

合理进行均值比较——含区组因素的两因素析因设计定量资料方差分析

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】 本文目的是介绍与均值比较有关的6个基本概念、计算方法、一个医学试验实例及其SAS实现。基本概念包括参数检验与非参数检验、前提条件、设计类型、交互作用以及方差分析的基本思想。计算方法涉及单因素两水平设计定量资料 t 检验、单因素多水平设计定量资料方差分析和含区组因素的析因设计定量资料方差分析。医学试验实例涉及如何考察可待因和针刺对男性受试者术后牙痛的影响。本文给出了采用SAS处理实例中定量资料的全过程, 包括单因素分析、两因素分析和三因素分析, 对输出结果作出解释, 并明确给出合理选择统计分析方法进行定量资料差异性分析的关键点。

【关键词】 区组因素; 析因设计; 参数检验; 非参数检验; 方差分析

中图分类号: R195.1

文献标识码: A

doi: 10.11886/scjsws20230201001

Reasonably carry out mean value comparison: ANOVA of the quantitative data collected from the two-factor factorial design with a block factor

Hu Chunyan¹, 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 6 basic concepts, calculation methods, a medical experiment example and its SAS implementation related to the mean comparisons. The basic concepts included the parametric test, non-parametric test, preconditions, design types, interaction and the basic ideas of the analysis of variance (ANOVA). The calculation method involved the t test of the quantitative data of single-factor two-level design, ANOVA of the quantitative data of single-factor multi-level design and factorial design with a block factor. An example of a medical experiment that examined the effects of codeine and acupuncture on postoperative toothache in male subjects. The paper gave the whole process of using SAS software to process the quantitative data in the example, including single-factor analysis, two-factor analysis and three-factor analysis, explained the output results, and the key points of reasonably selecting statistical analysis methods to analyze the differences of the quantitative data were clearly given.

【Keywords】 Block factor; Factorial design; Parametric test; Non-parametric test; Analysis of variance

在对定量资料进行差异性分析时, 首选的统计分析方法为参数检验法, 包括 Z 检验、 t 检验和方差分析。选用参数检验法要求定量资料需满足特定的前提条件, 还需要准确判定定量资料所取自的试验设计类型^[1-3]。本文在介绍有关基本概念的基础上, 针对一个多因素定量资料, 介绍如何采用SAS合理地进行多因素定量资料的差异性分析。本文中的定量资料取自含区组因素的两因素析因设计, 此设计的具体做法如下: 由两个试验因素的水平全面组合(设组合数为 k , 即形成 k 个处理组)构成一个两因素析因设计的框架, 将全部受试对象按某种属性

(例如, 本文实例中“耐受程度”, 它被称为区组因素)分成若干小组, 每个小组中有 k 个受试对象, 再将每个小组中的 k 个受试对象随机均分到前面已经形成的 k 个处理组中去。

1 基本概念

1.1 参数检验与非参数检验

所谓参数检验, 就是用于作假设检验的计算公式(即“检验统计量”)所依赖的分布规律中含有特定的参数(即刻画总体分布规律的特征数值), 如总

体均值、总体标准差等。非参数检验,就是在对总体分布形式不了解或知之甚少时进行推断的假设检验方法。在此类检验中,一般不对总体分布作假设,或只作一点诸如连续、对称之类的简单假设。

用于均值比较的常用参数检验有如下 3 类:基于正态分布理论构建的 Z 检验,基于 t 分布理论构建的 t 检验,基于 F 分布理论构建的 F 检验或称方差分析。

1.2 前提条件

使用参数检验要求定量资料需满足 3 个前提条件:独立性、正态性和方差齐性。所谓独立性,是指一批试验数据彼此之间是相互独立的,若通过重复测量获得某受试者同一指标的多个观测值,则这些观测值之间就是不独立的。正态性,是指整个试验中各个组的定量资料应当取自各自的正态分布总体,换言之,是要检验每个小组中的数据是否近似服从正态分布,这里的“小组”是指每个因素的每个水平组。方差齐性,是指每个影响因素各水平组的总体方差应相等。

关于检验“独立性”,没有具体的统计检验方法,可依据基本常识和专业判断进行判断。关于检验“正态性与方差齐性”,有多种统计检验方法,通常需借助计算机软件来实现。

1.3 设计类型

面对一个给定的多因素定量资料,合理选择统计分析方法的关键是正确判定该定量资料所取自的试验设计类型。

所谓设计类型,就是单个因素各水平与受试对象之间存在何种关系;多个因素相互之间以及诸因素与受试对象之间存在什么关系。具体地说,应考虑以下情形:因素在施加时是否有先后顺序、因素对定量结果的影响是否有主次关系、受试对象在某因素的不同水平组中是否相互独立、因素之间是否存在嵌套关系等。统计学家将前述各种情况归纳起来,概括为“设计类型”。

