の点を $\hat{\mu}_L$ とした場合,

$$\hat{\mu}_u = \hat{\mu} + d \cdot \hat{\sigma}, \quad \hat{\mu}_L = \hat{\mu} - d \cdot \hat{\sigma} \quad (5.19)$$

により求められる。

信頼区間 ($\hat{\mu}_L, \hat{\mu}_u$) の間に 0 を含まなければ, 有意水準 α でグループ群の間に差があることになる。実際の計算において, 一方のグループに含まれる群の個数を K , 他方を L とした場合, 比較に用いない群の平均値にかける重みは 0 とし,

前者の重み $= 1/K$, 後者の重み $= -1/L$

のようにグループ内では同一の重みに, 全体では重みの合計が 0 になるように設定すればよい。2 群の母平均値の検定には, それぞれ 1 と -1 の重みを付ければよいことになる。なお, 全体の標本数が大きい場合, (5.18) 式による d は, 自由度 $(k-1)$ の χ^2 分布の上側 α 点の値 $\chi_\alpha^2(k-1)$ を用いて,

$$d = \sqrt{\chi_\alpha^2(k-1)} \quad (5.20)$$

により近似できる。また, 上記のように信頼区間を設定するのは, 検定すべき群の組合せの個数が多い場合, やや面倒である。そのような場合には, $d \cdot \hat{\sigma}$ を求め, $\hat{\mu}$ の方が大きければ有意差ありとした方が計算しやすいだろう。すなわち,

$$|\hat{\mu}| > d \cdot \hat{\sigma} \quad (5.21)$$

のとき有意差ありとすればよい。

4. クルスカル-ウォリス (Kruskal-Wallis) 検定

例題 5・4

慢性前立腺炎患者に無作為に, A, B および C 薬をそれぞれ 25 例, 26 例, 25 例に投与し経過を観察したところ, 第 3 表のようなデータが得られた。このデータから, 3 薬の薬効に差があるといえるだろうか。

第 3 表 3 薬の慢性前立腺炎への効果

群	著効	有効	不変	悪化	計
A 薬	8	10	6	1	25
B 薬	3	9	10	4	26
C 薬	3	8	11	3	25
計	14	27	27	8	76

【検定法 5・4】

薬効は表に示したように, 4 カテゴリに分けられている。ここではクルスカル-ウォリス検定により, 薬効差を調べることにする。

各カテゴリには, 「著効」に 7.5, 「有効」に 28, 「不変」に 55, 「悪化」に 72.5 の中間順位が与えられる (本講座 I または II を参照)。

A 薬群の順位和 R_A は,

$$R_A = 8 \times 7.5 + 10 \times 28 + 6 \times 55 + 1 \times 72.5 \\ = 742.5$$

となり, 平均順位 $\bar{R}_A$ は,

$$\bar{R}_A = R_A / 25 = 29.7$$

となる。同様に, B 薬群の順位和 R_B と平均順位 $\bar{R}_B$ は,

$$R_B = 1114.5, \quad \bar{R}_B = 42.865$$

となり, C 薬群の順位和 R_C と平均順位 $\bar{R}_C$ は,

$$R_C = 1069, \quad \bar{R}_C = 42.76$$

となる。また, (5.23) 式から全体の平均順位 $\bar{R}$ は,

$$\bar{R} = (76 + 1) / 2 = 38.5$$

と求められる。

これらの数値から検定が行えるが, データにはタイがあるので, (5.25) 式からタイのある場合の修正項を求めておく。

$$D = 1 - \frac{(14^3 - 14) + (27^3 - 27) \times 2 + (8^3 - 8)}{76^3 - 76}$$

$$\div 0.903$$

となる。(5.24) 式から K_0 を求めると、

$$K_0 = \frac{12}{76 \times 77} \times \{25 \times (\bar{R}_A - \bar{R})^2 + 26 \times (\bar{R}_B - \bar{R})^2 + 25 \times (\bar{R}_C - \bar{R})^2\} / D$$

$$\div 6.551$$

と求められる。 K_0 は自由度 $3-1=2$ の χ^2 分布に従うものとみなせる。自由度 2 の χ^2 分布の上側 5% 点の値は 5.991, 1% 点は 9.210 である。 K_0 は 5% 点の値より大きい、1% 点の値より小さい。したがって、5% 水準で有意差が認められる。

