

骨科科研实验中的成组设计及其 SAS 分析

王琪, 胡良平

(军事医学科学院生物医学统计学咨询中心, 北京 100850)

关键词 成组设计; SAS; 中医骨伤科学

DOI: 10.3969/j.issn.1003-0034.2010.06.031

Group design and SAS analysis in orthopedics research experiments WAGN Qi, HU Liang-ping. *Biomedicine Statistics Consultation Center, Academy of military Medical Sciences, Beijing 100850, China*

Key words Group design; SAS; Orthopedics-traumatology

Zhongguo Gushang/China J Orthop & Trauma, 2010, 23(6): 477-480 www.zggszz.com

当每次实验仅涉及 1 个重要实验因素的 1 个特定水平, 所考察的因素确实是 1 个“单因素”, 其各水平之间只是量上的不同, 而没有本质的差别; 一切非实验因素对该因素各水平组的影响是平衡的, 选用单因素设计是合适的^[1]。在前面的统计讲座中, 我们详细介绍了单因素设计中的单组设计和配对设计及其 SAS 分析, 本文将重点介绍成组设计。当实验研究中只考察 1 个 2 水平的实验因素, 但受试对象不具备配对的条件时, 可选用此设计。

1 成组设计的定义和特点

两组受试对象接受不同处理, 这样安排实验的方法叫做成组设计。实验中, 仅涉及一个具有 2 水平的实验因素, 若实验因素独立于受试对象, 则可将全部受试对象完全随机地分

入该实验因素的 2 个水平组中去; 反之, 将从特定的 2 个子总体中随机地抽取受试对象^[2]。

实验中, 仅涉及 1 个具有 2 水平的实验因素, 未对其他任何重要非实验因素进行有计划的安排, 仅仅希望通过随机化分组来平衡所有非实验因素在两组间对观测结果的干扰和影响。

2 成组设计的资料性质分类和统计分析方法的合理选用

成组设计按资料性质可以分为成组设计定量资料和成组设计定性资料; 按观测指标的数量可以分为成组设计一元资料和成组设计多元资料。定量资料假设检验以两组样本分别代表的总体均值(中位数)之间差别是否具有统计学意义为目的, 常见的情况见表 1。

3 成组设计应用举例

【例 1】将 24 只 Wistar 大鼠随机分为对照(生理盐水,

通讯作者: 胡良平 E-mail: lphu812@126.com

- 540.
- [9] 谭明生. 胸腰椎骨折的分类与治疗选择. 中国骨伤, 2008, 21(1): 1-3.
- [10] 刘延东, 刘景生. 胸腰椎爆裂骨折机制及保守治疗的研究进展. 中国骨伤, 2004, 17(4): 254-255.
- [11] Singh K, Vaccaro AR, Eichenbaum MD, et al. The surgical management of thoracolumbar injuries. *J Spinal Cord Med*, 2004, 27(2): 95-101.
- [12] Eskenazi MS, Bendo JA, Spivak JM. Thoracolumbar spine trauma: evaluation and management. *Curr Opin in Orthoped*, 2000, 11: 176-185.
- [13] Vaccaro AR, Lehman RA Jr, Hurlbert RJ, et al. A new classification of thoracolumbar injuries: the importance of injury morphology, the integrity of the posterior ligamentous complex, and neurologic status. *Spine*, 2005, 30(20): 2325-2333.
- [14] 董健. 胸腰椎骨折的手术指征及方式的合理选择. 中国骨伤, 2009, 22(7): 485-487.
- [15] Kim DH, Vaccaro AR. Osteoporotic compression fractures of the spine: current options and considerations for treatment. *Spine J*, 2006, 6(5): 479-487.
- [16] 胡波. 综合法治疗胸腰椎压缩性骨折. 中国骨伤, 2001, 14(6): 381.
- [17] Domenicucci M, Preite R, Ramieri A, et al. Thoracolumbar fractures without neurosurgical involvement: surgical or conservative treatment. *J Neurosurg Sci*, 1996, 40(1): 1-10.
- [18] 江志东. 前置式脊柱骨折复位固定器治疗胸腰椎骨折. 世界中医骨科杂志, 2005, 7(2): 89-90.
- [19] 李勇强, 宋琦, 任树军, 等. 三维外固定架治疗老年胸腰椎楔形骨折. 中华实用中西医杂志, 2007, 20(8): 705.
- [20] Aligizakis A, Katonis P, Stergiopoulos K, et al. Functional outcome of burst fractures of the thoracolumbar spine managed non-operatively, with early ambulation, evaluated using the load sharing classification. *Acta Orthop Belg*, 2002, 68(3): 279-287.
- [21] 周辉, 彭亮, 韩勇, 等. 体外充气复位结合椎体成形术治疗胸腰椎压缩性骨折. 中国骨伤, 2007, 20(3): 155-157.
- [22] 彭亮, 张政宏, 倪飞, 等. 便携式充气复位仪治疗椎体压缩骨折的临床研究. 中国中医骨伤科杂志, 2008, 16(9): 13-15.
- [23] 何立成, 谭远超, 陶君. 非手术治疗胸腰段椎体前柱骨折 78 例体会. 中医正骨, 2008, 20(2): 56-57.
- [24] Mazanec DJ, Podichetty VK, Mompoin A, et al. Vertebral compression fractures: manage aggressively to prevent sequelae. *Cleve Clin J Med*, 2003, 70(2): 147-156.