与均值比较有关的常见设计类型包括以下几种:单因素设计(包括单组设计、单因素两水平或多水平设计、配对设计),随机区组设计,拉丁方设计,交叉设计,嵌套设计,裂区设计,析因设计,含区组因素的析因设计,正交设计,均匀设计等。

1.4 交互作用

以两因素为例,所谓两因素 A 与 B 之间的交互作用效应,就是因素 A 各水平对观测结果的影响将随着因素 B 水平的改变而改变。假定因素 A 有 A_1 和 A_2 两个水平,因素 B 有 B_1 和 B_2 两个水平。若试验显示出如下结果:在 B_1 条件下, A_1 的效果好于 A_2 (即 A_1B_1 条件下的结果好于 A_2B_1 条件下的结果);而在 B_2 条件下, A_1 的效果却差于 A_2 (即 A_1B_2 条件下的结果差于 A_2B_2 条件下的结果)。这种反应就被称为因素 A 与因素 B 之间存在交互作用效应。值得一提的是,对于交互作用效应的大小,不应仅凭样本均值的大小加以判断,也需要对交互效应进行假设检验。同理,可理解三因素之间的交互作用与四因素之间的交互作用。

1.5 方差分析的基本思想

所谓方差分析,实际上就是“方差比较”。一般来说,就是对某特定因素(或因素之间的交互作用)的方差与误差的方差进行比较,再依据 F 分布作出统计推断。然而,方差是离均差平方和与其自由度之商。在分析某给定的定量资料时,各待分析项的方差在数量上似乎没有必然的联系,但它们的分子(即离均差平方和)在数量上却有着非常明确的关系,因为它们各项之和等于总离均差平方和。由此可知,方差分析的基本思想是关于总离均差平方和的分解,即将全部数据关于总均值的离均差平方和分解成拟考察项数加一个误差项的离均差平方和。自由度也有类似的分解方法,各部分离均差平方和除以各自的自由度,就是各项的方差(或称均方),以误差的均方为分母,以拟考察的各项均方为分子,就构造出相应的检验统计量 F 。然后,依据 F 分布就可作出统计推断^[4-5]。

2 计算方法

2.1 单因素两水平设计定量资料 t 检验

检验假设(设 μ_1 和 μ_2 分别代表两总体的均值)为:

$$H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2, \alpha = 0.05.$$

其检验统计量分别见式(1)和式(2)。

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{S_{\bar{x}_1 - \bar{x}_2}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{(n_1 + n_2)}{n_1 n_2 (n_1 + n_2 - 2)} (SS_1 + SS_2)}} \quad (1)$$

$$SS_1 = \sum x_1^2 - (\sum x_1)^2 / n, SS_2 = \sum x_2^2 - (\sum x_2)^2 / n \quad (2)$$

式(1)定义的检验统计量 t 服从自由度 $df = n_1 + n_2 - 2$ 的 t 分布, 当 $t \geq t_{(n_1 + n_2 - 2)(1 - \alpha/2)}$ 时, 就有 $P \leq \alpha$; 当 $t < t_{(n_1 + n_2 - 2)(1 - \alpha/2)}$ 时, 就有 $P > \alpha$ 。

2.2 单因素多水平设计定量资料方差分析

单因素多水平设计定量资料, 其总离均差平方和 $SS_{总}$ 可按如下方式分解, 见式(3)、式(4)。

$$SS_{总} = SS_{组间} + SS_{误差} \quad (3)$$

$$MS_{组间} = \frac{SS_{组间}}{df_{组间}}, MS_{误差} = \frac{SS_{误差}}{df_{误差}} \quad (4)$$

$$df_{组间} = \text{组数} - 1, df_{误差} = N - 1 - df_{组间}, F = \frac{MS_{组间}}{MS_{误差}}$$

通过查方差分析用的 F 界值表, 可得 $F_{(df_{组间}, df_{误差})(1 - \alpha)}$ 。

若 $F \geq F_{(df_{组间}, df_{误差})(1 - \alpha)}$, 则 $P \leq \alpha$; 反之, 则有 $P > \alpha$ 。最后, 确定 P 值, 作出统计推断, 并结合专业知识给出专业结论。

