合理进行均值比较——单组和配对设计定量资料多元方差分析

胡纯严[1]，胡良平[1,2*]

（1. 军事科学院研究生院，北京 100850；
2. 世界中医药学会联合会临床科研统计学专业委员会，北京 100029
*通信作者：胡良平，E-mail：lphu927@163.com）

【摘要】 本文目的是介绍与单组设计和配对设计定量资料多元方差分析有关的基本概念、计算方法、两个医学实例以及SAS实现。基本概念包括单组设计与配对设计、均值向量、多元方差分析和前提条件；计算方法涉及一般检验统计量和Hotelling's T^2检验统计量；两个医学实例分别为“结核病患者营养状况的调查资料”和“石杉碱甲治疗增龄相关记忆障碍效果的试验资料”。借助SAS对两个医学实例中的定量资料分别进行一元和多元差异性分析，并对这两类分析方法的区别进行讨论。

【关键词】 单组设计；配对设计；均值向量；方差协方差矩阵；多元方差分析
中图分类号：R195.1 文献标识码：Ad doi：10.11886/scjsws20230319001

Reasonably carry out mean value comparison: MANOVA of the quantitative data collected from the single group design and the paired design

Hu Chunyan[1], *Hu Liangping*[1,2*]

(1. *Graduate School, Academy of Military Sciences PLA China, Beijing* 100850, *China*;
2. *Specialty Committee of Clinical Scientific Research Statistics of World Federation of Chinese Medicine Societies, Beijing* 100029, *China*
*Corresponding author: Hu Liangping, E-mail: lphu927@163.com)

【Abstract】 The purpose of this article was to introduce the basic concepts, calculation methods, two medical examples and SAS implementation related to the multivariate analysis of variance (MANOVA) for quantitative data with the single group design and the paired design. Basic concepts included the single group design and the paired design, mean vector, MANOVA and preconditions, calculation methods involved the general test statistics and Hotelling's T^2 test statistics, two medical examples were survey data on nutritional status of tuberculosis patients and trial data on the efficacy of huperzine A in the treatment of age-related memory impairment. With the help of SAS software, the univariate and multivariate difference analysis of quantitative data in two medical cases were carried out, and the differences between these two types of analysis approaches were discussed.

【Keywords】 Single group design; Paired design; Mean vector; Variance-covariance matrix; Multivariate analysis of variance

在医学研究中，由于所研究问题的复杂性，研究者不仅要考虑多个影响因素，还要观测多个定量指标的取值，并希望采用多元统计分析方法将多个定量指标同时纳入统计分析中。在多元统计分析中[1-2]，最简单的统计分析方法是多元方差分析，它是一元定量资料t检验或方差分析的推广。对于不同设计类型下收集的多元定量资料，需要采用相应的多元方差分析方法。本文将介绍2种最简单的设计类型（即单组设计和配对设计）下，多元定量资料的实例、多元方差分析的计算方法以及SAS实现。

1 基本概念

1.1 单组设计与配对设计

单组设计是指对一组符合研究目的的受试对象未按任何其他因素分组，仅在试验因素的某特定水平下观测m个定量指标的数值。若m=1，其定量资料称为单组设计一元定量资料；若m≥2，其定量资料则称为单组设计m元定量资料。

配对设计有三种情形：第一种情形为自身配对设计，就是对一组符合研究目的的受试对象在两个不同时间点或两个不同的试验条件下，对m个定量指标进行两次重复观测；第二种情形为同源配对设计，就是选择符合研究目的且每对来源相同的多组受试对象，将每对中的2个个体随机分配进入2个试验组，对m个定量指标进行观测；第三种情形，称为条件相近者配对设计，先确定配对条件（例如性别、年龄、血型、身高、体重等），选择符合研究目的且数量尽可能多的受试对象，将他们严格按事先确定的配对条件形成多个配对组，每个配对组中包含2个条件最接近的受试对象，对m个定量指标进行