(收稿日期: 2010-01-23 本文编辑: 王宏)

表 1 常用成组设计类型及统计分析方法

设计类型	可选用的统计分析方法
成组设计一元定量资料	成组设计 t 检验或 Wilcoxon 秩和检验
成组设计多元定量资料	成组设计多元方差分析
成组设计二值定性资料	一般 χ^2 检验或 Fisher 精确检验、RR 值和 OR 值及其假设检验
成组设计多值定性资料(有序)	秩和检验或 Ridit 分析或有序变量的 logistic 回归分析
成组设计多值定性资料(无序)	一般 χ^2 检验或 Fisher 精确检验

表 2 脊髓损伤 2 周两组大鼠的斜板维持率

组别	斜板维持率										
NS 组	0.351	0.342	0.321	0.361	0.307	0.369	0.379	0.392	0.381	0.312	0.463
ATP 组	0.425	0.417	0.406	0.517	0.483	0.397	0.436	0.459	0.483	0.491	0.406

NS)组和腺苷(ATP)组,每组 12 只。改良 Allen 方法致 T₁₁ 段脊髓不完全损伤,经腹腔注射 NS(1.5 ml)或 ATP(40 mg/kg),伤后 2 周进行斜板实验,计算斜板维持率。斜板维持率=(术后斜板角度/术前斜板角度)×100%,结果如表 2 所示^[3]。问:两组大鼠斜板维持率的差别有无统计学意义?

【分析与解答】该定量资料设计类型属于成组设计。因定量资料满足正态性和方差齐性(具体检验结果见 SAS 输出结果),故采用成组设计定量资料的 t 检验。

下面用 SAS 软件编程法进行分析:

```

DATA a;
INPUT g$;
DO i=1 to 11;
INPUT x @@;
OUTPUT;
END;
CARDS;
NS
0.351 0.342 0.321 0.361
0.307 0.369 0.379 0.392
0.381 0.312 0.463
ATP
0.425 0.417 0.406 0.517
0.483 0.397 0.436 0.459
0.483 0.491 0.406
;
RUN;
PROC SORT; BY g;
RUN;
PROC UNIVARIATE NORMAL;
VAR x; BY g;
RUN;
PROC TTEST COCHRAN;
CLASS g; VAR x;
RUN;
PROC NPAR1WAY WILCOXON;
CLASS g; VAR x;
RUN;