【说明】关于析因设计定量资料方差分析计算公式和含区组因素的析因设计定量资料方差分析计算公式都比较复杂, 因篇幅所限, 此处从略。

3 实例与 SAS 实现

3.1 问题与数据结构

3.1.1 一个临床试验问题及数据

【例1】文献[6-7]提供的数据来自一项考察可待因和针刺对男性受试者术后牙痛治疗效果影响的试验。两个试验因素都有两个水平。可待因(Codeine)的两个水平分别是糖胶囊(Codeine=1)与可待因胶囊(Codeine=2); 针刺(Acupuncture)的两个水平分别是两个活动穴位(Acupuncture=1)与两个非活动穴位(Acupuncture=2)。采用析因试验设计, 有4种不同的处理组合。根据疼痛耐受性(PainLevel)评估结果, 将32名受试者分为8组(编号为1~8, 假定编号数值越大, 代表耐受性越强), 每组4名。疼痛缓解程度为Relief, 其数值越大, 表明疗效越好。试验安排和结果见表1, 试分析两个试验因素(即两种胶囊、两类穴位)和区组因素(即耐受性程度)对牙痛治疗效果的影响是否有统计学意义。

3.1.2 对数据结构的分析

此试验数据所对应的数据结构为具有一个区组因素的两因素析因设计的数据结构。其形成过

程包括以下两步: 第一步, 由两个2水平试验因素的水平全面组合, 形成4个不同的试验条件组; 第二步, 将8个区组(即表1中的8行)的每个区组中的4名受试对象随机均分入4个试验条件组中去。这样, 每个试验条件组就有8次独立重复试验的结果, 即8个定量的试验数据(Relief的评分)。

表1 两个试验因素对男性受试者术后牙痛治疗效果的影响结果

Table 1 Results of the influence of two experimental factors on the treatment effect of postoperative toothache in male subjects

| 耐受程度 编号 | 疼痛耐受性评分 | | | | |
|------------|---------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| | A×C: | A ₁ C ₁ | A ₁ C ₂ | A ₂ C ₁ | A ₂ C ₂ |
| 1 | | 0.0 | 0.5 | 0.6 | 1.2 |
| ... | | ... | ... | ... | ... |
| 8 | | 1.2 | 1.7 | 1.6 | 2.4 |

注: “耐受程度编号”是区组因素; A和C是两个试验因素, A代表“针刺两类穴位”, A₁和A₂分别代表“针刺两个活动穴位”与“针刺两个非活动穴位”; C代表“胶囊种类”, C₁和C₂分别代表“糖胶囊”与“可待因胶囊”; “疼痛耐受性评分”是定量的结果

3.1.3 创建 SAS 数据集

设所需要的 SAS 数据步程序如下:

```

title1 'Randomized Complete Block With Two Factors';
data PainRelief;
input PainLevel Codeine Acupuncture Relief @@;
datalines;
1 1 1 0.0 1 2 1 0.5 1 1 2 0.6 1 2 2 1.2
2 1 1 0.3 2 2 1 0.6 2 1 2 0.7 2 2 2 1.3
3 1 1 0.4 3 2 1 0.8 3 1 2 0.8 3 2 2 1.6
4 1 1 0.4 4 2 1 0.7 4 1 2 0.9 4 2 2 1.5
5 1 1 0.6 5 2 1 1.0 5 1 2 1.5 5 2 2 1.9
6 1 1 0.9 6 2 1 1.4 6 1 2 1.6 6 2 2 2.3
7 1 1 1.0 7 2 1 1.8 7 1 2 1.7 7 2 2 2.1
8 1 1 1.2 8 2 1 1.7 8 1 2 1.6 8 2 2 2.4
;
run;
    
```

【变量说明】PainLevel代表“耐受程度编号”; Codeine代表“胶囊种类”; Acupuncture代表“针刺穴位的种类”; Relief代表“疼痛耐受性评分”。

3.2 用 SAS 实现统计分析

3.2.1 进行单因素分析

第一步, 分析可待因胶囊与糖胶囊对疼痛缓解程度的影响。

设所需要的 SAS 过程步程序如下:

```
proc ttest data=PainRelief;
class Codeine;
var Relief;
run;
```

【SAS 输出结果及解释】可待因胶囊与糖胶囊对应的两组疼痛缓解程度的总体方差的比较结果为: $F=1.350$ 、 $P=0.573$, 表明两组定量资料满足方差齐性的要求; 可待因胶囊与糖胶囊两组疼痛缓解程度的总体均值比较结果: $t=-2.730$ 、 $P=0.011$, 表明两总体均值之间差异有统计学意义。由于可待因胶囊与糖胶囊的均值分别为 1.425 与 0.888, 说明可待因胶囊比糖胶囊对疼痛缓解的效果更好。