观测。当m=1或m≥2时，所获得的定量资料分别称为配对设计一元或m元定量资料。

1.2 均值向量

将取自单组设计的m个定量指标分别求算术平均值（简称均值），再将这些均值按一定顺序排列起来，就称它们为样本均值向量，其表达见式（1）和式（2）。

$$\bar{\mathrm{X}}=\begin{bmatrix}\bar{x}_1\\ \vdots\\ \bar{x}_m\end{bmatrix} \tag{1}$$

$$\bar{\mathrm{X}}=\left[\bar{x}_1,\cdots,\bar{x}_m\right]' \tag{2}$$

式（1）和式（2）表达的向量分别为列向量和行向量，其中，式（2）中的“′”为转置运算（简称转置），就是将“列”转变成“行”。同理，可写出与式（1）和式（2）对应的总体均值向量，分别见式（3）和式（4）。

$$\mu=\begin{bmatrix}\mu_1\\ \vdots\\ \mu_m\end{bmatrix} \tag{3}$$

$$\mu=\left[\mu_1,\cdots,\mu_m\right]' \tag{4}$$

在对单组设计多元定量资料进行差异性分析时，需要为总体均值向量赋予一组确定的数值；而在对配对设计多元定量资料进行差异性分析时，由于要算出各定量指标在同一配对组中的差量，故在配对因素2个水平作用相同的假设之下，全部差量的总体均值向量为零向量，即向量的各分量都是0。

1.3 多元方差分析

在对一元定量资料进行差异性检验时，人们常用基于正态分布的Z检验、基于t分布的t检验和基于F分布的F检验（或称为一元方差分析，简称为方差分析）[3-4]；同理，在对多元定量资料进行差异性检验时，统计学家基于数学原理进行推导，得出了推广的t检验和方差分析，被称为T^2检验和Wilks′λ检验[5-6]。这两种用于多元方差分析的检验统计量与一元方差分析中的F检验统计量之间存在精确的数量关系，即可以基于所计算出来的T^2检验或Wilks′λ检验统计量的数值直接推算出F检验统计量的数值。于是，可以直接基于F分布来作出统计推断。

1.4 前提条件

采用Z检验、t检验、F检验时，所分析的一元定量资料应满足的前提条件是：独立性（即任何2个数据之间互相独立）、正态性（各组定量资料服从正态分布）和方差齐性（各组定量资料的方差相等）；而采用T^2检验和Wilks′λ检验时，独立性的含义与前述相同，但正态性应由“单变量正态分布”调整为“多元正态分布”，方差齐性应由“单变量在多组之间的方差相等”调整为“多变量在多组之间的方差-协方差矩阵相等”。

2 计算方法

2.1 将配对设计转变成单组设计

设样本含量为n，定量指标的个数为m，定量指标的变量名为X_{ijk}，其中，下标i代表第i个定量变量，j代表第j组（或试验因素的第j个水平，单组设计时，j始终取值1，配对设计时，j=1、2），k代表第k个样本（k=1，2，…，n）。配对设计时，用每对中同一个定量指标数值的差量作为一个新定量指标，并把它视为单组设计时的原始定量指标。于是，就可把配对设计定量资料转变成单组设计定量资料了。因此，所有的统计表达、描述和统计分析，都可以将单组设计与配对设计视为一种情形，故下面仅介绍单组设计下的各种做法。在分析配对设计定量资料时，通常假定配对因素的两个水平对定量指标的影响相同，故每个定量指标差量的理论均值为0，全部m个定量指标差量的理论均值向量为0向量。

2.2 单组设计多元定量资料的一般统计量

一般统计量包括样本均值向量、样本离均差平方和矩阵、样本方差和协方差矩阵，样本均值向量定义式见式（5）。

$$\bar{\mathrm{X}}=\left(\bar{x}_{11},\bar{x}_{21},\cdots,\bar{x}_{m1}\right)' \tag{5}$$

式（5）中，第i个元素的定义见式（6）。

$$\bar{x}_{i1}=\frac{1}{n}\sum_{k=1}^{n}x_{i1k},\quad i=1,2,\cdots,m \tag{6}$$

样本离均差平方和矩阵的定义见式（7）。

$$L=\begin{bmatrix}l_{11} & \cdots & l_{1m}\\ \vdots & \vdots & \vdots\\ l_{m1} & \cdots & l_{mm}\end{bmatrix} \tag{7}$$

式（7）中，第(i,t)个元素的定义见式（8）。

$$l_{it}=\sum_{k=1}^{n}\left(x_{i1k}-\bar{x}_{i1}\right)\left(x_{t1k}-\bar{x}_{t1}\right),\ i,t=1,2,\cdots,m \tag{8}$$

样本方差和协方差矩阵的定义见式（9）。

$$S=\frac{1}{n-1}L \tag{9}$$

式（9）中，等号右边L的计算见式（7）。

2.3 单组设计多元定量资料的检验统计量

Hotelling[7]于1931年基于学生比的推广导出

了 T^2 检验统计量，称为 Hotelling's T^2 检验统计量，见式(10)。

$$T^2 = n\left(\bar{X} - \mu_0\right)' S^{-1}\left(\bar{X} - \mu_0\right) \quad (10)$$