```

【SAS 程序说明】数据步,DATA 语句建立数据集 a,用 INPUT 语句输入数据,@@ 是 SAS 软件中规定的指针控制符。变量 g\$代表分组变量,g 的取值可以是字符串,x 为斜板维持率。“PROC SORT”以 g 作为分组变量对全部数据进行排序。“PROC UNIVARIATE NORMAL”进行单变量分析和正态性检验。“PROC TTEST”调用 ttest 分析过程进行 t 检验,“PROC NPAR1WAY WILCOXON”为进行秩和检验。

【主要输出结果】如下所示:

g=ATP		矩	
N	11	权重总和	11
均值	0.447 272 73	观测总和	4.92

g=NS

矩

N	11	权重总和	11
均值	0.361 636 36	观测总和	3.978

以上显示的是两组的均值。可见,ATP 组大鼠斜板维持率的均值高于 NS 组。

Obs	W	P
1	0.916 761	0.292 6
2	0.919 596	0.315 3

以上是正态性检验的结果,可知两组定量资料均服从正态分布(W 分别为 0.916 761、0.919 596;P 分别为 0.292 6、0.315 3,均大于 0.05)。

T-Tests

Variable	Method	Variances	DF	tValue	Pr> t
x	Pooled	Equal	20	4.70	0.000 1
x	Satterthwaite	Unequal	19.9	4.70	0.000 1
x	Cochran	Unequal	10	4.70	0.000 8

Equality of Variances

Variable	Method	Num DF	Den DF	F Value	Pr > F
x	Folded F	10	10	1.15	0.826 5

以上是两组方差齐性检验的结果,可知两组总体方差相等($F=1.15, P=0.826 5>0.05$)。

【统计和专业结论】该定量资料设计类型属于成组设计。因定量资料满足正态性和方差齐性,故采用成组设计定量资料的 t 检验: $t=4.70, P=0.000 1<0.05$,两个平均值之间的差别有统计学意义,两组大鼠平均斜板维持率不相等,由均值可知,ATP 组大鼠的斜板维持率高于 NS 组。

【例 2】选择 50 例活动期类风湿关节炎患者,分为对照组与治疗组。两组均给予基础治疗:帕夫林(白芍总苷)口服,600 mg,每日 2 次;妥抒(来氟米特)口服,20 mg/d。同时,治疗组应用松梅乐(鹿瓜多肽注射液)静脉点滴,16 mg/d,连续 2 周;对照组给予西乐葆(塞来昔布)口服,200 mg,每日 2 次,连续 2 周。具体数据见表 3^[4],试比较两组患者的疗效是否相同。

【分析与解答】该研究的观测对象分为两组,观测指标是定性的二值变量(疗效),因此它属于成组设计 2×2 表资料,可以采用一般 χ^2 检验或 Fisher 精确检验来处理。

下面用 SAS 软件编程法进行分析:

【SAS 程序说明】该程序关键语句是 CHISQ,目的是输出

表 3 两组患者的治疗效果比较(例)

分组	例数			合计
	治疗效果:	有效	无效	
对照组		8	13	21
治疗组		20	9	29
合计		28	22	50

```
DATA b;
DO A=1 TO 2;
DO B=1 TO 2;
INPUT F @@; OUTPUT;
END;
END;
CARDS;
8 13
20 9
;
RUN;
```

几种常用的 χ^2 检验统计量,列联表为 2×2 表时,还可以输出 Fisher 精确检验统计量。

【主要输出结果】如下所示:

A * B 表的统计量

统计量	自由度	值	概率
卡方	1	4.710 7	0.030 0
似然比卡方	1	4.759 0	0.029 1
连续校正卡方	1	3.541 2	0.059 9
Mantel-Haenszel 卡方	1	4.616 5	0.031 7
Phi 系数		-0.306 9	
列联系数		0.293 4	
Cramer 的 V		-0.306 9	

以上给出了若干 χ^2 检验统计量,通常查看一般 χ^2 检验结果即可,当总频数 $n > 40$ 且至少有一个理论频数 $1 < T \leq 5$ 时,可选用连续校正的 χ^2 检验。

Fisher 精确检验

单元格 (1,1) 频数 (F)	8
左侧 Pr <= F	0.029 7
右侧 Pr >= F	0.993 3
表概率 (P)	0.023 0
双侧 Pr <= P	0.044 2