第二步, 分析针刺两个非活动穴位与两个活动穴位对疼痛缓解程度的影响。

设所需要的 SAS 过程步程序如下:

```
proc ttest data=PainRelief;
class Acupuncture;
var Relief;
run;
```

【SAS 输出结果及解释】针刺两个非活动穴位与两个活动穴位对应的两组疼痛缓解程度的总体方差的比较结果为: $F=1.180$ 、 $P=0.757$, 表明两组定量资料满足方差齐性的要求; 针刺两个非活动穴位与两个活动穴位对应的两组疼痛缓解程度的总体均值比较结果为: $t=-3.500$ 、 $P=0.002$, 表明两总体均值之间差异有统计学意义。由于针刺两个活动穴位与针刺两个非活动穴位的均值分别为 1.481 与 0.831, 说明针刺两个活动穴位比针刺两个非活动穴位对疼痛缓解的效果更好。

第三步, 分析耐受性程度对疼痛缓解程度的影响。

设所需要的 SAS 过程步程序如下:

```
proc anova data=PainRelief;
class PainLevel;
model Relief=PainLevel;
means PainLevel;
run;
```

【SAS 输出结果及解释】8 种耐受性程度的受试者的缓解程度均值比较的方差分析结果为: $F=3.180$ 、 $P=0.016$, 表明 8 种耐受性程度的受试者的缓解程度均值之间差异有统计学意义(说明: 假定 8 组定量资料满足方差齐性要求)。

因篇幅所限, 8 种耐受程度的受试者的缓解程度均值和标准差的计算结果从略。

由输出结果可知: 除第 4 个耐受程度组外, 其他 7 组的均值都随着耐受性水平编号的增大逐渐增大, 说明耐受性强的受试者治疗后疼痛缓解的效果更好。

3.2.2 分析 Codeine 与 Acupuncture 两因素对疼痛缓解程度的影响

第一步, 不考虑两因素之间的交互作用。

设所需要的 SAS 过程步程序如下:

```
proc anova data=PainRelief;
class Codeine Acupuncture;
model Relief=Codeine Acupuncture;
run;
```

【SAS 输出结果及解释】当不考虑交互作用效应时, 两因素析因设计定量资料方差分析结果显示: 两个因素对疼痛缓解程度的影响均有统计学意义。与 Codeine 对应的计算结果为: $F=11.270$ 、 $P=0.002$; 与 Acupuncture 对应的结果为: $F=11.480$ 、 $P<0.01$ 。

第二步, 考虑两因素之间的交互作用。

设所需要的 SAS 过程步程序如下:

```
proc anova data=PainRelief;
class Codeine Acupuncture;
model Relief=Codeine|Acupuncture;
means Codeine*Acupuncture;
run;
```

【SAS 输出结果及解释】当考虑交互作用效应时, 两因素析因设计定量资料方差分析结果表明: 两个因素对疼痛缓解程度的影响均有统计学意义, 但它们之间的交互作用对疼痛缓解程度的影响无统计学意义。具体地说, 与 Codeine 对应的计算结果为: $F=10.960$ 、 $P=0.003$; 与 Acupuncture 对应的计算结果为: $F=16.030$ 、 $P<0.001$; 与 Codeine*Acupuncture 对应的计算结果为: $F=0.210$ 、 $P=0.648$ 。

因篇幅所限, 两因素各水平组合下疼痛缓解程度的均值和标准差的计算结果从略。由输出结果可知: 两个因素都取 1 水平(即用糖胶囊且针刺两个非活动穴位)时, 疗效最差; 两个因素都取 2 水平(即用可待因胶囊且针刺两个活动穴位)时, 疗效最好。

【说明】分析 Codeine 与 PainLevel 两因素对疼痛缓解程度的影响, 分析 Acupuncture 与 PainLevel 两因素对疼痛缓解程度的影响, 可仿照“第 3.2.2 节”

中的做法,完成这两种情况下的分析和结果解释。因篇幅所限,这两部分内容均从略。

3.2.3 进行三因素分析

第一步,不考虑两试验因素之间的交互作用。设所需要的SAS过程步程序如下:

```
proc anova data=PainRelief;
class PainLevel Codeine Acupuncture;
model Relief=PainLevel Codeine Acupuncture;
run;
```