式(10)中，$\bar{X}$、S 分别见上文式(5)和式(9)，μ_0 为给定的均值向量。式(10)定义的 T^2 检验统计量可以经过线性变换转变成 F 检验统计量[1,5]，见式(11)。

$$F = \left[\frac{(n-m)}{m(n-1)}\right]T^2 \quad (11)$$

式(11)中，n 为样本含量，m 为定量指标的个数，T^2 由式(10)定义。当满足前面提及的前提条件时，式(11)定义的 F 检验统计量近似服从 F 分布，见式(12)。

$$F \sim F(m, n-m) \quad (12)$$

式(12)中，m 和 $n-m$ 分别为服从 F 分布的随机变量的分子和分母上的自由度。

在实际应用中，通常是把 T^2 值转变为 F 值，再利用 F 检验统计量解决均值向量的假设检验问题。在 SAS 的输出结果中[8]，输出的检验统计量有 Wilks' Lambda、Pillai's Trace、Hotelling-Lawley Trace 和 Roy's Greatest Root。其中，Hotelling-Lawley Trace 的计算结果就是 T^2 值。

3 实例与 SAS 实现

3.1 问题与数据结构

3.1.1 两个临床试验问题及数据

【例 1】为探讨结核病患者的营养状况，某研究者对 2000 年收治的 45 例患者的营养状况进行调查，调查的指标包括白蛋白(ALB)和白蛋白/球蛋白的比值(A/G)，测定结果见表 1。已知白蛋白指标的下限为 35g/L，白蛋白/球蛋白指标的下限为 1.5。假定资料满足参数检验的前提条件，试评价结核病患者的营养状况[9]。

表 1 45 例结核病患者营养状况的两项定量指标测定结果

Table 1 Measurement results of two quantitative indicators of the nutritional status of 45 tuberculosis patients

患者编号	ALB	A/G	患者编号	ALB	A/G	患者编号	ALB	A/G
1	30.252 7	0.933 2	16	29.404 6	0.495 3	31	37.140 7	0.497 6
2	37.246 1	1.154 6	17	31.718 5	1.329 0	32	24.205 6	1.068 5
3	31.933 2	1.115 4	18	26.307 0	0.954 5	33	34.101 5	0.699 2
4	38.297 0	0.827 0	19	42.204 1	1.058 0	34	36.484 8	0.874 5
5	29.725 8	1.464 2	20	28.229 0	0.814 2	35	41.332 2	1.034 1
6	40.249 2	1.173 8	21	32.003 5	0.959 5	36	28.344 9	1.121 0
7	29.651 8	1.767 0	22	33.497 6	0.814 1	37	41.350 2	1.172 3
8	33.929 0	0.791 5	23	28.757 5	0.933 1	38	47.174 0	0.570 4
9	35.898 2	1.472 6	24	23.632 9	1.365 9	39	25.109 1	0.933 1
10	33.279 2	1.237 6	25	32.345 2	0.753 5	40	30.599 0	0.919 6
11	33.077 7	0.394 9	26	35.258 7	0.743 0	41	33.374 3	0.650 1
12	23.764 1	0.838 8	27	33.179 9	1.312 8	42	28.507 2	0.679 6
13	41.596 8	1.116 0	28	34.169 9	0.865 8	43	34.818 8	0.951 4
14	39.083 6	1.276 9	29	31.801 5	1.252 7	44	29.337 4	0.654 9
15	33.135 3	1.250 9	30	31.540 3	0.928 5	45	33.986 3	1.073 3

注：ALB 代表白蛋白，A/G 代表白蛋白/球蛋白的比值

【例 2】研究石杉碱甲治疗增龄相关记忆障碍的效果，选取 15 例增龄相关记忆障碍患者，年龄 60~80 岁，平均 72 岁，男性 9 例，女性 6 例，小学以上受教育程度，缓慢发生部分记忆减退 3 年以上，并按统一标准入选。其治疗前后长期记忆功能评分(包括“1~100 背数评分”“100~1 背数评分”和“1~19 累加评分”3 个定量指标)见表 2。假定资料满足参数检验的前提条件，试分析治疗前后各项长期记忆功能评分差异是否有统计学意义[9]。