以上是 Fisher 精确检验结果,在本例中 $P = 0.044 2$ 。当总频数 $n < 40$ 或有理论频数 $T < 1$ 时(SAS 输出结果会有提示),必须选用 Fisher 精确检验。

【统计和专业结论】本例属于成组设计中的一种特殊情况,是横断面研究设计 2×2 表资料,数据符合一般 χ^2 检验使用条件,因此只需要查看一般 χ^2 检验结果既可; $\chi^2 = 4.710 7, P = 0.03$,按照 0.05 的检验水准,拒绝原假设(两组有效率相等),接受备择假设(两组有效率不等),可以认为两组患者的治疗效果不同。结果表明,治疗组的有效率高于对照组。

【例 3】选择进行锚定法单开门椎管成形术并获得完整随访的患者 56 例(锚定法单开门, A 组),与 50 例同期进行传统单开门手术的患者(传统单开门, B 组)进行对比分析,对两组手术后轴性症状进行比较。具体数据见表 4^[5]。

表 4 两组手术后轴性症状比较(例)

组别	例数				合计	
	分级:	优	良	可		差
A 组		12	21	17	6	56
B 组		5	7	28	10	50
合计		17	28	45	16	106

【分析与解答】本例属于成组设计 2×C 表资料,结果变量为多值有序变量,因此可以采用秩和检验方法来处理。

下面用 SAS 软件编程法进行分析(见本页右下角):

【SAS 程序说明】数据步,首先建立名为 c 的数据集,建立数值变量 A、B、F,分别读入行号、列号、每格实际频数;然后建立名为 c2 的数据集,用 SET 语句把 c 复制到 c2 中,用循环 DO 语句重复执行 DO 和 END 之间的语句,并用 OUTPUT 语句把执行结果输出到数据集里面。过程步,调用单因素非参数检验过程 NPAR1WAY, WILCOXON 语句表示基于 WILCOXON 得分进行分析:两组比较时进行 Wilcoxon 秩和检验和 Kruskal-Wallis 检验,多组比较时执行 Kruskal-Wallis 检验;最后用 CLASS 指定分类变量,VAR 指定观测变量。

【主要输出结果】如下所示:

```
The NPAR1WAY Procedure
Wilcoxon Scores (Rank Sums) for Variable B
Classified by Variable A
A N Sum of Scores Expected Under H0 Std Dev Under H0 Mean
Score
1 56 2 516.50 2 996.0 149.697 809 44.937 50
2 50 3 154.50 2 675.0 149.697 809 63.090 00
Average scores were used for ties.
```

由以上结果可以看出,各组的平均秩和:A 组为 44.937 50, B 组为 63.090 00。

Wilcoxon Two-Sample Test

Statistic	3 154.500 0
Normal Approximation	
Z	3.199 8
One-Sided Pr > Z	0.000 7
Two-Sided Pr > Z	0.001 4

```
DATA c;
DO A=1 TO 2;DO B=1 TO 4;
INPUT F @@;
OUTPUT;
END;
END;
CARDS;
12 21 17 6
5 7 28 10
;
RUN;
```

表 5 睁眼时两组的重心移动图类型分布情况(例)

分组	人数							合计
	重心移动图类型:	多中心型	弥散型	前后型	中心型	斜向型	左右型	
颈性眩晕组		4	35	26	3	6	1	75
正常组		7	8	55	12	2	2	86
合计		11	43	81	15	8	3	161

t Approximation
 One-Sided Pr > Z 0.000 9
 Two-Sided Pr > |Z| 0.001 8
 Z includes a continuity correction of 0.5.
 以上是对两组患者手术后轴性症状分级进行 Wilcoxon 两样本秩和检验, 双侧检验 $P=0.0014 < 0.05$ 。

Kruskal-Wallis Test
 Chi-Square 10.260 0
 DF 1
 Pr > Chi-Square 0.001 4

以上是对两组患者手术后轴性症状分级进行 Kruskal-Wallis 秩和检验, $\chi^2=10.2600, P=0.0014 < 0.05$, 与 Wilcoxon 两样本秩和检验结果一致。