【SAS输出结果及解释】不考虑交互作用效应时,含区组因素的析因设计定量资料方差分析结果见表2。结果表明,三个因素对疼痛缓解程度的影响均有统计学意义。

表2 不考虑交互作用效应时含区组因素的析因设计定量资料方差分析结果

| 变异来源 | 自由度 | Anova SS | 均方 | F | Pr>F |
|-------------|-----|----------|-------|---------|--------|
| PainLevel | 7 | 5.599 | 0.800 | 50.450 | <0.010 |
| Codeine | 1 | 2.311 | 2.311 | 145.800 | <0.010 |
| Acupuncture | 1 | 3.380 | 3.380 | 213.220 | <0.010 |

第二步,考虑两个试验因素之间的交互作用。设所需要的SAS过程步程序如下:

```
proc anova data=PainRelief;
class PainLevel Codeine Acupuncture;
model Relief=PainLevel Codeine|Acupuncture;
run;
```

【SAS输出结果及解释】考虑交互作用效应时,含区组因素的析因设计定量资料方差分析结果见表3。结果显示,三个因素各自对疼痛缓解程度的影响均有统计学意义,但两个试验因素之间的交互作用对疼痛缓解程度的影响无统计学意义。

表3 考虑交互作用效应时含区组因素的析因设计定量资料方差分析结果

| 变异来源 | 自由度 | Anova SS | 均方 | F | Pr>F |
|---------------------|-----|----------|-------|---------|--------|
| PainLevel | 7 | 5.599 | 0.800 | 55.300 | <0.010 |
| Codeine | 1 | 2.311 | 2.311 | 159.790 | <0.010 |
| Acupuncture | 1 | 3.380 | 3.380 | 233.680 | <0.010 |
| Codeine*Acupuncture | 1 | 0.045 | 0.045 | 3.110 | 0.092 |

【说明】由于PainLevel是区组因素,一般来说,区组因素与试验因素之间的交互作用比较小,故通常不需要考虑;此外,三个因素各水平组合下没

有进行重复试验,因而也无法计算三因素之间的交互作用。

3.3 结论

对本例资料进行了三次单因素分析、两次两因素分析和两次三因素分析,结论如下:三个因素各自对疼痛缓解程度的影响均有统计学意义;二个试验因素之间的交互作用对疼痛缓解程度的影响无统计学意义。具体地说,可待因胶囊对疼痛缓解的效果优于糖胶囊,针刺两个活动穴位的效果优于针刺两个非活动穴位,耐受性强的受试者治疗效果更好。

4 讨论与小结

4.1 讨论

从前文结果可知:单因素分析与两因素分析和三因素分析所得到的结论是一致的,这是因为在本例资料中,两个试验因素之间的交互作用无统计学意义。事实上,当因素之间存在复杂的交互作用时,误用单因素分析和不恰当的多因素分析,其结果和结论的可信度均会大打折扣,甚至得出错误的结论。

在分析多因素定量资料时,若分析目的是均值之间的比较,合理选择统计分析方法的关键在于以下两点:①正确判定试验设计类型^[8];②检验前提条件是否满足。

4.2 小结

本文介绍了基于参数检验法进行均值比较的三个前提条件和两类参数检验方法。针对一个实例,呈现了采用SAS实现均值比较的方法,包括单因素两水平设计定量资料t检验和具有一个区组因素的两因素析因设计定量资料方差分析等。分析了以单因素分析取代多因素分析的弊端,并强调准确判定定量资料所来自的试验设计类型的重要性。

参考文献

- [1] 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.
- [2] 胡良平.如何正确运用Z检验:Z检验的基本概念与前提条件

- [J]. 四川精神卫生, 2020, 33(5): 418-421.
- Hu LP. How to use Z test correctly: the basic concepts and the preconditions of the Z test [J]. Sichuan Mental Health, 2020, 33(5): 418-421.
- [3] 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.
- [4] Krishnaiah PR. Handbook of statistics (volume 1): analysis of variance [M]. Amsterdam: North Holland Publishing Company, 1980: 1-132.
- [5] Armitage P, Colton T. Encyclopedia of biostatistics [M]. 2nd edition. New York: John Wiley & Sons, Inc, 2005: 184-199.
- [6] Neter J, Wasserman W, Kutner MH. Applied linear statistical models: regression, analysis of variance, and experimental designs [M]. 3rd edition. Homewood, IL: Irwin, 1990: 941.
- [7] SAS Institute Inc. SAS/STAT®15.1 user's guide [M]. Cary, NC: SAS Institute Inc, 2018: 1053-1108.
- [8] Montgomery DC. Design and analysis of experiments [M]. 6 版. 北京: 人民邮电出版社, 2007: 119-372.
- Montgomery DC. Design and analysis of experiments [M]. 6th edition. Beijing: Posts & Telecom Press, 2007: 119-372.

(收稿日期: 2023-02-01)

(本文编辑: 陈霞)