

分散分析と多重比較：資料

芳賀昭弘

1 はじめに

被験者全員に同じ CT 装置を使って治療前後の病変の大きさを調べた。母集団のデータは正規分布に従っているものと仮定する。治療前後の病変の大きさに違いがあるかどうかを調べるには、どのような検定が必要でしょう？

被験者を 2 群にわけて、第 1 群は CT 装置で、第 2 群は MRI 装置で病変の大きさを調べた。それぞれの群の母集団は正規分布に従うとする。装置の違いで検知される病変の大きさが異なるかどうかを調べるには、どのような検定が必要でしょうか？

ここまでは、2 群比較のための検定を行ってきました。ですが、上の 1 つ目の例で言えば、治療前、治療直後、治療後 1 ヶ月、、、（2 つ目の例でも、例えば MRI のシーケンスプロトコルを複数用意したり、PET や超音波などを含めた場合）のように 3 群以上の標本の比較を行いたい場合があります。3 群以上の比較を行いたいときに使う統計的方法として、分散分析（Analysis of variance: ANOVA）と呼ばれるものがあります*1。

分散分析において有意差が出た（＝どれか 2 群の間に差が出たということ）場合、次にどの群に差が出たのかを調べたいと思うのが常です。群間の比較を調べるには多重比較法を使います（なお、分散分析と絡めずに最初から多重比較法だけで行うほうが良いと主張する研究者も多くいますが、ここでは両方を学んでおきたいと思います）。

分散分析や多重検定を手計算で行うのは大変です。しかし、コンピュータではたった一行でそれを計算してくれるコマンドが存在します。現代のコンピュータ事情なら、こうした統計解析も非常に簡単にできます！ですが、計算結果が簡単に出来ることも大事ですが、中身の理解の方がもっと重要です。計算結果が出たとしても、分析方法が誤っていたら時間の無駄をしていることになるだけでなく、誤った解釈を導き出してそれを公表することで社会に大きな不利益を与えることにもなりかねません。知りたいことは何か、目的をはっきりさせて、その目的のために必要な分析方法を見つけるには、その分析方法をきちんと理解しておかなければならないのは言うまでもないでしょう。

*1 3 群以上の検定の場合でも、2 群の t 検定を 3 回行うことで同じようなことができる人があるかもしれませんが、これはやってはいけません！3 群以上のグループ比較を行いたいのであれば、必ず分散分析を行ってください（2 群の検定を 3 つの組み合わせで行うことを考えてみよう。3 つの検定それぞれが 95% となる組み合わせの実現確率は確率の乗法定理より $0.95 \times 0.95 \times 0.95 \sim 0.86$ となってしまいますね。

表 1. 施肥量と収量の関係

収量 (kg/a)	施肥量 (kg/a)			
	水準 1 (0)	水準 2 (10)	水準 3 (20)	水準 4 (30)
第 1 区画	9	13	17	16
第 2 区画	11	12	16	17
第 3 区画	10	14	15	18
平均	10	13	16	17

総平均 : 14

2 一元配置分散分析（対応なし）

表 1 をみよう。このデータには「要因による変動」と「誤差による変動」が含まれる。知りたいのは「要因による変動」が本当にあるのかどうか（施肥量が収量に影響するかどうか）。このために、誤差の変動に比べ、確かに要因による変動が大きいかどうかを調べる。

→ 誤差変動と要因変動の分散比により F 値を求めて有意差検定を実施する。

分散分析の検定：

$$F = \frac{\text{目的となる要因効果（群間変動）の分散}}{\text{誤差効果（群内変動）の分散}} \quad (1)$$

$F > 1$ の場合、要因変動 > 誤差変動

$F \leq 1$ の場合、要因変動 \leq 誤差変動

総変動（各値-総平均）

水準 1	水準 2	水準 3	水準 4
-5	-1	3	2
-3	-2	2	3
-4	0	1	4

→

群間変動（群平均-総平均）

水準 1	水準 2	水準 3	水準 4
-4	-1	2	3
-4	-1	2	3
-4	-1	2	3

+

群内変動（各値-群平均）

水準 1	水準 2	水準 3	水準 4
-1	0	1	-1
1	-1	0	0
0	1	-1	1

総変動の偏差平方				群間変動の偏差平方				群内変動の偏差平方			
水準 1	水準 2	水準 3	水準 4	水準 1	水準 2	水準 3	水準 4	水準 1	水準 2	水準 3	水準 4
25	1	9	4	16	1	4	9	1	0	1	1
9	4	4	9	16	1	4	9	1	1	0	0
16	0	1	16	16	1	4	9	0	1	1	1
偏差平方和 $S = 98$				$S_R = 90$				$S_E = 8$			
自由度 $nk - 1 = 11$				$k - 1 = 3$				$k(n - 1) = 8$			
分散 $V = 8.91$				$V_R = 30$				$V_E = 1$			

よって $F = \frac{V_R}{V_E} = 30$

まとめ：手順

1. 仮説の設定

帰無仮説 H_0 ：要因効果と誤差効果との間に差はない（要因効果なし）

対立仮説 H_1 ：要因効果と誤差効果との間に差はある（要因効果あり）

2. 検定統計量の選定

F 値

3. 有意水準（棄却域）の決定

有意水準 5%（上の例題では分子・分母それぞれの自由度は 3 と 8、F 分布表から有意水準 5% の境界値は

$F = 4.07$ ）

4. 帰無仮説の棄却 or 採択（保留）

F 値の実現値は 30 であったので、境界値より十分に大きい。帰無仮説は棄却され、要因効果があると考えら

れる。

表 2. 分散分析表

変動要因	平方和	自由度	分散	分散比	P 値	F 境界値
群間	90	3	30	30	0.0001	4.07
群内	8	8	1			
計	98	11				

一元配置分散分析の対応ありの場合や多元配置などもありますが、さらに計算が面倒になります。まずは上の一元配置分散分析（対応なし）をしっかりと理解して、教科書の多元配置の節を読んだりコンピュータで分析したりする、という学習が効果的でしょう。

3 多重比較法

上の分散分析では、要因に差があるという結論は出のですが、どの要因とどの要因に差が出たのか（またはその複数の組み合わせで差が出たのか）、わかりません。そこで多重比較によってさらに差が出た要因について調べてみます。前節の例では帰無仮説は、

帰無仮説 H_{12} ：水準 1 = 水準 2

帰無仮説 H_{13} ：水準 1 = 水準 3

帰無仮説 H_{14} ：水準 1 = 水準 4

帰無仮説 H_{23} ：水準 2 = 水準 3

帰無仮説 H_{24} ：水準 2 = 水準 4

帰無仮説 H_{34} : 水準 3 = 水準 4

です。それぞれは 2 群の比較なのですが、上で述べたように t 検定では検定結果が甘く出てきてしまいますので使えません。そこで、検定量を次の q 値に置き換えたもので検定を実施する方法があります；

$$q_{\bar{x}_2 - \bar{x}_1} = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{\hat{\sigma}_e^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (2)$$

ここで \bar{x} は平均値、 $\hat{\sigma}_e^2$ は「誤差変動（群内変動）」の分散。q 分布表から限界値を読んでこの値と比較することで検定を行います。この検定を Tukey の方法と言います。その他にも沢山の検定方法が提案されており、どの検定を使うべきかを判断するためには、やはりきちんとその方法を理解しなければなりません。