上記の結果から、少なくとも 1 つの薬剤の効果は他と異なるものと考えてよいことになる。

【解説 5・4】

特定の変数のデータが順位で与えられている場合を考えよう。グループ数を k とし、第 i 群での標本数を n_i , 順位データを $R_{i1}, R_{i2}, \dots, R_{in_i}$ とする。ただし、順位はグループ内でのものではなく、全体での順位を与えるものとする。また、順位にタイがある場合には、その中間順位を与えるものとする(本講座 I, II 参照)。ここでは、全体で t 種類のタイがあり、各タイにはそれぞれ $d_1, d_2, \dots, d_t$ 個の同順位のデータが含まれるものとする。

第 i 群の順位和を R_i とし、順位の平均を $\bar{R}_i$ とすると、

$$\bar{R}_i = (R_{i1} + R_{i2} + \dots + R_{in_i}) / n_i = R_i / n_i \quad (5.22)$$

となる。また、全体での平均値は全標本数を N としたとき、

$$\bar{R} = (N+1)/2 \quad (5.23)$$

となる。

群間に差があるか否かを調べるための統計量 K_0 は、

$$K_0 = \left\{ \frac{12}{N(N+1)} \cdot \sum_{i=1}^k n_i (\bar{R}_i - \bar{R})^2 \right\} / D$$

$$= \left\{ \frac{12}{N(N+1)} \cdot \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1) \right\} / D \quad (5.24)$$

により求められる。ここで、

$$D = 1 - \sum_{i=1}^k (d_i^3 - d_i) / (N^3 - N) \quad (5.25)$$

である。タイがなければ、 $D=1$ である。

適当な定数 c について、

$$K_0 \geq c$$

ならば、有意差ありとすればよい。しかし、検定のための一般的な数値表を、さまざまなグループ数や各群の標本数に応じて使用できるように作成するのは困難である(グループ数や標本数が小さい場合のものは作成されている)。検定のためには、 K_0 が自由度 $(k-1)$ の χ^2 分布に従うことを利用すればよい。すなわち、自由度 $(k-1)$ の χ^2 分布の上側 α (5% か 1%) 点の値 $\chi_{\alpha}^2(k-1)$ と K_0 を比べ、 K_0 が大きければ有意差ありとすればよい。 χ^2 分布による近似は、グループ数が 3 以上で、各群の標本数が 5 以上ならばかなりよいようである。

5. 順位和による多重比較

例題 5・5

例題 5・4 でみたように、慢性前立腺炎に対する 3 薬の効果には有意差が認められた。ここではどの薬群間に差があるのかを検定しよう。

【検定法 5・5】

ここでは、「シェッフエの方法」により順位和のデータをもとに多重比較を行う。

例題 5・4 で求めた各薬群の順位の平均 $\bar{R}_A, \bar{R}_B, \bar{R}_C$ とタイのある場合の修正項 D を用いて、(5.26) 式により検定に必要な計算を行う。

各検定に必要な計算のうち、共通する部分の計算をまず行うことにする。それは (5.26) 式の分母の一部である。すなわち、

$$V = D \times N \times (N+1) / 12 \div 440.393$$

である。なお、以下の方法により、検定全体の有意水準は 5% となる。

① A 薬群と B 薬群の薬効差について検定する。

検定のためには、A 薬群と B 薬群の重みを 1 と -1 とし計算すればよい。

$$\chi_{A \cdot B}^2 = \frac{(\bar{R}_A - \bar{R}_B)^2}{V \times \{1^2/25 + (-1)^2/26\}} \div 5.061$$

となる。この値と自由度2の χ^2 分布の上側5%点の値5.991と比べると、 $\chi^2_{A \cdot B}$ の方が小さいので有意とはいえない。

②A薬群とC薬群については、

$$\chi^2_{A \cdot B} = \frac{(\bar{R}_A - \bar{R}_C)^2}{V \times \{1^2/25 + (-1)^2/25\}} \div 4.841$$