3.1.2 对数据结构的分析

例 1 中样本含量 n=45，他们未按其他任何变量进行分组，属于同一个试验组，故属于单组设计；定量指标有 2 个，分别为白蛋白、白蛋白与球蛋白的比值，各定量指标均有下限值。由于这两个定量指标在临床上具有一定的关联性，需要将它们视为一个整体进行分析，故这是一个二元定量资料差异性分析问题。

例 2 中样本含量 n=15，以治疗前后作为配对条

件，即采取自身配对，在治疗前后，分别收集每位受试对象3个定量指标的取值，故这是一个自身配对设计三元定量资料差异性分析问题。评分越高，表明记忆力越好。若治疗后的评分大于治疗前的评分，表明治疗有效。

表2 石杉碱甲对增龄相关记忆障碍患者治疗前后长期记忆功能评分的测定结果

Table 2 Measurement results of huperzine A on long-term memory function score of patients with age-related memory impairment before and after treatment

患者编号	1~100背数评分		100~1背数评分		1~19累加评分	
	治疗前	治疗后	治疗前	治疗后	治疗前	治疗后
1	7.1	10.3	9.3	12.9	5.7	10.4
2	8.1	12.4	12.0	12.1	7.0	9.4
3	6.2	12.3	10.7	16.1	7.9	12.9
4	6.7	8.7	5.9	9.2	2.5	16.6
5	6.8	11.0	7.5	8.6	7.6	11.1
6	7.0	11.4	10.8	20.1	6.8	13.7
7	5.3	10.1	10.4	14.8	4.5	10.3
8	8.1	10.9	7.4	9.3	8.4	12.2
9	5.9	12.0	5.7	8.5	4.3	8.9
10	7.2	15.8	7.8	13.8	7.1	11.0
11	7.2	9.1	9.1	11.2	5.6	7.8
12	6.1	10.1	8.0	16.2	4.8	10.0
13	7.4	10.5	8.6	9.4	6.2	9.3
14	6.9	9.0	10.4	15.1	4.9	12.5
15	6.3	17.2	6.4	8.9	3.3	11.5

3.1.3 创建SAS数据集

分析例1资料，设所需要的SAS数据步（详细数据见表1）程序如下：

```
data a1;
input id x1 x2 @@;
y1=x1-35; y2=x2-1.5;
cards;
1 30.2527 0.9332 16 29.4046 0.4953 31 37.1407 0.4976
2 37.2461 1.1546 17 31.7185 1.3290 32 24.2056 1.0685
……
14 39.0836 1.2769 29 31.8015 1.2527 44 29.3374 0.6549
15 33.1353 1.2509 30 31.5403 0.9285 45 33.9863 1.0733
;
run;
```

【变量说明】id代表受试者编号，x1代表白蛋白，x2代表白蛋白/球蛋白的比值，y1代表x1－35之后的结果，y2代表x2－1.5之后的结果。

分析例2资料，设所需要的SAS数据步（详细数据见表2）程序如下：

```
data a2;
input id x1 x2 y1 y2 z1 z2;
xd=x2-x1; yd=y2-y1; zd=z2-z1;
cards;
1 7.1 10.3 9.3 12.9 5.7 10.4
2 8.1 12.4 12.0 12.1 7.0 9.4
… … … … … … …
14 6.9 9.0 10.4 15.1 4.9 12.5
15 6.3 17.2 6.4 8.9 3.3 11.5
;
run;
```

【变量说明】id代表受试者编号，(x1 x2)、(y1 y2)、(z1 z2)分别代表3个定量指标治疗前和治疗后的测定结果；xd、yd、zd分别是3个定量指标治疗后与治疗前测定结果的差值。

3.2 用SAS实现统计分析

3.2.1 分析例1中的资料

进行一元统计分析所需要的SAS过程步程序如下：

```
proc glm data=a1;
model y1-y2=/ss3;
run; quit;
```

【SAS输出结果及解释】对变量y1而言，对应的输出结果为：t=−2.400，P=0.021；对变量y2而言，对应的输出结果为：t=−12.080，P<0.001。由于两个t统计量的数值都是负值，说明x1（白蛋白）的均值小于35 g/L，x2（白蛋白/球蛋白的比值）的均值小于1.5。

进行二元统计分析所需要的SAS过程步程序如下：

```
proc glm data=a1;
model y1-y2=/nouni ss3;
manova h=intercept;
run; quit;
```

【SAS输出结果及解释】T^2=0.779，F=75.820，分子和分母的自由度分别为2和43，P<0.001，说明由两个定量指标组成的均值向量与给定的均值向量[35, 1.5]′之间差异有统计学意义。