【统计和专业结论】根据对两组进行秩和检验的结果, $P < 0.05$, 拒绝原假设(两组手术后轴性症状总体中位数相同), 接受备择假设(两组手术后轴性症状总体中位数不同或不完全相同), 可以认为两组患者手术后轴性症状分级的差别有统计学意义。结果表明, B 组症状明显者(症状分级为可与差者)的出现率高于 A 组。

【例 4】应用人体重心测试系统对 86 例正常人和 75 例颈性眩晕患者进行定量测试和定性诊断, 睁眼时两组的重心移动图类型的分布见表 5, 试比较睁眼时颈性眩晕组和正常组在重心移动图类型的分布上有无差别^[6]?

【分析与解答】本例属于成组设计 2×C 表资料, 结果变量为多值名义变量, 因此可以用一般 χ^2 检验或 Fisher 的精确检验来处理, 但因表 5 小于 5 的理论频数的个数超过了总格子数的 1/5, 所以应选用 Fisher 的精确检验。

下面用 SAS 软件编程法进行分析:

```

DATA d;
DO A=1 TO 2; DO B=1 TO 6;
INPUT F @@; OUTPUT;
END; END;
CARDS;
4 35 26 3 6 1
7 8 55 12 2 2
;
RUN;

PROC FREQ;
WEIGHT F;
TABLES A * B/CHISQ;
Exact Fisher;
RUN;

```

【SAS 程序说明】由于资料非 2×2 表, 故程序中须输入 Exact Fisher 要求系统输出 Fisher 精确检验结果。

【主要输出结果】如下所示:

A * B 表的统计量			
统计量	自由度	值	概率
卡方	5	35.301 0	<.000 1
似然比卡方	5	37.201 3	<.000 1
Mantel-Haenszel 卡方	1	4.533 0	0.033 2

Phi 系数 0.468 3
 列联系数 0.424 1
 Cramer 的 V 0.468 3

WARNING: 33% 的单元格的期望计数比 5 小。
 卡方可能不是有效检验。

Fisher 精确检验

表概率 (P) 1.640E-11
 Pr <= P 1.356E-07

当不满足一般 χ^2 检验前提条件时 (SAS 输出结果会提示), 需要查看 Fisher 精确检验结果。以上是 Fisher 精确检验结果, $P=1.356E-07$ 。

【统计和专业结论】本例资料经过 Fisher 精确检验, $P=1.356E-07 < 0.05$, 按照 0.05 的检验水准, 可认为睁眼时颈性眩晕组与正常组在重心移动图类型的分布上的差别有统计学意义, 即颈性眩晕组在重心移动图类型上以弥散型为主、以前后型为辅, 而正常组在重心移动图类型上以前后型多见。

参考文献

- [1] 胡良平. 口腔医学科研设计与统计分析. 北京: 人民军医出版社, 2007. 125.
- [2] 胡良平. 统计学三型理论在实验设计中的应用. 北京: 人民军医出版社, 2006. 52-57.
- [3] 王颖, 陈晓亮. 外源性腺苷对大鼠急性脊髓损伤后运动诱发电位和后肢运动功能的影响. 中国骨肿瘤骨病, 2005, 4 (5): 297-299.
- [4] 郭生玉, 林连捷, 楚川, 等. 鹿瓜多肽治疗类风湿关节炎临床疗效观察. 中国骨伤, 2008, 21 (8): 610-612.
- [5] 张学利, 王善金, 王云力, 等. 锚定法单开门椎管成形术对术后轴性症状和颈椎曲度影响的病例对照研究. 中国骨伤, 2008, 21 (10): 759-761.
- [6] 黄阿勇, 周卫, 钟红刚, 等. 人体重心测试系统诊断颈性眩晕的评价. 中国骨伤, 2008, 21 (5): 345-348.

(收稿日期: 2010-02-23 本文编辑: 连智华)