となり、5.991より小さいので有意ではない。

③B薬群とC薬群については、

$$\chi^2_{B \cdot C} = \frac{(\bar{R}_B - \bar{R}_C)^2}{V \times \{1^2/26 + (-1)^2/25\}} \div 0.000$$

となり、5.991より小さいので有意ではない。

④A薬群とB、C薬群をこみにした場合についての差を検定する。A薬群には1の重みを、B、C薬群にはそれぞれ $-1/2$ の重みをかけることになる。

$$\chi^2_{A \cdot BC} = \frac{\{\bar{R}_A - 0.5 \times (\bar{R}_B + \bar{R}_C)\}^2}{V \times \{1^2/25 + (-1/2)^2 \times (1/26 + 1/25)\}} \div 6.549$$

となる。この値は、5.991より大きいので有意である。

⑤B薬群とA、C薬群についても同様に、

$$\chi^2_{B \cdot AC} = \frac{\{\bar{R}_B - 0.5 \times (\bar{R}_A + \bar{R}_C)\}^2}{V \times \{1^2/26 + (-1/2)^2 \times (1/25 + 1/25)\}} \div 1.710$$

となり、5.991より小さいので有意ではない。

⑥C薬群とA、B薬群についても同様に、

$$\chi^2_{C \cdot AB} = \frac{\{\bar{R}_C - 0.5 \times (\bar{R}_A + \bar{R}_B)\}^2}{V \times \{1^2/25 + (-1/2)^2 \times (1/25 + 1/26)\}} \div 1.598$$

となり、5.991より小さいので有意ではない。

上記の結果から、各2薬群間に有意差は認められないが、B薬群とC薬群をこみにした場合、A薬群と有意差が認められた。したがって、A薬が3薬の中では最も効果があるものと考えてよいだろう。また、これは①②で求めた χ^2 値が有意ではないが、かなり大きな値となつていることとも符合している。

【解説 5・5】

クルスカール-ウォリスの検定結果が有意であつた場合、次の問題としてどのグループ間に差が認

められるかを調べる必要があるだろう。ここでは、すでに説明した「シェッフェの方法」により、順位和を用いた一般的な対比の方法について説明する。用いる記号などはすべて前節と同様である。

グループ数を k 、第 i 群の標本数を n_i とし、全データについて順位づけした場合の第 i 群における順位和を R_i 、その平均を $\bar{R}_i$ とする。各平均値に対する重みを、3節と同様に考え、以下の統計量を計算する。

$$\chi^2_0 = \frac{\left(\sum_{i=1}^k w_i \cdot \bar{R}_i \right)^2}{\left\{ \frac{D \cdot N(N+1)}{12} \cdot \sum_{i=1}^k \frac{w_i^2}{n_i} \right\}} \quad (5.26)$$

を求める。ここで、 D は(5.25)式で与えられるタイのある場合の修正項、 N は総標本数である。

ただし、 $\sum_{i=1}^k w_i = 0$ とする。

χ^2_0 が自由度 $(k-1)$ の χ^2 分布に従うことにより検定を行う。すなわち、自由度 $(k-1)$ の χ^2 分布の上側 α 点を $\chi^2_{\alpha^2}(k-1)$ としたとき、

$$\chi^2_0 \geq \chi^2_{\alpha^2}(k-1)$$

ならば、グループ(群)間に有意差ありとする。これにより、検定全体での有意水準は α となる。2つのグループを比較する場合、一方の重みを1とし、他方を -1 、比較に用いないものの重みは0とすればよい。基本的な考え方は3節と同様であり、そちらを参照して欲しい。

多重比較の方法について、紙面の都合もあり、すべての方法について説明を行うことはできない。ここで取り上げた方法や他の方法、理論についてより詳しく知りたい読者は、参考文献1~4を参照すればよいだろう。

参考文献

- 1) 豊川裕之、柳井晴夫編著：医学・保健学の例題による統計学，現代数学社，1982。
- 2) E.L. レーマン原著，鍋谷清治他訳：ノンパラメトリックス，森北出版，1978 (E.L. Lehmann: Nonparametrics, Holden-Day, 1975)
- 3) 佐久間昭：薬効評価一計画と解析Ⅰ，東京大学出版会，1977。
- 4) 佐久間昭：薬効評価一計画と解析Ⅱ，東京大学出版会，1981。