在本例中，一元差异性检验和二元差异性检验的结果是完全一致的，x1（白蛋白）的均值小于35 g/L，x2（白蛋白/球蛋白的比值）的均值小于1.5，说明结核病患者的营养状况不好，因为所考查的两项重要定量指标的均值尚未超过临床上规定的最低界限值。

3.2.2 分析例2中的资料

进行一元统计分析所需要的SAS过程步程序如下：

```
proc glm data=a2;
model xd yd zd=/ss3;
run; quit;
```

【SAS输出结果及解释】对变量xd而言，对应的输出结果为：t=7.030，P<0.001；对变量yd而言，对应的输出结果为：t=5.500，P<0.001；对变量zd而言，对应的输出结果为：t=7.000，P<0.001。由于3个t统计量的数值均为正值，说明治疗后评分的均值大于治疗前评分的均值，即治疗有效。

进行三元统计分析所需要的SAS过程步程序如下：

```
proc glm data=a2;
model xd yd zd=/nouni ss3;
manova h=intercept;
run; quit;
```

【SAS输出结果及解释】T^2=7.637，F=30.550，分子和分母的自由度分别为3和12，P<0.001，说明由3个定量指标差量组成的均值向量与假定的均值向量$[0,0,0]'$之间差异有统计学意义。

在本例中，一元差异性检验和三元差异性检验的结果是完全一致的，3个定量指标差量的均值都大于0，表明治疗是有效的。

4 讨论与小结

4.1 讨论

本文中两个实例的分析结果都表明，单变量分析结果与多变量分析结果是相同的，所得到的结论是完全一致的，但这种情况并非总是如此。通常，在单变量统计分析中，有些有统计学意义，而另一些无统计学意义；而多元方差分析的结果也会出现有或无统计学意义的情况。一般来说，单变量统计分析强调“局部情况”，而多元统计分析则强调“整体情况”。尤其是当多变量之间存在复杂的相关关系时，基于单变量统计分析结果来下结论，可能会存在问题，甚至可能得出错误的结论。此时，采用多元统计分析，不仅可以获得概括性结论，还有利于提高结论的科学性和可靠性。

4.2 小结

本文介绍了与单组和配对设计多元定量资料差异性分析有关的基本概念、计算方法和两个医学实例的统计分析。基本概念涉及单组设计和配对设计、均值向量、多元方差分析和前提条件；计算方法涉及T^2检验统计量和F检验统计量；两个医学实例分别涉及单组设计和配对设计多元定量资料及其统计分析。借助SAS实现了单组设计和配对设计定量资料的一元和多元方差分析。

参考文献

[2] Johnson DE. 应用多元统计分析方法[M]. 北京：高等教育出版社，2005：397-488.
Johnson DE. Applied multivariate methods for data analysis[M]. Beijing：Higher Education Press，2005：397-488.

[3] 高飞，刘媛媛，李长平，等. 如何正确运用t检验：t检验的基本概念与前提条件[J]. 四川精神卫生，2020，33(3)：211-216.
Gao F，Liu YY，Li CP，et al. How to use t test correctly：the basic concepts and preconditions of t test[J]. Sichuan Mental Health，2020，33(3)：211-216.

[4] 胡纯严，胡良平. 如何正确运用方差分析：方差分析概述[J]. 四川精神卫生，2022，35(1)：6-10.
Hu CY，Hu LP. How to use analysis of variance correctly：an overview of analysis of variance[J]. Sichuan Mental Health，2022，35(1)：6-10.

[5] Armitage P，Colton T. Encyclopedia of biostatistics[M]. 2nd edition. New York：John Wiley & Sons，2005：2667-2670.

[6] Krishnaiah PR. Handbook of statistics（volume 1）：analysis of variance[M]. Amsterdam：North Holland Publishing Company，1980：513-540.

[7] Hotelling H. The generalization of Student's ratio[J]. Annals of Mathematical Statistics，1931，2(3)：360-378.

[8] SAS Institute Inc. SAS/STAT®15.1 user's guide[M]. Cary，NC：SAS Institute Inc，2018：3957-4142.

[9] 胡良平. 面向问题的统计学：(3)试验设计与多元统计分析[M]. 北京：人民卫生出版社，2012：333-352.
Hu LP. Problem-oriented statistics：(3) experimental design and multivariate statistical analysis[M]. Beijing：People's Medical Publishing House，2012：333-352.

（收稿日期：2023-03-19）
（本文编辑：陈 霞）